A single-cell atlas of *Toxoplasma* sexual development identifies regulators of gametogenesis

Hisham S Alrubaye[1], Sarah M Reilly[1], Rafaela da Silva[1], NyJaee Washington[1] and Jon P Boyle[1]*

[1] Department of Biological Sciences, Dietrich School of Arts and Sciences, University of Pittsburgh

*To whom correspondence should be addressed:
University of Pittsburgh
4249 Fifth Avenue
Pittsburgh, PA. 15260
boylej@pitt.edu

Abstract

Toxoplasma gondii undergoes sexual development exclusively in the feline intestine, a process critical for genetic diversity and population expansion. Recent studies have identified genes critical in suppressing presexual development[1] and metabolic differences in felines that may promote sexual development[2], but to date the gene regulatory networks driving development in the cat are unknown. To investigate this, we performed single-cell transcriptomics on parasites isolated from cat intestines, using fluorescent reporter strains and flow cytometry. From 15,068 cells across two experiments, we identified rare populations, including cells that bear all of the hallmarks of gametes. Candidate genes emerging from this study were tested via CRISPR-Cas9 Perturb-seq, identifying AP2X6 as a regulator of macrogametocyte development. Our single-cell data extends what is known about gene expression changes throughout sexual development and should be useful to those in the field working towards inducing gametogenesis, mating, and oocyst production *in vitro*.

Main

T. gondii has a complex life cycle that includes both asexual and sexual reproduction. While asexual stages can develop in nearly all cells of warm-blooded animals, sexual development is restricted to the epithelial cells of the felid small intestine. The sexual phase of *T. gondii* is essential for generating genetic diversity through recombination and for large scale environmental dissemination. A single feline host is capable of shedding almost a billion oocysts after a single infection [3–6]. Shed oocysts mature further after completion of meiosis and one round of mitosis, resulting in the production of infectious sporozoites.

Among the developmental stages of *T. gondii*, the tachyzoite and bradyzoite have been studied the most due to their relative tractability *in vitro*, and a substantial amount of information is available on the mechanisms of tachyzoite to bradyzoite stage transition[7,8] as well as stage-specific cell cycle regulation[9,10], metabolism[11,12], and host-parasite interactions[13]. In contrast, due to challenges associated with cultivating pre-sexual and ultimately sexual stages, there is far less known about how this developmental program is triggered and executed by the parasite during infection of the feline host. Early foundational ultrastructural studies by Dubey and Frenkel in the 1970s began to differentiate between the asexual and sexual enteric stages of *T. gondii*[3,14]. These studies identified five morphologically distinct pre-gamete stages that develop following infection of cat enterocytes[14]. While these investigations filled critical gaps in knowledge and provided a developmental framework for understanding enteric development, the molecular mechanisms underlying these processes remained unknown until recent advances. First, global transcriptomic studies showed that *T. gondii* undergoing enteric development in the feline host have unique transcriptional profiles compared to other developmental stages. These bulk transcriptomic studies have been instrumental in uncovering significant transcriptional differences between

51 developmental stages in intermediate and definitive hosts, offering a comprehensive overview of
52 stage-specific transcriptional programs[15–17]. Furthermore, metabolic cues from the feline may help
53 trigger the transition to from asexual to presexual development in *T. gondii,* highlighting that host-
54 specific signals are essential in orchestrating parasite differentiation[18]. In addition, recent studies
55 have identified parasite transcription factors AP2XII-1 and AP2XI-2 as key suppressors of *T.*
56 *gondii* pre-sexual development, and their genetic depletion *in vitro* permits presexual development
57 [add citation]. These studies have paved the way for targeted investigations into the regulation of
58 this developmental program[1]. However, the depletion of those suppressors did not result in
59 parasites that were capable of forming gametes *in vitro*, indicating that additional factors are
60 required for gametogenesis.

61 Here, we interrogated the transcriptional program underlying the enteric stages of *T. gondii*
62 in the cat using single-cell RNA sequencing (scRNA seq) to identify previously unknown
63 regulators of enteric development. This approach enabled us to capture the transcriptional
64 heterogeneity of *T. gondii* parasites developing in the feline intestine at unprecedented resolution.
65 We identified multiple transcriptionally distinct clusters within the enteric stages, including
66 putative male and female gametocyte populations based on sex-specific expression signatures. To
67 functionally characterize key regulators of parasite development, we designed a targeted perturb-
68 seq screen of 28 candidate genes, leading to the identification of AP2X6 as a regulator of
69 macrogametocyte differentiation. Our data provides the field with novel gene markers for studying
70 distinct subpopulations of *T. gondii*, offering new targets for individual investigation to enhance
71 our understanding of the molecular mechanisms underlying the parasite development in the
72 definitive host.

73 **Results**

74 **Generation of a cat-competent *Toxoplasma gondii* reporter strain**

75 Transcriptomic analyses of *T. gondii* enteric stages in the feline intestine have thus far been
76 restricted to bulk RNA sequencing approaches, which lacks the resolution to capture the cellular
77 heterogeneity and obscures rare subpopulations such as transitional forms and gametes[16,17]. To
78 address this limitation, we interrogated the transcriptional program underlying enteric
79 development of *T. gondii* by generating single-cell transcriptomic profiles from parasites isolated
80 from the feline small intestine. To facilitate this, a *TgVEG* reporter line was engineered to express
81 GFP from a highly active tachyzoite promoter (GRA1) and mCherry from a hybrid promoter
82 (GRA11A + GRA1 upstream sequences) that resulted in a strain expressing GFP during tachyzoite
83 development and mCherry during both asexual and sexual development (Fig. 1A; Extended Data
84 Fig. 1).

85 To determine if our engineered strain retained the ability to complete its entire life cycle,
86 five brains from chronically infected CBA/J mice were fed to a cat. Tissue cysts from those brains
87 expressed both GFP and mCherry (Fig. 1B). During feline infection, fecal samples were collected
88 daily between days 3 and 17 post-infection. Unsporulated oocysts of TgVEG:GFP:mCh lacked
89 detectable GFP or mCherry fluorescence (Fig. 1C), whereas sporulated oocysts exhibited mCherry
90 fluorescence but no GFP fluorescence (Fig. 1D). Sporozoites were excysted from oocysts via
91 mouse infection[3], and parasites were recovered from the peritoneum by lavage and as expected
92 GFP and mCherry fluorescence were detected in TgVEG:GFP:mCh parasites growing as
93 tachyzoites *in vitro* (Fig. 1D). Because mCherry expression was observed in tachyzoites,
94 bradyzoites, and sporozoites, we hypothesize that constitutive expression results from the absence
95 of regulatory elements in the truncated GRA11A promoter, which has been proposed to be

96 occupied by MORC[19] , along with the physical proximity of this truncated GRA11A promoter to
97 GRA1 sequences in the plasmid driving GFP (see Extended Data Fig.1)
98
99

100 **A single cell atlas of *T. gondii* enteric development**

101 To identify the transcriptional program underlying the enteric development, we used
102 TgVEG:GFP:mCh parasites in feline infections. Thirty mice were infected with 1,000 low passage
103 TgVEG:GFP:mCh parasites, and 21 Days post-infection (DPI) mice were euthanized, and
104 deskinned skulls were divided into two groups and fed to two cats. Tissues were consumed within
105 30 min and infected cats began shedding oocysts on D4 PI. On D5 or D6 PI, feline intestinal
106 enterocytes were isolated from the small intestine by manual scraping, mechanically lysed, and
107 parasites were isolated by flow cytometry using expression of fluorescent proteins as proxies.
108 Single-cell transcriptomic profiles were generated for a total of 16,966 cells across two
109 independent experiments (Fig. 2A).

110 The single-cell RNA sequencing (scRNA-seq) transcriptomic profiles of *T. gondii* isolated
111 from feline enterocytes on 5 and 6 DPI differed significantly from *T. gondii* tachyzoites cultured
112 *in vitro* for 8 days post-excystation (Fig. 2B). As expected, the *in vitro* excysted parasites showed
113 high transcript abundance for canonical tachyzoite markers like SAG1 and SAG2A (Fig. 2C; top
114 panels)[20,21]. In contrast, parasites derived from the cat intestine exhibited significantly higher levels
115 of known merozoite markers like GRA80 and PNP (Fig. 2C; bottom panels) compared to
116 tachyzoites[1,16]. Overall, the gene expression profiles of intestine-derived parasites show many
117 unique transcripts that were distinct from those of tachyzoites as well as overlapping transcripts
118 (Fig. 2D), underscoring both the differences between life stages and the remarkable adaptability

119 of *T. gondii*, which enables it to thrive across various host environments. To identify any
120 differences between the cat samples, we downsampled to account for the differences in the cell
121 counts and performed differential gene expression analysis (Extended data Fig. 2). The enteric
122 samples from both cats are highly correlated (Fig. 2E)

123 **Identifying distinct clusters of *T. gondii* in enteric stages**

124 To interrogate the developmental program and identify distinct clusters in cat enteric
125 stages, we first identified and filtered out doublets using the scDblFinder algorithm[22]. The optimal
126 neighborhood size parameter (pK) for doublet detection was empirically determined as 0.08 to
127 maximize classification accuracy[23]. Artificial doublets were simulated from 25% of cells (pN =
128 0.25), and dimensionality reduction was performed using the top 25 principal components as
129 observed by the elbow method (Extended Fig. 3B). Predicted doublets were excluded prior to
130 downstream analysis. In addition, cells with fewer than 500 unique molecular identifiers (UMIs)
131 were removed based on the knee point detection in EmptyDrops (Extended Fig. 2A)[24]. After
132 filtering 15,068 of 16,966 cells passed these filters. Gap statistic analysis was then applied to
133 determine the optimal number of clusters in the scRNA-seq dataset[25,26]. This analysis identified K
134 = 10 as the optimal cluster number, which was used to partition cells into transcriptionally distinct
135 subpopulations (Fig. 3A). Feature plots of reported enteric stage-specific *T. gondii* genes show
136 variation in spatial distribution of transcript abundance (Fig. 3B).

137 To assess if any of the clusters were associated with previously reported canonical marker
138 genes, we examined the differentially expressed genes across all clusters and associated them with
139 cell-type based on previous studies. Cluster 1 showed enriched expression of canonical tachyzoite
140 markers like SAG2A and SAG1 (Fig. 3D). Both genes were found to be significantly enriched and
141 uniquely expressed in this cluster ($\log_2$FC >15; adj. p = 0.0). Therefore, we annotated cells in

142 cluster 1 as 'enteric tachyzoites' (Fig. 3C). It is possible that these parasites were in the process of
143 dissemination from the intestinal lining to other tissues within the host, since extraintestinal
144 development of *T. gondii* occurs even in the definitive host. Cluster 2 was defined by high
145 expression of genes like *Bradyzoite Rhoptry Protein 1* (*BRP1*; $\log_2$FC >9; adj. p = 1.43e-93) a
146 bradyzoite and merozoite marker, and AMA2 which is a protein involved bradyzoite invasion
147 ($\log_2$FC >9; adj. p = 1.43e-93)[7,27,28]; Extended Data Fig. 2C and 2D). While BRP1 expression was
148 expected in merozoites[29], it was actually restricted to cluster 2 rather than being broadly distributed
149 across all parasites. TGME49_275790 is a merozoite-associated gene that is also enriched in this
150 cluster ($\log_2$FC >11.5; adj. p = 0), along with other tachyzoite-associated genes like ALV3 and
151 ALV12 ($\log_2$FC >12; adj. p = 0; Fig. 3E). The presence of bradyzoite/merozoite and tachyzoite
152 markers in these cells distinguishes this cluster from the rest of the enteric parasite population and
153 suggests that these parasites may represent an intermediate state between known canonical stages.
154 Therefore, this cluster was annotated as "BRP1+" based on one of its defining markers (Fig. 4E).
155 We compared the transcript abundance of the top 100 genes in the BRP1+ cluster to *in vivo* bulk
156 RNAseq transcriptome data from *T. gondii* cat infections[17] and found that these genes were most
157 characteristic of EES2 and EES3 stages (Fig. 3F). This confirms that this cluster is not an anomaly
158 of this particular infection as it is consistent with this prior dataset from infected definitive host
159 stages.

160 Clusters 3, 4, and 5 were enriched for multiple known merozoite-associated transcripts
161 (Fig. 3G) including SRS15B, SRS22E, SRS22D, GRA11A, and GRA11B[1,16,17,19,30]. Cluster 3
162 comprised 1,317 cells and was defined by high expression of TGME49_287040 ($\log_2$FC = 5.91;
163 adj. p = 0) and TGME49_200270 ($\log_2$FC = 5.87; adj. p = 0), both previously detected in enteric-
164 stage parasites[17]. Cluster 4 was the largest cluster, containing 12,051 cells (~80% of all cells), and

165 was characterized by the expression of genes like IMC7 ($\log_2$FC = 1.04; adj. p = 0) and
166 TGME49_244440 ($\log_2$FC = 1.73; adj. p = 0)[17]. Cluster 5 is made up of 1,199 cells with multiple
167 differentially expressed transcripts including TGME49_311720 ($\log_2$FC = 1.8; adj. p = 0) and
168 TGME49_293870 ($\log_2$FC = 1.8; adj. p = 0)[17]. Although all three clusters shared core merozoite
169 markers like GRA11A, GRA11B, and GRA82, each exhibited unique gene signatures that
170 distinguish them from one another. The top 10 differentially expressed genes for each cluster are
171 shown in Fig. 3H. Based on these profiles, clusters 3, 4, and 5 were annotated as "Merozoite-A,"
172 "Merozoite-B" and "Merozoite-C" respectively.

173 Cluster 6 was enriched for transcripts associated with late-stage development, as defined
174 by comparative expression analysis with data from Ramakrishnan *et al*[17]. Among the top
175 differentially expressed genes were TGME49_266960 ($\log_2$FC = 3.38; adj. p = 0) and
176 TGME49_290200 ($\log_2$FC = 4.41; adj. p = 0; Fig. 4B). When compared to Ramakrishnan *et al*[17],
177 the top 100 genes in this cluster exhibited increased expression in late EES samples (EES4 and 5;
178 Fig. 4C). Based on data presented immediately below, we hypothesizes that cluster 6 represents a
179 pre-gamete stage of development. Therefore, this cluster was annotated as "Pre-gamete" (Fig. 4A).

180 To determine if any of the clusters were associated with canonical markers of female
181 gametocytes, we used the genes HOWP1 and AO2 which are known female gamete markers in *T.*
182 *gondii*[31]. We found transcripts for these genes to be exclusively present in cells of cluster 7
183 (Extended Data Fig 4A and 4B)[31]. To further confirm this cluster as being made up of female
184 gametocytes, we used comparative transcriptomics between recently published *Cryptosporidium*
185 *parvum* datasets, *Plasmodium* and our data to identify conserved female-specific markers[32]. Out
186 of 108 female gametocyte-specific genes identified in *T. gondii*, *C. parvum* and *P. berghei*[31–33],
187 we found that 28 of them were enriched in Cluster 7 (Fig. 4D; Extended Data Fig 4C ;

188 Supplementary Table 2), including *TGME49_272900*, whose *Plasmodium* ortholog
189 (*PF3D7_0816800)* encodes the female-specific meiotic recombination protein DMC1[34], and
190 *TGME49_294820*, which encodes a type I fatty acid synthase, a protein found in high abundance
191 in oocysts from *T. gondii* and *Eimeria* and is also enriched in *Eimeria tenella* gametocytes [35]. The
192 expression of reported female-specific genes supported the annotation of this cluster as 'Female',
193 thereby providing the first detailed transcriptional view of *T. gondii* macrogametocytes. Among
194 the top 20 genes in the female cluster, three belong to the family D protein family
195 (TGME49_236975, TGME49_316670, and TGME49_277572), a large family of tandemly
196 duplicated genes of mostly unknown function. Additionally, we found that a putative surface
197 antigen, SRS46, is uniquely expressed in the cells of the Female cluster (Extended Data Fig. 3D),
198 making SRS46 a candidate surface marker of macrogametocytes. Furthermore, we examined the
199 top 100 differentially expressed genes with the cluster and found that their transcript abundance
200 was increased in late developmental samples from Ramakrishnan *et al*[17] (Fig. 4E). Thus, the
201 transcriptional signature of the top genes in Cluster 7 exhibited a late temporal expression profile,
202 consistent with the timing of macrogamete development.

203 Similar to macrogametes, very little is known about gene expression in *T. gondii*
204 microgametes. To identify these cells in our scRNAseq-derived clusters, we used putative
205 microgamete markers from Ramakrishnan *et al.* including genes involved in axoneme assembly,
206 flagellar transport, cell motility, and fertilization[17]. Mapping these genes onto our scRNAseq
207 projection (Fig. 4F; Supplementary Table 3) revealed a small number of cells across two clusters
208 (9 cells in cluster 8 and 3 cells in cluster 9) that express microgamete marker genes. We recognize
209 the limitation with the number of cells, however, the relative rarity of microgametocytes compared
210 to macrogametes is consistent with the literature[36,37]. The enrichment of previously identified

211 male-specific genes in these clusters supported their annotation as 'Males-1' and 'Males-2',
212 respectively[17] (see below). Importantly, we observed a similar female-biased ratio in a subsequent
213 scRNAseq experiment, where we identified 11 male and 106 female cells (Extended Data Fig. 5A
214 and 5B).

215 The Male-1 cluster expresses Hapless 2 (HAP2), a male-specific gene essential for gamete
216 fusion and subsequent fertilization in both *T. gondii and Plasmodium* (Extended Data Fig.
217 4B)[17,38,39]. In addition, Male-1 are predominantly in in S phase of the cell cycle which may suggest
218 that these are the replicating gametocytes (Extended Data Fig. 7). In contrast, the Male-2 cluster
219 exhibits higher expression of genes related to motility, such as *sperm associated antigen 6* (or
220 *PF16;* TGME49_297820), as well as genes encoding for flagellum-associated proteins (Extended
221 Data Fig. 4C). The *PF16* protein has been observed to associate with the growing cytoplasmic
222 axonemes along the emerging flagella in *T. gondii*[19], and in *Plasmodium berghei*, PbPF16 is
223 expressed in the male gametes[40]. PbPF16 is also essential for maintaining the correct microtubule
224 structure in the axoneme's central apparatus. This shows that PbPF16 plays a critical role in
225 flagellar formation, a process essential for male motility. Hence, it is possible to speculate that the
226 Male-1 cluster represents intracellular stages (either from a developing macrogamont or a
227 microgamete that has invaded a cell in search of the macrogamete), and that the Male-2 cluster
228 consists of extracellular microgametes, potentially searching for host cells with macrogametes.
229 When we compared the top 100 genes from clusters 8 and 9 to the transcriptomic data reported by
230 Ramakrishnan *et al*[17]*.,* we observed a strong concordance with genes enriched in samples from
231 later developmental stages (Fig. 4G and 4H), consistent with the expected timing of
232 microgametocyte development.
233

234 Another branch emerging from the trifurcation point of the pre-gamete cluster contains
235 cluster 10. Marker genes that are significantly enriched within cluster 10 include SRS48Q ($\log_2$FC
236 > 6.7; adj. p = 9.22E-164) and SAG2D ($\log_2$FC > 8.23; adj. p = 0), which are known to be
237 associated with merozoites[17]. Although cluster 10 shares some of the same genes with the Female
238 and Males-1 clusters, such as TGME49_294640 and TGME49_211030, it lacks the expression of
239 sex-specific genes. Specifically, SRS46, AP2X6, and TGME49_293368 are not detected in this
240 cluster. The expression profile of cluster 10 may indicate that these are cells re-entering the
241 merogony cycle after failing to differentiate into macrogametes or microgametes. This could
242 suggest a dynamic process where cells have the potential to shift between pre-gamete and
243 merogony depending on specific regulatory signals. Therefore, we have named this cluster
244 "Transitioning-merozoites" reflecting its intermediate state between pre-gametes and merozoites
245 (Fig. 4A). Furthermore, the top 100 genes from cluster 10 do not show an obvious association with
246 EES1-5 from Ramakrishnan *et al.* (Fig. 4L), suggesting that this is potentially a transient
247 population.

248 We expanded our analysis to investigate the biological processes distinguishing identified
249 male and female clusters by using GO term enrichment analysis on ToxoDB. The top 50 genes in
250 the female cluster were enriched in GO terms such as 'carbohydrate catabolic process'
251 (GO:0016052), 'small molecule metabolic process' (GO:0044281), and 'carboxylic acid metabolic
252 process' (GO:0019752) suggesting unique changes in specific metabolic pathways (Fig. 4I). It is
253 possible that female gametes may rely more on carbohydrate and carboxylic acid catabolism, as
254 well as small-molecule metabolic pathways, potentially to support ATP production and
255 biosynthetic processes required for fertilization and oocyst development. In contrast, the top 50
256 genes in the Male-1 cluster revealed overrepresentation of GO terms related to 'microtubule-based

257 movement' GO:0007018' 'microtubule-based process' (GO:0007017), and 'protein polymerization'
258 (GO:0051258) (Fig. 4J). The prominent GO term, 'microtubule-based process,' similar enrichment
259 was reported in *P. falciparum*[41] microgametes, suggesting a potential role for these genes in
260 shaping microgametes or driving the motility of flagellated male gametes. Furthermore, GO term
261 enrichment analysis of the top 50 genes from the Male-2 cluster revealed enrichment in terms such
262 as 'cilium assembly' (GO:0060271), 'plasma membrane-bounded cell projection assembly'
263 (GO:0120031), and 'cell projection assembly' (GO:0030031), emphasizing the important role of
264 structural and motility-related processes in this male cluster (Fig. 4K).
265

266 ***In vivo* Perturb-seq screening identified AP2X6 as a repressor of asexual enteric**
267 **development in the cat.**

268 To genetically identify novel regulators of merogony and sexual commitment, we
269 conducted a targeted genetic screen, selecting 28 candidate genes based on their distinct
270 transcriptional profiles observed in our scRNA-seq dataset (Extended Fig. 4A). These genes
271 encode DNA- or RNA-binding proteins, surface antigens, and hypothetical proteins, and, more
272 importantly, 23 out of 28 exhibit unique expression profiles in sexual stages. Three different
273 protospacer sequences per gene were cloned into a *T. gondii* Perturb-seq vector using Gibson
274 assembly and validated by whole plasmid sequencing. The validated plasmids were pooled and
275 transfected into TgVEG:GFP:mCh parasites, generating a knockout pool that was used to infect
276 HFFs (Fig. 5A). To minimize bottleneck effects, each gene knockout was propagated as an
277 independent population, resulting in 28 distinct knockout parasite populations. These populations
278 were individually inoculated into mice at doses of 1,000 or 10,000 parasites per mouse (6–8 mice
279 per population). To control parasite growth, sulfadiazine treatment was administered to the mice

280 from 4 to 10-DPI. At 25 DPI, brain tissue containing bradyzoites in tissue cysts was harvested
281 from the chronically infected mice, and equal amounts (~16 g) were fed to two cats. Parasite
282 shedding was monitored, and upon detection parasites were isolated from the small intestine by
283 scraping and cell sorting for fluorescent parasites at 6 and 7 DPI (Fig. 5A).

284 To evaluate the complexity of the library prior to the *in vivo* infection, we performed
285 scRNA sequencing on the transfected parasites after drug selection, detecting protospacers in
286 23,285 of the 37,729 sequenced cells (~62%) and observing a sequencing saturation of 22.3%,
287 consistent with a highly complex library even at the low sequencing depth we used as our goal
288 was only to characterize library complexity and confirm the efficacy of gRNA capture using the
289 10X genomics system. To identify whether the targeted genes induced significant transcriptional
290 changes *in vitro*, we performed differential gene expression (DGE) analysis to compare the
291 expression profiles of each target gene across distinct populations. No significant transcriptional
292 differences were observed based on the presence of any single gRNA in tachyzoites (Extended
293 Data Fig. 8B). This lack of observable significance is consistent with the expression profile of
294 the targeted genes, which are predominantly active during the feline enteric stages and poorly
295 expressed in tachyzoites.

296 In the cat Perturb-Seq sample, 5,060 cells harboring single gRNAs targeting 18 different
297 genes were recovered (~19.5% of cells). This is lower than the expected recovery of cells
298 harboring gRNAs, which could be influenced by multiple factors including a bottleneck within the
299 cat small intestine. Cells expressing multiple gRNAs were filtered out, leaving 4,741 cells with
300 single gRNA sequences for downstream analysis. We found 5 specific perturbations in individual
301 cells that resulted in significant transcriptional changes in multiple genes (Fig. 5B). *AP2X6*

302 perturbation significantly increased transcript abundance for 121 genes relative to parasites with
303 other gRNAs, causing the largest perturbation of the parasite transcriptome across all detected
304 gRNAs. In cells harboring the SRS32-1 gRNA, we observed a significant decrease in transcript
305 abundance in 79 genes (Fig. 5B). Cells carrying the TGME49_311070-2 gRNA, a homolog of *fd1*
306 in *Plasmodium falciparum* 3D7[33] and associated with a female gametocytes (Extended Data Fig.
307 6A), had significant transcriptional changes in 36 genes. Furthermore, cells captured with
308 TGME49_297410 and SRS15B-2 gRNAs were found to have changes in transcript abundance in
309 73 and 5 genes, respectively. Cells harboring the gRNA corresponding to significant perturbations
310 are mapped on the UMAP (Fig. 5C). Some perturbations were observed to impact the same group
311 of genes, exposing a potential link in their function (Fig. 5E). However, follow-up experiments are
312 required to investigate the functions of these genes during enteric development and to confirm that
313 they participate in similar or connected gene regulatory networks.

314 AP2X6 is of a special interest due to its unique expression in the Female cluster.
315 Interrogating the top 50 differentially expressed genes in cells with the gRNAs targeting the female
316 specific *AP2X6* gene shows that the dysregulated genes are expressed primarily during the earlier
317 stages of merozoite development and particularly genes that are characteristic of the Merozoite-A
318 and BRP+ clusters (Extended Fig. 6B). This suggests that AP2X6 acts as a repressor of genes
319 express prior to gametogenesis in females and is potentially required to initiate and/or maintain
320 female gametocyte differentiation.

321 **Discussion**

322 The enteric stages of *T. gondii* are notoriously difficult to investigate due to their restriction
323 to feline hosts and the challenges associated with isolating them from cat intestinal epithelial cells.
324 Traditional bulk RNA sequencing approaches have been used on cat enteric stages[16,17] but these

325 bulk studies do not capture the heterogeneity and asynchronous development of the parasite. In
326 recent years, single-cell transcriptomics has emerged as a powerful tool, addressing these
327 challenges and enabling characterization of complex developmental trajectories in a wide array of
328 organisms including apicomplexan parasites[32,41–43]. In this study, we developed a constitutively
329 active reporter system using a hybrid promoter in *T. gondii* that enabled efficient sorting of enteric-
330 stage parasites from feline intestinal tissue using FACS, followed by scRNA-seq and
331 computational analysis to interrogate the transcriptional program underlying feline-specific
332 development. Using previously reported canonical markers, we identified 10 transcriptionally
333 distinct clusters, including intestinal tachyzoites, merozoites, transitional forms including pre-
334 gametes, and gametocytes. Evaluation of the top 100 genes from gametocyte clusters revealed
335 enrichment for genes found to have higher expression in later stages of development (i.e., EES5
336 samples in Ramakrishnan *et al.*) and supporting the late developmental identity of these clusters[17].
337

338 One of the most striking observations from our dataset is the disproportionate
339 representation of parasites undergoing merogony compared to those committed to sexual
340 development. Gametocytes were significantly outnumbered by merozoites, suggesting that only a
341 small fraction of parasites proceed toward gametogenesis. Our data support a model in which some
342 pre-gametes may re-enter the merogony cycle, forming what we term "transitioning merozoites."
343 Our observations suggest that not all cells within the cat enteric environment progress to
344 gametogenesis. Based on clustering and marker expression profiles, we hypothesize that the
345 majority of cells remain in an asexual enteric replication state rather than committing to sexual
346 development. This indicates that gamete formation may be a restricted or regulated process,
347 potentially requiring specific cues or conditions that are not detected by the majority of *T. gondii*

348 cells during development in cat enterocytes.. Such plasticity highlights the dominance of asexual
349 replication in apicomplexan life cycles and raises important questions about the regulatory
350 mechanisms that govern stage transitions. This skewed distribution is consistent with observations
351 in *P. falciparum*, where in controlled human malaria infections, fewer than 10% of parasites
352 commit to gametocyte formation[44,45]. It is reasonable to speculate that by maintaining a large
353 comparatively large asexual population it can sustain infection and prolong the window of
354 opportunity for mating, maximizing transmission potential. This strategy may be evolutionarily
355 conserved across apicomplexans.

356 Although our dataset constitutes significant progress towards gaining an understanding of
357 the gene expression dynamics in *T. gondii* enteric development, the comparatively low number of
358 male gametes indicate that we have yet to fully determine the spectrum of transcriptional profiles
359 in this life stage. This low number compared to female gametes is consistent with expectations
360 based on observations in *P. falciparum*[46,47]. Some male gametes may also be motile and either
361 invade host cells rapidly or be swept away in the cat gut and/or when we wash the intestine prior
362 to scraping. However, there is little doubt that these small clusters of cells represent male gametes.
363 For example, the centrin family protein (TGME49_237490) is associated with the basal structure
364 and function of flagella[48], sperm-associated antigen 6 (TGME49_297820) is a conserved marker
365 of male gametes, and *HAP2* is a well-characterized protein required for gamete fusion during
366 fertilization[49,40,19,17,31]. The specific and exclusive expression of these genes in the annotated male
367 gamete clusters supports their annotation, despite their comparative low numbers. To address the
368 challenge of low cell abundance, future studies could use male-specific markers identified in this
369 study to enrich for male gametocytes during isolation, allowing for deeper sequencing and
370 potentially functional studies with these poorly understood life stages.

371

372 Through scRNA-seq, we identified 28 candidate genes with potential contribution to
373 various aspects of enteric development and used Perturb-Seq to determine their impact on parasite
374 transcriptional profiles during enteric development. Among the genes with significant
375 transcriptional impact was TGME49_311070, which is a female-specific gene. TGME49_311070
376 encodes a putative RNA-binding protein and is a homolog of the fd1 gene in *P. berghei*, which,
377 when knocked out, results in infertile females lacking transcripts for a unique set of core female
378 markers[33]. AP2X6-targeted cells showed increased transcript abundance for genes that are
379 expressed earlier in development than AP2X6, suggesting that it functions as a repressor of specific
380 genes during cat enteric development. This is consistent with the fact that multiple AP2 factors
381 have been found to be repressors of transcription in *T. gondii*[1,50–52] rather than activators of
382 transcription. The *in vivo* Perturb-seq data provide compelling evidence that *AP2X6* functions as
383 a repressor of merozoite-specific genes in *T. gondii* during its enteric development. While we
384 recognize the low number of cells detected with the *AP2X6* protospacer, the observed
385 transcriptional differences between AP2X6-targeted cells and cells with other protospacers are
386 robust. AP2X6-targeted cells induced upregulation of cells associated mostly with Merozoite-A
387 and BRP1+ clusters. The expression of those genes decreased as we move down the developmental
388 trajectory into Merozoite-B, which aligns with AP2X6 expression which was detected initially in
389 Merozoite-B. The change in the transcriptional profile suggests that the loss of AP2X6 potentially
390 leads to a failure in properly repressing asexual gene expression within the female lineage. This
391 discovery constitutes a step forward in investigating gamete formation, even with the limited cell
392 numbers.

393 Although this screen constitutes progress toward gaining a functional understanding of
394 enteric development, it has some limitations. First, the gRNA capture rate was lower than expected
395 at ~19%. Therefore, we were unable to detect all the gRNAs from the input library. Second, the
396 resources on gene essentiality and lethality are not available for the enteric cycle. gRNA targeting
397 of any number of these genes could have decreased or increased parasite fitness relative to others
398 in the population, leading to poor detection of those gRNAs.

399 In summary, our single-cell data provide valuable insights into the gene expression
400 dynamics during the enteric development of *T. gondii* in cats. This work holds the potential to
401 uncover previously unknown genetic drivers of differentiation, particularly those involved in
402 gametogenesis. These findings contribute to ongoing efforts in the field to induce gametogenesis,
403 mating, and oocyst formation *in vitro*, paving the way for future progress in manipulating the
404 parasite's life cycle.

405

406

407 **Materials and methods**

408 **Parasite strain and growth conditions:**

409 *TgVEG* parasites were used in this study. These were early passage parasites (P6) that were
410 excysted from sporulated oocysts as described[53], and maintained in human foreskin fibroblast
411 (HFF) cultures incubated at 37°C and 5% CO in Dulbecco's Modified Eagle Medium (DMEM;
412 Gibco) supplemented with 10% fetal bovine serum (FBS) (Atlas Biologicals), 2 mM L-glutamine
413 (Thermo, Fisher Scientific, Waltham, MA), and 50 mg/mL penicillin-streptomycin and
414 cryopreserved in 10% DMSO.

415

416 **Plasmid construction**

417 The TgVEG:GFP:mCh reporter strain was generated by introducing a cassette containing
418 a truncated pGRA11A promoter upstream of mCherry into the pClick-GFP-Luc vector[54]. All
419 primers used in this study are listed in Supplementary Table 7. The pTrunGRA11A was amplified
420 from *TgVEG* genomic DNA, and mCherry was amplified from pCas9.T2A.HXGPRT[55]. The
421 amplified products of pTrunGRA11A, mCherry, and linearized pClick-GFP-Luc were combined
422 in a Gibson assembly reaction master mix (NEB E2621S) at a molar ratio of 0.014 pmol. The
423 Gibson reaction was incubated at 50°C for 1 hour. Plasmid sequences were verified by Oxford
424 Nanopore sequencing.

425 **Generation of a *T. gondii* VEG reporter strain**

426 *TgVEG* was transfected with the reporter cassette using standard approaches. Briefly,
427 parasites were mechanically released by needle passage (25 and 27-gauge needles) and then
428 pelleted (10 minutes at 800 x g). Approximately 2 x 10^7 parasites were re-suspended in 100µl
429 Amaxa Basic Parasite Nucleofector 1 (Lonza) along with 3µg of circular pSAG1::CAS9-
430 U6::sgUPRT (Addgene plasmid #54467)[56] and 5µg of the pTrunGRA11A-mCherry-pGRA1-GFP
431 cassette flanked by 20 bp of homology regions to the UPRT locus were co-transfected by
432 electroporation (Supplementary Table 7). Transgenic and FUDR-resistant parasites were obtained
433 by selection with 10 µM FUDR (Sigma-Aldrich, St. Louis, MO). We performed a diagnostic PCR
434 to verify homologous integration and gene deletion.

435 **Cat infections, oocyst isolation, and excystations**

436 For cat TgVEG:GFP:mCh parasites were released from HHF culture mechanically using
437 25 and 27-gauge needles. The released parasites were then pelleted at 800 x g for 10 minutes, then

438 resuspended in PBS. Ten Mice (CBA/J from Jackson lab) were infected with 1,000 parasites in
439 200 ul PBS intraperitoneally (IP). The levels of IFN-γ were used as an indication of parasite
440 infection and burden early in infection. Blood serum from infected mice was obtained via
441 submandibular bleed, centrifuged at 100 × g for 5 minutes, and then diluted at a 1:20 ratio. IFN-γ
442 concentrations were determined using BD OptEIA Mouse IFN-γ (AN-18) ELISA Set (Cat. #
443 551866) according to the manufacturer's instructions.
444
445 Mice were treated with Sulfadiazine (0.4g/L *ad libitum* in drinking water) from 6-10 DPI
446 to increase survival during the acute phage of infection. Throughout the course of infection, the
447 mice were monitored and weighed regularly. At 21DPI, a mouse was euthanized to determine the
448 cyst burden. The brain was harvested and homogenized in 1 ml of PBS (Sigma) and stained with
449 Biotin-conjugated Dolichos biflorus agglutinin (DBA) (1:500; Vector Laboratories, cat. no. RL-
450 1032) for 1 hour at room temperature. Subsequently, the sample was incubated with streptavidin
451 conjugated to Alexa Fluor 647. Cysts were visualized and counted using a 40x objective on an
452 Olympus IX83 microscope using cellSens software.
453
454 The remaining mice were euthanized, and the deskinned skulls and brains were fed to cats.
455 Feces were collected between days 3 and 17 post-infection and examined for oocyst production
456 using sucrose fecal floats. The collected feces were submerged in water and mixed with a wooden
457 tongue depressor to break down the materials. The emulsified mixture was then filtered through
458 840 μm and 250 μm sieves to further break down the particles. The filtered liquid was poured into
459 225 mL tubes and centrifuged at 2,000 × g for 10 minutes. The resulting pellet was resuspended
460 in a 1.56 M sucrose solution (30 mL sucrose per tube per 10 g of feces). The sucrose mixture was

461 centrifuged at 1,180 × g for 10 minutes, and the top 10 mL of the supernatant was collected for
462 oocyst quantification using hemocytometer. This oocyst collection step was repeated at least twice
463 to minimize oocyst loss during processing. The isolated sucrose pellet containing oocysts was
464 mixed with water at five times the volume and centrifuged at 1,180 × g for 10 minutes. The pellet
465 was then resuspended in 2% H_2SO_4 and kept at room temperature with shaking for 7 days to allow
466 for sporulation. These samples were stored at 4°C.
467
468 To excyst *T. gondii* oocysts, sporulated oocysts were treated with 1N NaOH (at a ratio of
469 3:5 oocyst volume) to neutralize the 2% H_2SO_4 solution, and then washed multiple times with
470 Hanks' balanced salt solution (HBSS; Life Technologies, 14175145) until the pH was ~7.2. pH
471 of the solution was measured using pH indicator strips. The oocysts were then pelleted at 1,000 x
472 g for 8 minutes at 24°C and resuspended in PBS to prepare for excystation in mice. CBA/J mice
473 were infected with 100 oocysts orally or IP. On days 4-7 post-infection, mice were euthanized,
474 and peritoneal lavage fluids were collected in PBS. Lavage fluid was palleted by centrifugation,
475 resuspended in cDMEM, and then placed on confluent HHFs monolayers for 18 hours. Media was
476 changed as necessary, and parasites were visualized using light and fluorescence microscopy.

477 **Single-cell transcriptomics from *T. gondii* cultured *in vitro***

478 For this experiment, we used parasites that were cultured *in vitro* for 8 days post-
479 excystation. The sporozoites were excysted from oocysts as previously described [57,58]. Post-
480 excystation, the sporozoites were resuspended in culture media and used to infect confluent
481 monolayers of HHFs, grown at 37°C and 5% CO_2. The parasites were passaged as necessary. At
482 8 days post-excystation, the parasites were counted and processed for library preparation using
483 10x Genomics Chromium technology. The libraries were then sequenced on a HiSeq 4000

484 platform (Illumina) with a targeted depth of approximately 30,000 reads per cell for the 3' mRNA
485 library.

486 **Single cell transcriptomics of *T. gondii* tachyzoites and cat enteric stages.**

487 For single-cell transcriptomics, cats were euthanized at 5 and 6-DPI. The small intestine
488 was isolated from the euthanized cats and divided into three segments. The small intestine was
489 then washed with PBS and opened. The parasites were isolated from the small intestine of the cat
490 by scraping the intestinal lining with glass microscope slide (Fisherbrand: cat#12-550-112). The
491 scraped materials were syringe-passaged through 18, 22, 25, and 27-gauge needles, followed by
492 sequential filtration through cell strainers (100 µm and 40 µm, Corning). The strained materials
493 were pelleted at 1,000 x g for 8 minutes and resuspended in PBS with 1% FBS. The
494 *Tg*VEG:GFP:mCh parasites in the scraped materials were then sorted based on mCherry-derived
495 fluorescence expression. Sorting was performed using BD FACSAria. The sorted parasites cells
496 were concentrated by centrifugation at 1,000 x g for 8 minutes at 4°C and resuspended in PBS
497 containing 1% FBS (Sigma). The cells were then processed for library preparation using 10x
498 Genomics Chromium technology. The libraries were then sequenced on a HiSeq 4000 platform
499 (Illumina) with a targeted depth of approximately 50,000 reads per cell for the 3' mRNA library.

500 **scRNA-seq data processing**

501 Sequencing reads from the 3' mRNA libraries were processed using Cell Ranger v8.0.1
502 (10x Genomics). Reads were aligned to a custom *Toxoplasma gondii* reference genome (ME49,
503 based on release 56), including intronic reads, while retaining all other settings at their default
504 values. The custom library also included mCherry fused to the 3' end of DHFR and GFP fused to

505 the 3' end of GRA2, as these were components of the reporter cassette. We sequenced 16,695 cells
506 with 42,871 mean reads per cell; with a median UMI of 561 counts per cell. Cell calling was
507 performed using the two-stage algorithm implemented in Cell Ranger, which distinguishes cell-
508 associated barcodes from background noise such as ambient RNA from lysed or dead cells. The
509 barcode rank plot was used to visualize this separation, and the knee point guided the selection of
510 a UMI threshold to exclude low-quality cells and empty droplets. EmptyDrops was additionally
511 applied to refine cell calling. The resulting UMI count matrices were imported into Seurat v5.1.0
512 or Loupe Browser v8.0.0 for downstream analysis. A UMI threshold of 500 UMIs per cell was
513 applied to filter low-quality cells. Doublets were identified and removed using the scDblFinder
514 algorithm. The neighborhood size parameter (pK) was empirically optimized to maximize
515 classification accuracy, and artificial doublets were simulated from a subset of cells.
516 Dimensionality reduction was performed using principal component analysis, with the number of
517 components selected by the elbow method. Clustering analysis was performed using the gap
518 statistics method to determine the optimal number of clusters in R.

519

520 **Generation of perturb-seq vector**

521 Capture sequence 1 (GCTTTAAGGCCGGTCCTAGCAAG) was inserted into the
522 pCAS9-T2A-DHFR vector using Q5 mutagenesis (NEB) to generate pCAS9-T2A-DHFR-CS1
523 (TgPerturb-seq vector) and the perturb-seq vector was verified by nanopore sequencing. To
524 generate the pool of perturb-seq vectors used in the screen, single-stranded DNA (ssDNA)
525 oligonucleotides encoding the protospacer sequences for each gene of interest were generated
526 using EuPaGDT[59] . Three ssDNA were synthesized for each gene with 20 bases of homology to
527 either side of the perturb-seq vector integration site (Supplementary Table 5). Each ssDNA was

528 integrated into the linearized TgPerturb-seq vector (NcoI and PacI) by Gibson assembly (vector:
529 insert = 1:5). The Gibson reaction was incubated for 30 mins at 50 °C.

530 **Perturb-seq experiment**

531 *Tg*VEG:GFP:mCh parasites were transfected with NotI-HF-digested (NEB) perturb-seq
532 vectors by electroporation using Amaxa Basic Parasite Nucleofector 1 (Lonza). For each
533 transfection, approximately 10^7 parasites were electroporated with 15 µg of pooled Perturb-seq
534 vectors, composed of three vectors expressing different gRNA targeting the same gene (5 µg each).
535 Twenty-four hours post-transfection, 3 µM pyrimethamine (Sigma-Aldrich, St. Louis, MO) was
536 added to the media to select for integration of the perturb-seq vectors into the genome for 12 days.
537 The diversity of perturb-seq vectors was validated by amplifying the gRNAs, followed by
538 sequencing the amplicons.

539 Mice were infected IP with 1,000 or 10,000 parasites in 200 µl PBS. The infected mice
540 were treated with Sulfadiazine (0.2g/L) from 4DPI to 10DPI. The treatment was employed to
541 reduce the parasite burden. They were monitored and weighed regularly throughout the
542 experiments. To assess the cyst burden in the brains of infected mice, two mice were euthanized
543 at 21DPI for analysis. The brains were homogenized in 1 ml PBS and subsequently stained with
544 Biotin-conjugated *Dolichos biflorus* agglutinin (DBA) (1:500; Vector Laboratories, cat. no. RL-
545 1032) for 1 hour at room temperature, followed by incubation with Biotin-conjugated to Alexa
546 Fluor 647. Cysts were visualized and counted using a 40x objective on an Olympus IX83
547 microscope using cellSens software.

548 Upon cyst verification, brain tissues from the remaining mice were collected and fed to
549 two pathogen-free cats. Feces were collected from 3 DPI until oocysts were detected at 5 DPI. On

550 6 and 7 DPI, the cats were euthanized, and parasites were isolated from the small intestine and
551 sorted as previously described above.

552 The pooled sorted cells were concentrated by centrifugation at 1,000 x g for 8 minutes at
553 4°C and resuspended in PBS containing 1% FBS (Sigma). The cells were then processed for library
554 preparation using 10x Genomics Chromium technology. The libraries were then sequenced on a
555 HiSeq 4000 platform (Illumina) with a targeted depth of approximately 45,000 reads per cell for
556 the 3' mRNA library and 10,000 reads per cell for the CRISPR sgRNA library.

557 **Perturb-seq data processing**

558 The sequencing reads from the 3' mRNA and CRISPR sgRNA libraries were processed
559 using Cell Ranger v7.2.0 (10x Genomics). The reads were aligned to both a custom *T. gondii*
560 transcriptome reference and another reference containing protospacer sequences. Intronic reads
561 were included in the alignment, while all other parameters were left at their default settings. Cell
562 Ranger assigned the reads to specific cell barcodes, mapped them to the corresponding genes or
563 protospacers, and counted the unique molecular identifiers (UMIs) for each gene or protospacer
564 within each cell barcode. This generated UMI × cell barcode count matrices, which were
565 downloaded for downstream analysis to identify gRNA-specific perturbations using in the single-
566 cell transcriptome Loupe Browser v8.0.0 and Seurat v5.1.0. In total, we sequenced 26,000 cells
567 with a median depth of 1,996 UMI. 5060 cells were detected with one or more protospacer. For
568 analysis, we filtered cells with multiple protospacers to retains cells with a single protospacer
569 detected, 4741 cells, for further analysis.
570

571 **Animal statement**

572 The animal experiments were performed according to the guidelines of the American
573 Veterinary Association. Euthanasia for mice was conducted using controlled CO_2 exposure.
574 Euthanasia for cats was conducted using Beuthanasia. Animal protocols were approved under the
575 Institutional Animal Care and Use Committee (IACUC) protocol no. 22061169.

576 References

577 1. Antunes, A. V. *et al.* In vitro production of cat-restricted Toxoplasma pre-sexual stages. *Nature*
578 **625**, 366–376 (2024).
579 2. Di Genova, B. M., Wilson, S. K., Dubey, J. P. & Knoll, L. J. Intestinal delta-6-desaturase
580 activity determines host range for Toxoplasma sexual reproduction. *PLoS Biol.* **17**, e3000364
581 (2019).
582 3. Frenkel, J. K. & Dubey, J. P. Toxoplasmosis and its prevention in cats and man. *J. Infect. Dis.*
583 **126**, 664–673 (1972).
584 4. Dubey, J. P. Oocyst Shedding by Cats Fed Isolated Bradyzoites and Comparison of Infectivity
585 of Bradyzoites of the VEG Strain Toxoplasma gondii to Cats and Mice. *J. Parasitol.* **87**, 215
586 (2001).
587 5. Dubey, J. P. TACHYZOITE-INDUCED LIFE CYCLE OF TOXOPLASMA GONDII IN
588 CATS. *J. Parasitol.* **88**, 713–717 (2002).
589 6. Dubey, J. P. Unexpected oocyst shedding by cats fed Toxoplasma gondii tachyzoites: In vivo
590 stage conversion and strain variation. *Vet. Parasitol.* **133**, 289–298 (2005).
591 7. Sokol-Borrelli, S. L. *et al.* A transcriptional network required for bradyzoite development in
592 Toxoplasma gondii is dispensable for recrudescent disease. *Nat. Commun.* **14**, 6078 (2023).

593 8. Waldman, B. S. *et al.* Identification of a Master Regulator of Differentiation in Toxoplasma.
594 *Cell* **180**, 359-372.e16 (2020).
595 9. Radke, J. Defining the cell cycle for the tachyzoite stage of Toxoplasma gondii. *Mol. Biochem.*
596 *Parasitol.* **115**, 165–175 (2001).
597 10. Radke, J. R., Guerini, M. N., Jerome, M. & White, M. W. A change in the premitotic period
598 of the cell cycle is associated with bradyzoite differentiation in Toxoplasma gondii. *Mol.*
599 *Biochem. Parasitol.* **131**, 119–127 (2003).
600 11. Blume, M. & Seeber, F. Metabolic interactions between Toxoplasma gondii and its host.
601 *F1000Research* **7**, 1719 (2018).
602 12. Gallego-López, G. M., Contreras Guzman, E., Desa, D. E., Knoll, L. J. & Skala, M. C.
603 Metabolic changes in *Toxoplasma gondii* -infected host cells measured by autofluorescence
604 imaging. *mBio* **15**, e00727-24 (2024).
605 13. Hakimi, M., Olias, P. & David, L. Toxoplasma Effectors Targeting Host Signaling and
606 Transcription. *Clin. Microbiol. Rev.* **30**, 615-645p38 (2017).
607 14. Dubey, J. P. & Frenkel, J. K. Cyst-Induced Toxoplasmosis in Cats*. *J. Protozool.* **19**, 155–
608 177 (1972).
609 15. Behnke, M. S., Zhang, T. P., Dubey, J. P. & Sibley, L. D. Toxoplasma gondii merozoite gene
610 expression analysis with comparison to the life cycle discloses a unique expression state during
611 enteric development. *BMC Genomics* **15**, (2014).
612 16. Hehl, A. B. *et al.* Asexual expansion of Toxoplasma gondii merozoites is distinct from
613 tachyzoites and entails expression of non-overlapping gene families to attach, invade, and
614 replicate within feline enterocytes. *BMC Genomics* **16**, 1–16 (2015).

615 17. Ramakrishnan, C. *et al.* An experimental genetically attenuated live vaccine to prevent
616 transmission of Toxoplasma gondii by cats. *Sci. Rep.* **9**, 1–14 (2019).
617 18. Martorelli Di Genova, B., Wilson, S. K., Dubey, J. P. & Knoll, L. J. Intestinal delta-6-
618 desaturase activity determines host range for Toxoplasma sexual reproduction. *PLOS Biol.* **17**,
619 e3000364 (2019).
620 19. Farhat, D. C. *et al.* A MORC-driven transcriptional switch controls Toxoplasma
621 developmental trajectories and sexual commitment. *Nat. Microbiol.* **5**, 570–583 (2020).
622 20. Ferguson, D. J. P. Use of molecular and ultrastructural markers to evaluate stage conversion
623 of Toxoplasma gondii in both the intermediate and definitive host. *Int. J. Parasitol.* **34**, 347–
624 360 (2004).
625 21. Manger, I. D., Hehl, A. B. & Boothroyd, J. C. The Surface of *Toxoplasma* Tachyzoites Is
626 Dominated by a Family of Glycosylphosphatidylinositol-Anchored Antigens Related to
627 SAG1. *Infect. Immun.* **66**, 2237–2244 (1998).
628 22. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet
629 identification in single-cell sequencing data using scDblFinder. *F1000Research* **10**, 979
630 (2022).
631 23. Sun, R. *et al.* TREM2 inhibition triggers antitumor cell activity of myeloid cells in
632 glioblastoma. *Sci. Adv.* **9**, eade3559 (2023).
633 24. participants in the 1st Human Cell Atlas Jamboree *et al.* EmptyDrops: distinguishing cells
634 from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63
635 (2019).
636 25. Tibshirani, R., Walther, G. & Hastie, T. Estimating the Number of Clusters in a Data Set Via
637 the Gap Statistic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63**, 411–423 (2001).

638 26. Guo, J. *et al.* Integrating imaging and genomic data for the discovery of distinct glioblastoma
639 subtypes: a joint learning approach. *Sci. Rep.* **14**, 4922 (2024).
640 27. Schwarz, J. A., Fouts, A. E., Cummings, C. A., Ferguson, D. J. P. & Boothroyd, J. C. A novel
641 rhoptry protein in Toxoplasma gondii bradyzoites and merozoites. *Mol. Biochem. Parasitol.*
642 **144**, 159–166 (2005).
643 28. Najm, R. *et al.* Invasion of *Toxoplasma gondii* bradyzoites: Molecular dissection of the moving
644 junction proteins and effective vaccination targets. *Proc. Natl. Acad. Sci.* **120**, e2219533120
645 (2023).
646 29. Schwarz, J. A., Fouts, A. E., Cummings, C. A., Ferguson, D. J. P. & Boothroyd, J. C. A novel
647 rhoptry protein in Toxoplasma gondii bradyzoites and merozoites. *Mol. Biochem. Parasitol.*
648 **144**, 159–166 (2005).
649 30. Ramakrishnan, C., Walker, R. A., Eichenberger, R. M., Hehl, A. B. & Smith, N. C. The
650 merozoite-specific protein, TgGRA11B, identified as a component of the Toxoplasma gondii
651 parasitophorous vacuole in a tachyzoite expression model. *Int. J. Parasitol.* **47**, 597–600
652 (2017).
653 31. Walker, R. A. *et al.* RNA Seq analysis of the Eimeria tenella gametocyte transcriptome reveals
654 clues about the molecular basis for sexual reproduction and oocyst biogenesis. *BMC Genomics*
655 **16**, 1–20 (2015).
656 32. Walzer, K. A. *et al.* Transcriptional control of the Cryptosporidium life cycle. *Nature* **630**,
657 174–180 (2024).
658 33. Russell, A. J. C. *et al.* Regulators of male and female sexual development are critical for the
659 transmission of a malaria parasite. *Cell Host Microbe* **31**, 305-319.e10 (2023).

660 34. Mlambo, G., Coppens, I. & Kumar, N. Aberrant Sporogonic Development of Dmc1 (a Meiotic
661 Recombinase) Deficient Plasmodium berghei Parasites. *PLoS ONE* **7**, e52480 (2012).
662 35. Bushkin, G. G. *et al.* Evidence for a Structural Role for Acid-Fast Lipids in Oocyst Walls of
663 *Cryptosporidium* , *Toxoplasma* , and *Eimeria*. *mBio* **4**, e00387-13 (2013).
664 36. Ferguson, D. J. P. Toxoplasma gondii and sex: essential or optional extra? *Trends Parasitol.*
665 **18**, 351–355 (2002).
666 37. Reece, S. E. & Read, A. F. Malaria sex ratios. *Trends Ecol. Evol.* **15**, 259–260 (2000).
667 38. Liu, Y. *et al.* The conserved plant sterility gene *HAP2* functions after attachment of fusogenic
668 membranes in *Chlamydomonas* and *Plasmodium* gametes. *Genes Dev.* **22**, 1051–1068 (2008).
669 39. Angrisano, F. *et al.* Targeting the Conserved Fusion Loop of HAP2 Inhibits the Transmission
670 of Plasmodium berghei and falciparum. *Cell Rep.* **21**, 2868–2878 (2017).
671 40. Straschil, U. *et al.* The Armadillo Repeat Protein PF16 Is Essential for Flagellar Structure and
672 Function in Plasmodium Male Gametes. *PLoS ONE* **5**, e12901 (2010).
673 41. Mohammed, M. *et al.* Single-cell transcriptomics reveal transcriptional programs underlying
674 male and female cell fate during Plasmodium falciparum gametocytogenesis. *Nat. Commun.*
675 **15**, 7177 (2024).
676 42. Xue, Y. *et al.* A single-parasite transcriptional atlas of Toxoplasma Gondii reveals novel
677 control of antigen expression. *eLife* **9**, e54129 (2020).
678 43. Real, E. *et al.* A single-cell atlas of Plasmodium falciparum transmission through the mosquito.
679 *Nat. Commun.* **12**, 3196 (2021).
680 44. Collins, K. A. *et al.* A controlled human malaria infection model enabling evaluation of
681 transmission-blocking interventions. *J. Clin. Invest.* **128**, 1551–1562 (2018).

682 45. Reuling, I. J. *et al.* A randomized feasibility trial comparing four antimalarial drug regimens
683 to induce Plasmodium falciparum gametocytemia in the controlled human malaria infection
684 model. *eLife* **7**, e31549 (2018).
685 46. Schall, J. J. The sex ratio of *Plasmodium* gametocytes. *Parasitology* **98**, 343–350 (1989).
686 47. Gruenberg, M. *et al.* qRT-PCR versus IFA-based Quantification of Male and Female
687 Gametocytes in Low-Density Plasmodium falciparum Infections and Their Relevance for
688 Transmission. *J. Infect. Dis.* **221**, 598–607 (2020).
689 48. Francia, M. E., Dubremetz, J.-F. & Morrissette, N. S. Basal body structure and composition in
690 the apicomplexans Toxoplasma and Plasmodium. *Cilia* **5**, 3 (2015).
691 49. Smith, E. F. & Lefebvre, P. A. PF16 encodes a protein with armadillo repeats and localizes to
692 a single microtubule of the central apparatus in Chlamydomonas flagella. *J. Cell Biol.* **132**,
693 359–370 (1996).
694 50. Radke, J. B. *et al.* ApiAP2 transcription factor restricts development of the *Toxoplasma* tissue
695 cyst. *Proc. Natl. Acad. Sci.* **110**, 6871–6876 (2013).
696 51. Fan, F., Xue, L., Yin, X., Gupta, N. & Shen, B. AP2XII-1 is a negative regulator of merogony
697 and presexual commitment in Toxoplasma gondii. *mBio* https://doi.org/10.1128/mbio.01785-
698 23 (2023) doi:10.1128/mbio.01785-23.
699 52. Srivastava, S., Holmes, M. J., White, M. W. & Sullivan, W. J. Toxoplasma gondii AP2XII-2
700 Contributes to Transcriptional Repression for Sexual Commitment. *mSphere* **8**, e00606-22
701 (2023).
702 53. Sokol, S. L. *et al.* Dissection of the in vitro developmental program of Hammondia hammondi
703 reveals a link between stress sensitivity and life cycle flexibility in Toxoplasma gondii. *eLife*
704 **7**, e36491 (2018).

705 54. English, E. D. & Boyle, J. P. Impact of Engineered Expression of Mitochondrial Association
706 Factor 1b on *Toxoplasma gondii* Infection and the Host Response in a Mouse Model. *mSphere*
707 **3**, e00471-18 (2018).
708 55. Young, J. *et al.* A CRISPR platform for targeted in vivo screens identifies Toxoplasma gondii
709 virulence factors in mice. *Nat. Commun.* **10**, 3963 (2019).
710 56. Shen, B., Brown, K. M., Lee, T. D. & Sibley, L. D. Efficient Gene Disruption in Diverse
711 Strains of Toxoplasma gondii Using CRISPR/CAS9. https://doi.org/10.1128/mBio.01114-14
712 (2014) doi:10.1128/mBio.01114-14.
713 57. Sokol, S. L. *et al.* Dissection of the in vitro developmental program of Hammondia hammondi
714 reveals a link between stress sensitivity and life cycle flexibility in Toxoplasma gondii. *eLife*
715 **7**, e36491 (2018).
716 58. Sokol, S. L., Wong, Z. S., Boyle, J. P. & Dubey, J. P. Generation of Toxoplasma gondii and
717 Hammondia hammondi Oocysts and Purification of Their Sporozoites for Downstream
718 Manipulation. in *Toxoplasma gondii* (ed. Tonkin, C. J.) vol. 2071 81–98 (Springer US, New
719 York, NY, 2020).
720 59. Peng, D. & Tarleton, R. EuPaGDT: a web tool tailored to design CRISPR guide RNAs for
721 eukaryotic pathogens. *Microb. Genomics* **1**, (2015).
722

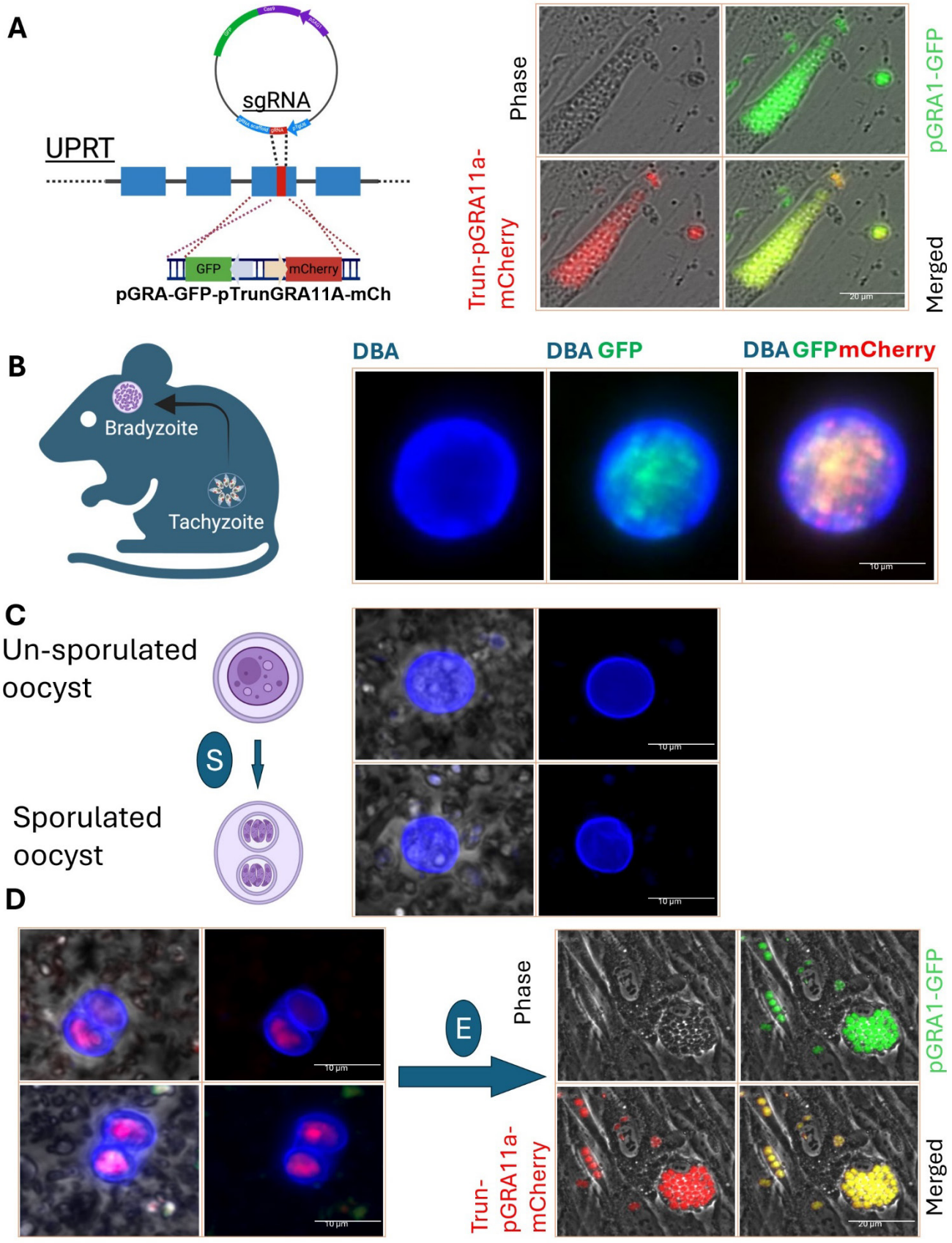
A
sgRNA
UPRT
GFP
mCherry
pGRA-GFP-pTrunGRA11A-mCh
Phase
pGRA1-GFP
Trun-pGRA11a-mCherry
Merged
B
Bradyzoite
Tachyzoite
DBA
DBA GFP
DBA GFP mCherry
C
Un-sporulated oocyst
S
Sporulated oocyst
D
E
Phase
pGRA1-GFP
Trun-pGRA11a-mCherry
Merged

Fig. 1: Generation of a cat-competent *Toxoplasma gondii* reporter strain.
A) A well-characterized oocyst-forming strain, TgVEG, was used in this study. A dual-reporter cassette was constructed with GFP driven by the *GRA1* promoter and mCherry under the control of a truncated *GRA11A* promoter. Both promoters were active in *T. gondii* tachyzoites. **B)** Tissue cysts derived from the TgVEG reporter strain were stained with Dolichos biflorus agglutinin (DBA; blue) and expressed both mCherry and GFP. **C)** TgVEG oocysts isolated from a cat experimentally infected with TgVEG reporter strain tissue cysts exhibited strong UV autofluorescence but lacked detectable GFP or mCherry expression. **D)** Sporulated oocysts of the TgVEG reporter strain expressed mCherry but not GFP. **E)** Sporulated oocysts were used to infect mice. Parasites excysted in the peritoneal cavity and were placed into culture, where they again expressed both mCherry and GFP.

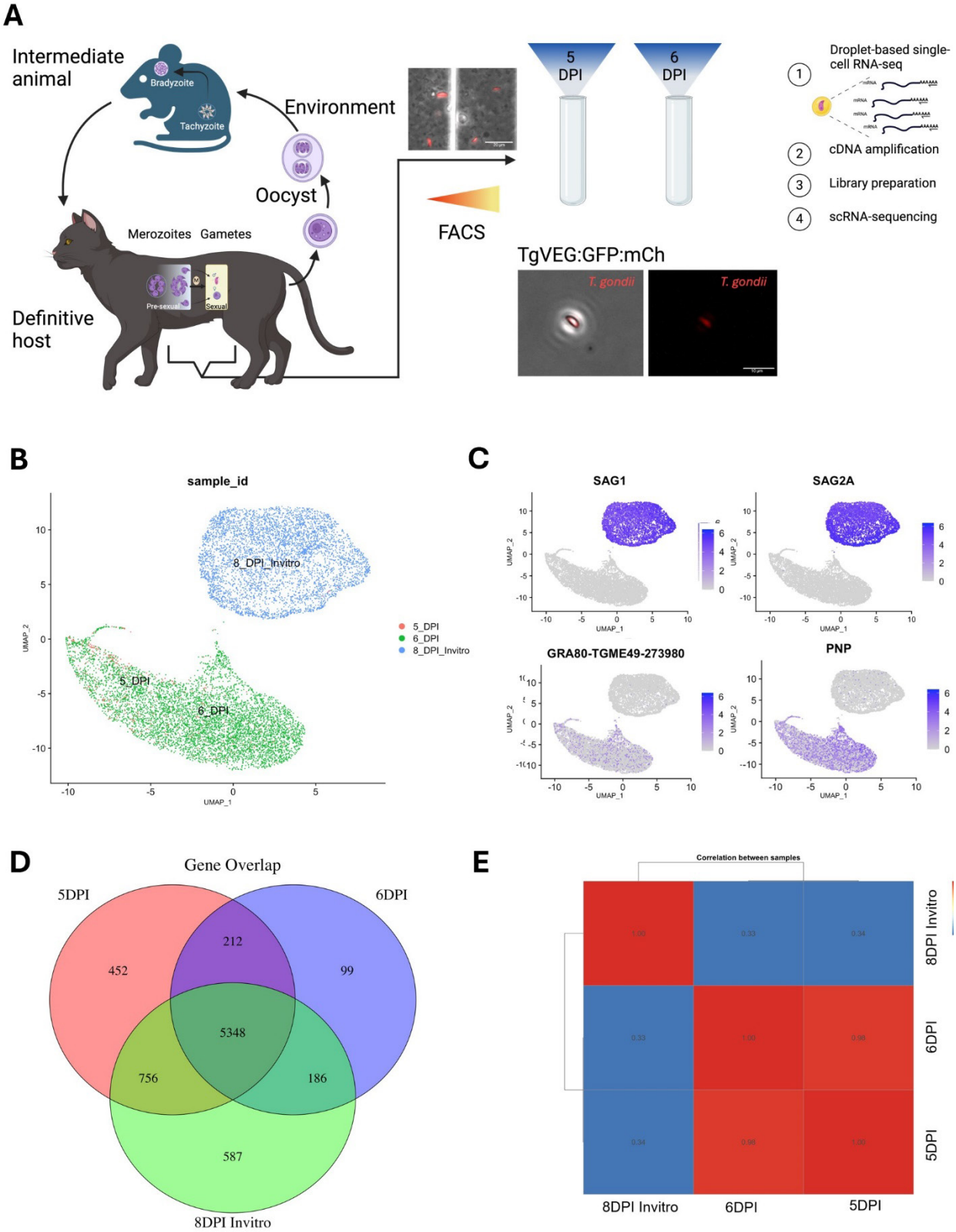
A
Intermediate animal
Bradyzoite
Tachyzoite
Environment
Oocyst
Merozoites
Gametes
Pre-sexual
Sexual
Definitive host
FACS
5 DPI
6 DPI
TgVEG:GFP:mCh
T. gondii
T. gondii
Droplet-based single-cell RNA-seq
1
2 cDNA amplification
3 Library preparation
4 scRNA-sequencing
B
sample_id
8_DPI_Invitro
5_DPI
6_DPI
UMAP_1
UMAP_2
C
SAG1
SAG2A
GRA80-TGME49-273980
PNP
D
Gene Overlap
5DPI
6DPI
8DPI Invitro
452
212
99
5348
756
186
587
E
Correlation between samples
8DPI Invitro
6DPI
5DPI
1.00
0.33
0.34
0.98

Fig. 2: scRNA-seq transcriptome landscape of *Toxoplasma gondii* sexual and asexual development. **A)** Experimental workflow: Mice were infected with the TgVEG reporter strain for 25 days. Brain tissue containing bradyzoite cysts from 28 chronically infected mice was fed to cats (n = 2). Enteroepithelial stages were isolated from the small intestine of a cat euthanized at 5–6 days post-infection (DPI), sorted by mCherry fluorescence, and subjected to single-cell RNA sequencing (10x Genomics). **B)** UMAP visualization of scRNA-seq data from cat enteric stages (D5 and D6 DPI) and tachyzoites cultured for 8 days post-excystation. Distinct clustering reflects transcriptional differences between developmental stages. **C)** Expression profiles of canonical tachyzoite markers (top) and merozoite markers (bottom), highlighting stage-specific transcript abundance. **D)** Venn diagram showing overlap of expressed transcripts among 5 DPI, 6 DPI, and 8 DPI samples. Shared and unique transcripts are indicated by overlapping and non-overlapping regions. **E)** Correlation analysis of transcriptomes across samples, illustrating the degree of similarity between *in vitro* tachyzoites and enteric stages at 5 and 6 DPI.

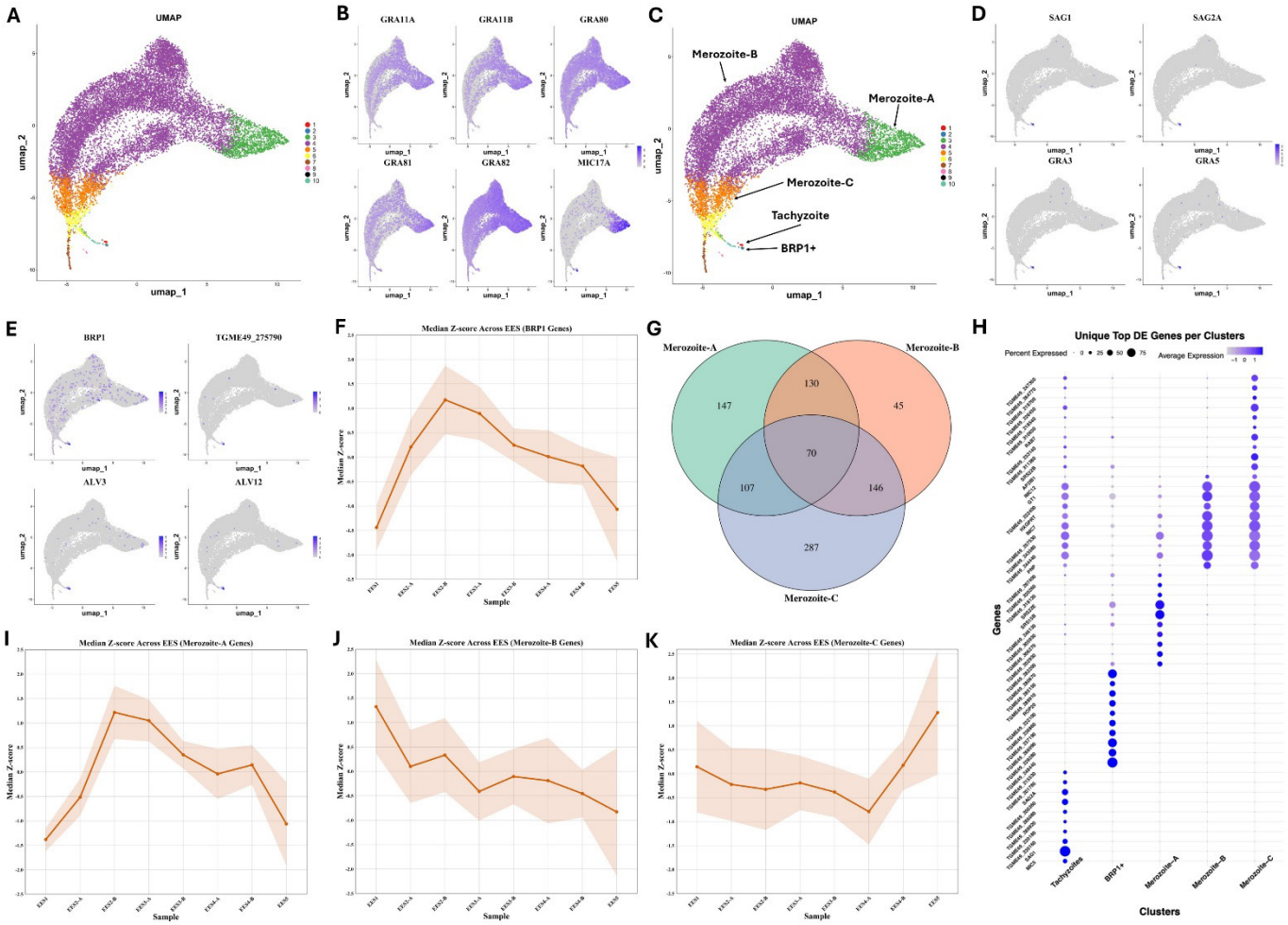

Figure 3. Single-cell transcriptomic profiling of *T. gondii* enteric stages.

A) UMAP visualization of scRNA-seq data from cat enteroepithelial stages at 5 and 6 days post-infection (DPI) (n = 15,068 cells). Predicted doublets were removed using scDblFinder. Dimensionality reduction was performed with the top 25 principal components prior to UMAP embedding, and cluster number was estimated using the gap statistic. **B)** Feature plots showing expression of canonical markers for enteroepithelial stages. **C)** Cluster annotation based on top differentially expressed genes. **D)** Feature plot of Cluster 1 highlighting expression of canonical tachyzoite markers.**E)** Cluster 2 annotated as BRP1+ based on enrichment of BRP1-associated genes. **F)** Line plot of median z-scores for the top 100 genes defining the BRP1+ cluster across EES samples from Ramakrishnan *et al.* Shaded ribbons indicate ±1 SD. **G)** Venn diagram of significantly expressed genes shared and unique to Merozoite-A, -B, and -C clusters. **H)** Dot plot of the top 10 differentially expressed genes for each cluster. Dot size indicates the proportion of cells expressing the gene; color intensity reflects average expression. **I–K)** Line plots of median z-scores for the top 100 genes defining Merozoite-A, -B, and -C clusters across EES samples. Shaded ribbons represent ±1 SD.

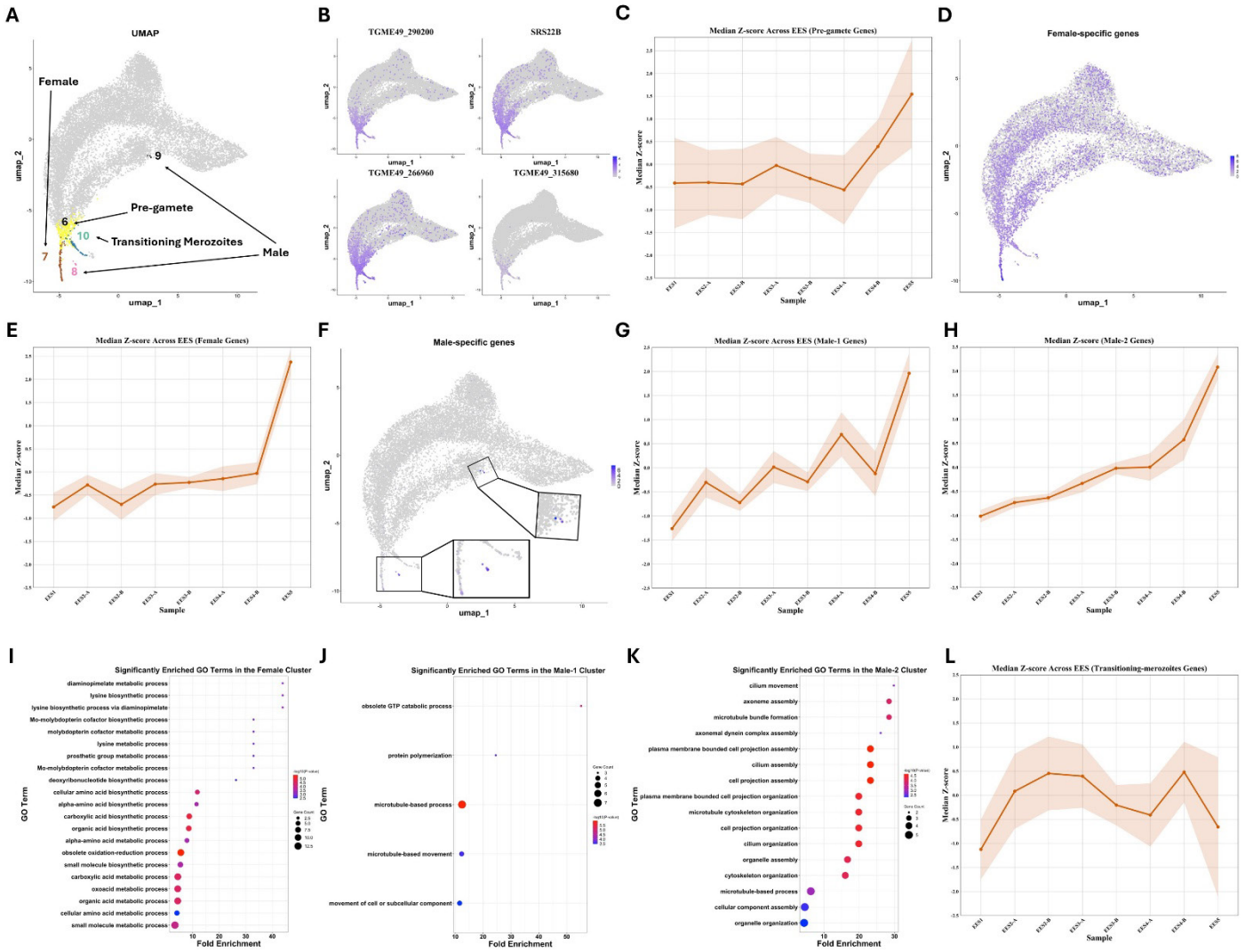

Fig. 4: Transcriptomic signatures of pre-gamete and gamete-associated clusters in *T. gondii*.

A) Annotation of Clusters 6–10 based on top differentially expressed genes. **B)** Feature plots showing genes enriched in the pre-gamete cluster. **C)** Median z-scores of the top 100 pre-gamete cluster genes across EES samples from Ramakrishnan *et al.* Lines represent median expression per sample; shaded ribbons indicate ±1 SD. **D)** 106 female-specific genes identified in *T. gondii*, *Plasmodium berghei*, and *Cryptosporidium parvum* exhibit transcriptional enrichment in cluster 7. **E)** Median z-scores of the top 100 female cluster genes, showing increased transcript abundance in late EES samples. **F)** Distribution of male-specific genes across clusters. Feature plot of 46 male-specific genes identified by Ramakrishnan *et al.*, enriched primarily in clusters 8 and 9. **G–H)** Median z-scores of the top 100 genes defining Male-1 and Male-2 clusters, respectively. Both clusters show increased expression in later EES samples. **I–K)** Gene Ontology (GO) enrichment analysis of female (cluster 7), Male-1 (cluster 8), and Male-2 (cluster 9) clusters. Female cluster terms are associated with metabolic processes; male clusters are enriched for microtubule-based movement and organization. **L)** Median z-scores of the top 100 genes defining the transitioning merozoite cluster.

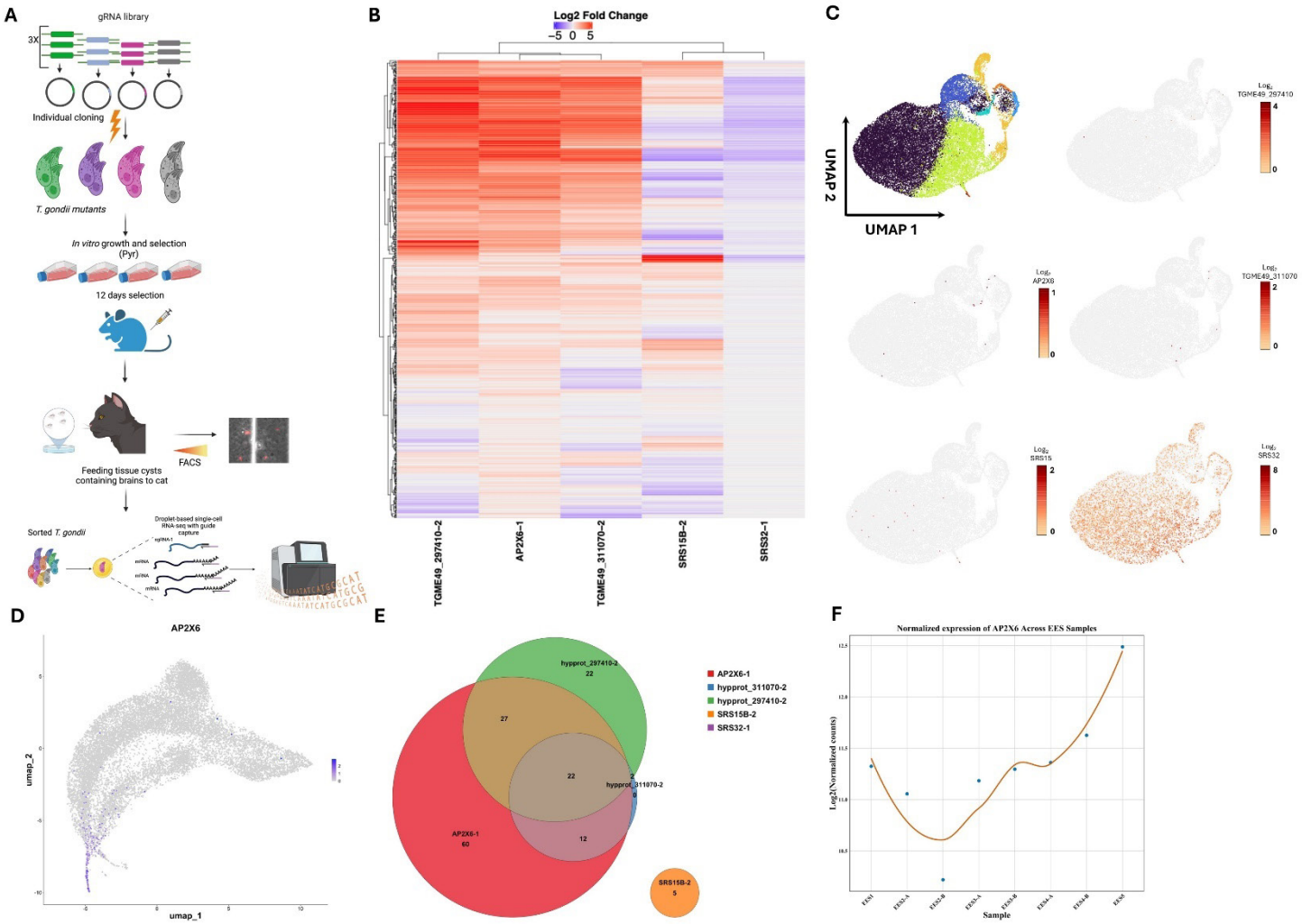

Fig. 5: Perturb-CRISPR screen and transcriptomic impact of candidate gene disruptions in *Toxoplasma gondii*.

A) Schematic of the perturb-CRISPR workflow. For each gene in the 18-candidate list, three distinct gRNAs targeting the same gene were transfected into TgVEG parasites to minimize bottleneck effects. At 24 h post-infection, cultures were switched to cDMEM containing 3 µM pyrimethamine for 12 days. Each knockout population was injected into 6–8 mice. After 25 days, brains were harvested for DNA extraction and for feeding to two cats. At 6–7 days post-infection, enteric stages were isolated from cat intestines, syringe-lysed, filtered (40 µm), and sorted by fluorescence before perturb-seq analysis. **B)** Heatmap of gRNAs associated with significant transcriptomic perturbations. **C)** UMAP visualization of single cells, highlighting those harboring gRNAs that caused significant transcriptional changes. **D)** Feature plot showing expression of AP2X6, a gene associated with the female cluster. **E)** Venn diagram illustrating shared and unique transcripts among cells harboring gRNAs that caused significant transcriptional changes. Only transcripts with consistent directional changes (all upregulated or all downregulated across compared groups) were included. Transcripts with opposing regulation in any group were excluded (e.g., *SRS32-1*). **F)** Expression dynamics of AP2X6 across EES samples. Blue points represent $\log_2$-normalized read counts; the red line indicates the smoothed trend.