# Single-cell RNA profiling of *Plasmodium vivax* liver stages reveals parasite- and host-specific transcriptomic signatures and drug targets

Anthony A Ruberto [a], Steven P Maher [a], Amélie Vantaux [b], Chester J Joyner [a,c], Caitlin Bourke [d], Balu Balan [d,e] , Aaron Jex [d,e], Ivo Mueller [d] , Benoit Witkowski [b], Dennis E Kyle [a]

[a] Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA, United States

[b] Malaria Molecular Epidemiology Unit, Institut Pasteur du Cambodge, Phnom Penh, Cambodia

[c] Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, GA, USA

[d] Population Health & Immunity Division, Walter and Eliza Hall Health Institute, Parkville, Victoria, Australia

[e] Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, Victoria, Australia

Running Head: Single-cell RNA profiling of *P. vivax* liver stages

Address correspondence to dennis.kyle@uga.edu

## ABSTRACT

The resilience of *Plasmodium vivax*, the most widely distributed malaria-causing parasite in humans worldwide, is attributed to its ability to produce dormant liver forms known as hypnozoites, which can activate weeks, months, or even years after an initial mosquito bite. The factors underlying hypnozoite formation and activation are poorly understood, as is the parasite's influence on the host hepatocyte. Here, we shed light on transcriptome-wide signatures of both the parasite and the infected host cell by sequencing over 1,000 *P. vivax* liver forms at single-cell resolution. We distinguish between replicating schizonts and hypnozoites at the transcriptional level, identifying key differences in transcripts encoding for RNA-binding proteins associated with cell fate. In infected hepatocytes, we show that genes associated with energy metabolism and antioxidant stress response were upregulated, and those involved in the host immune response downregulated, suggesting both schizonts and hypnozoites alter the host intracellular environment. The transcriptional markers in schizonts, hypnozoites, and infected hepatocytes revealed here pinpoint potential factors underlying dormancy and can inform drug targets against *P. vivax* liver-stage infection.

## IMPORTANCE

In humans, malaria-causing *Plasmodium* parasites invade and replicate in the liver before mounting an infection in the blood. A few *Plasmodium* species, including *P. vivax*, can cause relapsing infections due to forms that persist in the liver known as hypnozoites. While interventions targeting the hypnozoite can prevent clinical malaria, little is known about their biology at the molecular level. Transcriptome-wide assessment of individual hypnozoites has not been accessible due to the technical challenges associated with liver stage infections. To overcome technical limitations, we performed single-cell transcriptional profiling of liver stages generated using an *in vitro* infection platform. In elucidating the transcriptional signatures of *P. vivax* liver forms, as well as hepatocytes in response to infection, our data provide new insights into these host-parasite interactions. An improved understanding of these interactions will allow for novel drug or vaccine interventions targeting either parasite- or hepatocyte-specific metabolic pathways essential for successful liver stage development and disease progression.

## INTRODUCTION

Malaria is a significant global health burden, with an estimated 421 million cases and 627,000 deaths per year (World Health Organization, 2021). At least five *Plasmodium* species are known to infect humans, with *P. vivax* being responsible for the majority of cases outside Africa (World Health Organization, 2021). All *Plasmodium* parasites that infect humans have an obligatory developmental stage in the liver wherein the parasite undergoes asexual reproduction known as schizogony within a

49 hepatocyte before releasing thousands of merozoites into the blood. In *P. vivax*, a proportion of parasites
50 known as hypnozoites forgo immediate division in hepatocytes, and persist in the liver for weeks,
51 months, or years before activating to initiate schizogony resulting in a relapse blood infection (Adams
52 and Mueller, 2017; Krotoski et al., 1982). Hypnozoites are thought to cause up to 90% of cases in certain
53 geographical regions (Adekunle et al., 2015; Commons et al., 2020; Robinson et al., 2015), and
54 mathematical models predict that eliminating *P. vivax* malaria will be difficult without interventions that
55 target hypnozoite reservoirs (White et al., 2018).

56 The factors influencing the biogenesis, persistence, and activation of hypnozoites are poorly
57 understood. Studies have found that hypnozoites in temperate regions have long latencies and that
58 tropical ones have short latencies, suggesting a correlation between geographical region and relapse
59 patterns ((Battle et al., 2014; Huldén et al., 2008; Krotoski et al., 1986; White, 2011). These findings
60 suggest that the developmental trajectory is genetically encoded, and support the hypothesis that
61 developmental fate is somehow pre-programmed in the sporozoite (Lysenko et al., 1977; Ungureanu et
62 al., 1976); but other contributing factors from the host could play a role in determining whether the
63 sporozoite becomes a schizont or hypnozoite (Schäfer et al., 2021).

64 Our understanding of *P. vivax* liver stage biology at the cellular and molecular level is in its
65 nascency for two overarching reasons. For a long time, protocols to cryopreserve and continuously
66 maintain erythrocytic stage parasites *in vitro* were lacking, which necessitated the collection of
67 gametocyte-infected blood from human volunteers in *P. vivax* malaria-endemic regions to generate the
68 sporozoites required to initiate an infection. New laboratory models have recently been developed which
69 can partially alleviate the need to obtain gametocytes from human infected blood (Mikolajczak et al.,
70 2015; Schäfer et al., 2020). In addition to the sporozoite limitation, upon infecting the liver the high
71 ratio of noninfected versus infected hepatocytes has historically made it difficult to isolate and study
72 these forms in sufficient numbers. Robust *in vitro* pre-erythrocytic platforms capable of long-term
73 culture of sufficient numbers of *P. vivax* liver stages have only recently been developed (Maher et al.,
74 2021; Roth et al., 2018). Much of our knowledge of hypnozoite biology at the molecular level remains
75 limited to studies using *P. cynomolgi*—a parasite capable of causing relapsing infections non-human
76 primates, and for which molecular tools are available (Chua et al., 2019; Cubi et al., 2017; Voorberg-
77 van der Wel et al., 2017, 2020). However, it is unknown whether mechanisms regulating hypnozoite
78 formation and dormancy are consistent across different relapsing malaria parasites.

79 Transcriptome-wide profiling of *P. vivax* liver stages has previously been performed using a
80 bulk RNA sequencing strategy (Gural et al., 2018). While informative, the data only provide insights
81 into average transcript levels derived from either mixed (schizont and hypnozoite) or hypnozoite-only
82 populations, masking any transcriptional variation that might exist between individual parasites. Second,

83 the parasite's potential impact on the hepatocyte is not considered, barring analysis of host-pathogen
84 interactions. Nevertheless, these shortcomings highlight a need to better understand differences between
85 individual parasites, as well as how the parasite alters the hepatocyte. Advancements in these areas could
86 help elucidate parasite- and host- factors important for the parasite's liver stage development.

87 Single-cell RNA-sequencing (scRNA-seq) methods offer new ways of studying cell-to-cell
88 variation (Aldridge and Teichmann, 2020), and have already allowed for novel insights into host-
89 pathogen interactions in various disease contexts (Penaranda and Hung, 2019). In *Plasmodium* spp.,
90 scRNA-seq analyses have so far revealed that individual parasite forms exist on a spectrum
91 of transcriptomic states (Bogale et al., 2021; Howick et al., 2019; Poran et al., 2017; Real et al., 2021;
92 Reid et al., 2018; Ruberto et al., 2021a), and applying them to study the biology of *P. vivax* is a
93 burgeoning area of research (Mancio-Silva et al., 2021; Ruberto et al., 2021b; Sà et al., 2020). Droplet-
94 based techniques, such as 10x Genomics (Zheng et al., 2017), which use microfluidics, allow for the
95 sequencing of tens to hundreds of thousands of cells, increasing the chances of observing rare cell
96 populations (e.g., hepatocytes infected with *P. vivax*). We therefore presumed that applying single-cell
97 sequencing to *P. vivax* liver forms derived using an *in vitro* platform would prove a fruitful avenue of
98 inquiry into the parasite's biology cellular and molecular levels.

99 Using scRNA-seq, we characterize transcriptomic signatures of parasite and host in an *in vitro*
100 *P. vivax* liver stage model. We find differences in gene expression between replicating schizonts and
101 hypnozoites, and reveal variation, previously unobservable by bulk sequencing approaches, between
102 individual hypnozoites. In infected hepatocytes, we find that genes associated with energy metabolism
103 and antioxidant stress response are upregulated, and those involved in the host immune response
104 downregulated. Various host genes associated with these processes are found to be exclusively
105 upregulated during hypnozoite infection. Our study elucidates the transcriptional signatures among
106 infected hepatocytes and presents potential mechanisms *P. vivax* can use to either exploit its host for
107 rapid proliferation or to quietly persist, remaining undetected by the host immune system.

## 108 RESULTS

### 109 Strategy used to study *P. vivax* liver stage biology

110 Primary human hepatocytes were cultured and infected with *P. vivax* sporozoites on a 384-well plate
111 platform (Maher et al., 2021; Roth et al., 2018). **Figure 1A** outlines the workflow for the experiment.
112 Following infection with sporozoites, cultures were assessed at two endpoints: one containing both
113 schizonts and hypnozoites at day 5 post-infection, and the other containing solely hypnozoites at day 9
114 post-infection. Schizonts were eliminated from day 9 post-infection cultures by adding the

115 phosphatidylinositol 4-kinase (PI4K) inhibitor MMV390048 (1 μM) to the culture media on days 5, 6
116 and 7 post-infection (Gural et al., 2018; Paquet et al., 2017). For each collection time point, we
117 simultaneously initiated two sets of hepatocyte cultures: one infected with sporozoites for quantification
118 of immunofluorescent-labeled parasites by high-content imaging (HCI), and one infected with
119 sporozoites for transcriptional profiling using a high-throughput, droplet-based scRNA-seq workflow
120 (Zheng et al., 2017). As controls, a third set of naïve cultures were left uninfected and processed for
121 scRNA-seq alongside cultures at days 5 and day 9 post-infection. At each collection point, a target of
122 approximately 10,000 individual hepatocytes were partitioned into droplets prior to the generation of
123 RNA libraries. Each droplet contained primers with a unique cellular barcode that was ligated to
124 transcripts during the reverse transcription step, allowing for the identification of each transcript's
125 cellular origin post-sequencing. After sequencing the single-cell libraries, we mapped the resulting reads
126 onto the human (Ensembl, v103) and *P. vivax* PVP01 (PlasmoDB, v51) transcriptomes. We combined
127 the transcript information with the cellular barcodes to assess the number of transcripts in each cell, and
128 to determine whether a hepatocyte contained parasite material.

129 As expected, we detected hepatocytes infected with both forms of the parasite at 5 days post-
130 infection, and only hypnozoites at 9 days post-infection following exposure to MMV390048 (**Fig. 1B**
131 **and Fig. S1A**). **Figure 1C** displays the distribution of individual parasite growth areas, number of
132 parasites detected, and parasitemia levels at each endpoint. Across replicates, the parasitemia ranged
133 from 5.50-6.93% at 5 days post-infection and 1.76-2.77% at 9 days post-infection. Together, the high
134 infection rates coupled with the high-throughput capacity of 10x Genomics' single-cell isolation and
135 sequencing workflow allowed for sufficient material to be generated for successful profiling of both
136 parasite and host transcriptomes.

137 ***P. vivax* schizonts and hypnozoites are distinguishable at single-cell resolution**

138 We first analyzed the reads mapping to the *P. vivax* transcriptome to identify schizonts and hypnozoites,
139 as well as to assess the extent to which these forms differed at single-cell resolution. In total, we
140 characterized 1,438 parasite transcriptomes across two independent experimental replicates (**Fig. S1B**).
141 The number of parasite transcriptomes profiled corroborated the parasitemia levels quantified using
142 high-content imaging (**Table S1**). Low dimensional representation of the data revealed two populations
143 of parasites in samples profiled at 5 days post-infection (**Fig. 2A, left panel**). In contrast, and as
144 expected, samples profiled at 9 days post-infection contained one population (**Fig. 2A, right panel**).
145 Given that the population of cells profiled at day 9 post-infection, from which schizonts were chemically
146 removed, overlapped with one of the two cell populations profiled at 5 days post-infection, we deduced
147 that the non-overlapping population of cells encoded for schizonts. Consistent with our reasoning, an
148 unsupervised clustering algorithm (Traag et al., 2019) unbiasedly assigned the cells to one of the two

149 clusters (**Fig. 2B and Fig. S1C**). As expected, number of transcripts detected, unique molecular
150 identifier (UMI) counts (a readout for absolute transcript abundances (Kivioja et al., 2011), and the
151 transcript levels of the early liver stage development marker liver stage-specific protein, LISP2
152 (PVP01_0304700) (Gupta et al., 2019), all displayed greater quantities in the schizont cluster (C2)
153 relative to the hypnozoite cluster (C1) (**Fig. 1C**). Together, these findings show that the transcriptomes
154 of *P. vivax* schizonts and hypnozoites can be distinguished using our experimental workflow.

155 **Hypnozoites exhibit distinct transcriptional signatures compared to schizonts**

156 We next performed differential gene expression analysis to identify transcriptome-wide differences
157 distinguishing schizonts and hypnozoites. Using stringent inclusion cutoffs (adjusted p value < 0.01 &
158 absolute (average $\log_2$ fold change [FC]) > 0.5), we found significant differences in ~18% (836/4722)
159 of all transcripts detected (**Fig. S2A; Data Set S1A**). Relative to hypnozoites, we found an increase in
160 the expression levels of 368 genes in schizonts. As expected, we observed a significant increase of
161 LISP2 in schizonts (average $\log_2$ FC: 2.32) compared to hypnozoites. Shortly after invading the
162 hepatocyte, non-relapsing *Plasmodium* spp. liver forms undergo immediate replication. Studies in
163 rodent malaria models have shown that these forms are metabolically active and that various genes
164 involved in energy metabolism are upregulated (Caldelari et al., 2019; Tarun et al., 2008; Toro-Moreno
165 et al., 2020). Consistent with previous data, we found an increased transcription associated with
166 glycolysis (PVP01_0816000, PVP01_1229700, PVP01_1244000) and the TCA cycle
167 (PVP01_0106100, PVP01_1332400) in schizonts (**Fig. S2B; Data Set S1A**). More broadly, we found
168 increased transcription of various metabolic processes, as well as motility, DNA-, RNA-, and protein-
169 processing (**Fig. S2C; Data Set S1A**). Transcripts encoding for various proteasome subunits displayed
170 greater transcription in schizonts (**Fig. S3A; Data Set S1A**). To further show the importance of this
171 protein complex, we tested a known proteosome inhibitor, Carfilzomib (Kortuem and Stewart, 2013),
172 in our standard 12-day *P. vivax* small molecule screening assay (Maher et al., 2021), and found it
173 effectively killed schizonts ($IC_{50}$ 288nM) and hypnozoites ($IC_{50}$ 511nM) with some selectivity over host
174 hepatocytes ($IC_{50}$ 3.32 μM) (**Figs. S3B and S3D**).

175 The detection of transcripts associated with energy metabolism in hypnozoites, albeit at a lower
176 level relative to schizonts, suggested that these forms are also metabolically active (**Fig. S2B**). Relative
177 to replicating schizonts, we found an increase in gene expression of 468 transcripts in hypnozoites (**Data**
178 **Set S1A**). As represented in **Figure 3A**, genes with greater expression in hypnozoites encoded for a
179 diverse set of ontologies—including proteases, membrane proteins, transcription factors, and DNA-
180 /RNA-regulating proteins. Genes encoding for proteases vivapain-1, vivapain-2 (PVP01_1248900,
181 PVP01_0916200), and plasmepsin IV (PVP01_1340900) displayed the greatest differential expression
182 in these forms compared to schizonts (**Figs. 3A and 3B; Data Set S1A**). While limited research has

183 reported the functionality of these gene products in *P. vivax* (Byoung-Kuk et al., 2004), work in other
184 *Plasmodium* spp. has revealed that they localize to the food vacuole and play an important role in
185 hemoglobin catabolism during the blood stages of the parasite's life cycle (Dame et al., 2003; Drew et
186 al., 2008).

187 We also identified genes with greater expression associated with RNA regulatory mechanisms.
188 Notably, we found greater expression of CCR4 (PVP01_1429000), a transcript encoding for a protein
189 shown to have mRNA deadenylase activity and play a role in mRNA turnover (Tucker et al., 2001)
190 (**Data Set S1A**). In *Plasmodium* spp., it has been described as important in regulating the expression of
191 genes associated with invasion and egress in blood stages (Balu et al., 2011). RNA-binding proteins
192 (RBPs) also govern gene expression at the post-transcriptional level by transiently storing transcripts
193 until later processing (Hentze et al., 2018). Pumilio domain (PUF) proteins are a major family of RBPs
194 that regulate cellular fating through translational repression in eukaryotes. We found that the transcript
195 encoding for PUF1 (PVP01_1015200) was detected almost exclusively in hypnozoites (**Figs. 3A and**
196 **3B**; **Data Set S1A**). Furthermore, we found greater expression of 7-Helix-1 protein (PVP01_1008400)
197 (**Figs. 3A and 3B**; **Data Set S1A**)—a stress granule component that interacts with the RNA-binding
198 protein PUF2 and serves to be crucial for protein synthesis in *P. falciparum* (Miao et al., 2010).
199 Numerous studies have shown that *Plasmodium* spp. use various post-transcriptional regulatory
200 mechanisms to coordinate stage transitions and development (Baumgarten et al., 2019; Bunnik et al.,
201 2016; Foth et al., 2011; le Roch et al., 2004; Tarun et al., 2008; Vembar et al., 2016). Our data
202 corroborate these findings and suggest that similar strategies may be used in *P. vivax* to regulate liver
203 stage development. Moreover, these data suggest that RBP-mediated cellular fating and translational
204 repression may be involved in regulating hypnozoite dormancy.

205 Given the potential role in RBP-mediated cellular fating and translational repression governing
206 latency in hypnozoites, we performed further analyses focused on RBP-mediated regulation. To this
207 end, we scanned the *P. vivax* genome searching for genes encoding for cellular fating RBPs based on
208 ortholog groups (OGs) of known human cellular fating RBPs (348 RBPs; 227 OGs), identifying 89 OGs
209 (99 RBPs) in *P. vivax.* We detected 48 genes encoding for cellular fating RBPs with greater expression
210 in schizonts (**Fig. 3C, left panel**; **Data Set S1B**). Notably, we identified eukaryotic translational
211 initiation factors (PVP01_1238100, PVP01_0531200, PVP01_1431600, PVP01_0807600,
212 PVP01_1005600, PVP01_0519700, PVP01_1329200, PVP01_0825000, PVP01_1303500), and
213 splicing factors (PVP01_0716000, PVP01_0521000, PVP01_0804100, PVP01_1111800,
214 PVP01_0607600, PVP01_1012200) suggesting active translational events. Gene Ontology (GO) term
215 analysis revealed that these transcripts were associated with translation initiation (GO:0005852;
216 GO:0033290; GO:0016282) (adjusted p value $< 0.05$, **Fig. 3C**; **Data Set S1C**).

217 In hypnozoites, we detected 51 genes encoding for cellular fating RBPs with increased
218 expression relative to schizonts (**Fig. 3C, left panel**; **Data Set S1B**). These transcripts encode for
219 prominent P body markers, namely DOZI/DDX6 (PVP01_0819400), CCR4-NOT complex proteins
220 (PVP01_1429000, PVP01_1014400, PVP01_1331700, PVP01_1453400, PVP01_0929400), mRNA-
221 decapping enzyme subunit 1 (PVP01_0617200), mRNA-decapping enzyme subunit 2
222 (PVP01_1409900), UPF1 (PVP01_0805200), UPF2 (PVP01_0724300), UPF3 (PVP01_1226700),
223 PUF2 (PVP01_0526500), PABP (PVP01_1442500), Fibrillarin (PVP01_1341600) and trailer hitch
224 homolog (PVP01_1269800) (**Data Set S1B**). Transcripts encoding for cellular fating RBPs in
225 hypnozoite also represented GO terms corresponding to active biological condensate formations, such
226 as the CCR4-NOT complex (GO:0030014; GO:0030015) and the P body (GO:0000932) (adjusted p
227 value < 0.05, **Fig. 3C, right panel; Data Set S1C**). Together, these findings suggest that P body-
228 mediated cell fating could be important in regulating hypnozoite persistence.

229 **Transcriptomic heterogeneity exists amongst individual *P. vivax* hypnozoites**

230 Having described the transcriptomic signatures distinguishing schizonts and hypnozoites, we next
231 sought to assess the extent to which gene expression varied amongst hypnozoites. We thus parsed the
232 data into only hypnozoite transcriptomes. In total, 1,147 transcriptomes were assessed (**Fig. 4A, left**
233 **panel**). Low dimensional representation of the data revealed slight variation between samples collected
234 at days 5 and 9 post-infection (**Fig. 4A, right panel**). After regrouping the data using the same
235 unsupervised clustering algorithm described previously, we identified three clusters (**Fig. 4B, right**
236 **panel**). The proportion of hypnozoites in each cluster varied across the day of collection: specifically,
237 the proportion of hypnozoites in cluster 3 (C3) was greater in hypnozoites collected at day 9 compared
238 to day 5 post-infection (**Fig. 4B, left panel; Fig. S4A**). Notably, we also found greater gene expression
239 of LISP2 in C3 relative to the other two clusters (**Fig. 4C**).

240 We next identified markers that define the hypnozoites in each cluster using the FindAllMarkers
241 function in Seurat (Stuart et al., 2019). We defined a marker as a transcript displaying greater than 0.5
242 FC (average $\log_2$) compared to the other clusters. In total, we detected 421 markers (adjusted p value <
243 0.05) (**Fig. S4B**; **Data Set S2A**). Notable markers in C1 included transcripts associated with drug
244 resistance (PVP01_1447300, PVP01_1259100, and PVP01_1010900), energy metabolism
245 (PVP01_1435000 and PVP01_1122800), protein modification (PVP01_1248900 and
246 PVP01_0916200), and RNA-binding proteins (PVP01_0939900 and PVP01_0819400); in C2 those
247 associated with translation (PVP01_1255500 and PVP01_0822500); and in C3 those associated with
248 translation (PVP01_0213700 and PVP01_0203400), fatty acid metabolism (PVP01_1143000 and
249 PVP01_1022800), protein processing (PVP01_1308000 and PVP01_0927700) and ion transport
250 (PVP01_1407500) (**Fig. 4D**; **Data Set S2A**).

251 We next explored the data in the context of cellular fating to elucidate potential patterns in RBP
252 signatures. Among the hypnozoite markers found within our sub-clustering analysis, we identified 68
253 cellular fating RBPs, of which 54 RBPs were differentially expressed in at least in one of the three
254 clusters (adjusted p value < 0.05) (**Fig. 4E, top; Data Set S2B**). Approximately 4% (2/54) of the
255 differentially expressed genes in C3 encoded for cellular fating. In this cluster, the increased gene
256 expression of eukaryotic translation initiation factor 4E (PVP01_0203400) further suggests active
257 translation events may be occurring. C2 contained ~16% (9/54) of the cellular fating RBPs differentially
258 expressed. In this cluster, we found RBPs which are important determinants of cell proliferation from a
259 state of senescence such as eukaryotic translation initiation factor 5A (PVP01_1303500),
260 polypyrimidine tract binding protein (PVP01_1142400), GTP-binding nuclear protein RAN/TC4
261 (PVP01_0918300), transformer-2 protein homolog beta (PVP01_0802200). C1 contained the majority
262 of the differentially expressed RBPs (~80%, 43/54). Within this cluster, we identified prominent P body
263 markers with cellular fating ability, such as DOZI/DDX6 (PVP01_0819400), CCR4-NOT complex
264 proteins (PVP01_1429000, PVP01_1014400, PVP01_1331700, PVP01_1453400), UPF1
265 (PVP01_0805200), UPF2 (PVP01_0724300), PABP (PVP01_1442500), Fibrillarin (PVP01_1341600)
266 and trailer hitch homolog (PVP01_1269800). GO analysis revealed enrichment of cellular
267 compartments associated to biological condensate formations, namely CCR4-NOT complex
268 (GO:0030014), P granule (GO:0043186), and germplasm (GO:0060293) (**Fig. 4E; Data Set S2C**).

269 These findings, combined with the shift of hypnozoite proportions and increased transcript
270 levels of LISP2 from C1 to C3, allowed us infer the phenotype of hypnozoites in these clusters. As these
271 observations likely represent populations of hypnozoites in persisting and activating states, we defined
272 clusters C1, C2, and C3 to represent persisting, transitioning, and activating hypnozoites, respectively.
273 Together, these findings further our understanding of *P. vivax* hypnozoites through characterization of
274 gene usage and putative regulatory mechanisms in activating and persister forms. Furthermore, they
275 highlight a potential role of P body-mediated cellular fating governing persister hypnozoites.

276 **Strategy used to assess hepatocytes infected with *P. vivax* parasites**

277 After deciphering the identities of *P. vivax* transcriptomes representing schizonts and hypnozoites, we
278 next sought to characterize the transcriptional signatures of hepatocytes infected with either of these
279 forms. To this end, we integrated the data across all the conditions (**Figs. S5A and 5B)** and examined
280 whether distinct sub-populations of infected hepatocytes existed. As highlighted by the low dimensional
281 visualization of the data, we did not find a prominent pattern of infection (**Fig. 5**). Moreover, we detected
282 parasites in hepatocytes with different metabolic states (MacParland et al., 2018), as represented by their
283 transcriptional signatures associated with different hepatocyte zonation programs (**Figs. S5C and S5D**).

284 **Identification of transcriptional signatures in *P. vivax*-infected hepatocytes**

285 To identify host transcriptional signatures specific to infection status, we performed differential gene
286 expression analysis, comparing infected to non-infected hepatocytes. Among the hepatocytes infected
287 with schizonts, 106 genes displayed greater expression and 16 displayed decreased expression **(Data**
288 **Set S3A)**; while in hepatocytes infected with hypnozoites, 365 genes displayed greater expression and
289 41 genes displayed decreased expression (**Data Set S3B**) (adjusted p value < 0.05). As depicted in the
290 volcano plots in **Fig. 6A**, the magnitude of change in expression in the infected versus non-infected cells
291 (average $\log_2$ fold change) was greater in hepatocytes containing schizonts compared to those containing
292 hypnozoites.

293 To gain greater insights into the biological processes and pathways that characterize a schizont-
294 or hypnozoite-infected hepatocyte, we performed enrichment analysis (Raudvere et al., 2019)—scoring
295 18,594 gene sets against the genes differentially expressed in each of our transcriptional comparisons.
296 When considering both the schizont- and hypnozoite-infected hepatocyte populations, we detected a
297 total of 254 enriched gene sets (adjusted p value < 0.05) (**Fig. 6B**; **Data Set S3C**). Approximately 18%
298 (45/254) of these gene sets overlapped in hepatocytes infected with either schizonts or hypnozoites (**Fig.**
299 **S5E; Data Set S3D**). These data indicate that while some overlap exists in the gene networks the parasite
300 alters during infection, the majority of processes and pathways altered are specific to schizont versus
301 hypnozoite infected hepatocytes.

302 Using these gene sets, we then constructed network enrichment maps (Merico et al., 2010) to
303 provide a global overview of the prominent biological themes associated with each infection condition
304 (**Figs. 6C and 6D**). We detected upregulation of genes associated with pathways related to energy
305 metabolism (oxidative phosphorylation, mitochondrial function, and lipid catabolism) in schizont- and
306 hypnozoite-infected hepatocytes. Analysis of the most differentially expressed genes associated with
307 these pathways included *MT-ND2*, *CHCHD10*, and *AKR1B10* (**Data Sets S3A and S3B**). Corroborating
308 our transcriptomic data, we found an increase in protein levels of AKR1B10 in infected versus non-
309 infected hepatocytes (**Figs. S6A-D**). AKR1B10 has been implicated in cell survival through its
310 regulation of lipid metabolism and elimination of carbonyls (Martin and Maser, 2009; Qu et al., 2021;
311 Wang et al., 2009) and has primarily been described as a biomarker of cancer (Díez-Dacal et al., 2011;
312 Fukumoto et al., 2005; Jin et al., 2006). More recently, it has been shown that its expression is positively
313 correlated with non-alcoholic fatty liver disease (Govaere et al., 2020). Its role during *Plasmodium* spp.
314 liver stage infection warrants further research. In schizonts, the upregulation of genes linked to fatty
315 acid metabolism is consistent with observations in hepatocytes harboring developing forms from non-
316 relapsing *Plasmodium* spp. (Albuquerque et al., 2009; Itoe et al., 2014). We now reveal that infection

317 with hypnozoites also elicits enrichment of transcripts associated with fatty acid metabolism (**Fig. 6D**),
318 suggesting that these forms may be scavenging host resources to fuel their persistence.

319 Previous work assessing *Plasmodium* liver stage biology has revealed that the host is able to
320 detect the parasite (Liehl and Mota, 2012) and that type I interferon plays an integral role in this response
321 (Liehl et al., 2013; Miller et al., 2014). The decrease in transcripts associated with interferon signaling
322 (REAC:R-HSA-913531, GO:0034340, GO:0034341, GO:0071357, GO:0060337) in hepatocytes
323 infected with either schizonts or hypnozoites suggest that *P. vivax* manipulates the hepatocytes to evade
324 detection (**Figs. 6C and 6D; Data Sets S3C and S3D**). In hepatocytes infected with schizonts, we found
325 a decrease in expression of NFKB-regulated genes (*SOD2, TMEM45A, CCL20,* and *TPM1*) (**Data Set**
326 **S3A**), consistent with findings that *Plasmodium* spp. release signals to block proinflammatory response
327 (Singh et al., 2007). These findings, in addition to the decrease in the expression of *HLA-B*—a molecule
328 involved in antigen presentation to T cells that can trigger their activation when 'self' peptides are not
329 presented—and the neutrophil chemoattractant *CXCL8* suggest that replicating parasites alter
330 hepatocyte signaling pathways to minimize recruitment of innate immune cells and avoid detection by
331 T cells.

332 Similar to hepatocytes infected with schizonts, we found genes associated with interferon
333 signaling downregulated in hepatocytes infected with hypnozoites, namely: *UBB, UBC, HSP90AB1,*
334 *IFITM3, IFITM2,* and *HLA-A* (**Data Set S3B**). Furthermore, transcripts associated with the
335 downregulation of class major histocompatibility complex (*CD81* and *B2M*) were also decreased in
336 these hepatocytes (**Data Set S3B**). Together, these findings suggest that *P. vivax* infection of
337 hepatocytes may suppress the secretion of various chemokines and proinflammatory cytokines to thwart
338 the recruitment of inflammatory cells to the site of infection while simultaneously reducing the chances
339 they will be detected if these cells are recruited by downregulating MHC- I.

340 Enrichment analysis linking the changes in gene expression in infected hepatocytes with
341 transcriptional factors revealed distinct sets of putative transcriptional regulators. In total, we detected
342 35 transcription factors with potential roles in governing the gene expression changes in infected
343 hepatocytes (adjusted p value < 0. 05, **Fig. 7; Data Set S3E**). In infected hepatocytes, we found
344 enrichment of transcripts governed by NRF2 (**Fig. 7; Data Set S3E)**. The regulation of gene expression
345 through NRF2 in infected hepatocytes may serve to minimize reactive oxygen species in the cell (Hayes
346 and Dinkova-Kostova, 2014). Thus, in addition to the decreased expression of genes associated with
347 inflammatory response (**Data Set S3C**), NRF2-mediated regulation of antioxidant responsive genes may
348 protect hepatocytes from *Plasmodium*-induced stress and thus serve as a complementary means of
349 minimizing the host inflammatory response.

350 **DISCUSSION**

351 The infection of the liver by *P. vivax* constitutes an obligatory step in its life-cycle. *In vivo* studies have
352 shown that within hepatocytes, *P. vivax* can exist in one of two phenotypically distinct forms: schizonts
353 or hypnozoites (Cogswell, 1992; Mikolajczak et al., 2015). They have also been observed *ex vivo* by
354 infection of hepatocyte monolayers (Gural et al., 2018; Hollingdale et al., 1985; Roth et al., 2018;
355 Sattabongkot et al., 2006). A key problem hindering the study of *P. vivax* liver stage biology is the
356 difficulty involved in obtaining a sufficient number of infected hepatocytes to make meaningful
357 comparisons between them. We show that by coupling our *in vitro* liver stage model (Maher et al., 2021;
358 Roth et al., 2018) with a high throughput single-cell capture technology (Klein et al., 2015; Macosko et
359 al., 2015; Zheng et al., 2017) transcriptome-wide characterization of a large number of infected
360 hepatocytes is feasible. Our approach therefore provides much needed insight into the molecular
361 signatures of *P. vivax* hypnozoites and the host cells they infect.

362 First, we found clear transcriptional differences between schizonts and hypnozoites. Our
363 transcriptome-wide assessment supports the notion that hypnozoites are metabolically active (Adams
364 and Mueller, 2017), demonstrated by the transcription associated with glycolysis and fatty acid
365 metabolism. Second, we found that hypnozoites exist in various transcriptomic states. We speculate that
366 these differences might represent a spectrum of phenotypes, from persisting to activating. The
367 differential expression of genes encoding for cellular fating RBPs supports the notion that these forms
368 may use post-transcriptional mechanisms to control gene usage (Mueller et al., 2019). The biological
369 significance of the enrichment of transcripts associated with translational repression, found exclusively
370 in a subpopulation of hypnozoites, suggests that this mechanism inhibits activation, and warrants further
371 investigation.

372 Second, we found distinct transcriptomic signatures of infected versus non-infected
373 hepatocytes, and showed that signaling pathways associated with energy metabolism, antioxidant stress,
374 and immune response are altered in infected hepatocytes. We found no clear pattern of infection, as both
375 schizonts and hypnozoites are detected in hepatocytes with various metabolic signatures. The extent to
376 which individual hepatocytes respond to *P. vivax* schizonts and hypnozoites has previously been
377 unknown. A bulk RNA sequencing approach has been used to identify pathways altered in hepatocytes
378 in response to *Plasmodium* infection (LaMonte et al., 2019); however, it is limited insofar as it averages
379 gene expression from millions of cells, obscuring any parasite-specific effects on the infected cell.
380 Targeted approaches have revealed that *P. vivax* liver forms recruit host factors to facilitate growth
381 (LaMonte et al., 2019; Posfai et al., 2020), but how certain infected hepatocyte populations contribute
382 to disease transmission, and how hepatocyte subpopulations support development, are intriguing topics
383 for future studies (Vijayan et al., 2021). Our data identify potential host factors essential for *P. vivax*'s

384 successful liver stage development. Whether certain hepatocyte phenotypes influence the development,
385 especially activation of hypnozoites into schizonts, requires further investigation.

386 Several observations suggest protein degradation may be important for the parasite's liver
387 stages. Targeting this process may be an effective means of killing these forms. The ability of
388 hepatocytes to uptake and break down hemoglobin (Bissell et al., 1972) supports that this protein
389 substrate could be an essential amino acid source for the parasite. Our data suggest that several
390 mechanisms may allow *P. vivax* to utilize the amino acids derived from this protein. In hypnozoites, the
391 increased expression of genes encoding for proteases, such as vivapains, supports the idea that these
392 parasites digest host cytosolic proteins**.** The role of these proteins as hemoglobinases have been defined
393 extensively in *P. falciparum* blood stages (Drew et al., 2008; Rosenthal, 2011; Salas et al., 1995). A
394 similar role has been described for *P. vivax* (Byoung-Kuk et al., 2004), but there is limited information
395 about their significance in the liver stages. Vivapains are highly homologous to cathepsin protease L
396 (TgCpl) in the related apicomplexan *Toxoplasma gondii.* In these parasites, TgCpl is localized to its
397 plant-like vacuolar compartment—a lysosome-like organelle containing a variety of proteases (Parussini
398 et al., 2010)—and one of its functions is to aid in the digestion of host-derived proteins (Dou et al.,
399 2014). Genetic or chemical disruption of TgCpl negatively impacts the viability of bradyzoites, a
400 persisting form of this parasite (di Cristina et al., 2017). Given these orthologous proteins' roles in *P.*
401 *falciparum* and *T. gondii*, we speculate that the expression of vivapains may support the persistence of
402 *P. vivax* hypnozoites.

403 Persistence of hypnozoites is contingent on the longevity of the hepatocyte in which it resides.
404 The increased expression of genes associated with antioxidant stress (Scarpulla, 2002) is consistent with
405 the notion that hypnozoite infection induces changes in the hepatocyte to promote cell survival, in which
406 NRF-2 could be playing a regulatory role. We detected genes displaying greater expression in infected
407 hepatocytes related to NYFB-regulation, a transcription factor important for cell proliferation,
408 mitochondrial integrity, and cellular longevity (Tharyan et al., 2020). The increased expression of
409 mitochondria-related genes in hepatocytes infected with schizonts may be required to meet the parasite's
410 energy demands during replication, but the enrichment of mitochondria-related transcripts in cells
411 infected with hypnozoites could also aid in ensuring longevity of the hepatocyte. In addition to these
412 positive regulators of cell survival, we found negative regulation of transcripts through the decreased
413 expression of genes associated with cell death (*UBB, UBC,* and *HSP90AB1*). Their expression is thought
414 to be regulated by CEBPD, a transcription factor that has a role in governing cell death (Tsai et al., 2017;
415 Wang et al., 2019).

416 Persistence of hypnozoites also requires evading host detection, but how they are able to do so
417 has been elusive. Our analyses indicate that the parasite alters the host hepatocyte to evade detection by

418 immune cells. First, there is a decrease in expression of genes encoding for chemoattractants (*CCL20,*
419 *CXCL8, IL32*) in infected hepatocytes, and an increase in expression of PTGR1, a negative regulator of
420 the chemotactic factor leukotriene B4 (Yokomizo et al., 1993). These changes would result in decreased
421 secretion of chemokines that recruit innate immune cells to the liver, thereby reducing recruitment of
422 adaptive immune cells like CD8 T-cells that kill infected hepatocytes (Chakravarty et al., 2007).
423 Interestingly, there was also a decrease in expression of MHC-1 molecule (*HLA-A and -B*). The
424 presentation of parasite peptides to be detected by surveying CD8 T-cells would therefore be reduced
425 (Neefjes et al., 2011), creating an additional layer of protection should innate immune cells begin to be
426 recruited, for example upon lysis of the hepatocyte by schizonts. We propose that *P. vivax* infected
427 hepatocytes are re-programmed to prevent recruitment of innate immune cells on the one hand, and to
428 reduce the likelihood of their detection by the adaptive immune system on the other, all to enable
429 progress through schizogony and for hypnozoites to persist undetected. Understanding the mechanisms
430 leading up to this should be studied in the future, as their reversal may lead to improved clearance of
431 hypnozoites, thus reducing relapses. Indeed, a vaccine would need to decrease relapses to have any
432 impact on *P. vivax* transmission and disease (White et al., 2017). However, the decrease of MHC-I
433 expression in infected hepatocytes may mean that standard focuses to generate CD8 and CD4
434 polyfunctional T-cells in the liver may not be effective in eliminating *P. vivax* or other relapsing
435 parasites. Recent attempts to test anti-relapse vaccine candidates with *P. cynomolgi* did not prevent
436 relapses despite higher inducing IFNγ producing cells, a common metric used as a correlate of protection
437 (Kim et al., 2020), indicating that alternative strategies may be needed.

438 Our transcriptome-wide characterization of schizonts, hypnozoites, and the infected hepatocytes
439 highlights transcripts altered during infection, and provides a framework for examining the role of
440 infection dynamics of *P. vivax* liver stage infections. By examining two early liver stage time points at
441 single-cell resolution, our study lays out proof of principle and a technical advance demonstrating the
442 importance of *in vitro* models to shed light on the molecular basis of *P. vivax* liver stage development
443 and infection.

## 444 CONCLUSIONS

445 The overall aim of our research was to elucidate the molecular signatures of *P. vivax* liver infection, a
446 better understanding of which will aid in facilitating the development of new therapeutics to ultimately
447 reduce transmission. We anticipate that interventions could involve parasite- and/or host- directed small
448 molecule therapies, two approaches our group is actively pursuing. The comprehensive mapping of the
449 transcriptional landscape of both the host and *P. vivax* described here provides an important framework
450 for further investigation.

451 **MATERIALS AND METHODS**

452 **Ethics statement**

453 All research procedures were reviewed and approved by the Cambodian National Ethics Committee for
454 Health Research (approval numbers: #113NECHR, #104NECHR and #089NECHR).

455 **Blood samples, mosquitos, and infections**

456 The complete protocol used for mosquito rearing an infection has been published (Maher et al., 2021).
457 In brief, blood samples from symptomatic patients infected with *P. vivax* were collected at local health
458 facilities in Mondulkiri province in Cambodia from March 2020 to October 2021. Following a *P. vivax*
459 gametocyte-containing blood meal, *An. dirus* mosquitoes were maintained at ambient temperature on a
460 12:12 light: dark cycle and fed 10% sucrose + 0.5% PABA solution. *An. dirus* found positive for *P.*
461 *vivax* oocysts at six-days post-feeding were transported to the Institut Pasteur Cambodia Insectary
462 Facility in Phnom Penh, Cambodia where they were maintained under the same conditions described
463 above.

464 **Sporozoite isolation and primary human hepatocyte infections**

465 The complete protocol used for seeding primary human hepatocytes, harvesting sporozoites, and
466 infecting cultures has been published (Maher et al., 2021). In brief, *P. vivax* sporozoites were isolated
467 from the salivary glands of female *An. dirus* mosquitoes 18 days after an infectious blood-meal. Primary
468 human hepatocytes (Lot BGW, BioIVT) were seeded 2 days before infection with 30,000 sporozoites
469 (replicate 1), or 15,200 sporozoites (replicate 2) per well. Media was replaced with fresh CP media
470 (BioIVT) containing antibiotics (penicillin, streptomycin, and gentamycin) the day after infection and
471 every 2-3 days thereafter. Hepatocytes containing solely hypnozoites at 9 days post-infection were
472 treated with 1 µM MMV390048 on days 5, 6, and 7 post-infection to kill replicating schizonts. At days
473 5 and 9 post-infection, cultures were processed for single-cell sorting (described in the following
474 section) or immunofluorescence assay (IFA). For the latter, hepatocytes were fixed with 4%
475 paraformaldehyde in 1X PBS. Fixed cultures were then stained overnight at 4ºC with recombinant
476 mouse anti-*P. vivax* UIS4 antibody (Schäfer et al., 2018) diluted 1:25,000 in a stain buffer (0.03% Triton
477 X-100) and 1% (w/v) BSA in 1X PBS. The following day, cultures were washed three times with 1X
478 PBS and then stained overnight at 4ºC with 1:1,000 rabbit anti-mouse AlexaFluor ™488-conjugated
479 antibody diluted 1:1,000 in stain buffer. Cultures were washed three times with 1X PBS and
480 counterstained with 1 mg/mL Hoechst 33342 (Thermo Fisher) to detect parasite and host DNA. High
481 content imaging was performed on a Lionheart imaging system (Biotek). Quantification of the number

482 of nuclei and parasite liver forms was performed using Gen5 high content analysis software (Biotek).
483 Confocal images were obtained using an ImageXpress Micro (Molecular Devices).

484 **Single-cell partitioning, library preparation, and sequencing**

485 Hepatocyte cultures infected with *P. vivax* were processed prior to- and post- PI4K treatment (5- and 9-
486 days post-infection, respectively). Two replicate experiments, defined as an independent plating of
487 hepatocytes infected with sporozoites from a unique field isolate, were performed several months apart.
488 Additionally, and as a negative control, one replicate of hepatocyte cultures never exposed to *P. vivax*
489 (naïve) were processed similarly to those that were infected. At each endpoint for each replicate,
490 approximately 640,000 PHH hepatocytes (16 wells of a 384-well plate each having approximately
491 10,000 cells) were treated with trypsin (Corning 25-053-Cl) to facilitate their detachment from the plate.
492 Once detached, trypsin was inactivated with 1:1 complete media and the cell suspension was collected
493 into a 5mL protein-low bind centrifuge tube (Eppendorf). To remove residual trypsin, hepatocytes were
494 washed via three sequential rounds of pelleting (50 x g, 3 min, 4ºC) and resuspending in 1X PBS
495 containing 0.1% BSA. Cells were then passed through a 35µm strainer (Falcon 352235) to minimize or
496 remove hepatocyte clusters. Prior to partitioning cells on the Chromium controller (10x Genomics),
497 hepatocytes were counted and the cell concentration was assessed using a hemocytometer.
498 Approximately 10,000 hepatocytes were loaded on a Chromium Chip B (10x Genomics) as per
499 manufacturer's recommendations. Chips containing hepatocyte suspensions, Gel Beads in Emulsion
500 (GEMs), and reverse transcription reagents were loaded into the Chromium controller (10x Genomics)
501 for single-cell partitioning and cell barcoding. Barcoded cDNA libraries were generated according to
502 the Chromium Single Cell 3' gene expression protocol (version 3). In total, 6 cDNA libraries were
503 generated (two replicates at 5 days post-infection, two replicates at 9 days post-infection, one uninfected
504 control for day 5 cultures, and one uninfected control for day 9 cultures). Libraries were loaded on
505 individual flow cell lanes and sequenced at a depth of 400 million reads using a HiSeq X Ten platform
506 (Illumina) at Macrogen (Seoul, Korea).

507 **Computational analysis of single-cell RNA sequencing data**

508 Quality control checks on the raw sequencing data were performed using FASTQC (Andrews, 2010). A
509 multi-species transcriptome index—containing the transcriptomes for *P. vivax* P01 (PlasmoDB.org,
510 version 51) and *H. sapiens* GRCh8 (Ensembl version 103)—was created using the 'ref' function in kb-
511 python (0.26.3) (Bray et al., 2016; Melsted et al., 2021). Next, reads pseudoaligned and counted using
512 the 'count' function in kb-python with default parameters selected. The resulting unfiltered cell-gene
513 count matrices were read into R (v 4.1.1) for further processing.

514 **Filtering and normalization of scRNA-seq count matrices**

515 R scripts containing all computational analyses will be available on a dedicated GitHub repository upon
516 publication.

517 To distinguish between droplets containing cells and those only with ambient RNA, we used
518 emptyDrops function (Lun et al., 2019) with the lower total UMI count threshold (at or below which all
519 barcodes are assumed to correspond to empty droplets) set to 1,000. The droplets with cells (FDR <
520 0.001), more specifically, those containing barcoded transcripts derived from either human, parasite, or
521 a combination of the two—were kept for downstream processing.

522 **Assessment of *P. vivax* liver stage transcriptomes**

523 For the analyses of *P. vivax*, we filtered out the transcripts derived from human, leaving only transcripts
524 derived from the parasite quantifiable. For each sample, cells with less than 200 UMIs were removed.
525 Next, we filtered out low detected genes, retaining transcripts with 10 or greater UMIs. Post cell and
526 gene filtering, the data were normalized using Seurat (Stuart et al., 2019) 'LogNormalize' function with
527 the default parameters selected.

528 *Data integration and clustering*

529 The data were integrated using Seurat's SCTransform function (Hafemeister and Satija, 2019)
530 with the following parameters indicated: variable.features.n = NULL, variable.features.rv.th = 1.3.
531 Clustering was performed using Seurat's SNN-Leiden algorithm (Traag et al., 2019). The full
532 description of this analysis is found in the extended methods.

533 *Differential gene expression analysis*

534 To detect differentially expressed transcripts, the Seurat function FinalAllMarkers and
535 FindMarkers were used. For differential gene expression testing using the FindMarkers function, we
536 used the following parameters: logfc.threshold = 0.0, ident.1 = '1', ident.2 = '2', test.use = "wilcox",
537 min.pct = 0, only.pos = F, assay = "RNA". For differential gene expression testing using the
538 FindAllMarkers function, we used the following parameters: test.use = "wilcox", only.pos = T, assay =
539 "RNA", log2FC = 0.0.

540 *RBP analyses*

541 RBPs with cell fating potential reported in humans (Decker and Parker; Gerstberger et al., 2014;
542 Guallar and Wang, 2014; Voronina et al., 2011; Youn et al., 2019) were manually curated. The

543 orthologue's counterparts of these cell fating RBPs from the *P. vivax* PVP01 genome were mined using
544 OrthoMCL (Release-6.8)(Chen et al., 2006). Within the differentially expressed genes in *P. vivax*,
545 transcripts encoding for cell fating RBPs were filtered for further assessment.

546 *GO enrichment analyses*

547 GO enrichment analysis was performed on the PlasmoDB (v55) website with the following
548 input parameters selected: P value cutoff = 0.05; Limit to GO Slim terms = No; Evidence: Computed,
549 curated. Enrichment from GO cellular compartment (CC), biological processes (BP), and molecular
550 function (MF) were all assessed.

551 **Assessment of host hepatocyte transcriptomes**

552 For the analyses of the host, we removed all transcripts derived from *P. vivax*, leaving only the
553 transcripts derived from hepatocytes. We retained transcripts detected in at least 30 cells. Next, cells
554 with less than 1,500 UMIs were removed from further analyses**.** Post-cell and gene filtering, the data
555 from each replicate were normalized using Seurat's 'LogNormalize' function with the default
556 parameters.

557 *Data integration and clustering*

558 Filtered, normalized data matrices were integrated in a manner described in in the Seurat
559 (version 4) vignette, *Introduction to scRNA-seq integration*, available on the Sajita Lab's website
560 (https://satijalab.org/seurat/articles/integration_introduction.html). Briefly, integration features were
561 identified in each replicate using the 'SelectIntegrationFeatures' function. Next, integration anchors
562 were identified using the 'FindIntegrationAnchors'. Last, using these anchors, the four datasets were
563 integrated using the 'IntegrateData' function. Clustering was performed using Seurat's SNN-Leiden
564 algorithm (Traag et al., 2019).

565 *Gene Ontology and Reactome enrichment analyses*

566 Enriched biological processes and pathways were determined using gProfiler (Raudvere et al.,
567 2019) with the following options selected: ordered query, statistical domain scope—only annotated
568 genes, significance threshold—g:SCS threshold, user threshold—0.05, Gene Ontology—GO biological
569 processes, No electronic GO annotations, biological pathways—Reactome. Transcripts that displayed
570 significant changes in expression in infected hepatocytes versus non-infected hepatocytes were used as
571 input. EnrichmentMap (v 1.1.0) (Merico et al., 2010) for CytoScape (v 3.8.2) (Shannon et al., 2003) was
572 used to create the enrichment networks. Significantly enriched GO biological processes and Reactome

573 pathways (adjusted p value < 0.05) were used as input. Pathways were clustered and annotated using
574 the AutoAnnotate function (Kucera et al., 2016). Node size corresponds to the number of transcripts in
575 the corresponding set, and edge widths correspond to the number of transcripts shared between sets.

576 *Transcription factor enrichment analysis*

577 We downloaded curated sets of known transcription factors from the website
578 (http://amp.pharm.mssm.edu/Enrichr/). Transcription factor—target gene sets were obtained by ChIP-
579 X experiments from the ChEA (Lachmann et al., 2010) and ENCODE (Project Consortium, 2004)
580 projects. TF-target relationships using Chip-seq data from ENCODE and CHEA were used to identify
581 overrepresented pathways and TFs. Differentially expressed genes (adjusted p value < 0.05) in each
582 condition versus non-infected cells were used as inputs. Gene sets with an adjusted p value < 0.05 were
583 considered significantly enriched.

584 **Carfilzomib dose-response assessment**

585 A carfilzomib dose-response against *P. vivax* liver forms was obtained from three independent
586 experiments, in which a different *P. vivax* clinical isolate was allowed to infect BGW human
587 hepatocytes. The 12-day assay was used (Maher et al., 2021). All independent experiments resulted in
588 sufficient hypnozoite formation within each well for dose-response calculations. Liver stage parasite
589 growth metrics and compound dilutions were loaded into CDD Vault (Collaborative Drug Discovery)
590 and growth data normalized such that zero (negative) inhibition was set as the average of DMSO wells
591 and 100% (positive) inhibition was set to the effective doses of a nigericin control. Normalized values
592 were then combined in Graphpad Prism to fit a dose-response curve and calculated $IC_{50}$ values from all
593 replicates.

594 **Quantification of AKR1B10 protein expression**

595 To analyze AKR1B10 expression and localization, *P. vivax*-infected PHHs were fixed 12 days post-
596 infection. Infected wells were then permeabilized for 20 minutes with 0.2% TritonX-100, washed thrice
597 with 1X PBS, and blocked in 3% BSA for one hour. Primary antibodies rabbit anti-HsAKR1B10 (PA5-
598 30773, Thermo Fisher) and mouse anti-PvUIS4 were sequentially added, diluted 1:100 and 1:500,
599 respectively. Antibodies were incubated overnight at 4ºC and washed 3 times with 1X PBS before
600 addition of 1:400 Alexa- Fluor 568 donkey anti-rabbit for HsAKR1B10 or 1:400 AlexaFluor 488 goat
601 anti-mouse for PvUIS4. Secondary antibodies were incubated overnight at 4ºC. Cells were washed thrice
602 with 1X PBS before adding 0.5 mg/mL DAPI for 10 minutes. Cells were washed 3 final times with 1X

603 PBS before image collection with an ImageXpress inverted confocal microscope (Molecular Devices).
604 Image analysis was completed using MetaXpress analysis software (Molecular Devices).

605 **Contact for reagent and resource sharing**

606 Further information and requests for resources and reagents should be directed to and will be fulfilled
607 by the Lead Contact, Dennis Kyle (email: dennis.kyle@uga.edu).

608 **Data availability**

609 Raw sequencing files (FASTQ), processed data, scripts, and all host- and parasite- differentially
610 expressed genes will be made publicly available upon publication of the manuscript.

611 **Acknowledgements**

612 We thank the patients of Mondulkiri Province, Cambodia, for participating in this study. We thank the
613 Institut Pasteur du Cambodge's field site manager (Saorin Kim) for logistical assistance, insectary staff
614 (Makara Pring, Koeun Kaing, Nora Sambath) for the *An. dirus* mosquito colony maintenance, and
615 laboratory staff (Eakpor Piv, Chansophea Chhin, Sreyvouch Phen, Chansovandan Chhun, Sivcheng
616 Phal, Baura Tat) for assistance with the mosquito dissections and the *in vitro* assays.

617 High-content imaging data was produced in part by the Biomedical Microscopy Core at the
618 University of Georgia, supported by the Georgia Research Alliance. A.R.J., B.B. and I.M. acknowledge
619 support from the Victorian State Government Operational Infrastructure Support and Australian
620 Government National Health and Medical Research Council Independent Research Institute
621 Infrastructure Support Scheme. This work was also supported by the Agence Nationale de la Recherche
622 (ANR-17-CE13-0025 to A.A.R and I.M.); an Australian National Health and Medical Research Council
623 Leadership Fellowship (APP1194330 to A.R.J.); the Bill & Melinda Gates Foundation (OPP1023601
624 to D.E.K.); and Medicines for Malaria Venture (RD/2017/0042 to B.W. and A.V., RD/16/1082 and
625 RD/15/0022 to S.P.M. and D.E.K.).

626 A.A.R., S.P.M., A.V., I.M., B.W. and D.E.K conceived and designed the study; A.A.R, S.P.M.
627 and A.V. performed the experiments; A.A.R., C.B., B.B. and C.J. curated the data; A.A.R., S.P.M., B.B.,
628 A.R.J. and C.J. performed the formal analyses; S.P.M., A.V. and I.M., B.W. and D.E.K. coordinated the
629 research activities; S.P.M., A.V., I.M., B.W. and D.E.K. provided resources; S.P.M. and D.E.K.
630 acquired funding for the work; A.A.R wrote the initial draft of the article; all authors reviewed and made
631 contributions to the submitted article.

632 **REFERENCES**


633 Adams, J.H., and Mueller, I. (2017). The Biology of Plasmodium vivax. Cold Spring Harbor
634 Perspectives in Medicine *7*, a025585.

635 Adekunle, A.I., Pinkevych, M., McGready, R., Luxemburger, C., White, L.J., Nosten, F., Cromer, D.,
636 and Davenport, M.P. (2015). Modeling the Dynamics of Plasmodium vivax Infection and Hypnozoite
637 Reactivation In Vivo. PLOS Neglected Tropical Diseases *9*, e0003595.

638 Albuquerque, S.S., Carret, C., Grosso, A.R., Tarun, A.S., Peng, X., Kappe, S.H.I., Prudêncio, M., and
639 Mota, M.M. (2009). Host cell transcriptional profiling during malaria liver stage infection reveals a
640 coordinated and sequential set of biological events. BMC Genomics *10*, 1–13.

641 Aldridge, S., and Teichmann, S.A. (2020). Single cell transcriptomics comes of age. Nature
642 Communications 2020 11:1 *11*, 1–4.

643 Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available at:
644 Http://Www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc.

645 Balu, B., Maher, S.P., Pance, A., Chauhan, C., Naumov, A. v., Andrews, R.M., Ellis, P.D., Khan, S.M.,
646 Lin, J. wen, Janse, C.J., et al. (2011). CCR4-Associated Factor 1 Coordinates the Expression of
647 Plasmodium falciparum Egress and Invasion Proteins. Eukaryotic Cell *10*, 1257.

648 Battle, K.E., Karhunen, M.S., Bhatt, S., Gething, P.W., Howes, R.E., Golding, N., van Boeckel, T.P.,
649 Messina, J.P., Shanks, G.D., Smith, D.L., et al. (2014). Geographical variation in Plasmodium vivax
650 relapse. Malaria Journal *13*, 1–16.

651 Baumgarten, S., Bryant, J.M., Sinha, A., Reyser, T., Preiser, P.R., Dedon, P.C., and Scherf, A. (2019).
652 Transcriptome-wide dynamics of extensive m6A mRNA methylation during Plasmodium falciparum
653 blood-stage development. Nature Microbiology *4*, 2246.

654 Bissell, D.M., Hammaker, L., and Schmid, R. (1972). Hemoglobin and Erythrocyte Catabolism in Rat
655 Liver: The Separate Roles of Parenchymal and Sinusoidal Cells. Blood *40*, 812–822.

656 Bogale, H.N., Pascini, T. v., Kanatani, S., Sá, J.M., Wellems, T.E., Sinnis, P., Vega-Rodríguez, J., and
657 Serre, D. (2021). Transcriptional heterogeneity and tightly regulated changes in gene expression during

*Plasmodium berghei* sporozoite development. Proceedings of the National Academy of Sciences *118*, e2023438118.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology *34*, 525–527.

Bunnik, E.M., Batugedara, G., Saraf, A., Prudhomme, J., Florens, L., and le Roch, K.G. (2016). The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum. Genome Biology *17*.

Byoung-Kuk, N.A., Shenai, B.R., Sijwali, P.S., Choe, Y., Pandey, K.C., Singh, A., Craik, C.S., and Rosenthal, P.J. (2004). Identification and biochemical characterization of vivapains, cysteine proteases of the malaria parasite Plasmodium vivax. Biochemical Journal *378*, 529.

Caldelari, R., Dogga, S., Schmid, M.W., Franke-Fayard, B., Janse, C.J., Soldati-Favre, D., and Heussler, V. (2019). Transcriptome analysis of Plasmodium berghei during exo-erythrocytic development. Malaria Journal *18*, 330.

Chakravarty, S., Cockburn, I.A., Kuk, S., Overstreet, M.G., Sacci, J.B., and Zavala, F. (2007). CD8+ T lymphocytes protective against malaria liver stages are primed in skin-draining lymph nodes. Nature Medicine 2007 13:9 *13*, 1035–1041.

Chen, F., Mackey, A.J., Stoeckert, C.J., and Roos, D.S. (2006). OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Research *34*.

Chua, A.C.Y., Jie, J., Ong, Y., Malleret, B., Suwanarusk, R., Kosaisavee, V., Zeeman, A.-M., Cooper, C.A., Tan, K.S.W., Zhang, R., et al. (2019). Robust continuous in vitro culture of the Plasmodium cynomolgi erythrocytic stages. Nature Communications *10*, 1–13.

Cogswell, F.B. (1992). The hypnozoite and relapse in primate malaria. Clinical Microbiology Reviews *5*, 26–35.

Commons, R.J., Simpson, J.A., Watson, J., White, N.J., and Price, R.N. (2020). Estimating the Proportion of Plasmodium vivax Recurrences Caused by Relapse: A Systematic Review and Meta-Analysis. The American Journal of Tropical Medicine and Hygiene *103*, 1094–1099.

683 di Cristina, M., Dou, Z., Lunghi, M., Kannan, G., Huynh, M.-H., McGovern, O.L., Schultz, T.L.,
684 Schultz, A.J., Miller, A.J., Hayes, B.M., et al. (2017). Toxoplasma depends on lysosomal consumption
685 of autophagosomes for persistent infection.

686 Cubi, R., Vembar, S.S., Biton, A., Franetich, J.-F., Bordessoulles, M., Sossau, D., Zanghi, G., Bosson-
687 Vanga, H., Benard, M., Moreno, A., et al. (2017). Laser capture microdissection enables transcriptomic
688 analysis of dividing and quiescent liver stages of Plasmodium relapsing species. Cellular Microbiology
689 *19*, e12735.

690 Dame, J.B., Yowell, C.A., Omara-Opyene, L., Carlton, J.M., Cooper, R.A., and Li, T. (2003).
691 Plasmepsin 4, the food vacuole aspartic proteinase found in all Plasmodium spp. infecting man.
692 Molecular and Biochemical Parasitology *130*, 1–12.

693 Decker, C.J., and Parker, R. P-Bodies and Stress Granules: Possible Roles in the Control of Translation
694 and mRNA Degradation.

695 Díez-Dacal, B., Gayarre, J., Gharbi, S., Timms, J.F., Coderch, C., Gago, F., and Pérez-Sala, D. (2011).
696 Identification of aldo-keto reductase AKR1B10 as a selective target for modification and inhibition by
697 prostaglandin A(1): implications for antitumoral activity. Cancer Research *71*, 4161–4171.

698 Dou, C., Mcgovern, Z.L., Cristina, D. di, and Carruthers, M.B. (2014). Toxoplasma gondii ingests and
699 digests host cytosolic proteins. MBio *5*, 1188–1202.

700 Drew, M.E., Banerjee, R., Uffman, E.W., Gilbertson, S., Rosenthal, P.J., and Goldberg, D.E. (2008).
701 Plasmodium Food Vacuole Plasmepsins Are Activated by Falcipains. The Journal of Biological
702 Chemistry *283*, 12870.

703 Foth, B.J., Zhang, N., Chaal, B.K., Sze, S.K., Preiser, P.R., and Bozdech, Z. (2011). Quantitative time-
704 course profiling of parasite and host cell proteins in the human malaria parasite Plasmodium falciparum.
705 Molecular and Cellular Proteomics *10*.

706 Fukumoto, S.I., Yamauchi, N., Moriguchi, H., Hippo, Y., Watanabe, A., Shibahara, J., Taniguchi, H.,
707 Ishikawa, S., Ito, H., Yamamoto, S., et al. (2005). Overexpression of the aldo-keto reductase family
708 protein AKR1B10 is highly correlated with smokers' non-small cell lung carcinomas. Clinical Cancer
709 Research : An Official Journal of the American Association for Cancer Research *11*, 1776–1785.

710 Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nature
711 Reviews Genetics 2014 15:12 *15*, 829–845.

712 Govaere, O., Cockell, S., Tiniakos, D., Queen, R., Younes, R., Vacca, M., Alexander, L., Ravaioli, F.,
713 Palmer, J., Petta, S., et al. (2020). Transcriptomic profiling across the nonalcoholic fatty liver disease
714 spectrum reveals gene signatures for steatohepatitis and fibrosis. Science Translational Medicine *12*.

715 Guallar, D., and Wang, J. (2014). RNA-binding proteins in pluripotency, differentiation, and
716 reprogramming. Frontiers in Biology *9*, 389–409.

717 Gupta, D.K., Dembele, L., Voorberg-Van Der Wel, A., Roma, G., Yip, A., Chuenchob, V.,
718 Kangwanrangsan, N., Ishino, T., Vaughan, A.M., Kappe, S.H., et al. (2019). The plasmodium liver-
719 specific protein 2 (LISP2) is an early marker of liver stage development. ELife *8*.

720 Gural, N., Mancio-Silva, L., Miller, A.B., Galstian, A., Butty, V.L., Levine, S.S., Patrapuvich, R., Desai,
721 S.P., Mikolajczak, S.A., Kappe, S.H.I., et al. (2018). In Vitro Culture, Drug Sensitivity, and
722 Transcriptome of Plasmodium Vivax Hypnozoites. Cell Host & Microbe *23*, 395-406.e4.

723 Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq
724 data using regularized negative binomial regression. Genome Biology *20*, 1–15.

725 Hayes, J.D., and Dinkova-Kostova, A.T. (2014). The Nrf2 regulatory network provides an interface
726 between redox and intermediary metabolism. Trends in Biochemical Sciences *39*, 199–218.

727 Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding
728 proteins. Nature Reviews Molecular Cell Biology 2018 19:5 *19*, 327–341.

729 Hollingdale, M.R., Collins, W.E., Campbell, C., and Schwartz, A.L. (1985). In vitro culture of two
730 populations (dividing and nondividing) of exoerythrocytic parasites of Plasmodium vivax. The
731 American Journal of Tropical Medicine and Hygiene *34*, 216–222.

732 Howick, V.M., Russell, A.J.C., Andrews, T., Heaton, H., Reid, A.J., Natarajan, K., Butungi, H., Metcalf,
733 T., Verzier, L.H., Rayner, J.C., et al. (2019). The Malaria Cell Atlas: Single parasite transcriptomes
734 across the complete Plasmodium life cycle. Science (New York, N.Y.) *365*, eaaw2619.

735 Huldén, L., Huldén, L., and Heliövaara, K. (2008). Natural relapses in vivax malaria induced by
736 Anopheles mosquitoes. Malaria Journal *7*, 1–11.

737 Itoe, M.A., Sampaio, J.L., Cabal, G.G., Real, E., Zuzarte-Luis, V., March, S., Bhatia, S.N., Frischknecht,
738 F., Thiele, C., Shevchenko, A., et al. (2014). Host Cell Phosphatidylcholine Is a Key Mediator of Malaria
739 Parasite Survival during Liver Stage Infection. Cell Host & Microbe *16*, 778.

740 Jin, J., Krishack, P., and Cao, D. (2006). Role of aldo-keto reductases in development of prostate and
741 breast cancer. Frontiers in Bioscience : A Journal and Virtual Library *11*, 2767–2773.

742 Kim, Y.C., Dema, B., Rodriguez-Garcia, R., López-Camacho, C., Leoratti, F.M.S., Lall, A., Remarque,
743 E.J., Kocken, C.H.M., and Reyes-Sandoval, A. (2020). Evaluation of Chimpanzee Adenovirus and
744 MVA Expressing TRAP and CSP from Plasmodium cynomolgi to Prevent Malaria Relapse in
745 Nonhuman Primates. Vaccines *8*, 1–14.

746 Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., and Taipale, J. (2011).
747 Counting absolute numbers of molecules using unique molecular identifiers. Nature Methods *9*, 72–74.

748 Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A.,
749 and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic
750 stem cells. Cell *161*, 1187–1201.

751 Kortuem, K.M., and Stewart, A.K. (2013). Carfilzomib. Blood *121*, 893–897.

752 Krotoski, W.A., Collins, W.E., Bray, R.S., Garnham, P.C., Cogswell, F.B., Gwadz, R.W., Killick-
753 Kendrick, R., Wolf, R., Sinden, R., Koontz, L.C., et al. (1982). Demonstration of hypnozoites in
754 sporozoite-transmitted Plasmodium vivax infection. The American Journal of Tropical Medicine and
755 Hygiene *31*, 1291–1293.

756 Krotoski, W.A., Garnham, P.C.C., Cogswell, F.B., Collins, W.E., Bray, R.S., Gwasz, R.W., Killick-
757 Kendrick, R., Wolf, R.H., Sinden, R., and Hollingdale, M. (1986). Observations on early and late post-
758 sporozoite tissue stages in primate malaria. IV. Pre-erythrocytic schizonts and/or hypnozoites of
759 Chesson and North Korean strains of Plasmodium vivax in the chimpanzee. American Journal of
760 Tropical Medicine and Hygiene *35*, 263–274.

761 Kucera, M., Isserlin, R., Arkhangorodsky, A., and Bader, G.D. (2016). AutoAnnotate: A Cytoscape app
762 for summarizing networks with semantic annotations. F1000Research *5*.

763 Lachmann, A., Xu, H., Krishnan, J., Berger, S.I., Mazloom, A.R., and Ma'ayan, A. (2010). ChEA:
764 transcription factor regulation inferred from integrating genome-wide ChIP-X experiments.
765 Bioinformatics (Oxford, England) *26*, 2438–2444.

766 LaMonte, G.M., Orjuela-Sanchez, P., Calla, J., Wang, L.T., Li, S., Swann, J., Cowell, A.N., Zou, B.Y.,
767 Abdel-Haleem Mohamed, A.M., Villa Galarce, Z.H., et al. (2019). Dual RNA-seq identifies human
768 mucosal immunity protein Mucin-13 as a hallmark of Plasmodium exoerythrocytic infection. Nature
769 Communications *10*, 488.

770 Liehl, P., and Mota, M.M. (2012). Innate recognition of malarial parasites by mammalian hosts.
771 International Journal for Parasitology *42*, 557–566.

772 Liehl, P., Zuzarte-Luís, V., Chan, J., Zillinger, T., Baptista, F., Carapau, D., Konert, M., Hanson, K.K.,
773 Carret, C., Lassnig, C., et al. (2013). Host-cell sensors for Plasmodium activate innate immunity against
774 liver-stage infection. Nature Medicine 2013 20:1 *20*, 47–53.

775 Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., and Marioni, J.C. (2019). EmptyDrops:
776 distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome
777 Biology *20*, 63.

778 Lysenko Ja., A., Beljaev, A.E., and Rybalka, V.M. (1977). Population studies of Plasmodium vivax: 1.
779 The theory of polymorphism of sporozoites and epidemiological phenomena of tertian malaria. Bulletin
780 of the World Health Organization *55*, 541.

781 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R.,
782 Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of
783 Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

784 MacParland, S.A., Liu, J.C., Ma, X.-Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N.,
785 Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct
786 intrahepatic macrophage populations. Nature Communications *9*, 4383.

787 Maher, S., Vantaux, A., Cooper, C., Chasen, N., Cheng, W., Joyner, C., Manetsch, R., Witkowski, B.,
788 and Kyle, D. (2021). A Phenotypic Screen for the Liver Stages of Plasmodium vivax. BIO-PROTOCOL
789 *11*.

790 Mancio-Silva, L., Gural, N., Wadsworth, M.H., Butty, V.L., Hughes, T.K., March, S., Nerurkar, N.,
791 Roobsoong, W., Fleming, H.E., Whittaker, C., et al. (2021). Gene signatures and host-parasite
792 interactions revealed by dual single-cell profiling of Plasmodium vivax liver infection. BioRxiv
793 2021.09.01.454555.

794 Martin, H.J., and Maser, E. (2009). Role of human aldo–keto-reductase AKR1B10 in the protection
795 against toxic aldehydes. Chemico-Biological Interactions *178*, 145–150.

796 Melsted, P., Booeshaghi, A.S., Liu, L., Gao, F., Lu, L., Min, K.H. (Joseph), da Veiga Beltrame, E.,
797 Hjörleifsson, K.E., Gehring, J., and Pachter, L. (2021). Modular, efficient and constant-memory single-
798 cell RNA-seq preprocessing. Nature Biotechnology 2021 39:7 *39*, 813–818.

799 Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment Map: A Network-
800 Based Method for Gene-Set Enrichment Visualization and Interpretation. PLOS ONE *5*, e13984.

801 Miao, J., Li, J., Fan, Q., Li, X., Li, X., and Cui, L. (2010). The Puf-family RNA-binding protein PfPuf2
802 regulates sexual development and sex differentiation in the malaria parasite Plasmodium falciparum.
803 Journal of Cell Science *123*, 1039.

804 Mikolajczak, S.A., Vaughan, A.M., Kangwanrangsan, N., Roobsoong, W., Fishbaugher, M.,
805 Yimamnuaychok, N., Rezakhani, N., Lakshmanan, V., Singh, N., Kaushansky, A., et al. (2015).
806 Plasmodium vivax liver stage development and hypnozoite persistence in human liver-chimeric mice.
807 Cell Host & Microbe *17*, 526–535.

808 Miller, J.L., Sack, B.K., Baldwin, M., Vaughan, A.M., and Kappe, S.H.I. (2014). Interferon-Mediated
809 Innate Immune Responses against Malaria Parasite Liver Stages. Cell Reports *7*, 436–447.

810 Mueller, I., Jex, A.R., Kappe, S.H.I., Mikolajczak, S.A., Sattabongkot, J., Patrapuvich, R., Lindner, S.,
811 Flannery, E.L., Koepfli, C., Ansell, B., et al. (2019). Transcriptome and histone epigenome of
812 Plasmodium vivax salivary-gland sporozoites point to tight regulatory control and mechanisms for liver-
813 stage differentiation in relapsing malaria. International Journal for Parasitology *49*, 501–513.

814 Neefjes, J., Jongsma, M.L.M., Paul, P., and Bakke, O. (2011). Towards a systems understanding of
815 MHC class I and MHC class II antigen presentation. Nature Reviews Immunology 2011 11:12 *11*, 823–
816 836.

817 Paquet, T., le Manach, C., Cabrera, D.G., Younis, Y., Henrich, P.P., Abraham, T.S., Lee, M.C.S., Basak,
818 R., Ghidelli-Disse, S., Lafuente-Monasterio, M.J., et al. (2017). Antimalarial efficacy of MMV390048,
819 an inhibitor of Plasmodium phosphatidylinositol 4-kinase. Science Translational Medicine *9*.

820 Parussini, F., Coppens, I., Shah, P.P., Diamond, S.L., and Carruthers, V.B. (2010). Cathepsin L
821 Occupies a Vacuolar Compartment and is a Protein Maturase within the Endo/Exocytic System of
822 Toxoplasma gondii. Molecular Microbiology *76*, 1340.

823 Penaranda, C., and Hung, D.T. (2019). Single-Cell RNA Sequencing to Understand Host–Pathogen
824 Interactions. ACS Infectious Diseases *5*, 336–344.

825 Poran, A., Nötzel, C., Aly, O., Mencia-Trinchant, N., Harris, C.T., Guzman, M.L., Hassane, D.C.,
826 Elemento, O., and Kafsack, B.F.C. (2017). Single-cell RNA sequencing reveals a signature of sexual
827 commitment in malaria parasites. Nature *551*, 95–99.

828 Posfai, D., Maher, S.P., Roesch, C., Vantaux, A., Sylvester, K., Péneau, J., Popovici, J., Kyle, D.E.,
829 Witkowski, B., and Derbyshire, E.R. (2020). Plasmodium vivax Liver and Blood Stages Recruit the
830 Druggable Host Membrane Channel Aquaporin-3. Cell Chemical Biology.

831 Project Consortium, E. (2004). The ENCODE (ENCyclopedia of DNA Elements) Project. Science *306*,
832 636–640.

833 Qu, J., Li, J., Zhang, Y., He, R., Liu, X., Gong, K., Duan, L., Luo, W., Hu, Z., Wang, G., et al. (2021).
834 AKR1B10 promotes breast cancer cell proliferation and migration via the PI3K/AKT/NF-κB signaling
835 pathway. Cell & Bioscience *11*, 163.

836 Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler:
837 a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic
838 Acids Research *47*, W191–W198.

839 Real, E., Howick, V.M., Dahalan, F.A., Witmer, K., Cudini, J., Andradi-Brown, C., Blight, J., Davidson,
840 M.S., Dogga, S.K., Reid, A.J., et al. (2021). A single-cell atlas of Plasmodium falciparum transmission
841 through the mosquito. Nature Communications 2021 12:1 *12*, 1–13.

842 Reid, A.J., Talman, A.M., Bennett, H.M., Gomes, A.R., Sanders, M.J., Illingworth, C.J.R., Billker, O.,
843 Berriman, M., and Lawniczak, M.K.N. (2018). Single-cell RNA-seq reveals hidden transcriptional
844 variation in malaria parasites. ELife *7*, 1–29.

845 Robinson, L.J., Wampfler, R., Betuela, I., Karl, S., White, M.T., Li Wai Suen, C.S.N., Hofmann, N.E.,
846 Kinboro, B., Waltmann, A., Brewster, J., et al. (2015). Strategies for understanding and reducing the
847 Plasmodium vivax and Plasmodium ovale hypnozoite reservoir in Papua New Guinean children: a
848 randomised placebo-controlled trial and mathematical model. PLoS Medicine *12*.

849 le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson,
850 K.C., Holder, A.A., Carucci, D.J., et al. (2004). Global analysis of transcript and protein levels across
851 the Plasmodium falciparum life cycle. Genome Research *14*, 2308–2318.

852 Rosenthal, P.J. (2011). Falcipains and Other Cysteine Proteases of Malaria Parasites. Advances in
853 Experimental Medicine and Biology *712*, 30–48.

854 Roth, A., Maher, S.P., Conway, A.J., Ubalee, R., Chaumeau, V., Andolina, C., Kaba, S.A., Vantaux, A.,
855 Bakowski, M.A., Thomson-Luque, R., et al. (2018). A comprehensive model for assessment of liver
856 stage therapies targeting Plasmodium vivax and Plasmodium falciparum. Nature Communications *9*,
857 1837.

858 Ruberto, A.A., Bourke, C., Merienne, N., Obadia, T., Amino, R., and Mueller, I. (2021a). Single-cell
859 RNA sequencing reveals developmental heterogeneity among Plasmodium berghei sporozoites.
860 Scientific Reports *11*, 4127.

861 Ruberto, A.A., Bourke, C., Vantaux, A., Maher, S.P., Jex, A., Witkowski, B., Snounou, G., and Mueller,
862 I. (2021b). Single-cell RNA sequencing of Plasmodium vivax sporozoites reveals stage- and species-
863 specific transcriptomic signatures. BioRxiv 2021.11.24.469176.

864 Sà, J.M., Cannon, M. v., Caleon, R.L., Wellems, T.E., and Serre, D. (2020). Single-cell transcription
865 analysis of Plasmodium vivax blood-stage parasites identifies stage- and species-specific profiles of
866 expression. PLOS Biology *18*, e3000711.

867 Salas, F., Fichmann, J., Lee, G.K., Scott, M.D., and Rosenthal, P.J. (1995). Functional expression of
868 falcipain, a Plasmodium falciparum cysteine proteinase, supports its role as a malarial hemoglobinase.
869 Infection and Immunity *63*, 2120.

870 Sattabongkot, J., Yimamnuaychoke, N., Leelaudomlipi, S., Rasameesoraj, M., Jenwithisuk, R.,
871 Coleman, R.E., Udomsangpetch, R., Cui, L., and Brewer, T.G. (2006). Establishment of a human
872 hepatocyte line that supports in vitro development of the exo-erythrocytic stages of the malaria parasites
873 Plasmodium falciparum and P. vivax. Am. J. Trop. Med. Hyg *74*, 708–715.

874 Scarpulla, R.C. (2002). Transcriptional activators and coactivators in the nuclear control of
875 mitochondrial function in mammalian cells. Gene *286*, 81–89.

876 Schäfer, C., Dambrauskas, N., Steel, R.W., Carbonetti, S., Chuenchob, V., Flannery, E.L., Vigdorovich,
877 V., Oliver, B.G., Roobsoong, W., Maher, S.P., et al. (2018). A recombinant antibody against
878 Plasmodium vivax UIS4 for distinguishing replicating from dormant liver stages. Malaria Journal *17*,
879 370.

880 Schäfer, C., Roobsoong, W., Sattabongkot, J., Mikolajczak, S.A., and Kappe, S.H.I. (2020). A
881 Humanized Mouse Model for Plasmodium vivax to Test Interventions that Block Liver Stage to Blood
882 Stage Transition and Blood Stage Infection. IScience *23*, 101381.

883 Schäfer, C., Zanghi, G., Vaughan, A.M., and Kappe, S.H.I. (2021). Plasmodium vivax Latent Liver
884 Stage Infection and Relapse: Biological Insights and New Experimental Tools. Annual Review of
885 Microbiology *75*.

886 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B.,
887 and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular
888 Interaction Networks. Genome Research *13*, 2498–2504.

889 Singh, A.P., Buscaglia, C.A., Wang, Q., Levay, A., Nussenzweig, D.R., Walker, J.R., Winzeler, E.A.,
890 Fujii, H., Fontoura, B.M.A., and Nussenzweig, V. (2007). Plasmodium Circumsporozoite Protein
891 Promotes the Development of the Liver Stages of the Parasite. Cell *131*, 492–504.

892 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M.,
893 Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-
894 1902.e21.

895 Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., Daly, T.M., Bergman,
896 L.W., and Kappe, S.H.I. (2008). A combined transcriptome and proteome survey of malaria parasite
897 liver stages. Proceedings of the National Academy of Sciences *105*, 305–310.

898 Tharyan, R.G., Annibal, A., Schiffer, I., Laboy, R., Atanassov, I., Weber, A.L., Gerisch, B., and Antebi,
899 A. (2020). NFYB-1 regulates mitochondrial function and longevity via lysosomal prosaposin. Nature
900 Metabolism *2*, 387–396.

901 Toro-Moreno, M., Sylvester, K., Srivastava, T., Posfai, D., and Derbyshire, E.R. (2020). RNA-Seq
902 Analysis Illuminates the Early Stages of Plasmodium Liver Infection.

903 Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-
904 connected communities. Scientific Reports *9*, 1–12.

905 Tsai, H.H., Lai, H.Y., Chen, Y.C., Li, C.F., Huang, H.S., Liu, H.S., Tsai, Y.S., and Wang, J.M. (2017).
906 Metformin promotes apoptosis in hepatocellular carcinoma through the CEBPD-induced autophagy
907 pathway. Oncotarget *8*, 13832.

908 Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., and Parker, R. (2001). The
909 transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA
910 deadenylase in Saccharomyces cerevisiae. Cell *104*, 377–386.

911 Ungureanu, E., Killick-Kendrick, R., Garnham, P.C.C., Branzei, P., Romanescu, C., and Shute, P.G.
912 (1976). Prepatent periods of a tropical strain of Plasmodium vivax after inoculations of tenfold dilutions
913 of sporozoites. Transactions of the Royal Society of Tropical Medicine and Hygiene *70*, 482–483.

914 Vembar, S.S., Droll, D., and Scherf, A. (2016). Translational regulation in blood stages of the malaria
915 parasite Plasmodium spp.: systems-wide studies pave the way. Wiley Interdisciplinary Reviews. RNA
916 *7*, 772.

917 Vijayan, K., Wei, L., Glennon, E.K.K., Mattocks, C., Bourgeois, N., Staker, B., and Kaushansky, A.
918 (2021). Host-targeted Interventions as an Exciting Opportunity to Combat Malaria. Chemical Reviews
919 *121*, 10452–10468.

920 Voorberg-van der Wel, A., Roma, G., Gupta, D.K., Schuierer, S., Nigsch, F., Carbone, W., Zeeman,
921 A.M., Lee, B.H., Hofman, S.O., Faber, B.W., et al. (2017). A comparative transcriptomic analysis of
922 replicating and dormant liver stages of the relapsing malaria parasite plasmodium cynomolgi. ELife *6*.

923 Voorberg-van der Wel, A.M., Zeeman, A.M., Nieuwenhuis, I.G., van der Werff, N.M., Klooster, E.J.,
924 Klop, O., Vermaat, L.C., Kumar Gupta, D., Dembele, L., Diagana, T.T., et al. (2020). A dual fluorescent
925 Plasmodium cynomolgi reporter line reveals in vitro malaria hypnozoite reactivation. Communications
926 Biology *3*.

927 Voronina, E., Seydoux, G., Sassone-Corsi, P., and Nagamori, I. (2011). RNA Granules in Germ Cells.
928 Cold Spring Harbor Perspectives in Biology *3*.

929 Wang, C., Yan, R., Luo, D., Watabe, K., Liao, D.-F., and Cao, D. (2009). Aldo-keto Reductase Family
930 1 Member B10 Promotes Cell Survival by Regulating Lipid Synthesis and Eliminating Carbonyls. The
931 Journal of Biological Chemistry *284*, 26742–26748.

932 Wang, S.M., Lin, H.Y., Chen, Y.L., Hsu, T.I., Chuang, J.Y., Kao, T.J., and Ko, C.Y. (2019).
933 CCAAT/enhancer-binding protein delta regulates the stemness of glioma stem-like cells through
934 activating PDGFA expression upon inflammatory stimulation. Journal of Neuroinflammation *16*, 1–12.

935 White, N.J. (2011). Determinants of relapse periodicity in Plasmodium vivax malaria. Malaria Journal
936 *10*, 297.

937 White, M., Amino, R., and Mueller, I. (2017). Theoretical Implications of a Pre-Erythrocytic
938 Plasmodium vivax Vaccine for Preventing Relapses. Trends in Parasitology *33*, 260.

939 White, M.T., Walker, P., Karl, S., Hetzel, M.W., Freeman, T., Waltmann, A., Laman, M., Robinson,
940 L.J., Ghani, A., and Mueller, I. (2018). Mathematical modelling of the impact of expanding levels of
941 malaria control interventions on Plasmodium vivax. Nature Communications 2018 9:1 *9*, 1–10.

942 World Health Organization (2021). World malaria report 2021.

943 Yokomizo, T., Izumi, T., Takahashi, T., Kasamal, T., Kobayashis, Y., Satos, F., Taketanill, Y., and
944 Shimizu, T. (1993). Enzymatic Inactivation the Porcine Kidney O f Leukotriene B4 by a Novel Enzyme
945 Found in PURIFICATION AND PROPERTIES OF LEUKOTRIENE Bq 12-
946 HYDROXYDEHYDROGENASE*. THE JOURNAL OF BIOLOGICAL CHEMISTRY *268*, 18128–
947 18135.

948 Youn, J.Y., Dyakov, B.J.A., Zhang, J., Knight, J.D.R., Vernon, R.M., Forman-Kay, J.D., and Gingras,
949 A.C. (2019). Properties of Stress Granule and P-Body Proteomes. Molecular Cell *76*, 286–294.

950 Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler,
951 T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single
952 cells. Nature Communications *8*, 14049.

953

954

955 **SUPPLEMENTARY MATERIAL**

956 **Supplementary Table S1.** Comparison of metrics derived from scRNA-seq and high content imaging
957 data analyses.

958 **Figure S1. Assessment of *P. vivax* liver stages using high-content imaging and scRNA-seq.** (A)
959 Representative high content images of primary hepatocytes infected with *P. vivax* on day 5 (left) and
960 day 9 (right) post-infection. Images were obtained from one well from a 384-well plate with a 4x
961 objective. Inset: one field of view (orange box) from the same well captured with a 20x objective. Cells
962 were stained with DAPI (blue) and *Pv*UIS4 (green). White arrow: liver form assigned as a hypnozoite,
963 yellow arrow: liver form assigned as a schizont. White bar represents 1mm, grey bar represents 200µm.
964 Images are representative of replicate 2. (B) Number of *P. vivax* liver stage transcriptomes assessed
965 post-cell and gene filtering. (C) Number of *P. vivax* liver stage parasites assigned to cluster 1 and cluster
966 2 (hypnozoites and schizonts, respectively).

967 **Figure S2. *P. vivax* schizonts and hypnozoites have distinct transcriptomic signatures. (A)**
968 Scatterplot showing average $\log_2$ fold change versus the difference in the proportion of cells the
969 transcript is detected in clusters encoding for hypnozoites and schizonts. Negative values: greater
970 detection in hypnozoites; positive values: greater detection in schizonts. (B) Violin plots displaying the
971 expression of genes encoding for TCA- and glycolysis-related proteins. (C) Dot plots showing
972 transcripts with decreased expression in hypnozoites relative to schizonts. The size of the dot
973 corresponds to the percentage of cells expressing the gene, colored by average expression. Differentially
974 expressed transcripts were identified using Seurat's FindMarkers function. Wilcoxon rank-sum test,
975 adjusted p values $< 0.01$. Scale: normalized expression; pct. exp: percent of cells expressing the gene.
976 ldh: L-lactate dehydrogenase; gapdh: glyceraldehyde 3-phosphate dehydrogenase; atp-synthC: ATP
977 synthase subunit C; gdh1: NADP-specific glutamate dehydrogenase 1.

978 **Figure S3. Carfilzomib inhibits *P. vivax* schizonts and hypnozoites.** (A) Violin plots displaying the
979 distribution of expression of select genes encoding for proteasome subunits in schizonts and
980 hypnozoites. (B) Table of proteasome-associated transcripts differentially expressed between schizonts
981 and hypnozoites. Positive average log2FC: higher expression in schizonts. (C) Structure of carfilzomib.
982 (D) Dose-response curve of carfilzomib-induced inhibition of *P. vivax* schizont growth area (yellow,
983 IC50 288nM), hypnozoite quantity per well (blue, IC50 511nM), and hepatocyte nuclei quantity (orange,
984 IC50 3.32 µM). Data shown are pooled from three independent experiments, each containing duplicate
985 wells at each concentration. Bars represent ± SEM.

986 **Figure S4. Assessment of hypnozoite transcriptomes.** (A) UMAP of hypnozoites colored by cluster
987 and faceted by day post-infection. (B) Violin plot showing the average $log_2FC$ of differentially expressed
988 genes in each cluster. Markers were identified using Seurat's AllFindMarkers function. Transcripts with
989 an adjusted p value < 0.05 (Wilcoxon rank-sum) and average $log_2FC > 0.5$ were deemed markers. Red
990 dashed line: average $log_2FC = 0.5$.

991 **Figure S5. Analysis of host-pathogen transcriptional signatures in *P. vivax*- infected and non-**
992 **infected hepatocytes.** (A) Number of hepatocytes assessed. Colored by infection status and faceted by
993 sample (left); proportion of infected, exposed, or naive hepatocytes assessed on day 5 or day 9 post-
994 infection. (B) t-SNE plot of hepatocytes colored by day. (C) t-SNE plots of hepatocytes colored by
995 percent expression of zonation markers in human hepatocytes. Markers obtained from Macparland et al.
996 (2019). (D) t-SNE plots of hepatocytes colored by hypnozoite infection status. Grey, uninfected (naive
997 and exposed) hepatocytes; red, hepatocytes infected with C1 hypnozoites; blue, hepatocytes infected
998 with C2; teal, hepatocytes infected with C3 hypnozoites. (E) Unique and overlapping gene sets in
999 hepatocytes infected with schizonts or hypnozoites. Positive enrichment derived from genes displaying
1000 greater expression versus non-infected hepatocytes; negative enrichment derived from transcripts
1001 displaying decreased expression versus noninfected hepatocytes.

1002 **Figure S6. Human AKR1B10 is upregulated in hepatocytes infected with *P. vivax*.**
1003 (A) Representative confocal images of *P. vivax* parasites on day 12 post-infection of hepatocytes. Cells
1004 were stained with DAPI (blue), *Pv*UIS4 (green), and HsAKR1B10 (red). White bar represents 100µm.
1005 (B) Violin plots displaying the distribution of *AKR1B10* transcript levels in infected (schizonts and
1006 hypnozoites) and non-infected (exposed and naive) hepatocytes. Black dot: mean expression ($log_2$). (C)
1007 Violin plot displaying the distribution of AKR1B10 signal in hepatocyte cultures infected with *P. vivax.*
1008 (D) Associated with panel C, violins plots displaying the distribution of AKR1B10 signal across
1009 biological replicates. IHC: immunohistochemistry; SZ: hepatocyte containing schizont; HZ: hepatocyte
1010 containing hypnozoite; EX: hepatocyte not infected but exposed to *P. vivax.* ****: p <= 0.0001; *: p <=
1011 0.05; n.s: not significant, p >0.05.

1012 **Data Set S1.** Data associated with schizonts versus hypnozoites analyses. (A) Differentially expressed
1013 genes defining schizonts and hypnozoites. Positive Average log 2 FC: Upregulated in hypnozoites
1014 relative to schizonts. (B) Differentially expressed genes identified as encoding for cellular fating RNA-
1015 binding proteins. (C) GO Term enrichment of Hypnozoite or Schizont transcripts identified as cellular
1016 fating RBPs. Rows highlighted in light green are terms with Bonferroni adjusted p value < 0.05. Input
1017 parameters: P value cutoff = 0.05; Limit to GO Slim terms = No; Evidence: Computed, curated;
1018 PlasmoDB v55.

1019 **Data Set S2.** Data associated with analyses of hypnozoite-to-hypnozoite heterogeneity. (A) Markers
1020 defining each hypnozoite cluster. Marker defined as: Average log 2 FC > 0.5, Adjusted p value < 0.05.
1021 Rows colored green are transcripts that meet marker criteria. (B) Differentially expressed genes
1022 identified as encoding for cellular fating RNA-binding proteins in hypnozoite clusters. (C) GO Term
1023 enrichment of hypnozoite transcripts identified as cellular fating RBPs. Rows highlighted in light green
1024 are terms with Bonferroni adjusted p value < 0.05. P value cutoff = 0.05; Limit to GO Slim terms = No;
1025 Evidence: Computed, curated; PlasmoDB v55.

1026 **Data Set S3.** Data associated with host-specific transcriptional signatures in response to *P. vivax*
1027 infection. (A) Differentially expressed genes in hepatocytes containing schizonts versus non-infected
1028 cells. Positive average log2FC: greater expression in infected hepatocytes. (B) Differentially expressed
1029 genes in hepatocytes containing hypnozoites versus non-infected cells. Positive average log2FC: greater
1030 expression in infected hepatocytes. (C) Enrichment of GO biological processes and Reactome metabolic
1031 pathways. (D) Unique and overlapping enriched gene sets in hepatocytes infected with schizonts or
1032 hypnozoites. (E) Transcription factor enrichment analysis. Rows highlighted in green are sets that are
1033 significantly enriched, adjusted p value < 0.05.

Figure 1

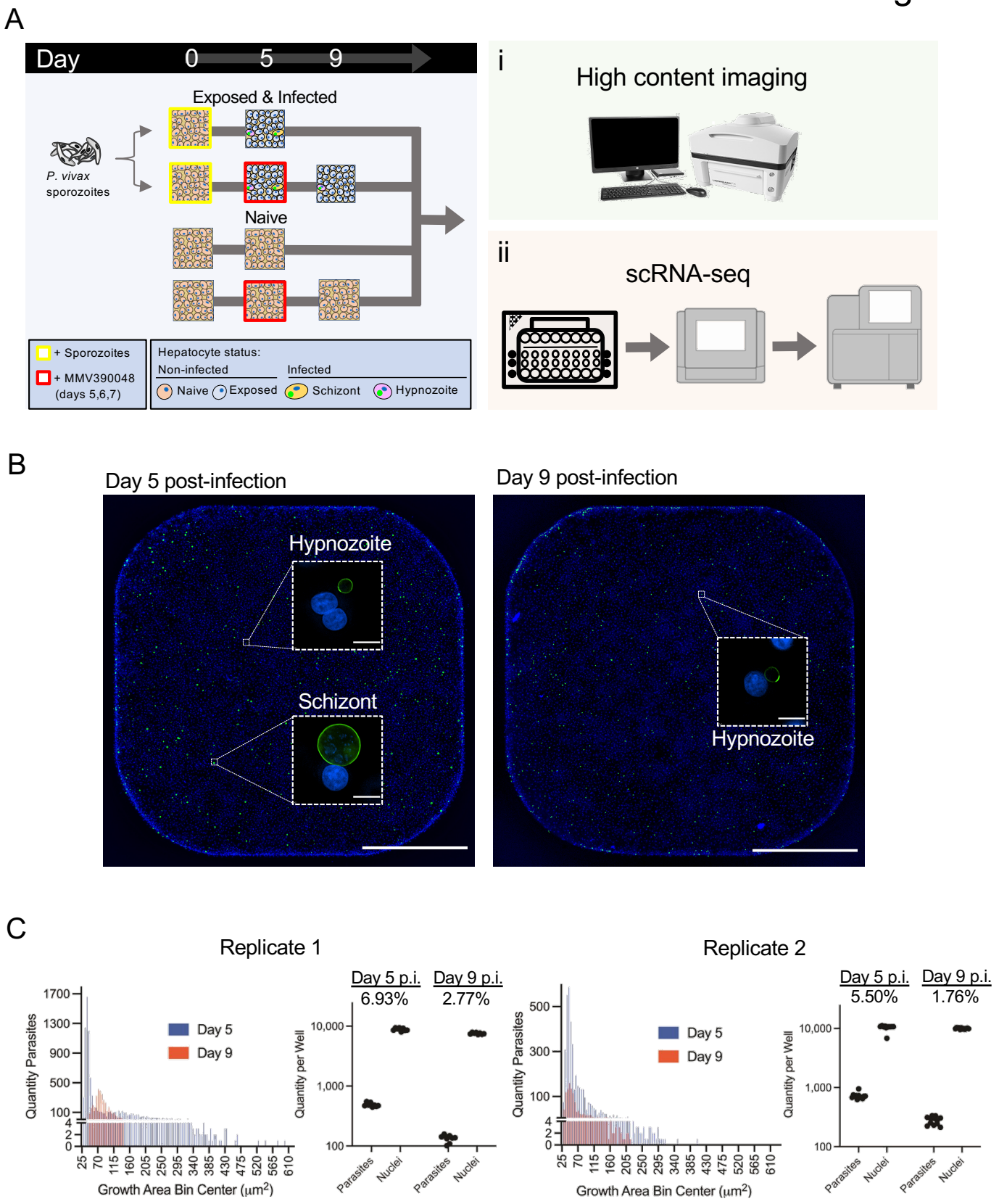



**Figure 1. Design and validation of detection of *P. vivax* liver stages using high-content imaging and scRNA-seq.** (A) Schematic illustrating the sample preparation and data generation pipeline used to assess *P. vivax* liver stage infection rates and gene expression. (B) Representative high content images of primary hepatocytes infected with *P. vivax* on day 5 (left) and day 9 (right) post-infection (p.i). Images were captured from one well from a 384-well plate with a 4x objective. Inset: one field of view (white dashed box) from the same well captured with a 20x objective. Cells were stained with DAPI (blue) and *Pv*UIS4 (green). White bar represents 1mm, inset bars represent 10µm. Images are representative of replicate 1. (C) Histogram displaying the distribution of parasite growth areas fixed at day 5 or day 9 post-infection (left) and net quantity of parasites and hepatocyte nuclei (right). Net infection rate is indicated for each endpoint. Growth and infection metrics are shown for both replicates.

Figure 2

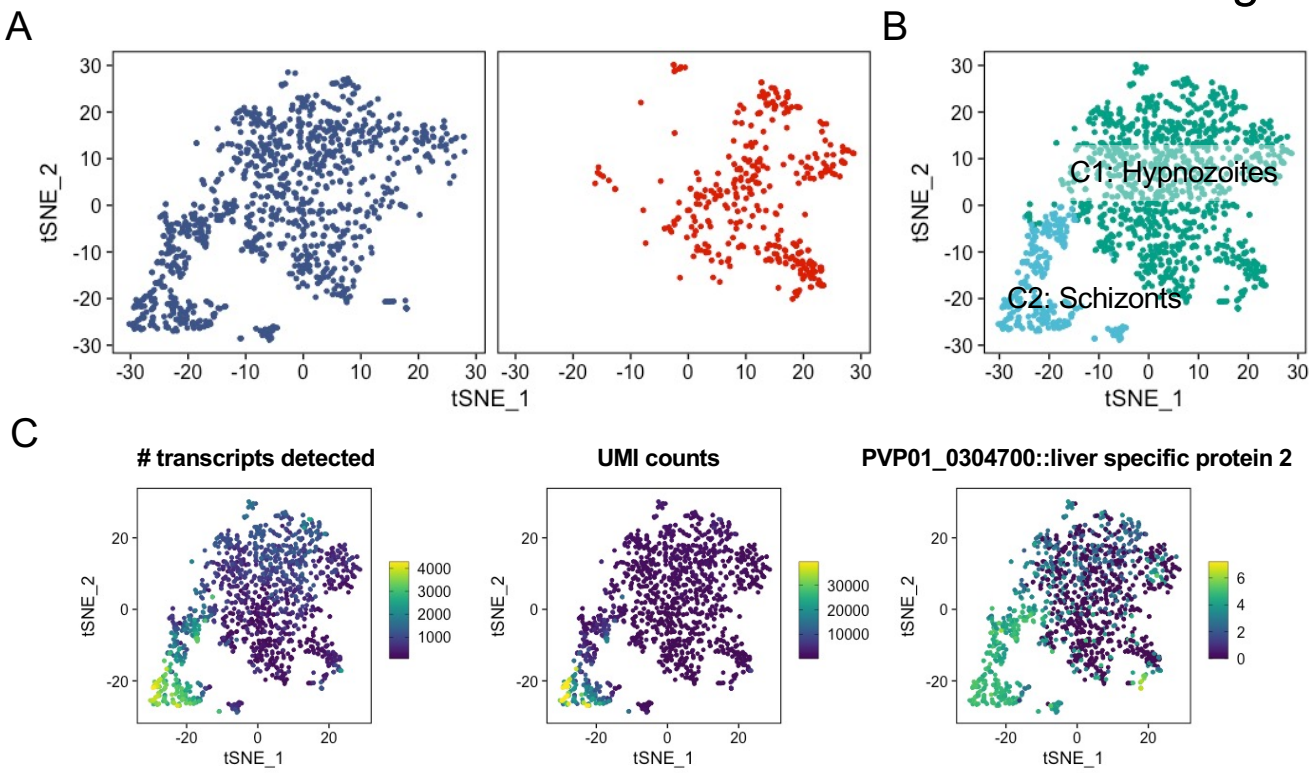
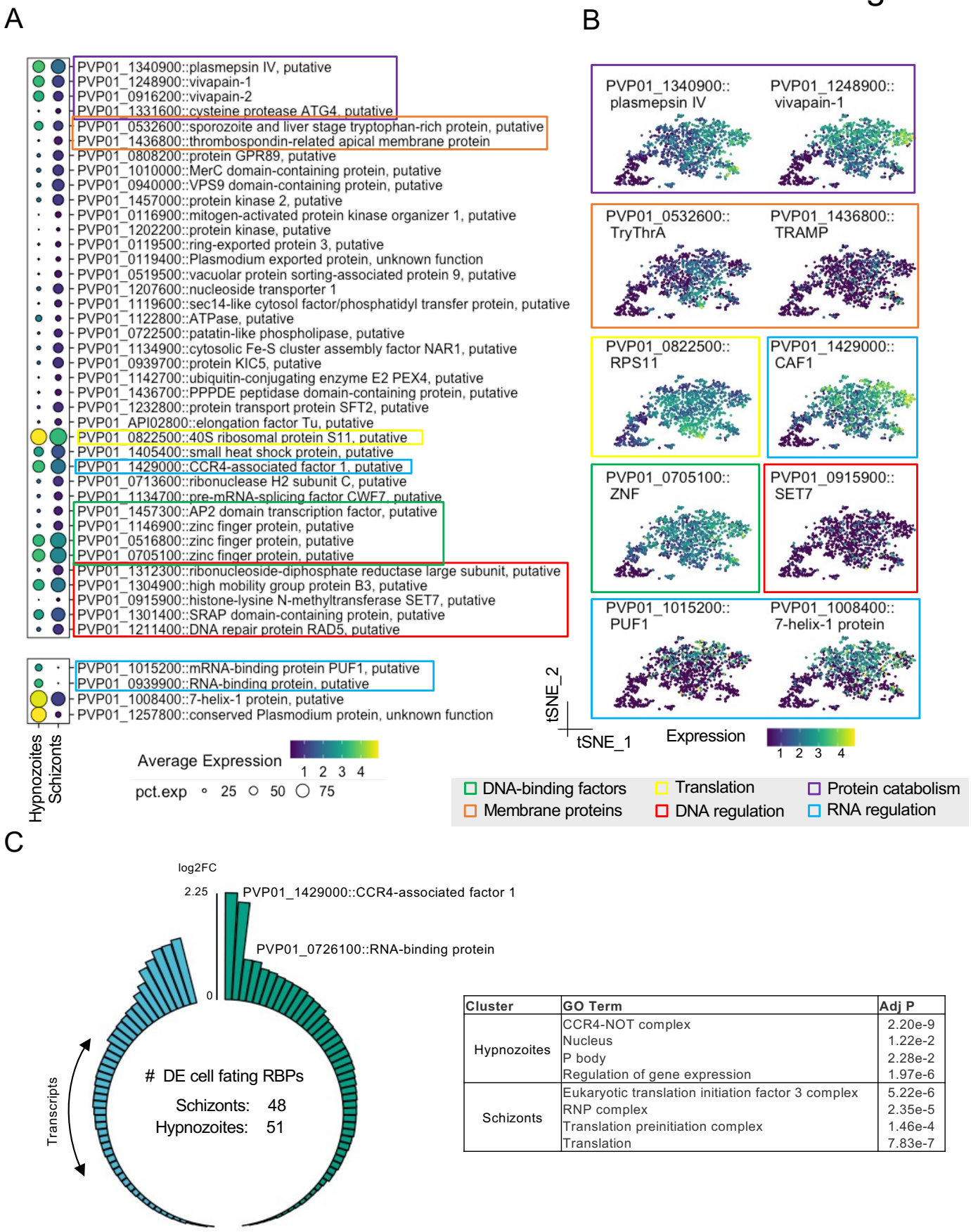



**Figure 2. Low dimensional representation of *P. vivax* schizont and hypnozoite transcriptomes** (A) t-SNE plots of integrated *P. vivax* liver stage scRNA-seq data faceted by parasites isolated on day 5 (left) and day 9 (right) post-infection. (B) t-SNE plot of *P. vivax* liver stage data colored by parasites encoding for putative hypnozoites (C1) and schizonts (C2). (C) t-SNE plots of *P. vivax* liver stage scRNA-seq data colored by number of transcripts detected (left), UMI counts (middle), and the expression of the schizont marker, LISP2 (PVP01_0304700) (right). Scales: absolute count (left, middle), normalized expression (right).

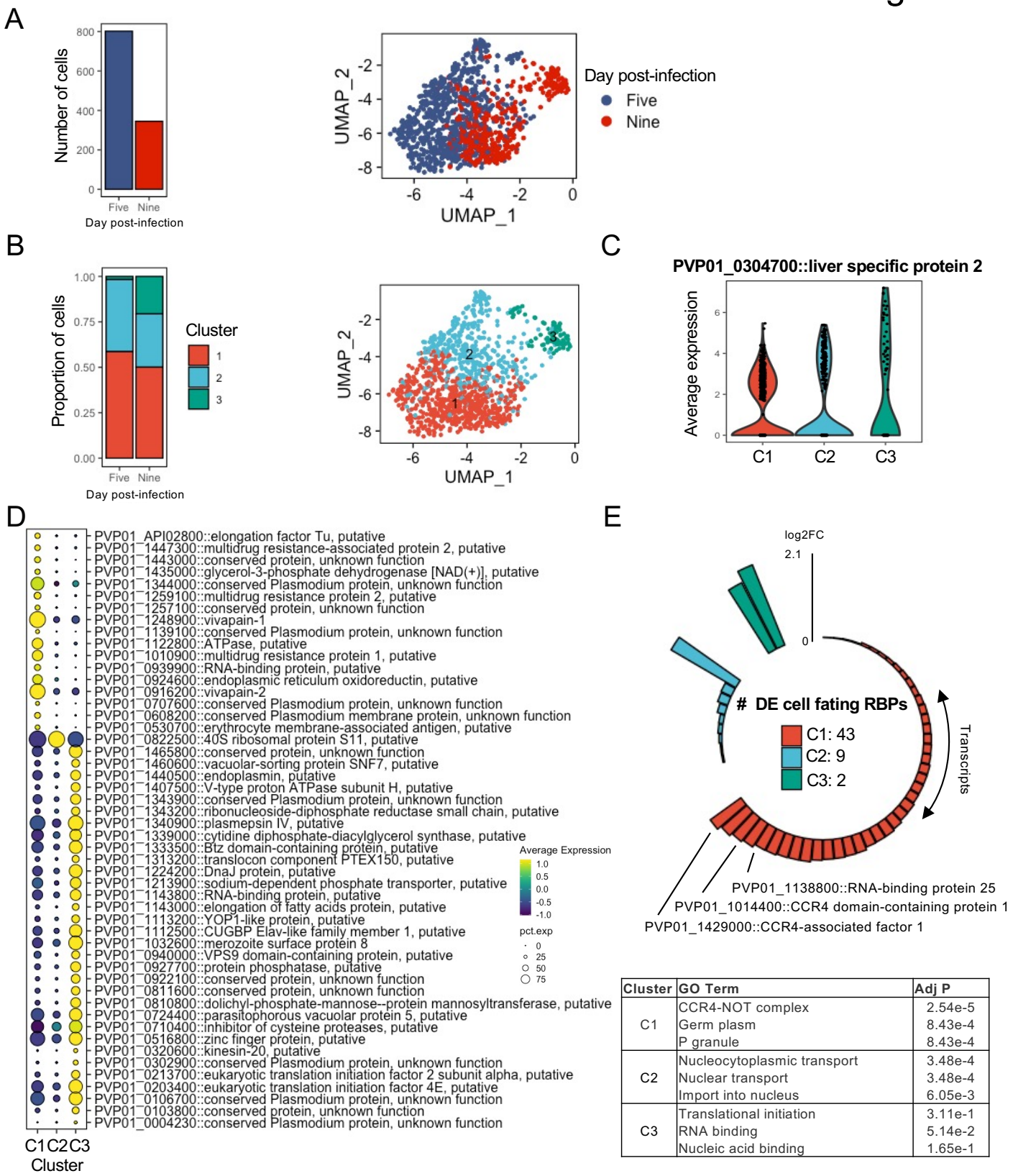
Figure 3




| Cluster | GO Term | Adj P |
|---|---|---|
| Hypnozoites | CCR4-NOT complex | 2.20e-9 |
| | Nucleus | 1.22e-2 |
| | P body | 2.28e-2 |
| | Regulation of gene expression | 1.97e-6 |
| Schizonts | Eukaryotic translation initiation factor 3 complex | 5.22e-6 |
| | RNP complex | 2.35e-5 |
| | Translation preinitiation complex | 1.46e-4 |
| | Translation | 7.83e-7 |

**Figure 3. *P. vivax* hypnozoites have distinct transcriptomic signatures**. (A) Dot plots showing transcripts that distinguish all hypnozoites from schizonts. The size of the dot corresponds to the percentage of cells that transcript was detected per condition, colored by average expression. Differentially expressed transcripts were identified using Seurat's FindMarker function (Wilcoxon rank-sum test, adjusted p values < 0.01). (B) tSNE plots of liver stages colored by transcripts displaying greater levels in hypnozoites. Colored boxes in A and B indicate the family or canonical biological process that the transcript is associated with. Scale: normalized expression; pct. exp: percent of cells expressing transcript. (C) Number of transcripts in schizonts and hypnozoites identified as cellular fating RBPs (left) and GO term enrichment of this subset. Adj P: Bonferroni adjusted p value.

Figure 4




| Cluster | GO Term | Adj P |
|---|---|---|
| C1 | CCR4-NOT complex | 2.54e-5 |
| | Germ plasm | 8.43e-4 |
| | P granule | 8.43e-4 |
| C2 | Nucleocytoplasmic transport | 3.48e-4 |
| | Nuclear transport | 3.48e-4 |
| | Import into nucleus | 6.05e-3 |
| C3 | Translational initiation | 3.11e-1 |
| | RNA binding | 5.14e-2 |
| | Nucleic acid binding | 1.65e-1 |

**Figure 4. Transcriptomic signatures of reactivating and persister hypnozoites.** (A) Number of hypnozoites assessed (left) and UMAP of cells identified as hypnozoites colored by day of collection. (B) Proportion of cells in clusters on days 5 and 9 post-infection (left) and UMAP of hypnozoites colored by cluster. (C) Violin plot showing the expression of LISP2 in each cluster. (D) Dot plot showing markers that distinguish hypnozoites in each cluster. The size of the dot corresponds the percentage of cells in the expressing the gene of interest, colored by scaled normalized expression of the transcript in each cluster. Marker transcripts were identified using Seurat's AllFindMarkers function. Wilcoxon rank-sum test, average $\log_2 FC > 1$, adjusted p value $< 0.01$. (E) Number transcripts in each cluster encoding for cellular fating RBPs sorted by their average log2FC relative to the other clusters (top) and GO Term enrichment of transcripts in each cluster (bottom). Adj P: Bonferroni adjusted p value.

Figure 5

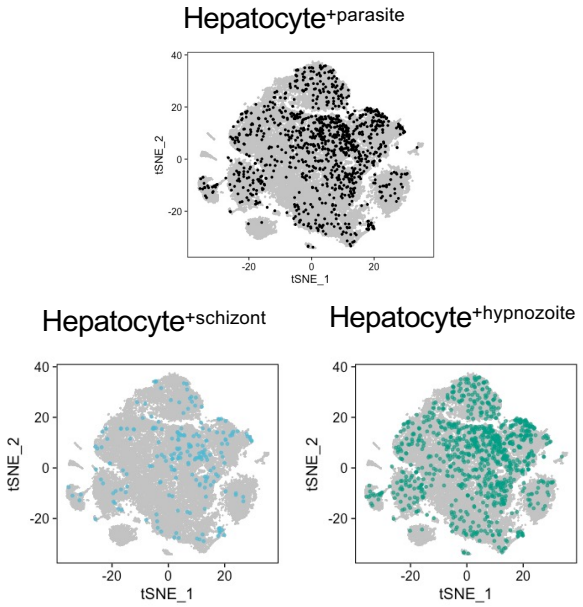



**Figure 5. Analysis of host-pathogen transcriptional signatures in *P. vivax*- infected and non-infected hepatocytes.** t-SNE plots of hepatocytes faceted by infection status. Grey, noninfected (naive and exposed) hepatocytes; light blue, hepatocytes infected with schizonts; teal, hepatocytes infected with hypnozoites.

Figure 6

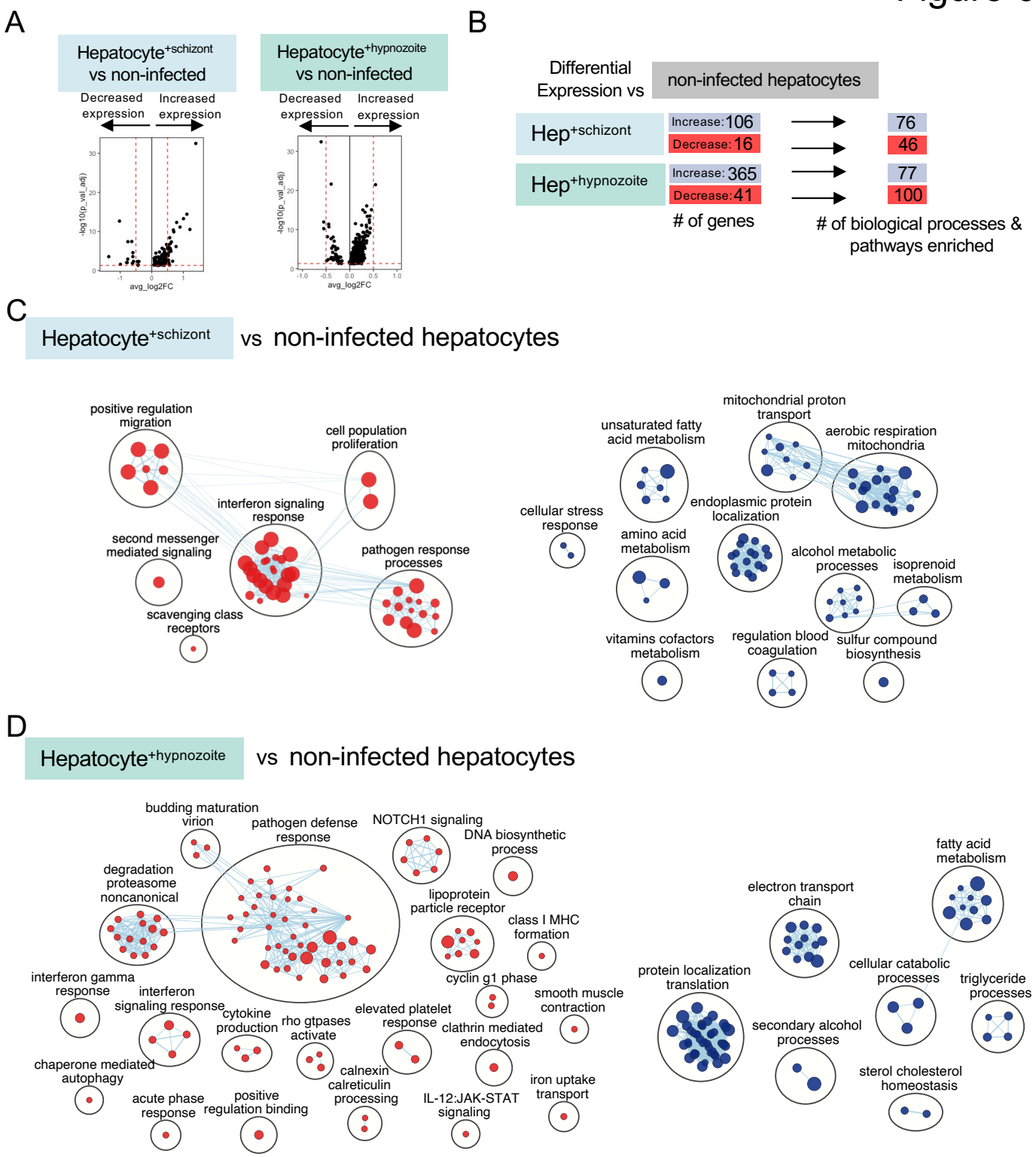



**Figure 6. Analysis of host transcriptional signatures in *P. vivax* infected hepatocytes.** (A) Volcano plots showing changes in gene expression in infected hepatocytes versus non infected (Wilcoxon rank sum, adjusted p value < 0.05). Positive fold change (FC) values represent genes with greater expression, and negative FC values represent genes with decreased expression. Dashed horizontal red lines: adjusted p value = 0.05; dashed vertical red lines: $\log_2 FC = 0.5$. (B) Summary of differential gene expression data shown in the volcano plots and of the enrichment analyses. Hep: hepatocyte. (C) Enrichment map of cellular processes and pathways associated with hepatocyte infection with *P. vivax* schizonts. (D) Enrichment map of cellular processes and pathways associated with hepatocyte infection with *P. vivax* hypnozoites. For the maps depicted panels C and D, node size is proportional to the number of genes identified in each gene set (minimum 3 genes/gene set); each node represents a distinct biological process or pathway derived from gene with decreased expression (red) or increased expression (blue) versus non-infected cells; and edges (blue lines) represent the number of genes overlapping between two processes or pathways.

Figure 7

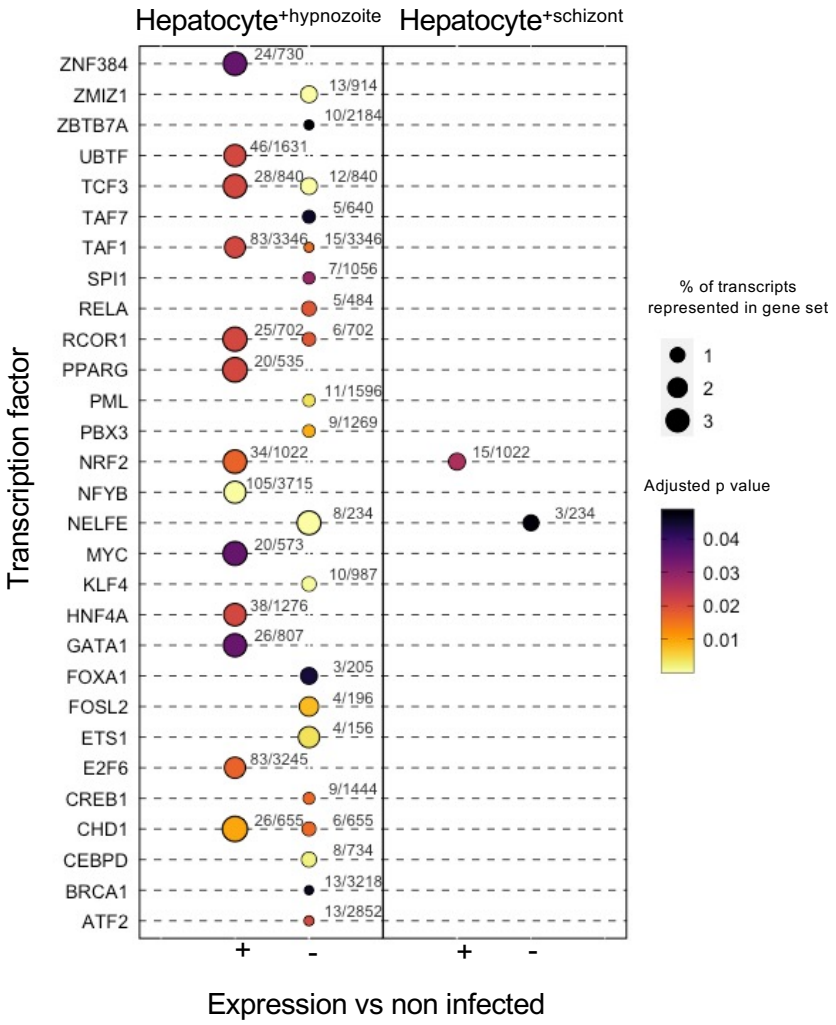



**Figure 7. Transcription factor enrichment in hepatocytes infected with schizonts or hypnozoites.** Dot plot of transcription factors enriched in Hepatocyte$^{+hypnozoite}$ (left) or Hepatocyte$^{+schizont}$ sets (right). Positive (+) enrichment derived from genes displaying greater expression versus non-infected hepatocytes; negative (-) enrichment derived from genes displaying decreased expression versus noninfected hepatocytes. ENCODE (Feingold et al., 2004) and ChEA (Lachmann et al., 2010) consensus transcription factors from ChIP-X (downloaded from Enrichr (Kuleshov et al., 2016)).