Edinburgh Research Explorer

Antibody-independent mechanisms regulate the establishment of chronic Plasmodium infection

1   # Antibody-independent mechanisms regulate the establishment of
2   chronic *Plasmodium* infection
3

4   Thibaut Brugat[1]*†, Adam James Reid[2]*†, Jingwen Lin[1], Deirdre Cunningham[1,] Irene

5   Tumwine[1], Garikai Kushinga[1], Sarah McLaughlin[1], Philip Spence[4]‡, Ulrike Böhme[2],

6   Mandy Sanders[2], Solomon Conteh[5], Ellen Bushell[2], Tom Metcalf[2], Oliver Billker[2],

7   Patrick E. Duffy[5], Chris Newbold[2,3], Matthew Berriman[2] & Jean Langhorne[1]*.

8

9   [1] The Francis Crick institute, London NW1 1AT, UK.

10   [2] Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.

11   [3] Weatherall Institute of Molecular Medicine, Oxford OX3 9DS, UK

12   [4] MRC National Institute for Medical Research, London NW7 1AA, UK.

13   [5] Laboratory of Malaria Immunology and Vaccinology, National Institute of Allergy and

14   Infectious Diseases, National Institutes of Health, Rockville, MD, United States of

15   America.

16

17   †These authors contributed equally to the work

18   *Corresponding Authors: Jean Langhorne and Thibaut Brugat, Francis Crick institute,

19   London NW1 1AT. jean.langhorne@crick.ac.uk and thibaut.brugat@crick.ac.uk, tel: +44

20   208 816 2558; Adam James Reid, Wellcome Trust Sanger Institute, Genome Campus,

21   Hinxton, Cambridgeshire, CB10 1SA. ar11@sanger.ac.uk, tel: +44 1223 494810.

22   ‡Present Address: Institute of Immunology and Infection Research (IIIR), School of

23   Biological Sciences, The University of Edinburgh, United Kingdom

24

Malaria is caused by parasites of the genus *Plasmodium*. All human-infecting *Plasmodium* species can establish long-lasting chronic infections[1-5], creating an infectious reservoir to sustain transmission[1,6]. It is widely accepted that maintenance of chronic infection involves evasion of adaptive immunity by antigenic variation[7]. However, genes involved in this process have been identified in only two of five human-infecting species: *P. falciparum* and *P. knowlesi*. Furthermore, little is understood about the early events in establishment of chronic infection in these species. Using a rodent model we demonstrate that only a minority of parasites from among the infecting population, expressing one of several clusters of virulence-associated *pir* genes, establishes a chronic infection. This process occurs in different species of parasite and in different hosts. Establishment of chronicity is independent of adaptive immunity and therefore different from the mechanism proposed for maintainance of chronic *P. falciparum* infections[7-9]. Furthermore, we show that the proportions of parasites expressing different types of *pir* genes regulate the time taken to establish a chronic infection. Since *pir* genes are common to most, if not all, species of *Plasmodium*[10], this process may be a common way of regulating the establishment of chronic infections.

42

Long-lasting *Plasmodium falciparum* infections are thought to be maintained by cytoadherence to avoid splenic clearance and antigenic variation of the adherent proteins to avoid clearance by antibodies. We wanted to know how chronic infections are established by malaria parasites lacking genes known to be involved in these processes (*var* and *sicavar*). We used, as an example, the rodent malaria parasite *Plasmodium chabaudi chabaudi* AS that gives rise to chronic infections in laboratory mice (Figure 1A). The acute infection is defined by a peak of parasitaemia (5-10% iRBC) approximately ten days after mosquito bite, and clearance of the majority of parasites by day fifteen. A chronic infection is then established, giving rise to several episodes of patent parasitemia (>0.01%) for up to eighty days. Comparison of the transcriptomes of parasite populations from the acute and chronic phases of infection showed that among the multigene families identified in the *P. chabaudi* AS genome, the *pir* family is the most differentially expressed (Figure 1A, Supplementary Tables 1 and 2). This multigene family is present in most *Plasmodium* genomes[11] and has been associated with virulence[12]. More than half of the *pir* genes were differentially expressed between acute and chronic phases. The majority of these genes had significantly higher expression in the acute phase of infection, whereas only a few had higher expression in the chronic phase.

A *Plasmodium*-specific antibody response is detectable from day seven of infection[13] and thus the change in *pir* expression might be the result of adaptive immunity-dependent selection of parasites expressing certain *pir* genes. However, we observed the same change in mice lacking B cells and antibodies (μMT; Figure 1B) or lacking T cells able to promote an antibody response (TCRα-/-; Figure 1C;

65   Supplementary Figure 1; Supplementary Table 1). Therefore the transcriptomic changes

66   observed between acute and chronic phases are independent of adaptive immunity and

67   distinct from the mechanism suggested to be involved in maintenance of chronic

68   infection in *P. falciparum*[9].

69   In the genomes of most malaria parasites, *pir* genes are located in subtelomeric

70   regions that are difficult to resolve based on the *de novo* assembly of short sequencing

71   reads. Using Single Molecule Real-Time (SMRT) sequencing, we generated a new

72   genome assembly for *P. chabaudi* AS comprising a complete set of fourteen

73   chromosomes with no gaps (Supplementary Table 3). We observed that some

74   subtelomeres had similar sets of *pir* genes arranged in clusters (Supplementary Figures 2

75   and 3*)*. In recent work, *pir* genes were classified into several short and long forms (S1-7

76   and L1-4, respectively)[14]. We found eight clusters with common structures: three were

77   enriched for the L1 subfamily and five were enriched for the S7 subfamily (Figure 2;

78   Supplementary Table 4). Interestingly, we found that the majority of *pir* expression in our

79   samples came from this small number of clusters, and that genes within clusters were co-

80   expressed during infection (Supplementary Figure 4). While the S7-rich clusters were

81   highly expressed during the acute phase, the chronic phase was dominated by expression

82   of L1-rich loci (Figure 2C; Supplementary Table 5). This pattern was replicated in the

83   αβT-cell receptor- and B-cell knockout mice (Figure 2C) reinforcing our earlier

84   conclusion that changes in *pir* gene expression are independent of selection by the

85   adaptive immune response. Individual L1- or S7-rich loci are hereafter termed

86   *Chronicity-Associated Pir Locus* (ChAPL), and *Acute-Associated Pir Locus* (AAPL),

87   respectively. We identified an intergenic motif, commonly occurring upstream of most

88   ChAPL genes, that might serve as a promoter sequence (Supplementary Figure 5).

89   However, as only a subgroup of ChAPL genes are expressed in each mouse (Figure 2C),

90   it is likely that epigenetic mechanisms regulate access of transcription factors to this

91   motif, as shown for subtelomeric gene families in *P. falciparum*[15]. Of the three ChAPLs,

92   we observed a preference for clusters on chromosome 6.

93   We recently showed that increased *P. chabaudi* virulence during serial blood

94   passage (SBP) is associated with a reduced repertoire of *pir g*ene transcripts compared

95   with parasites transmitted by mosquito (MT)[12]. The gene most highly upregulated in SBP

96   parasites was a *pir* from the L1 subfamily (PCHAS_1100300). Although not found in a

97   ChAPL, this gene is from the same subfamily that is most common in these loci. We

98   therefore hypothesized that chronic phase parasites expressing ChAPLs would have a

99   similar phenotype. To test this, we injected acute and chronic phase parasites into naïve

100   mice. Chronic parasites gave rise to significantly greater parasitaemias and pathology

101   than parasites derived from the acute phase of infection (Figure 3A; Supplementary

102   Figure 6A). This suggests that parasite populations derived from the chronic phase

103   expand better in the face of the response of a naive mouse than parasites derived from the

104   acute phase. However, as chronic parasites only reached high parasitaemias when

105   injected into naïve mice, our data also indicate that they are better controlled when the

106   host has previously experienced an acute infection. This pattern was also observed when

107   chronic parasites were collected from B-cell knockout mice (Supplementary Figure 6B).

108   Chronic infections therefore involve the emergence of phenotypically distinct parasites

109   expressing ChAPLs, independently of adaptive immunity.

110   Thus we have established that there are transcriptional and phenotypic changes

111   observed during chronic *Plasmodium* infections and that they do not stem from classical

112   antigenic variation. We then sought to determine whether these were due to a

113   transcriptional switch at the population level, or to a selection of pre-existing parasites

114   from the acute phase. We isolated single parasites from the acute and chronic phases of

115   infection, initiated clonal infections in C57Bl/6 mice and performed RNA sequencing

116   analyses of these cloned populations of parasites (Figure 3B). We found that one out of

117   ten acute-phase clones was more virulent than the population as a whole (Figure 3B;

118   Supplementary Figure 7). While avirulent clones expressed AAPLs, the virulent clone

119   expressed a ChAPL, as predicted by our finding that ChAPL-expressing parasites are

120   more virulent in naïve mice (Figure 3C; Supplementary Figure 8; Supplementary Table

121   6). As expected, the clones from the chronic phase were all virulent and expressed

122   ChAPLs (Figures 3B & 3C, Supplementary Figures 8 & 9). This suggests that the

123   establishment of chronic infection involves selection of a small proportion of parasites

124   expressing ChAPLs from a clonally variant population largely expressing AAPLs.

125   On the basis of these observations, we hypothesized that the proportion of virulent

126   parasites already in the acute phase might affect the success of establishing a chronic

127   infection. In our initial transcriptome experiment, we found that high expression of

128   ChAPLs during the acute phase was associated with a more rapid recrudescence

129   (Supplementary Table 7). We then tested whether chronic-like parasites, expressing L1

130   *pir* genes, outgrew acute parasites largely expressing AAPLs. We used SBP parasites as a

131   proxy for chronic parasites as they express the L1 *pir* PCHAS_1100300, and have a

132   similar phenotype to ChAPL-expressing chronic parasites. We performed mixed

133 infections in C57BL/6 mice with different ratios of SBP and acute MT (expressing

134 AAPLs) parasites tagged with different fluorescent markers (Figure 4, Supplementary

135 Figure 10 A-C; Supplementary Table 8). SBP parasites rapidly outcompeted MT

136 parasites (Figure 4; Supplementary Figure 10D), confirming that the more virulent,

137 chronic-like parasites had a selective advantage, as has been shown previously for

138 different parasite isolates[16-20]. The selection of a minority of iRBCs, expressing

139 virulence-associated *pir* genes (Figures 3B & 3C), can therefore explain the change in

140 expression from S- to L-*pir* genes observed in the parasite population between the acute

141 and chronic phases of infection (Figures 1 & 2). As we had hypothesized, higher doses

142 of L1-expressing parasites resulted in an earlier recrudescence confirming that these

143 parasites have a better capacity to establish a chronic infection (Figure 4, Supplementary

144 Figure 11).

145 When a similar infection mixing SBP and MT parasites was carried out in RAG1-/- mice

146 (lacking both T and B cells) and RAG-/- mice treated with chlodronate liposomes to

147 deplete macrophages (Supplementary Figure 12A), virulent parasites similarly rapidly

148 outgrew AAPL-expressing parasites. Thus, this early selection takes place even in the

149 absence of B cells, T cells, and macrophages indicating that other mechanisms are

150 operating (Supplementary Figure 12B).

151 Unexpectedly, when mice were infected with 10^4 virulent parasites plus 10^5

152 avirulent parasites, the acute parasitaemia and pathology were lower than in mice

153 infected with 10^4 virulent parasites alone (Figure 4B, Supplementary Figure 11). This

154 suggests that AAPL-expressing parasites act to lower the virulence of parasites

155 expressing largely virulence-associated L1 *pirs* and delay the onset of chronicity.

156   Therefore, the initial composition of parasites expressing different *pir* genes regulates

157   parasite recrudescence.

158   To determine whether our findings apply more generally, beyond *P. chabaudi* AS

159   infections of laboratory mice, we established chronic infections with alternative

160   host/parasite combinations. We used the more virulent *P. chabaudi* CB strain in C57BL/6

161   mice and *P. chabaudi* AS in the African thicket rat *Grammomys surdaster*. This is a

162   natural host of rodent *Plasmodium* (Supplementary Figures 13A & 13B).   Reduced

163   AAPL and increased ChAPL gene expression during chronic *P. chabaudi* infection were

164   observed in both cases, suggesting that our findings are a general feature of *P. chabaudi*

165   chronic infections irrespective of strain of parasite or host (Supplementary Figures 13C &

166   13D). We then investigated chronic *P. berghei* infections in Brown Norway rats

167   (Supplementary Figure 14A). The subtelomeric regions and *pir* repertoire of *P. berghei*

168   are quite different from those of *P. chabaudi* [14]. Loci resembling the AAPLs and ChAPLs

169   of *P. chabaudi* are absent. Nevertheless, we observed a general decrease in *pir* expression

170   during the chronic phase with an increase in expression of L2 *pir* genes (Supplementary

171   Table 9; Supplementary Figure 14B). The L2 *pir* genes of *P. berghei* are similar to the *P.*

172   *chabaudi* L1 genes we found in ChAPLs. They are present in recently copied loci

173   (Supplementary Figure 14C), contain long variable central domains and have a longer N-

174   terminal domain that may contain a second transmembrane helix (Supplementary Figure

175   15). This suggests our findings are a general feature of malaria parasites infecting rodents.

176   Our results show that mosquito transmission of *P. chabaudi* gives rise to a

177   clonally variant population of parasites, the majority expressing AAPLs and a minority

178   ChAPLs. During the acute phase of infection, AAPL-expressing parasites are controlled

179   by the host response, while ChAPL-expressing parasites survive and establish a chronic

180   infection. The more ChAPL-expressing parasites that are present early in infection, the

181   more rapidly the chronic infection is established.

182

183   While it is thought that maintenance of chronic *P. falciparum* infection depends

184   on the differential expression of *var* genes[7,8], how similar infections are established and

185   maintained by *P. vivax, P. malariae* and *P. ovale,* that lack these variant antigens, is

186   unknown. Here, we show that the establishment of chronic *Plasmodium* infection in

187   rodent models is reproducibly associated with phenotypically distinct parasites,

188   expressing particular *pir* gene clusters. Importantly, contrary to what is understood about

189   maintenance of chronic infection in *P. falciparum*[9], this change in the parasite population

190   is independent of adaptive immunity. Antigenic variation may still operate in maintaining

191   chronicity, but our results clearly show that other mechanisms allow for the establishment

192   of chronic infection in parasites that infect rodents.

193   Our finding that the initial population of parasites comprises individuals, each

194   expressing different *pir* genes, suggests that the chronic infection is established by a

195   small number of parasites which evade the early immune response. The selective

196   advantage during infection of ChAPL-expressing over AAPL-expressing parasites could

197   be of multiple origins. Although our previous results showed that growth rates of

198   avirulent MT parasites and virulent SBP parasite were similar in RAG1-/- mice[12], red cell

199   invasion efficiency might still play a role in selection of virulent clones for establishment

200   of chronic infection. We are also investigating rosette formation[21,22], red cell

201   deformability[23], sequestration[24], and the capacity to escape a macrophage-independent

202 innate immune response[25] as possibilities. It has been shown that the expression of

203 subsets of *P. falciparum rif* genes (which may be relatives of *pir* genes[10]) and *P. vivax pir*

204 genes are associated with cytoadherence of iRBCs to human endothelial cells[26,27] and red

205 blood cells[21,28,29], two processes implicated in parasite survival and pathogenicity[30].

206 Furthermore, it has previously been suggested that the same molecules are involved in

207 cytoadherence and immune evasion in *P. chabaudi*[9,24]. Differences in cytoadhesive

208 properties could therefore promote the expansion of more virulent parasites.

209 Characterizing this process and determining the function of *pir* genes *in vivo* may identify

210 mechanisms that universally regulate immune evasion by *Plasmodium* and lead to novel

211 approaches to reduce their survival and transmission.

Methods

Mouse and parasite lines used

All experiments were performed in accordance with the UK Animals (Scientific Procedures) Act 1986 (PPL 80/2538) and the guidelines provided by the National Institutes of Health (NIH) Animal Care and Use Committee (ACUC) and approved by the Francis Crick Institute Ethical Committee and by the Institutional Animal Care and Use Committee (IACUC). T cell receptor-α chain knockout (TCRα-/-)[31], mice homozygous for a targeted mutation of the transmembrane exon of the IgM μ chain, μMT[32], and V(D)J recombination activation gene RAG-1 knockout (RAG-1-/-) [33] on a C57Bl/6 background, and C57 Bl/6 Wildtype (WT) mice were obtained from the specific-pathogen free (SPF) unit and subsequently conventionally housed with autoclaved cages, bedding and food at the BRF, of the Francis Crick Institute. The *Grammomys surdaster* rats used in this study were captured in the wild at Fungurume and Lumata in the Katanga province, south of DR Congo, were bred and maintained in a conventional unit in sterile cages, with autoclaved water and bedding at NIH/NIAID/LMIV [34]. Inbred Brown Norway rats were obtained from Charles River Laboratories and housed at the Animal Holding Unit of the Wellcome Trust Sanger Institute. Experiments were performed with 6-8 week old female mice and *Grammomys surdaster* rats housed under reverse light conditions (light 19:00-07:00, dark 07:00-19:00) and 6-week old female Brown Norway rats housed under normal light conditions, at 20-22°C. Measurements of clinical pathology were taken at 11:00. Core body temperature was measured with a rectal thermometer and erythrocyte density was determined with a VetScan HMII haematology system (Abaxis). Changes in body

11

236   temperature, body weight and erythrocyte density were calculated relative to a baseline

237   measurement performed 1 day before infection. An infection was considered virulent

238   when it induced a statistically significant decrease in the mouse temperature, weight

239   and/or red blood cells compared with uninfected control mice.

240   Cloned lines of *Plasmodium chabaudi chabaudi* AS and CB parasites were used. To

241   initiate infection, mice were either injected intraperitoneally (IP) or intravenously (IV)

242   with 1, 10^4 or 10^5 infected erythrocytes or submitted to the bite of 20 infected mosquitoes

243   as previously described[12]. *Grammomys surdaster* rats were infected by I.V. injection of

244   2500 sporozoites of *Plasmodium chabaudi chabaudi* AS. Brown Norway rats were

245   infected by I.V. injection of 7200 sporozoites of *Plasmodium berghei* ANKA.

246   Parasitaemia was monitored by Giemsa-stained thin blood films on blinded samples. The

247   limit of detection for patent parasitaemia was 0.001% infected erythrocytes.

248

249   **Chronic infections in wild type and genetically altered rodents**

250   Out of 10 wild-type C57Bl/6 mice, 10 μMT mice and 10 TCRα-/- mice infected with *P.*

251   *chabaudi* AS via mosquito bite, 9 WT C57Bl/6 mice, 5 μMT mice and 3 TCRα-/- mice

252   were randomly selected for transcriptomic analysis. 3 C57Bl/6 mice infected with *P.*

253   *chabaudi* CB, 3 *Grammomys surdaster* rats infected with *P. chabaudi* AS and 3 Brown

254   Norway rats infected with *P. berghei* ANKA (all via mosquito bites or injection of

255   sporozoites) were also used for transcriptomic analysis. Our aim was to use a minimum

256   of three biological replicates per condition for all experiments, as this is considered to be

257   sufficient for general surveys of differential expression[35]. Blood was extracted from the

258   same rodents at two time points, therefore analysing the changes occurring over time in

259 the same parasite populations. During the acute phase of infection, 100µL of blood was

260 collected from the tail after the completion of seven cycles of schizogony. During the

261 chronic phase, exsanguination was performed between 30 and 40 days post-infection in

262 µMT mice, at 30 days post-infection in TCRα-/- mice, at day 15 post-infection in Brown

263 Norway rats or when the parasitaemia reached 0.1% between 27 and 40 days post-

264 infection in wild type C57Bl/6 mice and *G. surdaster* rats. After data collection, the acute

265 sample from wild-type mouse 454 and both acute and chronic samples from µMT mouse

266 433 were excluded, as they were outliers in a multi-dimensional scaling analysis of the

267 gene expression data.

268

269 **Phenotypic analysis of acute and chronic parasites**

270 Donor mice were infected via mosquito bites. Parasites from donor mice (*n*= 3) were

271 blood-passaged into recipient mice at 1, 3, 4 and 6 weeks post-mosquito transmission.

272 Recipient mice (n=5 per donor mouse, the minimum number required for an 80% chance

273 of significant differences (p<=0.05) in parasitemia, weight and erythrocyte counts) were

274 infected by IP injection of 10^5 infected erythrocytes. Parasitaemia and signs of pathology

275 were monitored in recipient mice on blinded samples.

276

277 **Cloning of single parasites from acute and chronic phases**

278 One mouse was infected via mosquito bite. Infected blood was then collected from the

279 tail at two time points: 1 week post-infection (acute phase) and six weeks post-infection

280 (chronic phase). Both times, parasites were cloned by dilution and single parasites were

281 injected IV in wild type mice. After expansion, each cloned line was blood passaged into

282 eight experimental mice by IP injection of 10^5 infected erythrocytes. Five mice were used

283 to monitor parasitaemia and signs of pathology and three mice were sacrificed and

284 exsanguinated after the completion of seven cycles of schizogony for RNA extraction.

285

286 **RNA extraction**

287 During all RNA extractions experiments, blood-stage parasites were isolated at 11:00,

288 when >90% of the parasites were at the late trophozoite stage of development (Table S8).

289 Parasite RNA was isolated as previously described[12]. Briefly, whole blood was depleted

290 of leukocytes by filtration (Plasmodipur, EuroProxima) and from globin RNA and red

291 cell debris by saponin lysis (Sigma) and centrifugation. Purified parasites were

292 resuspended in TRIzol, snap-frozen on dry ice and kept at -80°C until use. RNA was then

293 extracted, resuspended in water and its quantity and quality were determined on an

294 Agilent 2100 Bionalayzer RNA 6000 Nano or Pico chip.

295

296 **RNA-seq library prep and sequencing**

297 The majority of RNA was used to make 200-450bp fragment libraries using Illumina's

298 TruSeq RNA Sample Prep v2 kit, with 10 cycles of PCR amplification using KAPA Hifi

299 Polymerase rather than the kit-supplied Illumina PCR Polymerase. Two samples required

300 14 cycles of PCR (T1_46-4 and T1_47-4). Mosquito-transmitted *P. chabaudi* CB

301 samples (MTCB) were subjected to 15 cycles of PCR.

302 The libraries were sequenced using an Illumina HiSeq 2000 v3 with 100bp paired-end

303 reads (samples beginning 'T'), an Illumina HiSeq 4000 with 75bp paired-end reads

304 (MTCB samples) or an Illumina HiSeq 2500 with 75bp paired-end reads for the rest.

305   RNA from cloned parasites was used to make 100-300bp fragment libraries produced

306   using Illumina's TruSeq Stranded mRNA Sample Prep Kit with 10 cycles of PCR

307   amplification using KAPA Hifi Polymerase. These libraries were sequenced using an

308   Illumina HiSeq 2500 with 75bp paired-end reads. A description of the libraries can be

309   found in Supplementary Table 10.

310

311

312   **Analysis of gene expression**

313   Tag sequences were removed from sequenced reads. Where individual samples were run

314   over multiple lanes, the reads from these were merged before mapping. Reads were then

315   mapped against spliced gene sequences (exons, but not UTRs) from the v3 *P. c. chabaudi*

316   AS (this work) or the v3 *P. berghei* ANKA (Fougère et al., under review) reference

317   genomes using Bowtie2 v2.1.0[36] (-a -X 800 –x). Read counts per transcript were

318   estimated using eXpress v1.3.0[37], with default parameters. Genes with an effective

319   length cutoff below 10 in any sample were removed. Summing over transcripts generated

320   read counts per gene.

321   Differential expression analysis was performed using edgeR v3.8.6[38] on genes with ≥ 3

322   counts per million. The Fisher's exact test was used with cutoffs FDR < 0.01 and fold

323   change ≥ 2, except for the *P.berghei* analysis where FDR < 0.1 and fold change ≥ 1 was

324   used. Functional categories of genes were identified by orthology using GeneDB [39] from

325   several different *P. falciparum* datasets: invasion genes[40], sexual genes[41] and

326   subtelomeric (by manual inspection of chromosomes).

327   Gene expression clustering was used to identify which genes within ChAPLs and AAPLs

328   were co-expressed. This was done by generating a toroidal 5x5 Self-Organising Map

329   (SOM) using the Kohonen package in R[42]. FPKM values for each gene were logged (base

330   2) and mean-normalised per gene. The SOM was trained in 100 iterations.

331

332   **Sequencing and assembly of *P. c. chabaudi* AS v3 genome**

333   High molecular weight DNA was prepared as follows. Heparinised blood from two mice

334   infected with *P. chabaudi AS* (blood passaged parasites, days seven & eight post-

335   infection) was pooled, diluted 1:3 in PBS, and passed through a Plasmodipur (Euro-

336   Diagnostica) filter. Erythrocyte membranes were removed by saponin lysis (0.15% in ice

337   cold PBS). Parasitised erythrocytes were recovered by centrifugation (2000g 10 min

338   4 °C) and washed twice in ice cold PBS. The cell pellet was taken up in 50mM Tris HCl

339   pH 7.5, 50mM EDTA pH8.0, 100mM NaCl, 0.5% SDS and digested with RNase A

340   (1mg/ml, Life Technologies) for 30 min at 37°C. Proteinase K (Roche) was added to a

341   final concentration of 1mg/ml, incubated at 45°C overnight, followed by extraction with

342   phenol chloroform and ethanol precipitation.

343   For preparation of long-read sequencing libraries, 5μg of *P. c. chabaudi* AS genomic

344   DNA was sheared to 20-25kb by passing through a 25mm blunt ended needle. SMRT

345   bell template libraries were generated using the Pacific Biosciences issued protocol (20

346   kb Template Preparation Using BluePippin™ Size-Selection System). After 7kb-20kb

347   size-selection using the BluePippin™ Size-Selection System (Sage Science,

348   Beverly, MA) the library was sequenced using P6 polymerase and chemistry version 4

349   (P6C4) on 5 single-molecule real-time (SMRT) cells, each with a 240 minutes movie

350 length. The five SMRT cells were processed using a Pacific Biosciences RSII. Reads

351 were filtered using SMRT portal v2.2 with default parameters (minimum subread length

352 50, minimum polymerase read quality 75, minimum polymerase read length 50). The

353 yield was 258214 reads totaling 1.51Gb, with read N50 9003bp and read quality 0.833.

354 The reads were assembled using HGAP v2.2.0[43] with expected coverage 75 and other

355 parameters as default.

356 The 28 unitigs from the HGAP assembly were BLASTed against the *P. chabaudi* AS v2

357 assembly[14] to identify sequences relating to nuclear, mitochondrial and apicoplast

358 genomes. Four unitigs were found to be mouse contamination. Fourteen unitigs

359 represented the fourteen chromosomes from the v2 assembly including most of the

360 unplaced contigs from v2 and some additional new sequence. All fourteen had telomeric

361 repeats at both ends. The only major rearrangement between the assemblies was between

362 the ends of chromosomes 6 and 13. We mapped Quiver-corrected PacBio reads against

363 v2 and v3 assemblies and confirmed that the v3 assembly was correct. However, we

364 noticed base errors and in particular many indels in the v3 assembly due to the relatively

365 high error rate of PacBio sequencing. The v2 assembly had had the benefit of Sanger and

366 Illumina technologies such that at the base level it was more accurate than the PacBio

367 assembly. Therefore we generated 2x250bp paired-end pseudoreads from the v2

368 assembly with a fragment size of 600bp to a depth of 50x and used these to correct the v3

369 assembly using iCORN2[44]. We transferred gene models from the v2 assembly to the

370 corrected v3 assembly using RATT[44]. We then trained Augustus v2.5.5[45] using the v2

371 gene models with default parameters and predicted a new set of models for the corrected

372 v3 assembly. We kept only those, which did not overlap with the RATT-transferred gene

373   models. We then manually assessed the transferred models, annotated the new Augustus

374   models and renamed the locus tags. In general we added a zero to the end of the v2 locus

375   tags, but where a gene did not exist in v2 or had moved we added an odd number at the

376   end. We maintained numerical ordering of locus tags across each chromosome so that

377   they reflect relative chromosomal location.

378

379   **Sequence analysis of *pir* genes**

380   *Pir* gene sub clade assignments were initially made using Hidden Markov Models

381   (HMMs) built from the rodent malaria *pir* subclades published in Otto et al.[14]. There were

382   some discrepancies between the two sets, e.g. where the clade in the published tree did

383   not agree with the top HMM hit. In particular we identified three members of the S3

384   subclade, which was not previously thought to be present in *P. chabaudi*. We used

385   hmmsearch from HMMer i1.1rc3[46] with an E-value cutoff of 1e-10 and took the best hit

386   to a subclade. We thus identified 207 *pir* genes which we could assign to a rodent malaria

387   parasite (RMP) *pir* subclade, while a further five genes could not be classified into

388   existing subfamilies (PCHAS_0401400, PCHAS_0602000 and PCHAS_1247000).

389   HMM assignments of L2 *pirs* in *P. berghei* were consistent with the original tree-based

390   classification presented in Otto et al. (2014).

391   We used FIRE v1.1a[47] to identify putative promoter sequences upstream of genes in the

392   AAPLs and ChAPLs. Firstly we used RNA-seq data from the study to determine 5'

393   unstranslated regions (UTRs) of *pir* genes and then manually identified untranscribed

394   regions upstream of each UTR, reaching until the proximal UTR of the next upstream

395   gene. Where there was insufficient RNA-seq data or other ambiguity, we did not identify

396   a region. This was successful for 27/32 ChAPL genes and 22/28 AAPL genes. We then

397   defined AAPL and ChAPL genes as two distinct expression clusters and ran FIRE with --

398   exptype=discrete and --nodups=1 options.

399   Sequence alignments were performed using Muscle[48] with default parameters and

400   visualized using Jalview[49].

401

402   **Regression analysis of *pir* genes against parasitological parameters**

403   DESeq2 was used to determine whether variation in gene expression between parasite

404   populations in different mice was correlated with parasitological parameters such as

405   parasitaemia and time of recrudescence[50]. In particular we found that the expression level

406   during the acute phase of 152 genes was correlated with the time to recrudescence across

407   nine mice (p-value < 0.01; fold change >= 2). We discretized the time in days at

408   recrudescence as follows: mice 214 (32 days), 421 (31 days), 427 (30 days) were short,

409   2014 (34 days), 2114 (36 days), 452 (33 days) and 456 (34 days) were medium, 211 (43

410   days) and 212 (39 days) were long. We then used this single discrete factor in the

411   DESeq2 design formula. Those genes where expression level during the acute phase was

412   negatively correlated with time at recrudescence were, with a single exception, *pir* genes

413   from 3L, 6L and 6R subtelomeres (Table S6). These are the three ChAPLs and another

414   L1-rich locus (4R). Interestingly, the strongest effect was from an exported gene of

415   unknown function (PCHAS_1201300). The strongest positive effects were from *pir*

416   genes in AAPLs.

417

418   **Mixed infection**

419   Transgenic lines of *Plasmodium chabaudi* AS expressing mCherry or mNeonGreen under

420   the control of the constitutive promoter EF1α were generated by transfection with the

421   plasmid *Pc*EFp230p_mCherry and *Pc*EFp230p_mNeonGreen, respectively. These

422   plasmids targeted the silent *230p* genomic locus, and were generated by replacement of

423   the ssu rRNA locus of p*Pc*-mCH$_{CAM,}$ described in[51], with the *P. chabaudi* p230p locus

424   (*PCHAS_03_v3* 276,451-283,284). For *Pc*EFp230p_mNeonGreen, the mCherry tag was

425   replaced with mNeonGreen[52]. Parasites were cycled ON and OFF acid water containing

426   pyrimethamine to select for stably transfected parasites, and cloned by serial dilution. The

427   insertion was verified by PCR amplification and Southern blot analysis and fluorochrome

428   expression by Flow cytometry analysis (Figure S9). mNeonGreen-expressing parasites

429   were then serially blood passaged (SBP) 10 times while mCherry-expressing parasites

430   were mosquito transmitted (MT) into donor mice. Both parasites were then mixed at

431   different ratios and 10^5 infected erythrocytes were injected i.p. into ten recipient mice per

432   group. Competition between parasite lines as well as parasitaemia and pathology were

433   then monitored in recipient mice. To determine the ratio of mNeonGreen- and mCherry-

434   expressing parasites during infection, tail blood was collected and stained with Hoechst

435   33342 (New England Biolab). Samples were acquired on a BD LSRII flow cytometer and

436   data were analysed with FlowJo software (TreeStar).

437

438   **Clodronate liposomes treatment**

439   B6.RAG-1-/- mice were treated with 200uL i.v. of either saline (Sigma; negative control)

440   or clodronate liposomes (clodronateliposomes.org). Flow cytometric analysis was

441   carried out to assess the depletion of phagocytic cells. Single-cell suspensions of

442 splenocytes were prepared, , and cells enumerated 24h after *in vivo* treatment. After

443 staining with Zombie Aqua (Biolegend) for live/dead discrimination, cells were stained

444 with monoclonal antibodies using appropriate combinations of fluorochromes (CD11c-

445 BV786, CD169-BV605, F4/80-BV421, Ly6G-PerCPCy5.5, MHCII-FITC, Ter119-PE-

446 Cy7, CD68-PE, Ly6c-APC-Cy7, CD11b-AF647, all from Biolegend). The samples were

447 acquired on a BD Fortessa/X20 (BD Biosciences) using Diva acquisition software (BD

448 Biosciences). FlowJo (Tree Star) software was used to analyze the data.

449 To determine the role of T-, B- and phagocytic cells in the selection of virulent parasites,

450 B6.RAG-1-/- mice were treated with saline or clodronate liposomes as described above at

451 days -1 and 4 after i.p. infection with 10^5 iRBCs containing a mix of Neon-Green-

452 expressing SBP parasites and mCherry-expressing MT parasites (see "Mixed infections"

453 section).

454

455 **Data availability**

456 RNA-seq datasets used in this study have been submitted to the European Nucleotide

457 Archive (ENA) under study accession ERP017479. These datasets are described further

458 in Supplementary Table 10. The Pacific Biosciences RSII genomic sequencing reads,

459 used to generate the *Plasmodium chabaudi chabaudi* AS v3 genome sequence, have been

460 submitted to the ENA under experiment accession ERX613966. The assembled genome

461 sequence and annotation can be accessed from GeneDB

462 (http://www.genedb.org/Homepage/Pchabaudi) and PlasmoDB

463 (http://plasmodb.org/plasmo/). The authors declare that no competing interests exist.

464

References

1   Ashley, E. A. & White, N. J. The duration of Plasmodium falciparum infections. *Malar J* **13**, 500, doi:10.1186/1475-2875-13-500 (2014).

2   Assennato, S. M. *et al.* Plasmodium genome in blood donors at risk for malaria after several years of residence in Italy. *Transfusion* **54**, 2419-2424, doi:10.1111/trf.12650 (2014).

3   Brown, K. N. & Brown, I. N. Immunity to malaria: antigenic variation in chronic infections of Plasmodium knowlesi. *Nature* **208**, 1286-1288 (1965).

4   Bruce, M. C. *et al.* Genetic diversity and dynamics of plasmodium falciparum and P. vivax populations in multiply infected children with asymptomatic malaria infections in Papua New Guinea. *Parasitology* **121 (Pt 3)**, 257-272 (2000).

5   Vinetz, J. M., Li, J., McCutchan, T. F. & Kaslow, D. C. Plasmodium malariae infection in an asymptomatic 74-year-old Greek woman with splenomegaly. *The New England journal of medicine* **338**, 367-371, doi:10.1056/NEJM199802053380605 (1998).

6   Bousema, T., Okell, L., Felger, I. & Drakeley, C. Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nature reviews. Microbiology* **12**, 833-840, doi:10.1038/nrmicro3364 (2014).

7   Scherf, A., Lopez-Rubio, J. J. & Riviere, L. Antigenic variation in Plasmodium falciparum. *Annu Rev Microbiol* **62**, 445-470, doi:10.1146/annurev.micro.61.080706.093134 (2008).

8   Kyes, S., Horrocks, P. & Newbold, C. Antigenic variation at the infected red cell surface in malaria. *Annual review of microbiology* **55**, 673-707, doi:10.1146/annurev.micro.55.1.673 (2001).

9   Warimwe, G. M. *et al.* Plasmodium falciparum var gene expression is modified by host immunity. *Proc Natl Acad Sci U S A* **106**, 21801-21806, doi:10.1073/pnas.0907590106 (2009).

10   Janssen, C. S., Phillips, R. S., Turner, C. M. & Barrett, M. P. Plasmodium interspersed repeats: the major multigene superfamily of malaria parasites. *Nucleic Acids Res* **32**, 5712-5720, doi:10.1093/nar/gkh907 (2004).

11   Cunningham, D., Lawton, J., Jarra, W., Preiser, P. & Langhorne, J. The pir multigene family of Plasmodium: antigenic variation and beyond. *Mol Biochem Parasitol* **170**, 65-73, doi:10.1016/j.molbiopara.2009.12.010 (2010).

12   Spence, P. J. *et al.* Vector transmission regulates immune control of Plasmodium virulence. *Nature* **498**, 228-231, doi:10.1038/nature12231 (2013).

13   Ndungu, F. M. *et al.* Functional memory B cells and long-lived plasma cells are generated after a single Plasmodium chabaudi infection in mice. *PLoS Pathog* **5**, e1000690, doi:10.1371/journal.ppat.1000690 (2009).

14   Otto, T. D. *et al.* A comprehensive evaluation of rodent malaria parasite genomes and gene expression. *BMC Biol* **12**, 86, doi:10.1186/s12915-014-0086-0 (2014).

15   Rovira-Graells, N. *et al.* Transcriptional variation in the malaria parasite Plasmodium falciparum. *Genome Res* **22**, 925-938, doi:10.1101/gr.129692.111 (2012).

510   16   de Roode, J. C. *et al.* Virulence and competitive ability in genetically diverse
511   malaria infections. *Proc Natl Acad Sci U S A* **102**, 7624-7628,
512   doi:10.1073/pnas.0500078102 (2005).

513   17   Bell, A. S., de Roode, J. C., Sim, D. & Read, A. F. Within-host competition in
514   genetically diverse malaria infections: parasite virulence and competitive success.
515   *Evolution* **60**, 1358-1371 (2006).

516   18   De Roode, J. C., Read, A. F., Chan, B. H. & Mackinnon, M. J. Rodent malaria
517   parasites suffer from the presence of conspecific clones in three-clone
518   Plasmodium chabaudi infections. *Parasitology* **127**, 411-418 (2003).

519   19   Schneider, P. *et al.* Virulence, drug sensitivity and transmission success in the
520   rodent malaria, Plasmodium chabaudi. *Proc Biol Sci* **279**, 4677-4685,
521   doi:10.1098/rspb.2012.1792 (2012).

522   20   Wargo, A. R., de Roode, J. C., Huijben, S., Drew, D. R. & Read, A. F.
523   Transmission stage investment of malaria parasites in response to in-host
524   competition. *Proceedings. Biological sciences / The Royal Society* **274**, 2629-
525   2638, doi:10.1098/rspb.2007.0873 (2007).

526   21   Goel, S. *et al.* RIFINs are adhesins implicated in severe Plasmodium falciparum
527   malaria. *Nat Med* **21**, 314-317, doi:10.1038/nm.3812 (2015).

528   22   Moll, K., Palmkvist, M., Ch'ng, J., Kiwuwa, M. S. & Wahlgren, M. Evasion of
529   Immunity to Plasmodium falciparum: Rosettes of Blood Group A Impair
530   Recognition of PfEMP1. *PLoS One* **10**, e0145120,
531   doi:10.1371/journal.pone.0145120 (2015).

532   23   Huang, X. *et al.* Differential Spleen Remodeling Associated with Different Levels
533   of Parasite Virulence Controls Disease Outcome in Malaria Parasite Infections.
534   *mSphere* **1**, doi:10.1128/mSphere.00018-15 (2016).

535   24   Gilks, C. F., Walliker, D. & Newbold, C. I. Relationships between sequestration,
536   antigenic variation and chronic parasitism in Plasmodium chabaudi chabaudi--a
537   rodent malaria model. *Parasite Immunol* **12**, 45-64 (1990).

538   25   Stevenson, M. M. & Riley, E. M. Innate immunity to malaria. *Nat Rev Immunol* **4**,
539   169-180, doi:10.1038/nri1311 (2004).

540   26   Carvalho, B. O. *et al.* On the cytoadhesion of Plasmodium vivax-infected
541   erythrocytes. *J Infect Dis* **202**, 638-647, doi:10.1086/654815 (2010).

542   27   Claessens, A. *et al.* A subset of group A-like var genes encodes the malaria
543   parasite ligands for binding to human brain endothelial cells. *Proc Natl Acad Sci*
544   *U S A* **109**, E1772-1781, doi:10.1073/pnas.1120461109 (2012).

545   28   Niang, M. *et al.* STEVOR is a Plasmodium falciparum erythrocyte binding
546   protein that mediates merozoite invasion and rosetting. *Cell Host Microbe* **16**, 81-
547   93, doi:10.1016/j.chom.2014.06.004 (2014).

548   29   Yam, X. Y. *et al.* Characterization of the Plasmodium Interspersed Repeats (PIR)
549   proteins of Plasmodium chabaudi indicates functional diversity. *Sci Rep* **6**, 23449,
550   doi:10.1038/srep23449 (2016).

551   30   Smith, J. D., Rowe, J. A., Higgins, M. K. & Lavstsen, T. Malaria's deadly grip:
552   cytoadhesion of Plasmodium falciparum-infected erythrocytes. *Cell Microbiol* **15**,
553   1976-1983, doi:10.1111/cmi.12183 (2013).

554   31   Philpott, K. L. *et al.* Lymphoid development in mice congenitally lacking T cell
555   receptor alpha beta-expressing cells. *Science* **256**, 1448-1452 (1992).

556 32   Kitamura, D., Roes, J., Kuhn, R. & Rajewsky, K. A B cell-deficient mouse by
557   targeted disruption of the membrane exon of the immunoglobulin mu chain gene.
558   *Nature* **350**, 423-426, doi:10.1038/350423a0 (1991).
559 33   Mombaerts, P. *et al.* RAG-1-deficient mice have no mature B and T lymphocytes.
560   *Cell* **68**, 869-877 (1992).
561 34   Conteh, S. *et al.* Grammomys surdaster, the natural host for Plasmodium berghei
562   parasites, as a model to study whole organism vaccines against malaria. .
563   *American Journal of Tropical Medicine and Hygiene* (in press).
564 35   Conesa, A. *et al.* A survey of best practices for RNA-seq data analysis. *Genome*
565   *Biol* **17**, 13, doi:10.1186/s13059-016-0881-8 (2016).
566 36   Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*
567   *Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
568 37   Roberts, A. & Pachter, L. Streaming fragment assignment for real-time analysis
569   of sequencing experiments. *Nat Methods* **10**, 71-73, doi:10.1038/nmeth.2251
570   (2013).
571 38   Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor
572   package for differential expression analysis of digital gene expression data.
573   *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).
574 39   Logan-Klumpler, F. J. *et al.* GeneDB--an annotation database for pathogens.
575   *Nucleic Acids Res* **40**, D98-108, doi:10.1093/nar/gkr1032 (2012).
576 40   Hu, G. *et al.* Transcriptional profiling of growth perturbations of the human
577   malaria parasite Plasmodium falciparum. *Nat Biotechnol* **28**, 91-98,
578   doi:10.1038/nbt.1597 (2010).
579 41   Young, J. A. *et al.* The Plasmodium falciparum sexual development
580   transcriptome: a microarray analysis using ontology-based pattern identification.
581   *Mol Biochem Parasitol* **143**, 67-79, doi:10.1016/j.molbiopara.2005.05.007 (2005).
582 42   Wehrens, R. & Buydens, L. M. C. Self- and super-organizing maps in R: The
583   kohonen package. *J Stat Softw* **21**, 1-19 (2007).
584 43   Chin, C. S. *et al.* Nonhybrid, finished microbial genome assemblies from long-
585   read SMRT sequencing data. *Nature Methods* **10**, 563-+,
586   doi:10.1038/Nmeth.2474 (2013).
587 44   Swain, M. T. *et al.* A post-assembly genome-improvement toolkit (PAGIT) to
588   obtain annotated genomes from contigs. *Nature Protocols* **7**, 1260-1284,
589   doi:10.1038/nprot.2012.068 (2012).
590 45   Stanke, M., Schoffmann, O., Morgenstern, B. & Waack, S. Gene prediction in
591   eukaryotes with a generalized hidden Markov model that uses hints from external
592   sources. *Bmc Bioinformatics* **7**, doi:Artn 62
593 10.1186/1471-2105-7-62 (2006).
594 46   Eddy, S. R. Accelerated Profile HMM Searches. *Plos Comput Biol* **7**, doi:ARTN
595   e1002195
596 10.1371/journal.pcbi.1002195 (2011).
597 47   Elemento, O., Slonim, N. & Tavazoie, S. A universal framework for regulatory
598   element discovery across all genomes and data types. *Molecular cell* **28**, 337-350,
599   doi:10.1016/j.molcel.2007.09.027 (2007).
600 48   Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high
601   throughput. *Nucleic Acids Res* **32**, 1792-1797, doi:10.1093/nar/gkh340

602   32/5/1792 [pii] (2004).
603   49   Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J.
604   Jalview Version 2--a multiple sequence alignment editor and analysis workbench.
605   *Bioinformatics* **25**, 1189-1191, doi:10.1093/bioinformatics/btp033 (2009).
606   50   Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
607   dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, doi:ARTN 550
608   10.1186/s13059-014-0550-8 (2014).
609   51   Spence, P. J. *et al.* Transformation of the rodent malaria parasite Plasmodium
610   chabaudi. *Nat Protoc* **6**, 553-561, doi:10.1038/nprot.2011.313 (2011).
611   52   Shaner, N. C. *et al.* A bright monomeric green fluorescent protein derived from
612   Branchiostoma lanceolatum. *Nature Methods* **10**, 407-+, doi:10.1038/Nmeth.2413
613   (2013).
614

Acknowledgements

Author contributions

630 T.B., A.J.R., C.N., M.B. and J.L. (last author) designed the study. T.B. designed and

631 performed the mouse experiments with the help of I.T., G.K., S.M. and P.S.. A.J.R.

632 designed the sequencing experiments and performed the bioinformatic analyses. M.S.

633 coordinated sequencing experiments. U.B. manually annotated the *Plasmodium chabaudi*

634 AS genome sequence. D.C. and J.L. (7th author) created transgenic parasites. P.E.D. and

635 S.C performed thicket rat experiments, T.M, E.B. and O.B carried out the experiments in

636 the Brown Norway rats. T.B., A.J.R., C.N., M.B. and J.L. (last author) wrote the

637 manuscript.

638

639 Competing financial Interests

640 The authors declare no competing financial interests.
641
642
643 Corresponding author
644
645 Correspondence to Jean Langhorne, Thibaut Brugat and Adam Reid. Requests for

646 parasites, mice and reagents should go to Jean Langhorne. Requests for

647 sequencing/bioinformatic data should go to Adam Reid.

648

Figure legends

650   **Figure 1: Chronic infections modify the transcriptome of *P. chabaudi* independently**

651   **of the adaptive immune response.**

652   Parasitaemias over the course of infection in (A) 10 wild-type C57Bl/6 mice, (B) 10

653   B6.μMT-/- mice and (C) 10 TCRα-/- mice are shown on the left panels. Parasite mRNAs

654   were collected from 9 wild-type C57Bl/6 mice, 5 B6.μMT-/- mice and 3 TCRα-/- mice

655   selected randomly at the time points indicated by the red arrows. On the right panels, hot

656   pie diagrams show expression levels (black and white inner circles) and fold changes

657   (coloured outer circles) for genes expressed more highly (FDR <= 0.01, fold change >=

658   2) in the chronic and acute phases. Genes are classified into groups according to several

659   categories including: subtelomeric genes (subtel.); genes associated with red blood cell

660   invasion (invasion); genes associated with gametocytogenesis (sexual) and *pir* genes

661   (highlighted in brown).

662

663   **Figure 2: Chronic infections are characterized by the expression of distinctive**

664   **clusters of *pir* genes in *P. chabaudi*.**

665   (A) Chronic Associated *Pir* Loci (ChAPL), and (B) Acute Associated *Pir* Loci (AAPL)

666   are shown in the context of the subtelomeric region from the telomeric repeats to the last

667   *pir* or *fam-a* gene. Those genes co-expressed across mice over the course of infection are

668   highlighted in orange. (C) Heatmap of *pir* gene expression at ChAPL and AAPL loci,

669   showing mean-normalised RPKM values for 10 wild type C57Bl/6 mice, four B6.μMT-/-

670   mice and three TCRα-/- mice infected by mosquito bites. Each column represents values

671   obtained from an individual mouse. Samples were collected from the same mice during

672 the acute and chronic (chro.) phases of infection. On the side, numbers indicate

673 chromosomes and letters indicate the location on that chromosome: L for left hand end, R

674 for right hand end. ChAPLs and AAPLs are highlighted. For each gene, red represents its

675 maximal expression, green its minimal expression. The absolute maximum and minimum

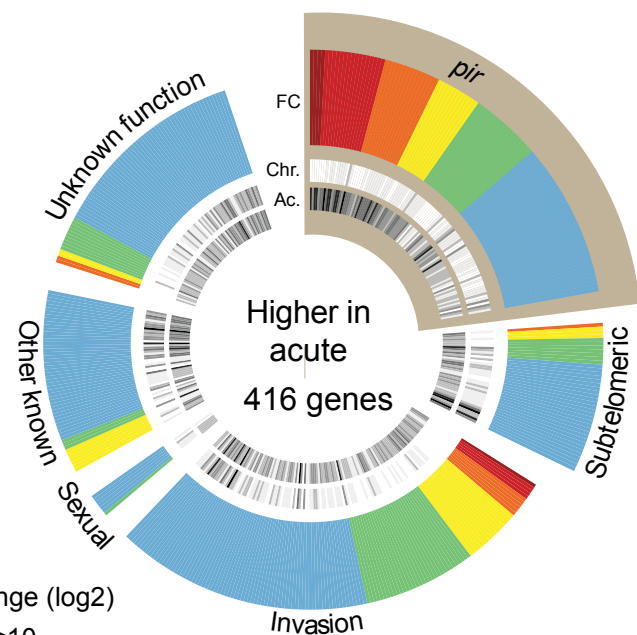676 values for each gene are shown in a separate heatmap.

677

678 **Figure 3: Chronic infections select for virulent *P. chabaudi* parasites expressing**

679 **distinctive clusters of *pir* genes.**

680 Donor mice were infected by mosquito bites. (A) Parasites were subsequently passaged

681 into naïve recipient mice after one, three, four or six weeks (n=15 mice per group). (B)

682 Individual parasites were cloned during the acute and chronic phases of infection. After

683 their expansion, clonal populations of parasites were used to infect recipient mice (n= 5

684 mice per clone). In (A) and (B), we show maximum parasitaemia, temperature change

685 and weight change for recipient mice compared to mice directly infected by mosquito

686 bites (MT; blue dotted line) or with serially blood passaged parasites (SBP; red dotted

687 line). Each dot represents the mean ± standard error of the mean (SEM) of 5 individual

688 mice. Each group has been compared to mice infected with 10^5 MT parasites (*$P<0,05$;

689 **$P<0,01$; ***$P<0,001$, two-sided Mann Whitney Test). (C) Heatmap of *pir* gene
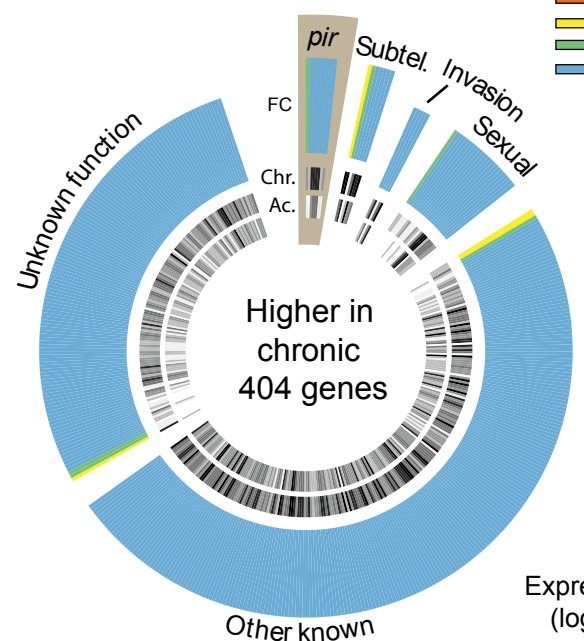
690 expression at ChAPLs and AAPLs in three acute and three chronic clones. The first three

691 columns represent cloned parasite lines derived from from the acute phase (f, a, b), the

692 next three columns represent cloned lines from the chronic phase (k, n, p) from 3

693 individual mice. On the side, chromosomes numbers are indicated as well as the location

694 on that chromosome: L for left hand, R for right hand. ChAPLs and AAPLs are

695   highlighted. For each gene, red represents its maximal expression, green its minimal

696   expression. The absolute maximum and minimum values for each gene are shown in a

697   separate heatmap.

698

699   **Figure 4: Chronicity and virulence of infection are dictated by the initial**

700   **composition of *P. chabaudi* population.**

701   Parasites tagged with mCherry (Fluo 1) were Mosquito-Transmitted (MT), and

702   mNeonGreen (Fluo 2) tagged parasites were Serially Blood-Passaged (SBP). Different

703   proportions of the two were used to infect recipient mice by i.p. injection of 10^5 parasites

704   (n= 10 mice per group). As a control, a group of 10 recipient mice was infected by i.p.

705   injection of 10^4 SBP parasites. The mean percentage of parasites being SBP (as

706   determined by flow cytometry) is indicated (upper panel). We show maximum

707   parasitaemia, temperature change, weight change (middle panel) and time of

708   recrudescence (lower panel) for each group compared to mice infected with 100% MT

709   parasites (blue dotted line) or 100% SBP parasites (red dotted line). Each dot represents

710   the mean ± standard error of the mean (SEM) for each group of 10 mice. Each group has

711   been compared to mice infected with 10^5 SBP parasites (*$P<0,05$; **$P<0,01$;

712   ***$P<0,001$, two-sided Mann Whitney Test).

A C57 Bl/6

B B6.µMT-/-
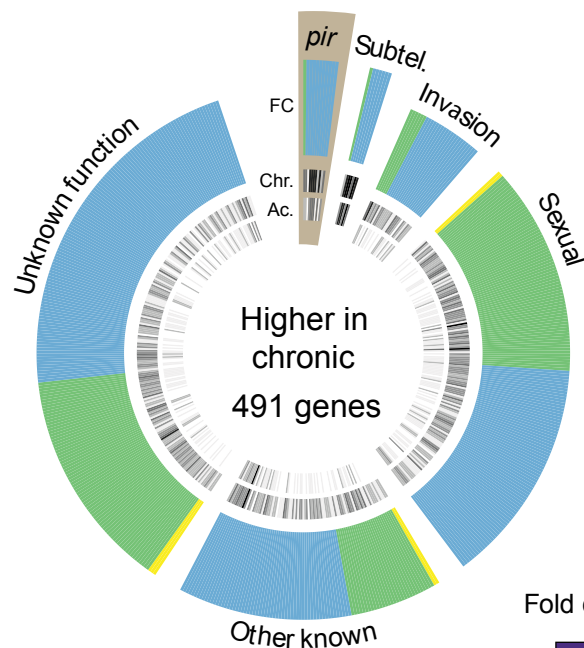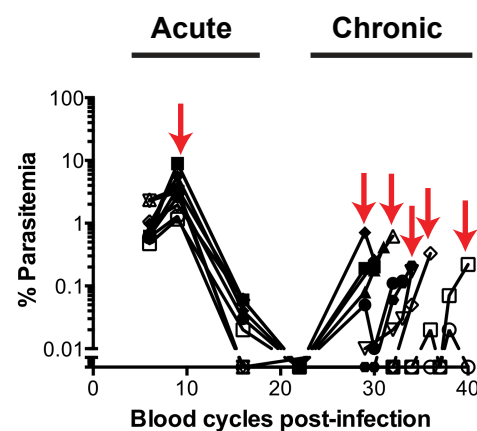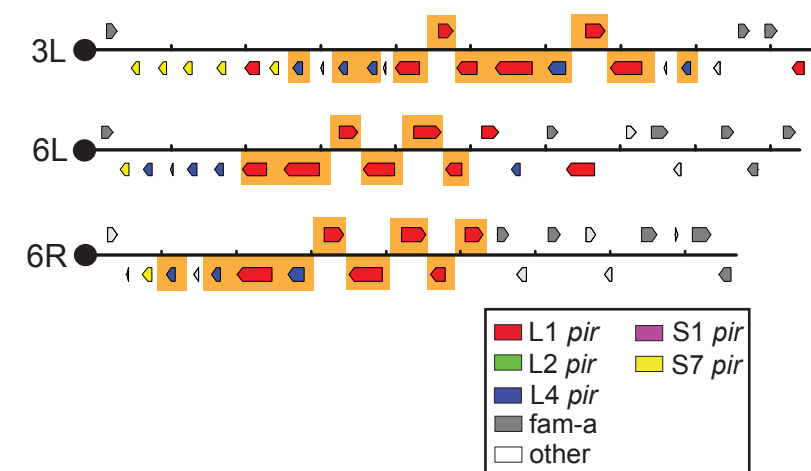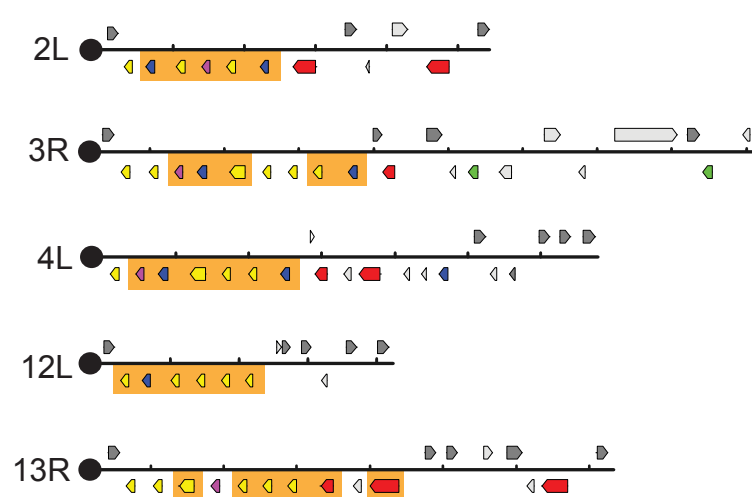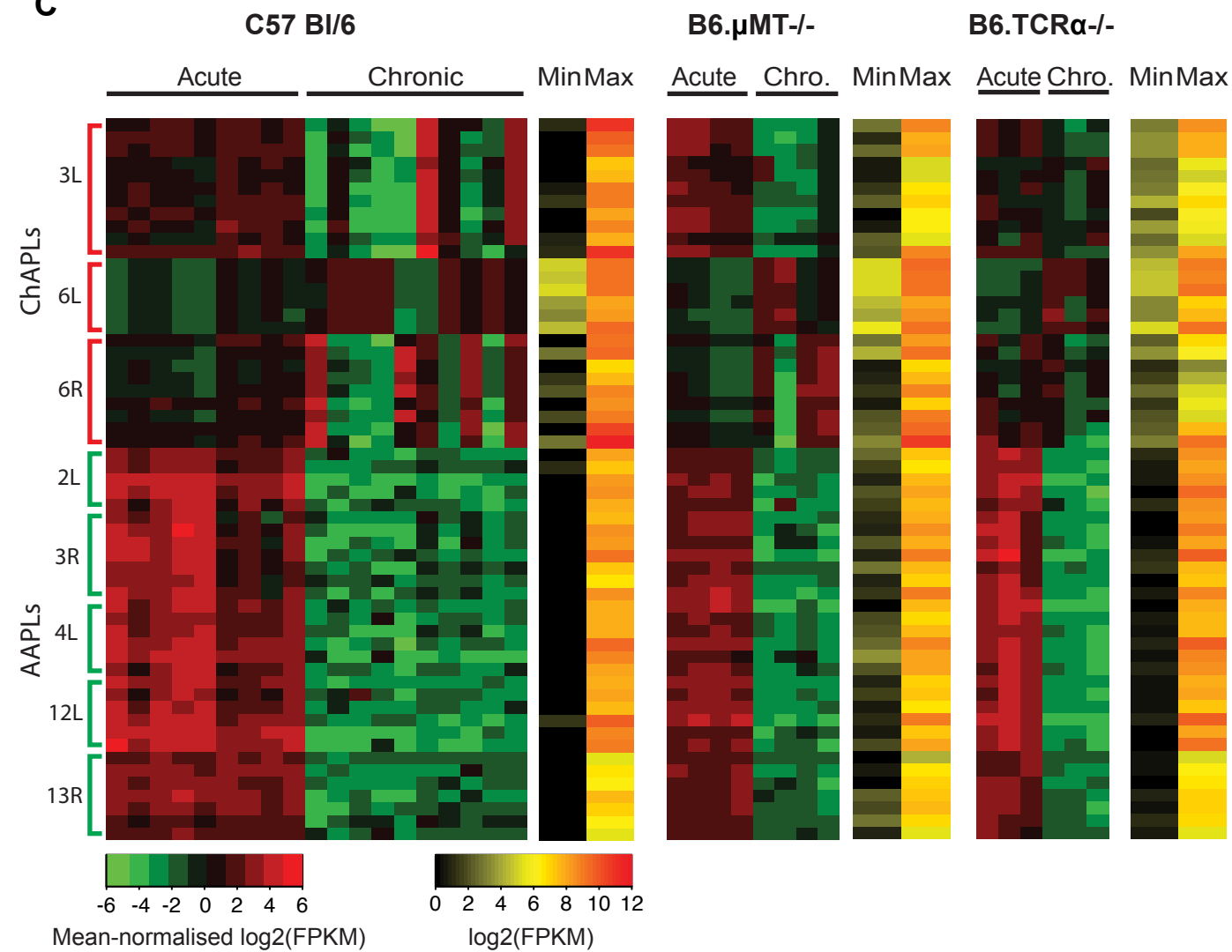
C B6.TCRα-/-

Fold change (log2)
- ≥10
- 8 - <10
- 6 - <8
- 5 - <6
- 4 - <5
- 3 - <4
- 1 - <3
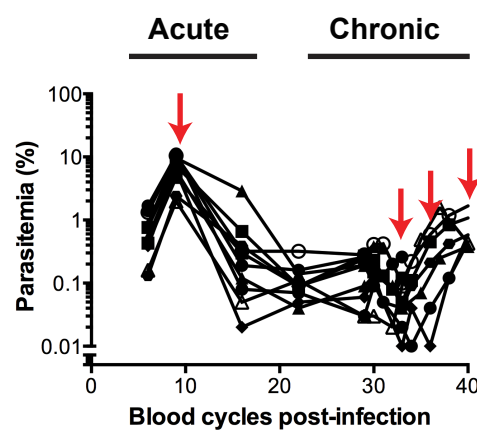
Expression level (log2 FPKM)
- ≥9
- 7.5 - <9
- 6 - <7.5
- 4.5 - <6
- 3 - <4.5
- 1.5 - <3
- 0 - <1.5

A ChAPLs

B AAPLs

Legend:
- L1 *pir* (red)
- L2 *pir* (green)
- L4 *pir* (blue)
- fam-a (grey)
- other (white)
- S1 *pir* (purple)
- S7 *pir* (yellow)

C

C57 Bl/6 — Acute | Chronic | MinMax

B6.µMT-/- — Acute | Chro. | MinMax

B6.TCRα-/- — Acute | Chro. | MinMax

ChAPLs: 3L, 6L, 6R

AAPLs: 2L, 3R, 4L, 12L, 13R

Mean-normalised log2(FPKM): -6 -4 -2 0 2 4 6

log2(FPKM): 0 2 4 6 8 10 12
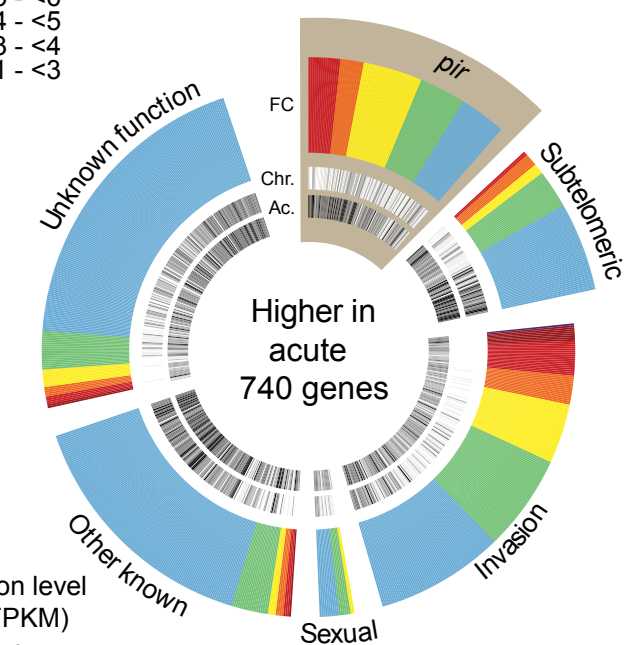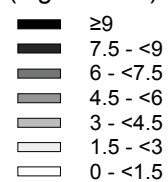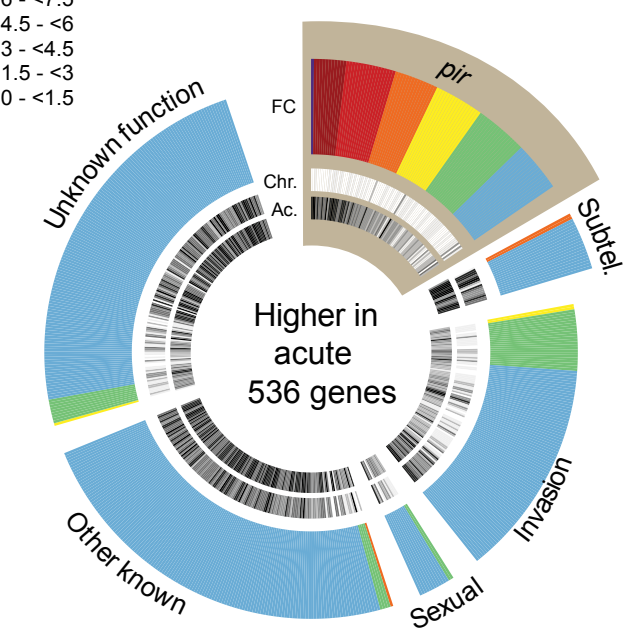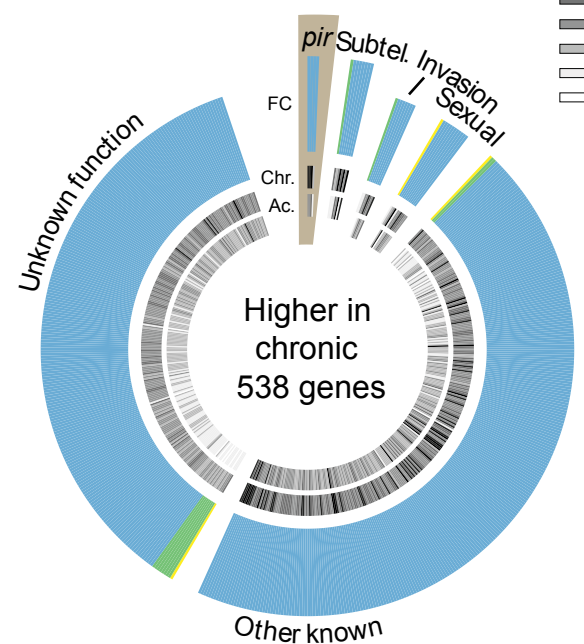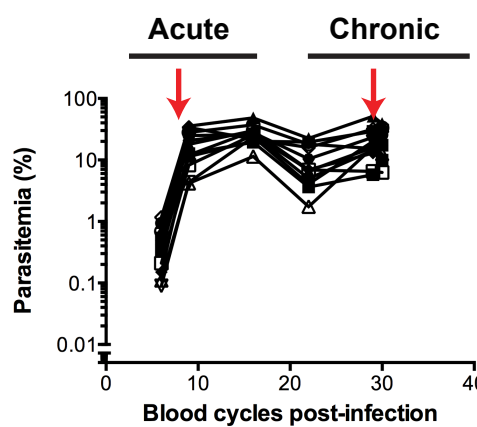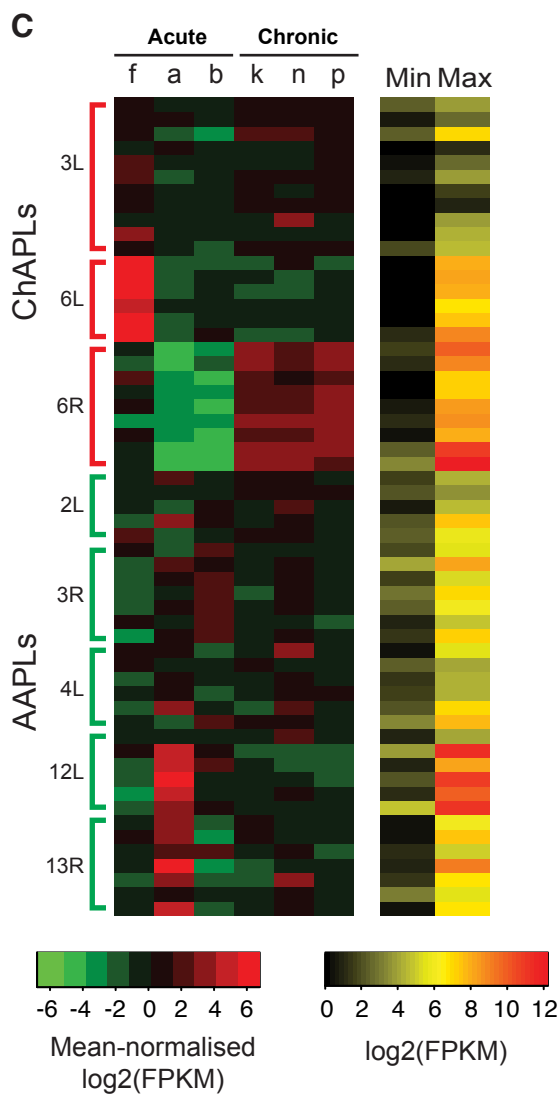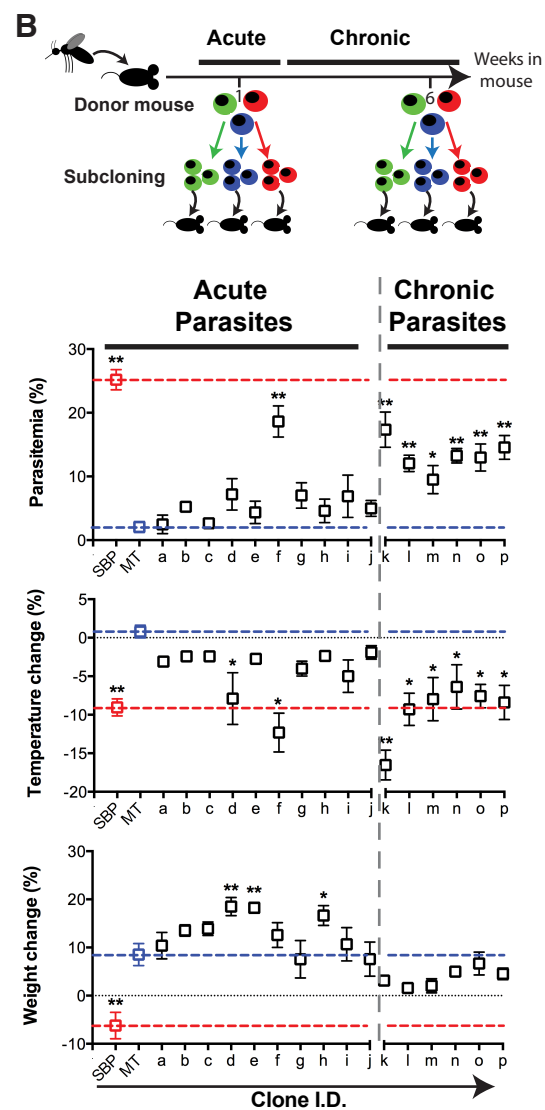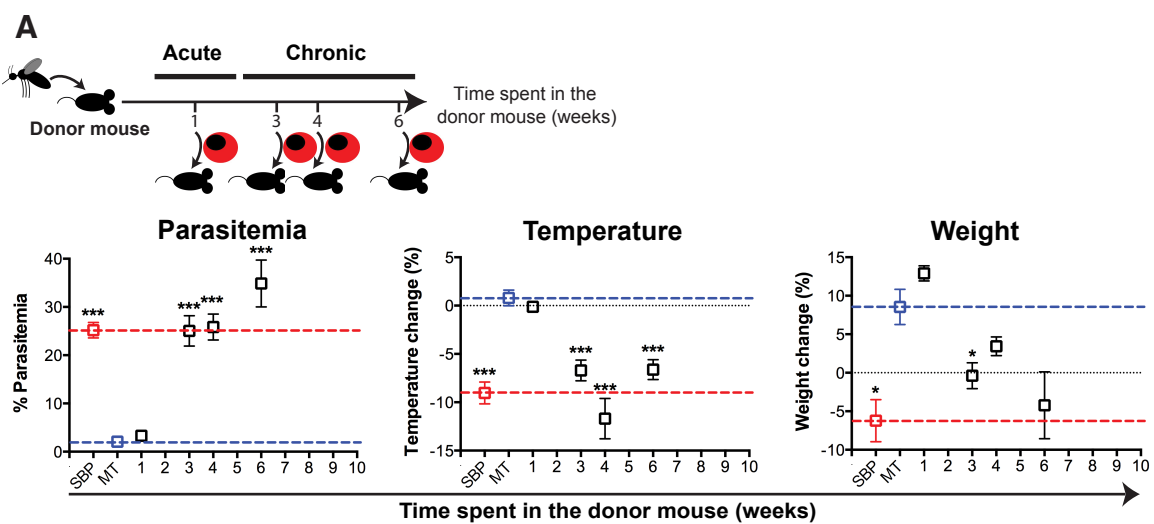
Legend:
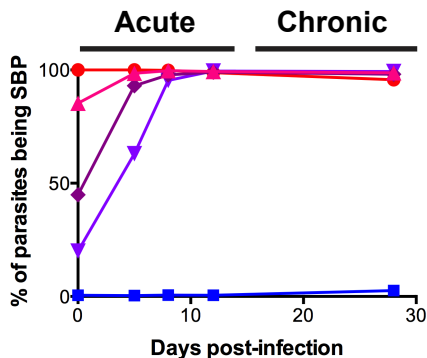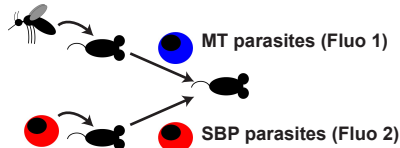- ● 100% SBP (10^5 iRBC)
- ○ 100% SBP (10^4 iRBC)
- ▲ 10% MT / 90% SBP
- ◆ 50% MT / 50% SBP
- ▼ 90% MT / 10 % SBP
- ■ 100% MT