1

Cytotoxic T cells are silenced to induce disease tolerance in human malaria

2 3 4 5	Diana Muñoz Sandoval, ^{1,2*} Florian A. Bach, ^{1*} Alasdair Ivens, ¹ Adam C. Harding, ¹ Natasha L. Smith, ¹ Michalina Mazurczyk, ³ Yrene Themistocleous, ⁴ Nick J. Edwards, ⁴ Sarah E. Silk, ^{4,5} Jordan R. Barrett, ^{4,5} Graeme J.M. Cowan, ¹ Giorgio Napolitani, ³ Nicholas J. Savill, ¹ Simon J. Draper, ^{4,5,6} Angela M. Minassian, ^{4,5,6} Wiebke Nahrendorf ^{1†} and Philip J. Spence ^{1†}
6	¹ Institute of Immunology and Infection Research, University of Edinburgh, UK
7	² Instituto de Microbiologia, Universidad San Francisco de Quito, Ecuador,
8	³ Weatherall Institute of Molecular Medicine. University of Oxford, UK.
9	⁴ The Jenner Institute, University of Oxford, UK.
10	⁵ Dept. of Biochemistry and Kavli Institute for Nanoscience Discovery, University of Oxford, UK.
11	⁶ NIHR Oxford Biomedical Research Centre, Oxford, UK.
12	these authors contributed equally
13	[†] these authors share senior authorship
14	correspondence: wnahrendorf@ed.ac.uk and pspence@ed.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

15 Abstract

16 Immunity to severe malaria is acquired quickly, operates independently of pathogen load and represents a highly effective form of disease tolerance. The mechanism that underpins tolerance 17 18 remains unknown. We developed a human re-challenge model of falciparum malaria in which 19 healthy adult volunteers were infected three times over a 12 month period to track the development 20 of disease tolerance in real-time. We found that parasitemia triggered a hardwired emergency host 21 response that led to systemic inflammation, pyrexia and hallmark symptoms of clinical malaria 22 across the first three infections of life. In contrast, a single infection was sufficient to reprogramme T cell activation and reduce the number and diversity of effector cells upon re-challenge. Crucially, 23 24 this did not silence stem-like memory cells but instead prevented the generation of cytotoxic 25 effectors associated with autoinflammatory disease. Tolerised hosts were thus able to prevent collateral tissue damage in the absence of anti-parasite immunity. 26

27 Introduction

The epidemiology of human malaria clearly shows that immunity develops in two distinct phases. 28 29 First, individuals acquire protection against severe life-threatening disease and in areas of high transmission this occurs very quickly (often before 12 months of age)^{1, 2, 3, 4}. Then after many years of exposure protection against clinical malaria is established, which promotes the transition to 31 asymptomatic infection (usually in adolescence)⁵. This temporal dissociation between clinical 32 immunity and immunity to severe malaria suggests that they are underpinned by different 33 mechanisms of host defense. In agreement, clinical immunity often coincides with control of 34 35 parasitemia (and can therefore be supported by mechanisms of host resistance) whereas immunity to severe malaria is acquired independently of pathogen load and is a form of disease tolerance¹. 36 One leading hypothesis suggests that broadly neutralising antibodies that recognise variant surface 37 antigens associated with severe malaria (such as group A/DC8 PfEMP1) could prevent severe 38 disease without affecting total pathogen load². In this scenario, immunity to severe malaria would 39 depend upon the rapid production of antibodies that can specifically eliminate pathogenic variants. 40 At present, there is limited *in vivo* evidence that such broad cross-reactivity can be achieved or that 41 neutralising antibodies can be produced within the first year of life to inhibit cytoadherence and 42 43 reduce sequestration^{6, 7}. The mechanism that underpins disease tolerance in human malaria 44 therefore remains unclear.

An alternative explanation is that the host response to infection is guickly modified to minimise the 45 harm caused by malaria parasites. It is well known that metabolic adaptations are induced during 46 the blood cycle to increase host fitness^{8, 9, 10} and control of inflammation might provide an additional 47 path towards disease tolerance¹¹. In support of this argument, inflammation decreases with 48 49 exposure in children (even at high parasite densities)¹² and pathogenic immune responses can be silenced to minimise tissue stress and toxicity in mice¹³. Importantly, host control of inflammation 50 51 as a defense strategy would not rule out a role for variant surface antigens in severe disease. After all, inflammation, tissue damage and hypoxia could create the right conditions for endothelium 52 53 activation and the selection of pathogenic variants¹⁴. As such, reducing inflammation may minimise 54 the preferential expansion of parasites associated with severe malaria - this would represent a highly effective route to disease tolerance that would not be influenced by parasite strain or 55 genotype. 56

Controlled human malaria infection (CHMI)¹⁵ offers a unique opportunity to investigate mechanisms 57 of disease tolerance in vivo. Healthy malaria-naive volunteers are inoculated with live (non-58 59 attenuated) parasites and their response is tracked throughout infection by repeated sampling (inc. all-important pre-infection samples); volunteers are inoculated with the same clonal malaria parasite 60 61 (with known variant surface antigen expression) to remove parasite genotype/phenotype as 62 confounding variables; and infections can be terminated at the same parasitemia to maintain a consistent pathogen load. Importantly, this threshold can be safely set at 5,000 - 10,000 parasites 63 m¹ to challenge the immune system with an enormous antigen load (more than 10⁷ parasites per 64 litre of blood at the peak of infection)¹⁴. What's more, adults have a far higher incidence of severe 65 disease than children during a first-in-life malaria episode, which means we can study the most at 66 risk group of individuals^{16, 17}. We therefore developed a human re-challenge model of falciparum 67 malaria to track the development of disease tolerance in real-time. 68

69 Results

70 The risk of severe malaria decreases exponentially with exposure

71 We first asked how guickly mechanisms of disease tolerance could be established in an area of endemicity. We re-analysed data from a prospective cohort study undertaken in an area of high 72 73 transmission in Tanzania¹. Here the authors performed longitudinal sampling with active case 74 detection in more than 800 infants and recorded each independent infection (including pathogen 75 load and disease severity) from birth to 4 years of age. Crucially, there was no reduction in pathogen 76 load across the study period as measured by microscopy (circulating parasitemia) or HRP2 ELISA (total parasite biomass). These data emphasise the absence of host resistance mechanisms in early 77 life. Nonetheless, the authors described a rapid decrease in the incidence of severe malaria 78 79 consistent with acquired immunity. To ask how quickly the risk of severe malaria decreased we 80 performed maximum likelihood estimation to select the best model fit for these data (Extended Data Figure 1). We asked whether the risk of severe malaria remained constant; decreased linearly or 81 82 exponentially with exposure; or decreased suddenly after n infections (where n = 1, 2, 3 etc.). The latter stepwise model performed poorly and was removed from further analysis. In contrast, the 83 84 linear and exponential models provided a good fit and performed better than a constant risk model. 85 To increase our sample size we repeated this analysis and included cases of moderately severe malaria - we collectively describe these episodes of malaria as complicated as 87.7% led to hospitalisation. Once again, the linear and exponential models provided a good fit with the latter 87 performing better by both log likelihood and AIC. Taken together, our results demonstrate that the risk of severe or complicated malaria is highest during the first infection of life and decreases 89 90 exponentially thereafter (Figure 1A-B). Mechanisms of disease tolerance are therefore beginning to be established after a single malaria episode. 91

92 Developing a human re-challenge model of malaria

To identify host adaptations that could quickly reduce the risk of severe disease we developed a 93 94 homologous re-challenge model of malaria using the most virulent human parasite Plasmodium falciparum. Ten healthy malaria-naive adult volunteers were recruited and infected by intravenous injection of parasitised red blood cells during the VAC063A and VAC063B clinical trials; all returned 96 97 for re-challenge between 4 and 8 months later (during the VAC063B and VAC063C trials, 98 respectively). Six volunteers who took part in both VAC063A and VAC063B were infected for a third 99 time during VAC063C. A blood challenge model was chosen because it standardises the infectious 100 dose, prolongs the period of blood-stage infection (cf mosquito challenge) and removes liver-stage immunity as a possible confounding factor^{18, 19}. Importantly, we used a recently mosquito-101 transmitted parasite line (< 3 blood cycles from liver egress)²⁰ since mosquitoes have been shown 102 103 to reset *Plasmodium* virulence²¹. Volunteers attended clinic the day before infection (baseline), every 104 12 hours from the day after infection until diagnosis (the peak of infection) and then during the period of drug treatment, which was initiated within 12 hours of diagnosis. These frequent visits 105 106 allowed for regular blood sampling to construct a detailed longitudinal time-course of each infection. Remarkably, we found that the parasite multiplication rate (and peak parasitemia) were 107 108 comparable between the first, second and third malaria episode (Figure 1C and Supplementary 109 Table 1). Furthermore, we found no significant change in symptoms or the frequency or severity of 110 pyrexia, anaemia, lymphopenia or thrombocytopenia (Figure 1D-G). These data thus show that

healthy adults do not acquire mechanisms of resistance (to reduce their pathogen load) and remain
 susceptible to clinical malaria. Our homologous re-challenge model therefore recapitulates the key

features of endemic malaria in early life.

114 Infection triggers a hardwired emergency host response

115 The absence of anti-parasite immunity means that any change in the host response to infection can not be attributed to a reduced number of circulating parasites. Our model therefore provides the 116 117 ideal setting in which to investigate mechanisms of disease tolerance. One potential route to 118 tolerance might be to reduce systemic inflammation - this correlates with clinical immunity in 119 endemic regions¹² but whether it also coincides with immunity to severe disease is not known. To 120 capture the acute phase response we used whole blood RNA-sequencing and DESeg2²² to identify 121 differentially expressed genes at diagnosis in first, second and third infection. We found a 122 remarkably similar pattern of interferon-stimulated gene expression regardless of infection number 123 (Extended Data Figure 2A-C). Furthermore, functional gene enrichment analysis showed that the 124 hierarchy of GO terms was near-identical (Figure 2A). These transcriptional signatures were 125 consistent with the rapid recruitment of activated monocytes and neutrophils into peripheral blood, which has been extensively described in naive hosts infected with *P. falciparum*^{23, 24} and *P. vivax*^{25,} 126 127 ²⁶, but it was surprising to see no obvious change upon re-challenge. Nevertheless, by analysing 128 each infection independently it was possible that we were missing important quantitative 129 differences and so we performed direct pairwise comparisons between first, second and third 130 infection. Initially, we compared each pre-infection time-point to identify season-dependent shifts in baseline gene expression - we found zero differentially expressed genes between infections (adj 131 132 p < 0.05 and absolute fold-change > 1.5). When we then compared each diagnosis time-point to 133 identify adaptations in the host response we again found zero differentially expressed genes (Figure 134 2B); evidently, the first three infections of life trigger a hardwired emergency response that is not 135 influenced by season or previous exposure.

Nonetheless, host control of inflammation may not be transcriptionally regulated and so we directly 136 measured systemic inflammation at protein level using a highly multiplexed custom bead assay (39 137 138 plasma analytes indicative of inflammation, coagulation, oxidative stress and metabolism). By analysing the concentration of each analyte through time we could fit mixed-effects models to test 139 140 the hypothesis that inflammation was attenuated upon re-challenge (Extended Data Figure 2D). 141 What we actually found, however, was that many of the prototypical products of monocyte and neutrophil activation (such as CXCL10, IL-1RA & MPO) were increased in second and third infection 142 143 (Figure 2C). The same was true for hallmark cytokines associated with innate lymphoid cell (ILC) or T cell activation (such as IFN_γ). Collectively, these data demonstrate that *P. falciparum* triggers a 144 145 hardwired emergency host response throughout the first three infections of life. And crucially, we 146 find no evidence that systemic inflammation is guickly attenuated.

147 A single malaria episode attenuates T cell activation

We therefore moved on to examine adaptive T cells, which are inherently plastic, proliferative and long-lived, and uniquely placed to quickly and permanently alter the host response to infection. The acute phase response to malaria causes extreme lymphopenia leading to a 30 - 70% loss of circulating cells at the peak of infection (Figure 1F). The majority of these are recruited to the

inflamed spleen^{27, 28} and so it is difficult to assess T cell activation and differentiation at diagnosis. 152 Instead, we need to analyse T cell activation after drug treatment when the emergency response 153 begins to resolve and T cells return to the circulation²⁹. At this time-point, analysing T cell 154 phenotypes in peripheral blood can provide a readout of tissue-specific immune responses. Post-155 156 treatment blood samples were not available from the VAC063A or VAC063B clinical trials but were 157 collected during VAC063C. In this trial, we recruited new malaria-naive controls to provide time-158 matched samples from first infection and infected 11 volunteers contemporaneously (three first 159 infection, two second infection and six third infection) (Supplementary Table 1). As such, we switched to a cross-sectional analysis of VAC063C to incorporate post-treatment time-points. 160

161 Six days after drug treatment (designated T6) lymphopenia had completely resolved and all other 162 clinical symptoms of malaria (inc. fever) had receded. Furthermore, markers of systemic inflammation had returned almost entirely to baseline and this was observed in every volunteer 163 regardless of infection number (Extended Data Figure 3A). It was therefore a surprise to find a large 164 transcriptional signature in whole blood at T6 (Extended Data Figure 3B). This signature did not 165 166 overlap with the emergency response captured at diagnosis and instead the differentially expressed 167 genes had unique functional enrichment terms relating to cell cycle and nuclear division (Extended Data Figure 3C). Remarkably, this proliferative burst was only observed in volunteers undergoing 168 169 their first infection of life (Extended Data Figure 3D).

170 Myeloid cells are generally terminally differentiated and do not proliferate after their release from 171 the bone marrow; it therefore seemed likely that our whole blood RNA-sequencing data were 172 capturing the return of activated T cells to the circulation. To explore this further we sorted CD4⁺ T cells with an effector or effector memory phenotype (CCR7^{neg} CD45RA^{neg}), which have been shown 173 to dominate the adaptive response to malaria in a naïve host²⁹. Sorting was performed on whole 174 175 blood one day before challenge and six days after drug treatment (Supplementary Data File 1). Analysis of the cell surface markers used for sorting revealed dramatic activation of CD4⁺ T cells in 176 177 first infection but not second or third infection (Figure 3A) and this striking observation was repeated in a separate CHMI study (VAC069) that infected volunteers with Plasmodium vivax, a different 178 species of human malaria parasite that is evolutionarily divergent from *P. falciparum*³⁰ (Figure 3B 179 180 and Extended Data Figure 4A). Our data therefore reveal that the unique feature of a first-in-life malaria episode is fulminant CD4⁺ T cell activation and that this response is attenuated as the risk 181 182 of severe disease begins to exponentially decrease.

183 T_{H} 1 polarisation is a unique feature of first infection

To explore the transcriptional landscape of activated CD4⁺ T cells we undertook RNA-sequencing 184 185 on the sorted effector (effector memory) CD4⁺ T cells obtained from volunteers infected with P. 186 falciparum. We found almost 6000 differentially expressed genes at T6 in first infection (adj p < 0.05187 and > 1.5 fold-change) and functional gene enrichment analysis showed that these cells were 188 proliferative and had increased their capacity for oxidative phosphorylation (Figure 3C and 189 Extended Data Figure 4B). Furthermore, they had upregulated each of the major costimulatory and 190 inhibitory receptors required to control their fate, and increased their expression of the signature 191 chemokine receptors and transcription factors associated with T_H1 polarisation. What's more, the cytokines IFN γ and IL-21 were both strongly induced, which could indicate that infection stimulates 192 double-producers³¹ or that follicular helper T cells were also released from the spleen (Figure 3D). 193

194 In third infection, IFNy and IL-21 were once again significantly upregulated at T6 together with $T_{H}1$ -195 associated chemokine receptors; CD4⁺ T cells are thus transcriptionally responsive to re-challenge (Figure 3C-D). Yet remarkably, the transcription factors that drive $T_{H}1$ differentiation (T-bet and 196 197 STAT1) were no longer induced, and neither were the costimulatory molecules or inhibitory receptors observed in first infection. We therefore asked what modifies CD4⁺ T cell activation to 198 199 avert T_{H1} polarisation upon re-challenge. We found no transcriptional evidence that they were 200 quiescent or anergic in third infection (Extended Data Figure 4C) and crucially the transcription 201 factors that epigenetically enforce exhaustion were not upregulated (Figure 3C). Furthermore, there 202 was no evidence for activation-induced cell death (Extended Data Figure 4D) and TCR^β sequencing 203 revealed near-identical repertoires before and after infection, which shows that attenuation was not 204 caused by the clonal deletion of activated T cells after the first or second malaria episode (Figure 3E). There was also no evidence that activated CD4⁺ T cells were diverted towards a regulatory 205 206 fate; for example, hallmarks of T_{B1} differentiation (such as Eomes and IL-10) were absent (Figure 207 3C-D).

An alternative explanation could be suppression by conventional regulatory CD4⁺ T cells but our data showed that Tregs were activated in first (not third) infection (Extended Data Figure 4E). It therefore appears that malaria-experienced hosts can launch a specialised adaptive T cell programme that maintains cytokine production without causing extensive activation, proliferation or T_H1 polarisation. Importantly, we can identify this modified T cell response by transcriptional profiling of whole blood (Extended Data Figure 4F) and it will therefore be possible to identify tolerised hosts in an endemic setting without the need for complex cell isolation protocols.

215 Stem-like memory CD4⁺ T cells respond to re-challenge

216 A major limitation of bulk RNA-sequencing is that there is very little power to detect transcriptional changes if only a small proportion of cells are responding, as seemed to be the case in third 217 218 infection. The transcriptional landscape of re-activated cells therefore remained largely unclear. To overcome this limitation we used single cell RNA-sequencing to examine individual CD4⁺ T cells, 219 which were sorted at baseline and T6 from three volunteers undergoing first infection and three 220 undergoing third infection; samples were barcoded with oligo-tagged antibodies and superloaded 221 onto the 10X Chromium platform³². Three libraries were then prepared and sequenced (cell surface, 222 223 5' gene expression and V(D)J) and after quality control (including doublet exclusion) our dataset 224 contained approximately 25,000 cells per infection with a median of 75,000 reads and 1000 genes 225 per cell (Extended Data Figure 5A-B). Initially, we concatenated all data to identify the heterogeneity of CD4⁺ T cells across the dataset and uncovered 13 unique clusters (7 with a non-naive phenotype) 226 227 (Figure 4A and Extended Data Figure 5C). We then split the data by volunteer and time-point to examine cluster abundance by linear regression and found that in third infection a single cluster of 228 229 CD4⁺ T cells (cluster 10) had expanded at T6 (Figure 4B). Importantly, this cluster was 230 transcriptionally high for the canonical activation marker CD38 (Extended Data Figure 5D).

We therefore examined the signature genes associated with cluster 10 and found enrichment of the transcription factor TCF1 (Supplementary Table 2); this has been associated with circulating T_{FH} cells in a mouse model of malaria³³ but we found no significant enrichment for PD1, CXCR5 or BCL6 (Extended Data Figure 5E). Instead, we found that cluster 10 was enriched for LEF1, SELL, TRIB2 and PELI1 (Figure 4C), which have all been linked to the maintenance of stem-like properties in T

cells^{34, 35, 36}. Indeed, we found that BLIMP1, whose induction is required for terminal differentiation, 236 was repressed in cluster 10 (Extended Data Figure 5F) and BLIMP1 is a known target of TCF1^{37, 38}. 237 Cluster 10 was also found to expand in first infection but was exceeded by cluster 11, which was 238 239 similarly high for CD38 but otherwise transcriptionally distinct (Figure 4B-C and Extended Data 240 Figure 5D). Cells in cluster 11 were enriched for signature genes associated with T_{H} polarisation, 241 terminal differentiation and cytotoxicity, including NKG7 and KLRB1 (CD161), which are hallmarks of rapidly responding short-lived effector cells^{39, 40}. Notably, cluster 11 downregulated TCF1 and 242 this could also be seen in our bulk RNA-sequencing data (Figure 3C and Extended Data Figure 5F). 243

244 There was no evidence, however, that cells in cluster 11 were descendants of cluster 10 and instead 245 trajectory analysis indicated bifurcation along the TCF1/BLIMP1 axis (Figure 4D and Extended Data 246 Figure 5G). In agreement, TCR V gene usage in cluster 10 shared remarkable overlap with all other T cell clusters but diverged in cluster 11 (Figure 4E). For example, cells in cluster 11 increased 247 248 TRAV16 / TRBV18 and minimised TRBV5-1 leading to an overall reduction in repertoire diversity (Figure 4F and Supplementary Data File 2). These data are consistent with the rapid clonal 249 250 expansion of short-lived effectors in first infection, which are silenced upon re-challenge. In 251 contrast, stem-like memory T cells are re-activated and increase their repertoire diversity (to 252 resemble naive T cells) during third infection (Figure 4F). This suggests that cluster 10 is selected 253 for a polyclonal repertoire across multiple malaria episodes and is consistent with reverse TCR evolution⁴¹. Crucially, the maintenance of stem-like memory appears to support long-lived humoral 254 255 responses as we observe remarkable stepwise boosting of class-switched IgG antibodies against 256 MSP1 and AMA1 (Figure 4G).

257 Cytotoxic T cells are silenced for a minimum 8 months

The terminal differentiation of cytotoxic CD4⁺ T cells has been described in autoinflammatory 258 disease, the tumour microenvironment and chronic (or latent) viral infection^{42, 43, 44}. We were 259 nevertheless surprised that our bulk and single cell RNA-sequencing data both suggested that T_H1 260 polarised CD4⁺ T cells acquired transcriptional features of cytotoxicity during a first-in-life malaria 261 episode (Figure 3D and 4C). Furthermore, the scale of activation (up to 80% of all circulating cells 262 with an effector or effector memory phenotype (Figure 3A-B)) was staggering, and far exceeded 263 observations made in other human challenge models of acute infection, such as influenza⁴⁵ or 264 typhoidal salmonella⁴⁶. Our data therefore suggested that rather than the response to re-challenge 265 266 being unusual, the host response to a first infection was excessive and disproportionate. To explore 267 this idea we used mass cytometry so that we could examine more cells, extend our analysis to all major T cell subsets and quantify the heterogeneity of activated cells at protein level. To this end 268 269 we designed an antibody panel that prioritised key markers of T cell function and fate 270 (Supplementary Table 3).

Whole blood samples were preserved in Cytodelics Stabilisation buffer within 30 minutes of blood draw at baseline, diagnosis and T6 (as well as during convalescence) during VAC063C. As before, we concatenated all data to identify the heterogeneity of T cells across the entire dataset and uncovered 49 unique clusters (Supplementary Data File 3). As expected most of the diversity was observed within the non-naive CD4⁺ and CD8⁺ T cell subsets (Figure 5A). Tracking the frequency of each cluster through time then resolved dynamic changes in the T cell compartment (Supplementary Data File 4) and we performed linear regression on cell count data using edgeR⁴⁷

to identify the differentially abundant clusters at each time-point (FDR < 0.05 and absolute foldchange > 2).

In first infection eighteen T cell clusters increased in abundance at T6 and all had an activated 280 281 (CD38^{hi} Bcl2^{lo}) phenotype; these were comprised of twelve adaptive and six innate-like clusters that 282 spanned every T cell lineage (Figure 5B-D). In sum, these clusters accounted for approximately 283 40% of the T cell compartment (Extended Data Figure 6A) and included cytotoxic effectors belonging to the CD8⁺, gamma delta and NK T cell subsets - all were characterised by upregulation 284 of the chemokine receptor CX3CR1 (Figure 5E). CD4⁺ T cells nevertheless dominated this response 285 and we observed enormous heterogeneity with the expansion of nine distinct clusters. There were 286 287 common traits (such as expression of the memory marker CD45RO) and most had an effector 288 (CCR7^{neg}) phenotype but the expression of other activation and differentiation markers (inc. T-bet) was highly variable (Figure 5E). The largest cluster of activated CD4⁺ T cells had an unusual CD27^{neg} 289 CX3CR1^{pos} cytotoxic phenotype and there were other unexpected features of terminal 290 differentiation such as the appearance of perforin-expressing CD161^{pos} effectors. We therefore 291 292 used limma to complement our analysis of cluster abundance and modelled changes in marker 293 expression through time; this confirmed significant upregulation of both granzyme B and perforin 294 in CD4⁺ T cells six days after drug treatment (Extended Data Figure 6B). Evidently, cytotoxicity is a 295 cardinal feature of the entire T cell response to a first-in-life malaria episode.

296 In third infection the T cell response was primarily driven by activated innate-like T cells but their abundance was substantially reduced (compared to first infection) and the induction of granzyme 297 298 B and perforin was suppressed (Figure 5C-D and Extended Data Figure 6B). The number of cytotoxic CD8⁺ T cells was also reduced but the most dramatic change was observed within the 299 CD4⁺ T cell compartment. Here all but one of the nine clusters that expanded in first infection was attenuated and the previously dominant CD27^{neg} CX3CR1^{pos} cytotoxic cluster was completely 301 silenced. These changes were not associated with an increase in regulatory T cells (Foxp3^{hi} CD39^{hi} 302 303 cells did not expand after re-challenge) or an upregulation of suppressor molecules (Extended Data 304 Figure 6B). In fact, only two clusters increased in size in third compared to first infection - a small subset of double negative T cells and a cluster of T-bet^{lo} Eomes^{neg} effector memory CD4⁺ T cells 305 306 (Figure 5C and 5E). The majority of activated CD4⁺ T cells belonged to this cluster and in contrast 307 to first infection these cells had already increased in abundance at diagnosis (Supplementary Data File 4); these presumably represent the stem-like memory cells identified by single cell RNA-308 309 sequencing. In sum, approximately 10% of the T cell compartment was activated after re-challenge, 310 which more closely aligns with the scale of T cell activation observed in other febrile human 311 infectious diseases.

312 Controlling T cell activation protects host tissues

Collectively, these data show that the T cell response to a first-in-life malaria episode is dominated by heterogeneous CD4⁺ T cells that present with unusual features of terminal differentiation and cytotoxicity. To directly test whether T cell activation could be pathogenic we measured biomarkers of collateral tissue damage, a common histological feature of severe malaria. The blood-stage of infection frequently causes liver injury in naive hosts^{48, 49, 50} and auto-aggressive T cells can directly kill primary human hepatocytes⁵¹ - we therefore measured alanine aminotransferase (ALT) to provide a readout of hepatocellular death *in vivo*. In first infection, two out of three volunteers had

abnormal ALT (more than the upper limit of the reference range) when activated T cells were released from inflamed tissues (Extended Data Figure 7A). This was accompanied by increased 321 322 gamma-glutamyl transferase (GGT) and aspartate aminotransferase (AST) leading to moderate or 323 severe adverse events in both volunteers. In contrast, there was little evidence of any deviation from 324 baseline in liver function tests after re-challenge. To expand our sample size, we performed a meta-325 analysis using a previously published surrogate dataset; specifically, we examined post-treatment 326 ALT measurements in almost 100 volunteers experiencing a first-in-life infection as part of a human challenge study⁴⁹. Importantly, we only included CHMI trials that were directly comparable to our 327 own re-challenge study - that is they used the same clonal parasite genotype (3D7 or the parental 328 329 NF54 line); parasites had recently been mosquito transmitted; and similar end-points were applied (treatment at around 10,000 parasites ml⁻¹) (Figure 6A). Remarkably, we found that the prevalence 331 of abnormal ALT was reduced from 75% during first infection to 25% upon re-challenge (Figure 6B). And in those rare cases where ALT was increased in second or third infection adverse events 332 333 were mild (not moderate or severe) even though pathogen load was increased.

334 We then looked more closely at the relationship between T cell activation and liver injury in our 335 cohort. We found a strong positive correlation between ALT and the frequency of activated effector or effector memory CD4⁺ T cells, regulatory T cells and cytotoxic T cells (Figure 6C and 337 Supplementary Table 4). Of note, these all had a strong negative correlation with antibody titres. What's more, they had no discernible relationship with markers of systemic inflammation or clinical 338 339 symptoms of malaria, which were instead closely associated with each other but located in a 340 separate clade. Finally, we co-cultured PBMC from first infection with HepG2 cells (an immortalised 341 hepatocyte cell line) and could show enhanced killing at T6 (compared to baseline) in vitro (Figure 6D and Extended Data Figure 7B-C). This was particularly prominent for v315 who had by far the 342 largest population of activated CD27^{neg} CX3CR1^{pos} cytotoxic CD4⁺ T cells (Extended Data Figure 343 344 6A). Taken together, these data show that the risk of tissue damage and injury can be significantly 345 reduced in the absence of parasite control, which provides the first *in vivo* evidence that long-lived 346 mechanisms of disease tolerance operate in human malaria and can be acquired after a single 347 infection. Moreover, protection does not require the attenuation of systemic inflammation but 348 instead coincides with host control of T cell activation and cytotoxicity.

349 Discussion

It has long been recognised that immunity to severe malaria is acquired early in life, offers protection against all manifestations of severe disease and usually precedes clinical immunity (the transition 351 to asymptomatic infection) by more than a decade^{2, 3, 4, 5}. We also know that immunity to severe 352 malaria does not require improved parasite control and is thus underpinned by acquired 353 354 mechanisms of disease tolerance¹. This is an important distinction to make if we want to understand 355 how to reduce malaria mortality. Host control of inflammation could provide a rapid route to disease 356 tolerance but our data show that malaria parasites trigger a hardwired emergency response across 357 the first three infections of life. This correlates closely with symptomatology in our cohort and is consistent with repeated episodes of fever in children living in endemic settings. In contrast, T cell 358 activation is guickly modified to limit the number and diversity of effector cells, and to avoid cell 359 fates associated with collateral tissue damage and autoinflammatory disease. Our data can therefore begin to explain why large multi-centre genome-wide association studies have repeatedly 361 362 failed to identify immune loci that associate with severe malaria⁵² - variation in human T cells is almost exclusively driven by non-heritable factors⁵³. 363

364 If we want to understand how the immune response contributes to severe disease we instead need 365 to ask how malaria triggers such widespread and indiscriminate activation of cytotoxic T cells. These are enriched for markers of terminal differentiation and likely represent short-lived effectors 367 (as evidenced by the downregulation of TCF1 and de-repression of BLIMP1 in CD4⁺ T cells). What remains unclear, however, is whether these are malaria-specific cells or activated via bystander 368 369 (TCR-independent) mechanisms. If the latter, these could be derived from pre-existing memory cells, which have a far lower threshold for activation than naive cells^{54, 55}. This would certainly explain 370 371 the magnitude of the response, and may also explain the increased susceptibility of adults to severe 372 disease during a first-in-life infection (adults have a larger memory pool than infants and children). Alternatively, malaria may activate autoreactive T cells in the same way that it can activate B cells 373 that produce autoantibodies^{56, 57}. Central tolerance (which deletes self-reactive clones in the 374 375 thymus) is only partially effective and there is enormous degeneracy in TCR reactivity⁵⁸. Systemic 376 infection with a pathogen carrying more than 5000 protein-coding genes may therefore lead to considerable cross-reactivity between parasite and host. In either case, the activation of bystander 377 378 or cross-reactive T cells would explain our observation that the TCRB repertoire is essentially 379 unchanged after a first malaria episode.

So how would activation of a large heterogeneous pool of T cells promote severe disease? One possibility is that T cell activation causes extensive activation of the endothelium leading to 381 enhanced parasite cytoadherence and increased pathogen load (an important contributing factor 382 for severe malaria). This has yet to be formally tested but could be assessed ex vivo using 3D 383 models of the human microvasculature⁵⁹. This is an important next step because group A/DC8 384 variants are known to associate with severe disease but do not have an intrinsic growth advantage *in vivo*¹⁴; a host factor must therefore promote their selection. An alternative (though not mutually exclusive) explanation is that cytotoxic T cells disrupt the endothelial barrier in critical organs and 387 there is emerging evidence for this pathological process in children and adults with cerebral 388 malaria^{60, 61}. Furthermore, cytotoxicity may damage parenchymal tissue and our data support a 389 direct role for activated T cells in hepatocyte death. Remarkably, we found that PBMC-mediated 390 391 killing of HepG2 cells in vitro did not require peptide loading and we therefore propose that tissue

392 damage is the result of TCR-independent off-target cytotoxicity. This would explain why we only observe raised transaminases after parasite clearance - activated T cells traffic to the liver during 393 resolution of the immune response to facilitate their clearance⁶². In this scenario, liver injury would 394 be attenuated after re-challenge because the size of the T cell response is diminished. This could 395 396 be tested in vivo using fine needle liver aspirates, which have recently been obtained during CHMI to examine tissue resident memory⁶³. In the meantime, we can conclude that explosive cytotoxic T 397 398 cell activation and tissue damage are the unique features of a first-in-life infection and both are silenced after a single malaria episode to promote tissue health. Importantly, this host adaptation 399 persists for at least 8 months, which would be sufficient to maintain immunity against severe malaria 400 401 through a prolonged dry season.

402 It is therefore apparent that we need to start asking how the T cell response to malaria is modified so quickly if we want to obtain a mechanistic understanding of disease tolerance. One possibility is 403 that infection initiates heritable epigenetic programmes that reduce T cell responsiveness. Our data 404 do not provide transcriptional evidence of anergy or exhaustion but this hypothesis needs to be 405 406 directly tested. An alternative explanation could be suppression by conventional (thymus-derived) 407 regulatory T cells but we found no evidence for their activation in second or third infection - it seems 408 that Treqs are redundant at this point because there is no explosive effector response to keep in 409 check. In much the same way, we found no evidence for the development of IL-10 producing T_{R1} cells, which have been extensively described in endemic settings^{64, 65, 66}. At first glance, it could be 410 assumed that this is because T_R1 cells are generated in the context of chronic stimulation and CHMI 411 is a model of acute infection. However, even in CHMI (where drug treatment is usually initiated 412 413 within 8 to 12 days of infection) malaria chronically stimulates the immune system. This is because once-infected red cells carry parasite-derived surface antigens and continue to circulate for weeks 414 after drug treatment⁶⁷; follicular dendritic cells present antigens within follicles for months (possibly 415 years)⁶⁸; and hemozoin (an insoluble by-product of infection) persists in lymphoid tissues 416 indefinitely⁶⁹. Indeed, evidence of chronic stimulation can clearly be seen in our study - two clusters 417 of activated (CD38^{hi} Bcl2^{lo}) gamma delta T cells are still expanded 45 days after infection 418 (Supplementary Data File 4). It is this persistence of parasite-derived material that we believe is 419 responsible for the long-term protection afforded by a single malaria episode. After all, hemozoin-420 loaded dendritic cells have a reduced capacity to stimulate T cell proliferation in vitro⁷⁰ and a similar 421 422 mechanism operating within the spleen could provide the simplest explanation for tolerance.

423 In much the same way, antigen presenting cells can be modified by chronic viral infection and this 424 has been shown to preferentially promote the generation of memory CD8⁺ T cells with stem-like properties⁷¹. A similar stem-like fate has recently been described for CD4⁺ T cells⁷² and these appear 425 to share remarkable overlap with the TCF1⁺ cluster that is re-activated in our cohort upon re-426 challenge. TCF1 endows memory CD4⁺ T cells with the capacity for asymmetric division⁷³ to 427 promote self-renewal and the differentiation of T_{FH} precursors^{37, 74}, which are primed to enter 428 germinal centres or else support B cell responses in the extra follicular pathway. Crucially, 429 430 asymmetric division is the most efficient way to generate long-lived immunity and the activation of stem-like memory CD4⁺ T cells in third infection (8 months after second infection) coincides with 431 432 the boosting (MSP-1) and diversification (AMA-1) of parasite-specific class-switched antibodies. A second requirement for long-lived immunity when pathogens persist is reverse TCR evolution, 433 which gradually selects for low affinity T cell clones; this is beneficial because memory cells with a 434 high affinity TCR are driven towards senescence by chronic stimulation⁴¹. Importantly, effective anti-435

parasite immunity (and the transition to asymptomatic infection) does not require high affinity antibodies but instead antibody diversification^{75, 76, 77}. The polyclonal repertoire of stem-like memory CD4⁺ T cells in third infection may therefore represent the selection of low affinity clones with increased longevity and an enhanced capacity to support a broad repertoire of B cells. As such, we argue that the slow acquisition of anti-parasite immunity in endemic settings (and in our own CHMI study) is not a failure of T cell help.

442 So do equivalent data from endemic settings support our findings? An important caveat here is that in almost every case samples are collected around the time of drug treatment when activated T 443 444 cells are still sequestered in inflamed tissues. Nevertheless, we can ascertain that circulating 445 cytotoxic (granzyme B^{pos}) CD4⁺ T cells do not expand in children or adults during an uncomplicated episode of febrile disease⁷⁸. In this study, all of these patients would be expected to have acquired 446 447 immunity to severe malaria because they have lived for at least 2 years in a perennial high transmission setting (183 infective bites per person per year). We can therefore conclude that there 448 449 is no evidence of cytotoxicity when the risk of severe disease is low. On the other hand, adults with effective anti-parasite immunity have an expanded population of CD161^{pos} central memory CD4⁺ T 450 451 cells⁷⁹, which may share some similarities with the stem-like memory cells identified here. That's 452 because stem-like, T_{CM} and T_{FH} cells are all situated on the same branch of the TCF1-BLIMP1 axis (with cytotoxic, T_{H1} and T_{R1} cells on the other side)^{37, 38, 74, 80}. And crucially, the frequency of CD161^{pos} 453 T_{CM} CD4⁺ T cells correlates with naturally acquired antibody titres against a broad range of targets 454 (including RH5). These data are therefore consistent with our findings that exposure to malaria 455 456 quickly silences pathogenic T cells but does not prevent the maintenance of stem-like memory. 457 That this can be achieved through CHMI (a short drug-cured infection) means that interventions 458 that reduce pathogen load but don't completely eliminate blood-stage parasites could reduce the 459 incidence of clinical malaria in the short-term and have the added benefit or providing long-lived immunity to severe malaria. Moreover, our data indicate that this protection would persist even if 460 461 these interventions were stopped (drug cover or monoclonal antibodies) or lose efficacy (blood-462 stage vaccines).

463 Methods

464 Modelling the risk of severe malaria

465 To examine the risk of severe disease as a function of prior exposure we used data from a Tanzanian birth cohort¹. In brief, the authors examined 882 children for *P. falciparum* every 2 - 4 weeks and 466 recorded parasite density and disease severity for each independent infection. Data showing the 467 468 incidence of severe or complicated malaria (the latter summing severe and moderately severe 469 cases) were extracted from Figure 2C and Supplementary Figure 3, respectively, as shown in the published article. We then used these data to plot the total number of cases of malaria (including 470 471 mild or uncomplicated episodes); impute missing values by least squares regression (as shown in Extended Data Figure 1); and model risk (as a function of exposure) using the fraction of severe or 472 complicated cases divided by the total number of cases at each order of infection. Binomial 473 474 regression models were then fit using maximum likelihood methods to test whether the empirical data was best explained by a constant risk (no change), a linear decrease in risk or an exponential 475 476 decay with increasing exposure. To compare model performance we used the Akaike Information 477 Criterion (AIC) to balance goodness of fit with model complexity, and viewed a reduction in AIC > C478 2 as indicating a superior model. Maximum likelihood coefficients were estimated 479 with stats4::mle() and stats::glm().

480 Clinical trial design - Plasmodium falciparum

All volunteers were healthy malaria-naive adults aged between 18 and 50 years and were enrolled in 481 482 up to three CHMI studies - VAC063A (November 2017), VAC063B (March 2018) and VAC063C (November 2018). The VAC063A and VAC063B trials evaluated vaccine efficacy of the recombinant 483 484 blood-stage malaria protein RH5.1 in AS01_B adjuvant (GSK) whereas the VAC063C trial investigated 485 the durability of anti-parasite immunity. We collected whole blood samples in each trial from the 486 non-vaccinated (control) volunteers and analysed 8 volunteers in VAC063A (all receiving their first 487 infection), 10 volunteers in VAC063B (8 receiving their second infection and 2 their first infection) and 11 volunteers in VAC063C (6 receiving their third infection, 2 receiving their second infection 488 and 3 their first infection). Our study design and laboratory analysis plan are shown in 489 Supplementary Table 1. All clinical trials received ethical approval from the UK NHS Research Ethics 490 491 Service - Oxfordshire Research Ethics Committee A for VAC063A and VAC063B (reference 492 16/SC/0345) and South Central Oxford A for VAC063C (reference 18/SC/0521) - and were registered at ClinicalTrials.gov (NCT02927145 and NCT03906474). VAC063A-C were sponsored by 493 494 the University of Oxford, carried out in the UK at the Centre for Vaccinology and Tropical Medicine and conducted according to the principles of the current revision of the Declaration of Helsinki 2008 495 496 (in full conformity with the ICH Guidelines for Good Clinical Practice). Volunteers signed written 497 consent forms and consent was checked prior to each CHMI. Details of volunteer recruitment, 498 inclusion/exclusion criteria and group allocation can be found in Minassian et al. (for VAC063A and VAC063B)⁸¹ and Salkeld et al. (for VAC063C)⁸². 499

500 During each CHMI all volunteers were infected with *P. falciparum* (clone 3D7) blood-stage parasites 501 by direct intravenous infusion. The inoculum was thawed and prepared under aseptic conditions 502 and volunteers received between 452 and 857 infected red cells in a total volume of 5 ml saline. 503 Starting one day after challenge volunteers attended clinic every 12 hours for assessment and blood

504 sampling, and parasite density was measured in real-time by gPCR. The parasite multiplication rate was then calculated by fitting linear models to log10 transformed gPCR data, as previously 505 described⁸³. In VAC063A thick blood films were evaluated at each time-point by experienced 506 microscopists and diagnosis required volunteers fulfil two out of three criteria: a positive thick blood 507 film (one viable parasite in 200 fields) and/or qPCR data showing at least 500 parasites ml⁻¹ blood 508 509 and/or symptoms consistent with malaria. In VAC063B and C microscopy was dropped to minimise 510 the variation in parasitemia that was observed between volunteers at diagnosis; importantly, this protocol change did not impact volunteer safety and the average parasitemia at diagnosis remained 511 comparable to VAC063A. The new criteria for diagnosis were: asymptomatic with any gPCR result 512 above 10,000 parasites ml⁻¹ or symptomatic with a gPCR result above 5000 parasites ml⁻¹. In all 513 514 cases, volunteers were treated with artemether and lumefantrine (Riamet) except where its use was 515 contraindicated and atovaguone and proguanil (Malarone) were given instead. In our analysis, we refer to the blood sample taken immediately before drug treatment (< 30 minutes) as the diagnosis 516 517 time-point.

518 Clinical symptoms of malaria (malaise, fatigue, headache, arthralgia, back pain, myalgia, chills, 519 rigor, sweats, nausea, vomiting and diarrhoea) were recorded by volunteers on diary cards or during clinic visits. All symptoms were recorded as adverse events and assigned a severity score: 0 520 521 (absent), 1 (transient or mild discomfort), 2 (mild to moderate limitation in activity) or 3 (severe limitation in activity requiring assistance). Pyrexia was scored as absent (≤ 37.5°C), mild (37.6 -522 523 38.2°C), moderate (38.3 - 38.9°C) or severe (\geq 39°C). And full blood counts and blood chemistry 524 (inc. electrolytes, urea, creatinine, bilirubin, alanine aminotransferase, alkaline phosphatase and 525 albumin) were evaluated at the Churchill and John Radcliffe Hospital in Oxford.

526 Clinical trial design - Plasmodium vivax

In VAC069A six healthy malaria-naive adults were enrolled to test the infectivity of a new 527 528 cryopreserved stabilate containing a clonal field isolate of *P. vivax* (PvW1), which had been carefully prepared for use in CHMI⁸⁴. Volunteers were infected with blood-stage parasites by direct 529 intravenous infusion (as above) and the immune response to PvW1 was comprehensively 530 531 characterised and compared to P. falciparum (3D7). In VAC069B three of these volunteers returned for a homologous re-challenge (8-months after VAC069A) and two additional malaria-naive adults 532 received their first challenge with PvW1. In both clinical trials treatment was initiated once two 533 534 diagnostic criteria were met: a positive thick blood smear, more than 5000 or 10,000 parasites ml⁻¹ blood and/or symptoms consistent with malaria. Treatment consisted of artemether and 535 536 lumefantrine (Riamet) or atovaquone and proguanil (Malarone) if Riamet was contraindicated. Whole blood sampling, processing and downstream analyses were performed analogously to the VAC063 537 538 trials (see below). VAC069A and B were sponsored by the University of Oxford, received ethical approval from the UK NHS Research Ethics Service - South Central Hampshire A (reference 539 18/SC/0577) - and were registered at ClinicalTrials.gov (NCT03797989). The trials were conducted 540 541 in line with the current version of the Declaration of Helsinki 2008 and conformed with the ICH Guidelines for Good Clinical Practice. 542

543 Processing whole blood for RNA and plasma

544 Venous blood was drawn into K₂EDTA-coated vacutainers (BD #367835). To preserve RNA for transcriptional analysis 1 ml whole blood was mixed thoroughly with 2 ml Tempus reagent 545 546 (ThermoFisher #4342792) and samples were stored at -80°C. No more than 2 hours passed 547 between blood draw and RNA preservation. To obtain platelet-depleted plasma 3 ml whole blood 548 was divided into two 2 ml Eppendorf tubes and centrifuged at 1000 xg for 10 minutes (at 4°C) to pellet cellular components. Working on ice, plasma was then carefully transferred to a new 2 ml 549 tube and centrifuged at 2000 xg for 15 minutes (at 4°C) to pellet platelets. Cell-free platelet-depleted 550 551 plasma was aliquoted into 1.5 ml Eppendorf tubes, snap frozen on dry ice and stored at -80°C.

552 Whole blood RNA-sequencing

553 RNA was extracted from whole blood using the Tempus spin RNA isolation reagent kit (ThermoFisher #4380204) according to the manufacturer's instructions. To account for the reduced 554 starting volume and to maintain Tempus stabilizing reagent at the correct final concentration we 555 added just 1 ml PBS to each sample after thawing. Diluted samples were then centrifuged at 3000 556 xg for 30 minutes (at 4°C) to pellet nucleic acids. Pellets were resuspended in RNA purification 557 558 resuspension solution and centrifuged on a silica column to remove non-nucleic acid contaminants. 559 After washing, the column was incubated for 2 minutes at 70°C before eluting nucleic acids. 40 µl eluate was then subjected to DNA digestion using the RNA clean and concentrator-5 kit (Zymo 560 Research #R1013). Purified RNA was eluted using 30 µl DNase/RNase-free water, quantified using 561 562 a Qubit Fluorometer (HS RNA assay kit, ThermoFisher #Q32852) and RNA integrity assessed using an Agilent Bioanalyzer 2100 (RNA 6000 nano kit, Agilent #5067-1511). The average RNA integrity 563 564 number (RIN) was 9 (98% of samples > 8, lowest RIN 7). Libraries were prepared by Edinburgh Genomics (United Kingdom) using the TruSeg stranded mRNA library prep kit (Illumina #20020595). 565 Stranded library preparation allows transcript expression to be estimated more accurately; in 566 567 particular it is more effective in quantifying anti-sense gene expression, properly assigning 568 transcripts to putative coding genes and to resolve ambiguity in reads from overlapping genes. 569 Libraries were sequenced using the NovaSeg 6000 Illumina platform yielding 50 bp paired end (PE) 570 reads. These short reads are sufficient to accurately capture gene expression thanks to the well 571 annotated human transcriptome. The average number of reads per sample passing QC across all 572 samples was 8.45×10^7 .

573 Data analysis of whole blood RNA-sequencing

574 Quality and content of FASTQ files, which contain raw PE sequencing reads, were assessed using 575 FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). A single sample (volunteer 1040 baseline second infection) failed quality control and we therefore excluded all RNAseq data 576 577 from this volunteer's second infection during analysis. TruSeq primer sequences were removed using Cutadapt v1.9, reads were aligned to the Ensembl release 96 Homo sapiens transcript set 578 with Bowtie2⁸⁵ v2.2.7 (parameters – very-sensitive -p 30 – no-mixed – no-discordant – no-unal) 579 and reads that mapped to globin transcripts were discarded (on average 11.7% of mapped reads, 580 range 1.6 - 38.1%). We opted for bioinformatic globin depletion as it is highly sensitive and 581 reproducible whereas depleting globin during RNA preparation can compromise the amount and 582 quality of RNA recovered^{86, 87}. Following these steps, the average alignment rate to the human 583

584 transcriptome across all samples was 88%. A matrix of normalised counts for each transcript was 585 obtained from sorted, indexed BAM files using Samtools idxstats (http://www.htslib.org/doc/samtools.html). Transcript imported into 586 counts were the R/Bioconductor environment (v3.6) and differential gene expression analyses (pairwise group 587 comparisons) were performed using functions within the DESeg2²² package. If cShrink was applied 588 to the output of each pairwise comparison (type normal) and lists of differentially expressed 589 590 transcripts were subsequently analysed in R. Non-coding transcripts were removed and multiple transcripts annotated to the same gene were consolidated by keeping the transcript with the 591 highest absolute fold-change. Only protein-coding transcripts with an adjusted p value (adj p) < 592 0.05 and a fold-change > 1.5 were considered differentially expressed. Volcano/radar/dot plots and 593 heatmaps were generated using the ggplot2⁸⁸ package. The gene lists used for cell cycle analysis 594 were manually compiled from published datasets^{89, 90} and we allocated genes to a single phase 595 based on cell cycle transcriptional network analysis^{91, 92}. 596

597 Functional gene enrichment analysis using ClueGO

Lists of differentially expressed genes were imported into ClueGO^{93, 94} v2.5.7. ClueGO identified 598 significantly enriched GO terms (Biological Process and Molecular Function) associated with these 599 genes and placed them into a functionally organised non-redundant gene ontology network based 600 on the following parameters: adj p cutoff = 0.01; correction method used = Bonferroni step down; 601 602 min. GO level = 5; max. GO level = 11; number of genes = 3; min. percentage = 5; GO fusion = true; 603 sharing group percentage = 40; merge redundant groups with > 40% overlap; kappa score 604 threshold = 0.4; and evidence codes used [All]. Each of the functional groups was assigned a unique colour and a network was then generated using an edge-weighted spring-embedded layout based 605 606 on kappa score - groups were named by the leading GO term (lowest adj p with min. GO level 5 or 607 6). Merged networks were constructed by inputting two lists of differentially expressed genes; for each GO term information on what fraction of associated genes were derived from each list was 608 609 retained. Any GO term containing > 60% associated genes from a single list was considered to be 610 enriched in that group, otherwise GO terms were considered to be shared.

611 Multiplexed plasma protein analysis

612 The concentration of 39 analytes was measured in plasma samples collected at baseline, during 613 infection, diagnosis, 6 days after drug treatment (T6) and 28 or 45 days post-challenge 614 (convalescence). Plasma was thawed on ice and centrifuged at 1000 xg for 1 minute (at 4°C) to remove potential protein aggregates. We customised 4 LEGENDplex panels from BioLegend and 615 616 performed each assay on filter plates according to the manufacturer's instructions. Samples and 617 standards were acquired on an LSRFortessa flow cytometer (BD) and FCS files were processed 618 using LEGENDplex software (version 7.1), which automatically interpolates a standard curve using the plate-specific standards and calculates analyte concentrations for each sample. All samples 619 620 from v1040 were excluded after failing QC and downstream data analysis was performed in R (v3.6). To determine which plasma proteins varied significantly through time we used the Ime4 package to 621 622 fit a separate linear mixed-effects model for each analyte. All available time-points were included 623 and models were fit to log2 transformed data with time-point as a categorical fixed effect and 624 volunteer as a random effect. A Kenward-Roger approximation was used to calculate p values 625 (using the pbkrtest package), which were adjusted for multiple testing using the BenjaminimedRxiv preprint doi: https://doi.org/10.1101/2021.08.19.21262298; this version posted October 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in

perpetuity. It is made available under a CC-BY 4.0 International license .

626 Hochberg method. Results were visualised with ggplot2 and analytes with at least a 1.5 fold-change from baseline to diagnosis or T6 are shown. To determine which plasma proteins varied between 627 628 first infection and re-challenge we used Ime4 to fit mixed-effects models that included time-point 629 and infection number as categorical fixed effects (with volunteer as a random effect). In this case, 630 linear hypothesis testing was performed using multcomp's glht function (with Benjamini-Hochberg correction for multiple testing). Only analytes that were significant in both second and third infection 631 (versus first infection) are shown (adj p < 0.05 and fold-change > 1.5). 632

633 Flow-sorting CD4⁺ T cell subsets

634 We flow-sorted CD4⁺ T cell subsets for bulk RNA-sequencing from 3 ml whole blood collected in K₂EDTA vacutainers. After red cell lysis (erythrocyte lysis buffer, eBioscience #00-4300-54) 635 636 leukocytes were washed in PBS supplemented with 2% fetal bovine serum (heat inactivated and 637 0.22 µm filtered FBS Premium Plus, Gibco #16000044) containing 5 mM EDTA (Life Technologies 638 #AM9260G). We then blocked Fc receptors (human TruStain FcX, BioLegend #422302) and incubated leukocytes for 20 minutes (at 4°C) with the following fluorophore-conjugated antibodies: 639 CD3 (clone OKT3), CD4 (clone OKT4), CD127 (clone A019D5), CD25 (clone M-A251), CCR7 (clone 640 G043H7), CD45RA (clone HI100), CD38 (clone HIT2) and HLA-DR (clone L243) (all from BioLegend). 641 642 Cells were washed with cold PBS (containing 2% FBS and 5 mM EDTA) and filtered through a 40 643 µm cell strainer (BD #352340) immediately before sorting. We used a FACSAria III or Fusion cell sorter with a 70 µm nozzle running FACS Diva v8 software (sort setting = purity) to simultaneously 644 sort 10,000 naive (CD127^{pos} CCR7^{pos} CD45RA^{pos}), regulatory (CD25^{hi} CD127^{neg}) and effector or 645 effector memory (CD127^{pos} CCR7^{neg} CD45RA^{neg}) CD4⁺ T cell subsets. Cells were sorted directly into 646 647 RNase-free sterile 1.5 ml screw cap tubes (ThermoFisher #11529924) containing 1 ml TRIzol 648 Reagent (ThermoFisher #15596026), and incubated for 5 minutes at room temperature. Samples were then stored at -80° C. Note that sort purity was > 95% for every sample - this was assessed 649 650 in real-time by simultaneously sorting naive CD4⁺ T cells into cold PBS (containing 2% FBS and 5 651 mM EDTA) and re-acquiring these on the cell sorter.

652 Bulk RNA-sequencing of CD4⁺ T cell subsets

RNA was extracted using a modified phenol-chloroform protocol⁹⁵ with 1-Bromo-3-653 654 chloropropane (Sigma-Aldrich #B62404) and Isopropanol (Acros Organics #423835000). Total RNA 655 was guantified and integrity assessed using an Agilent Bioanalyzer 2100 (RNA 6000 pico chip, Agilent #5067-1513); all sequenced samples had a RIN value above 8. cDNA was generated from 656 657 1 ng total RNA using the SMART-Seq v4 ultra low input RNA kit (Takara Bio #634894) and amplified using 12 cycles of PCR. Amplified cDNA was purified using AMPure XP beads (Beckman Coulter 658 659 #A63880) and guantified on a Qubit Fluorometer (dsDNA high sensitivity kit, ThermoFisher 660 #Q32851). The quality of the amplified cDNA was then assessed by Bioanalyzer (DNA high sensitivity kit, Agilent #5067-4626). Libraries were constructed from 150 pg cDNA using the Nextera 661 XT DNA library preparation kit (Illumina #FC-131-1024) according to the manufacturer's 662 instructions. As above, libraries were quantified by Qubit and quality was assessed by Bioanalyser 663 664 to measure the fragment size distribution. Using this information, samples were combined to create equimolar library pools that were sequenced on a NextSeg 550 Illumina platform to yield 75 bp PE 665 666 reads; the average number of PE reads per sample passing QC across all samples was 2.74 x 10⁷.

RNA-sequencing analysis of CD4⁺ T cell subsets

667

Quality and content of FASTQ files were assessed using FastQC and reads were aligned to the 668 Ensembl release 96 Homo sapiens transcript set with Bowtie2 v2.2.7 (parameters -very-sensitive 669 670 -p 30 -- no-mixed -- no-discordant -- no-unal). Our median alignment rate was 66.4% 671 after removing primer and adapter traces (SMART-Seg v4 primers, polvG, polvT and polvN) with 672 Cutadapt v1.9. The sorted, indexed BAM files were then used to obtain a matrix of normalised 673 counts for each transcript using Samtools idxstats. Transcript counts were imported into the 674 R/Bioconductor environment (v3.6) and differential gene expression analyses (pairwise group 675 comparisons) were performed using functions within the DESeg2 package. If cShrink was applied 676 to the output of each pairwise comparison (type normal) and lists of differentially expressed 677 transcripts were filtered by removing all non-coding transcripts and retaining only those with an adj 678 p < 0.05 and fold-change > 1.5. Multiple transcripts annotated to the same gene were consolidated by keeping the transcript with the highest absolute fold-change. Heatmaps and circular stacked 679 bar plots were generated using the ggplot2 package. 680

681 Cryopreservation of peripheral blood mononuclear cells

682 Peripheral blood mononuclear cells (PBMC) were isolated from 5 ml whole blood, which was diluted 683 with an equal volume of Dulbecco's phosphate buffered saline (D-PBS) containing 2% fetal bovine 684 serum (FBS) (STEMCELL Technologies #07905) and layered onto 3.5 ml Lymphoprep in a 15 ml SepMate tube (STEMCELL Technologies #07851 and #07905). SepMate tubes were centrifuged at 685 686 1200 xg for 10 minutes (at room temperature), the upper plasma-containing layer was discarded and the remaining supernatant above the insert (which contains the PBMC) was poured into a 15 687 688 ml tube. Cells were then washed twice in D-PBS containing 2% FBS and cell concentration was determined using a CASY cell counter. Finally, PBMC were resuspended in FBS (heat inactivated 689 690 and 0.22 µm filtered FBS Premium Plus, Gibco #16000044) containing 10% Dimethyl sulfoxide (DMSO, Sigma-Aldrich #D8418) at a concentration of 5 x 10^6 cells ml⁻¹ and transferred to cryovials. 691 These were placed in a Corning CoolCell freezing container, which was subsequently placed at -692 693 80°C. After 24 hours, PBMC-containing cryovials were transferred to liquid nitrogen where they were stored long-term. To prepare cryopreserved PBMC for use in experiments cryovials were 694 695 guickly thawed in a 37°C water bath and diluted by adding 10x volume thawing media (RPMI 1640 696 supplemented with 10% FBS and 20 units ml⁻¹ DNase I (Sigma-Aldrich #D4513)) one drop at a time 697 with continuous agitation. Cells were centrifuged at 350 xg for 10 minutes (at room temperature) and washed twice in thawing media. After washing, viable cell counts were calculated using a 698 699 hemocytometer.

700 *T* cell receptor repertoire sequencing

We analysed the TCRβ repertoire of the six volunteers who underwent third infection during VAC063C, assessing their bulk repertoire prior to challenge (baseline) and 28 days post-challenge (convalescence) during their first (VAC063A) and second (VAC063B) malaria episode. RNA was extracted from thawed PBMC using the Quick-RNA miniprep plus kit (Zymogen) as per the manufacturer's instructions, including DNase treatment. A modified protocol from Mamedov *et al.*⁹⁶ was then used to synthesise cDNA, which was treated with 1 µl Uracyl DNA glycosylase (5 units µl⁻ 1). A nested PCR then generated TCRβ V region amplicons with outer primers using indexed forward

708 primers composed of the SMART synthesis oligo sequence fused to a P7 Illumina tag (CAAGCAGAAGACGGCATACGAGATXXXXXGGCGAAGCAGTGGTATCAACGCAGAGT) and a 709 710 within the TCR region fused P5 Illumina reverse primer С to а tag (AATGATACGGCGACCACCGAGATCTACACACACSTTKTTCAGGTCCTC). 711 Library QC was performed by Nanodrop and Bioanalyzer, and asymmetric 400 bp + 100 bp sequencing was then 712 713 performed on an Illumina NovaSeg using custom read primers (ACTCTGCGTTGATACCACTG index 714 with CGAGATCTACACACACSTTKTTCAGGTCCTC for read 1 and 715 GGCGAAGCAGTGGTATCAACGCAGAGT for read 2). The average number of functional TRBV reads was 2.58 x 10⁶ per sample resulting in an average of 13,410 unique CDR3 sequences per 716 sample. Quality control was performed on the raw FASTQ files using FastQC and sequencing data 717 that incorporated UMI was demultiplexed using MiGEC⁹⁷. All sequences were aligned using 718 MiXCR⁹⁸ software utilising IMGT⁹⁹ nomenclature. MiGEC and MiXCR standard error correction 719 720 thresholds were used (including a minimal nucleotide quality score of 20 within the target gene region) and only in-frame functional CDR3 sequences were included in downstream analyses. 721 Custom pipelines of Python scripts were then used to analyse and plot the MiXCR output and 722 723 proportional TRBV gene usage was calculated for each sample. Note that for visualisation, TRBV 724 genes that had a zero count at any given time-point were replaced with the minimum gene count 725 prior to log10 transformation.

726 Single cell RNA-sequencing of CD4⁺ T cells

Cryopreserved PBMC from volunteers 313, 315 and 320 (first infection) and 1061, 1068 and 6032 727 728 (third infection) were prepared 2 days before infection (baseline) and 6 days post-treatment (T6) during VAC063C. To undertake single cell RNA-sequencing these samples were thawed, washed, 729 counted and resuspended at a concentration of 3 x 10⁶ cells ml⁻¹ in PBS supplemented with 2% 730 731 fetal bovine serum (heat inactivated and 0.22 µm filtered FBS Premium Plus, Gibco #16000044) and 732 5 mM EDTA (Life Technologies #AM9260G). Cells were then stained with surface antibodies exactly 733 as described above (see Flow-sorting CD4⁺ T cell subsets) but with the addition of TotalSeq-C 734 oligo-tagged barcoding antibodies that recognise CD298 and β2-microglobulin (a combination of 735 clones LNH-94 and 2M2, available from BioLegend). For each sample we flow-sorted 300,000 CD4+ 736 T cells (CD3^{pos} CD4^{pos}) into 15 ml conical tubes containing 5 ml RPMI 1640 supplemented with 10% FBS Premium Plus. At the same time we sorted an equal number of CD4⁺ T cells into 5 ml cold PBS 737 738 (containing 2% FBS and 5 mM EDTA) and re-acquired these on the cell sorter to check purity, which 739 was > 95% for every sample.

740 After sorting and purity checks, six individually barcoded samples were pooled (one pool for first infection and a separate pool for third infection) and the concentration of each pool was carefully 741 742 measured and adjusted to 1262 cells µl⁻¹. Each pool was then loaded into 2 wells of an A Chip, 743 which itself was loaded onto a 10X Genomics Chromium Controller. Note that we intentionally superloaded the controller in order to capture ~ 30,000 singlets per pool based on the workflow 744 745 described in Stoeckius et al.³² Captured cells were then tagged with 10X cell barcodes in GEMs, cDNA was amplified and libraries were constructed according to the manufacturer's instructions 746 747 (Chromium single cell V(D)J reagent kit with feature barcoding technology for cell surface protein protocol CG000186 revision C). From each GEM three libraries were constructed and separately 748 749 indexed: i) the cell surface barcode, ii) 5' gene expression and iii) the T cell receptor α and β chains 750 (after amplification of the V(D)J regions). Following QC and quantification (using the Agilent

Bioanalyzer and TapeStation as well as Qubit) the libraries were pooled at a ratio of 8% cell surface, 84% gene expression and 8% V(D)J and then sequenced on a NovaSeq 6000 Illumina platform to yield at least 66,000 150 bp PE reads per cell. In total we obtained 5931 million PE reads (2848 and 3083 million PE reads for first and third infection, respectively) of which 487 million PE reads (8.2%) were derived from the cell surface library, 4858 million PE reads (81.9%) from the 5' gene expression library and 586 million PE reads (9.9%) from the V(D)J library.

757 Single cell RNA-sequencing analysis

Cell Ranger multi (v6.0.2) was used to align 5' gene expression and V(D)J sequencing reads to the 758 759 GRCh38 reference genome. This initial step was performed independently for first and third 760 infection since both pools used the same set of six TotalSeq-C barcodes - C0251 761 (GTCAACTCTTTAGCG), C0252 (TGATGGCCTATTGGG), C0253 (TTCCGCCTCTCTTG), C0254 762 (AGTAAGTTCAGCGTA), C0255 (AAGTATCGTTTCGCA) and C0256 (GGTTGCCAGATGTCA) (all 763 BioLegend). Feature-Barcode matrices were then loaded into R (v4.2.0) using Seurat (v4.3.0) and oligo count matrices were generated from the cell surface library using CITE-seg-Count (v1.4.2) 764 765 (https://doi.org/10.5281/zenodo.2590196). These datasets were filtered to only include 10X cell 766 barcodes that were detected in both the 5' gene expression and cell surface libraries, and after 767 normalisation and scaling PCA-initialised tSNE was used to visualise the data (first seven principal components, perplexity = 100). For each pool of samples, inspection of the oligo expression levels 768 769 revealed six large clusters that corresponded to single positive cells (or singlets). As such, we could use nearest neighbour identification (PCA 1-6) and Louvain clustering to assign each singlet to a 770 771 sample and exclude cells with multiple oligo tags (doublets). At this stage the data from first and 772 third infection were combined, and a standard Seurat workflow was followed for quality control. 773 Briefly, we excluded cells with more than 2500 feature counts (or less than 200) and cells with more 774 than 5% mitochondrial transcripts. Data were log normalised and the top 2000 most variable 775 features were selected (after variance stabilisation) for computing of the principal components (PC). Harmony¹⁰⁰ (v0.1.0) was used to integrate data and reduce individual variation, and Seurat's 776 wrapper for the Louvain algorithm (FindClusters) was used to cluster cells based on the first twelve 777 778 PC (with default settings at a resolution of 0.8). An initial pass of this analysis identified a persistent 779 and hyperexpanded cluster of TRBV11-3^{pos} CD4⁺ T cells. These cells were unique to v1068 (the only 780 CMV seropositive volunteer) and did not respond to malaria; we therefore excluded this cluster from 781 all downstream steps to avoid confounding signatures of malaria-induced activation.

782 Next differential cluster abundance was modelled using negative binomial regression with glm.nb() 783 from the MASS package. Time-point was fit as a categorical variable, p values were adjusted (Benjamini-Hochberg method) and an FDR < 0.05 was considered significant. Slingshot¹⁰¹ was then 784 used to perform pseudotime analysis on Harmony-derived PC (default settings); the leaf node was 785 set to either cluster 10 or 11 (CD38^{hi} cells) but no trajectory passed through both clusters. Finally, 786 787 to analyse TCR diversity we used createHTOContigList() to filter and match Cell Ranger's annotations on the assembled contigs to those extracted from scRepertoire¹⁰² (v1.8.0). V gene 788 789 usage (TRAV and TRBV) was analysed using standard immunarch (v0.9.0) 790 (https://github.com/immunomind/immunarch) workflows (excluding ambiguous assignments) and the amino acid sequences of the TCR in each cluster were used to calculate their Gini coefficient 791 (using convenience functions). Data were visualised using ComplexHeatmap¹⁰³ or ggplot2. 792

793 Processing whole blood for mass cytometry

Venous blood was collected in K₂EDTA-coated vacutainers, stabilised within 30 minutes of blood 794 795 draw in whole blood stabilisation buffer (Cytodelics #hWBCS002) and stored at -80°C. For antibody staining samples were quickly thawed in a 37°C water bath and then fixed and red cell 796 797 lysed for 15 minutes using a whole blood preservation kit (Cytodelics #hC002). Next cells were 798 permeabilised with Maxpar barcode permeabilisation buffer (Fluidigm #201057) and each sample 799 was barcoded using Cell-ID 20-plex palladium barcodes (Fluidigm #201060). Samples were then pooled and stained with our T cell focussed surface antibody mix (see Supplementary Table 3) for 800 801 30 minutes. After washing, cells were fixed and permeabilised with the Maxpar nuclear antigen 802 staining buffer set (Fluidigm #201063) and incubated with the nuclear antibody mix for 45 minutes. 803 Cells were then washed and fixed for 10 minutes in 1.6% formaldehyde diluted in PBS (ThermoFisher #28906). After a final round of washes, cells were resuspended at a concentration 804 of 3 x 10⁶ cells ml⁻¹ in 72.5 nM Cell-ID Intercalator Ir solution (Fluidigm #201192A) and stored 805 overnight at 4°C. Samples were acquired the next day on a freshly tuned Helios mass cytometer (acquisition rate 300-500 events per second) using the WB injector and 10% EQ four element 807 808 calibration beads (140Ce, 151Eu, 165Ho and 175Lu, Fluidigm #201078).

809 Mass cytometry data analysis

The Fluidigm CvTOF software (version 6.7) generated FCS files, which were normalised¹⁰⁴ and 810 debarcoded¹⁰⁵ using the R package CATALYST¹⁰⁶. Samples were compensated using single stained 811 beads¹⁰⁷. After exclusion of normalisation beads and doublets we gated on CD45^{pos} CD3^{pos} T cells 812 using the Cytobank web portal (https://www.cytobank.org). We then inspected the intensity 813 814 distribution of each channel and removed those with low variance (CD16, CD69, CXCR5, GATA3, 815 ROR γ t, TCR $\gamma\delta$ and TIM3). The remaining 30 markers were used for UMAP and FlowSOM clustering. UMAP¹⁰⁸ was used to generate a 2D projection of this high-dimensional dataset. Here 816 phenotypic similarity of cells within and between populations is preserved in the Euclidean distance 817 of the projection. We used its R implementation in the scater¹⁰⁹ package, which in turn relies on 818 uwot (https://www.github.com/ilmelville/uwot). Features were scaled to unit variance and the 15 819 820 nearest neighbours were considered for embedding. UMAP coordinates were then exported for visualisation using ggplot2. FlowSOM¹¹⁰ uses self-organising maps (SOM) to efficiently categorise 821 822 cytometry data into non-overlapping cell populations and was performed using CATALYST (default 823 parameters, target: 100 clusters, 50 metaclusters). After manual inspection we merged two phenotypically similar clusters to avoid overclustering¹¹¹ and ended up with 49 discrete T 824 cell clusters. The R/Bioconductor package ComplexHeatmap was used to visualise T 825 cell phenotypes and the arcsine transformed signal intensity of each marker was independently 827 scaled using a 0-1 transformation across all clusters.

To analyse differential cluster abundance we used the workflow laid out by Nowicka et 828 829 al.¹¹² FlowSOM cluster cell counts were modelled linearly with time-point as a dependent categorical variable and volunteer as a fixed effect using the diffcyt¹¹³ implementation of edgeR⁴⁷. 830 The edgeR functions automatically normalise cluster counts for the total number of cells and 831 improve statistical power by sharing information on cluster count variance between clusters. 832 833 Pairwise comparisons were performed relative to baseline, and clusters with an FDR < 0.05 and 834 absolute fold-change > 2 were deemed to vary significantly through time. We assessed differential

835 cluster abundance independently for volunteers receiving their first or third infection. We also assessed whether marker expression varied significantly through time and to do this we 836 merged clusters belonging to the same T cell lineage according to their expression of CD4, 837 838 CD8, CD56, V δ 2 and V α 7.2. Adaptive CD4⁺ and CD8⁺ T cells were then split into naive, effector, effector memory and central memory subsets based on their expression of the markers CD45RA. 839 840 CD45RO, CD57 and CCR7. All regulatory T cells (CD4^{pos} CD25^{hi} CD127^{neg}) were merged into a single 841 cluster. Linear models derived from the limma package, which is optimised for continuous data, 842 were then used to independently assess differential marker expression relative to baseline using 843 pairwise comparisons with moderated t-tests; a shift in median expression of at least 10% and an 844 FDR < 0.05 were required for significance. Results were visualised using ComplexHeatmap with row-wise z-score transformed marker intensities shown for each lineage (or subset) of T cells. 845

846 Meta-analysis of liver injury

A surrogate dataset from Reuling et al.⁴⁹ was used to assess the risk of liver injury during a first-in-847 848 life infection compared to re-challenge. We extracted data from every CHMI study that used the 3D7 P. falciparum clone (or its parental NF54 line), initiated infection via mosquito bite or direct 849 blood challenge and that had a treatment threshold based on thick smear positivity (estimated to 851 be at least 5000 parasites ml⁻¹ blood). This produced data for 95 volunteers across 7 CHMI studies. 852 Notably, Reuling et al. used longitudinal data to show that liver function test (LFT) abnormalities 853 peaked up to 6 days post-treatment in line with our own T6 time-point. And furthermore, in every CHMI study (including our own) LFT abnormalities were graded using the same adaptation of the 854 855 WHO adverse event grading system. An LFT reading > 1.0 but < 2.5 times the upper limit of normal was graded as mild; a reading > 2.5 but < 5.0 times the upper limit was graded moderate; and a 857 reading > 5.0 times the upper limit was graded severe. For ALT, the upper limit of normal was 35 858 or 45 units litre⁻¹ for female and male volunteers, respectively. Data from the 95 volunteers in Reuling et al. and the 3 volunteers undergoing first infection in our VAC063C study were pooled for analysis 859 860 and we calculated a weighted peak parasitemia across the cohort by using the mean number of parasites ml⁻¹ and the number of volunteers in each of the 8 CHMI studies. To statistically test 861 862 whether an abnormal ALT reading was more prevalent in the 98 volunteers experiencing their first malaria episode compared to the 8 volunteers undergoing re-challenge (either second or third 863 infection) we used Barnard's test, which examines the association of two independent categorical 864 variables in a 2 x 2 contingency table. A p value below 0.05 was considered significant. 865

866 Pearson correlation analysis

867 To assess the relationship between liver injury, T cell activation, systemic inflammation and the clinical symptoms of malaria we constructed a Pearson correlation matrix using our VAC063C 868 dataset. Importantly, this analysis was agnostic of infection number (all volunteers were included). 869 We input the concentration of ALT and the percentage of activated effector (effector memory) CD4+ 870 T cells, regulatory T cells and cytotoxic (granzyme B^{pos}) T cells (all at T6); the fold-change of 871 872 lymphocytes, haemoglobin and all significant plasma analytes (those shown in Extended Data 873 Figure 3A), which were calculated at diagnosis or T6 (relative to baseline) according to their largest absolute fold-change; the maximum parasite density and maximum core temperature (at any time-874 875 point up to 48 hours post-treatment); and the titres of AMA1/MSP1-specific class-switched 876 antibodies (measured 28 days after challenge). All data were log2 transformed and Pearson

correlation was performed in R using the corrplot function. Correlation coefficients were used forunsupervised hierarchical clustering by Euclidean distance.

879 In vitro cytotoxicity

The hepatoma cell line HepG2 was kindly provided by Shaden Melhem (University of Edinburgh). 881 HepG2 cells were maintained at 37°C and 5% CO2 in RPMI 1640 supplemented with 10% heatinactivated fetal bovine serum (FBS Premium Plus, Gibco #16000044), 1 x GlutaMAX (Gibco 882 #35050061), 100 units ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin (Gibco #15140122). Expansion 883 884 was performed in tissue culture treated culture flasks (Corning #430641U) and cells were detached 885 using TrypLE Express Enzyme (Gibco #12604013). For cytotoxicity assays, HepG2 cells were 886 seeded at a density of 5.5 x 10⁵ cells per well in 96 well tissue culture treated plates pre-coated with Poly-D-Lysine (Gibco #A3890401) and incubated for 24 hours to allow formation of a confluent 887 monolayer. 888

889 The next day PBMC were thawed, washed, counted and resuspended at a concentration of 3 x 10⁶ 890 cells ml⁻¹ in RPMI 1640 supplemented with 10% fetal bovine serum (FBS Premium Plus), 2 mM L-Glutamine (#25030081), 1 mM sodium pyruvate (#11360070), 10 mM HEPES (#15630080), 500 µM 891 β-mercaptoethanol (#31350010) and 1 x MEM non-essential amino acids (#M7145) (all Gibco 892 except amino acids from Sigma-Aldrich). Cells were transferred to ultra low attachment plates 893 894 (Corning #3473) and stimulated with cell activation cocktail (BioLegend #423302) for 3 hours at 895 37°C and 5% CO₂ - the working concentration of PMA and ionomycin used was 40.5 nM and 669.3 896 nM, respectively. PBMC were then harvested, washed twice and added onto HepG2 monolayers at increasing ratios of effector (PBMC) to target cells (HepG2). After 24 hours of co-culture the 897 release of lactate dehydrogenase (LDH) from damaged HepG2 cells was measured in the culture 898 899 supernatant using the CyQUANT LDH cytotoxicity assay according to the manufacturer's instructions (Invitrogen #C20301). Background LDH release was quantified in wells containing only 900 901 HepG2 cells. Absorbance was read on a FLUOstar Omega plate-reader.

902 Acknowledgements

903 We thank Paul Sopp for T cell sorting during the VAC063 trial and Neil Ashley for support with single 904 cell RNA-sequencing. Flow-sorting, 10X Genomics and CyTOF data were generated within the flow 905 cytometry, single cell and mass cytometry facilities at the Weatherall Institute of Molecular Medicine (University of Oxford), which are supported by MRC Human Immunology Unit core funding 906 (MC UU 00008) and the Oxford Single Cell Biology Consortium (OSCBC). CD4⁺ T cell subset 907 908 RNAseg libraries were prepared and sequenced by the Edinburgh Clinical Research Facility at the 909 University of Edinburgh, which receives financial support from NHS Research Scotland (NRS). Whole blood RNAseq libraries were prepared and sequenced by Edinburgh Genomics, which is 910 supported through core grants from NERC (R8/H10/56), MRC UK (MR/K001744/1) and BBSRC 911 (BB/J004243/1). We would like to extend our thanks to the VAC063 and VAC069 clinical and 912 laboratory teams for assistance, and to all of the volunteers who participated in this study. 913

914 Funding

915 The VAC063 trial was supported in part by the Office of Infectious Diseases, Bureau for Global 916 Health, U.S. Agency for International Development (USAID) under the terms of the Malaria Vaccine Development Program (MVDP) contract AID-OAA-C-15-00071, for which Leidos, Inc. was the prime 917 918 contractor. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the USAID. The VAC063 trial was also supported in part by the National Institute for 919 920 Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those 921 of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The VAC069 trial was supported by funding from the European Union's Horizon 2020 research and 922 923 innovation programme under grant agreement for MultiViVax (number 733073). The single cell RNAsequencing experiment was supported by the Human Infection Challenge Network for Vaccine 924 Development (HIC-Vac) funded by the GCRF Networks in Vaccines Research and Development, 925 which was co-funded by the MRC and BBSRC. This UK-funded award is part of the EDCTP2 927 programme supported by the European Union. DMS is the recipient of a Darwin Trust of Edinburgh 928 PhD studentship and FB, ACH and NLS are each the recipient of a Wellcome Trust PhD studentship (grant no. 203764/Z/16/Z, 226857/Z/23/Z and 204511/Z/16/A, respectively). SJD was the recipient 929 of a Wellcome Trust Senior Fellowship (grant no. 106917/Z/15/Z) and is a Jenner Investigator. This project was supported by the Wellcome Trust-University of Edinburgh Institutional Strategic 931 Support Fund, and PJS is the recipient of a Sir Henry Dale Fellowship jointly funded by the 932 933 Wellcome Trust and the Royal Society (grant no. 107668/Z/15/Z).

934 Conflict of interest

The authors have no conflict of interest to declare and the funders had no role in study design, data interpretation or the decision to submit the work for publication.

937 Data availability

All bulk RNAseg data (whole blood and sorted T cell subsets) have been deposited in NCBI's Gene 938 939 Expression Omnibus and are accessible through GEO SuperSeries accession number GSE172481. 940 Single cell RNA-sequencing data have also been deposited in NCBI's Gene Expression Omnibus and are available through accession number GSE275092. Bulk TCRβ sequencing data have been 941 deposited in the European Nucleotide Archive and are available through accession number 942 PRJEB71976. And CyTOF (mass cytometry) data have been deposited at flowrepository.org - these 943 can be accessed through experiment numbers FR-FCM-Z47Z (VAC063C); FR-FCM-Z3HA 944 (VAC069A); and FR-FCM-Z465 (VAC069B). 945



Figure 1. The risk of severe malaria decreases exponentially with exposure. Data were extracted 947 from Gonçalves et al.¹ to examine the frequency of severe (A) or complicated (B) malaria during the 948 first 14 infections of life in infants living in a hyperendemic setting. We performed maximum likelihood estimation to select the best model fit for these data; the black line shows the best fit and 949 grey shading represents the 95% confidence intervals. In both cases, an exponential decay provided a better fit than either a linear decay or constant risk. In (A) n = 102 severe episodes and 951 in (B) n = 199 complicated episodes of malaria (see Extended Data Figure 1 for case imputation 952 and model performance). (C-G) Healthy malaria-naive adults were infected up to three times with 953 954 P. falciparum (clone 3D7) by direct blood challenge. Repeated sampling before, during and after 955 each infection allowed us to track the development of disease tolerance in real-time. (C) Parasite 956 growth curves for first, second and third infection; each line represents a volunteer and lines are colour-coded by infection number. Parasite density was measured in peripheral blood by gPCR 957 every 12 hours. The grey box represents the lower limit of quantification (20 parasites ml⁻¹) and the 958 treatment threshold of 10,000 parasites ml⁻¹ is denoted by the black line. Maximum core body 959 temperature (D) and minimum haemoglobin (E) recorded during each infection (up to 48 hours posttreatment). Each symbol represents one volunteer with a line shown at the median. Grey shading 961 indicates normal range (115 - 165 g litre⁻¹ haemoglobin for female and 130 - 180 g litre⁻¹ haemoglobin 962 for male volunteers). Lymphocytes (F) and platelets (G) were quantified in circulation the day before 963 964 infection (baseline), 6 days after challenge (c6), at the peak of infection (diagnosis) and approx. 1 month after drug treatment (convalescence). Boxplots show the median and IQR, and whiskers represent the 95% confidence intervals. Sample size (n) is 10 for first and second infection and n = 6 for third infection. There was no statistically significant difference between groups using a significance threshold of 5% (Kruskal-Wallis test).



Figure 2. Infection triggers a hardwired emergency host response. (A) RNA-sequencing was used 970 to identify differentially expressed genes in whole blood at diagnosis (versus baseline) (adj p < 0.05and > 1.5 fold-change). ClueGO was then used for functional gene enrichment analysis and placed 971 972 significant GO terms into functional groups by relatedness. Shown are the leading GO terms from 15 non-redundant groups with the lowest adj p value in first infection. The same GO terms are 973 974 plotted in second and third infection (note that each infection was analysed independently). (B) Radar plots (or 3-way volcano plots) show the number of differentially expressed genes in whole 975 976 blood between each infection - the left plot compares all baseline samples and the right plot 977 diagnosis. Dashed lines represent the centre-point for each volcano plot and the position of each 978 dot relative to this line shows up- or downregulation. There were no differentially expressed genes in any of the six pairwise comparisons (adj p < 0.05 and > 1.5 fold-change). (C) 39 plasma analytes 979 were quantified before and during each infection using a highly multiplexed bead-based assay. The 981 log2 fold-change of each analyte is shown relative to baseline on day 6 and 8 post-challenge (c6 and c8, respectively) and at diagnosis. Analytes are ordered by log2 fold-change and are shown in 982 bold if they varied significantly during both second and third (compared to first) infection (adj p < 983 984 0.05 by linear regression with Benjamini-Hochberg correction for multiple testing). In (A-B) n = 10(first infection), 9 (second infection) and 6 (third infection). v1040 was excluded from RNAsequencing analysis in second infection because their baseline sample failed QC. In (C) n = 9 (first and second infection) and 5 (third infection). v1040 was excluded from plasma analysis because all 987 samples failed QC.



Figure 3. A single malaria episode attenuates T cell activation. (A) The percentage of activated CD38^{hi} effector or effector memory (EM) CD4⁺ T cells was analysed by flow cytometry at baseline 990 (grey dots) and 6 days after parasite clearance (coloured dots) during the VAC063C study. Data are 991 shown for volunteers undergoing their first, second or third infection of life (see Supplementary Data File 1 for gating strategies). (B) The percentage of activated CD38^{hi} effector or effector memory 993 CD4⁺ T cells at baseline (grey dots) and 6 days after parasite clearance (coloured dots) during the 994 995 VAC069 study. (C-D) RNA-sequencing was used to identify differentially expressed genes in flowsorted effector (effector memory) CD4⁺ T cells (T6 versus baseline) in first and third infection during 996 VAC063C (adj p < 0.05 and > 1.5 fold-change). The heatmaps in (C) show the log2 fold-change of 997 998 markers of T cell activation and exhaustion and the master transcription factors that shape T cell 999 fate (plus associated chemokine receptors). Non-significant genes are displayed with a log2 foldchange of zero and square brackets indicate that common gene names have been used. The 1000 1001 stacked circular bar chart in (D) shows the log2 fold-change of cytokines and cytotoxic effector molecules. (E) Correlation between log10 transformed T cell receptor beta variable (TRBV) gene 1002 frequency at baseline and 28 days post-challenge (convalescence) showing linear regression (black 1003 line) with 99% confidence intervals (grey shaded area). These data represent V gene usage across 1004 the entire T cell compartment and were obtained after one or two malaria episodes from volunteers 1005 1006 who were subsequently infected for a third time during VAC063C. In (A) n = 3 for first infection, n =1007 2 for second infection and n = 6 for third infection. In (B) n = 8 for first infection and n = 3 for second infection. In (C-D) n = 2 or 3 for first infection (T6 and baseline, respectively) and n = 6 for third 1008 1009 infection (v313 was excluded at T6 because this sample failed QC). In (E) n = 5 in first and second infection (v1065 convalescence samples failed QC). 1010



1011 Figure 4. Stem-like memory CD4⁺ T cells respond to re-challenge. Droplet-based single cell RNA-1012 sequencing was carried out during VAC063C on flow-sorted CD4⁺ T cells obtained at baseline and 1013 T6 from volunteers undergoing their first or third infection of life. (A) Data from all volunteers and time-points was concatenated for UMAP analysis and FlowSOM identified 13 discrete clusters of 1014 1015 CD4⁺ T cells across the dataset (each given a unique colour). (B) Heatmap showing the differential 1016 abundance of non-naive CD4⁺ T cell clusters at T6 (versus baseline) in first and third infection (FDR < 0.05). Note that non-significant clusters are shown with a log2 fold-change of zero and the 1017 1018 identification of non-naive clusters is shown in Extended Data Figure 5. (C) Dot plot showing differentially expressed signature genes in clusters 10 and 11 (adj p < 0.05). Square brackets 1019 indicate that common gene names have been used. (D) Trajectory inference with Slingshot revealed 1020 1021 clusters 10 and 11 as discrete non-overlapping endpoints of CD4+ T cell activation and differentiation (analysis was performed on concatenated data and CD38^{hi} cells were set as the 1022 1023 endpoint; Slingshot identified two possible non-overlapping routes). (E) Spearman correlation 1024 matrix showing shared V gene usage (TRAV and TRBV) across all CD4⁺ T cell clusters (the order of features was determined by unsupervised hierarchical clustering). (F) Gini plot showing the equality 1025 of V gene usage in each CD4⁺ T cell cluster in first and third infection; zero denotes perfect equality, 1026 1027 which indicates a diverse TCR repertoire. (G) Class-switched antibodies (IgG) recognising the 1028 malaria antigens MSP1 and AMA1 were quantified in serum by ELISA at baseline (grey dots) and one month after challenge (coloured dots). Samples were obtained from volunteers undergoing their 1029 1030 first, second or third infection during VAC063A, VAC063B and VAC063C, respectively. In (A-F) n = 3 for first and third infection whereas in (G) n = 10 for first and second infection and n = 6 in third 1031 1032 infection.



1033 Figure 5. Cytotoxic T cells are silenced for a minimum 8 months. Whole blood was preserved within 1034 30 minutes of blood draw at baseline, diagnosis and T6 during VAC063C as well as 45 days postchallenge (convalescence). Samples were stained with a T cell focussed antibody panel (see 1035 1036 Supplementary Table 3) and acquired on a Helios mass cytometer. After exclusion of normalisation beads and doublets we gated on CD45^{pos} CD3^{pos} T cells for downstream steps. (A) 1037 Data from all volunteers and time-points was concatenated for UMAP analysis and FlowSOM 1038 1039 identified 49 discrete clusters of T cells across the dataset (each given a unique colour). The major 1040 T cell subsets are labelled according to expression of lineage, memory and activation markers. (B) UMAP showing the T cell clusters that are differentially abundant at T6 (versus baseline) in first and 1041 1042 third infection (FDR < 0.05 and > 2 fold-change). Clusters that are not significant are shown in black. (C) The mean frequency of each T cell cluster that is differentially abundant in first and/or third 1043 infection is shown as a proportion of all CD45^{pos} CD3^{pos} T cells at T6. (**D**) Pies show the relative size 1044 1045 of each differentially abundant cluster. (E) Heatmap showing the normalised median expression 1046 values of all markers used for clustering in each of the differentially abundant T cell clusters. The order of features was determined by unsupervised hierarchical clustering. Colour codes to the left 1047 1048 of the heatmap indicate cluster identity and show whether clusters were significant in first infection or first and third infection. Note that no cluster was unique to third infection. In (A-E) n = 3 for first 1049 1050 infection and n = 6 for third infection.

medRxiv preprint doi https://doi.org/10.1101/2021.08.19.21262298; this version posted October 4, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in

		number of petuityabnormal able under a CC-BY 40 internat			max. parasite		
	114344412442112			mild	moderate	severe	density*
	Reuling et al.	95	70	40	16	14	
first infection	VAC063C	3	2	0	0	2	
	total	98	72	40	16	16	17,587

VAC063C re-challenge 8 2 2 0 0 36,630 *weighted mean number of parasites per ml blood







IL-12(p70) max. parasite density MMP9 CXCL8 TNFRII CXCL10 D-Dimer IL-1RA CXCL9 max. temperature cortisol anti-MSP1 IgG anti-AMA1 IgG act. regulatory T cells

act. effector (EM) CD4+ T cells act. cytotoxic T cells

1051 Figure 6. Controlling T cell activation protects host tissues. (A) A surrogate dataset from Reuling et al.49 was used to extract information on the frequency and severity of abnormal alanine 1052 aminotransferase (ALT) during a first-in-life infection (up to 6 days post-treatment). All volunteers 1053 1054 were infected with P. falciparum (3D7 or NF54) as part of a CHMI trial that used equivalent endpoints to our own study and in every case abnormal ALT was scored using the same adaptation of 1055 1056 the WHO adverse event grading system (see methods). Data from 95 volunteers in Reuling et al. (those enrolled in the EHMI-3, LSA-3, EHMI-8B, EHMI-9, ZonMw2, TIP5 and CHMI-trans1 studies) 1057 and the 3 first infection volunteers in VAC063C were pooled for analysis. (B) Frequency and severity 1058 1059 of liver injury; Barnard's test was used to statistically determine whether an abnormal ALT reading was more prevalent during a first-in-life infection compared to second or third infection (a p value 1060 1061 below 0.05 was considered significant). (C) Pearson correlation matrix showing the fold-change of 1062 differentially abundant plasma analytes, lymphocytes and haemoglobin during VAC063C. Fold-1063 change was calculated either at diagnosis or T6 (relative to baseline) according to when this was largest for each feature. Also included is maximum parasite density, maximum core temperature 1064 1065 (up to 48 hours post-treatment) and class-switched antibody titres (28 days after challenge). Finally, circulating ALT is shown at T6 together with the frequency of activated (CD38^h) effector (effector 1066 memory) CD4⁺ T cells, regulatory T cells and cytotoxic T cells (defined as granzyme B^{pos}). All data 1067 were log2 transformed and the order of features was determined by unsupervised hierarchical 1068 clustering. (D) Peripheral blood mononuclear cells (PBMC) were isolated during VAC063C from 1069 volunteers undergoing their first infection of life, re-stimulated in vitro with PMA/ionomycin and co-1070 1071 cultured with HepG2 cells for 24 hours. Cytotoxicity was measured by the release of lactate dehydrogenase (LDH). Experiments were performed using baseline and T6 samples, and data are 1072 1073 shown as baseline subtracted values (i.e. absorbance at T6 minus absorbance at baseline). Curves were fit using a cubic polynomial function (the shaded areas represent 95% confidence intervals). 1074 1075 Note that LDH release was specific to HepG2 cells as shown in Extended Data Figure 7B and all 1076 assays were run in duplicate. In (A-B) n = 98 for first infection and n = 8 for re-challenge (2 second infection and 6 third infection). In (C) n = 10 (3 first infection, 2 second infection and 5 third infection) 1077 1078 and in (D) n = 3 (first infection only).



I = intercept SE (I) = SE intercept $\beta_1 = order of infection$ $SE (\beta_1) = SE order of infection$ logL = log likelihood AIC = akaike information criterion GLM = generalised linear model

top no. = risk of *severe* malaria bottom no. = risk of complicated malaria

1079 **Extended Data Figure 1.** The risk of severe malaria decreases exponentially with exposure. Data 1080 were extracted from Gonçalves *et al.*¹ to examine the frequency of severe or complicated malaria during the first 14 infections of life in infants living in a hyperendemic setting. (A) We first plotted 1081 the total number of cases of malaria (including mild or uncomplicated episodes) and used least 1082 squares regression to impute missing values (for infection number 9, 11, 12 and 13). Imputed values 1083 are shown as a grey triangle whereas filled circles indicate data as reported by Gonçalves et al. The 1084 total number of children experiencing at least 1 episode of malaria was n = 715. (B) We then plotted 1085 1086 the incidence of severe or complicated malaria at each order of infection (as shown in Figure 1A-B) and performed maximum likelihood estimation to select the best model fit for these data; log 1087 likelihood (logL) and AIC both show that an exponential decay in risk provides the best fit. 1088



extended data figure 2

1089 **Extended Data Figure 2.** Systemic inflammation is not attenuated upon re-challenge. (A) 1090 Proportion of eosinophils, lymphocytes, monocytes and neutrophils in whole blood at baseline and diagnosis. The mean frequency is shown for each time-point and infection. Note that the loss of 1091 lymphocytes at diagnosis is comparable in all three infections. (B) RNA-sequencing was used to 1092 identify differentially expressed genes in whole blood at diagnosis (versus baseline) (adj p < 0.051093 and > 1.5 fold-change). Volcano plots show all differentially expressed genes (coloured dots) and 1094 1095 the dashed lines represent the significance/fold-change cut-offs (genes that are not significant are shown in grey). The top 10 differentially expressed genes (lowest adj p) in each infection are labelled 1096 (first, second and third infection were analysed independently). (C) The log2 fold-change of 1097 1098 signature genes associated with interferon signaling and type I inflammation are shown at diagnosis (versus baseline) in first, second and third infection. Square brackets indicate that common gene 1099 names have been used. (D) Linear regression was used to identify plasma analytes that vary 1100 1101 significantly between infections at diagnosis. Multiple test correction was performed using the Benjamini-Hochberg method and analytes that were significant in both second and third infection 1102 (versus first infection) are displayed (adj p < 0.05). Box (median and IQR) and whisker (1.5x upper 1103 or lower IQR) plots are shown with outliers as dots. In (A) n = 10 (first and second infection) and n 1104 = 6 (third infection). In (B-C) n = 10 (first infection), n = 9 (second infection) and n = 6 (third infection). 1105 1106 v1040 was excluded from RNA-sequencing analysis in second infection because their baseline 1107 sample failed QC. In (D) n = 9 (first and second infection) and n = 5 (third infection). v1040 was 1108 excluded from plasma analysis because all samples failed QC.



extended data figure 3

1109 **Extended Data Figure 3.** The acute phase response is followed by proliferation in first infection. 1110 (A) Mixed-effects models and linear regression were used to identify plasma analytes that vary significantly at diagnosis and/or T6 across the entire VAC063C dataset (all volunteers regardless of 1111 infection number). Kenward Roger approximation was used to calculate p values and multiple test 1112 correction was performed using the Benjamini-Hochberg method; significance (adj p < 0.05) is 1113 indicated by coloured box plots (purple at diagnosis and turguoise at T6). Box (median and IQR) 1114 and whisker (1.5x upper or lower IQR) plots are shown with outliers as dots. (B) RNA-sequencing 1115 was used to identify differentially expressed genes in whole blood at diagnosis and T6 during first 1116 infection (versus baseline) (adj p < 0.05 and > 1.5 fold-change). Volcano plots show all differentially 1117 1118 expressed genes (coloured dots) and the dashed lines represent the significance/fold-change cut-1119 offs (genes that are not significant are shown in grey). The top 10 differentially expressed genes (lowest adi p) at each time-point are labelled. (C) Differentially expressed genes at T6 and diagnosis 1120 1121 were combined for GO analysis of first infection and ClueGO was used to construct a merged functional gene ontology network. Each node represents a GO term and nodes are coloured 1122 according to whether their associated genes were majoritively (> 60%) derived from the diagnosis 1123 or T6 time-point. GO terms that were shared between time-points are coloured grey (see methods). 1124 Four leading GO terms (each from a unique functional group) are labelled for each time-point. (D) 1125 1126 Heatmaps show differentially expressed genes at diagnosis and T6 (versus baseline) during first 1127 and third infection (adj p < 0.05 and > 1.5 fold-change). The log2 fold-change of key genes associated with each phase of the cell cycle is shown. Non-significant genes are displayed with a 1128 log2 fold-change of zero. In (A) n = 10 (3 first infection, 2 second infection and 5 third infection). 1129 v1040 was excluded from plasma analysis because all samples failed QC. In (B - C) n = 3 (first 1130 infection only) and in (D) n = 3 (first infection) and n = 6 (third infection). 1131



1132 **Extended Data Figure 4.** Re-challenge does not induce CD4⁺ T cell anergy, exhaustion or clonal deletion. (A) Healthy malaria-naive adults were infected up to two times with *P. vivax* (clone PvW1) 1133 by direct blood challenge during the VAC069 study. Parasite growth curves are shown for first and 1134 second infection; each line represents a volunteer and lines are colour-coded by infection number. 1135 Parasite density was measured in peripheral blood by gPCR every 12 - 24 hours. The grey box 1136 represents the lower limit of quantification (20 parasites ml⁻¹) and the treatment threshold of 10,000 1137 parasites ml⁻¹ is denoted by the black line. (B) RNA-sequencing was used to analyse transcriptional 1138 1139 regulation of fatty acid β -oxidation, the tricarboxylic acid (TCA) cycle and oxidative phosphorylation (oxphos) in flow-sorted effector (effector memory) CD4⁺ T cells during first infection with P. 1140 1141 falciparum (VAC063C). The circular bar charts show the log2 fold-change of each major enzyme 1142 involved in fatty acid β -oxidation and the TCA cycle clockwise in reaction order 6 days after parasite clearance (T6 versus baseline). The vertical bar chart shows the proportion of oxphos enzymatic 1143 1144 subunits that are transcriptionally upregulated at T6 - all subunits required to form complex I to IV 1145 in the electron transport chain and ATP synthase are shown. The key molecules that connect these 1146 metabolic pathways are labelled. (C - D) RNA-sequencing was used to identify differentially expressed genes in flow-sorted effector (effector memory) CD4⁺ T cells in first and third infection 1147 during VAC063C (T6 versus baseline) (adj p < 0.05 and > 1.5 fold-change). Heatmaps show the 1148 1149 log2 fold-change of (C) hallmark genes indicating T cell anergy or quiescence and (D) markers of activation-induced cell death. (E) Flow-sorted regulatory T cells (CD4^{pos} CD25^{hi} CD127^{neg}) were also 1150 1151 analysed by RNA-sequencing to identify differentially expressed genes relating to their activation and suppressor function (T6 versus baseline) in first and third infection (adj p < 0.05 and > 1.5 fold-1152 1153 change). (F) RNA-sequencing was used to identify differentially expressed genes in whole blood at diagnosis and T6 (relative to baseline) in first and third infection during VAC063C (adj p < 0.05 and 1154 > 1.5 fold-change). Signature genes associated with T cell activation, T_H1 differentiation and 1155 1156 cytokine production are shown. In (A) n = 8 for first infection and n = 3 for second infection. In (B -E) n = 2 or 3 for first infection (T6 and baseline, respectively) and n = 6 for third infection (v313 was 1157 1158 excluded at T6 because this sample failed QC). In (F) n = 3 for first infection and n = 6 for third infection. In (C-F) non-significant genes are displayed with a log2 fold-change of zero. 1159



В	aligned reads	mapped to genome	mapped to transcriptome	single cells	median reads per cell	median genes per cell	total genes detected	median UMIs per cell
first infection	2.021x10 ⁹	90%	75%	23,227	77,682	954	20,848	3,011
third infection	2.026x10 ⁹	87%	73%	26,891	67,914	922	21,042	2,676







[TCF1]





5

F

UMAP 2

5

0

-5

[BLIMP1]

1160 **Extended Data Figure 5.** Activated CD4⁺ T cells bifurcate along the TCF1/BLIMP1 axis. (A) 1161 VAC063C single cell RNA-sequencing workflow: each sample of flow-sorted CD4⁺ T cells was barcoded using TotalSeq-C oligo-tagged antibodies; samples from all volunteers and time-points 1162 were pooled (separately for first and third infection); and pooled samples were superloaded onto a 1163 10X Chromium Controller (we aimed to capture 30,000 singlets per pool). Gel beads in emulsion 1164 1165 (GEMs) encapsulating a single cell (or doublets) were then generated and from each GEM three libraries were produced: i) the cell surface barcode, ii) 5' gene expression and iii) T cell receptor 1166 (after amplification of the V(D)J regions). Libraries were pooled at the specified ratios and 1167 1168 sequenced. Finally, we used PCA-based clustering to debarcode all samples and remove doublets (see methods). (B) Cell Ranger was used to align 5' gene expression and V(D)J sequencing reads 1169 1170 (independently for first and third infection). Shown is the output of Cell Ranger after removing doublets and performing QC. (C - F) Data from all volunteers and time-points was concatenated for 1171 UMAP analysis. The expression intensity of markers for memory (C), activation (D) and T_{FH} 1172 differentiation (E) are shown across the UMAP. The blue line represents the split between naive and 1173 1174 memory cells whereas the green line represents the split between memory and activated cells. In (F) the expression intensity of the master transcription factors associated with terminal 1175 1176 differentiation (BLIMP1) versus the maintenance of stem-like properties (TCF1) are shown. In all 1177 cases, each UMAP is equivalent to those shown in Figure 4 (for cross-reference) and square brackets indicate that common gene names have been used. (G) Proposed model of T cell 1178 1179 activation during a first-in-life malaria episode. The maintenance of stem-like T cells is essential for long-lived memory; this requires sustained expression of TCF1 to repress BLIMP1 and prevent the 1180 1181 terminal differentiation of short-lived effector cells. In (A - F) n = 3 for first and third infection.





0 2 4

first & third infection

Α

extended data figure 6

1182 Extended Data Figure 6. Cytotoxic T cells are silenced after a single malaria episode. (A) Stacked bar chart showing the frequency of activated (CD38^{hi} Bcl2^{lo}) T cells at baseline, diagnosis and T6 1183 as well as 45 days post-challenge (convalescence) during VAC063C. Each bar represents one 1184 volunteer and individual T cell clusters are colour-coded to match Figure 5 (note that only 1185 differentially abundant clusters are included). The major CD4⁺ and CD8⁺ T cell clusters with 1186 cytotoxic features are highlighted with a black border in the key to the left of the plot. (B) Differential 1187 marker expression through time in each major T cell subset. First, T cell clusters belonging to the 1188 same lineage were merged and then CD4⁺ and CD8⁺ T cells were split into naive, effector, effector 1189 memory (EM) and central memory (CM) subsets. Next, linear models were used to independently 1190 1191 assess differential marker expression in each subset at each time-point (relative to baseline); a shift in median expression of at least 10% and an FDR < 0.05 were required for significance. Shown are 1192 1193 all subset/marker pairs that were called as significant at T6 and data are presented as row-wise zscore marker intensities. Colour codes to the left of the heatmap indicate whether markers were 1194 differentially expressed during first infection, third infection or both infections. 1195



1196 **Extended Data Figure 7.** Raised serum transaminases are unique to first infection. (A) Blood 1197 chemistry measured the concentration of alanine aminotransferase (ALT) through infection and convalescence in the VAC063C study (baseline, 6 days post-challenge (c6), diagnosis, 6 - 18 days 1198 after drug treatment (T6 - T18) and 45 or 90 days post-challenge (convalescence 1 and 2, 1199 respectively)). Each line represents one volunteer and lines are colour-coded by infection number. 1200 (B - C) Peripheral blood mononuclear cells (PBMC) were isolated during VAC063C from volunteers 1201 undergoing their first infection of life, re-stimulated in vitro with PMA/ionomycin and co-cultured 1202 with HepG2 cells for 24 hours. Cytotoxicity was measured by the release of lactate dehydrogenase 1203 1204 (LDH). (B) PBMC were co-cultured with or without HepG2 cells to determine their relative 1205 contribution to LDH release. (C) PBMC isolated at baseline (grey) or T6 (turquoise) were co-cultured with HepG2 cells at different effector to target cell ratios. Each plot represents the raw data from 1206 one volunteer, the black line shows the mean absorbance in control wells (HepG2 cells only) and 1207 volunteers are ordered according to their frequency of activated cytotoxic CD4⁺ T cells at T6 1208 (highest to lowest (top to bottom) - see Extended Data Figure 6A). In (A) n = 11 (3 first infection, 2 1209 1210 second infection and 6 third infection). In (B) a representative experiment is shown and in (C) n = 31211 (first infection only). Note that all cytotoxicity assays were performed in duplicate or triplicate and curves were fit using a cubic polynomial function (the shaded areas represent 95% confidence 1212 1213 intervals).

Supplementary Data File 1. Gating strategy for sorting CD4⁺ T cell subsets during VAC063C. CD4⁺ T cells were sorted *ex vivo* (within 2 hours of blood draw) into TRIzol for downstream RNAsequencing - cells with a naive, effector (effector memory) or regulatory phenotype were sorted as shown at baseline and T6. Note that we did not use CD38 for sorting but subsequently used this marker to assess the level of activation within each subset at both time-points.

Supplementary Data File 2. T cell receptor V gene usage in CD4⁺ T cell clusters. Droplet-based single cell RNA-sequencing was carried out during VAC063C on flow-sorted CD4⁺ T cells obtained at baseline and T6 from volunteers undergoing their first or third infection of life. Data from all volunteers and time-points was concatenated to examine V gene usage (in TCR α and TCR β chains) in each cluster. Note that clusters are colour-coded exactly as shown in Figure 4 and it is clusters 10 (self-renewal) and 11 (terminal differentiation) that are activated (CD38^{hi}) by malaria (n = 3 in first and third infection).

Supplementary Data File 3. T cell cluster phenotypes by mass cytometry. Heatmap showing the normalised median expression values of all markers used for clustering in each of the 49 T cell clusters. Names were assigned manually using activation, lineage and memory markers to broadly categorise each T cell cluster; when more than one cluster was placed into the same category (e.g. activated EM CD4) clusters were given an accessory label to highlight their unique phenotype or property (e.g. T-bet^{lo} Eomes^{neg}). The order of features was determined by unsupervised hierarchical clustering.

Supplementary Data File 4. T cell cluster frequencies by mass cytometry. Each cluster is shown as a proportion of all CD45^{pos} CD3^{pos} T cells at each time-point. Clusters are shown in the same order as Supplementary Data File 3 (left to right and top to bottom). Box (median and IQR) and whisker (1.5x upper or lower IQR) plots are shown (with outliers as dots) and significance (FDR < 0.05 and > 2 fold-change) is indicated by colour (dark blue for first infection and bright blue for third). In all plots n = 3 for first infection and n = 6 for third infection.

Supplementary Table 1. Demographics of volunteers infected and re-challenged with *Plasmodium falciparium* (3D7) during VAC063A-C; includes genetic and non-genetic variables known to influence human immune variation *in vitro*. Also shown is the parasite multiplication rate (detailed methodology for the modelling of qPCR data can be found in Minassian *et al.*⁸¹) and maximum circulating parasite density. There were no statistically significant differences between first, second and third infection (Kruskal-Wallis test, p = 0.3240 for parasite multiplication rate and p = 0.5975 for maximum parasite density).

Supplementary Table 2. Signature genes associated with terminal differentiation and self-renewal 1246 in CD4⁺ T cells. Tab one lists the signature genes for each of the 13 clusters of CD4⁺ T cells identified 1247 1248 by single cell RNA-sequencing during VAC063C. Note that clusters are colour-coded exactly as 1249 shown in Figure 4 and for each cluster genes are ordered by adj p value. Tab two lists the 1250 differentially expressed genes identified by direct pairwise comparison between activated (CD38^{hi}) 1251 clusters 10 and 11. Genes that are more highly expressed in cluster 10 have a positive log2 foldchange. Note that differential gene expression analysis was performed across the entire dataset 1252 1253 (i.e. data from all volunteers and time-points was concatenated) (n = 3 in first and third infection).

- Supplementary Table 3. Mass cytometry antibody panel for T cell fate and function in VAC063C;
 includes information on antibody clone, source and heavy metal conjugate.
- Supplementary Table 4. Coefficient and p values underlying the Pearson correlation matrix shownin Figure 6C.

1258 References

1261

1264

1267

1270

1273

1277

1280

1283

1289

1292

1296

1299

1302

1305

- 12591.Goncalves, B.P. et al. Parasite burden and severity of malaria in Tanzanian children. N Engl1260J Med **370**, 1799-1808 (2014).
- Moxon, C.A., Gibbins, M.P., McGuinness, D., Milner, D.A., Jr. & Marti, M. New Insights into Malaria Pathogenesis. *Annu Rev Pathol* **15**, 315-343 (2020).
- 12653.Snow, R.W. & Marsh, K. New insights into the epidemiology of malaria relevant for disease1266control. Br Med Bull 54, 293-309 (1998).
- Snow, R.W. *et al.* Risk of severe malaria among African infants: direct evidence of clinical protection during early infancy. *J Infect Dis* **177**, 819-822 (1998).
- 1271 5. Doolan, D.L., Dobano, C. & Baird, J.K. Acquired immunity to malaria. *Clin Microbiol Rev* 22, 13-36 (2009).
- 12746.McLean, F.E. et al. Detection of naturally acquired strain-transcending antibodies against1275rosetting Plasmodium falciparum strains in humans. bioRxiv, 2024.2001.2011.5752521276(2024).
- 1278 7. Tuju, J. *et al.* Antigenic cartography of immune responses to Plasmodium falciparum 1279 erythrocyte membrane protein 1 (PfEMP1). *PLoS Pathog* **15**, e1007870 (2019).
- Cumnock, K. *et al.* Host Energy Source Is Important for Disease Tolerance to Malaria. *Curr Biol* 28, 1635-1642 e1633 (2018).
- Ramos, S. *et al.* Renal control of disease tolerance to malaria. *Proc Natl Acad Sci U S A* **116**, 5681-5686 (2019).
 1286
- 1287 10. Wang, A. *et al.* Glucose metabolism mediates disease tolerance in cerebral malaria. *Proc* 1288 *Natl Acad Sci U S A* **115**, 11042-11047 (2018).
- Medzhitov, R., Schneider, D.S. & Soares, M.P. Disease tolerance as a defense strategy.
 Science 335, 936-941 (2012).
- 129312.Ademolue, T.W., Aniweh, Y., Kusi, K.A. & Awandare, G.A. Patterns of inflammatory1294responses and parasite tolerance vary with malaria transmission intensity. *Malar J* 16, 1451295(2017).
- 1297 13. Nahrendorf, W., Ivens, A. & Spence, P.J. Inducible mechanisms of disease tolerance provide 1298 an alternative strategy of acquired immunity to malaria. *Elife* **10** (2021).
- 130014.Milne, K. *et al.* Mapping immune variation and var gene switching in naive hosts infected1301with Plasmodium falciparum. *Elife* **10** (2021).
- 130315.Stanisic, D.I., McCarthy, J.S. & Good, M.F. Controlled Human Malaria Infection:1304Applications, Advances, and Challenges. Infect Immun 86 (2018).
- 130616.Baird, J.K. *et al.* Adult Javanese migrants to Indonesian Papua at high risk of severe disease1307caused by malaria. *Epidemiol Infect* **131**, 791-797 (2003).
- 130917.Baird, J.K. *et al.* Age-dependent susceptibility to severe disease with primary exposure to1310Plasmodium falciparum. J Infect Dis **178**, 592-595 (1998).

1312 18. Draper, S.J. *et al.* Malaria Vaccines: Recent Advances and New Horizons. *Cell Host Microbe*1313 24, 43-56 (2018).
1314

1311

1320

1323

1326

1329

1333

1337

1340

1344

1348

1351

1354

1357

1360

- 131519.Duncan, C.J. & Draper, S.J. Controlled human blood stage malaria infection: current status1316and potential applications. Am J Trop Med Hyg 86, 561-565 (2012).
- 1317
 1318 20. Cheng, Q. *et al.* Measurement of Plasmodium falciparum growth rates in vivo: a test of malaria vaccines. *Am J Trop Med Hyg* **57**, 495-500 (1997).
- 1321 21. Spence, P.J. *et al.* Vector transmission regulates immune control of Plasmodium virulence.
 1322 Nature 498, 228-231 (2013).
- 132422.Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for1325RNA-seq data with DESeq2. Genome Biol **15**, 550 (2014).
- 132723.Feintuch, C.M. *et al.* Activated Neutrophils Are Associated with Pediatric Cerebral Malaria1328Vasculopathy in Malawian Children. *mBio* **7**, e01300-01315 (2016).
- Loughland, J.R. *et al.* Transcriptional profiling and immunophenotyping show sustained
 activation of blood monocytes in subpatent Plasmodium falciparum infection. *Clin Transl Immunology* 9, e1144 (2020).
- Antonelli, L.R. *et al.* The CD14+CD16+ inflammatory monocyte subset displays increased
 mitochondrial activity and effector function during acute Plasmodium vivax malaria. *PLoS Pathog* 10, e1004393 (2014).
- 133826.Rocha, B.C. *et al.* Type I Interferon Transcriptional Signature in Neutrophils and Low-Density1339Granulocytes Are Associated with Tissue Damage in Malaria. *Cell Rep* **13**, 2829-2841 (2015).
- 134127.Hviid, L. *et al.* Rapid reemergence of T cells into peripheral circulation following treatment1342of severe and uncomplicated Plasmodium falciparum malaria. *Infect Immun* 65, 4090-40931343(1997).
- 134528.Mandala, W.L., Ward, S., Taylor, T.E. & Wassmer, S.C. Characterization of Lymphocyte1346Subsets in Lymph Node and Spleen Sections in Fatal Pediatric Malaria. Pathogens 111347(2022).
- Bach, F.A. *et al.* A systematic analysis of the human immune response to Plasmodium vivax.
 J Clin Invest 133 (2023).
- 135230.Sharp, P.M., Plenderleith, L.J. & Hahn, B.H. Ape Origins of Human Malaria. Annu Rev1353Microbiol 74, 39-63 (2020).
- 135531.Carpio, V.H. *et al.* T Helper Plasticity Is Orchestrated by STAT3, Bcl6, and Blimp-1 Balancing1356Pathology and Protection in Malaria. *iScience* **23**, 101310 (2020).
- 135832.Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet1359detection for single cell genomics. *Genome Biol* **19**, 224 (2018).
- 136133.Lonnberg, T. *et al.* Single-cell RNA-seq and computational analysis using temporal mixture1362modelling resolves Th1/Tfh fate bifurcation in malaria. Sci Immunol 2 (2017).
- 136434.Cao, W. et al. TRIB2 safeguards naive T cell homeostasis during aging. Cell Rep 42, 1121951365(2023).

- 1367 35. Galletti, G. et al. Two subsets of stem-like CD8(+) memory T cell progenitors with distinct 1368 fate commitments in humans. Nat Immunol 21, 1552-1562 (2020).
- 1370 36. Ko, C.J. et al. The E3 ubiquitin ligase Peli1 regulates the metabolic actions of mTORC1 to suppress antitumor T cell responses. EMBO J 40, e104532 (2021). 1371
- 1373 Choi, Y.S. et al. LEF-1 and TCF-1 orchestrate T(FH) differentiation by regulating 37. 1374 differentiation circuits upstream of the transcriptional repressor Bcl6. Nat Immunol 16, 980-1375 990 (2015).
- 38. Zhao, X., Shan, Q. & Xue, H.H. TCF1 in T cell immunity: a broadened frontier. Nat Rev 1377 1378 Immunol 22, 147-157 (2022).
- 1380 39. Mathewson, N.D. et al. Inhibitory CD161 receptor identified in glioma-infiltrating T cells by 1381 single-cell analysis. Cell 184, 1281-1298 e1226 (2021).
- 1383 40. Ng, S.S. et al. The NK cell granule protein NKG7 regulates cytotoxic granule exocytosis and 1384 inflammation. Nat Immunol 21, 1205-1218 (2020).
- 1386 41. Schober, K. et al. Reverse TCR repertoire evolution toward dominant low-affinity clones 1387 during chronic CMV infection. Nat Immunol 21, 434-441 (2020). 1388
- 1389 42. Appay, V. et al. Characterization of CD4(+) CTLs ex vivo. J Immunol 168, 5954-5958 (2002).
- 1391 43. Mattoo, H. et al. Clonal expansion of CD4(+) cytotoxic T lymphocytes in patients with IgG4-1392 related disease. J Allergy Clin Immunol 138, 825-838 (2016).
- 1394 44. Oh, D.Y. et al. Intratumoral CD4(+) T Cells Mediate Anti-tumor Cytotoxicity in Human 1395 Bladder Cancer. Cell 181, 1612-1625 e1613 (2020).
- 1397 45. Rahil, Z. et al. Landscape of coordinated immune responses to H1N1 challenge in humans. 1398 J Clin Invest 130, 5800-5816 (2020).
- 1400 46. Napolitani, G. et al. Clonal analysis of Salmonella-specific effector T cells reveals serovar-1401 specific and cross-reactive T cell responses. Nat Immunol 19, 742-754 (2018).
- 1403 47. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for 1404 differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140 1405 (2010).
- 1407 48. Chughlay, M.F. et al. Liver Enzyme Elevations in Plasmodium falciparum Volunteer Infection 1408 Studies: Findings and Recommendations. Am J Trop Med Hyg 103, 378-393 (2020).
- Reuling, I.J. et al. Liver Injury in Uncomplicated Malaria is an Overlooked Phenomenon: An 49. 1410 1411 Observational Study. EBioMedicine 36, 131-139 (2018).
- 50. 1413 Woodford, J., Shanks, G.D., Griffin, P., Chalon, S. & McCarthy, J.S. The Dynamics of Liver 1414 Function Test Abnormalities after Malaria Infection: A Retrospective Observational Study. 1415 *Am J Trop Med Hyg* **98**, 1113-1119 (2018).
- 1417 51. Dudek, M. et al. Auto-aggressive CXCR6(+) CD8 T cells cause liver immune pathology in 1418 NASH. Nature 592, 444-449 (2021).
- 1419

1416

1366

1369

1372

1376

1379

1382

1385

1390

1393

1396

1399

1402

1406

1409

- 1420 52. Malaria Genomic Epidemiology, N. Insights into malaria susceptibility using genome-wide 1421 data on 17,000 individuals from Africa, Asia and Oceania. Nat Commun 10, 5732 (2019).
- 1423 53. Patin, E. et al. Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. Nat Immunol 19, 302-314 (2018). 1424
- 1426 54. Raue, H.P., Beadling, C., Haun, J. & Slifka, M.K. Cytokine-mediated programmed 1427 proliferation of virus-specific CD8(+) memory T cells. Immunity 38, 131-139 (2013).
- 1429 55. Weng, N.P., Liu, K., Catalfamo, M., Li, Y. & Henkart, P.A. IL-15 is a growth factor and an 1430 activator of CD8 memory T cells. Ann N Y Acad Sci 975, 46-56 (2002).
- 1432 Greenwood, B.M. Possible role of a B-cell mitogen in hypergammaglobulinaemia in malaria 56. 1433 and trypanosomiasis. Lancet 1, 435-436 (1974).
- 1435 57. Mourao, L.C., Cardoso-Oliveira, G.P. & Braga, E.M. Autoantibodies and Malaria: Where We 1436 Stand? Insights Into Pathogenesis and Protection. Front Cell Infect Microbiol 10, 262 (2020).
- 1438 58. Welsh, R.M. & Selin, L.K. No one is naive: the significance of heterologous T-cell immunity. 1439 Nat Rev Immunol 2, 417-426 (2002).
- 1441 59. Bernabeu, M. et al. Binding Heterogeneity of Plasmodium falciparum to Engineered 3D Brain 1442 Microvessels Is Mediated by EPCR and ICAM-1. mBio 10 (2019).
- 1444 60. Riggle, B.A. et al. CD8+ T cells target cerebrovasculature in children with cerebral malaria. 1445 J Clin Invest 130, 1128-1138 (2020).
- Wassmer, S.C., de Koning-Ward, T.F., Grau, G.E.R. & Pai, S. Unravelling mysteries at the 1447 61. 1448 perivascular space: a new rationale for cerebral malaria pathogenesis. Trends Parasitol 40, 1449 28-44 (2024).
- 1451 62. Crispe, I.N., Dao, T., Klugewitz, K., Mehal, W.Z. & Metz, D.P. The liver as a site of T-cell 1452 apoptosis: graveyard, or killing field? Immunol Rev 174, 47-62 (2000).
- 1454 63. Noe, A. et al. Deep Immune Phenotyping and Single-Cell Transcriptomics Allow 1455 Identification of Circulating TRM-Like Cells Which Correlate With Liver-Stage Immunity and 1456 Vaccine-Induced Protection From Malaria. Front Immunol 13, 795463 (2022).
- 1458 64. Edwards, C.L. et al. IL-10-producing Th1 cells possess a distinct molecular signature in 1459 malaria. J Clin Invest 133 (2023).
- 1461 65. Jagannathan, P. et al. IFNgamma/IL-10 co-producing cells dominate the CD4 response to 1462 malaria in highly exposed children. PLoS Pathog 10, e1003864 (2014).
- 1464 66. Walther, M. et al. Distinct roles for FOXP3+ and FOXP3- CD4 T cells in regulating cellular 1465 immunity to uncomplicated and severe Plasmodium falciparum malaria. PLoS Pathog 5, 1466 e1000364 (2009).
- Newton, P.N. et al. A comparison of the in vivo kinetics of Plasmodium falciparum ring-1468 67. 1469 infected erythrocyte surface antigen-positive and -negative erythrocytes. Blood 98, 450-457 1470 (2001). 1471
- 1472 68. Heesters, B.A. et al. Characterization of human FDCs reveals regulation of T cells and 1473 antigen presentation to B cells. J Exp Med 218 (2021).
- 1474

1422

1425

1428

1431

1434

1437

1440

1443

1446

1450

1453

1457

1460

1463

69. Clark, H.C. & Tomlinson, W.J. The pathologic anatomy of malaria. *Malariology* **2**, 874-903

1475

1476

1480

1484

1487

1490

1493

1497

1501

1505

1511

1514

1525

(1949).

- 1477
 1478 70. Osii, R.S., Otto, T.D., Garside, P., Ndungu, F.M. & Brewer, J.M. The Impact of Malaria
 1479 Parasites on Dendritic Cell-T Cell Interaction. *Front Immunol* **11**, 1597 (2020).
- 148171.Snell, L.M. *et al.* CD8+ T Cell Priming in Established Chronic Viral Infection Preferentially1482Directs Differentiation of Memory-like Cells for Sustained Immunity. *Immunity* **49**, 678-6941483e675 (2018).
- 148572.Zou, D. et al. CD4+ T cell immunity is dependent on an intrinsic stem-like program. Nat1486Immunol 25, 66-76 (2024).
- 148873.Nish, S.A. *et al.* CD4+ T cell effector commitment coupled to self-renewal by asymmetric1489cell divisions. J Exp Med **214**, 39-47 (2017).
- 149174.Xu, L. *et al.* The transcription factor TCF-1 initiates the differentiation of TFH cells during1492acute viral infection. Nat Immunol **16**, 991-999 (2015).
- Kimingi, H.W. *et al.* Breadth of Antibodies to Plasmodium falciparum Variant Surface
 Antigens Is Associated With Immunity in a Controlled Human Malaria Infection Study. *Front Immunol* **13**, 894770 (2022).
- 149876.Osier, F.H. *et al.* Breadth and magnitude of antibody responses to multiple Plasmodium1499falciparum merozoite antigens are associated with protection from clinical malaria. Infect1500Immun **76**, 2240-2248 (2008).
- 150277.Rono, J. et al. Breadth of anti-merozoite antibody responses is associated with the genetic1503diversity of asymptomatic Plasmodium falciparum infections and protection against clinical1504malaria. Clin Infect Dis 57, 1409-1416 (2013).
- Furtado, R. *et al.* Cytolytic memory CD4+ T cell clonotypes are expanded during
 Plasmodium falciparum infection. *bioRxiv*, 2021.2007.2021.453277 (2021).
- 150979.de Jong, S.E. *et al.* Systems analysis and controlled malaria infection in Europeans and1510Africans elucidate naturally acquired immunity. Nat Immunol 22, 654-665 (2021).
- 151280.Soon, M.S.F. *et al.* Transcriptome dynamics of CD4(+) T cells during malaria maps gradual1513transit from effector to memory. *Nat Immunol* **21**, 1597-1610 (2020).
- 1515 81. Minassian, A.M. *et al.* Reduced blood-stage malaria growth and immune correlates in humans following RH5 vaccination. *Med (N Y)* 2, 701-719 (2021).
 1517
- 1518 82. Salkeld, J. *et al.* Repeat controlled human malaria infection of healthy UK adults with blood1519 stage Plasmodium falciparum: Safety and parasite growth dynamics. *Front Immunol* 13, 984323 (2022).
 1521
- 152283.Payne, R.O. et al. Demonstration of the Blood-Stage Plasmodium falciparum Controlled1523Human Malaria Infection Model to Assess Efficacy of the P. falciparum Apical Membrane1524Antigen 1 Vaccine, FMP2.1/AS01. J Infect Dis **213**, 1743-1751 (2016).
- Minassian, A.M. *et al.* Controlled human malaria infection with a clone of Plasmodium vivax
 with high-quality genome assembly. *JCI Insight* 6 (2021).

- 1529 85. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
- 153286.Harrington, C.A. *et al.* RNA-Seq of human whole blood: Evaluation of globin RNA depletion1533on Ribo-Zero library method. Sci Rep **10**, 6271 (2020).

1531

1534

1537

1539

1547

1550

1556

1560

1567

1570

1573

- 153587.Shin, H. *et al.* Variation in RNA-Seq transcriptome profiles of peripheral whole blood from1536healthy individuals with and without globin depletion. *PLoS One* **9**, e91041 (2014).
- 1538 88. Wickam, H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York* (2016).
- Mizuno, H., Nakanishi, Y., Ishii, N., Sarai, A. & Kitada, K. A signature-based method for
 indexing cell cycle phase distribution from microarray profiles. *BMC Genomics* **10**, 137 (2009).
- 154490.Pena-Diaz, J. et al. Transcription profiling during the cell cycle shows that a subset of1545Polycomb-targeted genes is upregulated during DNA replication. Nucleic Acids Res 41,15462846-2856 (2013).
- 1548 91. Dominguez, D. *et al.* A high-resolution transcriptome map of cell cycle reveals novel connections between periodic genes and cancer. *Cell Res* **26**, 946-962 (2016).
- 1551 92. Giotti, B., Joshi, A. & Freeman, T.C. Meta-analysis reveals conserved cell cycle 1552 transcriptional network across multiple human cell types. *BMC Genomics* **18**, 30 (2017). 1553
- 155493.Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene1555ontology and pathway annotation networks. *Bioinformatics* **25**, 1091-1093 (2009).
- 155794.Mlecnik, B. *et al.* Functional network pipeline reveals genetic determinants associated with1558in situ lymphocyte proliferation and survival of cancer patients. Sci Transl Med 6, 228ra2371559(2014).
- 1561 95. Chomczynski, P. & Sacchi, N. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc* 1, 581-585 (2006).
 1564
- 156596.Mamedov, I.Z. *et al.* Preparing unbiased T-cell receptor and antibody cDNA libraries for the
deep next generation sequencing profiling. *Front Immunol* **4**, 456 (2013).
- 1568 97. Shugay, M. *et al.* Towards error-free profiling of immune repertoires. *Nat Methods* **11**, 653-1569 655 (2014).
- 157198.Bolotin, D.A. *et al.* MiXCR: software for comprehensive adaptive immunity profiling. Nat1572Methods 12, 380-381 (2015).
- 1574 99. Lefranc, M.P. *et al.* IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res*1575 27, 209-212 (1999).
 1576
- 1577 100. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony.
 1578 Nat Methods 16, 1289-1296 (2019).
- 1580 101. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell 1581 transcriptomics. *BMC Genomics* **19**, 477 (2018). 1582

102. Borcherding, N., Bormann, N.L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Res* **9**, 47 (2020).

1583

1584

1585

1588

1591

1594

1597

1600

1603

1607

1614

- 1586 103. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in 1587 multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).
- 1589 104. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytometry A* **83**, 483-494 (2013).
- 1592105.Zunder, E.R. *et al.* Palladium-based mass tag cell barcoding with a doublet-filtering scheme1593and single-cell deconvolution algorithm. Nat Protoc **10**, 316-333 (2015).
- 1595 106. Crowell, H., Zanotelli, V., Chevrier, S. & Robinson, M. CATALYST: Cytometry dATa 1596 anALYSis Tools. *R/Bioconductor* version 1.14.0 (2020).
- 1598 107. Chevrier, S. *et al.* Compensation of Signal Spillover in Suspension and Imaging Mass 1599 Cytometry. *Cell Syst* **6**, 612-620 e615 (2018).
- 1601 108. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection 1602 for Dimension Reduction. *arXiv:1802.03426* (2020).
- 1604109.McCarthy, D.J., Campbell, K.R., Lun, A.T. & Wills, Q.F. Scater: pre-processing, quality1605control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 33,16061179-1186 (2017).
- 1608 110. Van Gassen, S. *et al.* FlowSOM: Using self-organizing maps for visualization and 1609 interpretation of cytometry data. *Cytometry A* **87**, 636-645 (2015). 1610
- 1611 111. Saeys, Y., Van Gassen, S. & Lambrecht, B. Response to Orlova et al. "Science not art: statistically sound methods for identifying subsets in multi-dimensional flow and mass cytometry data sets". *Nat Rev Immunol* **18**, 78 (2017).
- 1615 112. Nowicka, M. *et al.* CyTOF workflow: differential discovery in high-throughput high-1616 dimensional cytometry datasets. *F1000Res* **6**, 748 (2017).
- 1618 113. Weber, L.M., Nowicka, M., Soneson, C. & Robinson, M.D. diffcyt: Differential discovery in 1619 high-dimensional cytometry via high-resolution clustering. *Commun Biol* **2**, 183 (2019).