

RESEARCH ARTICLE SUMMARY

MALARIA

A single-cell atlas of sexual development in *Plasmodium falciparum*

Sunil Kumar Dogga†, Jesse C. Rop†, Juliana Cudini†, Elias Farr, Antoine Dara, Dinkorma Ouologuem, Abdoulaye A. Djimdé, Arthur M. Talman, Mara K. N. Lawniczak*

INTRODUCTION: *Plasmodium falciparum* malaria is a major contributor to mortality and morbidity in developing nations. The extensive spatiotemporal genetic diversity of *P. falciparum* presents a challenge to the development of effective diagnostics, drugs, and vaccines. Co-infection with different *P. falciparum* strains occurs in more than 70% of infections in malaria-endemic populations. Understanding and monitoring parasite evolution in these settings will be necessary for effective control strategies. Additionally, resources that support deeper understanding of the sexual commitment and development of *P. falciparum* are needed to assist in the development of strategies that block parasite transmission.

RATIONALE: We have generated and explored a complete atlas of intraerythrocytic development of *P. falciparum*, enabling greater insight into its sexual development. This high-resolution reference atlas has great utility for profiling cell types found in natural infections. Single-cell RNA-seq (scRNA-seq) of parasites from natural infections enables extensive characterization of multiple aspects of variation between parasites including differences in gene expression, isoform diversity, genetic variation, relatedness, and transcriptional variation between strains in each stage. These different kinds of variation mediate parasite adaptation to dif-

ferent host conditions and likely underpin parasite evolution.

RESULTS: We generated both short- and long-read scRNA-seq data from ~37,000 laboratory malaria parasite cells covering the asexual and sexual developmental stages. Cell and gene clustering revealed a topology reflecting the intraerythrocytic stages, with sexual developmental stages branching off from the asexual replication cycle, progressing to form female and male gametocytes. Notably, a cell cluster at the base of the sexual stages corresponded to parasites expressing sexual commitment signatures. Using trajectory inference analysis, we investigated distinct expression modules as gametocytes differentiated, developed, and matured into late male and female stages. Furthermore, we identified discernible male and female characteristic signatures within the seemingly transcriptionally similar gametocyte developmental stalk. Using long-read data, we captured the intraerythrocytic cell cycle topology, identified novel isoforms, and discovered exon usage differences between life cycle stages. We also profiled and investigated ~8000 parasites obtained from four naturally infected malaria carriers, each of whom carried multiple genotypic strains. Within the sexual stages of these natural infections, we identified cell clusters that showed significantly reduced ex-

pression as the parasites aged. Notably, we observed transcriptional differences between genetically distinct strains within the same host in male and female gametocytes. These differences were mostly in genes that mediate host-parasite interactions. Integrating these natural infection scRNA-seq datasets into the laboratory reference atlas, we created a combined atlas comprising 45,691 cells. This integrated resource will support the mapping and analysis of both donor and laboratory malaria parasites, offering a comprehensive view for the interrogation of gene expression.

CONCLUSION: We characterized the intraerythrocytic cycle in lab strains with a focus on sexual development, exploring distinct expression modules underlying gametocyte development. Investigating natural infections at single-cell resolution enabled strain and stage assignment for each parasite and revealed unexpected transcriptomic clusters and differential expression between strains even within the same host. Long-read scRNA-seq exposed differential isoform usage between stages in lab strains and natural infections. The integrated dataset, comprising cells from laboratory strains and natural infections, spanning asexual and sexual development is presented as a new chapter in the interactive Malaria Cell Atlas data resource (malariaatlas.org). Single cell evaluations of malaria parasites from natural infections will enhance our understanding of malaria parasite persistence, pathology, and transmission dynamics, and this atlas will be a key resource underpinning future work. ■

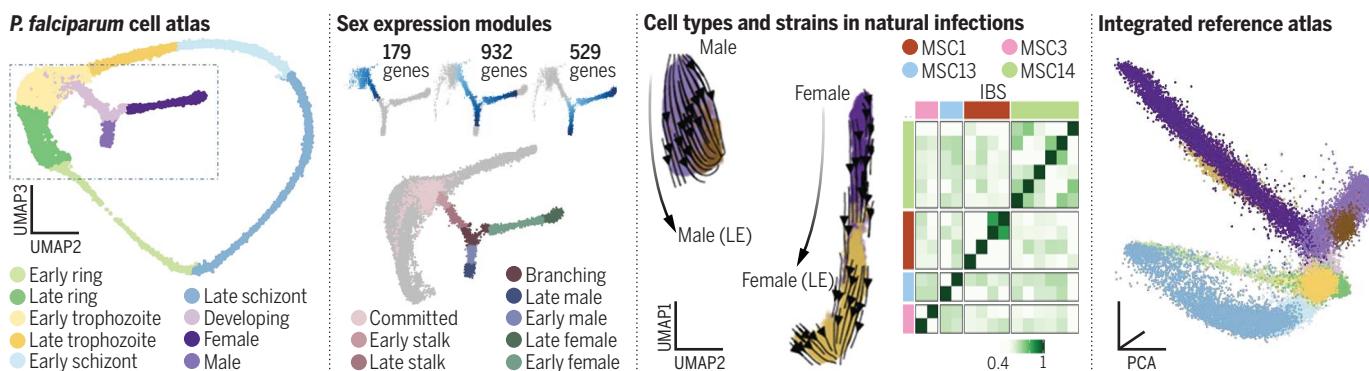
The full list of author affiliations can be found online.

*Corresponding author. Email: mara@sanger.ac.uk

†These authors contributed equally to this work.

Cite this article as S. K. Dogga et al., *Science* **384**, eadj4088 (2024). DOI: [10.1126/science.adj4088](https://doi.org/10.1126/science.adj4088)

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.adj4088>



Single-cell transcriptional and genotypic landscape of sexual and asexual stages in *Plasmodium falciparum*. The atlas of *P. falciparum* intraerythrocytic stages facilitated exploration of gene expression modules and clusters along sexual development. Profiling parasites from natural infections uncovered unexpected cell types, differential exon usage, and enabled examination of genetic relatedness between strains. The integrated atlas comprising cells from laboratory culture and natural infections provides an essential reference for malaria researchers (malariaatlas.org).

RESEARCH ARTICLE

MALARIA

A single cell atlas of sexual development in *Plasmodium falciparum*

Sunil Kumar Dogga^{1†}, Jesse C. Rop^{1†}, Juliana Cudini^{1†}, Elias Farr^{1,2}, Antoine Dara³, Dinkorma Ouologuem³, Abdoulaye A. Djimdé³, Arthur M. Talman⁴, Mara K. N. Lawniczak^{1*}

The developmental decision made by malaria parasites to become sexual underlies all malaria transmission. Here, we describe a rich atlas of short- and long-read single-cell transcriptomes of over 37,000 *Plasmodium falciparum* cells across intraerythrocytic asexual and sexual development. We used the atlas to explore transcriptional modules and exon usage along sexual development and expanded it to include malaria parasites collected from four Malian individuals naturally infected with multiple *P. falciparum* strains. We investigated genotypic and transcriptional heterogeneity within and among these wild strains at the single-cell level, finding differential expression between different strains even within the same host. These data are a key addition to the Malaria Cell Atlas interactive data resource, enabling a deeper understanding of the biology and diversity of transmission stages.

Malaria parasites undergo asexual proliferation in the human host, resulting in disease; however, transmission from human to human can only occur when parasites successfully reproduce sexually in the mosquito vector. Sexual reproduction is initiated in the human host and depends on a facultative commitment to sex. This decision is controlled by the epigenetically regulated transcription factor AP2-G (1, 2) and depends on environmental factors, including competition, drug treatment, and host metabolic factors (3–8). Once sexually committed, human malaria parasites develop into male and female gametocytes that circulate in the bloodstream until taken up by a blood-feeding *Anopheles* mosquito. Preventing sexual reproduction by stopping the parasite's development at any point during its sexual cycle breaks the cycle of transmission and will contribute to malaria control.

Patterns of gene expression following sexual commitment to development

To describe transcriptomic patterns observed during sexual commitment and sexual development for *Plasmodium falciparum*, we performed 10x single-cell RNA sequencing using both short Illumina reads and full-length PacBio IsoSeq on laboratory parasites sampled along several time points after gametocyte induc-

tion. After quality control (QC) and integration of short-read data, we retained 37,624 cells of two strains of *P. falciparum*, NF54 and 7G8, covering the asexual cycle, sexual commitment, and the development and differentiation of the sexes (Fig. 1A, fig. S1, and Materials and Methods). Although parasite development and differentiation are continuous processes, researchers have defined discrete stages based on morphology and gene expression. We deployed a consensus stage prediction method using a combination of clustering, reference mapping to both single-cell and bulk data sets, and identification of marker genes to label cells and annotate stages of asexual and sexual development (figs. S2 to S4). We explored the data using UMAP (Uniform Manifold Approximation and Projection) (9) dimensionality reduction using a projection that portrays asexual cycling of parasites as a circle and the differentiation into sexual forms as a stalk and two branches (Materials and Methods). This projection was the most common one observed when exploring parasite development at single cell resolution and has also been observed in other single cell malaria datasets (10–13).

Gene expression signatures in sexually committed parasites

We labeled 8958 sexual cells in greater resolution to explore key cell fate transitions in detail during the developmental decisions that drive sexual commitment and sex determination (Fig. 1 and figs. S2 and S4). We define sexual commitment as the point at which asexuals are committed to exit the asexual cycle into the sexual trajectory, whereas sexual development spans the period as early gametocytes develop and then differentiate into clearly distinct male and female trajectories. We observe a cluster at the base of the UMAP

stalk shape emerging from the asexual trophozoites (Fig. 1B) that is transcriptionally distinct from trophozoites and marked by expression of early gametocyte markers such as *AP2-G*, *Pfs16*, *G27/25*, and *Pfg14-748*, as well as several previously reported markers of sexually committed rings and schizonts such as *HAD1*, *LSD2*, *REX4*, and *SURF8.2* among others (Fig. 1, B and C, and data S1) (14, 15). We conclude that this cluster corresponds to the earliest sexually committed state sampled in our data and termed it “committed.” Committed cells show a reduced repertoire of several genes encoding for proteins associated with knobs and Maurer's cleft that are involved in sequestering the asexual stages from the peripheral circulation by adherence to host endothelium, as previously reported (16). Among the 76 genes that are down-regulated, 11 overlap with those reported to localize to these morphological features (17). Committed cells also express a set of *ETRAMPs* and *GEXPs* that may assist in homing and retaining these early stages in the extravascular spaces of the bone marrow and other tissues in natural infections (data S1).

Modules of expression from commitment to differentiated sexes

After sexually committed cells have left the asexual cycle, in both PCA and UMAP analysis they become “stalk cells,” which appear to be neither male nor female. Subsequently stalk cells follow either male or female developmental “branches” (Fig. 1, A and B, and figs. S2 and S4). To understand the developmental decisions that lead to these transitions, we performed trajectory analyses and examined transcriptional changes along three paths: first the stalk, then along the branching clusters leading to male and female, and finally the progression through male and female development (Fig. 1D). For each of these three trajectories, we defined expression modules as described below and explored their underlying genetic programs.

Gene expression patterns within the stalk gametocytes

First, in examining the stalk cells as they depart from the asexual cycle, we identified distinct gene expression patterns that we classified as stalk (S) modules S1 to S3 (Fig. 1D, data S2, and fig. S5). In module S1, we find several Maurer's cleft proteins implicated in cytoadherence (*CBP2/GEXP07*, *REX1-3*, and *MAHRP1*) (18) that show a decrease in transcriptional capacity following sexual commitment, as do *AP2-G* and *GEXP02* (an early target of *AP2-G*) (19, 20). Module S2 shows transient expression of exported proteins (*ETRAMP5*, *ETRAMP9*, *MESA*, *GEXP14*, *GBP130*, and *ACBP2*) as well as two “plasmodium exported proteins” of unknown function, perhaps indicating a role in gametocyte sequestration (21–24). This is

¹Wellcome Sanger Institute, Hinxton CB10 1SA, UK. ²Institute for Computational Biomedicine, University of Heidelberg, Im Neuenheimer Feld 130.3, 69120 Heidelberg, Germany.

³Malaria Research and Training Center (MRTC), Faculty of Pharmacy, Université des Sciences, des Techniques et des Technologies de Bamako (USTTB), Point G, P.O. Box, 1805 Bamako, Mali. ⁴MIVEGEC, University of Montpellier, IRD, CNRS, Montpellier, France.

*Corresponding author. Email: mara@sanger.ac.uk

†These authors contributed equally to this work.

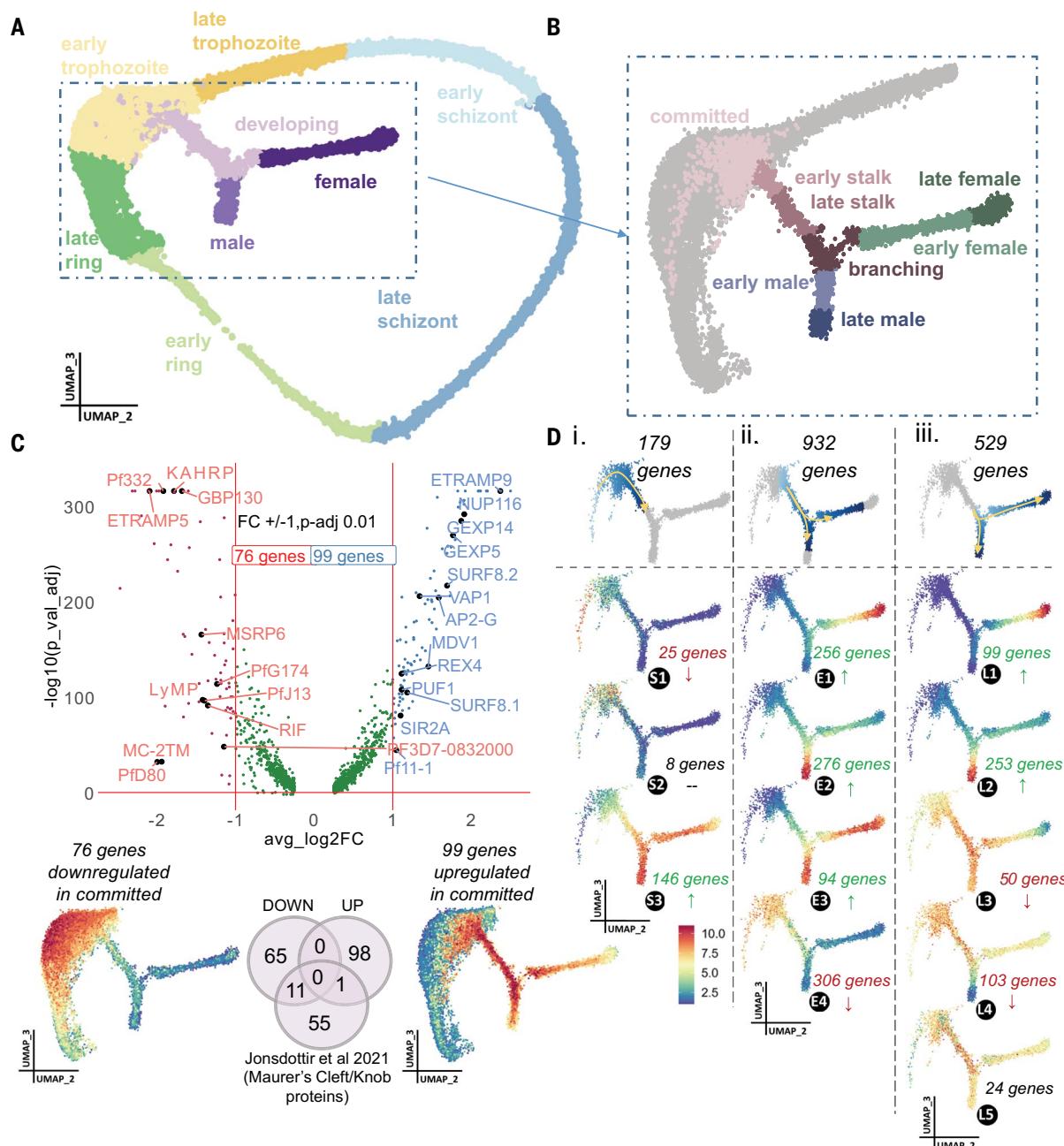


Fig. 1. An atlas of gametocyte development. (A) UMAP of 37,624 single cells colored by their assigned life cycle stage and (B) the subset of 8958 differentiating sexual cells. (C) Expression in the committed cells was explored in comparison to the early trophozoites from which the sexual stages project. The volcano plot shows the genes identified through MAST (presented in data S1). The panel below shows averaged expression of up-regulated and down-regulated genes on the cell UMAP. Comparison of differentially expressed candidates with proteins localized to Maurer's Cleft/Knobs (17) shows that committed cells down-regulate 11 of these

genes, as depicted in the Venn diagram. (D) Patterns of gene expression from commitment to differentiated sexes are explored and categorized into modules (i) following commitment, (ii) as the stalk gametocytes develop and branch into males or females, and (iii) following bifurcation, where males and females progress along their distinct sexual developmental trajectories. Modules are determined based on gene expression increases or decreases along the sexual lineages. Top UMAPs display the cells included in each comparison and the other UMAPs show the mean expression of all genes in each module. Complete results are presented in data S2.

followed by a transcriptional increase in factors affecting chromosome organization (H3, H2A, and the H3 variant H2B) and up-regulation of several transcription factors (including c-myc, HDPI, and ApiAp2 family proteins) in module S3 (fig. S5 and data S2). HDPI is a key player

for gametocyte maturation, facilitating transcription of genes essential for gametocyte development, including genes that are critical for the inner membrane complex and thereby gametocyte morphology (5, 25). Md1, which has been recently characterized as a key player in

sex determination and male development, is present in module S3 (26). Other uncharacterized nucleotide binding proteins within this module might similarly play roles in regulating gene expression at this critical juncture of sexual development.

Gene expression patterns as stalk gametocytes branch into female or male gametocytes

Second, we defined early (E) sex modules E1 to E4 in the stalk gametocytes as they branched into males or females. Sex-specific (modules E1 and E2) and sex-shared (module E3) transcripts increased (Fig. 1D, data S2, and fig. S6) with male gametocytes accumulating factors necessary for DNA replication, flagellar motility, and chromatin condensation (module E2), and female gametocytes accumulating mRNAs that become translationally repressed as they are later necessary for zygote and ookinete development in the mosquito vector (module E1) (27) (fig. S6 and data S2). The sex-shared module E3, while only represented by “microtubule-based movement” in GO enrichment analysis (data S2), comprises critical players for female and male gametocyte development (MDV1, MD1, MGET, and EBI) (data S2). We also find transient expression in exported proteins (module E4) putatively involved in erythrocyte remodeling in earlier gametocyte stages (21), suggesting that their continued expression may not be required for continued gametocyte sequestration (fig. S6 and data S2).

Gene expression patterns along maturation of female and male gametocytes

Finally, we examined transcriptional patterns after bifurcation is complete in the cell UMAP as the sexes mature, across late (L) sex modules L1 to L5 (Fig. 1D, data S2, and fig. S7). Sex-specific genes in modules L1 and L2—the majority of which are also represented in modules E1 and E2—appear to be initially transcribed within the branching gametocytes and continue to increase in expression through maturation. Modules L3 and L4 include genes that were shared between the developing gametocytes prior to sexual bifurcation but appear to maintain their expression only along male (L3) or female (L4) lineages following bifurcation (fig. S7 and data S2). Recent studies uncovered the role of an AT-rich interaction domain containing protein, PfARID, in regulating male gametogenesis and female fertility (28), and the knockdown of its *P. berghei* ortholog (*md4*) affects male gene expression (10). Notably, genes in modules E4, L3, and L4 that show transient expression along either or both sexual lineages are distinguished by a higher AT content in their putative promoter regions, suggesting a possible shared mode of transcriptional regulation of these modules, perhaps with the involvement of PfARID, though this requires further testing (fig. S8).

Investigating maleness or femaleness within the stalk gametocytes

The developmental stage at which a parasite commits to become either a male or female gametocyte remains unclear; i.e., the decision of whether to become male or female could be

established upon sexual commitment or could take place as a second step along the route to becoming a mature gametocyte (29–31). Additionally, one sex could be the default sex with additional factors required to push the default into the other sexual state. Although the cells sampled here reveal a developmental stalk in which gametocytes appear to be transcriptionally homogeneous before branching into male or female lineages, the UMAP is just a projection and may fail to reveal subtle signatures of maleness or femaleness among these stalk cells. Therefore, we investigated this late stalk to understand whether there were any transcriptional patterns consistent with sex already being determined in the stalk cells. We selected 10 genes from module S3 that show increased expression in the late stalk in the male lineage compared with the female lineage or vice versa, before bifurcation between the sexes is observable in the cell UMAP (figs. S5 and S9). We used the expression of these 10 genes to rank and group cells in the late stalk into putative male-like and female-like developing gametocytes (fig. S9). Differential expression analysis between these male-like and female-like gametocytes identified 44 genes with likely sex-dependent increases in expression well before bifurcation is observed in the cell UMAP. These include several known male and female marker genes (fig. S9 and data S3 and S4). The increase in transcription of these sex-specific genes in the stalk stages suggests that gametocytes become transcriptionally dimorphic earlier than at the point of bifurcation in the cell UMAP, implying that the sex of at least some of the gametocytes is determined at the time of divergence from the asexual cycle.

Global view of gene expression based on coexpression patterns

Identification of gene coexpression patterns across intraerythrocytic development using UMAP reduction

Functionally unannotated genes still remain the majority in the *P. falciparum* genome and classifying them based on their shared expression patterns with functionally annotated genes can reveal something about their functional relevance. To explore this further, we used UMAP to collapse a normalized, transposed cell-by-gene matrix and used k-means clustering to group these genes into 15 clusters (Fig. 2A and data S5; interactive version available at malariaatlas.org). The gene graph revealed a notable structure with most genes clustering according to their use across life stages (Fig. 2A and data S5). The gene graph displays three gene clusters specifically associated with sexual development and differentiation (2, 5, and 9) that also share genes with many of the modules described above (Fig. 2B). In some cases, these modules also contain genes found in other clusters, perhaps indicating recycled

gene usage in asexual and sexual stages or indicating genes that play a role in pushing asexual cells onto the sexual trajectory (Fig. 2B). Together, these three clusters contain the majority of sex-specific genes including 40 of 41 previously identified “gold standard” sexual markers (32) (Fig. 2C). Notably, these three clusters have a higher proportion of genes of unknown function than all other clusters and genes contained within them should be explored as potential transmission blocking targets (Fig. 2D).

Gene ontology (GO) analysis reveals similarly informative functional associations, as illustrated by clusters 1 and 8 (top GO cellular component terms, “apical part of cell” and “rhoptry”) comprising a significant number of proteins used in invasion and egress of the intraerythrocytic forms (data S5). Genes in cluster 9 predominantly exhibit expression in male gametocytes, aligning with the presence of gold standard sexual markers (PF3D7_0115100, PfMR5, PF3D7_1325200, and PUF1) and are enriched in GO terms related to cytoskeleton and motility. Notably, cluster 2, which exhibits consistent expression throughout all sexual stages (consisting of gametocyte markers P48/45, PfS230, Pfg27, MDV1, and PfS16) and cluster 5 with a predominant expression in female gametocytes (consisting of female gametocyte markers PfS25, pCCP5, PFs47, NEK2, and NEK4) present GO terms related to monocarboxylic and fatty acid biosynthetic processes as top terms, warranting further exploration toward possible metabolic targets for transmission-blocking interventions (data S5).

Inferring interactions of transcription factors with sexual gene clusters using gene regulatory networks

We next employed an unbiased gene regulatory network (GRN) inference approach to decipher the interactions of 96 experimentally validated and manually curated transcription factors (33) in regulating these 15 clusters using GENIE3 (34). We illustrate the top eight transcription factors that have multiple regulatory connections to the sex clusters (Fig. 2, E and F, and data S6). FD3 and MD4/PfARID show a high degree of connectivity to the female and male clusters, respectively, and play a role in gametocyte fertility in *P. berghei* and *P. falciparum* (10, 28, 35, 36). Disruption of HMGB2, AP2-O3, and PF3D7_0522900 leads to reduced growth in the ookinete or oocyst stages, indicating that downstream players of these factors may affect transmission (37–41). Finally, we find an association of the male cluster and AP2-G5, which is known to regulate AP2-G and is essential for gametocyte maturation (1, 42). Although GRN inferences have limitations (43), these are promising candidates for potential roles in sexual determination and development.

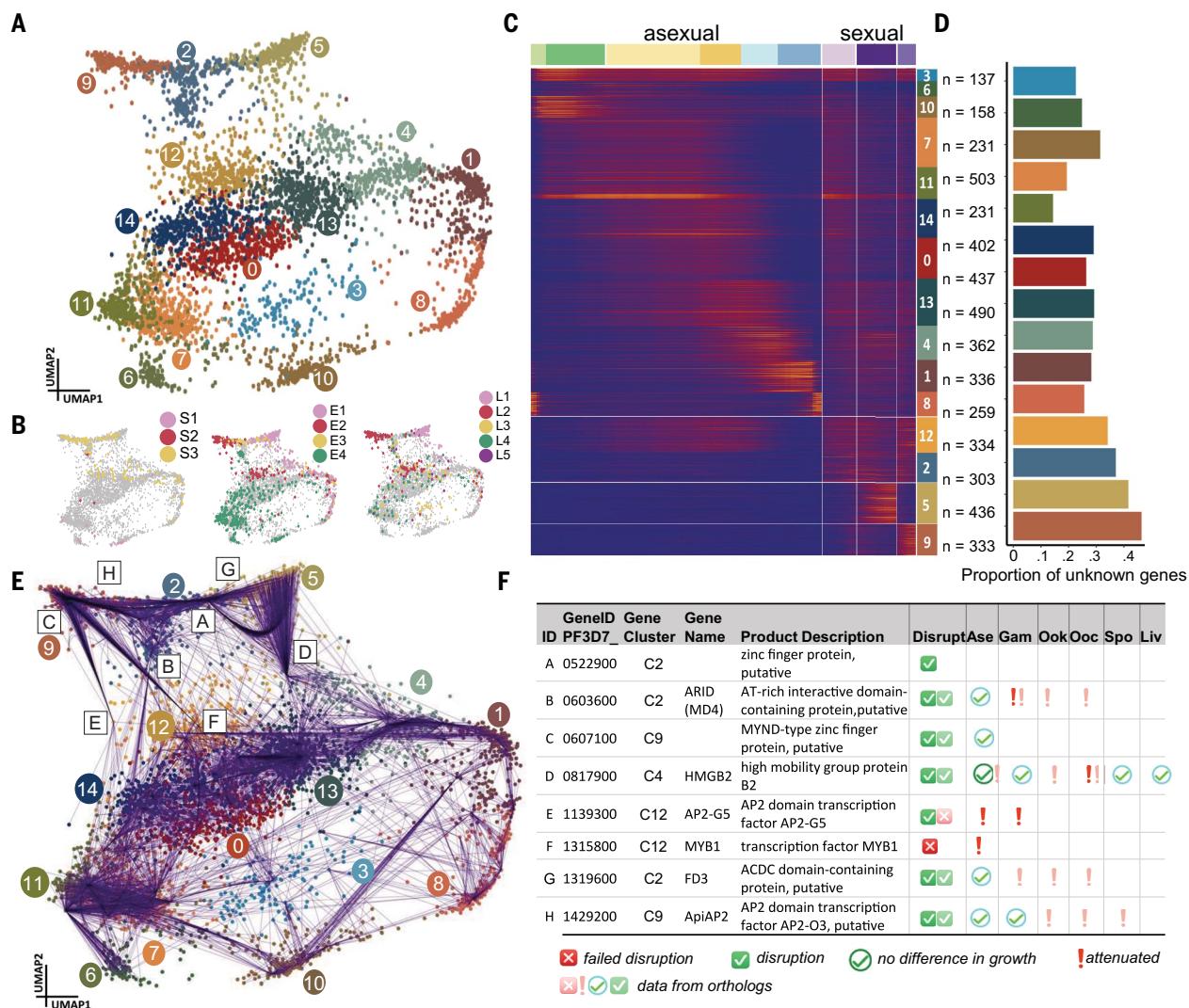


Fig. 2. Patterns of gene coexpression across the life cycle stages. (A) UMAP of 5087 genes (each dot is a gene) based on their expression in each cell. Genes are colored by their cluster assignment, using k-means ($k = 15$) clustering on the predicted UMAP dimensions. (B) Transcriptional modules described in Fig. 1D show considerable overlap with sex-specific gene clusters 2, 5, and 9. (C) Heatmap of gene expression clusters, where cells were ordered after pseudotime and assembled to 300 cell pseudobulks. (D) Proportion of genes in each cluster that have no known function; this is higher for the sex-specific clusters. (E) GENIE3-based gene regulatory network

to infer potential regulators of the sexual gene clusters, showing the strongest regulatory network connections. (F) Top eight candidates that appear to regulate genes in the sexually associated clusters, along with phenotypic information from Phenoplasm (36). (X, refractory to deletion; boxed tick, successfully disrupted; circled tick, no difference in asexual growth; exclamation mark, attenuated phenotype in the life cycle stages. Ase, asexual; gamm, gametocyte; ook, ookinete; ooc, oocyst; spo, sporozoite; liv, liver. Lighter colored icons indicate data from orthologs). The full list of genes ranked by number of connections is presented in data S6.

Interrogating sexual development using single-cell IsoSeq

Identification and classification of stage-specific isoforms by long-read sequencing

Distinct isoform usage can drive developmental decision making, and alternative splicing often plays an important role in sex determination and differentiation (26, 44). To explore whether stage- and sex-specific isoforms might underlie critical decision points in the parasite, we generated single-cell IsoSeq data to capture full length transcripts for each of the five samples. After QC, for each sample we recovered more than 1 million unique transcripts with

an average length of 1 kb (table S1 and fig. S10). Reads from the five IsoSeq runs were merged and 37,614 cells were assigned stage labels based on short-read data. For each stage, we typically obtained between ~200K and 1M total reads and between ~50 and 230 reads per cell depending on the stage (fig. S10). We observed a good concordance of gene/UMI detection between long- and short-read data (fig. S10). Gene expression signatures and cell type clusters were also consistent (Fig. 3, A and B). Following QC and chaining of stage-specific isoforms to obtain unique isoforms (Methods), isoform diversity was explored within each

stage and across all stages using SQANTI3 (figs. S10 to S12, Methods) and presented as a data resource (data S7).

Revealing stage-specific exon usage, isoforms, and novel splicing events using long-read sequencing

The generation of cDNA in the 10x protocols should be full length but the relatively short length of our long-read data [1203 base pairs (bp) on average] compared with the expectation from gene annotations (3022 bp average transcript length on PlasmoDB) indicated that many reads do not fully span genes (explored

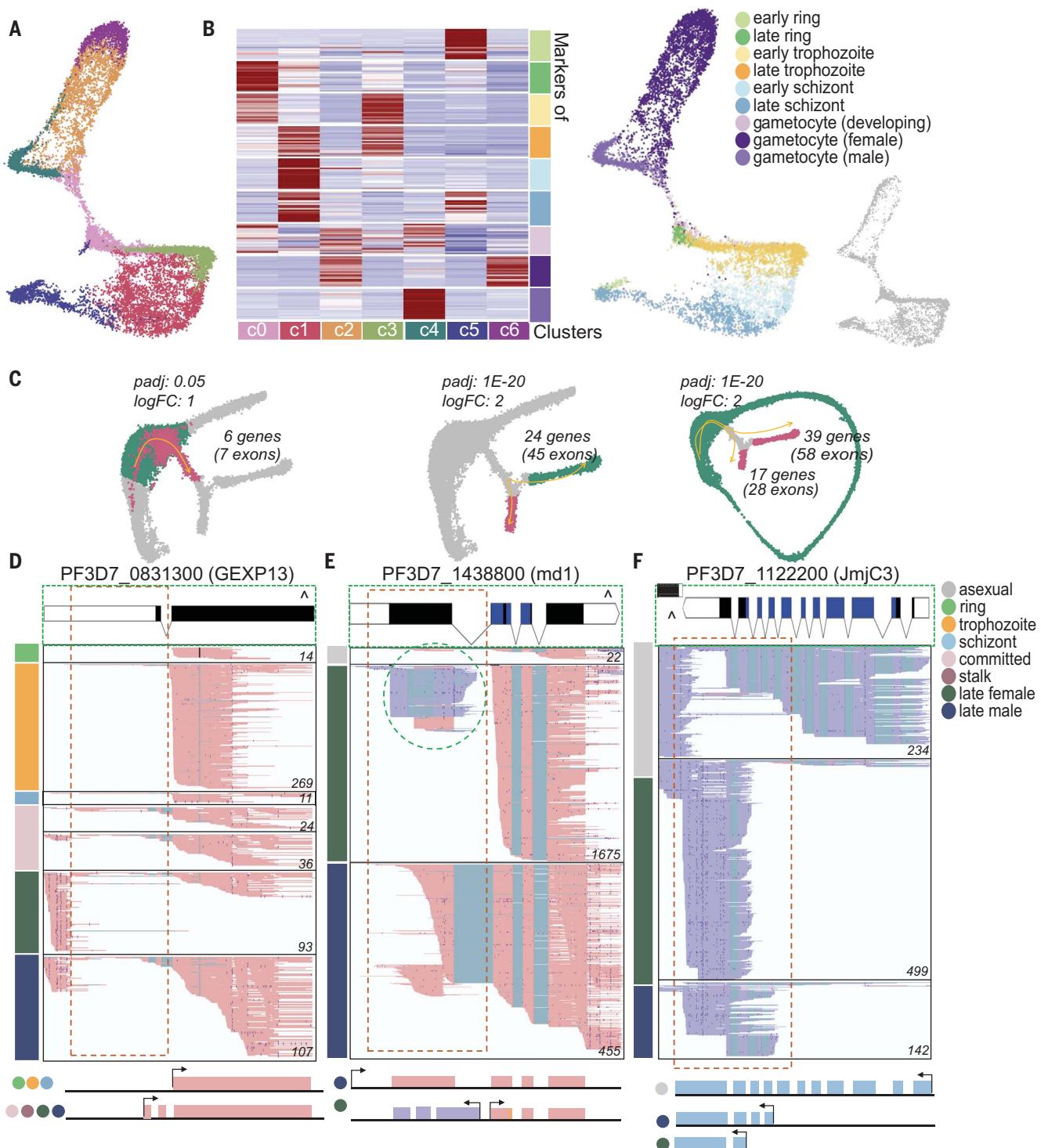


Fig. 3. Long-read data reveals stage-specific exon usage. (A) UMAP embedding of single-cell gene expression using long-read data echoes the topology inferred using short-read data. 17067 cells, with number of reads and genes >20, are colored according to the graph-based clusters determined using Seurat v4.1.0. (B) Identity of the above clusters is explored using the top 20 stage-specific markers from short-read data. The heatmap in the left panel shows the summed expression across all cells of each cluster for each stage-specific marker gene present. UMAP plots in the right panel show cells (12183 cells) in (A) colored by stage labels assigned using short-read data. Cells not present in the filtered short-read dataset (4884 cells) are

displayed in a separate panel in gray. (C) Differential exon usage (DEU) between early trophozoite versus committed (left); female versus male (middle); and asexual versus female and asexual versus male (right) was completed. In (D to F), long-read data is shown using IGV for three example genes, with PlasmoDB gene model on the top and inferred exon usage models on the bottom. PFAM domains in the gene model are displayed in blue. Summed coverage from IGV tracks is presented in the lower right corner. (D) Asexual stages present a novel isoform lacking exon 1 in GEXP13, whereas committed gametocytes and subsequent sexual stages show a previously undescribed splice site in the 5'UTR of the gene (data S7).

(E) In md1, female gametocytes lack exon 1 expression and display antisense expression of a spliced lncRNA (indicated by the dashed circle). The dominant isoform in the female stages appears to produce a fragmented translated product still containing the region coding for the OST-HTH

associated domain (data S7). **(F)** jmjC3 shows usage of all annotated exons in the asexual stages while altering its profile in both the males and females, where transcripts only spanning the terminal exons of the JmjC domain are present.

in Supplementary Text). To account for this technical challenge of incomplete full-length data, we used DEXseq (45) to investigate differential exon usage rather than isoform usage within and between the sexual and asexual stages. We found differential exon usage in six genes between committed and asexual trophozoites, in 24 genes between males and females, in 39 genes between females and asexuals, and in 17 genes between males and asexuals (Fig. 3C, fig. S13, and data S8). We present several representative illustrations of observed patterns as snapshots of IGV (integrative genomics viewer) visualizations (46) (Fig. 3, D to F, and fig. S14). For example, GEXP13—an exported protein potentially in the PHIST family (47)—shows a novel isoform appearing as the asexual stages commit to and proceed into sexual stages (Fig. 3D). This alternative isoform usage could influence sequestration or cell remodeling differences between sexuals and asexuals.

We also observe distinct exon usage in Md1 (Fig. 3E), as recently reported (26). Md1 plays important and distinct roles in sex determination and development; female gametocytes lack exon1 expression of Md1 altogether and instead show antisense transcription driven by a specific alternative bidirectional promoter (26). Such antisense expression, associated with decreased exon usage, was also observed in other examples (figs. S15 and S16), suggesting a common mode of regulation of stage-specific processes. Another notable candidate belonging to the Jumonji-type demethylase family gene jmjC3, which shows distinct exon usage profiles between gametocytes and asexuals, has the exon encoding its cupin-like domain not expressed in the sexual stages (Fig. 3F). Jumonji inhibitors have been shown to disrupt gametocyte development and formation in *P. falciparum* (48). Among the candidate genes explored through IGV visualization, we noticed several cases where current gene models in the annotation appear to be inaccurate or incomplete (fig. S16). Conclusions from the long-read data need to be carefully evaluated as we detected some differentially expressed exons that are likely the result of spurious reverse transcription due to polyA stretches (fig. S17). DEXseq analysis between NF54 and 7G8 strains within the female and male gametocytes revealed very few genes (15 and 3, respectively) that showed no clear differential exon usage profile when explored by IGV visualization (fig. S13 and data S8). DEXseq applied to the asexual stages also revealed several genes showing differential exon usage between rings, trophozoites, and

schizonts, which should be further explored for their influence on asexual development (fig. S18). It is clear that single-cell long-read data reveal novel and dynamic isoform usage that may be important in developmental decisions, and that improvements to the annotation and to the technical ability to represent complete full-length cDNA in single cells will enhance our understanding of the prevalence and importance of isoforms in parasite development.

Interrogating natural infections at single cell resolution

High-resolution stage and strain delineation in natural infections using single cell transcriptomics

We used the complete blood stage developmental time course discussed above as a reference atlas to investigate natural malaria infections profiled at the single cell level using 10x. We generated short and long-read scRNASeq data from asexual and sexual stage parasites collected from four naturally infected asymptomatic Malian donors aged 6 to 10 with a range of parasitemias and gametocytias (fig. S19). In areas of high endemicity such as Mali, people can carry different strains in a single infection. We aim to better understand the diversity of strains within infections, whether all strains contribute to the transmissible reservoir, and explore transcriptomic variability between different strains in this reservoir. For two donors (MSC1, 14) both asexual and sexual parasites were present and profiled, and for the other two donors (MSC3, 13) only sexual parasites were present and profiled (fig. S19). We used the short-read scRNASeq genotypes and transcriptomes to first perform quality control (QC) and to assign a strain (49) and a life cycle stage (50) to each parasite (figs. S20 and S21, and Methods). After QC and annotation we retained between 1123 and 3031 cells comprising the stages expected (early asexuals, late sexuals) from the peripheral venepuncture of asymptomatic natural infections (fig. S21) (51). Parasites from each donor were estimated to be derived from between 2 and 8 different genotypic strains distributed across the stages profiled (figs. S20 and S21).

Investigating the biology of cell types within natural infections

Unsupervised transcriptomic clustering of parasites from each donor resulted in three to five clusters per donor (figs. S20 and S21). Individual clusters from each donor were confidently assigned as asexual, male, or female based on

scmap to the lab atlas and expression of marker genes (fig. S21). However, in MSC3, 13, and 14, additional discrete cell clusters had convincing expression of known male or female markers but much weaker assignment scores to lab males or females in comparison to the canonical sexual clusters found in the same donors (32) (fig. S21). These weakly assigned clusters expressed fewer genes overall and these genes were expressed at greatly reduced levels compared with the canonical male and female clusters that were highly similar to lab cells (fig. S21 and data S9). Moreover, these clusters did not appear to be an artifact of the ex vivo purification protocol given that we did not observe weakly assigned gametocytes among mixed lab cultured asexuals and day 12 gametocytes processed in a mock field trial used to develop the ex vivo protocol (fig. S22). We labeled the weakly assigned male and female clusters as low expression (LE) for each donor separately (fig. S23). Notably, among donors with LE sexual stages, all strains that had any sexual stages also exhibited LE parasites, suggesting that the LE phenotype is not strain-specific (fig. S23). We explored the LE gametocytes further using donor MSC13 and 14, each of which had both male LE and female LE clusters, and MSC3 which only had the latter (fig. S23). For each LE cluster, gene expression across donors was strongly correlated, suggesting that LE cells are similar irrespective of donor (fig. S23). The LE males and LE females clearly expressed a subset of the core genes (defined as those detected in at least 25% of cells in that stage/cluster) expressed in their canonical counterparts but did not express any unique genes, which was not true of any other life stage (fig. S23 and data S9). GO analysis on core genes exclusive to all females (canonical and LE) showed enrichment for the crystalloid component among other GO terms (fig. S23 and data S9). Crystalloids are limited to the ookinete and oocyst but some genes encoding their proteins are expressed early on in female gametocytes (52). Core genes unique to canonical females were also enriched in the maternally associated apicoplast and mitochondrial components, and additionally, RNA-associated processes that are predictive of active transcript processing (53) (fig. S23 and data S9). GO terms attributed to the unique canonical male core genes and also the core genes shared by both canonical male and male LE included cilium, microtubule, and dynein complex components known to be essential for male gamete function (54) (fig. S23 and data S9). Integration of

gametocytes from all donors indicated that the LE clusters might be on a developmental continuum with the canonical gametocytes (Fig. 4, A to C, and fig. S24). RNA velocity analysis on the integrated dataset suggested that the parasites might be transitioning from the canonical gametocytes to the LE forms (Fig. 4C and fig. S24). Upon trajectory inference analysis to examine transcriptional changes along females, we observed a consistent decrease in the gene content from canonical to LE females (Fig. 4B, fig. S24, and data S10). Notably, a fraction of the genes enriched among those encoding for translationally repressed transcripts (27) appear to remain stable (fig. S24). Further work to determine the viability of these LE stages will be challenging given that they are found together with canonical stages in all donors explored thus far, but serial sampling post asexual clearance using drugs might enable a study of LE cell viability if they are indeed indicative of aged parasites.

Exploring strain and stage distribution and relatedness between strains within and between donors

For MSC14—the only donor from which we obtained considerable numbers of asexual parasites—we observed an imbalance in strain distribution across the asexual and sexual stages with one strain (SC1) contributing the majority of asexual parasites and only two strains (SC2 and SC6) convincingly present (>30 cells) in both sexual and asexual stages (fig. S25). The imbalanced distribution of strains in the asexual and sexual stages could be due to incomplete sampling, clearance of specific fractions by the immune system, differential propensities to sexually commit, or asynchronous sequestration. Exploring the female to male ratio of strains within donors revealed a consistent pattern across all strains in MSC3 and MSC14, being lower in the former and higher in the latter (fig. S25). However, in MSC1 and MSC13, some strains had more females whereas others had more males (fig. S25). The lower ratio of females to males in MSC3 is thought to be atypical, though male-biased sex ratio does occur in low-density infections (55, 56). Moreover, as shown in the rodent malaria model *Plasmodium chabaudi*, sex ratio may become male-biased with an increase in the number of genotypes (57), though we do not see this correlation among this small sample size of four donors.

To understand strain relatedness between and within donors while considering stage, we first validated souporcell genotype assignments using a custom pseudobulk genotyping pipeline (fig. S25 to S28, Methods). Souporcell has been shown to accurately identify minority clusters down to 25 cells (49). Using this cutoff we removed some strain-stage pseudobulk clusters in MSC1 (SC1, SC2, and SC4

asexuals) and some in MSC14 (SC1, SC3, and SC8 gametocytes, SC5 and SC7 asexuals, and SC4 asexuals and gametocytes) before pseudobulk genotyping and therefore all subsequent downstream strain analysis (fig. S25). The remaining strain-stage pseudobulk clusters had at least 30 parasites (fig. S25). After genotyping we also removed SC8 asexuals in MSC14 and SC5 gametocytes in MSC1 which had outlier genotype profiles similar to that of doublets (fig. S26 and data S9). From the resultant high-quality genotypes we observed high genetic similarity between at least two strains within MSC1 and 14, but not within MSC3 or 13 which only had two strains present (Fig. 4D, fig. S27, and data S11 and S12). We did not see any strong genetic similarity between parasites from donors despite the samples being collected over three weeks from the same region, indicative of high transmission (Fig. 4D, fig. S27, and data S11).

To better understand relatedness of strains within an infection, we focused on MSC14 which had the greatest strain diversity (fig. S28). For SC2 and SC6, the asexual and sexual genotypes were in agreement with souporcell assignments confirming that asexual and gametocytes assigned to the same strain were indeed genetically identical, despite clustering separately on the genotype PCA plot (figs. S25, S27, and S28 and data S11). We then estimated the genome proportions and locations that were identical by descent (IBD) between every pair of strains with validated genotypes (figs. S28 and S29 and data S12). Several strains shared considerable fractions of their genome in IBD tracts, suggesting very recent shared ancestry: SC1 shared 37% of its genome with SC2; SC2 shared 42% with SC6; and SC6 shared 45% with SC7 (figs. S28 and S29 and data S12). Despite this, SC7 shared no recent ancestry with SC2 nor indeed with any other strain, apart from SC6 (figs. S28 and S29 and data S12). Other pairs had IBD ranging from 0 to 25% (fig. S28 and data S12). IBD results reflected souporcell's PCA clustering: SC7 was separated from all strains along PC1 and was closest to the SC6 cluster, with which it shared IBD (figs. S25, S28, and data S12). SC1, 2, and 5 were separated along PC3 reflecting their levels of IBD (figs. S25 and S28, and data S12). Given the complex IBD patterns along the genome with all strains showing long stretches of relatedness to at least one other strain (fig. S29 and data S12), these parasites likely represent recombinants arising from the bite of a single mosquito. Recent results show a correlation between the total sporozoite load and the number of sporozoites released in a single bite and estimate that each *P. falciparum* oocyst typically releases more than 10,000 sporozoites (58). If a single mosquito bite was responsible for the strains detected in MSC14, this mosquito must have had at least 8 oocysts all of which arose from different parental strain

combinations and all of which must have contributed sporozoites to the inoculum.

Differential expression (DE) between strains in female and male gametocytes

To perform DE while accounting for developmental stage, we integrated the canonical sexual parasites from all donors into the lab-based reference atlas gametocytes (fig. S30). We excluded the LE gametocytes because these were quite different to the lab parasites based on scmap scores and would therefore not integrate accurately into the developmental trajectory of the lab gametocytes. Gene counts per cell can be different between samples as a result of either biology or variable cell numbers and/or reads per cell as a result of being processed in different batches (fig. S30). Different sequencing depth per cell for each sample due to these factors made it difficult to carry out DE analysis between donors and thus we focused on DE within each donor separately. We downsampled the field strains and sexes within each donor, ensuring that at least two strains had approximately equal numbers; we also did the same for the lab strains (fig. S31). From these downsampled cells, we further selected only strain/sex combinations that had at least 370k reads, as many combinations had very few total reads and thus would be unlikely to result in robust DE analyses (fig. S32). For the donor parasites, three strains in MSC14 females and two strains in MSC3 males passed this threshold, and we performed DE between pairwise combinations of these strains while accounting for pseudotime (Fig. 4, E to I, and fig. S32). This resulted in 152 DEGs for MSC14 females and 89 DEGs for MSC3 males (Fig. 4, F and I, fig. S32, and data S13). For MSC14 females where we had more than one pairwise strain comparison, the number of DEGs from these comparisons was negatively correlated with the degree of genetic relatedness with SC2 ($N = 23$) versus SC5 ($N = 42$) with IBD of 0.25 resulting in 29 genes and SC2 ($N = 23$) versus SC7 ($N = 43$) IBD of 0 giving 72 genes (Fig. 4F and fig. S28). SC5 ($N = 42$) versus SC7 ($N = 43$) IBD of 0 resulted in 111 genes but the higher number of parasites in this comparison may explain the higher number of DEGs compared with SC2 versus SC5 despite similar IBD levels (Fig. 4F and fig. S28). Notable genes among those with the strongest genotype-specific patterns of expression in MSC14 females included *Pfs16* and *Zip1*, both of which are implicated in parasite development within the mosquito (Fig. 4G) (59, 60). In MSC3 males, notable genes among the top DE genes included TREP, which has been linked to gliding motility of sporozoites in invading salivary glands (61, 62), and the Plasmodium-specific MyoE, highly expressed in all invasive stages with a potential role in sporozoite traversal

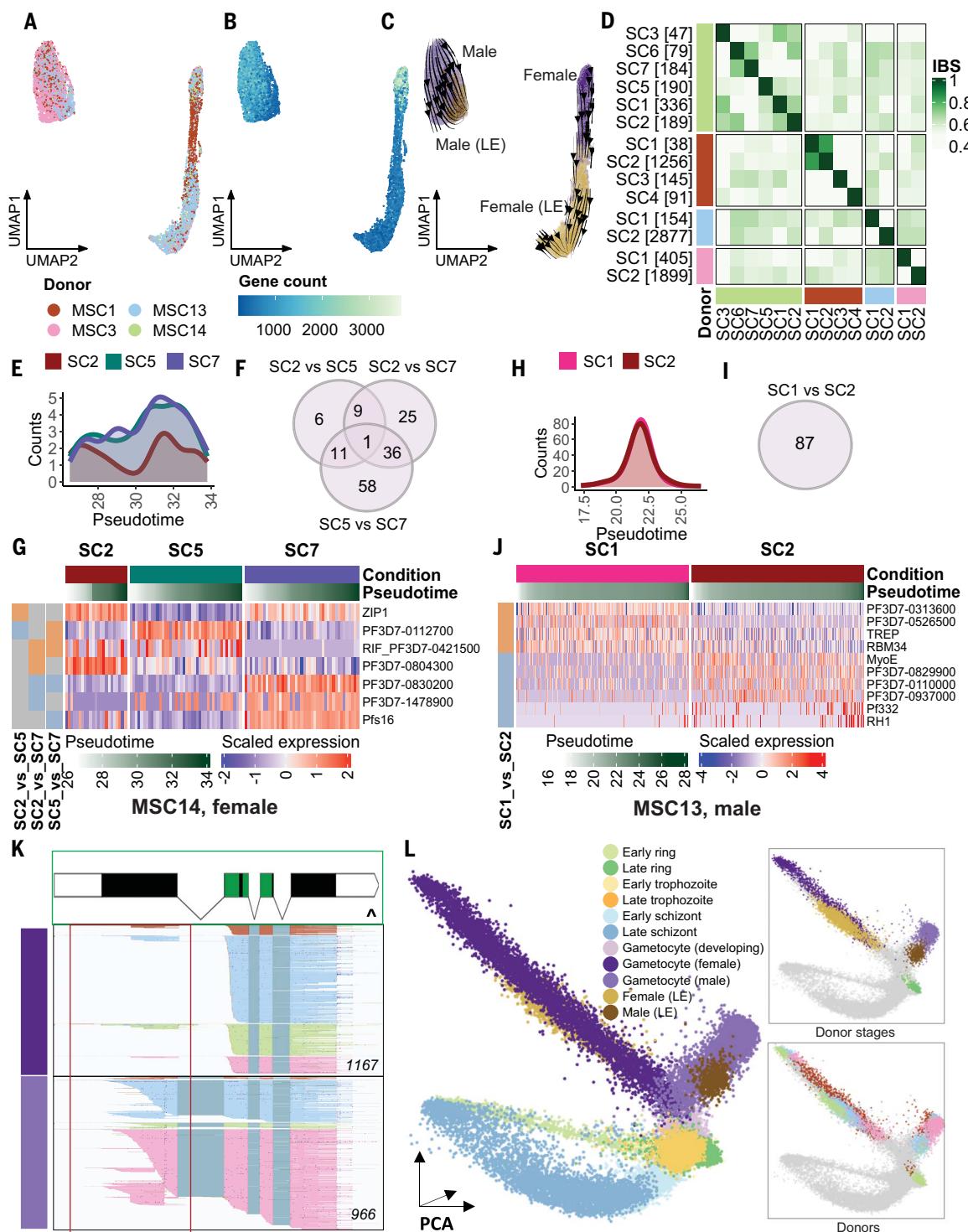


Fig 4. Stage and strain delineation of natural infections. (A) UMAP of the Seurat integrated gametocytes from four donors (MSC1, 3, 13, and 14), where each dot is a parasite colored by (A) donor (see fig. s24 for better donor resolution) (B) the number of genes expressed, and (C) stage. In (C) the arrows indicate the developmental progression based on RNA velocity as estimated by scvelo. (D) Genetic similarity estimation based on identity by state (IBS) measurements between pairwise comparison of strains between and across donors with the increasing similarity corresponding to increasing intensity of the green color. Within each donor the strains are arranged based on hierarchical clustering of the genetic similarity. The number inside the square brackets next

to the strain label corresponds to the number of parasites in that strain. (E, F, and G) DE of the strains containing sufficient numbers of parasites (SC2, SC5, and SC7) for MSC14 females. (E) Distribution of MSC14 female donor parasites that are subjected to differential expression colored by strain along pseudotime. (F) Total counts and overlap of all genes that come up as being differentially expressed (DE) (Adjusted MAST p-value <0.05) from each pairwise strain comparison. (G) Heatmap of the 10 highest log fold change DE genes (in rows) between pairwise strain comparisons of canonical females), with cells arranged in columns according to pseudotime and grouped by strain. The bars on the left are colored orange if the gene is highly expressed

in the first listed strain and blue if vice versa. (**H**, **I**, and **J**) The DE results for the MSC13 male SC1 and SC2 pairwise comparisons with similar description as for (E, F, and G), respectively. (**K**) Differential exon usage between males and females of donor parasites in exon 1 of *md1*, with reads colored by donor as

denoted in (A). (**L**) Integrated PCA plot containing all lab and field parasites, totalling 45,691. LE gametocytes have been labelled separately because they may reflect new biology and are highlighted together with other field stages in the top right panel, with the donors labeled in the bottom right panel.

to the salivary glands (63) (Fig. 4J). MyoE has been shown to coimmunoprecipitate with CDPK4 (64), a critical player in male gamete development and gamete motility (65), and strain-specific differences in its expression might indicate differences in transmission success once the gametocytes activate in the mosquito.

GO terms enriched in both female and male DE genes included host related processes, components and functions indicating a possible role in parasite adaptation to the host (fig. S32 and data S13). A separate analysis for the 7G8 versus NF54 grown in different cultures but processed in the same batch gave a different set of genes from those that came up in the donor strains with only 7.7% overlapping in female and only 6% in male (fig. S33 and data S13); however, GO enrichment analysis of these lab DE genes did not result in any significant hits. This is a notable difference between the GO results between strains within a host versus lab strains cultured separately and may reflect strain-specific responses to the host environment. We did explore DE among the strains that had reduced power as a result of fewer overall reads per cell, and these analyses yielded some VAR genes as well as ZIP1 and Pfs16 as seen above (fig. S34 and data S14). Altogether, the transcriptional heterogeneity we observed between strains within the same host highlights the importance of extending malaria parasite transcriptomic studies to natural settings to uncover critical players in host-parasite interactions that may mediate pathogenesis and transmission.

Exploring stage-specific isoforms and exon usage within field parasites

Single-cell IsoSeq on the parasites from each donor resulted in 0.8 to 1.65M unique long transcripts derived from parasites, most of which came from canonical male and female cells (table S2 and fig. S35). We generated SQANTI3-based isoform classification of these transcripts within each assigned stage across all donors merged together (fig. S35 and S36 and data S15). Upon differential exon usage (DEU) analysis between males and females across all donors, over 75 genes showed a different profile, 38 of which were also present within the comparison of laboratory strain sexual stages using similar thresholds, and here we present a few of these examples (Fig. 4K, fig. S37, and data S16). Visual exploration of DEU between the canonical and LE females revealed that LE females tend to harbor longer transcripts whereas canonical females are more

variable in their transcript lengths, with some showing more 3' degradation (fig. S38). This phenomenon is also reflected in the length distribution of isoforms and proportion of full splice and incomplete splice match isoforms generated by SQANTI3 (fig. S38 and data S15). Notably, 24 of 41 (58%) genes detected by long reads in LE females are among those known to be translationally repressed, compared with 60 of the 129 (46%) genes detected in canonical females (fig. S38). This difference is equally reflected in short-read data (fig. S38). Our observations suggest that canonical gametocytes may be undergoing active transcription, a finding that is supported by apparent marked 3' degradation compared with LE gametocytes, probably a result of active maintenance of RNA homeostasis. However, the possibility remains that the LE gametocytes might be fully mature or dying female gametocytes that retain the longer transcripts in messenger ribonucleoprotein (mRNP) particles. mRNPs are macromolecular complexes that harbor nontranslating mRNAs bound by various RNA-binding proteins regulating translation, localization, and turnover (27, 66). Given that gametocytes can remain in circulation and maintain transmissibility for months after clearance of asexual parasites through drug treatment (67), we suspect that the LE gametocytes found in donors are likely to be aged gametocytes and future studies are needed to investigate their competency for transmission.

Integrated Malaria Cell Atlas as a reference for natural infections

Although using single-cell approaches to study parasites derived from natural infections is in its infancy (11, 68), unexpected observations underscore the importance of exploring natural infections at single cell resolution. We found signatures indicating that males and females may display substantially reduced gene expression as they age. We also observed transcriptional diversity between genotypically distinct strains within the same host at the same time, perhaps driven by genotype-specific host-parasite interactions. It is clear that natural infections are revealing new biology likely relevant to parasite persistence, pathology, and transmission. We have therefore integrated these first natural infection scRNAseq datasets into the reference atlas to create a combined resource that supports mapping of both donor and laboratory *P. falciparum* parasites together with a more complete view on interrogation of gene expression (Fig. 4L and fig. S39). The integrated object comprising 45,691 cells from

laboratory and natural infections covering asexual and sexual differentiation is presented as a new chapter to the interactive Malaria Cell Atlas data resource (malariaatlas.org). Single-cell approaches now enable us to study strain composition and expression behavior including differential expression and isoform usage in natural infections, and have enormous potential toward improving our understanding of malaria parasite biology and transmission in natural infections.

Methods summary

Full details of the materials and methods used are supplied in the Supplementary Materials.

Plasmodium falciparum laboratory strains were captured using the 10x Genomics Single Cell 3' v3.1 kit on multiple days following gametocyte induction, allowing us to profile the asexual and sexual developmental intraerythrocytic stages of the parasite. Similarly, circulating intraerythrocytic stages from four natural *P. falciparum* malaria donors were captured following the depletion of leukocytes and enrichment of parasitized erythrocytes. The cDNA obtained from 10x single cell emulsions underwent library preparation for both short-read Illumina sequencing and long-read PacBio Iso-Seq sequencing.

Using short-read sequencing data, we used 10x Genomic's cellranger to map reads onto the *P. falciparum* reference genome and generate gene-cell matrix counts. Subsequently, Seurat was employed for quality control, cell clustering, and UMAP embedding. A reference atlas comprising ~37,000 cells from the laboratory strains was created and annotated with life cycle stages using a combination of clustering and mapping to existing RNA-seq datasets with scmap. Slingshot was used to infer pseudotime and developmental lineage trajectories. Additionally, tradeSeq and MAST were used to infer differential gene expression changes within and between lineages and cell types. Parasites from natural infections were similarly analyzed and assigned stages using the laboratory atlas generated above as a reference. Souporcell was used to delineate strain composition within each donor and custom scripts were employed for pseudobulk genotyping and the curation of SNPs used to infer relatedness between strains. snpRelate and hmmIBD were applied to estimate identity by state (IBS) and identity by descent (IBD) respectively.

Long-read RNA sequencing data was analyzed using PacBio's Iso-Seq analysis workflow and read clustering at the isoform level.

SQANTI3 was used for quality control, filtering, and classification of these isoforms. DEXseq was employed to infer exon usage differences between various life cycle stages within the laboratory dataset and natural infections.

Cells from the laboratory and natural infection strains were integrated using Harmony, providing an essential reference for malaria researchers.

REFERENCES AND NOTES

- B. F. C. Kafsack et al., A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248–252 (2014). doi: [10.1038/nature12920](https://doi.org/10.1038/nature12920); pmid: [24572369](https://pubmed.ncbi.nlm.nih.gov/24572369/)
- A. Sinha et al., A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. *Nature* **507**, 253–257 (2014). doi: [10.1038/nature12970](https://doi.org/10.1038/nature12970); pmid: [24572359](https://pubmed.ncbi.nlm.nih.gov/24572359/)
- N. M. B. Brancucci et al., Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite *Plasmodium falciparum*. *Cell* **171**, 1532–1544.e15 (2017). doi: [10.1016/j.cell.2017.10.020](https://doi.org/10.1016/j.cell.2017.10.020); pmid: [29129376](https://pubmed.ncbi.nlm.nih.gov/29129376/)
- H. P. Portugaliza et al., Artemisinin exposure at the ring or trophozoite stage impacts *Plasmodium falciparum* sexual conversion differently. *eLife* **9**, e60058 (2020). doi: [10.7554/eLife.60058](https://doi.org/10.7554/eLife.60058); pmid: [33084568](https://pubmed.ncbi.nlm.nih.gov/33084568/)
- A. Poran et al., Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature* **551**, 95–99 (2017). doi: [10.1038/nature24280](https://doi.org/10.1038/nature24280); pmid: [29094698](https://pubmed.ncbi.nlm.nih.gov/29094698/)
- P. Schneider et al., Adaptive plasticity in the gametocyte conversion rate of malaria parasites. *PLOS Pathog.* **14**, e1007371 (2018). doi: [10.1371/journal.ppat.1007371](https://doi.org/10.1371/journal.ppat.1007371); pmid: [30427935](https://pubmed.ncbi.nlm.nih.gov/30427935/)
- A. I. Abdi et al., *Plasmodium falciparum* adapts its investment into replication versus transmission according to the host environment. *eLife* **12**, e85140 (2023). doi: [10.7554/eLife.85140](https://doi.org/10.7554/eLife.85140); pmid: [36916164](https://pubmed.ncbi.nlm.nih.gov/36916164/)
- C. T. Harris et al., Sexual differentiation in human malaria parasites is regulated by competition between phospholipid metabolism and histone methylation. *Nat. Microbiol.* **8**, 1280–1292 (2023). doi: [10.1038/s41564-023-01396-w](https://doi.org/10.1038/s41564-023-01396-w); pmid: [37277533](https://pubmed.ncbi.nlm.nih.gov/37277533/)
- L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv*:1802.03426 [stat.ML] (2018).
- A. J. C. Russell et al., Regulators of male and female sexual development are critical for the transmission of a malaria parasite. *Cell Host Microbe* **31**, 305–319.e10 (2023). doi: [10.1016/j.chom.2022.12.011](https://doi.org/10.1016/j.chom.2022.12.011); pmid: [36634679](https://pubmed.ncbi.nlm.nih.gov/36634679/)
- V. M. Howick et al., The Malaria Cell Atlas: Single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science* **365**, eaaw2619 (2019). doi: [10.1126/science.aaw2619](https://doi.org/10.1126/science.aaw2619); pmid: [31439762](https://pubmed.ncbi.nlm.nih.gov/31439762/)
- F. Hentzschel et al., Host cell maturation modulates parasite invasion and sexual differentiation in *Plasmodium berghei*. *Sci. Adv.* **8**, eabm7348 (2022). doi: [10.1126/sciadv.abm7348](https://doi.org/10.1126/sciadv.abm7348); pmid: [35476438](https://pubmed.ncbi.nlm.nih.gov/35476438/)
- C. Nötzel, B. F. C. Kafsack, There and back again: Malaria parasite single-cell transcriptomics comes full circle. *Trends Parasitol.* **37**, 850–852 (2021). doi: [10.1016/j.pt.2021.07.011](https://doi.org/10.1016/j.pt.2021.07.011); pmid: [34391665](https://pubmed.ncbi.nlm.nih.gov/34391665/)
- S. K. Prajapati et al., The transcriptome of circulating sexually committed *Plasmodium falciparum* ring stage parasites forecasts malaria transmission potential. *Nat. Commun.* **11**, 6159 (2020). doi: [10.1038/s41467-020-19988-z](https://doi.org/10.1038/s41467-020-19988-z); pmid: [3268801](https://pubmed.ncbi.nlm.nih.gov/3268801/)
- O. Llorà-Batlle et al., Conditional expression of PfAP2-G for controlled massive sexual conversion in *Plasmodium falciparum*. *Sci. Adv.* **6**, eaaz5057 (2020). doi: [10.1126/sciadv.aaz5057](https://doi.org/10.1126/sciadv.aaz5057); pmid: [32577509](https://pubmed.ncbi.nlm.nih.gov/32577509/)
- M. Tibúrcio et al., Early gametocytes of the malaria parasite *Plasmodium falciparum* specifically remodel the adhesive properties of infected erythrocyte surface. *Cell. Microbiol.* **15**, 647–659 (2013). doi: [10.1111/cmi.12062](https://doi.org/10.1111/cmi.12062); pmid: [23114006](https://pubmed.ncbi.nlm.nih.gov/23114006/)
- T. K. Jonsdottir, M. Gabriela, S. C. Crabb, T. F. de Koning-Ward, P. R. Gilson, Defining the Essential Exportome of the Malaria Parasite. *Trends Parasitol.* **37**, 664–675 (2021). doi: [10.1016/j.pt.2021.04.009](https://doi.org/10.1016/j.pt.2021.04.009); pmid: [33985912](https://pubmed.ncbi.nlm.nih.gov/33985912/)
- E. Mundwiler-Pachlatko, H.-P. Beck, Maurer's clefts, the enigma of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 19987–19994 (2013). doi: [10.1073/pnas.1309247110](https://doi.org/10.1073/pnas.1309247110); pmid: [24284172](https://pubmed.ncbi.nlm.nih.gov/24284172/)
- H. P. Portugaliza, O. Llorà-Batlle, A. Rosanas-Urgell, A. Cortés, Reporter lines based on the gexp0 promoter enable early quantification of sexual conversion rates in the malaria parasite *Plasmodium falciparum*. *Sci. Rep.* **9**, 14595 (2019). doi: [10.1038/s41598-019-50768-y](https://doi.org/10.1038/s41598-019-50768-y); pmid: [31601834](https://pubmed.ncbi.nlm.nih.gov/31601834/)
- G. A. Josling et al., Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nat. Commun.* **11**, 1503 (2020). doi: [10.1038/s41467-020-15026-0](https://doi.org/10.1038/s41467-020-15026-0); pmid: [32198457](https://pubmed.ncbi.nlm.nih.gov/32198457/)
- F. Silvestrini et al., Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell. Proteomics* **9**, 1437–1448 (2010). doi: [10.1074/mcp.M900479-MCP200](https://doi.org/10.1074/mcp.M900479-MCP200); pmid: [20332084](https://pubmed.ncbi.nlm.nih.gov/20332084/)
- T. Spielmann, D. J. P. Ferguson, H.-P. Beck, etramps, a new *Plasmodium falciparum* gene family coding for developmentally regulated and highly charged membrane proteins located at the parasite-host cell interface. *Mol. Biol. Cell* **14**, 1529–1544 (2003). doi: [10.1091/mbc.e02-04-0240](https://doi.org/10.1091/mbc.e02-04-0240); pmid: [12686607](https://pubmed.ncbi.nlm.nih.gov/12686607/)
- T. Nessel et al., EXP1 is required for organisation of EXP2 in the intraerythrocytic malaria parasite vacuole. *Cell. Microbiol.* **22**, e13168 (2020). doi: [10.1111/cmi.13168](https://doi.org/10.1111/cmi.13168); pmid: [31990132](https://pubmed.ncbi.nlm.nih.gov/31990132/)
- R. L. Coppel, S. Lustigman, L. Murray, R. F. Anders, MESA is a *Plasmodium falciparum* phosphoprotein associated with the erythrocyte membrane skeleton. *Mol. Biochem. Parasitol.* **31**, 223–231 (1988). doi: [10.1016/0166-6851\(88\)90152-1](https://doi.org/10.1016/0166-6851(88)90152-1); pmid: [3065643](https://pubmed.ncbi.nlm.nih.gov/3065643/)
- R. A. C. Morillo et al., Publisher Correction: The transcriptional regulator HDPI controls expansion of the inner membrane complex during early sexual differentiation of malaria parasites. *Nature Microbiology* **7**, 289–299 (2022). doi: [10.1016/0166-6851\(88\)90152-1](https://doi.org/10.1016/0166-6851(88)90152-1); pmid: [3065643](https://pubmed.ncbi.nlm.nih.gov/3065643/)
- A. R. Gomes et al., A transcriptional switch controls sex determination in *Plasmodium falciparum*. *Nature* **612**, 528–533 (2022). doi: [10.1038/s41586-022-05509-z](https://doi.org/10.1038/s41586-022-05509-z); pmid: [36477538](https://pubmed.ncbi.nlm.nih.gov/36477538/)
- E. Lasonder et al., Integrated transcriptomic and proteomic analyses of *P. falciparum* gametocytes: Molecular insight into sex-specific processes and translational repression. *Nucleic Acids Res.* **44**, 6087–6101 (2016). doi: [10.1093/nar/gkw536](https://doi.org/10.1093/nar/gkw536); pmid: [27282255](https://pubmed.ncbi.nlm.nih.gov/27282255/)
- S. Kumar et al., PfARID Regulates *P. falciparum* Malaria Parasite Male Gametogenesis and Female Fertility and Is Critical for Parasite Transmission to the Mosquito Vector. *mBio* **13**, e0057822 (2022). doi: [10.1128/mbio.00578-22](https://doi.org/10.1128/mbio.00578-22); pmid: [35638735](https://pubmed.ncbi.nlm.nih.gov/35638735/)
- C. Bancells et al., Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*. *Nat. Microbiol.* **4**, 144–154 (2019). doi: [10.1038/s41564-018-0291-7](https://doi.org/10.1038/s41564-018-0291-7); pmid: [30478286](https://pubmed.ncbi.nlm.nih.gov/30478286/)
- F. Silvestrini, P. Alano, J. L. Williams, Commitment to the production of male and female gametocytes in the human malaria parasite *Plasmodium falciparum*. *Parasitology* **121**, 465–471 (2000). doi: [10.1017/S003182099006691](https://doi.org/10.1017/S003182099006691); pmid: [11287979](https://pubmed.ncbi.nlm.nih.gov/11287979/)
- T. G. Smith, P. Lourenco, R. Carter, D. Walliker, L. C. Randolph-Cartwright, Commitment to sexual differentiation in the human malaria parasite, *Plasmodium falciparum*. *Parasitology* **121**, 127–133 (2000). doi: [10.1017/S003182099006265](https://doi.org/10.1017/S003182099006265); pmid: [11085232](https://pubmed.ncbi.nlm.nih.gov/11085232/)
- L. Meerstein-Kessel et al., Probabilistic data integration identifies reliable gamete-specific proteins and transcripts in malaria parasites. *Sci. Rep.* **8**, 410 (2018). doi: [10.1038/s41598-017-18840-7](https://doi.org/10.1038/s41598-017-18840-7); pmid: [29323249](https://pubmed.ncbi.nlm.nih.gov/29323249/)
- R. Sardar et al., ApicoTFdb: The comprehensive web repository of apicomplexan transcription factors and transcription-associated co-factors. *Database* **2019**, baq094 (2019). doi: [10.1093/database/baz094](https://doi.org/10.1093/database/baz094); pmid: [31529106](https://pubmed.ncbi.nlm.nih.gov/31529106/)
- V. A. Huynh-Thu, A. Irthrum, L. Wehenkel, Y. Saeyns, P. Geurts, "Regulatory network inference with GENIE3: application to the DREAM5 challenge" in 3rd Annual Joint Conference on Systems Biology, Regulatory Genomics, and Reverse Engineering Challenges (2010); <https://orbi.ulg.ac.be/handle/2268/10758>.
- M. Yuda, I. Kaneko, Y. Murata, S. Iwanaga, T. Nishi, Mechanisms of triggering malaria gametocytogenesis by AP2-G. *Parasit. Int.* **84**, 102403 (2021). doi: [10.1016/j.parint.2021.102403](https://doi.org/10.1016/j.parint.2021.102403); pmid: [3419684](https://pubmed.ncbi.nlm.nih.gov/3419684/)
- T. Sanderson, J. C. Rayner, PhenoPlasm: A database of disruption phenotypes for malaria parasite genes. *Wellcome Open Res.* **2**, 45 <https://phenoplasm.org/> (2017). doi: [10.12688/wellcomeopenres.11896_2](https://doi.org/10.12688/wellcomeopenres.11896_2); pmid: [28748223](https://pubmed.ncbi.nlm.nih.gov/28748223/)
- S. Kumar, S. H. I. Kappe, PfHMG2B has a role in malaria parasite mosquito infection. *Front. Cell. Infect. Microbiol.* **12**, 1003214 (2022). doi: [10.3389/fcimb.2022.1003214](https://doi.org/10.3389/fcimb.2022.1003214); pmid: [36506024](https://pubmed.ncbi.nlm.nih.gov/36506024/)
- M. Gissot et al., High mobility group protein HMGB2 is a critical regulator of plasmodium oocyst development. *J. Biol. Chem.* **283**, 17030–17038 (2008). doi: [10.1074/jbc.M801637200](https://doi.org/10.1074/jbc.M801637200); pmid: [18400754](https://pubmed.ncbi.nlm.nih.gov/18400754/)
- K. Modrynska et al., A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. *Cell Host Microbe* **21**, 11–22 (2017). doi: [10.1016/j.chom.2016.12.003](https://doi.org/10.1016/j.chom.2016.12.003); pmid: [28018440](https://pubmed.ncbi.nlm.nih.gov/28018440/)
- Z. Li et al., Plasmodium transcription repressor AP2-O3 regulates sex-specific identity of gene expression in female gametocytes. *EMBO Rep.* **22**, e51660 (2021). doi: [10.1525/embr.202051660](https://doi.org/10.1525/embr.202051660); pmid: [33665945](https://pubmed.ncbi.nlm.nih.gov/33665945/)
- T. Nishi, I. Kaneko, S. Iwanaga, M. Yuda, PbAP2-FG2 and PbAP2-R2 function together as a transcriptional repressor complex essential for Plasmodium female development. *PLOS Pathog.* **19**, e1010890 (2023). doi: [10.1371/journal.ppat.1010890](https://doi.org/10.1371/journal.ppat.1010890); pmid: [36780562](https://pubmed.ncbi.nlm.nih.gov/36780562/)
- X. Shang et al., A cascade of transcriptional repression determines sexual commitment and development in *Plasmodium falciparum*. *Nucleic Acids Res.* **49**, 9264–9279 (2021). doi: [10.1093/nar/gkab683](https://doi.org/10.1093/nar/gkab683); pmid: [34365503](https://pubmed.ncbi.nlm.nih.gov/34365503/)
- K. Akers, T. M. Murali, Gene regulatory network inference in single-cell biology. *Curr. Opin. Syst. Biol.* **26**, 87–97 (2021). doi: [10.1016/j.coisb.2021.04.007](https://doi.org/10.1016/j.coisb.2021.04.007)
- L. M. Yeoh et al., Alternative splicing is required for stage differentiation in malaria parasites. *Genome Biol.* **20**, 151 (2019). doi: [10.1186/s13059-019-1756-6](https://doi.org/10.1186/s13059-019-1756-6); pmid: [31370870](https://pubmed.ncbi.nlm.nih.gov/31370870/)
- S. Anders, R. Ewers, W. Huber, Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012). doi: [10.1101/gr.133744.111](https://doi.org/10.1101/gr.133744.111); pmid: [22722343](https://pubmed.ncbi.nlm.nih.gov/22722343/)
- H. Thorsvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013). doi: [10.1093/bib/bbs017](https://doi.org/10.1093/bib/bbs017); pmid: [22517427](https://pubmed.ncbi.nlm.nih.gov/22517427/)
- J. D. Warncke, I. Vakonakis, H.-P. Beck, Plasmodium Helical Interspersed Subtelomeric (PHIST) Proteins, at the Center of Host Cell Remodeling. *Microbiol. Mol. Biol. Rev.* **80**, 905–927 (2016). doi: [10.1128/MMBR.00014-16](https://doi.org/10.1128/MMBR.00014-16); pmid: [27582258](https://pubmed.ncbi.nlm.nih.gov/27582258/)
- K. A. Matthews et al., Disruption of the *Plasmodium falciparum* Life Cycle through Transcriptional Reprogramming by Inhibitors of Junonji Demethylases. *ACS Infect. Dis.* **6**, 1058–1075 (2020). doi: [10.1021/acsinfecdis.9b00455](https://doi.org/10.1021/acsinfecdis.9b00455); pmid: [32272012](https://pubmed.ncbi.nlm.nih.gov/32272012/)
- H. Heaton et al., Souporcell: Robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* **17**, 615–620 (2020). doi: [10.1038/s41586-020-0820-1](https://doi.org/10.1038/s41586-020-0820-1); pmid: [32366989](https://pubmed.ncbi.nlm.nih.gov/32366989/)
- V. Y. Kiselev, A. Yu, M. Hemberg, scmap: Projection of single-cell RNA-seq data across data sets. *Nat. Methods* **15**, 359–362 (2018). doi: [10.1038/nmeth.4644](https://doi.org/10.1038/nmeth.4644); pmid: [29608555](https://pubmed.ncbi.nlm.nih.gov/29608555/)
- R. Thomson-Luque et al., *Plasmodium falciparum* transcription in different clinical presentations of malaria associates with circulation time of infected erythrocytes. *Nat. Commun.* **12**, 4711 (2021). doi: [10.1038/s41467-021-25062-z](https://doi.org/10.1038/s41467-021-25062-z); pmid: [34330920](https://pubmed.ncbi.nlm.nih.gov/34330920/)
- J. T. Dessens, A. Z. Trempl, S. A. Saeed, Crystalloids: Fascinating Parasite Organelles Essential for Malaria Transmission. *Trends Parasitol.* **37**, 581–584 (2021). doi: [10.1016/j.pt.2021.04.002](https://doi.org/10.1016/j.pt.2021.04.002); pmid: [33941493](https://pubmed.ncbi.nlm.nih.gov/33941493/)
- J. M. Santos et al., Maternally supplied S-acyl-transferase is required for crystalloid organelle formation and transmission of the malaria parasite. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7183–7188 (2016). doi: [10.1073/pnas.1522381113](https://doi.org/10.1073/pnas.1522381113); pmid: [27303037](https://pubmed.ncbi.nlm.nih.gov/27303037/)
- S. M. Khan et al., Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell* **121**, 675–687 (2005). doi: [10.1016/j.cell.2005.03.027](https://doi.org/10.1016/j.cell.2005.03.027); pmid: [15935755](https://pubmed.ncbi.nlm.nih.gov/15935755/)
- N. B. Henry et al., Biology of *Plasmodium falciparum* gamete sex ratio and implications in malaria parasite transmission. *Malar. J.* **18**, 70 (2019). doi: [10.1186/s12936-019-2707-0](https://doi.org/10.1186/s12936-019-2707-0); pmid: [30866941](https://pubmed.ncbi.nlm.nih.gov/30866941/)
- J. Bradley et al., Predicting the likelihood and intensity of mosquito infection from sex specific *Plasmodium falciparum* gamete density. *eLife* **7**, e34463 (2018). doi: [10.7554/eLife.34463](https://doi.org/10.7554/eLife.34463); pmid: [29848446](https://pubmed.ncbi.nlm.nih.gov/29848446/)
- S. E. Reece, D. R. Drew, A. Gardner, Sex ratio adjustment and kin discrimination in malaria parasites. *Nature* **453**, 609–614 (2008). doi: [10.1038/nature06954](https://doi.org/10.1038/nature06954); pmid: [18509435](https://pubmed.ncbi.nlm.nih.gov/18509435/)
- C. Andolina et al., A transmission bottleneck for malaria? Quantification of sporozoites expelling from laboratory and natural *P. falciparum* infections. *bioRxiv* 2020.06.12.130195 [Preprint] (2023) doi: [10.1101/2023.08.03.55175](https://doi.org/10.1101/2023.08.03.55175)

59. C. P. Sayers, V. Mollard, H. D. Buchanan, G. I. McFadden, C. D. Goodman. A genetic screen in rodent malaria parasites identifies five new apicoplast putative membrane transporters, one of which is essential in human malaria parasites. *Cell. Microbiol.* **20**, (2018). doi: [10.1111/cmi.12789](https://doi.org/10.1111/cmi.12789); pmid: [28902970](#)
60. J. Ramelow, Y. Keleta, G. Niu, X. Wang, J. Li. Plasmodium parasitophorous vacuole membrane protein Pfs16 promotes malaria transmission by silencing mosquito immunity. *J. Biol. Chem.* **299**, 104824 (2023). doi: [10.1016/j.jbc.2023.104824](https://doi.org/10.1016/j.jbc.2023.104824); pmid: [37196765](#)
61. A. Combe *et al.*, TREP, a novel protein necessary for gliding motility of the malaria sporozoite. *Int. J. Parasitol.* **39**, 489–496 (2009). doi: [10.1016/j.ijpara.2008.10.004](https://doi.org/10.1016/j.ijpara.2008.10.004); pmid: [19000911](#)
62. K. Beyer *et al.*, Limited Plasmodium sporozoite gliding motility in the absence of TRAP family adhesins. *Malar. J.* **20**, 430 (2021). doi: [10.1186/s12936-021-03960-3](https://doi.org/10.1186/s12936-021-03960-3); pmid: [34717635](#)
63. R. J. Wall *et al.*, Systematic analysis of Plasmodium myosins reveals differential expression, localisation, and function in invasive and proliferative parasite stages. *Cell. Microbiol.* **21**, e13082 (2019). doi: [10.1111/cmi.13082](https://doi.org/10.1111/cmi.13082); pmid: [31283102](#)
64. H. Fang *et al.*, Epistasis studies reveal redundancy among calcium-dependent protein kinases in motility and invasion of malaria parasites. *Nat. Commun.* **9**, 4248 (2018). doi: [10.1038/s41467-018-06733-w](https://doi.org/10.1038/s41467-018-06733-w); pmid: [30315162](#)
65. O. Billker *et al.*, Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. *Cell* **117**, 503–514 (2004). doi: [10.1016/S0092-8674\(04\)00449-0](https://doi.org/10.1016/S0092-8674(04)00449-0); pmid: [15137943](#)
66. G. R. Mair *et al.*, Regulation of sexual development of Plasmodium by translational repression. *Science* **313**, 667–669 (2006). doi: [10.1126/science.1125129](https://doi.org/10.1126/science.1125129); pmid: [16888139](#)
67. T. Bousema *et al.*, Revisiting the circulation time of *Plasmodium falciparum* gametocytes: Molecular detection methods to estimate the duration of gametocyte carriage and the effect of gametocytocidal drugs. *Malar. J.* **9**, 136 (2010). doi: [10.1186/1475-2875-9-136](https://doi.org/10.1186/1475-2875-9-136); pmid: [20497536](#)
68. A. Dara *et al.*, Tackling malaria transmission at a single cell level in an endemic setting in sub-Saharan Africa. *Nat. Commun.* **13**, 2679 (2022). doi: [10.1038/s41467-022-30268-w](https://doi.org/10.1038/s41467-022-30268-w); pmid: [35562353](#)

ACKNOWLEDGMENTS

This work was supported by a Medical Research Council research grant MR/S02445X/1 (to M.K.N.L.), which supports S.D. This work was also supported by Wellcome core funding to the Wellcome Sanger Institute (Grant 206194/Z/17/Z), which supports M.K.N.L. and J.C.R. Part of the field work in Mali was supported by the BMGF (INV-001927) through RCA 20_356 between Sanger and USTTB. We thank M. Jones and the staff of the Wellcome Sanger Institute Scientific Operations for their contribution to library preparation and sequencing; Y. Gu for help with running cellranger on raw data, D. Serre, B. Hazzard, L. Chappell, and L. Tseng who advised on single-cell IsoSeq, H. Heaton who advised on strain analysis, and V. Howick, who assisted with one of the sampling days. We also thank M.K. Trinh who advised on IBS analysis and M. Carrasquilla who advised on IBD analysis. We thank J. Smith and the Core Software Team at Wellcome Sanger Institute for their contribution to the MCA web resource. We also thank the team that runs the clinic in Faladie, Mali for their contributions to sampling natural infections as well as the study population. **Funding:** Medical Research Council research grant MR/S02445X/1 (M.K.N.L.) Wellcome core funding to the Wellcome Sanger Institute 206194/Z/17/Z Bill and Melinda Gates Foundation INV-001927 through RCA 20_356 between WSI and USTTB **Author contributions:** conceptualization: M.K.N.L., A.M.T., and A.A.D. Data Curation: S.K.D., J.C.R., J.C., E.F., and A.D. Formal Analysis: S.K.D., J.C.R., J.C., and E.F. Methodology: A.M.T., S.K.D., and J.C.R.

Investigation: S.K.D., J.C.R., A.D., E.F., J.C., and D.O. Visualization: S.K.D., J.C.R., and E.F. Funding acquisition: M.K.N.L., A.A.D., and A.M.T. Project administration: M.K.N.L. and A.A.D. Supervision: M.K.N.L., A.M.T., and A.A.D. Writing – original draft: S.K.D., J.C.R., E.F., and J.C. Writing – review and editing: S.K.D., J.C.R., E.F., J.C., A.D., D.O., A.M.T., A.A.D., and M.K.N.L. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All raw sequencing data have been deposited in the European Nucleotide Archive at European Molecular Biology Laboratory European Bioinformatics Institute (www.ebi.ac.uk/ena/) under accession numbers PRJEB58790. Stage and expression data across the life cycle stages profiled can be visualized at the Malaria Cell Atlas website (www.malariacellatlas.org). Metadata, raw counts and normalized counts for the cells are also available at the same website for download. Expression matrices, analysis code and supplementary files are archived on Zenodo (83). **License information:** Copyright © 2024 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adj4088

Materials and Methods

Supplementary Text

Figs. S1 to S39

Tables S1 to S3

References (69–120)

Data S1 to S16

MDAR Reproducibility Checklist

Submitted 12 July 2023; accepted 14 March 2024
10.1126/science.adj4088