Article

Transcriptional control of the *Cryptosporidium* life cycle

Katelyn A. Walzer[1], Jayesh Tandel[1], Jessica H. Byerly[1], Abigail M. Daniels[1], Jodi A. Gullicksrud[1], Eoin C. Whelan[2], Stephen D. Carro[1], Elise Krespan[1], Daniel P. Beiting[1] & Boris Striepen[1✉]

The parasite *Cryptosporidium* is a leading agent of diarrhoeal disease in young children, and a cause and consequence of chronic malnutrition[1,2]. There are no vaccines and only limited treatment options[3]. The parasite infects enterocytes, in which it engages in asexual and sexual replication[4], both of which are essential to continued infection and transmission. However, their molecular mechanisms remain largely unclear[5]. Here we use single-cell RNA sequencing to reveal the gene expression programme of the entire *Cryptosporidium parvum* life cycle in culture and in infected animals. Diverging from the prevailing model[6], we find support for only three intracellular stages: asexual type-I meronts, male gamonts and female gametes. We reveal a highly organized program for the assembly of components at each stage. Dissecting the underlying regulatory network, we identify the transcription factor Myb-M as the earliest determinant of male fate, in an organism that lacks genetic sex determination. Conditional expression of this factor overrides the developmental program and induces widespread maleness, while conditional deletion ablates male development. Both have a profound impact on the infection. A large set of stage-specific genes now provides the opportunity to understand, engineer and disrupt parasite sex and life cycle progression to advance the development of vaccines and treatments.

Cryptosporidiosis is an important waterborne disease, with increasing incidence even in countries with robust water safety infrastructure[7]. The infection is transmitted by ingestion of oocysts, which are highly resistant to chemical disinfection and the product of the parasite's single-host sexual life cycle. Sporozoites emerge from the oocyst and invade intestinal epithelial cells, with parasites propagating asexually for three consecutive cycles of invasion, intracellular replication and egress before differentiating into male and female gametes following an intrinsic developmental program[4,8]. Sexual reproduction then generates new oocysts, which exit the host for transmission or excyst in the gut leading to autoinfection and a life cycle reset. Sex is therefore critical for both continuous infection and transmission. Sex also drives genetic diversity and the parasite's ability to infect new hosts and environments as demonstrated by recent studies on the adaption of *Cryptosporidium hominis* and *C. parvum* to humans in the United States and in Europe[9,10]. However, the fundamental mechanisms of parasite sex remain enigmatic, including the progression from the asexual to the sexual phase, gamete sex determinations, and gamete recognition and fertilization[11]. Here we sought to define the molecular components of all life cycle stages to reveal the underlying mechanisms.

### Transcriptional analysis of *Cryptosporidium*

Earlier transcriptomic studies using *C. parvum*-infected cultures found that most reads map to the host and not the parasite genome.

We attempted to fractionate cells to isolate intracellular stages[12] but found yields to be insufficient. We therefore engineered parasites to express a fluorescent reporter in all stages (Extended Data Fig. 1a) and infected HCT-8 cell cultures with these parasites (Fig. 1a). The cultures were trypsinized to generate a cell suspension, from which infected cells were isolated using flow cytometry[8] and analysed using single-cell RNA sequencing (scRNA-seq; Fig. 1b,c and Extended Data Fig. 2). We also isolated, sorted and sequenced parasite-infected enterocytes from *Ifng*[−/−] mice. In total, we sequenced 5,879 parasites from culture and 3,219 parasites from mice passing quality control (Supplementary Table 1) and detected a median of 351 and 718 genes per parasite, for in vitro or in vivo samples, respectively (Extended Data Fig. 2).

Just-in-time expression of parasite components

We first sought to understand intracellular development in the asexual phase of the life cycle. *Cryptosporidium* replicates through a 12 h merogony cycle[4]: intracellular establishment is followed by growth, three nuclear divisions, the assembly of eight motile merozoites, merozoite egress and reinvasion[4] (Fig. 1g). To resolve the underlying transcriptional program, we performed nearest-neighbour graph-based clustering analysis of cells collected at 24 and 36 h, two timepoints at which parasites replicate exclusively asexually in culture[8]. Visualization using uniform manifold approximation and projection (UMAP) placed the parasite transcriptomes into a circle with nine distinct identity classes

[1]Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Biomedical Sciences, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA. ✉e-mail: striepen@upenn.edu
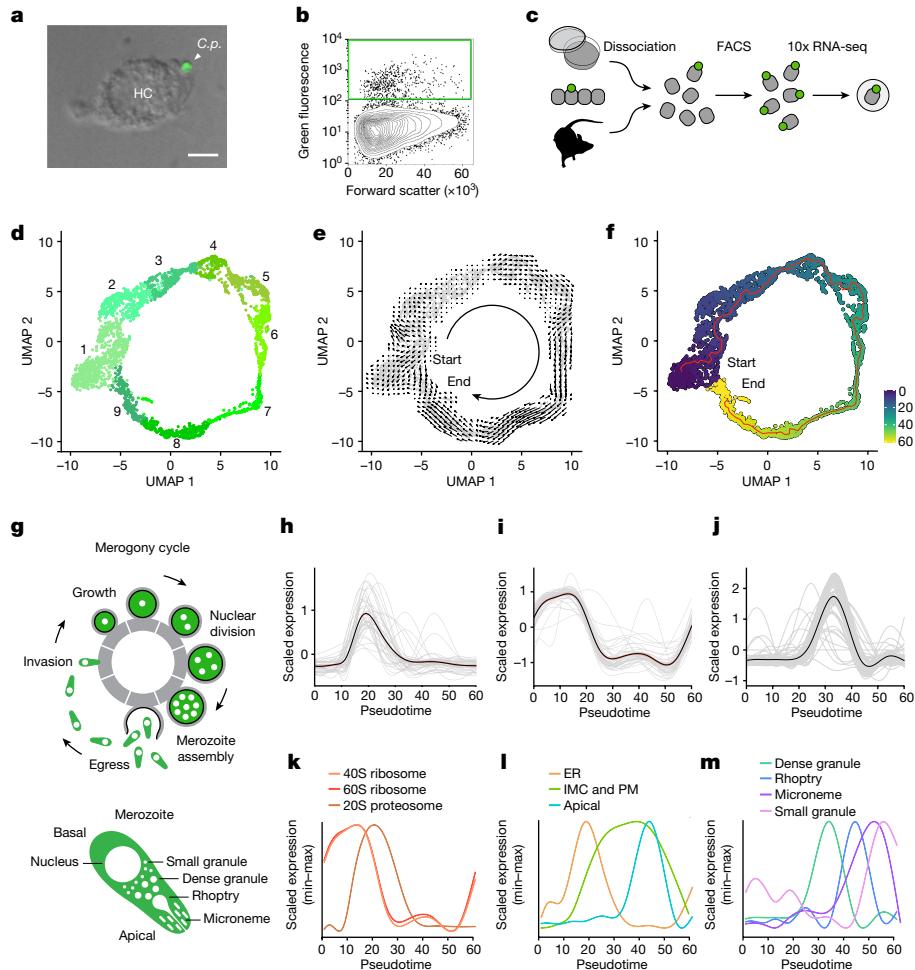
Fig. 1 | The transcriptional program of the asexual cycle. a, Representative image of an engineered green fluorescent *C. parvum* (*C.p.*) parasite shown within its intestinal host cell (HC). Scale bar, 5 μm. **b**, Infected host cells were isolated by cell sorting based on parasite fluorescence; the sorting gate is indicated by the green box. **c**, Outline of the collection, enrichment and encapsulation of cells infected with *C. parvum*. An element of the diagram was adapted from ref. 39, Springer Nature Limited. **d**–**f**, UMAP plots of 2,989 *C. parvum* single-cell transcriptomes representing nine distinct clusters (**d**), the direction of transcriptional change (**e**) and pseudotime (**f**). The start and end of the asexual cycle are noted. **g**, Schematic highlighting the asexual merogony cycle and the merozoite, which contains four distinct secretory organelles discharged during invasion. **h**–**j**, Scaled expression of genes across pseudotime: DNA replication (**h**; 39 proteins), 60S ribosome (**i**; 42 proteins) and dense granules (**j**; 79 proteins). Individual genes (grey), the mean (black) and the 95% confidence interval (red) are shown. **k**–**m**, The mean-scaled gene expression of organellar proteins plotted across pseudotime on an individually normalized min–max scale. Plots represent components for synthesis and growth (**k**), assembly (**l**) and secretion (**m**). ER, endoplasmic reticulum; IMC, inner-membrane complex; PM, plasma membrane.

(Fig. 1d). We used RNA velocity[13] to determine the direction of transcriptional change, and to demarcate parasite invasion and egress (genes with introns (12–16%) are evenly distributed across clusters). This revealed clockwise progression (Fig. 1e), and we interpreted the point of arrow inflection between clusters 9 and 1 as the start and end of the cycle. Applying a pseudotime trajectory, we placed individual parasite transcriptomes onto a timeline (Fig. 1f), enabling us to plot the temporal expression of all detected genes across intracellular development. When we plotted the expression of genes that are required for DNA replication (Fig. 1h), we noted a coordinated wave matching the observation of parasite mitoses visualized by live-cell microscopy[4]. To overcome the limited functional annotation of the *Cryptosporidium* genome, we used a spatial proteomic dataset that established the subcellular localization of 1,107 proteins[14] to generate biologically meaningful categories. Figure 1i,j shows mRNA abundance plots for proteins experimentally assigned to the 60S ribosome and dense granules as examples (all ten organellar plots and a heat map are shown in Extended Data Fig. 3, and gene lists are provided in Supplementary Table 2). This revealed a high degree of compartmental co-expression, and we

observed a notable temporal program that orchestrates intracellular development and the assembly of new infectious forms (Fig. 1k–m). Newly invaded trophozoites transcribed genes encoding the ribosome immediately followed by the proteasome, therefore setting the stage for protein synthesis and growth (Fig. 1k). This is followed by the endoplasmic reticulum, various membrane systems and, lastly, the structure of the apical tip of new merozoites that anchors the invasion machinery (Fig. 1l). *Cryptosporidium* host cell invasion relies on multiple secretory organelles to deliver a battery of pathogenesis factors[11,14]. Our transcriptional analysis demonstrated coordinated successive waves for dense granules, rhoptries, micronemes and small granules (Fig. 1m), suggesting a temporally organized secretory pathway resulting in the assembly of a highly complex infection system. In turn, clustering data (Fig. 1d and Supplementary Table 2) can now be used to associate uncharacterized proteins with organelles and pathways and provide further insights into their probable biological function. Not all groups derived from spatial proteomics[14] form a single sharp expression peak, and we point to the membrane/inner-membrane complex and endoplasmic reticulum groups as examples.
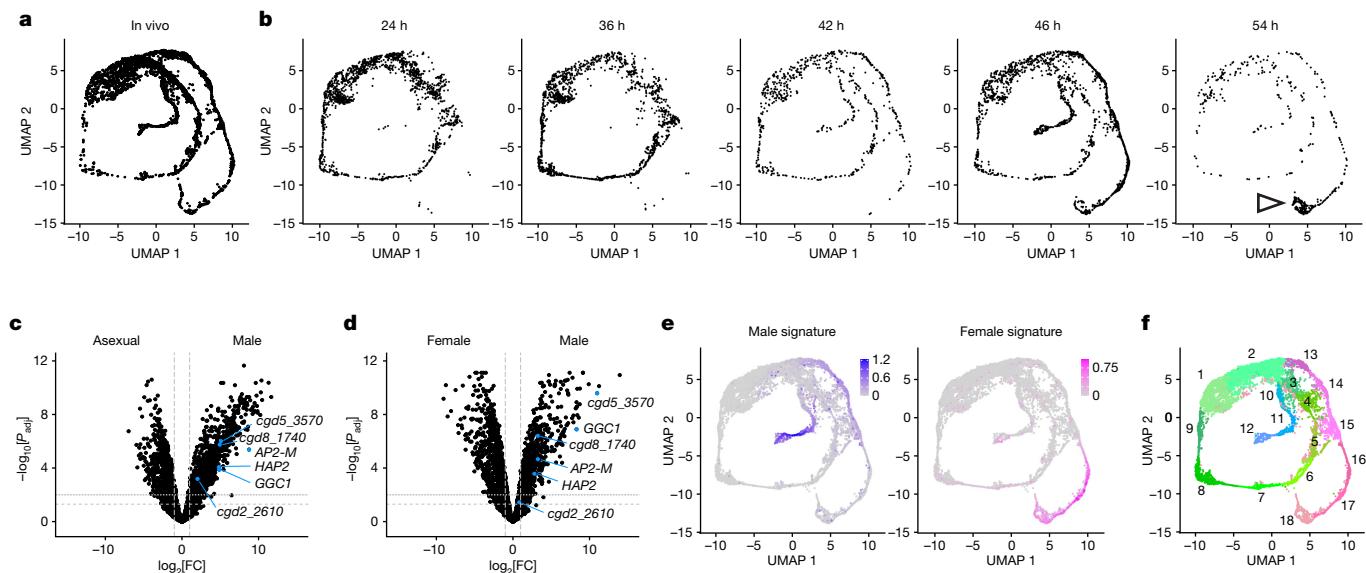
Fig. 2 | The transcriptome of the full *C. parvum* life cycle. a,b, UMAP plots of single-cell transcriptomes of *C. parvum* obtained from an infected mouse (**a**) or from infected HCT-8 cultures sampled at the indicated time (**b**). **c,d**, Differentially expressed genes when comparing bulk RNA samples of asexual and male parasites from culture (**c**) or female parasites from mice and male parasites from culture (**d**). $n = 4$ biological replicates per group. Each point represents a single *C. parvum* gene. The horizontal dashed lines show a false-discovery rate (FDR) of 0.05 (longer dashes) or 0.01 (shorter dashes), and the vertical dashed lines indicate a $\log_2$[fold change (FC)] of −1 and 1. **e**, UMAP analysis of the entire *C. parvum* life cycle, showing male and female gene signatures. Bulk comparisons across asexual, male and female samples generated 221 and 130 genes significantly upregulated in males and females, respectively ($\log_2$[fold change] ≥ 2, $P < 0.01$). These were used to generate sex-specific gene signatures, which were painted onto the UMAP for male (blue) and female (pink). The UMAP shows 9,098 individual *C. parvum* transcriptomes and represents the total analysis of samples depicted in **a** and **b**. **f**, UMAP projection coloured for clustering across asexual (green), male (blue) and female (pink) parasites.

Only three intracellular *Cryptosporidium* stages

The *Cryptosporidium* life cycle most widely shown proposes a type-1 meront that produces eight merozoites and is responsible for asexual replication, and a distinct type-2 meront yielding four merozoites preadapted to sexual development[6,15]. However, the existence of the type-2 meront has been challenged recently[4,8]. Figure 2a,b shows UMAP plots of parasite transcriptomes derived from cells isolated from the intestine of infected mice or collected from cultures at different timepoints. We observed the circular pattern of parasite transcriptomes after asexual development and noted the appearance of two paths of differing length protruding from the same point on the circle. In culture, the two additional transcriptional programmes were first observed at 42 h and became prominent at 46 h. This timing matched that observed for the appearance of gametes in culture[4,8,16].

To assign developmental fates, we used bulk transcriptome samples derived from infected host cells carrying parasites enriched for specific stages. Data for asexual type-1 meronts and female gametes were already available[8]. To obtain a male-enriched transcriptome, we infected cultures with parasites engineered to express a fluorescent reporter driven by the promoter of the recently identified male transcription factor AP2-M[17] and analysed the sorted cells using RNA-seq. Differential expression analysis of this transcriptome with asexual or female datasets showed enrichment for the two established male markers HAP2 and AP2-M (Fig. 2c,d and Supplementary Table 3) and yielded lists of transcripts associated with the male or female transcriptome (Supplementary Table 4; $P < 0.01$). Expression of these gene sets was painted onto the single-cell atlas (Fig. 2e) and identified male and female fate with high confidence. Both fates diverge from the asexual type-1 meront circle after cluster 2. The entire life cycle transcriptome consists of 18 clusters and 2,880 differentially expressed genes between these clusters (Fig. 2f, Extended Data Fig. 4 and Supplementary Table 5). We also analysed the expression of parasite non-coding RNAs[18] in our dataset, and found numerous examples of stage-specific expression

(Extended Data Fig. 5 and Supplementary Tables 6 and 7). Note that fertilization occurs in vivo but is blocked in culture[8]. This phenomenon was reflected in our scRNA-seq data, with females from in vitro culture accumulating into a terminal stage at 54 h (Fig. 2b (arrowhead)), whereas, in vivo, females rejoined the life cycle (Fig. 2a). This probably indicates the presence of a transcriptional checkpoint that responds to fertilization. We further noted that differences between the 24 h and 36 h datasets are modest in number and amplitude and limited to housekeeping genes highly expressed at both timepoints. Differential gene expression analysis did not detect transcripts uniquely associated with distinct types of meronts (Supplementary Table 8). Overall, the transcriptome data argue for the direct development of gametes from type-1 meronts as recently observed by time-lapse microscopy[4]. This model is consistent with the simple three-stage life cycle proposed in the initial description of the parasite[19] in contrast to the two-step model that prevails in the current literature[15].

Female gametes prepare for transmission

The female gamete grows rapidly, while maintaining a single nucleus throughout its development[4] (Fig. 3a). Previous studies found that females express genes related to carbohydrate metabolism, meiosis and oocyst wall formation[8]. Analysis at the single-cell level revealed six stage-specific clusters (Fig. 3b), with the first three not observed in our previous bulk RNA-seq analysis due to use of a late female reporter (Fig. 2e). In total, 773 genes were enriched in the female transcriptome ($P < 1 \times 10^{-50}$), of which 187 were unique to females (Supplementary Table 9), and pseudotime analysis established their temporal order along the female developmental trajectory (Fig. 3c). Females build the *Cryptosporidium* oocyst wall, which is a multilayered structure consisting of an external glycocalyx, a rigid bilayer of acid-fast lipids and an inner layer of proteins[20]. It is responsible for the parasite's considerable resistance to water chlorination; however, how it is assembled is unclear. We found that the oocyst proteome[14] was expressed in
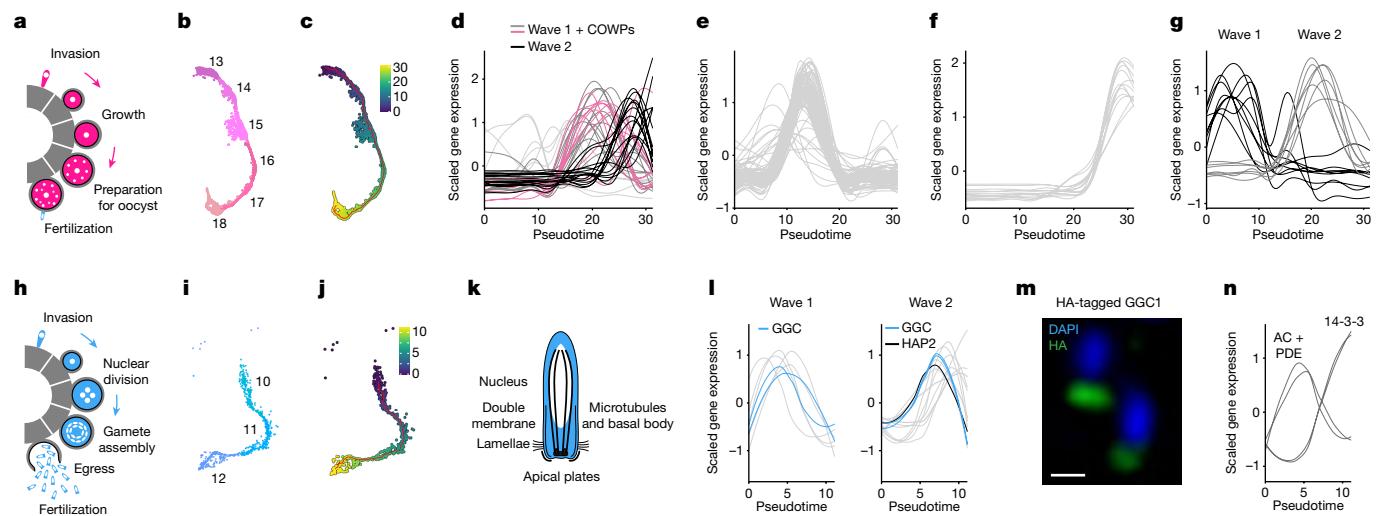
Fig. 3 | Female and male gametes use sex-specific gene sets. a, Schematic of female development. Wall-forming bodies are shown as smaller white or grey vesicles. **b,c**, UMAP plots of the female trajectory (2,136 single-cell transcriptomes). Development progresses through six transcriptionally distinct phases (**b**; clusters 13–18) with individual parasite transcriptomes ordered by pseudotime (**c**). **d**–**g**, Scaled expression of genes across pseudotime encoding components of the oocyst wall (44 proteins; **d**), dense granules (79 proteins; **e**) and crystalloid body (13 proteins; **f**), established by spatial proteomic analysis, as well as genes for secretory and membrane proteins (15 proteins; **g**). Each line represents a single gene (COWP-family genes are highlighted in pink and genes encoding proteins of two waves of oocyst components are shown in dark grey and black) (**d**). The expression of candidate female surface receptors also falls into two waves (**g**). **h**, Schematic of male development. Male gamonts undergo four nuclear divisions before assembling

16 gametes for release. **i,j**, UMAP plots of the male trajectory, encompassing 731 single-cell transcriptomes in three stage-specific clusters (**i**; clusters 10–12) and ordered by pseudotime (**j**). **k**, Diagram of the male gamete. The complex apical end contains the presumptive fertilization machinery, the components of which remain largely unclear. **l**, Scaled expression of male-specific secretory and membrane proteins (20 proteins) across pseudotime, which fall into two transcriptional waves. Genes from the GGC protein family are labelled in blue and HAP2 is shown in black. **m**, The GGC protein encoded by *cgd7_5500* was epitope-tagged and visualized using an immunofluorescence assay at the apical end of the male gamete. HA is shown in green and DAPI is shown in blue. Scale bar, 0.5 μm. **n**, Scaled expression of key male-exclusive signalling genes across pseudotime: adenylate cyclase (AC), 3′5′-cyclic nucleotide phosphodiesterase (PDE) and two 14-3-3 proteins. Lists of all of the genes shown are provided in Supplementary Tables 9 and 10.

two waves (Fig. 3d), potentially reflecting the two distinct classes of wall-forming bodies observed by electron microscopy[6]. The first wave was dominated by members of the *Cryptosporidium* oocyst wall proteins family (COWPs)[20–22] and further included multiple female-specific proteases and oxidoreductases, with probable roles in maturation and cross-linking of the proteinaceous portion of the wall[23,24]. The second wave of the wall proteome contained a secreted glycosyltransferase, multiple uncharacterized mucins and a C-type lectin suggesting it to yield the glycoconjugate-rich portion of the wall. The female transcriptome further included waves for the assembly of dense granules (Fig. 3e) and the crystalloid body (Fig. 3f), which is essential for oocyst formation[17], and also revealed meiosis-specific genes (Extended Data Fig. 6). How the female attracts and interacts with the male gamete is unclear. We note two waves of transcription, one preceding and one coinciding with the wall proteome. These included transporters, signalling and membrane fusion components, as well as numerous yet-to-be-characterized secretory proteins and receptors that may be critical for gamete interaction (Fig. 3g).

Candidates for the male fertilization machinery

In *Cryptosporidium*, intracellular male stages known as gamonts assemble and then release 16 gametes into the intestinal lumen (Fig. 3h). How these gametes find and fuse with their female counterpart hidden within a host cell is unclear. They lack flagella or amoeboid movement, and our analysis showed them to also lack the gliding and invasion machinery used by asexual stages (Extended Data Fig. 7a–c), suggesting a divergent motility and fertilization mechanism[11]. The male path of the single-cell atlas revealed three distinct expression clusters (Fig. 3i) with a total of 389 genes with an enrichment $P < 1 \times 10^{-50}$, including at least 51 exclusively male genes (Supplementary Table 10). Only two

male proteins were previously characterized in *Cryptosporidium*—the transcription factor AP2-M[17] and HAP2, which has a conserved role in gamete fusion[8]. The tiny (1.2 × 0.4 μm) male gamete contains a condensed bullet-shaped nucleus and a complex apical apparatus[8,14] that engages the female-containing cell (Fig. 3k). HAP2 localizes to this apical end and is the only known component of this fertilization machinery[8]. Its expression falls into the mid phase of male development, which we found is characterized by two waves of secretory and membrane proteins (Fig. 3l), including four paralogous GGC[25] proteins (Extended Data Fig. 7d). To test whether these proteins like HAP2 are part of the apical apparatus, we introduced an epitope tag into the locus of *cgd7_5500* (encoding GGC1; Extended Data Fig. 1b). Immunofluorescence studies found that the protein was expressed exclusively in males and labelled a bar-like structure likely at the apex of male gametes, the presumptive male adhesion site (Fig. 3m and Extended Data Fig. 7e). Among the non-secretory proteins expressed at this time, we noted a PAN/Apple-domain protein, an Armadillo protein and a calcium-regulated microtubule-binding protein. Importantly, we also identified a male-specific membrane-associated adenylate cyclase and a corresponding 3′5′-cyclic nucleotide phosphodiesterase (Fig. 3n). Cyclic AMP signalling triggers gamete fusion after contact in the unicellular alga *Chlamydomonas*[26], in which HAP2 is essential for gamete fusion[27,28]. Genes encoding signalling pathways are the most prominent late components of the male transcriptome. Two male-specific 14-3-3 proteins are the most highly induced genes (Fig. 3n), followed by multiple serine/threonine protein kinases, two cyclins, a CDC2-like protein kinase, CDPK5 and a papain and a cysteine protease. Lastly, multiple proteins are expressed in both male and female gametes, but not in the asexual forms, and therefore probably function in sex. This includes a zinc-finger RNA-binding protein, a START-domain lipid transporter, a PH-domain lipid-binding protein and the membrane-fusion factor synaptobrevin.
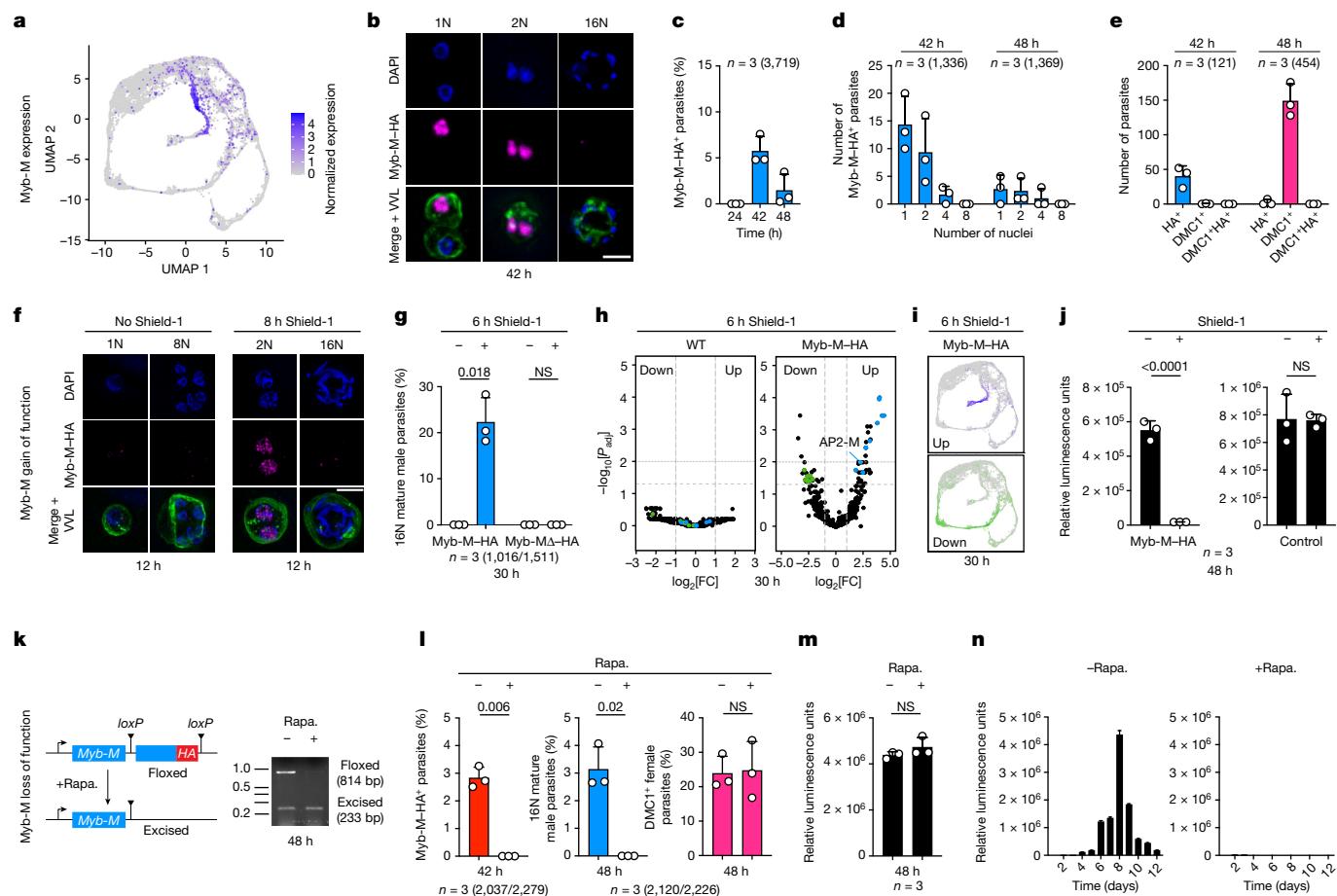
Fig. 4 | The transcription factor Myb-M commits parasites to a male fate.
a, UMAP showing *cgd6_2250* (encoding Myb-M) expression. **b**, Myb-M was tagged with the HA epitope and visualized using immunofluorescence after 42 h in culture (parasites were counterstained with *Vicia villosa* lectin (VVL)). Scale bar, 2 µm. **c**,**d**, HCT-8 cultures were infected with Myb-M–HA parasites, fixed at 24, 42 and 48 h, and the HA positivity (**c**) and staging of HA positive parasites (**d**) were assessed. **e**, Cultures that were infected with Myb-M–HA parasites were fixed at 42 and 48 h and were stained for HA and the female protein DMC1[16] (pink). **f**, Cultures were infected with inducible Myb-M parasites and treated with vehicle or Shield-1 at 4 h and fixed at 12 h. **g**, Inducible Myb-M–HA and Myb-M-domain-mutant (Myb-MΔ–HA) parasites were treated with Shield-1 and scored for mature male parasites. **h**, The transcriptional response of WT and inducible Myb-M–HA to Shield-1 (horizontal lines show FDR of 0.05 or 0.01, vertical lines show the log$_2$[fold change]). Asexual and male transcripts are highlighted in green and blue. **i**, Gene signatures derived from significantly

changed genes in **h** were mapped onto the transcriptional atlas. **j**, Cultures that were infected with inducible Myb-M or control parasites were treated at 24 h, and growth was measured at 48 h. **k**, Parasites were constructed to carry a floxed *Myb-M* gene and a rapamycin (Rapa.)-inducible Cre recombinase. Rapamycin induces excision (note some excision before treatment[17]; the source gel is shown in Supplementary Fig. 1). **l**, Rapamycin-treated or untreated *Myb-M^loxP DiCre* cultures were scored for Myb-M–HA-positive, 16N mature male, or DMC1-positive female parasites. **m**, Growth assay of Myb-M mutants with and without rapamycin. **n**, Groups of three *Ifng^−/−* mice were infected with Myb-M mutants, treated with vehicle or rapamycin, and oocyst shedding was measured using a luciferase assay. One out of two replicates is shown; *n* = 12 mice in total. Data are mean ± s.d. of the indicated number of biological replicates, and the total number of cells scored is shown in parentheses. Significance was evaluated using unpaired *t*-tests (**j**, **l** pink and **m**) or Welch's *t*-tests (**g** and **l** red and blue).

Myb-M is an early male nuclear protein

Sex determination in Apicomplexa does not rely on sex chromosomes. In *Plasmodium*, sexual commitment and sex determination are two independent fate decisions, with commitment governed by the transcription factor AP2-G[29,30] and determination controlled by a bistable switch at the *md1* locus[31,32]. However, *Cryptosporidium* does not have a morphologically[4] or transcriptionally (Fig. 2b,f) identifiable sexual progenitor, and Md1, which associates with cytoplasmic ribonucleic acid granule-like structures, is not conserved[32]. *Cryptosporidium* gametes do express sex-specific AP2 factors[8,17] and presumptive DNA-binding motifs have been previously identified[33]. We define two male and three female factors (Extended Data Figs. 8 and 9a) and identify cluster enrichment patterns for some of their associated motifs (Supplementary Table 11). However, our analysis shows that all of them are transcribed only after fate commitment. By contrast, *cgd6_2250* (encoding Myb-M) is the earliest and most significantly upregulated male gene

(Fig. 4a). Myb transcription factors control cell fate in plant development[34,35] and the switch from acute to chronic stage in *Toxoplasma*[36]. We therefore hypothesized that this protein may act as a transcriptional activator committing parasites to maleness. We tagged Myb-M with an epitope and imaged cultures infected with transgenic organisms (Fig. 4b). Fluorescent parasites were observed at 42 h and 48 h but not at 24 h (Fig. 4c). Labelling was nuclear localized and found largely in intracellular parasites with one or two nuclei (Fig. 4b,d). Female parasites marked by DMC1 as well as mature male gametes (16 nuclei) showed no staining (Fig. 4b,e). We conclude that Myb-M is expressed in early male gamonts and degraded before their full maturation.

Myb-M expression induces maleness

To understand the function of Myb-M, we introduced a second copy of the gene for gain-of-function studies. Myb-M transcription was driven by a promoter expressed in all stages of the life cycle, but the protein

was degraded due to a fused destabilization domain[37,38]. Cultures were infected with this strain, treated with vehicle or the stabilizer Shield-1, and fixed 12 h after infection. Shield-1, but not vehicle treatment, resulted in nuclear Myb-M staining (Fig. 4f) and, importantly, we observed mature male gamonts with 16 bullet-shaped nuclei at a timepoint that usually lacks sexual stages[4,8,16]. To quantify the induction of maleness by Myb-M, we engineered a reporter strain in which the late male transcription factor AP2-M[17] carried a MYC epitope (Extended Data Fig. 10). Parasites were grown for 18 h, the last six with or without Shield-1. Shield-1 induced MYC labelling and, when we scored 8N parasites, the stage in which AP2-M is detectable[17], we found more than 60% of parasites positive for MYC, in contrast to none in the untreated samples (Extended Data Fig. 10; $P = 0.0018$), demonstrating robust male commitment through Myb-M induction. We obtained similar results when treating for 4 h beginning at 24 h ($P = 0.0044$). We also constructed a control strain lacking the DNA-binding domain and quantified 16N mature male parasites. Mature males (around 20% of all cells) are observed only after Shield-1 treatment, and this depends on the Myb domain (Fig. 4g and Extended Data Fig. 10f; both proteins were expressed at comparable level and nuclear targeted).

To understand the relationship between Myb-M induction and stage-specific gene expression, we measured the transcript level for exemplary genes that are exclusively expressed in asexual, male or female parasites by quantitative PCR. Treatment resulted in enhanced transcription of male genes at the expense of asexual (or female) genes depending on timing (Extended Data Fig. 10g,h). We also conducted bulk RNA-seq analysis at 30 h when wild-type (WT) parasite cultures are still exclusively asexual. Induction of Myb-M resulted in significant gene expression changes, whereas WT control cultures remained unaffected (Supplementary Table 12). When we derived signatures of the changed genes and mapped them onto the single-cell atlas, we found that those that were induced were male genes, and those that were repressed were asexual genes (Fig. 4h,i). *Cryptosporidium* proliferation depends on asexual amplification through merogony. We therefore reasoned that diversion into premature and exclusive maleness may impact parasite growth. To test this, we treated cultures infected with the inducible Myb-M strain with vehicle or Shield-1 24 h after infection and measured growth using a luciferase assay at 48 h. This resulted in a profound block of growth (Fig. 4j; $P < 0.0001$). Note that Shield-1 is not toxic to *Cryptosporidium*, as a similar experiment with a control strain showed no impact.

Loss of Myb-M results in loss of males

Attempts to ablate the Myb-M gene did not recover viable transgenics (Extended Data Figs. 1d and 9b,c). *Cryptosporidium* transgenesis relies on the recovery of oocysts from infected mice[39], and mutants that have lost the ability to form oocysts through sex will be lost. We therefore examined several conditional gene-ablation strategies[17,40,41]. We ultimately found success by flanking a critical portion of the Myb-M gene with loxP sites followed by sexual cross[42] with parasites engineered to express a rapamycin-inducible split Cre recombinase (DiCre). Rapamycin treatment of cultures infected with this parasite strain resulted in efficient gene excision (Fig. 4k) and loss of the protein as monitored by immunofluorescence assay at 42 h (Fig. 4l). We next scored gametes at 48 h. Rapamycin treatment led to a complete loss of males, whereas female gametes were unaffected (Fig. 4l). As discussed, males do not contribute to in vitro growth and, in contrast to premature male induction, we did not observe a growth defect after male ablation (Fig. 4m). However, in contrast to Shield-1, rapamycin can be readily used in infected mice[17], providing the opportunity to test the importance of Myb-M in vivo. Rapamycin treatment of infected mice ablated oocyst production as quantified by a faecal luciferase assay[39] (Fig. 4n; rapamycin alone does not inhibit oocyst shedding in control strains[17,42]).

In summary, the *C. parvum* single-cell transcriptome revealed a highly ordered program for intracellular development and life-cycle progression, with a direct conversion to gametes from the asexual cycle. We found important roles for transcription factors in sex choice and for specific downstream aspects of gamete function[17] but, notably, not in the transition from asexual to sexual cycle. This may hint at an epigenetic switch, or a mechanism based on the decay or accumulation of a protein, RNA or metabolite. Sex is essential to continued infection and transmission and is therefore an attractive target for prevention and intervention. The single-cell atlas presented here provides a comprehensive list of the components of male and female gametes to further dissect the fundamental biology of sex in *Cryptosporidium*. It also represents a powerful resource for the ongoing effort to identify antigens for vaccines and targets for drug treatment, respectively.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07466-1.

1. Kotloff, K. L. et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* **382**, 209–222 (2013).
2. Khalil, I. A. et al. Morbidity, mortality, and long-term consequences associated with diarrhoea from *Cryptosporidium* infection in children younger than 5 years: a meta-analyses study. *Lancet Glob. Health* **6**, e758–e768 (2018).
3. Checkley, W. et al. A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for *Cryptosporidium*. *Lancet Infect. Dis.* **15**, 85–94 (2015).
4. English, E. D., Guerin, A., Tandel, J. & Striepen, B. Live imaging of the *Cryptosporidium parvum* life cycle reveals direct development of male and female gametes from type I meronts. *PLoS Biol.* **20**, e3001604 (2022).
5. Striepen, B. Parasitic infections: time to tackle cryptosporidiosis. *Nature* **503**, 189–191 (2013).
6. Current, W. L. & Reese, N. C. A comparison of endogenous development of three isolates of *Cryptosporidium* in suckling mice. *J. Protozool.* **33**, 98–108 (1986).
7. Gharpure, R. et al. *Cryptosporidiosis* Outbreaks—United States, 2009–2017. *MMWR Morb. Mortal. Wkly Rep.* **68**, 568–572 (2019).
8. Tandel, J. et al. Life cycle progression and sexual development of the apicomplexan parasite *Cryptosporidium parvum*. *Nat. Microbiol.* **4**, 2226–2236 (2019).
9. Huang, W. et al. Multiple introductions and recombination events underlie the emergence of a hyper-transmissible *Cryptosporidium hominis* subtype in the USA. *Cell Host Microbe* **31**, 112–123 (2023).
10. Nader, J. L. et al. Evolutionary genomics of anthroponosis in *Cryptosporidium*. *Nat. Microbiol.* **4**, 826–836 (2019).
11. Guerin, A. & Striepen, B. The biology of the intestinal intracellular parasite *Cryptosporidium*. *Cell Host Microbe* **28**, 509–515 (2020).
12. Kissinger, J. C., Hermetz, K. E., Woods, K. M. & Upton, S. J. Enrichment of *Cryptosporidium parvum* from in vitro culture as measured by total RNA and subsequent sequence analysis. *Mol. Biochem. Parasitol.* **220**, 5–9 (2018).
13. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
14. Guerin, A. et al. Cryptosporidium uses multiple distinct secretory organelles to interact with and modify its host cell. *Cell Host Microbe* **31**, 650–664 (2023).
15. Vetterling, J. M., Jervis, H. R., Merrill, T. G. & Sprinz, H. *Cryptosporidium wrairi* sp. n. from the guinea pig *Cavia porcellus*, with an emendation of the genus. *J. Protozool.* **18**, 243–247 (1971).
16. Jumani, R. S. et al. A suite of phenotypic assays to ensure pipeline diversity when prioritizing drug-like *Cryptosporidium* growth inhibitors. *Nat. Commun.* **10**, 1862 (2019).
17. Tandel, J. et al. Genetic ablation of a female-specific Apetala 2 transcription factor blocks oocyst shedding in *Cryptosporidium parvum*. *mBio* **14**, e0326122 (2023).
18. Li, Y., Baptista, R. P., Sateriale, A., Striepen, B. & Kissinger, J. C. Analysis of long non-coding RNA in *Cryptosporidium parvum* reveals significant stage-specific antisense transcription. *Front. Cell Infect. Microbiol.* **10**, 608298 (2021).
19. Tyzzer, E. E. An extracellular Coccidium, *Cryptosporidium Muris* (Gen. Et Sp. Nov.), of the gastric glands of the common mouse. *J. Med. Res.* **23**, 487–510 (1910).
20. Samuelson, J., Bushkin, G. G., Chatterjee, A. & Robbins, P. W. Strategies to discover the structural components of cyst and oocyst walls. *Eukaryot. Cell* **12**, 1578–1587 (2013).
21. Templeton, T. J. et al. The *Cryptosporidium* oocyst wall protein is a member of a multigene family and has a homolog in *Toxoplasma*. *Infect. Immun.* **72**, 980–987 (2004).
22. Spano, F., Puri, C., Ranucci, L., Putignani, L. & Crisanti, A. Cloning of the entire COWP gene of *Cryptosporidium parvum* and ultrastructural localization of the protein during sexual parasite development. *Parasitology* **114**, 427–437 (1997).
23. Katrib, M. et al. Stage-specific expression of protease genes in the apicomplexan parasite, *Eimeria tenella*. *BMC Genomics* **13**, 685 (2012).

Article

24. Belli, S. I., Wallach, M. G., Luxford, C., Davies, M. J. & Smith, N. C. Roles of tyrosine-rich precursor glycoproteins and dityrosine- and 3,4-dihydroxyphenylalanine-mediated protein cross-linking in development of the oocyst wall in the coccidian parasite *Eimeria maxima*. *Eukaryot. Cell* **2**, 456–464 (2003).
25. Abrahamsen, M. S. et al. Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* **304**, 441–445 (2004).
26. Pasquale, S. M. & Goodenough, U. W. Cyclic AMP functions as a primary sexual signal in gametes of *Chlamydomonas Reinhardtii*. *J. Cell Biol*. **105**, 2279–2292 (1987).
27. Liu, Y. et al. The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in *Chlamydomonas* and *Plasmodium* gametes. *Genes Dev*. **22**, 1051–1068 (2008).
28. Snell, W. J. Uncovering an ancestral green menage a trois: contributions of *Chlamydomonas* to the discovery of a broadly conserved triad of plant fertilization proteins. *Curr. Opin. Plant Biol*. **69**, 102275 (2022).
29. Kafsack, B. F. et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248–252 (2014).
30. Sinha, A. et al. A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature* **507**, 253–257 (2014).
31. Russell, A. J. C. et al. Regulators of male and female sexual development are critical for the transmission of a malaria parasite. *Cell Host Microbe* **31**, 305–319 (2023).
32. Gomes, A. R. et al. A transcriptional switch controls sex determination in *Plasmodium falciparum*. *Nature* **612**, 528–533 (2022).
33. Oberstaller, J., Joseph, S. J. & Kissinger, J. C. Genome-wide upstream motif analysis of *Cryptosporidium parvum* genes clustered by expression profile. *BMC Genomics* **14**, 516 (2013).
34. Brownfield, L. et al. A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genet*. **5**, e1000430 (2009).
35. Dubos, C. et al. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci*. **15**, 573–581 (2010).
36. Waldman, B. S. et al. Identification of a master regulator of differentiation in *Toxoplasma*. *Cell* **180**, 359–372 (2020).
37. Banaszynski, L. A., Chen, L. C., Maynard-Smith, L. A., Ooi, A. G. & Wandless, T. J. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* **126**, 995–1004 (2006).
38. Herm-Götz, A. et al. Rapid control of protein level in the apicomplexan *Toxoplasma gondii*. *Nat. Methods* **4**, 1003–1005 (2007).
39. Vinayak, S. et al. Genetic modification of the diarrhoeal pathogen *Cryptosporidium parvum*. *Nature* **523**, 477–480 (2015).
40. Choudhary, H. H., Nava, M. G., Gartlan, B. E., Rose, S. & Vinayak, S. A conditional protein degradation system to study essential gene function in *Cryptosporidium parvum*. *mBio* **11**, e01231-20 (2020).
41. Xu, R., Beatty, W. L., Greigert, V., Witola, W. H. & Sibley, L. D. Multiple pathways for glucose phosphate transport and utilization support growth of *Cryptosporidium parvum*. *Nat. Commun*. **15**, 380 (2024).
42. Shaw, S. et al. Genetic crosses within and between species of *Cryptosporidium*. *Proc. Natl Acad. Sci. USA* **121**, e2313210120 (2024).

Methods

Animal ethics statement
All animal experimentation was approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania (protocol 806292).

Mice
Ifng[-/-] mice (002287) were purchased from Jackson Laboratory and maintained as a breeding colony at the University of Pennsylvania. Mice were housed under a 12 h–12 h dark–light cycle, at a temperature of 65–73 °F, and a humidity level of between 30% and 40%. Mice used for experiments were aged from 5 to 8 weeks. Both male and female mice were used to generate and propagate *C. parvum* strains and did not exhibit a difference in oocyst shedding. Mice were initially chosen at random but were sex and age matched for each experiment. Experiments were not performed in a blinded manner.

Maintenance of cell lines and parasite strains for infection
Human colorectal adenocarcinoma HCT-8 cells (ATCC: CCL-224TM) were maintained in RPMI-1640 supplemented with 10% fetal bovine serum (FBS), 1× penicillin–streptomycin, 1× gentamicin and 1× fungizone at 37 °C and 5% CO_2. Before infection, the medium was switched to complete RPMI-1640 containing 1% FBS (infection medium). *C. parvum* oocysts were purchased from Bunchgrass Farms and are of the Iowa II strain background, a IIa subtype. For in vitro infection, oocysts were incubated in a 1:4 bleach:PBS solution for 5 min at 4 °C, centrifuged at 16,000*g* and washed three times in PBS[43]. Oocysts were resuspended in infection medium then added directly to HCT-8 cells. For most experiments with Myb-M overexpression (those with induction after the first cycle) and all experiments with the *Myb-M*[loxP]*DiCre*, excystation was performed after washing oocysts[43]. In brief, oocysts were treated with 0.2 mM sodium taurodeoxycholate for 10 min at 15 °C, centrifuged at 16,000*g*, resuspended in PBS and incubated for 1 h at 37 °C before pelleting, resuspending in infection medium and adding to HCT-8 cells. Note that HCT-8 cells were not authenticated and were not tested for mycoplasma.

Plasmid construction
Plasmids for genetic manipulation using CRISPR–Cas9 were generated as previously described[39,43]. In brief, annealed oligonucleotides (Sigma-Aldrich), designed using the CRISPR guide design tool on VEuPathDB[44], were ligated into the BbsI-digested *C. parvum* Cas9/U6 plasmid. The Cas9/U6 plasmids with the thymidine kinase guide RNA (g1) or AP2-M guide RNA (g6) were previously published[8,17]. Plasmids encoding the homology repair regions were constructed by Gibson assembly using the NEB Gibson Assembly Master Mix (New England Biolabs). The linear homology repair templates were generated by PCR using the high-fidelity PrimeSTAR Max DNA Polymerase (R045A, Takara), with homology arms ranging in length from 35 to 50 bp. A list of primers is provided in Supplementary Table 13.

Generation of transgenic parasites
Transgenic parasites were generated as previously described[39,43,45]. In brief, 2.5×10^7 *C. parvum* oocysts were bleached and washed before excystation with 0.2 mM sodium taurodeoxycholate for 10 min at 15 °C followed by incubation for 1 h at 37 °C to obtain sporozoites. Sporozoites were resuspended in SF transfection buffer containing 50 μg each of CRISPR plasmid and PCR repair template. Sporozoites were nucleofected with an Amaxa 4D nucleofector (Lonza) using the program EH100. Transfected sporozoites were resuspended in cold PBS. *Ifng*[-/-] mice were treated with antibiotics in the 5–7 days before infection, and their stomach acid was neutralized with 100 μl of 8% sodium bicarbonate administered through oral gavage immediately before infection. Mice were infected with 100 μl of transfected sporozoites through oral gavage, and stable transformants were selected with 16 mg ml[-1] paromomycin in drinking water. Parasite shedding was monitored by measuring the nanoluciferase activity in the faeces of infected mice[39]. Oocysts were purified from collected faeces by sucrose flotation followed by a caesium chloride gradient and stored long term in PBS at 4 °C (ref. 43).

Faecal or purified oocyst DNA was extracted using the Quick-DNA Faecal/Soil Microbe Microprep Kit (Zymo Research). This DNA was used as PCR template to evaluate proper homologous recombination of the repair template. Integration primers were designed to anneal outside of the 5' and 3' homology arms with internal primers binding to the integrated region (Extended Data Fig. 1). Control primers were designed to anneal to an adjacent gene.

Genetic cross to generate the loss-of-function mutant
To generate a conditional knockout mutant for *Myb-M*, we crossed the *cgd6_2250*[loxP]*-HA* (Extended Data Fig. 1h) strain with the *cre* driver strain parasite pheRS[r]-Nluc-DiCre[42]. Four *Ifng*[-/-] mice were each infected with 10,000 oocysts of both strains. Treatment with 16 mg ml[-1] paromomycin in drinking water and 10 mg per kg BRD7929 by daily oral gavage began on day 4 of infection. BRD7929 was dissolved in 7:3 PEG-400:5% glucose and administered to mice in 100 μl doses for 7 days. Oocysts purified from days 13 to 16 of infection were passaged by infection of a cage of *Ifng*[-/-] mice under drug selection.

Rapamycin treatment
Rapamycin (Thermo Fisher Scientific) was dissolved in 95% ethanol at a stock concentration of 50 mg ml[-1]. For in vitro experiments, rapamycin was used at a final concentration of 100 nM, using equally diluted vehicle as a control. Both were added at the start of infection. DNA was extracted from infected cultures using the DNeasy Blood and Tissue Kit (Qiagen). This DNA was used as a PCR template to evaluate Cre-mediated excision. For in vivo experiments, rapamycin stock was further diluted in water to deliver 10 mg per kg body weight in a 100 μl daily dose through oral gavage, with treatment beginning 2 days before infection.

Nanoluciferase assays
To measure oocyst shedding in faeces, 20 mg of faecal material was dissolved in 1 ml of faecal lysis buffer and vortexed for 1 min with 10 glass beads. The supernatant was mixed 1:1 with nanoluciferase buffer supplemented with 1:50 nanoluciferase substrate (NanoGlo Luciferase Assay Kit, Promega, N1110) in a white opaque plate and luminescence was measured with a Promega GloMax Plate Reader[43].

Parasite growth in vitro was measured by luminescence assay. HCT-8 cells in a 96-well plate were infected with 10,000 oocysts. For Myb-M gain-of-function studies, vehicle (0.1% ethanol) or Shield-1 (Takara Bio, 632189) was added to a final concentration of 0.5 μM after 24 h. For Myb-M loss-of-function studies, vehicle or 100 nM rapamycin (Thermo Fisher Scientific) was added at the start of infection. At 48 h, infection medium was removed from infected HCT-8 cells and 100 μl of faecal lysis buffer was added to each well. A pipette tip was used to scrape cells, which were pipetted up and down to mix. These cells sat for 5 min at room temperature. Then, 100 μl of supernatant was mixed 1:1 with nanoluciferase buffer supplemented with 1:50 nanoluciferase substrate in a white opaque plate. Luminescence was measured using the Promega GloMax Plate Reader.

Immunofluorescence assay
Timepoints for immunofluorescence assays are noted in figure legends. Infections for super-resolution microscopy were set up in 24-well plates, while time-course infections were set up in 24-well or 96-well plates. For the Myb-M-overexpression experiments, cultures were treated with vehicle (0.1% ethanol) or 0.5 μM Shield-1 (Takara Bio, 632189) at the noted times.

For all experiments, infected HCT-8 cells were washed once in PBS before fixing in 4% paraformaldehyde for 10 min. Cells were washed three times in PBS before permeabilization with 0.25% Triton X-100 for 10 min. Coverslips were blocked with 3% BSA for 30 to 60 min at room temperature. Primary antibodies were then added for 1 h, and for this study included rat anti-HA (1:1,000, Roche, 3F10), VVL-FITC (1:1,000, Vector, FL1231), mouse anti-DMC1 (1:10, a gift from C. Huston[16]), mouse anti-MYC (1:300, Roche or 1:1,000, Cell Signaling Technology, 9B11) and mouse anti-α-tubulin (1:1,000, a gift from J. Gaertig). Coverslips were washed three times in PBS and incubated with a 1:1,000 dilution of secondary Alexa Fluor antibodies (Abcam) for 45 min. These antibodies included Alexa Fluor 488 goat anti-mouse IgG (H+L) (A11029), Alexa Fluor 568 goat anti-mouse IgG (H+L) (A11004), Alexa Fluor 594 goat anti-mouse IgG (H+L) (A11005), Alexa Fluor 488 goat anti-rat IgG (H+L) (A11006) and Alexa Fluor 594 goat anti-rat IgG (H+L) (11007). Coverslips were then washed once in PBS, once in PBS containing 1:10,000 DAPI for 5 min, and then once more in PBS before mounting on slides with Fluorogel (Electron Microscopy Science) or Vectashield (Vector Laboratories). Microscopy for parasite counts was performed on the Leica DM6000 Widefield microscope (Penn Vet Imaging Core) using the ×100 objective and Leica Application Suite X software. Super-resolution microscopy was performed on the GE DeltaVision OMX (Penn Vet Imaging Core) system using the ×60 objective and Acquire SR Acquisition software. Images were processed and analysed using GE softWoRx, Leica Application Suite X and Fiji for Mac OS X software.

Flow sorting of infected cells for sequencing

For in vitro infections used for scRNA-seq, HCT-8 cells in a six-well plate were infected with 500,000 (24 h), 300,000 (36, 42 and 46 h) and 200,000 (54 h) oocysts of tdNeon-RPL22-HA (constitutive tdNeon reporter strain; Extended Data Fig. 1a), with wells set up in triplicate for the 24 h, 36 h, 42 h and 54 h timepoints. To collect infected cells, the infection medium was removed, 500 μl of room temperature PBS was added for one wash and 500 μl of TrypLE Express (Thermo Fisher Scientific) was added for 15 min. Cells were pelleted at 400g for 5 min and washed three times in PBS before passage through a 40 μm filter (BD Biosciences) placed in a 50 ml Falcon tube. These cells were transferred into a 15 ml Falcon tube and pelleted at 400g for 5 min. Cells were resuspended in 750 μl of Live/Dead Aqua (L34957, Invitrogen) made in fluorescence-activated cell sorting (FACS) buffer in a 1:200 dilution from stock and incubated in the dark for 20 min. Cells were pelleted and washed once in 5 ml of PBS before resuspension in 500 μl of FACS buffer for sorting in FACS tubes. Infected cells were sorted on the BD FACS Jazz sorter (Penn Cytomics and Cell Sorting Resource Laboratory) into 500 μl of 40% CCS complete medium, which was further diluted in PBS (up to 1 ml total volume) after sorting. Then, 10 μl of this solution was used for counting on a haemocytometer to determine the total number of cells.

For in vitro infections used for bulk RNA-seq, four replicates of HCT-8 cells in a six-well plate were infected with 300,000 oocysts of AP2-M-tdNeon (male reporter strain)[17]. At 48 h, cells were dissociated as described above. A total of 10,000 positive cells was sorted directly into 350 μl of RLT lysis buffer from the RNeasy minikit (Qiagen). For in vivo infections used for scRNA-seq, three Ifng[−/−] mice were infected with 20,000 oocysts of the tdNeon-RPL22-HA strain (Extended Data Fig. 1a), and parasites were selected with 16 mg ml[−1] paromomycin in the drinking water. On day 6, the parasite burden was checked by nanoluciferase assay of the faecal matter and the mice were euthanized. The small intestines were resected in ice-cold PBS, cut into small pieces and digested as follows. Intestinal pieces (2.5 cm) were stored in 10 ml of cold buffer A (HBSS with 5% FBS and 10 mM HEPES). Intestinal pieces were incubated in 20 ml of prewarmed buffer C (HBSS with 5% FBS, 5 mM EDTA and 1 mM dithiothreitol) for 20 min on a shaker at 37 °C. The intestinal pieces were then transferred into 20 ml of buffer B (HBSS with 2 mM EDTA and 10 mM HEPES) with vigorous shaking by hand. This

step was repeated twice, with mucus and fat layers removed as they separated from the intestinal piece. The supernatants were collected from buffers B and C and were centrifuged at 1,500 rpm for 5 min at 4 °C. The supernatant was discarded, and the pellet was resuspended in 10 ml of buffer A. This cell suspension was filtered through a 70 μm filter (BD Biosciences) placed into a 50 ml Falcon tube. The filter was washed with an additional 5–10 ml of buffer A. The cell suspension was centrifuged at 1,500 rpm for 5 min at 4 °C, the supernatant was discarded, and the pellet was resuspended in 5 ml of buffer A. This cell suspension was filtered through a 40 μm filter (BD Biosciences) placed in a 50 ml Falcon tube. The cell suspension was centrifuged at 1,500 rpm for 5 min at 4 °C, and the supernatant was discarded. Cells were resuspended in 750 μl of Live/Dead Aqua (L34957, Invitrogen) made in FACS buffer and were stained with anti-CD45.2 eFluor 450 (48-0454-82) and anti-EPCAM PE (118206, BioLegend) antibodies. This cell suspension was incubated in the dark for 20 min. Cells were pelleted and washed once in 5 ml of PBS before resuspension in 500 μl of FACS buffer for sorting in FACS tubes. Infected cells were sorted on the BD FACS Jazz sorter into 500 μl of 40% CCS complete medium, which was further diluted in PBS (up to 1 ml total volume) after sorting. One mouse sample was used to set the gates using the BD FACS software and all of the intestinal epithelial cells from the most highly infected mouse, with a nanoluciferase reading of 2.96×10^6, were sorted for sequencing. A total of 10 μl of this solution was used for counting on a haemocytometer to determine the total number of cells. Note that cells may be infected with both a male and a female parasite, or that a female parasite may carry a male gamete as a consequence of fertilization.

Bulk RNA-seq and analysis

For sorted AP2-M-tdNeon-infected cells, RNA was extracted using the RNeasy Mini kit according to the manufacturer's protocol (Qiagen). The Clontech SMART-seq HT kit was used for cDNA synthesis, and the Illumina Nextera XT Library Prep Kit was used to prepare libraries for sequencing. Total RNA and libraries were quality checked and quantified on the Agilent Tapestation 4200 (Agilent Technologies) and Qubit 3 (Thermo Fisher Scientific) systems, respectively. Replicates were pooled and single-end reads were generated on the Illumina NextSeq 500 sequencer.

To sequence Myb-M gain-of-function mutants, HCT-8 24-well cultures were infected with 250,000 oocysts of WT Myb-M or Myb-M–HA-overexpression parasites (Extended Data Fig. 1e), and treated with vehicle or 0.5 μM Shield-1 24 h after infection (each group in three replicates). At 30 h, RNA was extracted using the RNeasy Mini kit and the Illumina stranded mRNA preparation ligation kit (20040534) was used for cDNA generation and library preparation. Replicates were pooled and single-end reads were generated on the Illumina NextSeq 2000 sequencer using the NextSeq 1000/2000 P2 Reagents (100 Cycles) flow cell kit (20046811).

Raw reads were mapped to the C. parvum Iowa II reference (CryptoDB, release 46) using Kallisto (v.0.45.0)[46]. Analyses were performed in R (v.3.6.1) in RStudio 1.2.5001 and Bioconductor. In brief, transcript-level quantification data were summarized to genes using the tximport package, and data were normalized using the trimmed mean of M values method, implemented in EdgeR[47]. For stage-specific sorted samples, only genes with more than ten counts per million in at least four samples were retained for analysis. For Myb-M gain-of-function samples, only genes with more than one count per million in at least three samples were retained for analysis. Precision weights were applied to each gene on the basis of the mean-variance relationship using the VOOM function in limma[48]. Linear modelling and Bayesian statistics carried out in limma were used to identify differentially expressed genes with an FDR-adjusted $P \le 0.01$ (stage-specific sorted samples) or FDR-adjusted $P \le 0.05$ (Myb-M gain-of-function samples) and an absolute log$_2$-transformed fold change of ≥1 after adjustment with Benjamini–Hochberg correction. A list of differentially expressed genes from

asexual, male and female comparisons is provided in Supplementary Table 3, and a list of differentially expressed genes from overexpression of Myb-M during the asexual cycle is provided in Supplementary Table 12. To obtain male and female marker genes, limma was implemented with an FDR-adjusted $P \leq 0.01$ and an absolute $\log_2$-transformed fold change of ≥ 2 after adjustment with Benjamini–Hochberg correction. Genes that were upregulated against male and asexual or female and asexual were called as female and male markers, respectively. For the female transcriptome, we chose in vivo data (out of two available datasets[8]) owing to the clearer distinction that it made between male and female transcripts. A list of these markers is provided in Supplementary Table 4. Code used for this analysis is provided in Supplementary Data 1 and 2.

scRNA-seq and analysis

After counting samples on the haemocytometer, sorted cells were pelleted at 16,000g for 5 min at 4 °C in a 1.5 ml Eppendorf tube and resuspended in the 10x Genomics recommended resuspension solution of 1× PBS (calcium and magnesium free) containing 0.04% weight/volume BSA (400 µg ml^{-1}). The amount of resuspension solution added was determined by the targeted cell recovery concentration of 1,200 cells per µl with the exception of the 54 h sample, which was 700 cells per µl. These single-cell suspensions were loaded onto a microfluidic chip for encapsulation using the 10x Genomics platform, which uses Gel Beads-in-emulsion (GEMs) technology to generate cDNAs with a cell-specific barcode for each encapsulated cell. 10x Chromium Single Cell 3′ Library v3 chemistry was used according to the manufacturer's instructions for cDNA amplification and library construction. Amplified cDNA and sequencing libraries were quality checked and quantified on the Agilent Tapestation 4200 (Agilent Technologies) and Qubit 3 (Thermo Fisher Scientific) systems, respectively. Using the concentrations from Qubit 3, the average insert size and the expected number of cells, individual libraries were normalized and pooled to a final concentration of 4 nM. The NextSeq 500/550 High Output Kit v2.5 (75 cycles) was used to obtain paired-end reads on the Illumina NextSeq 500 sequencer.

Cell Ranger (v.3.1.0)[49] was used to process sequencing reads and build a reference genome for *C. parvum* Iowa II (CryptoDB, release 46). Processed reads from each sample were aligned to this reference genome, and the output filtered feature–barcode matrix was read into Seurat (v.3.1.4)[50] for filtering, normalization and clustering. These analyses were performed in R (v.3.6.1) in RStudio 1.2.5001 and Bioconductor. R projects were set up to analyse the asexual transcriptome at 24 and 36 h (code is provided in Supplementary Data 3) and the whole life-cycle transcriptome (code is provided in Supplementary Data 4). Low-quality *C. parvum* cells with less than 100 detected genes and greater than 60% rRNA content were removed from further analyses. The maximum cut-off for each sample was determined on the basis of the distribution of detected genes and RNA transcripts. For the in vivo sample, the maximum number of detected genes was 1,800 and the maximum number of unique molecular identifiers (UMIs) was 7,500. For timepoints 24, 36 and 42 h, the maximum number of detected genes was 1,200 and the maximum number of UMIs was 4,000. The maximum number of UMIs was 1,000 for 46 and 54 h, with the number of detected genes set to 500 and 400 for each. This left the following number of high-quality cells per sample for analyses: 1,483 (24 h), 1,506 (36 h), 619 (42 h), 1,828 (46 h), 443 (54 h) and 3,219 (in vivo) (Extended Data Fig. 2c–e and Supplementary Table 1). Data were log-normalized, with the scale factor set to 10,000. All 4,020 *C. parvum* genes were used for integration of multiple samples, and canonical correlation analysis (CCA) was used to find anchors and dimensions 1–30 were used to specify the neighbour search space and weighting procedure.

Principal component analysis was run on both integrated datasets. For the asexual dataset, PCs 1–20 were chosen for clustering and UMAP visualization. For the life-cycle atlas dataset, PCs 1–33 were used, as a significant drop in P value was observed after PC 33 on the basis of

JackStraw analysis. Cells were clustered by using the shared nearest neighbour (SNN) method in the function FindClusters, applying the default parameters. Resolutions of 0.4 and 1.0 were used for the asexual and life-cycle atlas datasets, respectively.

To identify differentially expressed genes between clusters or between samples, Wilcoxon rank-sum tests were implemented with the function FindAllMarkers on the RNA assays, using the following parameters: only.pos = TRUE, min.pct = 0.25, logfc.threshold = 0.25. Note that FindAllMarkers is not exclusive, and a gene may be differentially expressed and identified as a marker in multiple clusters. Adjacent clusters with similar marker genes were combined. Once the proper order from invasion to egress was defined, clusters were assigned new identities and relevelled in numerical order. FindAllMarkers was re-run on the clusters using the parameters above for both the ordered asexual and total atlas datasets, obtaining the differentially expressed genes across the life cycle (Supplementary Tables 2 and 5). AddModuleScore was used to obtain male and female signature scores with the default parameters. AddModuleScore was also used to get signature scores for glideosome, rhoptry and microneme genes across the life cycle, with gene lists obtained from previous published datasets[8,14]. FeaturePlot was used on the RNA assay to visualize the expression of individual genes. Marker genes for male and female clusters with $P < 1 \times 10^{-50}$ were examined for their exclusive expression in males or females and are highlighted in colour in Supplementary Tables 9 and 10.

Further analyses for *C. parvum* life cycle directionality and timing were performed on the RStudio server (Ghost Orchid release) with R v.4.1.3 (code is provided in Supplementary Data 5). The command line tool velocyto and package velocyto.R[13] were used to estimate and visualize RNA velocity of the asexual cycle by distinguishing unspliced and spliced mRNAs. Loom files were generated from the command line and read into R for filtering and analysis. The UMAP reductions from the initial Seurat analysis were intersected with the splicing information from the velocity analysis, and the function RunVelocity was performed on the new object with parameters deltaT = 1, kCells = 25 and fit.quantile = 0.02.

Individual parasites were placed along a developmental trajectory in pseudotime using Monocle3[51–53]. The script was run on the asexual dataset and on the subsets of male and female clusters from the total atlas. For building the asexual trajectory, the parameters k = 20 and resolution = 5e-4 were used for the function cluster_cells and ncenter = 200 and minimal_branch_len = 12 were used for the function learn_graph. The trajectory start was chosen based on the predicted egress and invasion point obtained from clustering and RNA velocity analyses.

To obtain unbranched trajectories for sexual stages, 359 late female parasites clustering along the asexual circle were filtered out of the dataset. They were selected using the function CellSelector and eliminated from the Seurat object by subset. The male and female cells were then subset by cluster identity, with males identified by clusters 10, 11 and 12, and females identified by clusters 13, 14, 15, 16, 17 and 18. For building the male trajectory, the parameters k = 20 and resolution = 5e-3 were used for the function cluster_cells and ncenter = 200 and minimal_branch_len = 12 were used for the function learn_graph. For building the female trajectory, the parameters k = 40 and resolution = 5e-4 were used for the function cluster_cells and ncenter = 130 and minimal_branch_len = 8 were used for the function learn_graph. The starting points of these trajectories were manually chosen to begin at cluster 10 for male and cluster 13 for female. Unless noted above, all of the other parameters were default.

To examine transcription across development, normalized expression data were pulled for specific gene groups from individual cells (gene lists are provided in Supplementary Tables 2, 9 and 10). These data were scaled and plotted as smoothed curves with the stat_smooth function, showing waves of gene expression across pseudotime. The LOPIT groups were previously published[14], while the DNA replication genes were obtained from Gene Ontology term GO:0006260 on CryptoDB.

Article

Other male- and female-specific genes plotted as waves were identified in this Article (Supplementary Tables 9 and 10). For the asexual data, the mean expression of genes from a group was calculated per single cell and plotted as a curve with a 95% confidence interval using stat_smooth. The means for multiple LOPIT groups were also plotted on a min–max scale across pseudotime to show the successive waves of expression of these organelle proteins.

To analyse the expression of long non-coding RNAs (lncRNAs) across the life cycle, we used the *C. parvum* IOWA-ATCC genome assembly with updated annotations[54] and used Cell Ranger v.3.1.0[49] to build a reference genome. Data were processed and analysed in Seurat as outlined above with the following notable changes (code is provided in Supplementary Data 6). Percentage of rRNA content was not used as a filtering parameter. All 4,426 *C. parvum* genes were used for integration of multiple samples, with CCA used to find anchors and dimensions 1–50 used to specify the neighbour search space and weighting procedure. PCs 1–44 were chosen for clustering and UMAP visualization, with a resolution of 1.0. Differentially expressed genes and the top lncRNAs per cluster are found in Supplementary Tables 6 and 7.

Motif analysis

To identify stage-specific motifs, we obtained the 200 bp upstream regions of cluster marker genes with a $P < 1 \times 10^{-50}$. We used simple enrichment analysis from MEME Suite (v.5.5.5)[55] to determine which previously published motifs[33] are over-represented in specific clusters. Shuffled input sequences were used as the control and the *e* value was set as ≤10. Motifs were designated as enriched if their $P < 0.001$ (Supplementary Table 11). To uncover novel motifs, we used STREME. Shuffled input sequences were used as the control and motifs were set with a minimum width of 6 and a maximum width of 16. A *P* value of 0.001 was used to limit the search. Only highly significant motifs meeting this threshold are listed along with their matching genes (Supplementary Table 11).

qPCR

Infections were set up in a 24-well plate, with 200,000 oocysts per well for the 18 h timepoint and 100,000 oocysts per well for the 48 h timepoint. Excystation was performed before infection, and vehicle (0.1% ethanol) or 0.5 μM Shield-1 (Takara Bio, 632189) was added at either 12 h or 36 h. RNA was extracted at 18 h or 48 h by direct lysis in the well with 350 μl of RLT lysis buffer from the RNeasy minikit (Qiagen). The manufacturer's instructions were followed, and RNA was eluted in 40 μl of RNase-free water. RNA concentrations were measured on the Qubit 3 (Thermo Fisher Scientific) system. Then, 1 μg of cDNA was prepared per sample using SuperScript IV reverse transcriptase (Thermo Fisher Scientific). The cDNA was diluted 1:10 before setting up the quantitative PCR (qPCR) in a 10 μl reaction, which included SsoAdvanced Universal SYBR Green Supermix (Bio-Rad). The reaction was loaded into the ViiA 7 Real Time PCR system (Thermo Fisher Scientific) with the following conditions: initial incubation at 95 °C for 3 min, 95 °C for 15 s and 60 °C for 30 s for 40 cycles, and a single melt curve. The $\Delta\Delta C_t$ method[56] was used to determine the relative expression with 18S rRNA as the control gene. A list of all of the primers is provided in Supplementary Table 13, many of which were previously published[57].

Statistical methods

GraphPad PRISM and Excel were used for statistical analyses. Standard *t*-tests were used to measure quantitative differences between two parasite populations, and Welch's correction was applied to parasite counts due to differences in sample size. Error bars are defined in the figure legends.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The reference genomes for *C. parvum* Iowa II and *C. parvum* IOWA-ATCC were obtained from CryptoDB (https://cryptodb.org/cryptodb/app/downloads). The *C. parvum* IOWA-ATCC non-coding RNAs were acquired from GenBank accessions CP044415–CP044422. Sorted asexual and female RNA-seq samples were downloaded from the GEO under accession number GSE129267. RNA-seq data generated in this study are available at the GEO under accession number GSE232438. The *C. parvum* single-cell atlas is available online (https://CryptoDB.org/).

Code availability

All code is found in Supplementary Data 1–6 and at GitHub (https://github.com/katelyn-walzer/Cryptosporidium_single_cell_atlas).

43. Pawlowic, M. C., Vinayak, S., Sateriale, A., Brooks, C. F. & Striepen, B. Generating and maintaining transgenic *Cryptosporidium parvum* parasites. *Curr. Protoc. Microbiol.* **46**, 20B.22.21–20B.22.32 (2017).
44. Alvarez-Jarreta, J. et al. VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center in 2023. *Nucleic Acids Res.* **52**, D808–D816 (2024).
45. Sateriale, A., Pawlowic, M., Vinayak, S., Brooks, C. & Striepen, B. Genetic manipulation of *Cryptosporidium parvum* with CRISPR/Cas9. *Methods Mol. Biol.* **2052**, 219–228 (2020).
46. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
47. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
48. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
49. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
50. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
51. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
52. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
53. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
54. Baptista, R. P. et al. Long-read assembly and comparative evidence-based reanalysis of *Cryptosporidium* genome sequences reveal expanded transporter repertoire and duplication of entire chromosome ends including subtelomeric regions. *Genome Res.* **32**, 203–213 (2022).
55. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).
56. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C_T method. *Nat. Protoc.* **3**, 1101–1108 (2008).
57. Mauzy, M. J., Enomoto, S., Lancto, C. A., Abrahamsen, M. S. & Rutherford, M. S. The *Cryptosporidium parvum* transcriptome during in vitro development. *PLoS ONE* **7**, e31715 (2012).

Author contributions K.A.W., J.T., D.P.B. and B.S. conceived the study. K.A.W. and J.T. generated transgenic parasite strains. J.T. and J.A.G. assisted in cell sorting. E.C.W. and E.K. contributed to single-cell sequencing. K.A.W. performed bioinformatic analyses with support from E.C.W. and D.P.B.; A.M.D., S.D.C. and B.S. contributed to stage-specific gene annotation. K.A.W., J.T., J.H.B. and A.M.D. performed genotypic and phenotypic characterization of parasites. K.A.W. and J.H.B. conducted animal experiments. K.A.W. and B.S. secured funding and wrote the paper.
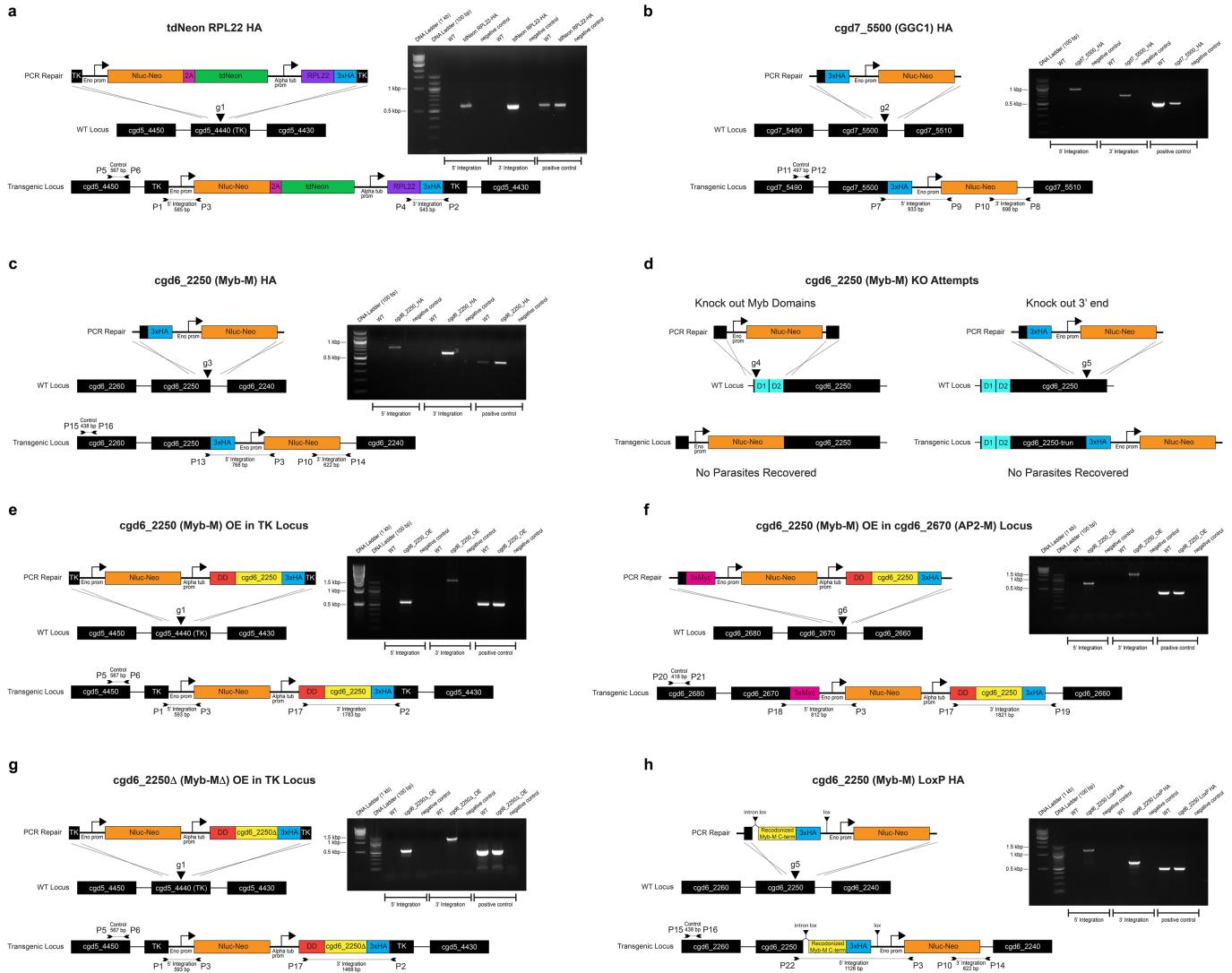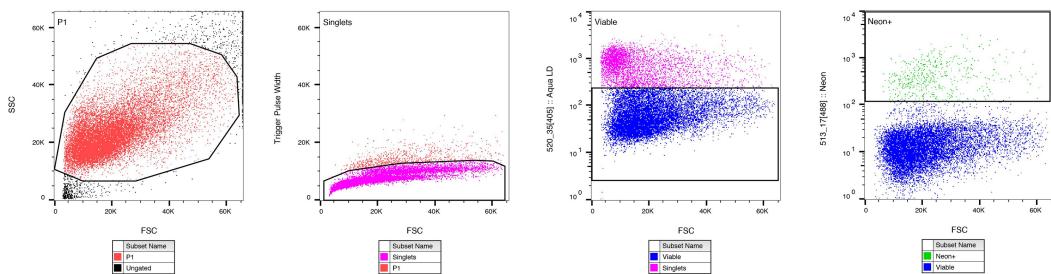
**Extended Data Fig. 1 | Generation of transgenic parasite strains.** Integration maps show the homology repair, the native locus, and the modified locus, with the CRISPR/Cas9 induced break for each transfection marked by an arrowhead. Diagnostic PCR products are represented as inverted arrows on the integration maps and are shown on the corresponding gel, demonstrating successful insertion of the homology repair. Transgenic strains include **a**, tdNeon RPL22 HA, **b**, cgd7_5500 (GGC1) HA, **c**, cgd6_2250 (Myb-M) HA, **d**, cgd6_2250 (Myb-M) knockouts (no parasites recovered), **e**, cgd6_2250 (Myb-M) overexpression in thymidine kinase locus, **f**, cgd6_2250 (Myb-M) overexpression in cgd6_2670 (AP2-M) locus, **g**, cgd6_2250Δ (Myb-MΔ) overexpression in thymidine kinase locus, and **h**, cgd6_2250 (Myb-M) LoxP HA. All guide RNAs and primers noted are listed in Supplementary Table 13. Viable transgenic strains were each generated once, with the exception of cgd6_2250 (Myb-M) HA which was generated three separate times.
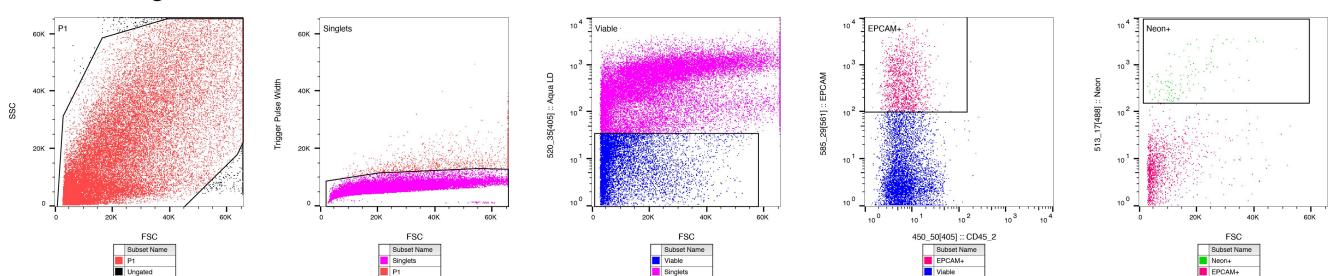
a

Flow Sorting Scheme *In Vitro*
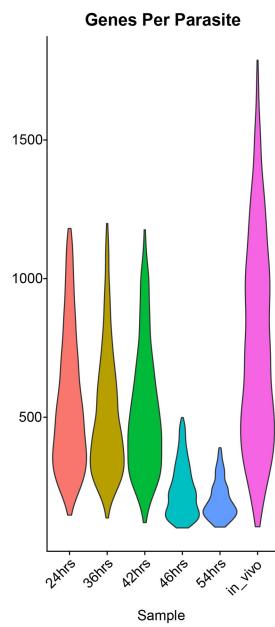
b
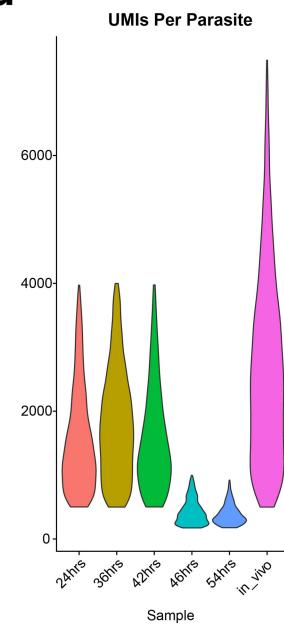
Flow Sorting Scheme *In Vivo*
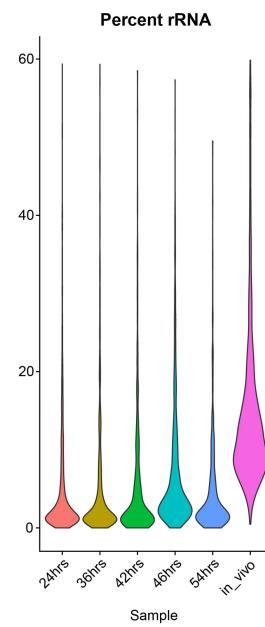
c

d

e
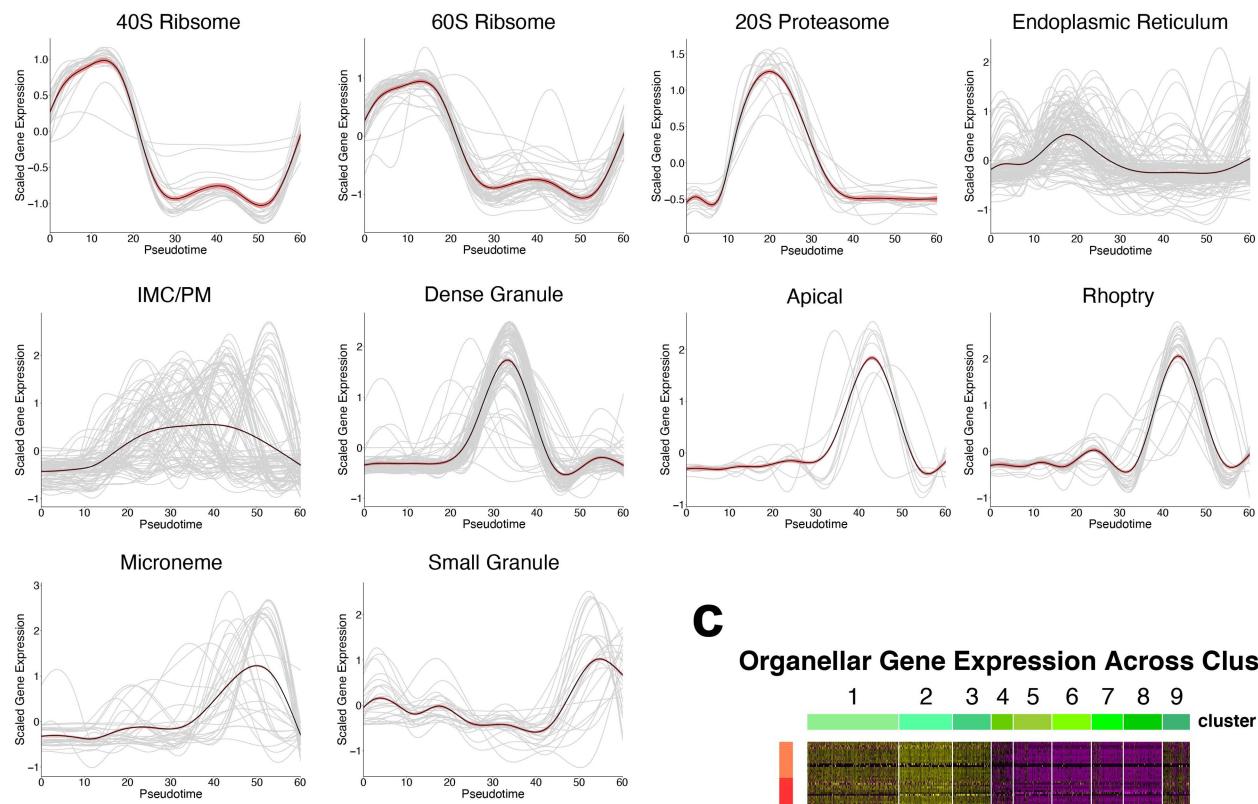
Extended Data Fig. 2 | Enrichment of *C. parvum* infected cells and quality control for scRNA-seq. a, **b**, Cell populations from **a**, culture or **b**, the ileum of an IFN-$\gamma^{-/-}$ mouse were sorted to enrich for viable cells infected with fluorescent green *C. parvum*. Cells sorted from the mouse ileum were also enriched for intestinal epithelial cells based on EPCAM positivity. Sorting gates are outlined in black. The plots were generated with FlowJo 10.8.1. **c-e**, Violin plots show the distribution of **c**, genes detected, **d**, UMIs, and **e**, percentage of transcripts mapping to rRNA genes across samples after filtering. Filtering cutoffs included a minimum of 100 detected genes and less than 60 percent of transcripts mapping to rRNA. The maximum cutoffs were set per sample according to the pre-processed distribution of detected genes and UMIs. For the in vivo sample, the maximum number of detected genes was 1800 and the maximum number of UMIs was 7500. For timepoints 24, 36, and 42 h, the maximum number of detected genes was 1200 and the maximum number of UMIs was 4000. For timepoints 46 and 54 h, the maximum number of detected genes was 500 and 400 and the maximum number of UMIs was 1000.
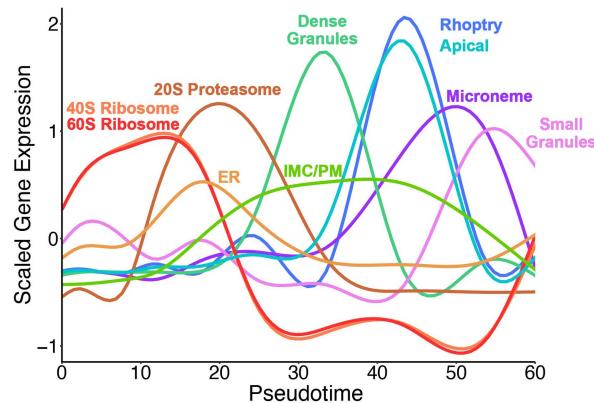
a
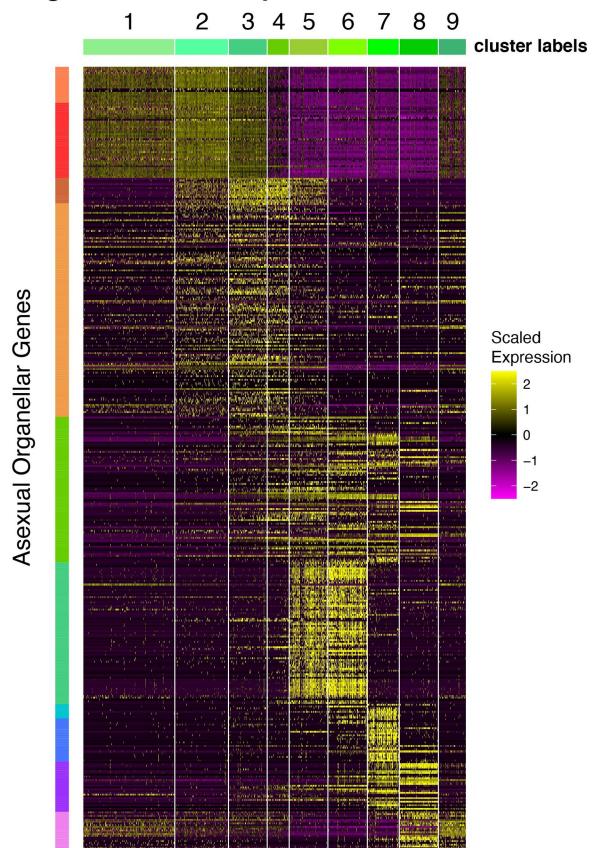
Asexual Organellar Genes Identified by Spatial Proteomics

c

Organellar Gene Expression Across Clusters

b

Asexual Organellar Mean Gene Expression
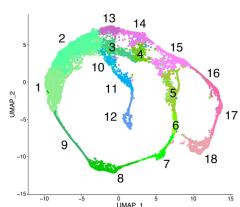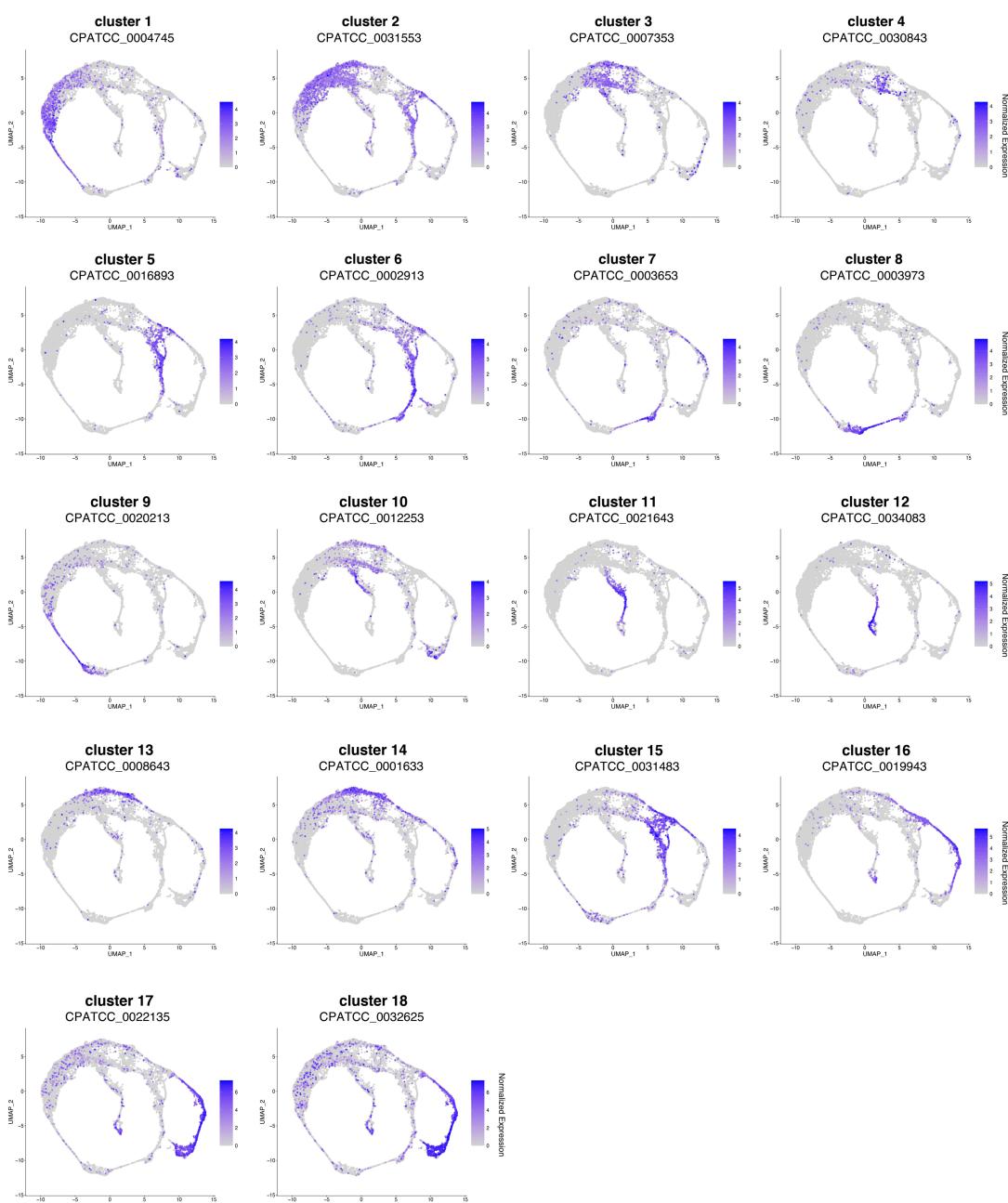
Extended Data Fig. 3 | Temporal gene expression across the asexual cycle.
a, **b**, Scaled expression of organellar genes[14] across pseudotime. Individual genes are shown in grey, the mean in black, and the 95 percent confidence interval in red (**a**). The means from each organelle (black in **a**) are plotted in colour (**b**). **c**, A heatmap shows the scaled expression of the organellar genes across asexual clusters. The colour for each gene group corresponds to that used in (**b**). These data represent the parasite transcriptomes captured at 24 and 36 h (see Supplementary Table 2 for gene lists).
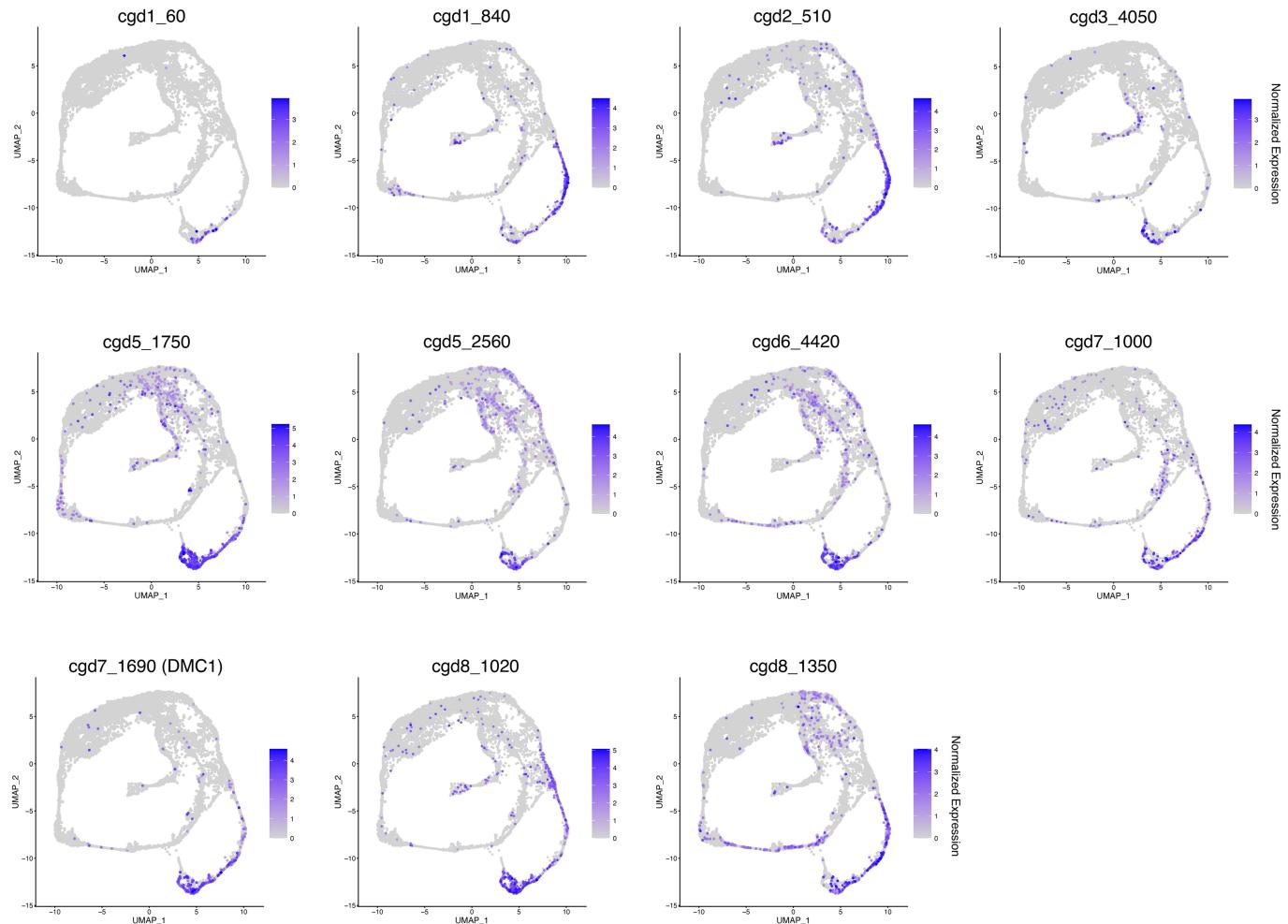
Extended Data Fig. 4 | Expression of cluster markers across the *C. parvum* life cycle. A heatmap shows the scaled expression of 2880 genes that were identified as markers for one or more clusters defining the *C. parvum* transcriptome (see Supplementary Table 5 for gene lists).

a   **Cluster analysis with lncRNAs**

**b**   **Top expressed lncRNA for each cluster**

Extended Data Fig. 5 | The transcriptome of the full *C. parvum* life cycle including lncRNAs. a, The UMAP shows 8962 individual *C. parvum* transcriptomes and is coloured for clustering across asexual (green), male (blue), and female (pink) parasites. **b**, Each UMAP shows the top lncRNA gene for each cluster with its normalized expression in blue.
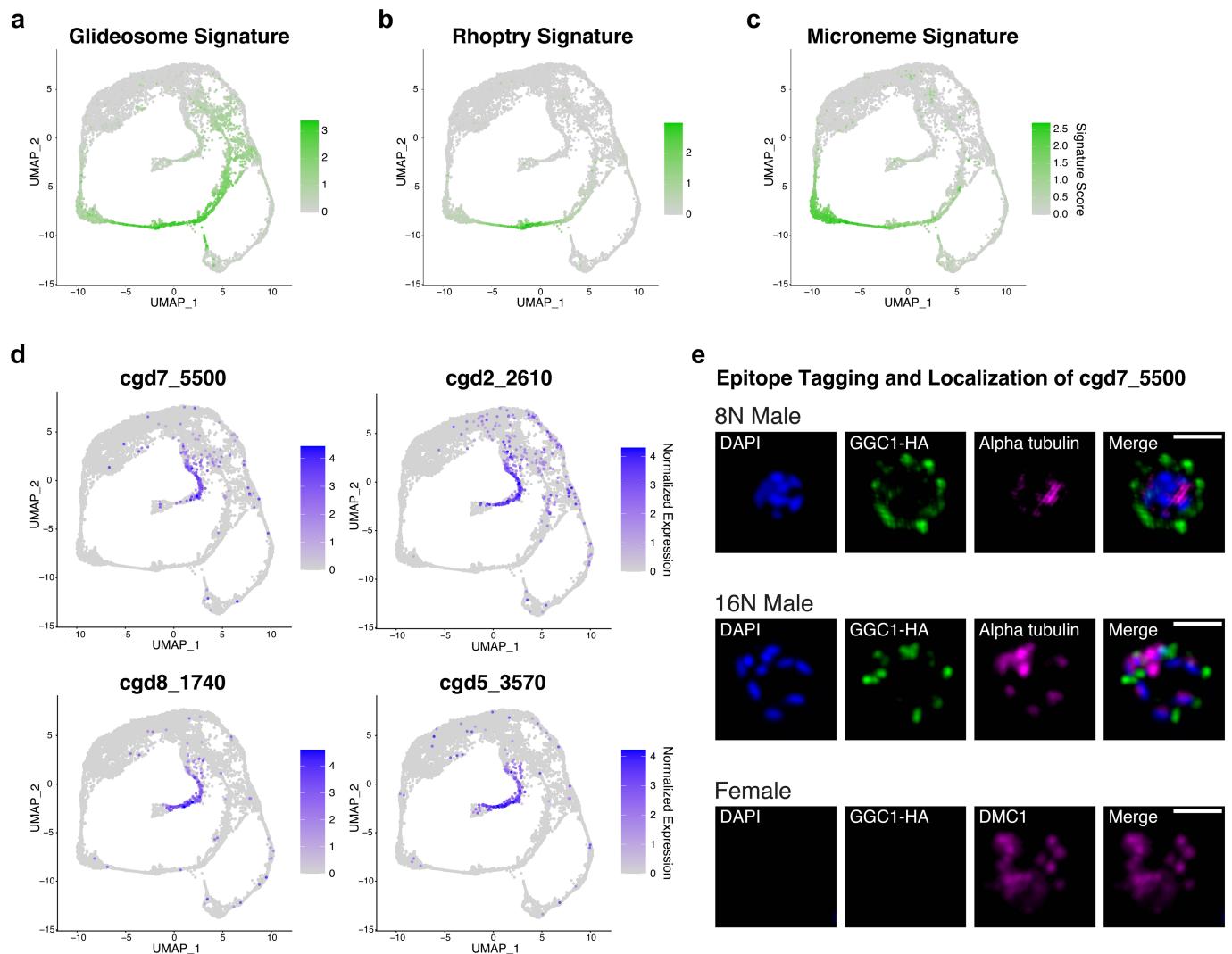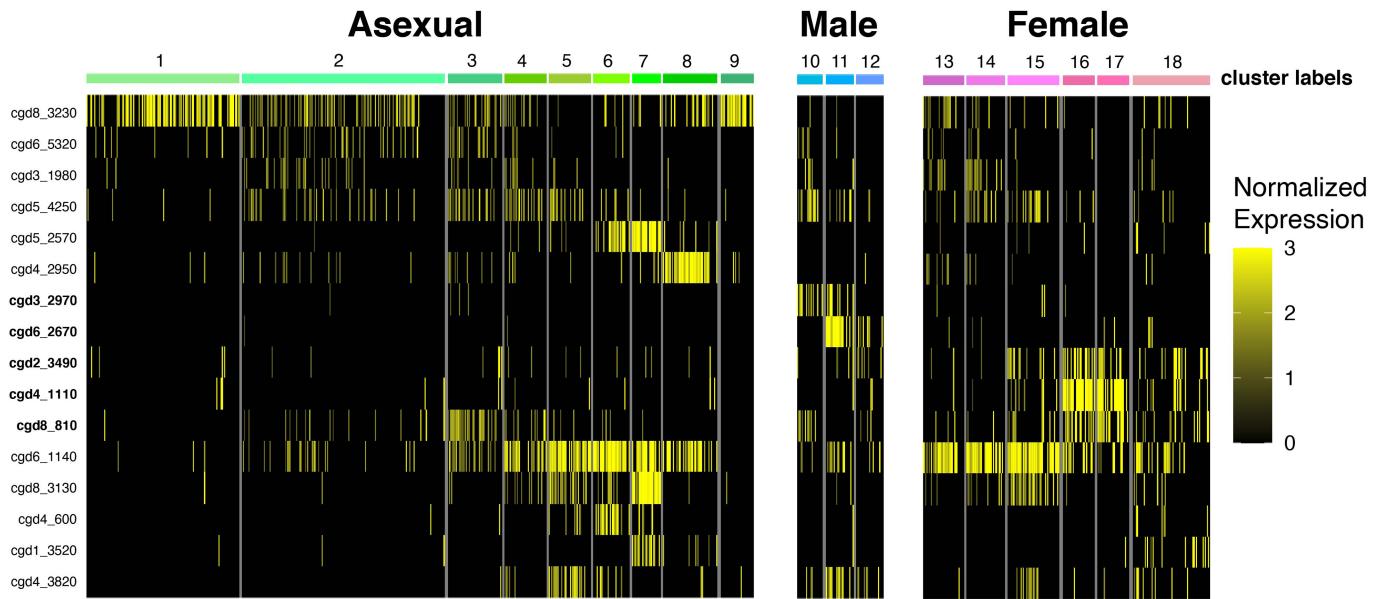
Extended Data Fig. 6 | Identification of meiosis-specific genes. Previously published meiosis and DNA repair genes[8] were examined for their female-specific expression, as expression during other stages would indicate that they are not specific for meiosis. Eight genes were identified as female-specific from the previous list and were designated as meiosis genes. Further examination of female-specific genes (Supplementary Table 9) revealed three additional meiosis genes not included in the previously published list. UMAPs show their normalized expression levels in blue.

a Glideosome Signature

b Rhoptry Signature

c Microneme Signature

d cgd7_5500   cgd2_2610

cgd8_1740   cgd5_3570

e Epitope Tagging and Localization of cgd7_5500

8N Male — DAPI | GGC1-HA | Alpha tubulin | Merge

16N Male — DAPI | GGC1-HA | Alpha tubulin | Merge

Female — DAPI | GGC1-HA | DMC1 | Merge
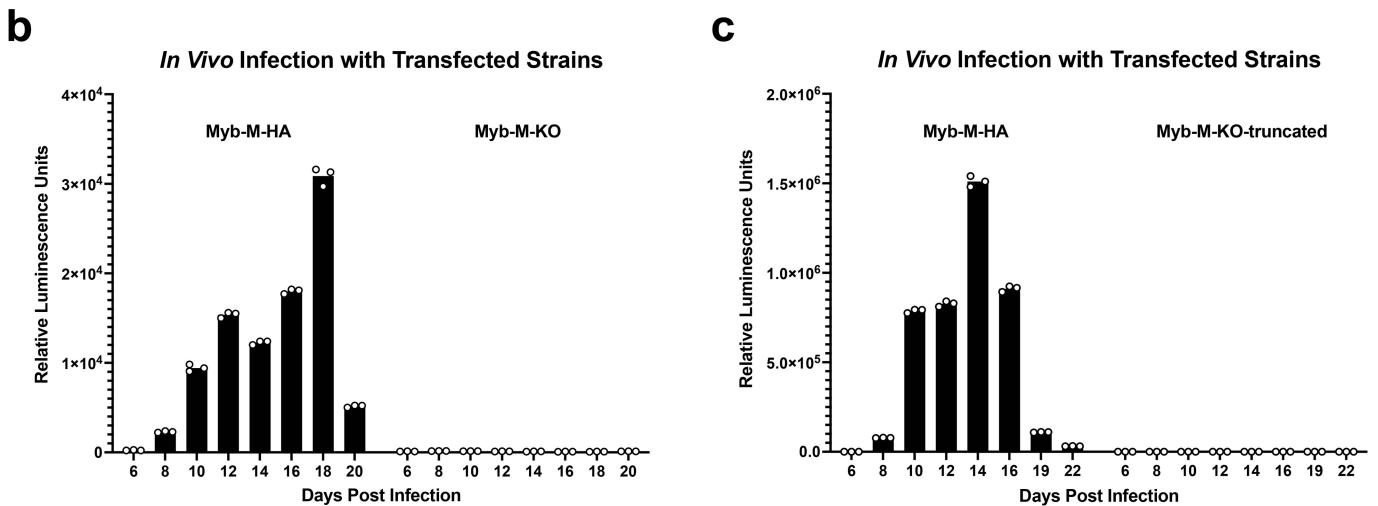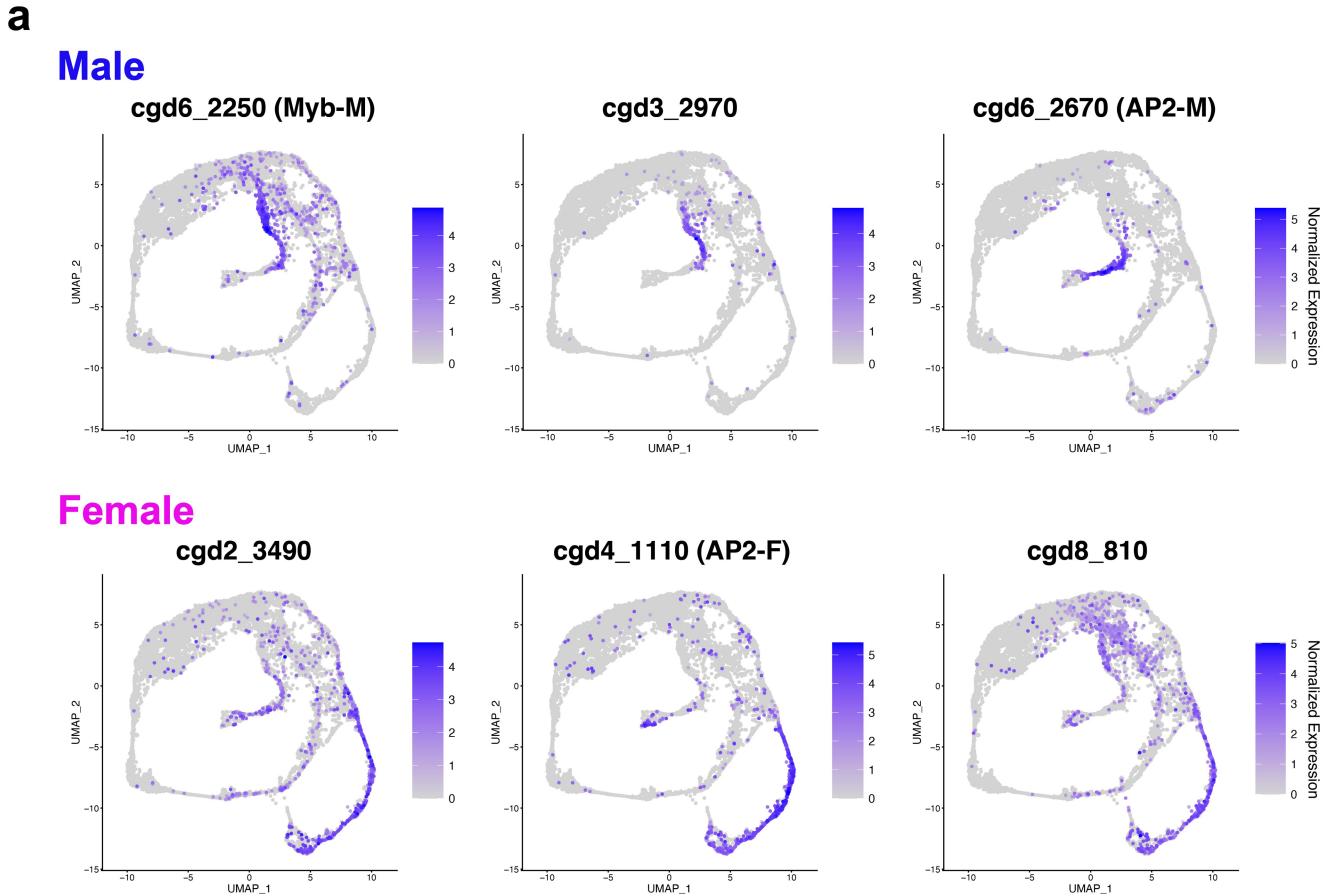
Extended Data Fig. 7 | Male gametes lack gliding and invasion machinery but exclusively express genes of the GGC exported protein family.
a-c, Signature scores were obtained for glideosome[8] (**a**), rhoptry[14] (**b**), and microneme[14] (**c**) genes and were painted onto the UMAP in green. Note that following successful fertilization and meiosis, in vivo females form sporozoites which are indistinguishable from merozoites with respect to their motility machinery. Late in vivo females therefore exhibit expression of glideosome genes. **d**, UMAPs show the normalized expression levels (in blue) of four paralogous GGC proteins. **e**, GGC1 (cgd7_5500) was epitope tagged and visualized in immunofluorescence assays after 54 h of growth in culture. Tagged GGC1 was stained with an antibody to HA, shown in green, while the male nucleus was stained with DAPI in blue. Males were stained with alpha tubulin, shown in magenta, and females were stained with DMC1, also shown in magenta, in two separate assays. GGC1 localizes to the periphery of the 8 N male gamont, surrounding the nuclei, and then localizes to the apical end of mature male gametes. It is not expressed in females (scale bar = 2 µm, data representative of two experiments).
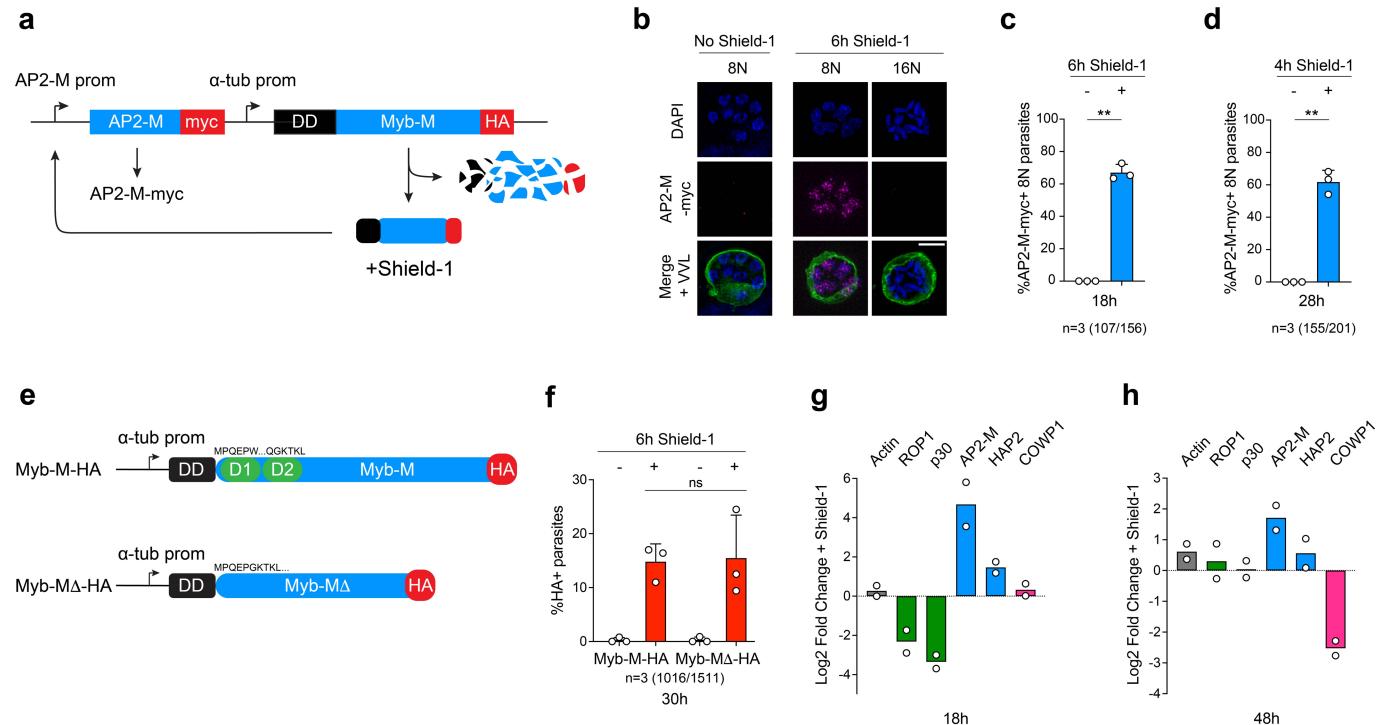
Extended Data Fig. 8 | Expression of AP2 transcription factors across the *C. parvum* life cycle. A heatmap shows the normalized expression of 16 AP2 transcription factors across the *C. parvum* transcriptome. AP2s are largely stage- and sex-specific. Male and female expressed AP2s are shown in bold.

a

Male

cgd6_2250 (Myb-M)

cgd3_2970

cgd6_2670 (AP2-M)

Female

cgd2_3490

cgd4_1110 (AP2-F)

cgd8_810

b

In Vivo Infection with Transfected Strains

Myb-M-HA   Myb-M-KO

c

In Vivo Infection with Transfected Strains

Myb-M-HA   Myb-M-KO-truncated

Extended Data Fig. 9 | Myb-M is the earliest male transcription factor and the locus is refractory to disruption. a, UMAPs show the normalized expression levels (in blue) of male- and female-specific transcription factors. Myb-M is expressed the earliest of all factors with female AP2s expressed particularly late in female developmental progression. **b, c**, Myb-M was targeted for deletion at the predicted N-terminal DNA binding domains (**b**) or after for C-terminal truncation (832 base pairs into the coding sequence, **c**) by CRISPR/ Cas9 mediated marker insertions. Myb-M-HA was used as a control in parallel.

Sporozoites transfected with a Cas9+gRNA plasmid and homology repair template (see Extended Data Fig. 1 for design) were given to mice via oral gavage. Oocyst shedding was monitored by faecal luminescence measurements. Data is represented as the mean of three technical replicates. Note that while epitope tagged mutants are readily recovered, no viable transgenics were isolated using the deletion and truncation constructs, suggesting that the gene is likely essential.

Extended Data Fig. 10 | Ectopic expression of Myb-M drives parasites to a male fate. a, Diagram showing conditional Myb-M overexpression AP2-M reporter strain. **b**, **c**, Infected cultures were treated with vehicle or Shield-1 at 12 h, fixed at 18 h, and AP2-M was detected with an antibody to Myc (scale bar = 2 μm) and quantified in 8 N parasites (**c**). Error bars represent the standard deviation of the mean from three biological replicates (****P** = 0.0018, two-tailed unpaired t-test with Welch's correction). **d**, Infected cultures were treated with vehicle or Shield-1 at 24 h, fixed at 28 h, and AP2-M was quantified as in **c** (****P** = 0.0044, two-tailed unpaired t-test with Welch's correction). **e**, Schematic comparing the conditional Myb-M-HA and Myb-MΔ-HA overexpression constructs. The conserved DNA-binding domains predicted by AlphaFold are marked in green, with the amino acid sequence annotated above. 315 nucleotides were deleted from conditional Myb-MΔ-HA, and the resulting open reading frame is shown indicating the amino acid sequence of the deletion. **f**, Inducible Myb-M-HA and Myb-MΔ-HA parasites were treated with Shield-1 and scored for HA staining. Error bars represent the standard deviation of the mean from three biological replicates. Significance was evaluated by a two-tailed Welch's t-test. **g**, **h**, HCT-8 cultures infected with the inducible Myb-M-HA strain were treated with vehicle or Shield−1 at either 12 h (**g**) or 36 h (**h**). RNA was harvested at 18 h (**g**) or 48 h (**h**) and representative constitutive (grey), asexual (green), male (blue), or female (pink) transcripts were measured by qPCR. The log2 fold change is relative to the 18S rRNA control and vehicle-treated samples. Data is representative of two independent biological repeats and is plotted as the mean.