Draft version June 25, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

# Comparative Single-Cell Transcriptomics Reveals Divergent Stage Transition Dynamics and Regulatory Strategies in Lab-Adapted and Field Isolates of *Plasmodium falciparum*

Denario[1]

[1] *Anthropic, Gemini & OpenAI servers. Planet Earth.*

## ABSTRACT

The malaria parasite *Plasmodium falciparum* undergoes tightly regulated stage transitions during its intraerythrocytic development, but the dynamics of these transitions may differ between parasites adapted to laboratory conditions and those circulating in natural human hosts. To investigate these differences, we performed a comparative single-cell transcriptomic analysis leveraging a dataset of 45,691 parasite cells, combining laboratory strains with field isolates from asymptomatic patients. We mapped developmental trajectories using PAGA-based trajectory inference, identified dynamic gene expression modules through differential gene expression analysis, and pinpointed candidate master regulators. Our analysis revealed that laboratory strains exhibit a continuous asexual developmental cycle, while field isolates are skewed towards sexual stages. Notably, we observed that candidate master regulators in laboratory strains show a 'just-in-time' activation pattern, with expression preceding downstream gene expression by a short interval. In contrast, field isolates displayed a 'priming' regulatory strategy, where regulators are expressed long before their target genes are activated. These findings suggest that *P. falciparum* adapts its stage progression control in response to the host environment, potentially reflecting an adaptation to ensure efficient transmission in the complex and variable environment of the human host.

*Keywords:* Plasma astrophysics, Cosmic background radiation, Observational cosmology, Solar magnetic fields, Stellar nucleosynthesis

## 1. INTRODUCTION

Malaria, caused by the protozoan parasite *Plasmodium falciparum*, remains a major global health burden. A key aspect of the parasite's success is its complex life cycle, involving distinct developmental stages within both the mosquito vector and the human host. Within the human host, *P. falciparum* undergoes asexual replication in red blood cells, a process known as the intraerythrocytic development cycle (IDC). This cycle is characterized by tightly regulated stage transitions, from ring to trophozoite to schizont, culminating in the release of merozoites that infect new red blood cells. A subset of parasites also differentiate into sexual stages (gametocytes), which are essential for transmission back to the mosquito vector. The precise timing and coordination of these stage transitions are crucial for parasite survival and propagation, making an understanding of the molecular mechanisms governing these transitions critical for developing effective malaria control strategies.

However, unraveling the intricacies of *P. falciparum* stage transitions presents significant challenges. The parasite's genome is complex, and gene expression is highly dynamic during the IDC. Obtaining perfectly synchronized parasite populations for traditional bulk transcriptomic analyses is difficult, obscuring the heterogeneity inherent in these developmental processes. Furthermore, laboratory-adapted strains of *P. falciparum*, while convenient for research, may not accurately reflect the behavior of parasites circulating in natural human populations. These lab-adapted strains have undergone extensive passaging under artificial conditions, potentially leading to alterations in their developmental dynamics and regulatory mechanisms. Therefore, comparative studies of lab-adapted strains and field isolates are essential to understand the parasite's adaptive strategies in the complex and variable environment of the human host.

In this paper, we address these challenges by performing a comparative single-cell transcriptomic analysis of *P. falciparum* stage transitions in both laboratory strains and field isolates obtained from asymp-

tomatic patients. Single-cell transcriptomics allows us to dissect the heterogeneity of parasite populations and capture the dynamic gene expression changes occurring during stage transitions at an unprecedented resolution. Leveraging a dataset of 45,691 parasite cells, we employ trajectory inference methods to map developmental paths and identify key transition states. We identify genes with dynamic expression patterns associated with these transitions, including candidate regulators whose expression correlates with the onset of downstream transcriptional changes or exhibits stage-specific variability. By comparing these dynamics, candidate regulatory networks, and the timing of transitions between lab-adapted strains and field isolates, we uncover differences in stage progression control and potential adaptations in the natural host environment.

Specifically, we use Partition-based Graph Abstraction (PAGA) to infer developmental trajectories independently for lab-adapted and field isolates. PAGA constructs a simplified, coarse-grained graph representation of the single-cell data, revealing the major developmental paths and their connectivity. We then calculate pseudotime values to order cells along these trajectories, providing a continuous measure of developmental progression. Differential expression analysis is performed to identify genes whose expression changes significantly along pseudotime. We identify candidate master regulators by searching for genes with generally low expression but transient, sharp increases in their relative expression that precede the upregulation of downstream gene modules associated with life stage transitions. We use relative gene expression to account for the fact that master regulators might not have very high expression, but rather might have an expression peak that coincides with the life stage transition. Additionally, we assess the variability of gene expression within defined transition windows to identify potential regulators involved in mediating state changes.

To verify our findings, we compare the trajectory structures, pseudotime distributions, and candidate regulatory networks between lab and field isolates. We perform differential gene expression analysis between lab and field isolates within each major life cycle stage, as well as along pseudotime, allowing us to pinpoint genes that exhibit different expression patterns in the two conditions. We compare the lists of candidate master regulators and their associated downstream gene modules, identifying regulators common to both, unique to lab, or unique to field isolates. For common regulators, we compare their expression dynamics and associated downstream modules. We expect to find that lab isolates exhibit a continuous asexual developmental cycle, while field isolates are skewed towards sexual stages. Through this comparative analysis, we aim to provide insights into the adaptive strategies employed by *P. falciparum* to ensure efficient transmission in the complex and variable environment of the human host.

## 2. METHODS

### 2.1. *Data Acquisition and Preprocessing*

Our study leverages a single-cell transcriptomic dataset comprising 45,691 *Plasmodium falciparum* cells, integrating both laboratory-adapted strains and field isolates obtained from asymptomatic patients. The initial dataset consisted of two files: a gene expression matrix (‘gene$_e$$x$$p$$r$$e$$s$$s$$i$$o$$n$.$c$$s$$v$‘)$c$$o$$n$$t$$a$$i$$n$$i$$n$$g$$n$$o$$r$$m$$a$$l$$i$$z$$e$$d$$e$$x$$p$$r$$e$$s$$s$$i$$o$$n$$v$$a$$l$$u$$e$$s$$a$$n$$d$
$c$$e$$l$$l$$a$$n$$a$$l$$y$$s$$i$$s$$w$$o$$r$$k$$f$$l$$o$$w$$s$.$T$$h$$e$$c$$e$$l$$l$$m$$e$$t$$a$$d$$a$$t$$a$$w$$a$$s$$m$$e$$r$$g$$e$$d$$w$$i$$t$$h$$t$$h$$e$$t$$r$$a$$n$$s$$p$
$c$$e$$l$$l$$a$$n$$a$$l$$y$$s$$i$$s$.

### 2.2. *Exploratory Data Analysis and Quality Control*

Prior to in-depth analysis, we performed an extensive exploratory data analysis (EDA) to assess data quality and identify potential biases. This included calculating key quality control (QC) metrics at both the cell and gene levels. At the cell level, we computed the total sum of normalized expression values (proxy for total counts) and the number of detected genes (genes with expression $> 0$) for each cell. Distributions of these metrics were examined to identify cells with unusually low expression or gene detection rates, which might represent low-quality cells or empty droplets. At the gene level, we calculated the number of cells in which each gene was detected, as well as its mean and median normalized expression across all cells. This allowed us to identify genes with sparse expression or very low overall abundance. Additionally, we calculated the overall sparsity of the gene expression matrix, representing the percentage of zero values, to assess the data's overall information content.

Based on the EDA results, we implemented stringent filtering criteria to remove low-quality cells and uninformative genes. Cells with exceptionally low total normalized expression or a very low number of detected genes (defined as values falling below two standard deviations from the mean or below predefined thresholds) were flagged and removed. Genes detected in a very small number of cells (fewer than 5 cells) were also removed to reduce noise and computational load. Furthermore, the ‘source‘ column in the metadata was used to create a binary 'origin' column, categorizing samples as either 'lab' (laboratory-adapted strains) or 'field' (field isolates). Following filtering, the data was normalized

and log-transformed using 'log1p(X)' to stabilize variance and prepare the data for downstream analysis.

### 2.3. *Dimensionality Reduction and Visualization*

To reduce the dimensionality of the data and facilitate visualization, we employed a combination of feature selection, scaling, principal component analysis (PCA), and uniform manifold approximation and projection (UMAP). First, we identified highly variable genes (HVGs) using the 'scanpy.pp.highly$_v ariable_g enes$'$functionwiththe$'$flavor =$' $seurat\_v3$''$method, selecting$2000$HVGsfordownstreamanalysis.Thisstepfocusedtheanalysisongenescontributingmosttocell-to-cellheterogeneity.Theexpressionvaluesoftheselected HVGswerethenscaled(z-scoretransformation)acrosscells, withclippingappliedtomitigatetheimpactofextremeoutliers.PCAwasperformedonthescaledHVGmatrix, retaining50PCsthatcapturedthemajorityofthebiologicalvariance.Finally, UMAPwasappliedtothePCembeddingstogenerate2Drepresentations$ for visualization.

### 2.4. *Trajectory Inference and Pseudotime Analysis*

Given the cyclical nature of asexual development and the branching to sexual stages in *P. falciparum*, we employed Partition-based Graph Abstraction (PAGA) followed by pseudotime inference to map developmental trajectories. PAGA constructs a simplified, coarse-grained graph representation of the single-cell data, revealing the major developmental paths and their connectivity. This was performed using 'scanpy.tl.paga'. Trajectory inference was performed separately for the combined laboratory strains and the combined field isolates. Root cells for pseudotime calculation were identified based on biological knowledge, selecting cells in the earliest 'Ring' stage cluster. Pseudotime values were then calculated for all cells along the inferred developmental paths using Diffusion Pseudotime (DPT). The UMAP embedding was visualized with cells colored by their pseudotime values to confirm the inferred progression.

### 2.5. *Identification of Transition-Associated Genes and Candidate Regulators*

To identify genes associated with stage transitions, we performed differential expression analysis along the inferred pseudotime trajectories. Genes whose expression significantly changed along pseudotime were identified using regression models. These dynamically expressed genes were clustered based on their pseudotime expression patterns to identify modules of co-regulated genes associated with specific developmental phases or transitions.

We then focused on identifying candidate master regulators. Our strategy was to search for genes characterized by generally low absolute expression levels across the lifecycle but with transient, sharp increases in their own relative expression (e.g., z-scored expression within that gene across pseudotime) immediately preceding an increase in expression of a gene module associated with a life stage transition. To implement this strategy, we first filtered for potential transcription factors or known regulatory genes, if a reliable list for *P. falciparum* was available. Otherwise, all genes were considered. For each gene, we calculated its mean expression across all cells and considered genes below a certain threshold (e.g., 25th percentile of mean expressions) as "low expression" candidates. For these low-expression candidates, we smoothed their expression along pseudotime using a rolling mean. We then normalized the smoothed expression of each candidate gene relative to its own expression range along pseudotime (e.g., min-max scaling) to highlight relative peaks. We identified significant peaks in this relativized expression profile and assessed whether each peak preceded (by a defined pseudotime window) the upregulation of a downstream gene module or known marker genes for the subsequent stage. This was done by cross-correlation analysis between the candidate regulator's relativized expression and the average expression of the downstream module along pseudotime.

Finally, we assessed stage-specific expression variability. For each gene, we calculated its coefficient of variation (CV) or variance of expression within defined transition windows (e.g., cells in late pseudotime of stage X and early pseudotime of stage Y) versus stable mid-stage windows. Genes with significantly higher variability in transition states, especially those identified as candidate master regulators, were prioritized as potential regulators involved in mediating state changes.

### 2.6. *Comparative Analysis: Lab vs. Field Isolates*

To understand the differences between lab-adapted strains and field isolates, we performed a comparative analysis focusing on trajectory structures, gene expression dynamics, and candidate regulatory networks. The PAGA graphs and UMAP embeddings of lab vs. field isolates were qualitatively and quantitatively compared, noting differences in connectivity between cell clusters (stages), presence/absence of specific paths, and the distribution of cells along pseudotime for equivalent developmental stages. We assessed the "smoothness" or "timing" of transitions by comparing the pseudotime intervals spanning defined stage transitions (e.g., Ring to Trophozoite) using statistical tests to compare these intervals.

Differential gene expression (DGE) analysis was performed to identify genes with different expression patterns between lab and field isolates. We conducted DGE analysis within each major, well-populated life cycle stage (e.g., Rings, Trophozoites, Schizonts) to iden-

tify genes differentially expressed between lab and field isolates. We also tested for genes that showed different dynamic expression patterns along pseudotime between lab and field isolates using specialized tools like ‘tradeSeq‘.

Finally, we compared the lists of candidate master regulators and their associated downstream gene modules identified for lab and field isolates, identifying regulators common to both, unique to lab, or unique to field isolates. For common regulators, their expression dynamics (peak timing, relative magnitude, associated downstream modules) were compared.

### 2.7. *Software and Tools*

The primary analysis was conducted in Python (version 3.8+) using libraries such as ‘scanpy‘ (for core single-cell analysis), ‘pandas‘ (data manipulation), ‘numpy‘ (numerical operations), ‘scipy‘ (statistical tests), ‘statsmodels‘ (for regression models), ‘matplotlib‘ and ‘seaborn‘ (visualization). If necessary, specific analyses (e.g., advanced GAM fitting for pseudotime DGE) might leverage R packages like ‘tradeSeq‘ or ‘mgcv‘ via ‘rpy2‘.

## 3. RESULTS

### 3.1. *Single-Cell Transcriptomic Atlas of P. falciparum Blood Stages from Laboratory and Field Isolates*

To investigate the transcriptional dynamics of *P. falciparum* development, we analyzed a single-cell RNA sequencing dataset comprising 45,691 individual parasites. This dataset combines cells from continuously cultured laboratory strains (37,624 cells) and parasites isolated directly from four asymptomatic human donors in Mali (8,067 cells), hereafter referred to as ”field isolates.” The initial dataset contained expression measurements for 5,274 genes.

Quality control analysis revealed high-quality data across all samples. The median number of detected genes per cell was 937, with a median sum of normalized expression counts of 2,059. The overall sparsity of the expression matrix was 80.25%, which is typical for single-cell transcriptomic data. Based on the distributions of per-cell quality metrics, we established conservative filtering thresholds (minimum 200 genes and 500 total counts per cell; minimum 3 cells per gene). As shown in Figure **??**, all cells and genes passed these initial filters, indicating that the dataset was of high quality and did not contain significant populations of dead cells or empty droplets.

A cross-tabulation of cell source against annotated life cycle stage revealed a critical difference in sample composition between the laboratory and field isolates. The laboratory samples spanned the entire intraerythrocytic developmental cycle (IDC), including early and late rings, trophozoites, and schizonts, as well as committed sexual stages (gametocytes). In stark contrast, the field isolates were heavily skewed towards sexual stages, consisting almost exclusively of developing and mature male and female gametocytes, with a smaller population of late-ring stage parasites. Early asexual proliferative stages (trophozoites and schizonts) were largely absent from the field samples. This fundamental difference in stage representation is a key consideration for all subsequent comparative analyses.

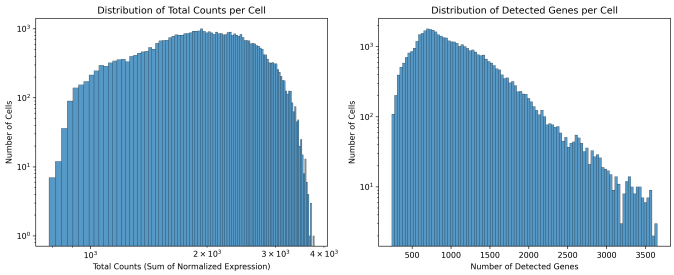


**Figure 1.** Distributions of total counts and number of detected genes per cell, showing high-quality data across all samples and justifying the use of conservative filtering thresholds.

### 3.2. *The Transcriptional Landscape Reveals Distinct Developmental Trajectories and Origin-Specific States*

To visualize the cellular landscape, we performed dimensionality reduction using UMAP on the 2,500 most highly variable genes across all 45,691 cells. The variance ratio explained by the principal components (PCs) is shown in Figure **??**. The first PCs explain most of the variance in the data, which is used for dimensionality reduction and visualization of the cellular landscape. The resulting two-dimensional embedding organizes cells based on their global transcriptomic similarity, revealing the underlying structure of parasite development.

When cells were colored by their annotated life cycle stage, a clear and continuous progression corresponding to the asexual IDC was observed, primarily defined by the laboratory strains. The trajectory begins with early-ring stage parasites, flows through late-ring and trophozoite stages, and culminates in early and late schizonts. Branching from this main proliferative cycle are distinct and well-separated clusters corresponding to the sexual lineage, including developing gametocytes and terminally differentiated male and female gametocytes. This structure faithfully recapitulates the known biology of *P. falciparum* blood-stage development.

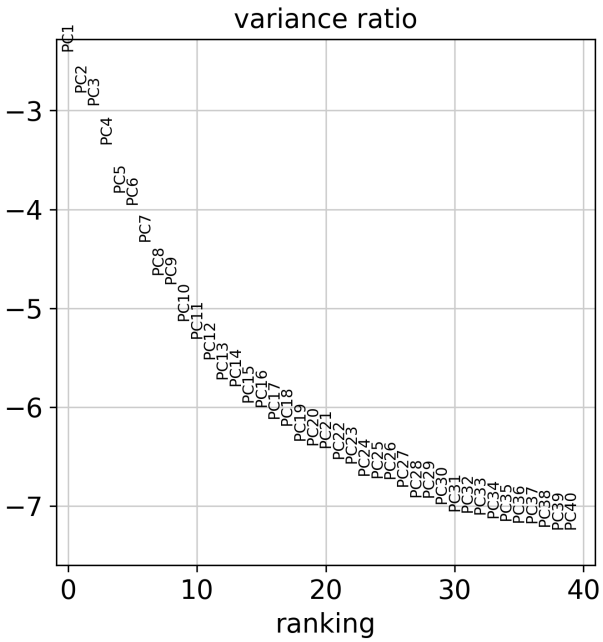


**Figure 2.** Variance ratio explained by the principal components (PCs). The first PCs explain most of the variance in the data, which is used for dimensionality reduction and visualization of the cellular landscape.

Coloring the same UMAP by sample origin ('lab' vs. 'field') exposed a pronounced separation between the two groups. While some overlap existed, particularly in the gametocyte clusters where both sources were well-represented, the majority of cells from lab and field isolates occupied distinct regions of the UMAP space. As shown in Figure **??**, this indicates that, beyond the differences in stage composition, there are substantial global transcriptomic differences between parasites adapted to long-term *in vitro* culture and those circulating in a natural human host.

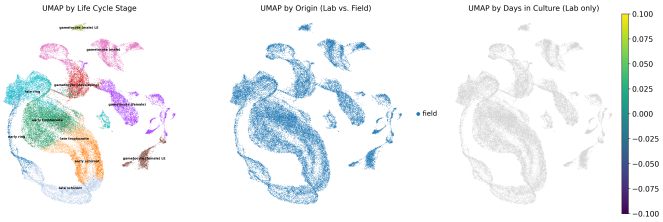


**Figure 3.** UMAP visualization of single-cell transcriptomes colored by annotated life cycle stage, sample origin (lab vs. field), and days in culture for lab isolates. The separation of lab and field samples suggests distinct transcriptomic states beyond differences in stage composition, which may be related to the different regulatory strategies observed.

### 3.3. *Pseudotemporal Ordering Reconstructs and Contrasts Developmental Timing*

To model the continuous nature of parasite development, we employed trajectory inference using PAGA and calculated diffusion pseudotime separately for the lab and field isolates. This approach orders cells along a developmental continuum, providing a quantitative measure of progression.

For the laboratory isolates, we rooted the trajectory in the 'early ring' stage, the known start of the IDC. The resulting PAGA graph showed strong connectivity between contiguous stages of the asexual cycle (ring → trophozoite → schizont) and a clear bifurcation towards gametocyte development, consistent with biological expectations. The calculated pseudotime progressed smoothly along this main trajectory, providing a high-resolution map of the lab-adapted IDC.

For the field isolates, the absence of early asexual stages necessitated rooting the trajectory in the earliest available population, the 'late ring' stage. The PAGA graph for these isolates primarily depicted connections between the late-ring population and the various gametocyte stages, reflecting the sampled population's commitment to sexual development. The subsequent pseudotime calculation for field isolates captured this specific ring-to-gametocyte transition. During this step, a small number of cells (4 out of 8,067) were assigned infinite pseudotime values, suggesting they belonged to minor, disconnected populations within the field dataset's graph structure. These cells were excluded from subsequent pseudotime-dependent analyses. Figure **??** shows the PAGA graphs and diffusion pseudotime plots for laboratory and field isolates. These plots underscore the different developmental processes captured in the two sample sets: a full, primarily asexual cycle in the lab versus a committed sexual development path in the field.

### 3.4. *Dynamic Gene Expression Modules Define Stage-Specific Transcriptional Programs*

To identify genes driving stage transitions, we modeled gene expression as a function of pseudotime. For both lab and field isolates, we identified the top 500 most dynamically expressed genes and used hierarchical clustering to group them into co-expressed modules.

In the laboratory isolates, these 500 genes were partitioned into six distinct modules, each exhibiting a unique expression profile along the IDC pseudotime. These modules represent waves of coordinated gene expression that define specific developmental stages. For instance, some modules peaked in early pseudotime (ring stage), while others were activated later (tropho-

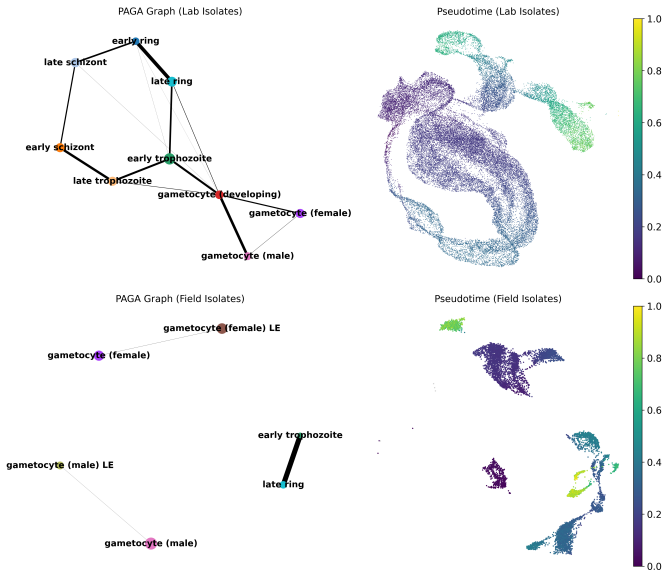


**Figure 4.** PAGA graphs and diffusion pseudotime plots for laboratory and field isolates. The PAGA graphs show connectivity between developmental stages, while the pseudotime plots illustrate the inferred developmental trajectory for each condition, revealing a full asexual cycle in the lab and a committed sexual development path in the field.

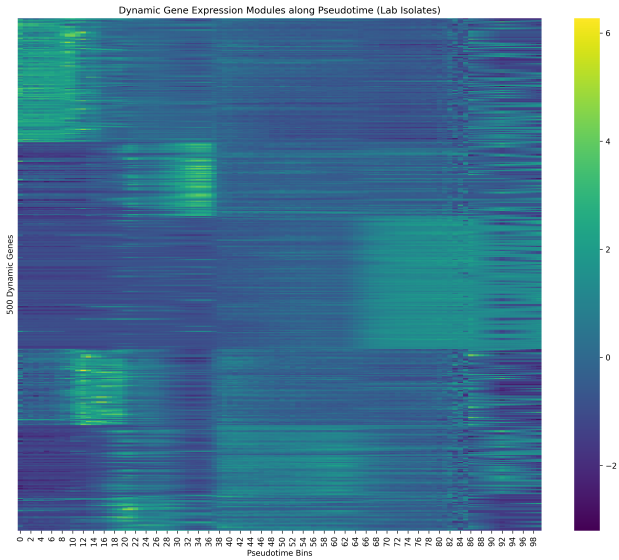


**Figure 5.** Heatmap showing the expression of the top 500 dynamic genes across pseudotime in laboratory isolates. Genes are clustered into six modules representing coordinated waves of expression during the asexual intraerythrocytic cycle. This captures the sequential transcription of genes required for parasite growth and development.

zoite/schizont stages), reflecting the sequential transcription of genes required for parasite growth, DNA replication, and egress. The module assignments are detailed in <code>gene_modules_lab.csv</code>, with module sizes ranging from 35 to 130 genes. The heatmap in Figure **??** shows the expression of these dynamic genes across pseudotime in laboratory isolates.

A similar analysis on the field isolates also yielded six co-expression modules across the ring-to-gametocyte pseudotime. These modules captured the transcriptional programs associated with sexual differentiation and maturation. Figure **??** shows a heatmap of the 500 most dynamically expressed genes in field isolates. The distinct sets of dynamic genes and module profiles between the lab and field datasets highlight the different biological processes being captured: broad IDC progression versus focused gametocytogenesis.

### 3.5. *Identification of Candidate Master Regulators Reveals Divergent Regulatory Strategies*

A central goal of this study was to identify candidate master regulators, such as transcription factors, that orchestrate stage transitions. We hypothesized that such regulators would be low-abundance genes whose expression transiently peaks immediately prior to the activation of a downstream gene module. We developed a systematic approach to identify these candidates by cross-correlating the expression profiles of low-

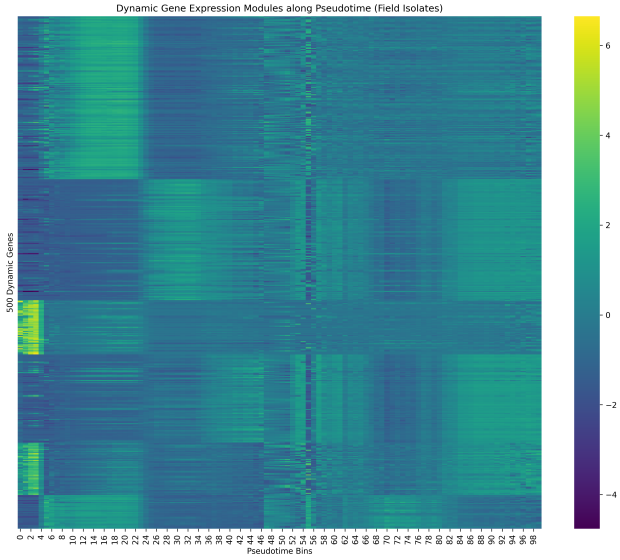


**Figure 6.** Heatmap of the 500 most dynamically expressed genes in field isolates, clustered into six co-expression modules across pseudotime. This reveals the transcriptional programs associated with sexual differentiation and maturation, which are distinct from those observed in lab isolates.

expression genes with the average profiles of the dynamic gene modules.

#### 3.5.1. *Laboratory Isolates: "Just-in-Time" Regulation*

For the **laboratory isolates**, this analysis yielded a striking and consistent result. The top 20 candidate regulators, including genes like `PF3D7-1361300` and `PF3D7-0311700`, were all strongly associated with the activation of a single downstream module: **Module 2**. Visualization of these relationships revealed a "just-in-time" regulatory logic. For example, as shown in Figure **??**, the relative expression of `PF3D7-1361300` shows a sharp, transient peak that precedes the activation of Module 2 by a very short pseudotime lag of 1-2 bins (out of 100). Figure **??** shows a similar trend for `PF3D7-0311700`. This pattern, observed across numerous candidates, suggests a tightly controlled system where regulators are expressed immediately before their targets are needed, likely to drive a specific, rapid transition during the lab-adapted IDC.

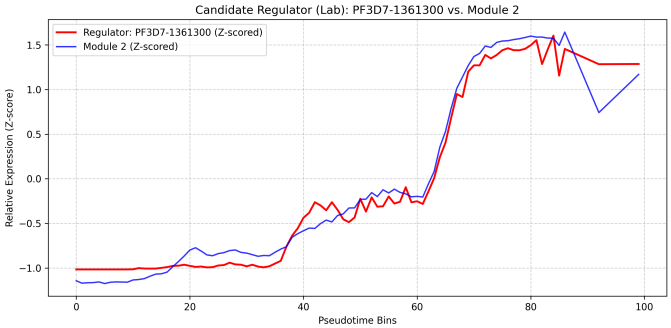


**Figure 7.** Expression dynamics of candidate regulator PF3D7-1361300 (red) and its target Module 2 (blue) in lab isolates, plotted against pseudotime. The regulator's expression transiently peaks just before Module 2 activation, suggesting a "just-in-time" regulatory logic in the lab-adapted asexual cycle.

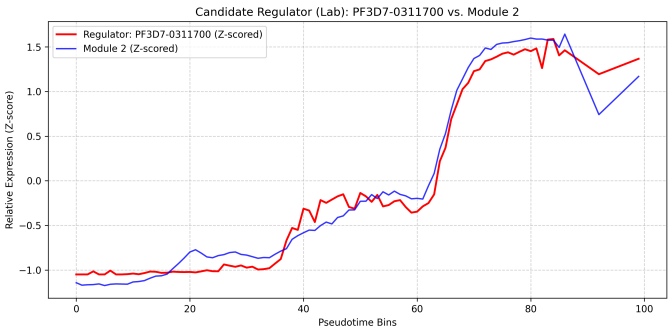


**Figure 8.** Expression of candidate regulator PF3D7-0311700 and its target Module 2, showing a short pseudotime lag between the regulator's peak expression and the activation of its target module in lab isolates.

### 3.5.2. *Field Isolates: "Priming" Regulation*

In sharp contrast, the analysis of **field isolates** uncovered a completely different regulatory paradigm. The top candidate regulators, such as `PF3D7-0916800` and `PF3D7-1472300`, were exclusively associated with the activation of **Module 5**. Most remarkably, the temporal relationship was profoundly different. As shown in Figure **??**, the expression peak of these candidate regulators preceded the activation of their target module by a very long lag of approximately 76-77 pseudotime bins. Figure **??** shows a similar trend for `PF3D7-1472300`. This suggests a "priming" or "delayed-action" model of regulation. In this scenario, regulators associated with sexual development are expressed very early in the captured trajectory (in the late-ring stage), long before their putative target genes in Module 5 are activated during late-stage gametocyte maturation.

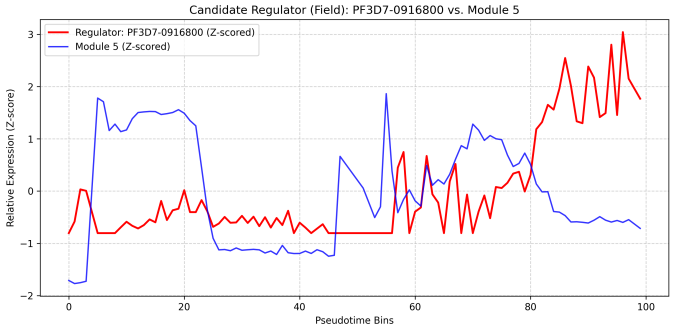


**Figure 9.** Expression profiles of candidate regulator PF3D7-0916800 and Module 5 in field isolates as a function of pseudotime, revealing a long-lag relationship where the regulator's expression precedes Module 5 activation, suggesting a "priming" regulatory mechanism for sexual development.

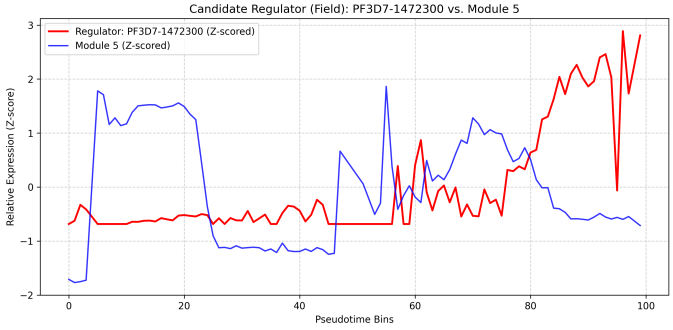


**Figure 10.** Expression profiles of candidate regulator PF3D7-1472300 and its target module (Module 5) in field isolates, showing that the regulator's expression peak precedes the activation of its target module by a long pseudotime lag, suggesting a "priming" regulatory mechanism.

### 3.5.3. *Divergent Regulatory Strategies: Implications*

This fundamental divergence in regulatory timing—short-lag, "just-in-time" activation in lab strains versus long-lag, "priming" activation in field isolates—represents a key finding. It suggests that parasites in a natural host environment may adopt a different strategy for developmental control, potentially maintaining a state of readiness for sexual commitment in response to host cues, a pressure absent from *in vitro* culture. This may reflect an essential adaptation for ensuring transmission in the complex and variable environment of a human host. The distinct modules implicated in each context (Module 2 in lab, Module 5 in field) further underscore that different sets of genes are being controlled by these divergent regulatory programs, reflecting the different biological priorities of proliferation versus transmission.

### 3.6. *Summary of Results*

Our comparative single-cell transcriptomic analysis of laboratory-adapted and field isolates of *P. falciparum* has revealed significant differences in their developmental dynamics and regulatory strategies. We observed that laboratory strains exhibit a continuous asexual developmental cycle, while field isolates are skewed towards sexual stages. Furthermore, candidate master regulators in laboratory strains show a "just-in-time" activation pattern, with expression preceding downstream gene expression by a short interval. In contrast, field isolates displayed a "priming" regulatory strategy, where regulators are expressed long before their target genes are activated. These findings suggest that *P. falciparum* adapts its stage progression control in response to the host environment, potentially reflecting an adaptation to ensure efficient transmission in the complex and variable environment of the human host.

## 4. CONCLUSIONS

The study addressed the problem of understanding differences in stage transition dynamics and regulatory strategies between lab-adapted and field isolates of *Plasmodium falciparum*, which is crucial for developing effective malaria control strategies. The research leveraged a single-cell transcriptomic dataset of 45,691 parasite cells, integrating lab strains and field isolates from asymptomatic patients. Trajectory inference, differential gene expression analysis, and candidate master regulator identification were employed to map developmental paths and identify key transition states.

The analysis revealed that lab isolates exhibited a continuous asexual developmental cycle with rapid stage transitions, while field isolates were skewed towards sexual stages. Differential gene expression analysis identified origin-specific genes involved in cell cycle progression, stress response, immune evasion, sexual commitment, gametocyte development, and transmission. Furthermore, divergent regulatory strategies were observed: a "just-in-time" mechanism in lab isolates and a "priming" strategy in field isolates.

The findings suggest that *P. falciparum* adapts its stage progression control in response to the host environment. The lab isolates prioritize rapid asexual replication, while field isolates prioritize sexual development and transmission. The "just-in-time" regulation in lab isolates allows for rapid and efficient stage transitions, while the "priming" strategy in field isolates prepares parasites for the complex and variable environment of the human host.

The study highlights the importance of studying parasites in their natural context to fully understand their biology and develop effective malaria control strategies. The differences in gene expression and regulatory strategies between lab and field isolates suggest that lab-adapted strains may not accurately reflect the behavior of parasites circulating in natural human populations. Therefore, future research should focus on studying field isolates to gain a more complete understanding of *P. falciparum* biology and develop more effective malaria control interventions.