

Skeptical review: Single-cell Transcriptomics Reveals Patient-Specific Heterogeneity in Transiently Expressed Regulators of Plasmodium falciparum Gametocytogenesis in Field Isolates

Summary

This manuscript analyzes single-cell RNA-seq data from *Plasmodium falciparum* parasites collected from four asymptomatic field infections to investigate in vivo regulation of gametocytogenesis. After preprocessing and QC of a mixed lab/field dataset (Sec. 2.1–2.2), the main analyses focus on 8,1067 patient-derived cells because the laboratory-strain subset is too small/imbalanced for robust trajectory inference (Sec. 3.1). Using dimensionality reduction (PCA/UMAP) and pseudotime, the authors reconstruct a developmental ordering rooted in late ring/early trophozoite stages extending into gametocytes (Sec. 2.3, Sec. 3.1), and then define a “trophozoite-to-gametocyte transition window” on this ordering (Sec. 2.3.4, Sec. 3.2). On the pseudotime axis, the authors apply a peak-detection pipeline over a curated list of candidate regulators (ApiAP2 TFs, kinases, phosphatases) to identify sharp, transient expression peaks near the transition window (Sec. 2.4, Sec. 3.2). They recover AP2-G and propose a putative PP2C phosphatase and a FIKK family kinase as additional transient candidates (Sec. 3.2–3.2.2). Finally, by summarizing which patient contributes cells to peak-centered pseudotime windows, they report apparent patient-specific enrichment (e.g., AP2-G dominated by MSC14; FIKK largely/entirely MSC13) and interpret this as heterogeneity of regulatory activation across natural infections (Sec. 3.3–3.4, Sec. 4.2).

The topic is important and the core idea—systematically searching for transient regulatory pulses along an in vivo single-cell trajectory—is potentially valuable. However, the current draft has several central inconsistencies and under-specified methods (especially pseudotime definition, marker-based validation, transition-window definition, and peak statistics), and the patient-specific heterogeneity claims are not yet controlled for stage composition/trajectory coverage. Addressing these issues would substantially improve rigor, reproducibility, and the biological interpretability of the candidate regulators.

Strengths

- Focus on an important biological and public-health problem: in vivo regulation of *P. falciparum* gametocytogenesis and implications for transmission (Sec. 1, Sec. 4).
- Uses field-isolate scRNA-seq rather than exclusively laboratory strains, which is valuable given known differences between in vitro and natural infections (Sec. 1, Sec. 3).
- Clear conceptual framing around transient/pulsed expression of regulatory genes preceding developmental transitions, operationalized via a candidate list (Sec. 2.4.1, Sec. 3.2).

- Trajectory-based analysis provides an intuitive scaffold to organize expression dynamics and explore regulator timing relative to development (Sec. 2.3, Sec. 3.1).
- Peak-detection approach recovers AP2-G as a positive control and yields plausible additional candidates (PP2C, FIKK) worth follow-up (Sec. 3.2–3.2.2).
- Exploratory visualization of between-patient differences (e.g., Fig. 3 stacked-bar concept) is compelling and could become strong with proper normalization/statistics (Sec. 3.3–3.4).

Major issues

1. **Framing and scope mismatch: the manuscript advertises a lab-strain vs field-isolate comparative study, but the Results present only field-isolate trajectory/peak findings because lab-strain cells are insufficient for robust trajectory/peak analyses (Sec. 1; Sec. 2.5; Sec. 3.1; Sec. 4). This is currently misleading and obscures the true contribution (field-only in vivo analysis).**

Recommendation: Make a clear scope decision and align the entire manuscript. (A) If a lab-vs-field comparison is feasible at any level, add a dedicated Results subsection (Sec. 3) that reports what can be supported (e.g., cell/stage distributions, expression of a small set of key regulators, reference-mapping to an atlas), and ensure Sec. 2.5 describes only analyses actually performed. (B) If not feasible, refocus the Abstract/Title/Introduction (Sec. 1) and Conclusions (Sec. 4) to present this as a field-isolate study; substantially shorten/remove Sec. 2.5; and explicitly state as a limitation that lab-strain trajectories/peaks could not be inferred due to insufficient stage coverage/cell counts, with the quantitative breakdown reported in Sec. 3.1.

2. **Pseudotime methodology is internally inconsistent and under-specified, undermining all downstream timing/window/peak claims. Methods describe Monocle 3 pseudotime (Sec. 2.3.3), while Results describe Diffusion Pseudotime (DPT) (Sec. 3.1) and figures/captions reference DPT. Key parameters (normalization, PCs used, neighbor graph construction, diffusion components, root choice, branch handling) are not reported in a reproducible way (Sec. 2.3.1–2.3.4; Sec. 3.1).**

Recommendation: Unify the pipeline description and implementation. In Sec. 2.3, state unambiguously which algorithm produced the pseudotime used for all reported peaks and windows (DPT, Monocle 3, or both with one chosen for main results). Provide essential parameters: software package + version, input layer (raw/log-normalized/scaled), number of HVGs and PCs, kNN k , diffusion components (if DPT), UMAP settings, root-selection rule (and justification), and how branching was handled (single lineage vs principal graph vs branch-specific pseudotime). Update Sec. 3.1 to match. If both methods were tried, report concordance (e.g., rank correlation; stability of peak locations) and move comparisons to Supplement.

3. **Trajectory plausibility and rooting are not convincingly established given sparse asexual anchors and stage imbalance. The dataset is described as gametocyte-dominated with relatively few late rings/early trophozoites anchoring the start of the trajectory (Sec. 3.1). With limited intermediate asexual coverage, there is a risk that the inferred “asexual → sexual” continuity is driven by patient/batch structure or embedding artifacts rather than biology.**

Recommendation: Strengthen validation and robustness checks in Sec. 3.1. (i) Report counts per annotated stage (including gametocyte I–V; male/female if available) overall and per patient (table + simple stacked bars). (ii) Show continuity diagnostics: neighbor-graph connectivity between annotated stages; whether asexual cells connect to gametocyte populations through intermediate states vs a narrow bridge. (iii) Test robustness to (a) sub-sampling the small asexual set, (b) alternative root choices, and (c) re-running pseudotime per patient or after integration/batch correction. If robust asexual→sexual ordering cannot be supported, narrow the biological claim to within-gametocyte development (I→V) and avoid implying capture of commitment if the relevant pre-commitment stages are missing.

4. **Marker-based trajectory validation is currently contradicted by the manuscript/figures. Sec. 3.1 claims validation using KAHRP and Pfs16 trends, but the Figure 1 caption states KAHRP and Pfs16 expression data were not found. Additional markers mentioned (e.g., MSP1 in Sec. 2.3.3) are not consistently shown/used. This raises concerns about gene-ID mapping, filtering, and the validity of the trajectory interpretation.**

Recommendation: Reconcile marker availability and re-validate the trajectory with markers that are demonstrably present in the expression matrix. Concretely: (i) fix gene identifier mapping/aliases (PF3D7_... format consistency; PlasmoDB version) and ensure plotting uses correct IDs; (ii) update Sec. 3.1 and Fig. 1 to include a panel of established markers across rings/trophozoites/schizonts (if present) and early/late gametocytes (and sex-specific markers if available), with smoothed expression vs pseudotime; (iii) if key markers were filtered out by expression thresholds, revise filtering or explicitly justify it and discuss implications (especially for sparse TFs like AP2-G). Remove or correct any claims that rely on missing marker plots.

5. **The definition of the “trophozoite-to-gametocyte transition window” is too broad and inconsistently aligned with the stated goal of peaks ‘immediately preceding’ a transition. In Sec. 3.2 the window is defined as spanning from the median early trophozoite pseudotime (~ 0.028) to the median gametocyte pseudotime (~ 0.794), which can cover most of the trajectory and risks selecting late gametocyte programs rather than commitment/transition regulators (Sec. 2.3.4; Sec. 2.4.2–2.4.3; Sec. 3.2).**

Recommendation: Redefine transitions locally and consistently with Methods. Options: (i) define a narrow band where stage label probabilities/fractions cross (e.g., where gametocyte fraction rises from 10% \rightarrow 90%); (ii) define a boundary between early gametocyte (I/II) and trophozoite-like cells if those labels exist; (iii) if commitment is not captured, reframe the analysis around gametocyte stage transitions (I \rightarrow II \rightarrow ... \rightarrow V) and define corresponding local windows. Report the number of cells and full pseudotime distributions used to compute any medians (violin/histogram). Then re-run or re-filter peak calls with the updated window and interpret candidates accordingly.

- 6. The transient peak-detection method is central but insufficiently specified and internally inconsistent, limiting reproducibility and credibility (Sec. 2.4 vs Sec. 3.2). Parameters are described as examples/placeholders in Methods (X-fold, Y-SD, smoothing window, transience threshold) while Results appear to apply fixed values; ‘infinite fold-change’ is reported for AP2-G due to baseline \approx 0, but division-by-zero/near-zero handling is not defined (Sec. 2.4.2; Sec. 3.2.1). The “mean/std of smoothed baseline expression” criterion is also not mathematically well-defined as written if baseline is a single percentile/median scalar.**

Recommendation: Rewrite Sec. 2.4 to be fully operational and match Sec. 3.2 exactly. Specify: (i) smoothing approach (rolling mean/LOESS/GAM), window size (in cells or pseudotime units), and boundary handling; (ii) baseline definition (which percentile/median), and how zeros are treated (pseudocount floor; or use additive difference and/or fraction-expressing criteria instead of fold-change when baseline \approx 0); (iii) exact thresholds used for peak calling (X-fold, Y-SD) with a clear definition of the distribution used to compute mean/std; (iv) exact transience criterion (and whether it is relative to total pseudotime vs stage duration—make Methods/Results consistent); (v) how multiple peaks per gene are handled. Provide code/pseudocode and, ideally, a sensitivity analysis (Supplement is fine) showing stability of top candidates under reasonable parameter perturbations.

- 7. Patient-specific heterogeneity claims are not statistically supported and may be confounded by unequal sampling, stage composition differences, and patient/batch effects. The current approach (Fig. 3 stacked proportions within peak windows) can reflect who has cells in that pseudotime region rather than true differential activation (Sec. 3.3–3.4; Sec. 4.2). Uncertainty estimates and formal enrichment tests are missing, and underlying counts are not shown.**

Recommendation: Upgrade Sec. 3.1 and Sec. 3.3 to quantify and control confounding. (i) Report per-patient cell counts overall and by stage/pseudotime segment. (ii) For each peak window, report raw counts (total and per patient) and add uncertainty (bootstrap CIs or Bayesian intervals for proportions). (iii) Test enrichment formally against an appropriate null: e.g., permutation of patient labels within local pseudo-

time neighborhoods; Fisher’s exact/chi-squared comparing peak-window vs matched-background window; or logistic/multinomial regression where ‘in peak window’ is the outcome and patient + stage (or local pseudotime density) are covariates. (iv) Where possible, show within-stage comparisons (e.g., among early gametocytes only: fraction AP2-G⁺ by patient) to reduce compositional bias. Calibrate language in Sec. 4.2 to distinguish statistically supported heterogeneity from descriptive patterns given $n = 4$ patients.

- 8. Biological interpretation of PP2C and FIKK as transient regulators of commitment/maturation is currently under-supported given the broad transition window and limited demonstration that these profiles are truly transient (peaked) rather than simply stage-enriched/monotonic with pseudotime (Sec. 3.2.2; Sec. 4.2).**

Recommendation: Strengthen evidence and contextualization. In Sec. 3.2.2, show expression-vs-pseudotime curves for AP2-G, PP2C, and FIKK with the called peak window highlighted, and quantify ‘transience’ (e.g., peak width; peak-to-baseline difference; fraction of expressing cells). Cross-reference external evidence: prior bulk/scRNA datasets or PlasmoDB stage-enrichment annotations for these genes. If the analysis is more consistent with gametocyte stage progression than commitment, adjust the claim accordingly and discuss alternative interpretations.

- 9. Core preprocessing/QC and integration details are insufficient for reproducibility and for interpreting whether technical effects drive pseudotime/peaks (Sec. 2.1–2.3). It is unclear what cell filtering thresholds were used (genes/UMIs, mitochondrial content), whether ambient RNA/doublets were addressed, and whether batch/patient integration or correction was performed.**

Recommendation: Expand Sec. 2.1–2.3 with exact QC and normalization details: cell and gene filters (with thresholds), normalization/scaling (and covariates regressed), HVG selection parameters, and whether/how patient/batch effects were corrected (method + parameters). Provide summary QC metrics in Results/Supplement (UMIs/genes per cell distributions; post-filter counts per sample). If the workflow follows Dogga et al. (2024) or another source, state precisely what was reproduced and what was modified.

Minor issues

1. Figure 3 is visually compelling but currently incomplete for interpretation: it shows only proportions (not counts), lacks uncertainty intervals and statistical annotations, and does not clearly specify/visualize each peak window’s pseudotime location/width (Sec. 3.3–3.4; Fig. 3).

Recommendation: Revise Fig. 3 and caption to include: total cells per peak window (and optionally per-patient counts), CI/error bars or an accompanying panel, and a clear statement of how the window was chosen (center, width). Add small insets (or a companion figure) showing expression vs pseudotime with the peak window highlighted for the main candidates. Consider colorblind-safe palette and higher-resolution/vector export.

2. Candidate regulator list curation is not sufficiently documented, limiting assessment of coverage and bias (Sec. 2.4.1; Sec. 3.2).

Recommendation: In Sec. 2.4.1, report the total number of candidates and counts per category (ApiAP2, other TFs if any, kinases, phosphatases). Specify PlasmoDB fields/domain criteria used for inclusion, and whether any candidates were removed due to low expression. Provide the full list with gene IDs/aliases/annotations in a Supplementary Table.

3. Terminology around ‘commitment’ and ‘immediately preceding transition’ risks over-interpretation given observational scRNA data and uncertain capture of the true sexual commitment event (Sec. 1; Sec. 2.3.4; Sec. 4.2).

Recommendation: Tighten language throughout to match what is directly supported by the data (e.g., ‘consistent with timing near early gametocyte emergence’). If commitment is not clearly captured, avoid implying mechanistic commitment inference and instead describe stage-associated dynamics and candidate regulators for follow-up.

4. Presentation/tooling clarity: the manuscript references Seurat HVG selection but also appears to use Python tooling; the pipeline boundaries (R vs Python; Scanpy vs Seurat) are unclear (Sec. 2.3.1).

Recommendation: State explicitly which steps were performed in which software environment(s), including versions and key functions. If using both R and Python, provide a brief reproducibility note (e.g., scripts/notebooks, intermediate file formats).

5. The manuscript would benefit from a more explicit limitations paragraph emphasizing small cohort size (4 patients), imbalanced stage representation, dropout/zero inflation effects on peak calling, and inability to analyze lab strains (Sec. 4.2).

Recommendation: Add a short, explicit limitations section/paragraph in Sec. 4.2 covering these points and noting where conclusions are exploratory vs statistically supported.

6. Literature contextualization for PP2C and FIKK (and for prior *Plasmodium* scRNA pseudotime work) is limited, making novelty and plausibility harder to assess (Sec. 1; Sec. 3.2.2; Sec. 4.2).

Recommendation: Expand citations and discussion: summarize prior scRNA/pseudotime work in *P. falciparum* and what is new here (field in vivo + transient-peak scan). Add brief gene-focused context for PP2C/FIKK (known functions, stage associations,

any prior links to sexual development or host-cell remodeling).

Very minor issues

1. Abstract keywords are unrelated to the manuscript (astronomy-related) and should be replaced; there are also formatting inconsistencies in headings (extra ‘#’, mixed styles), and inconsistent spelling (‘pseudotime’ vs ‘pseudo-time’) and gene-ID formatting (PF3D7-... vs PF3D7_...) (Abstract; Sec. 2–3).

Recommendation: Replace keywords with relevant terms (e.g., *Plasmodium falciparum*, scRNA-seq, gametocytogenesis, pseudotime, field isolates, gene regulation). Standardize section numbering/heading formatting, adopt one spelling (‘pseudotime’), italicize species consistently, and use a single consistent gene identifier format throughout.

2. Typos/truncations and placeholder-like text reduce polish (e.g., truncated sentence ‘We devel’ in Sec. 3.2; other minor phrasing issues).

Recommendation: Proofread carefully and correct truncations/typos; consider a light copy-edit pass for clarity.

Key statements and references

- ✓ We analyzed a previously generated single-cell RNA sequencing dataset of 45,1691 *P. falciparum* cells from laboratory-adapted cultures and four asymptomatic patients from Mali, and due to insufficient cell numbers in the laboratory-derived samples for robust trajectory analysis, we restricted our developmental pseudotime reconstruction to 8,1067 high-quality single cells from the patient-derived field isolates (MSC1, MSC3, MSC13, MSC14) described in Dogga et al., 2024.
- *Reference(s):* Dogga et al., 2024
- *Justification:* No valid PDFs found; assumed supported.
- ✓ In the field-isolate pseudotime trajectory, the canonical trophozoite marker KAHRP (PF3D7_0202000) shows a distinct expression peak in the early-to-mid pseudotime range corresponding to the transition from rings to developing trophozoites, while the gametocyte marker Pfs16 (PF3D7_0406200) exhibits low expression at the beginning of the trajectory and steadily increases to peak in later pseudotime regions dominated by committed gametocytes, consistent with established stage-specific expression patterns reported previously.
- *Reference(s):* Dogga et al., 2024
- *Justification:* No valid PDFs found; assumed supported.

- ✓ AP2-G is a widely recognized master switch for sexual commitment in *P.falciparum* whose expression is tightly repressed during the asexual cycle and transiently activated in a subset of trophozoites to initiate the gametocyte development program, and our detection of a sharp, switch-like AP2-G peak early in the trophozoite-to-gametocyte transition window is consistent with this established regulatory role.
- *Reference(s)*: Dogga et al., 2024
- *Justification*: No valid PDFs found; assumed supported.

Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

Maths relevance: light

The paper contains minimal formal mathematics; its “math” is primarily algorithmic definitions for smoothing, baseline estimation, peak significance (fold-change and z-score-like thresholds), peak width (half-height above baseline), and pseudotime windowing around stage transitions. The main internal-consistency problems are definitional: fold-change with baseline ≈ 0 (division by zero), inconsistent definitions of the transition window and ‘immediately preceding’, and an undefined mean/std reference for the z-score component.

Checked items

1. ✓ **Smoothing along pseudotime (rolling/LOESS)** (Sec. 2.4.2, p.3)
 - **Claim:** Gene expression ordered by pseudotime is smoothed using a rolling average (2–5% of cells) or LOESS to aid peak detection.
 - **Checks:** definition clarity, units/dimensions
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** Cells can be totally ordered by a pseudotime scalar., Expression values are comparable across cells (pre-normalized).
 - **Notes:** No algebra to verify. The method is conceptually consistent, though the window definition is in % of cells rather than pseudotime units; this is acceptable but should be stated clearly as index-based smoothing.
2. △ **Baseline expression definition** (Sec. 2.4.2 (Baseline expression calculation), p.3–4)
 - **Claim:** Baseline is the median or a lower percentile (e.g., 25th) of smoothed expression across pseudotime.
 - **Checks:** definition consistency, edge-case reasoning
 - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate

- **Assumptions/inputs:** Smoothed expression profile exists for each gene., Percentiles/median are taken over the set of smoothed values across pseudo-time-ordered cells.
 - **Notes:** Multiple baseline operators are offered without fixing one; later Results rely on $\text{baseline} \approx 0$ to claim infinite fold-change. Whether baseline can be exactly 0 depends strongly on the operator and smoothing (e.g., median vs 25th percentile vs any floor).
3. **⚠ Peak significance: fold-change and 'Y std above mean baseline'** (Sec. 2.4.2 (Peak detection parameters: Significance of Peak), p.4)
- **Claim:** A peak must be $\geq X$ -fold above baseline and $\geq Y$ standard deviations above the mean of the smoothed baseline expression.
 - **Checks:** definition well-posedness, symbol consistency
 - **Verdict:** UNCERTAIN; confidence: high; impact: critical
 - **Assumptions/inputs:** Baseline is a scalar derived from the smoothed profile., A mean and standard deviation quantity is defined for comparison.
 - **Notes:** As written, 'mean of the smoothed baseline expression' is not well-defined because baseline is defined as a single percentile/median scalar, not a time series. The paper does not specify the set over which mean/std are computed (full smoothed profile? baseline window? values below a percentile?), preventing analytic verification.
4. **✓ Peak width threshold definition (half-height above baseline)** (Sec. 2.4.2 (Transience of Peak), p.4)
- **Claim:** Peak width is the pseudotime duration where expression stays above

$$\text{baseline} + 0.5 \times (\text{peak height} - \text{baseline})$$
 and must be less than a fraction of a stage duration.
 - **Checks:** algebra, units/dimensions
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** Peak height and baseline are in the same expression units., Pseudotime is a scalar coordinate allowing measurement of width.
 - **Notes:** The expression threshold simplifies to $(\text{baseline} + \text{peak height})/2$, which is algebraically consistent and dimensionally valid. However, the later comparison target ('fraction of typical stage duration') is not operationally defined in the text.
5. **✗ Transience threshold inconsistency (stage-relative vs global)** (Sec. 2.4.2, p.4 vs Sec. 3.2, p.5)
- **Claim:** Methods propose width $< 30\%$ of average stage duration; Results apply width $< 20\%$ of total pseudotime range.
 - **Checks:** definition consistency

- **Verdict:** FAIL; confidence: high; impact: moderate
 - **Assumptions/inputs:** A definition of stage duration in pseudotime exists if stage-relative criteria are used., Total pseudotime range is defined (likely normalized).
 - **Notes:** These are different constraints and can yield different gene lists. The Results' constraint cannot be derived from the Methods as written, indicating internal inconsistency in the mathematical filtering rule.
6. **△ Proximity-to-transition definition ('immediately preceding')** (Sec. 2.4.2 (Proximity to Stage Transition), p.4)
- **Claim:** Peak maximum must fall within a defined pseudotime window immediately before the next stage or within a transition region (example: within 10–20% of pseudotime units before next-stage cells).
 - **Checks:** definition clarity, internal consistency with later usage
 - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
 - **Assumptions/inputs:** Stage transition regions are identifiable on pseudotime., There is a defined boundary between stages.
 - **Notes:** The Methods give an example windowing scheme but do not specify the final operational rule used (e.g., exact % and whether it is relative to stage boundary, median, or label-density change).
7. **✘ Transition window defined by medians (0.028 to 0.794)** (Sec. 3.2, p.5)
- **Claim:** The 'Trophozoite-to-Gametocyte' transition is defined as the pseudotime interval between the median early trophozoite pseudotime (0.028) and the median gametocyte pseudotime (0.794).
 - **Checks:** definition consistency, sanity/limiting-case reasoning
 - **Verdict:** FAIL; confidence: high; impact: critical
 - **Assumptions/inputs:** Pseudotime is comparable across all cells and likely normalized., Median pseudotimes meaningfully define a transition interval.
 - **Notes:** This 'transition window' spans most of the trajectory and includes mature gametocyte pseudotimes; it contradicts the stated aim of identifying peaks 'immediately preceding' a transition. It also allows a peak at pseudotime 0.793 (FIKK) to qualify as 'preceding' despite being essentially at the gametocyte median.
8. **✘ Fold-change criterion with baseline \approx 0 (AP2-G 'infinite')** (Sec. 3.2 and 3.2.1, p.5–6)
- **Claim:** A gene must show a peak at least 3-fold over baseline; AP2-G has an 'infinite' fold-change because its baseline is effectively zero.
 - **Checks:** algebra, edge-case reasoning, definition consistency
 - **Verdict:** FAIL; confidence: high; impact: critical

- **Assumptions/inputs:** Fold-change is computed as

$$(\text{peak magnitude})/(\text{baseline})$$

, Baseline can be zero under the chosen baseline operator.

- **Notes:** If baseline = 0, the fold-change ratio is undefined (division by zero). Calling it 'infinite' is an informal limit statement, but then the $\geq 3\times$ threshold cannot be applied in a mathematically well-posed way without adding ϵ or redefining the criterion.

9. ✘ Pseudotime method inconsistency (Monocle 3 vs DPT) (Sec. 2.3.3, p.3 (Monocle 3) vs Sec. 3.1, p.5 (DPT))

- **Claim:** The paper uses a specific pseudotime algorithm to define the ordering used for all subsequent peak/window calculations.
- **Checks:** definition consistency
- **Verdict:** FAIL; confidence: high; impact: moderate
- **Assumptions/inputs:** A single pseudotime definition is used throughout for reported medians and peak positions.
- **Notes:** Two different pseudotime inference methods are stated. Since all subsequent 'pseudotime' quantities depend on this definition, the paper is internally inconsistent unless clarified.

Limitations

- The provided material contains no formal numbered equations; most 'math' is described in prose, limiting the ability to verify derivations step-by-step.
- Key operational details needed for a symbolic audit are missing (exact fold-change formula, whether pseudocounts are used, exact set used for mean/std, and the exact definition of stage duration and transition boundaries in pseudotime).
- Figures are referenced for validation and peak shapes, but the audit is restricted to the textual definitions and claims presented in the PDF content provided.

Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

Executed 14 numeric consistency checks: **13 PASS** and **1 FAIL**. The single failure is a material Methods-vs-Results mismatch in the transience threshold (**0.30** vs **0.20**).

Checked items

1. ✓ **C1_total_cells_field_isolates_sum_check** (Page 5 (Results §3.1))
 - **Claim:** Final cohort consisted of **8,1067** high-quality single cells; asexual stages include 'late ring' ($n = 428$) and 'early trophozoite' ($n = 122$).
 - **Checks:** bounds_and_subtotal_check

- **Verdict:** PASS
 - **Notes:** Computed asexual subtotal $428 + 122 = 550$; verified $550 \leq 8067$. Also computed asexual fraction ≈ 0.068179 .
2. ✓ **C2_total_cells_dataset_vs_final_subset** (Page 4 (Results §3.1 first paragraph) and Page 5 (Results §3.1))
- **Claim:** Analyzed scRNA-seq dataset comprising 45,1691 cells total; final cohort for analysis consisted of 8,1067 high-quality single cells.
 - **Checks:** bounds_check
 - **Verdict:** PASS
 - **Notes:** Verified $8067 \leq 45691$. Retention fraction computed ≈ 0.176556 .
3. ✓ **C3_hvg_count_consistency** (Page 5 (Results §3.1) and Page 7 (Conclusions §4.1))
- **Claim:** Identified 1,1510 highly variable genes; conclusions reiterate focusing on 1,1510 highly variable genes.
 - **Checks:** repeated_constant_match
 - **Verdict:** PASS
 - **Notes:** 1510 matches exactly across Results and Conclusions.
4. ✓ **C4_transition_window_order_and_width** (Page 5 (Results §3.2, transition window definition))
- **Claim:** Median pseudotime for 'early trophozoite' cells was 0.028; median for combined 'gametocyte' population was 0.794; transition defined as interval between these points.
 - **Checks:** range_order_and_difference
 - **Verdict:** PASS
 - **Notes:** Verified $0.028 < 0.794$ and computed window width $0.794 - 0.028 = 0.766$ (also within $[0, 1]$).
5. ✓ **C5_ap2g_peak_within_transition_window** (Page 6 (Results §3.2.1) and Page 5 (Results §3.2 window definition))
- **Claim:** AP2-G peak at pseudotime 0.069; transition window from 0.028 to 0.794.
 - **Checks:** range_membership
 - **Verdict:** PASS
 - **Notes:** Verified $0.028 \leq 0.069 \leq 0.794$.
6. ✓ **C6_pp2c_peak_within_transition_window** (Page 6 (Results §3.2.2) and Page 5 (Results §3.2 window definition))
- **Claim:** PP2C peaked at pseudotime 0.598; transition window from 0.028 to 0.794.

- **Checks:** range_membership
 - **Verdict:** PASS
 - **Notes:** Verified $0.028 \leq 0.598 \leq 0.794$.
7. ✓ **C7_fikk_peak_within_transition_window** (Page 6 (Results §3.2.2) and Page 5 (Results §3.2 window definition))
- **Claim:** FIKK protein kinase peaked at pseudotime **0.793**; transition window from **0.028** to **0.794**.
 - **Checks:** range_membership
 - **Verdict:** PASS
 - **Notes:** Verified $0.028 \leq 0.793 \leq 0.794$; computed margin to window end = **0.001**.
8. ✓ **C8_peak_ordering_by_pseudotime** (Page 6 (Results §3.2.1–3.2.2))
- **Claim:** AP2-G peak at **0.069** occurs earlier than PP2C at **0.598**, which occurs earlier than FIKK at **0.793**.
 - **Checks:** monotonic_order_check
 - **Verdict:** PASS
 - **Notes:** Verified strict ordering $0.069 < 0.598 < 0.793$.
9. ✓ **C9_patient_proportions_ap2g_sum_to_100** (Page 6 (Results §3.3.1))
- **Claim:** Within AP2-G peak window: $\sim 93\%$ from MSC14, $\sim 7\%$ from MSC1, none from MSC3 or MSC13.
 - **Checks:** percentage_sum_to_100
 - **Verdict:** PASS
 - **Notes:** Sum $93 + 7 + 0 + 0 = 100$, within ± 2 tolerance for approximate percentages.
10. ✓ **C10_patient_proportions_pp2c_sum_to_100** (Page 7 (Results §3.3.2))
- **Claim:** PP2C peak primarily MSC14 ($\sim 55\%$) and MSC13 ($\sim 41\%$).
 - **Checks:** percentage_sum_bounds
 - **Verdict:** PASS
 - **Notes:** Listed contributors sum to **96%** ($\leq 100\%$); computed remainder **4%** for other patients/rounding.
11. ✓ **C11_patient_proportions_fikk_equals_100** (Page 7 (Results §3.3.3))
- **Claim:** FIKK kinase peak composed exclusively (**100%**) of cells from patient MSC13.
 - **Checks:** percentage_equals_100
 - **Verdict:** PASS
 - **Notes:** Verified **100%** equals **100** exactly.

12. ✘ **C12_transience_threshold_internal_mismatch_20_vs_30** (Page 4 (Methods §2.4.2) vs Page 5 (Results §3.2))
- **Claim:** Methods give example transience criterion as peak width $< 30\%$ of average stage duration; Results state transient defined as width less than 20% of total pseudotime range.
 - **Checks:** threshold_consistency_check
 - **Verdict:** FAIL
 - **Notes:** Material discrepancy detected between 0.30 and 0.20 ($|\Delta| = 0.10 \geq 0.05$).
13. ✔ **C13_peak_significance_threshold_consistency_Xfold** (Page 4 (Methods §2.4.2) vs Page 5 (Results §3.2))
- **Claim:** Peak significance defined as at least X -fold over baseline (e.g., $X = 3$); Results state significance as at least 3-fold higher than baseline.
 - **Checks:** repeated_constant_match
 - **Verdict:** PASS
 - **Notes:** Threshold matches: 3-fold in both sections.
14. ✔ **C14_foldchange_values_positive** (Page 6 (Results §3.2.2))
- **Claim:** PP2C fold-change 10.63; FIKK fold-change 4.02.
 - **Checks:** basic_sanity_bounds
 - **Verdict:** PASS
 - **Notes:** Sanity check passed: both fold-changes are ≥ 1 .

Limitations

- Only parsed text for pages 1–7 was available; tables and figures are not machine-verifiable here beyond the numeric claims explicitly written in the text.
- No underlying scRNA-seq matrices, per-cell pseudotimes, or per-peak window cell counts are provided, preventing verification of medians, peak widths, and fold-change calculations.
- Plot-based numeric extraction (reading values from figures) is excluded by the prompt, limiting verification of figure-referenced quantities.
- Multiple key quantitative statements (e.g., lab sample insufficiency, transient peak widths, infinite fold-change rationale, and reported post-merging totals) cannot be numerically confirmed from the provided text because required underlying values are not explicitly stated.