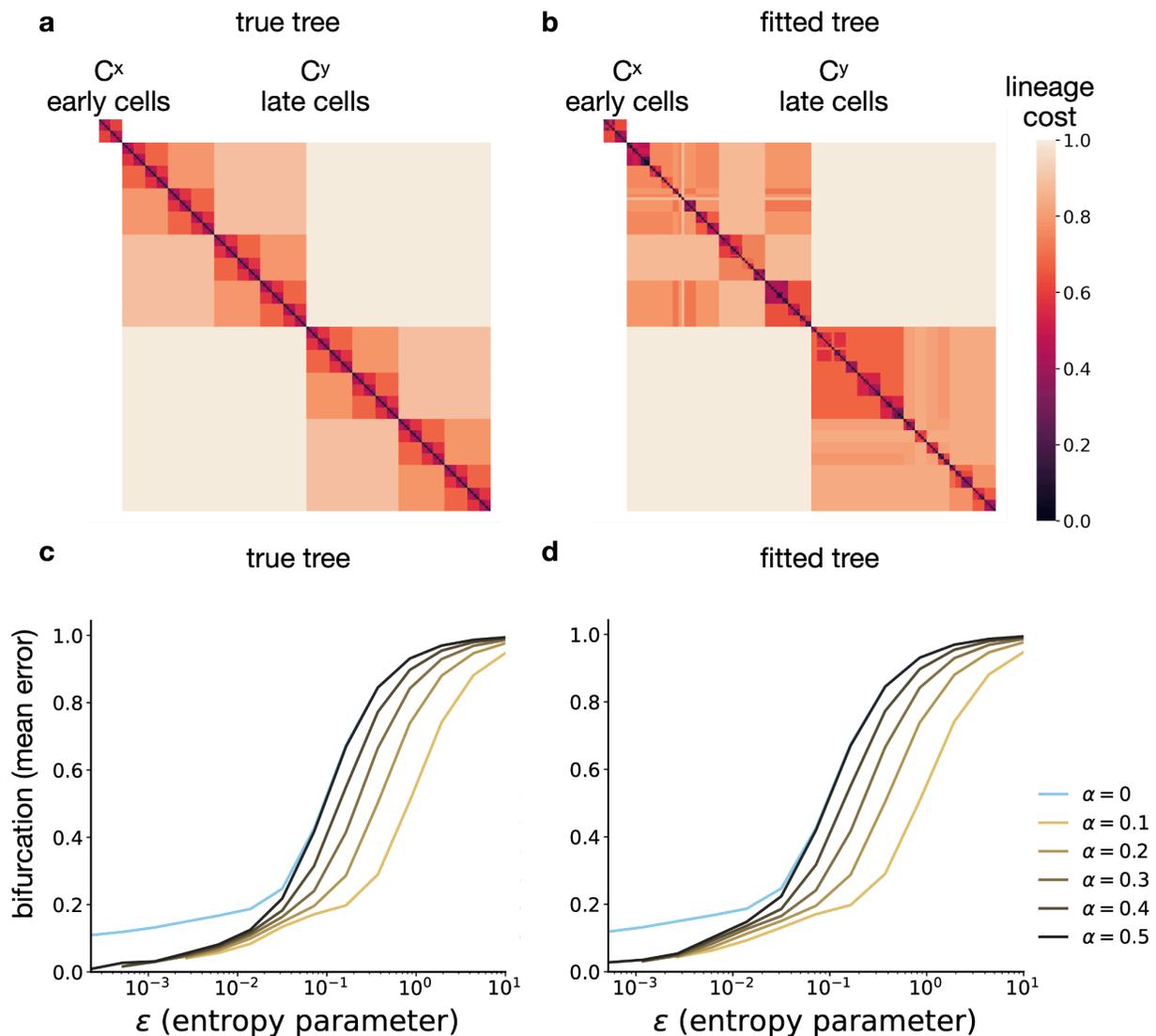


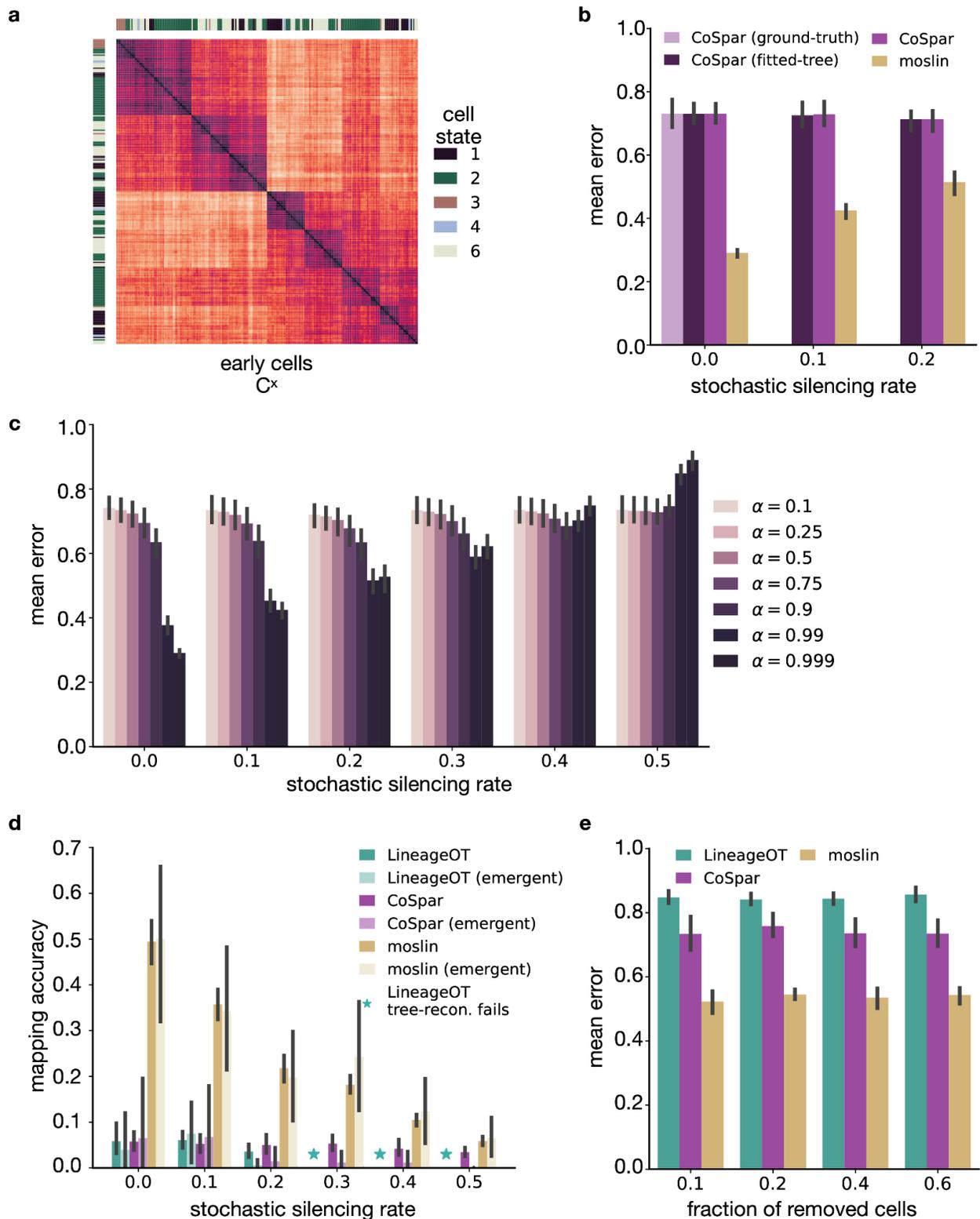
## Additional file 1: Supplementary Figures



**Fig. S1 | Performance over simulated datasets.**

The bifurcation trajectory as an illustrative example of moslin's performance. **a,b.** Heatmaps, visualizing the lineage distance cost matrices  $C^x$  and  $C^y$ , corresponding to lineage distance between early and late cells, respectively, as used by moslin in the bifurcation trajectory setting for the true tree (**a**) and the fitted tree (**b**). Higher cost values imply cells are further apart in the tree. **c,d.** Line plots, illustrating moslin's bifurcation-trajectory for the true tree (**c**) and the fitted tree (**d**), across a range of  $\alpha$  values. The  $\alpha = 0$  case corresponds to only using gene expression information. The x- and y-axis display the entropy parameter and the mean over ancestor and descendant errors, respectively.

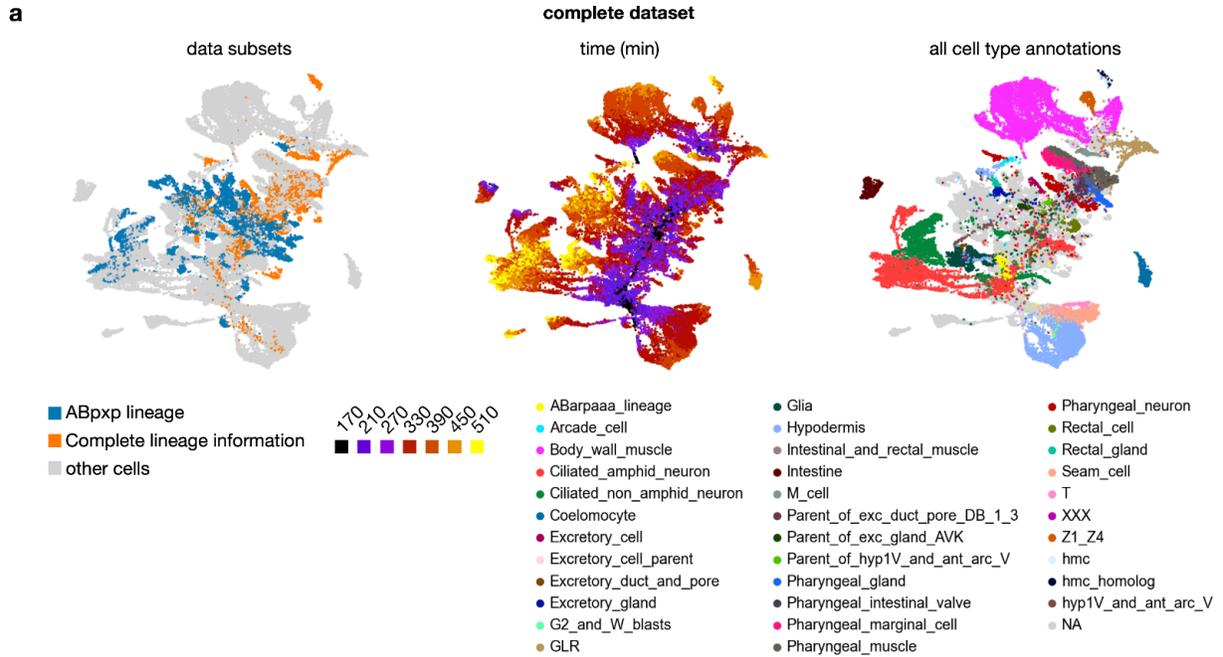




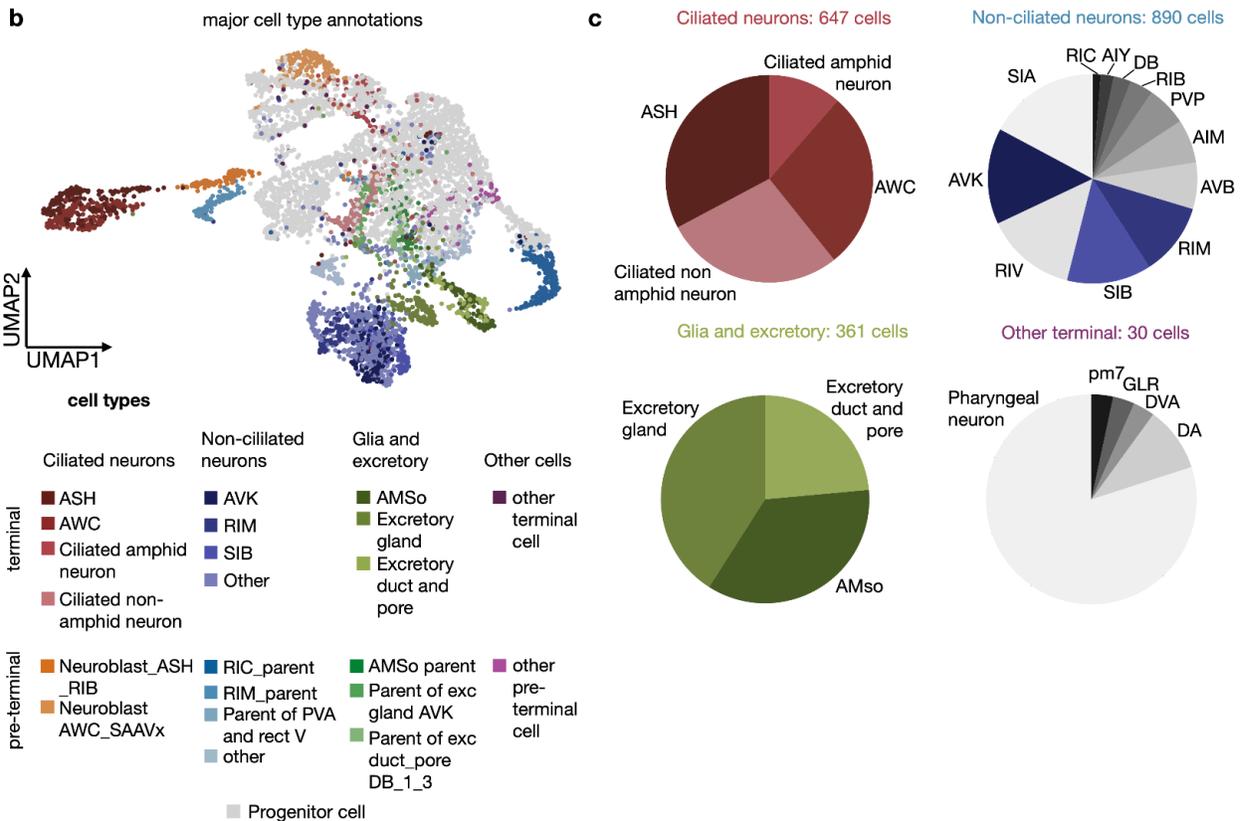
**Fig. S2 | Performance over the TedSim simulated datasets.**

**a.** Heatmaps, visualizing the Hamming distance-based lineage cost matrix in a TedSim simulation. Cells are sorted according to order of appearance in the true tree. Cell states are

depicted as provided by the simulation data. **b.** Mean prediction error of CoSpar based on different forms of lineage information, compared to moslin over the TedSim simulated dataset. The different forms of lineage information presented for CoSpar are the original ground-truth coupling, available only for  $ssr = 0$  (CoSpar, ground-truth), LineageOT tree fitting (CoSpar, fitted-tree), and barcode-based distances (CoSpar). **c.** Mean prediction error of moslin across a range of  $\alpha$  values, where  $\alpha = 0$  implies only using gene expression information and  $\alpha = 1$  only lineage information. **d.** The mapping accuracy of emergent states, cell states that only appear at the later time point, compared to non-emergent states. For a cell at a late time point an accurate mapping is a mapping in which the most probable ancestor corresponds to its ancestor in the ground truth tree. **e.** The mean prediction error of moslin compared to CoSpar and LineageOT, as a function of the number of removed cells from the late time point. Results are shown for  $ssr = 0.2$ . Error bars depict the 95% confidence interval across 10 random simulations.



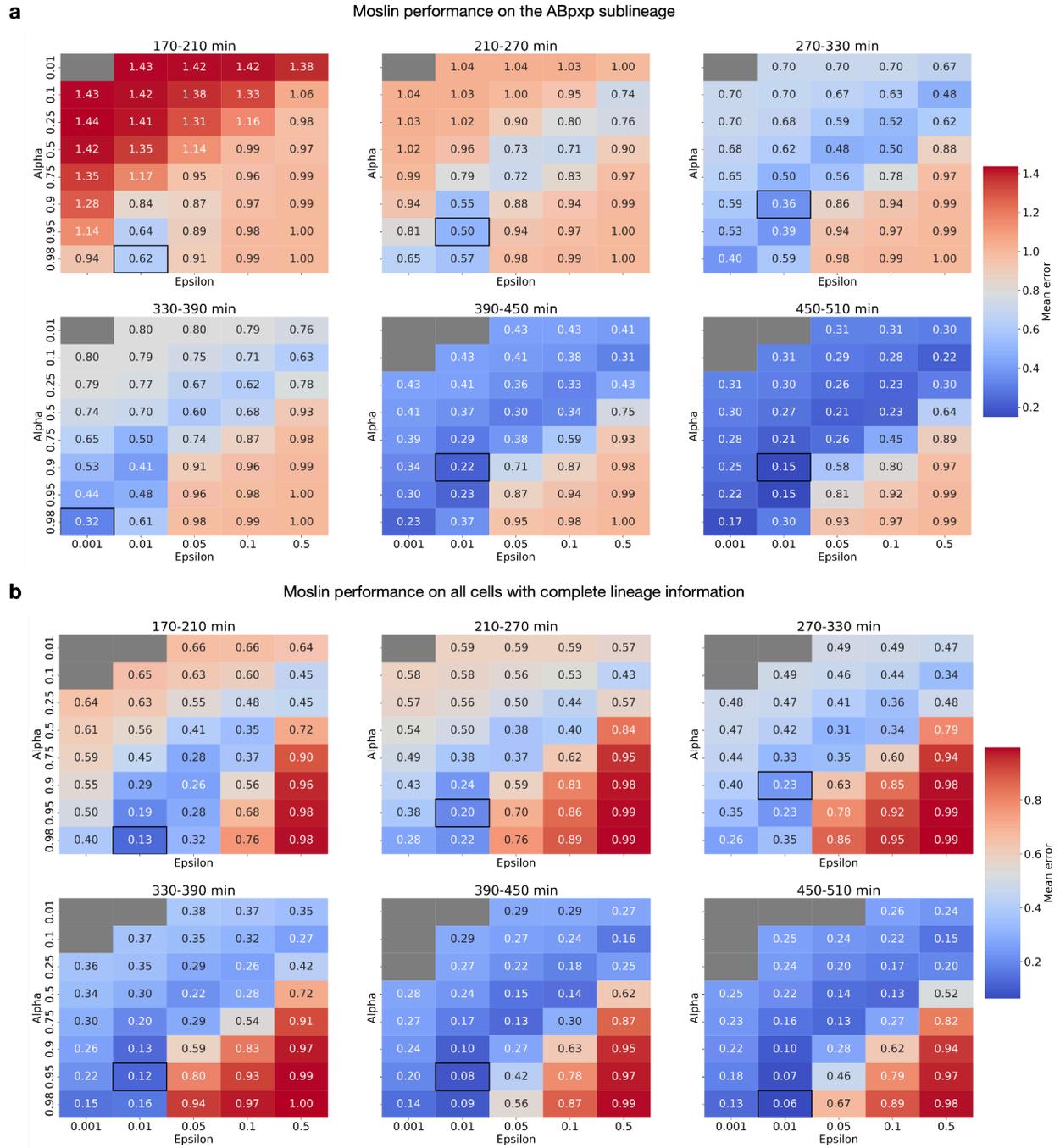
**focus on the ABpxp lineage**



**Fig. S3 | Cluster and lineage labels for the *C. elegans* data.**

**a.** UMAPs of 45,732 *C. elegans* cells with at least partial lineage information, colored by data subset used in this study (left), time point (middle), and cell type annotations from ref.<sup>7</sup> (right). **b.**

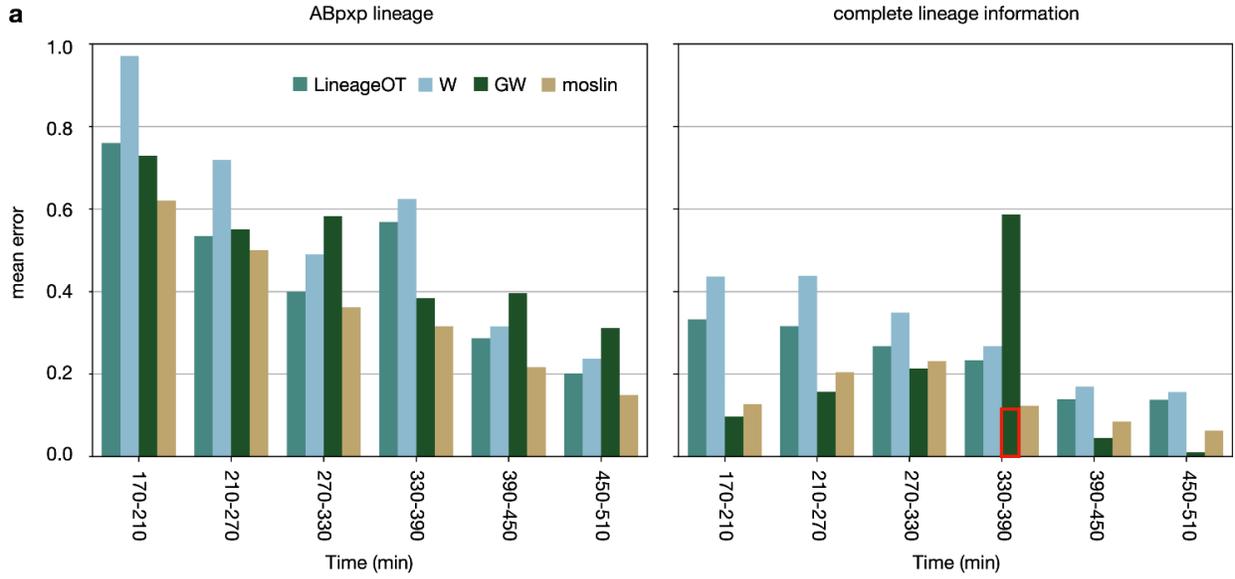
UMAP for cells of the ABpxp lineage, colored by cluster annotations for Ciliated neurons, Non-ciliated neurons and Glia and excretory cells **c.** Pie charts visualizing the makeup of terminal Ciliated neurons, Non-ciliated neurons, Glia and excretory, and other terminal cells.



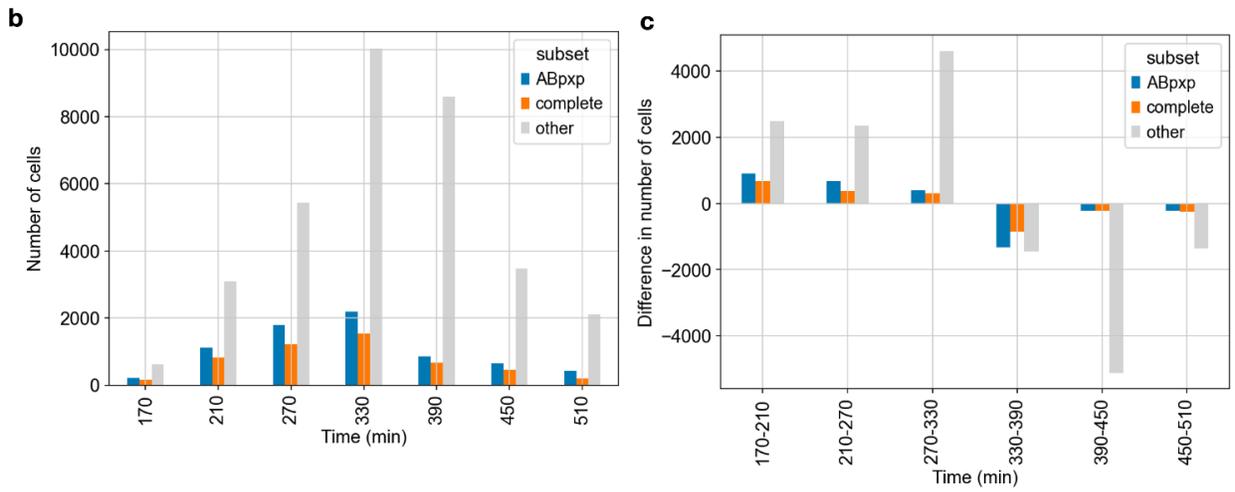
**Fig. S4 | Optimal hyperparameters for moslin on *C. elegans* embryogenesis.**

**a-b.** Heatmaps displaying the mean over ancestor and descendant errors for moslin on cells from the ABpxp sublineage (**a**) and cells with complete lineage information (**b**). Grey color indicates non-converged runs, black boxes highlight optimal hyperparameter combinations.

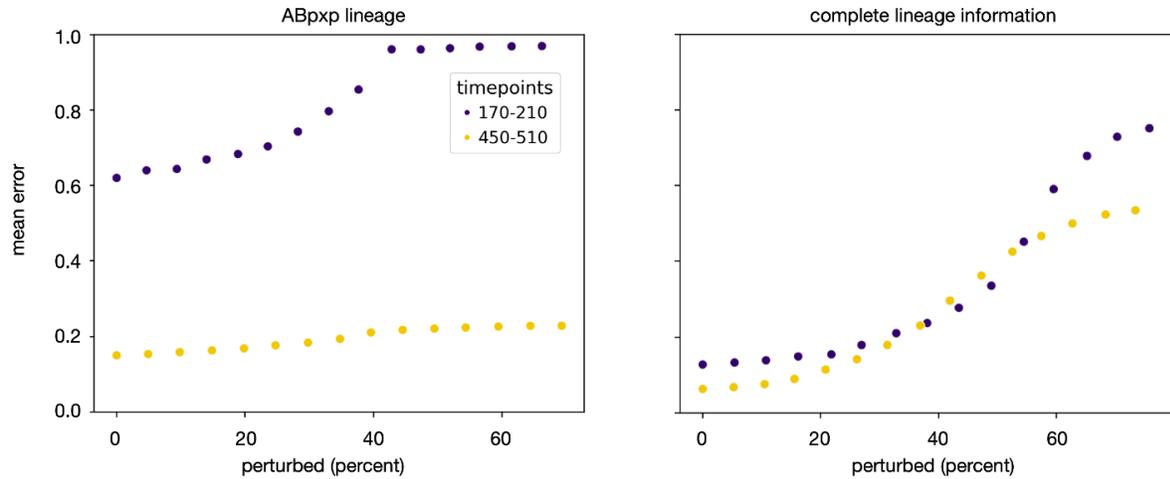
benchmark results



cell numbers



perturbing cost matrix elements

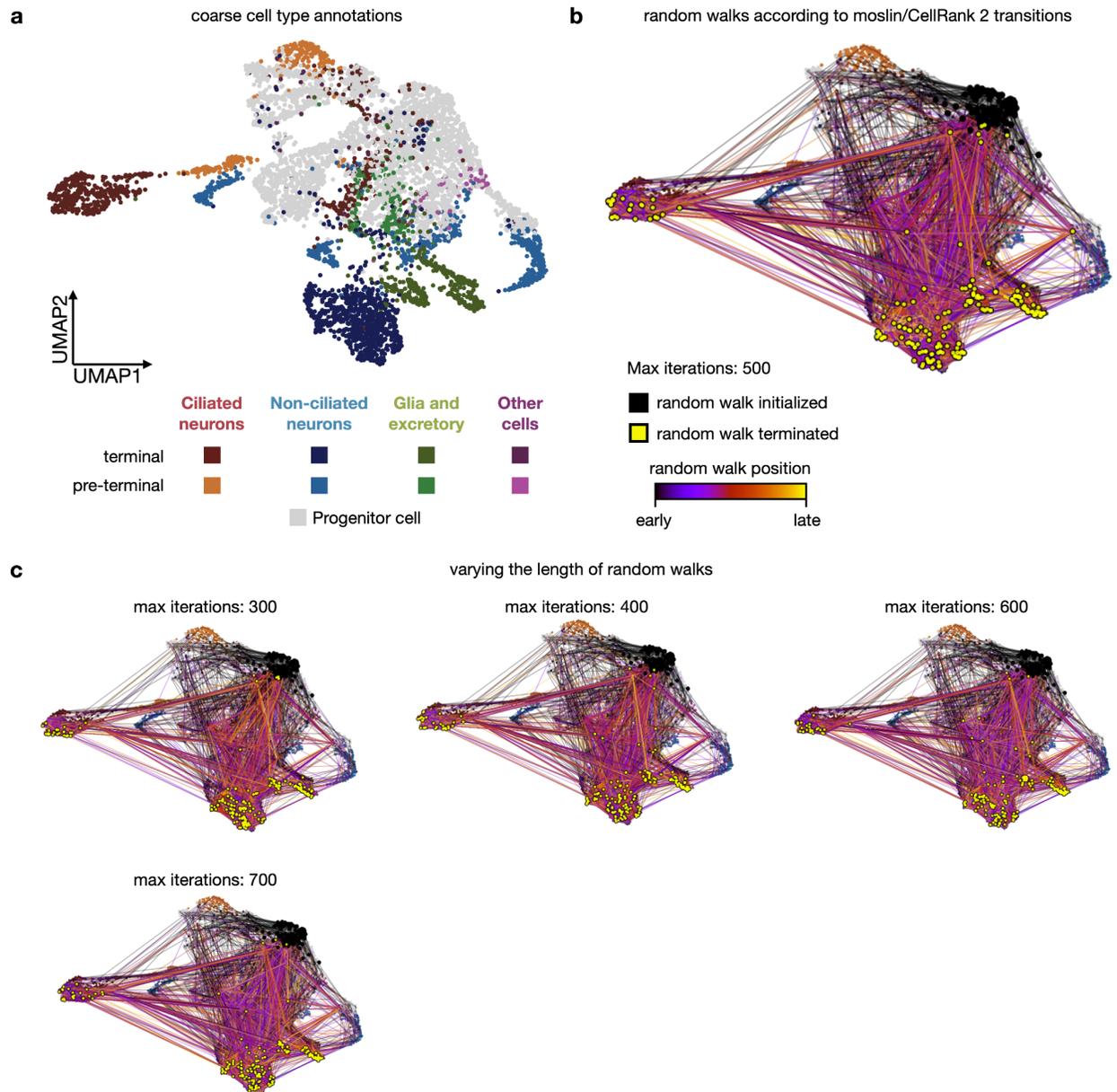


**Fig. S5 | moslin accurately recovers lineage relationships on *C. elegans* data**

**a.** Bar chart of the mean error for LineageOT<sup>21</sup>, W, GW, and moslin across time points for ABpxp cells (left) and cells with complete lineage information (right). For cells with complete lineage information, the red box on the 330/390 min time point pair shows GW's performance when initialized with the W solution; GW in that case slightly outperforms moslin in terms of the mean error (0.11 to 0.12). **b.** Bar chart over the number of cells per time point, grouped by data subset. **c.** Bar chart over the difference in cell number between two adjacent time points, grouped by data subset. **d.** Scatter plots, visualizing the dependency of the mean error on the fraction of permuted lineage-distance cost matrix elements for ABpxp cells (left) and cells with complete lineage information (right) across two pairs of time points.

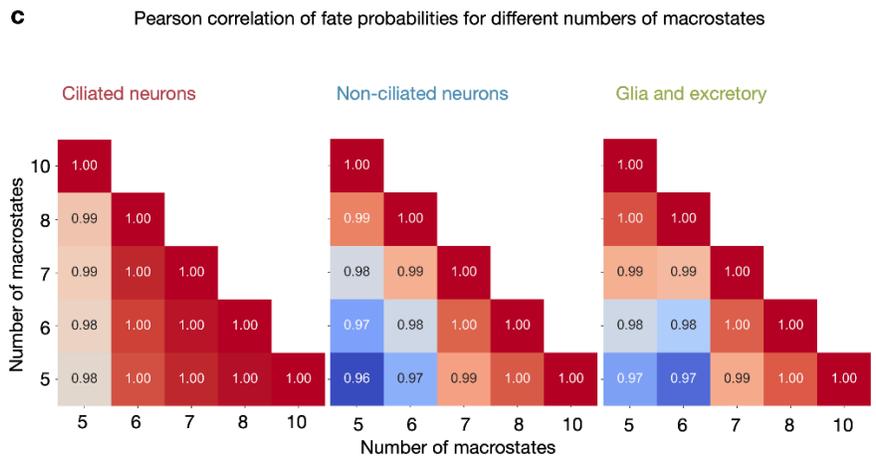
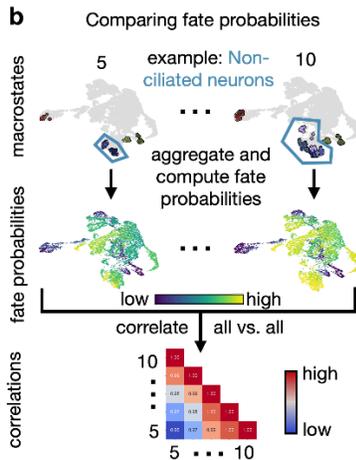
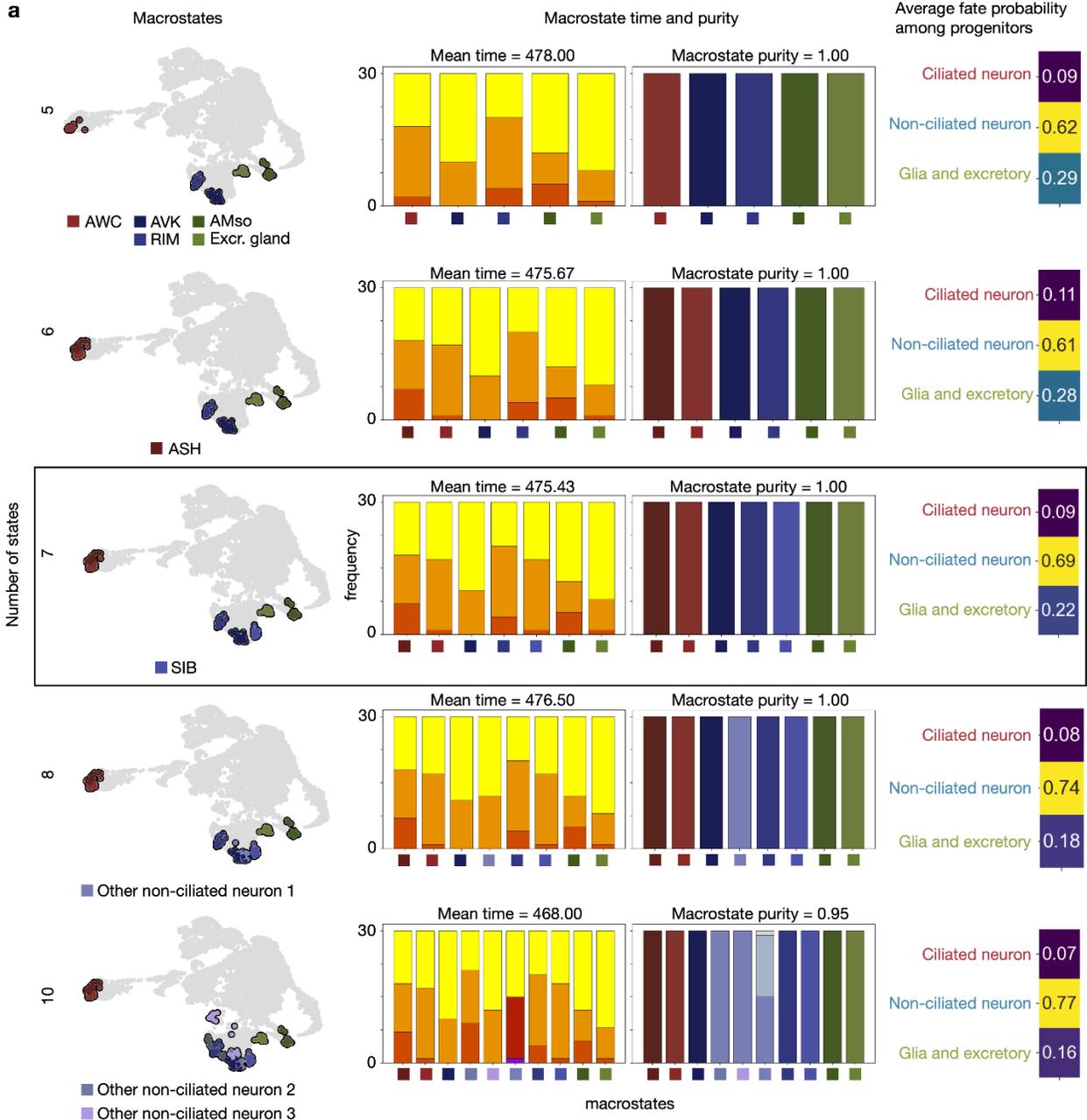


(leftmost) and predicted (right) 330 min ancestors of 390 min RIM cells by different methods. 390 min cells are colored in light gray, 330 min cells are colored according to their ancestor likelihood. Pie charts visualize the aggregated ancestor distribution over the clusters shown in (a), using the color scheme of (a). “error” indicates the aggregated ancestor error over 390 min RIM cells.



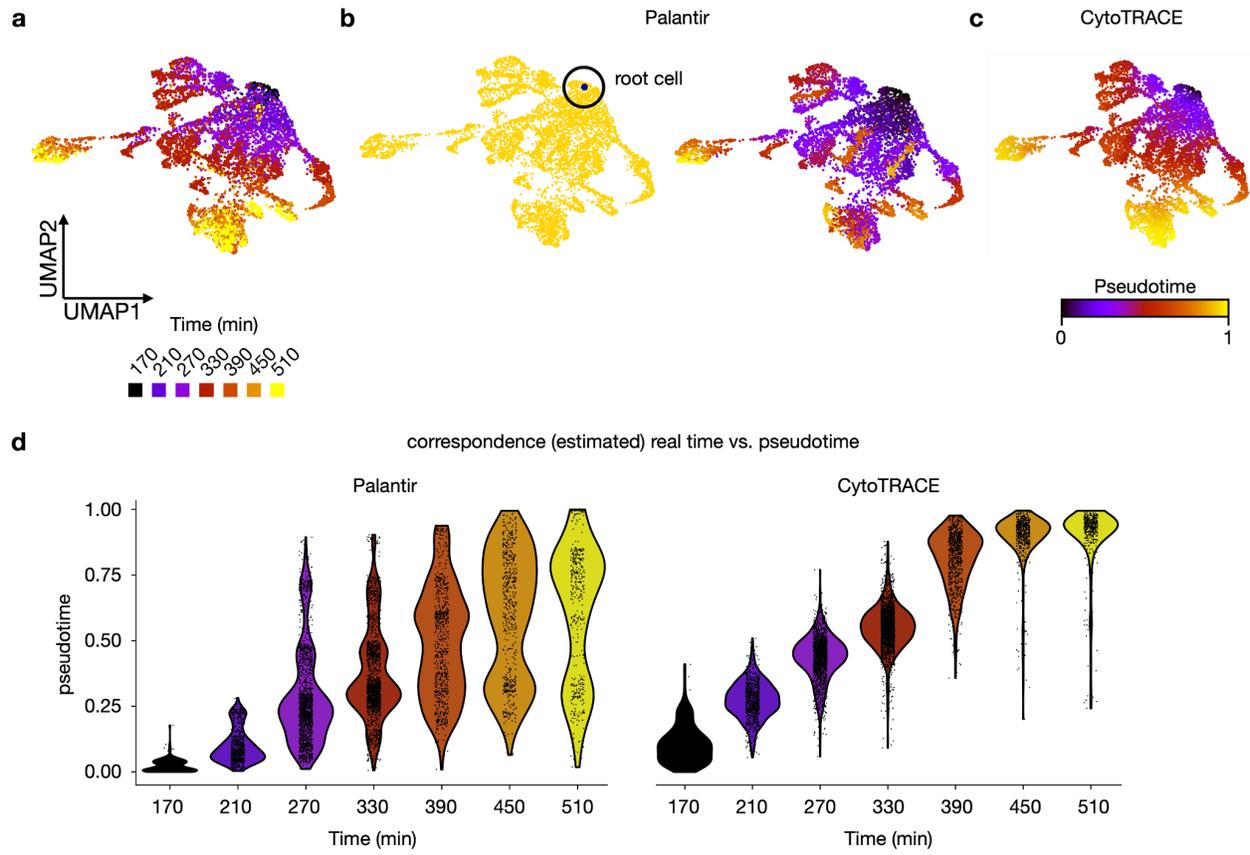
**Fig. S7 | Random walks recapitulate the known developmental hierarchy.**

**a-c.** UMAP embeddings of the ABpxp lineage, colored by coarse cell-type annotations. **b.** Additionally, we show 500-step random walks on the moslin/CellRank 2 computed transition matrix, initialized in 170 min cells (Methods). Random walks progress towards the expected terminal populations. **c.** Same as **(b)**, but varying the number of steps per random walk to showcase robustness.



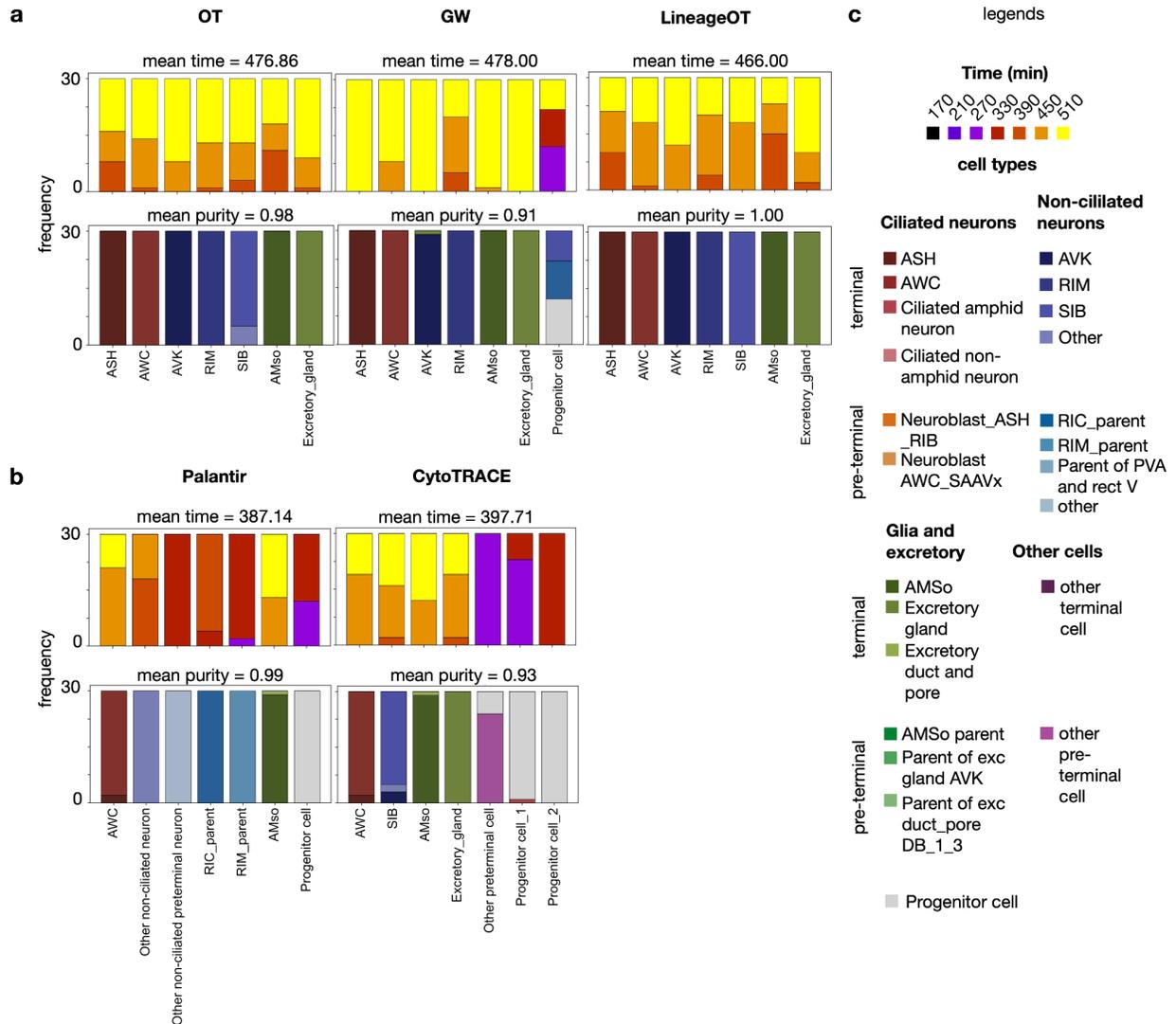
**Fig. S8 | moslin/CellRank 2 terminal states.**

**a.** Visualizing terminal states (left), their composition (middle), and average fate probabilities (right) for different numbers of macrostates (rows). UMAPs in the left column show the top 30 cells per moslin/CellRank 2 computed terminal state. For row, we indicate which macrostates emerged in addition to existing macrostates. Bar charts in the middle column show how the top 30 cells per terminal state distribute across time points and clusters. Timepoints and clusters are colored according to Fig. 3. Heatmaps in the left column show average fate probability among progenitor cells. The black box highlights seven terminal states, used throughout this manuscript. **b.** Cartoon, illustrating how we compare fate probabilities towards aggregated terminal states (example: Non-ciliated neurons) by correlating fate probabilities towards that state, computed over varying numbers of terminal states. **c.** Heatmaps, displaying Pearson correlation among fate probabilities towards aggregated terminal states, computed for varying numbers of macrostates as in **(a)**, following the procedure outlined in **(b)**.



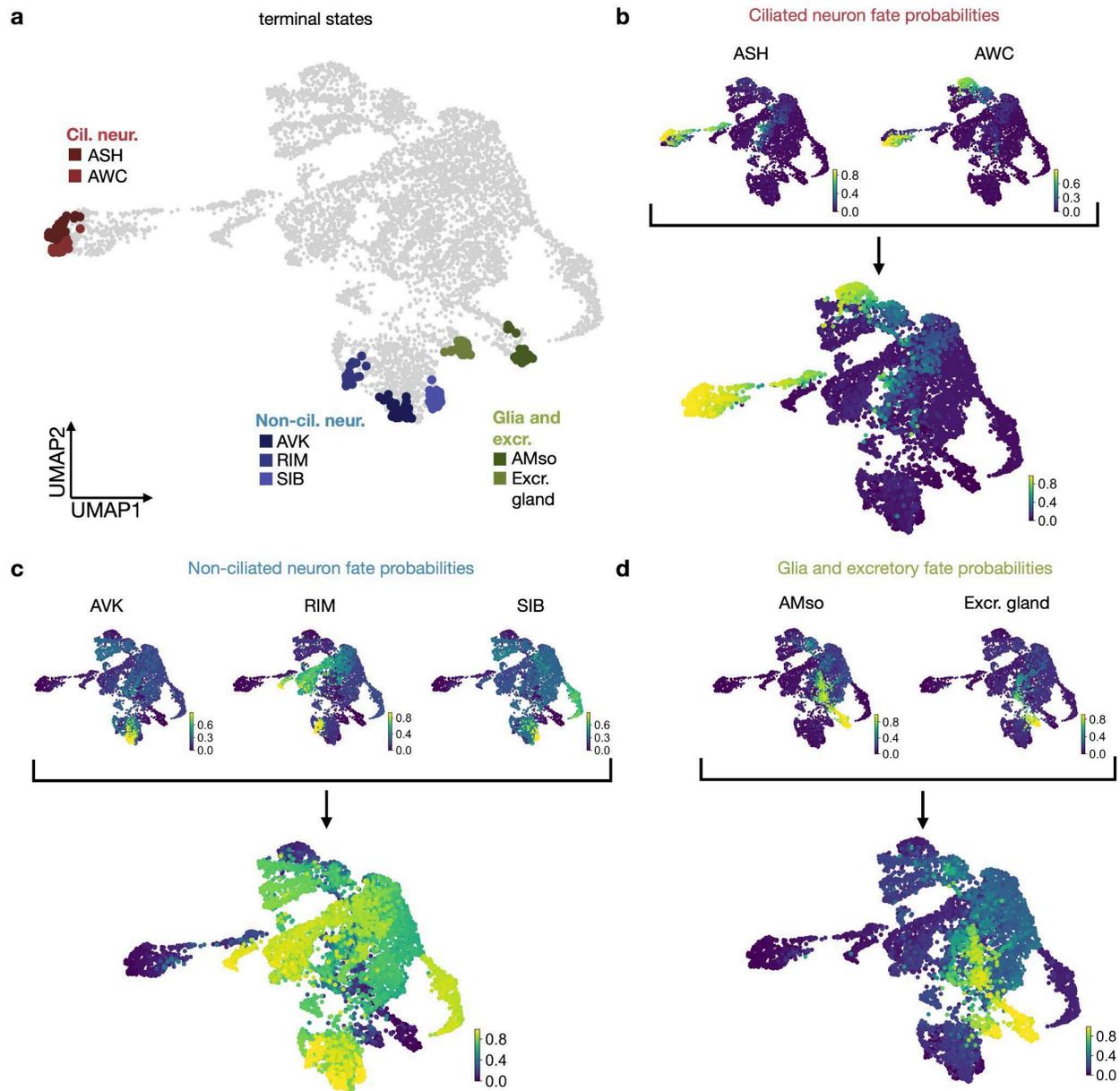
**Fig. S9 | Computing pseudotimes using Palantir and CytoTRACE.**

**a-c.** UMAPs of the ABxpx lineage, colored by (from left): **(a)** estimated real time points, **(b)** 170 min cell passed to Palantir as root cell (left), and the Palantir pseudotime (right)<sup>57</sup>, **(c)** the CytoTRACE pseudotime<sup>71,72</sup> (Methods). **d.** Violin plot over the Palantir (left) and CytoTRACE (right) pseudotimes, grouped by estimated real time points.



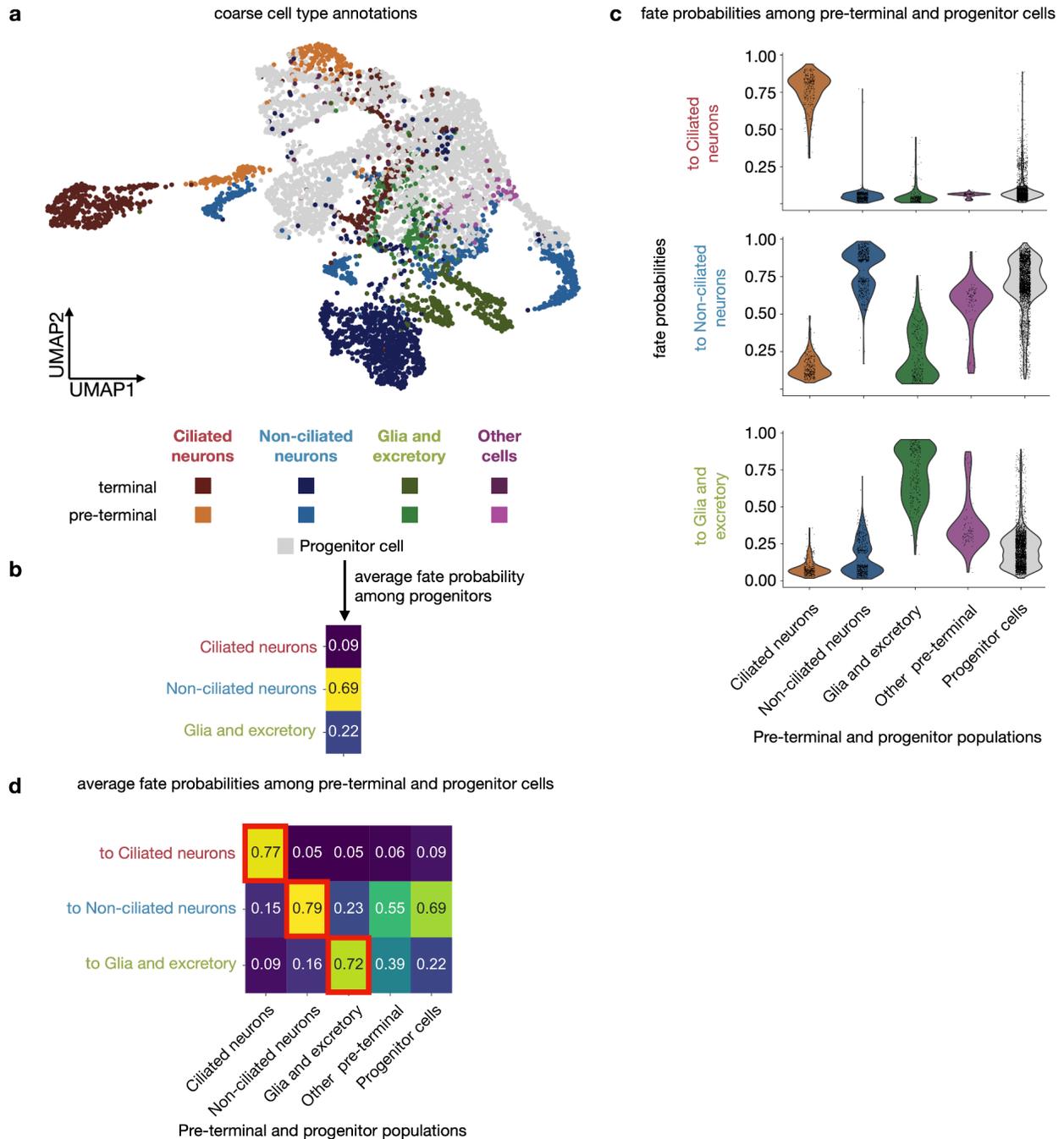
**Fig. S10 | Benchmarking terminal state cell distributions.**

**a,b.** Bar charts, showing how the top 30 cells per terminal state distribute across time points (top rows) and clusters (bottom rows) for methods that use (a) or do not use (b) time point information. In the titles, we indicate the mean macrostate time and purity (Methods). **c.** Time point (top) and cell type (bottom) legends.



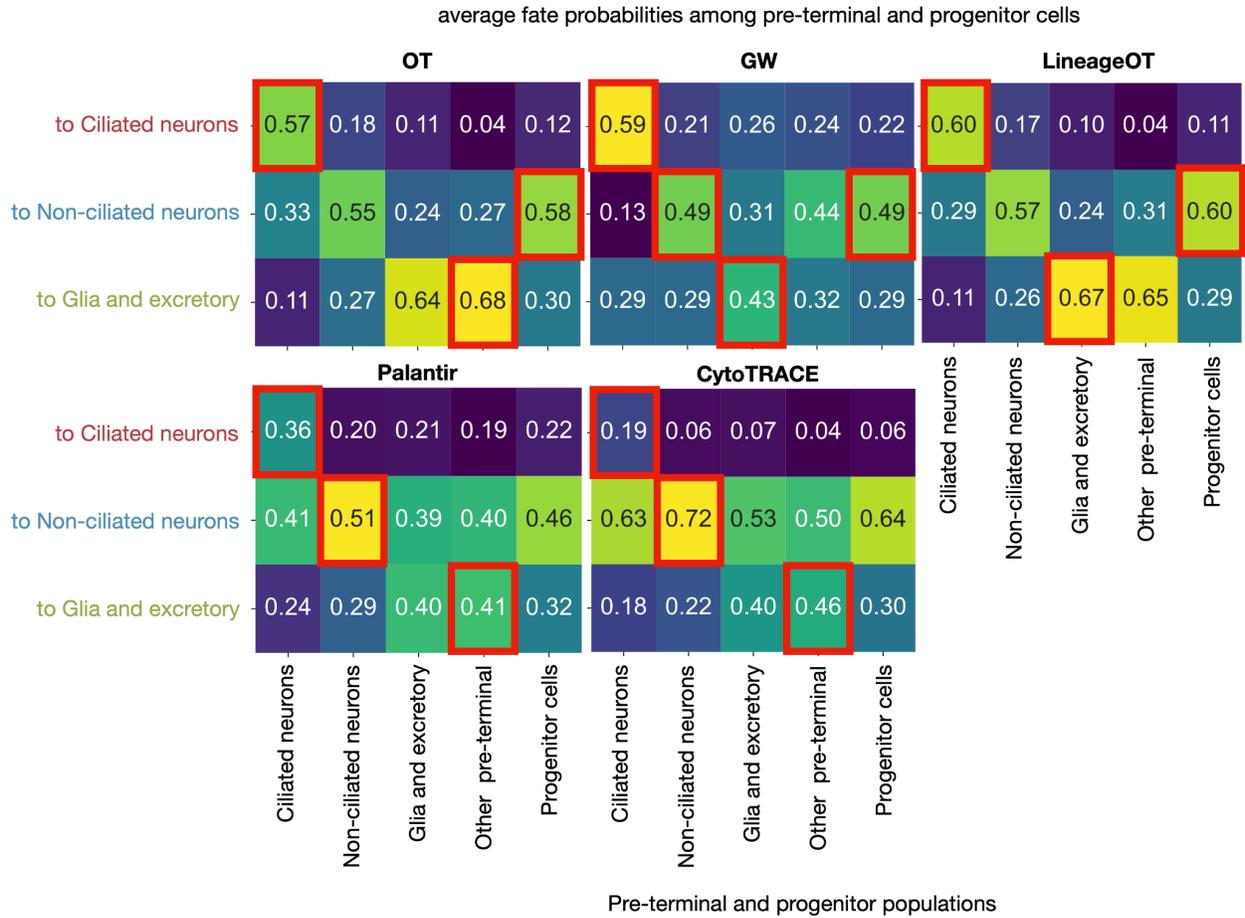
**Fig. S11 |Aggregating fate probabilities towards the main terminal groups.**

**a.** UMAP of the ABpxp lineage, with CellRank 2 computed terminal states indicated, as in Fig. 3.  
**b-d.** Aggregation of individual fate probabilities towards Ciliated neurons (**b**), Non-ciliated neurons (**c**), and Glia and excretory cells (**d**).



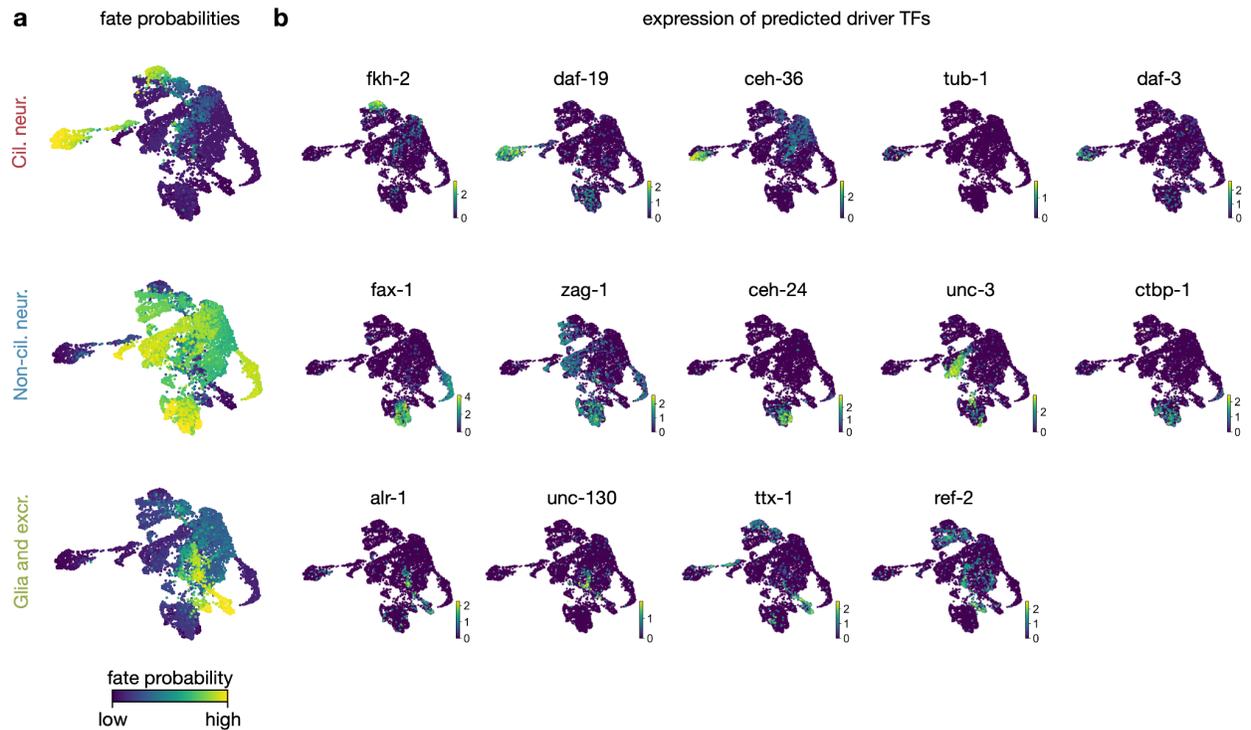
**Fig. S12 | Fate probabilities are higher in corresponding pre-terminal populations.**

**a.** UMAP of the ABxpx lineage, colored by coarse cluster annotations. **b.** Heatmap, displaying average fate probability among progenitors cells. **c.** Violin plots over the distribution of fate probabilities towards Ciliated neurons (top), Non-ciliated neurons (center), and Glia and excretory cells (bottom), grouped by pre-terminal cell states as shown in **(a)**. **d.** Heatmap, displaying the mean over the fate distributions in **(c)**. For each terminal state, moslin correctly predicts the corresponding pre-terminal group to have highest average fate probabilities (indicated with red boxes).



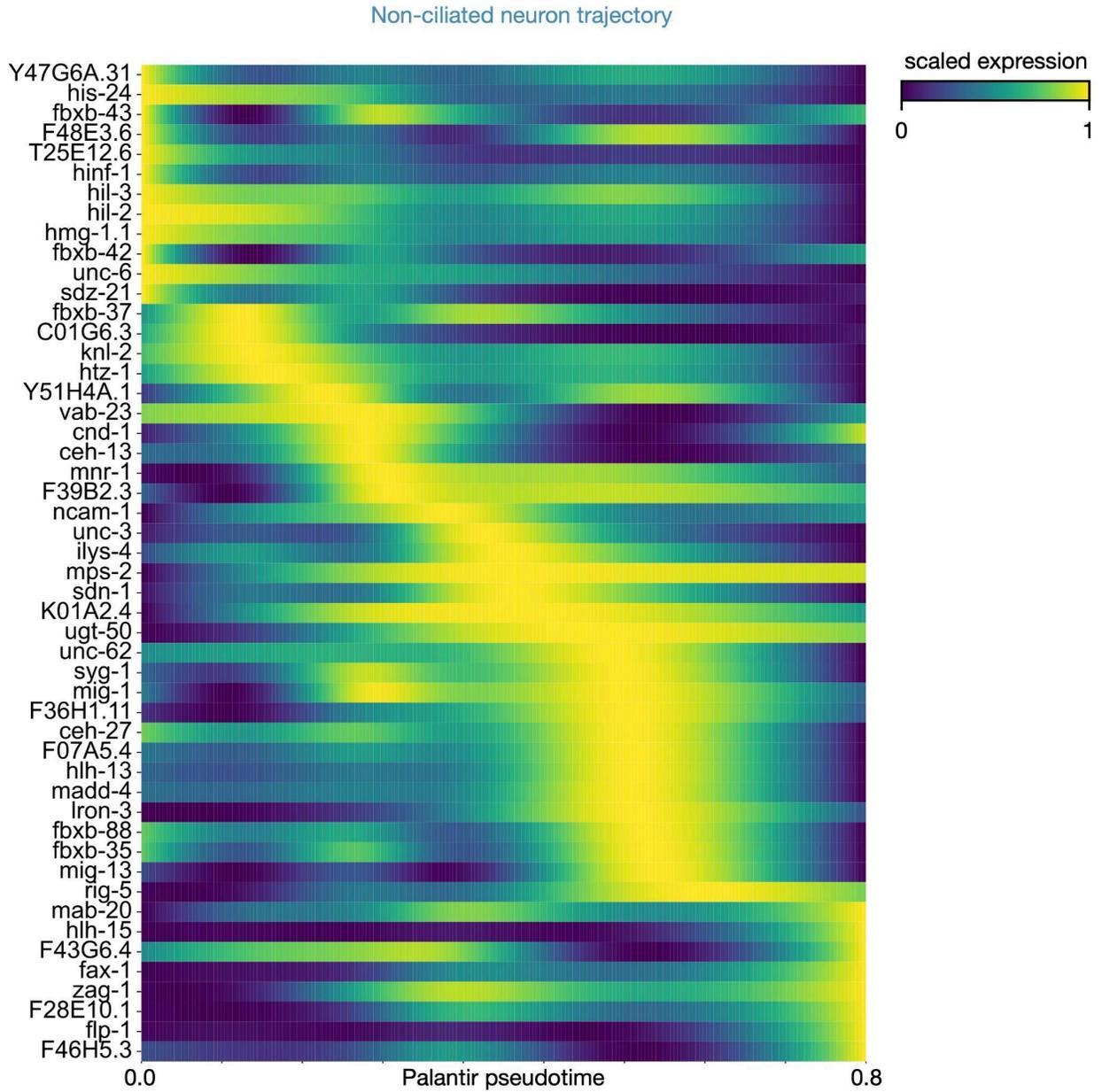
**Fig. S13 | Predicted mean fate probabilities from competing approaches.**

We show the same heatmap as for moslin in Additional file 1: Fig. S12d for competing approaches. Red boxes highlight the maximum mean fate probability per row.



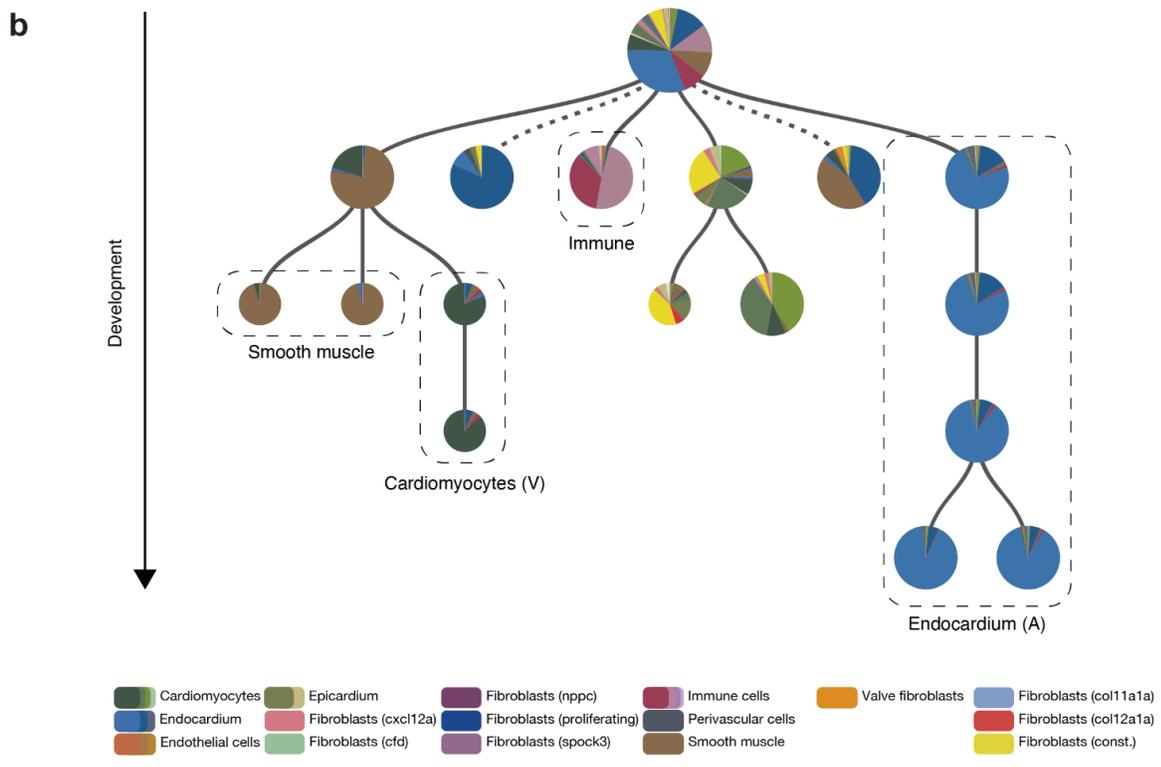
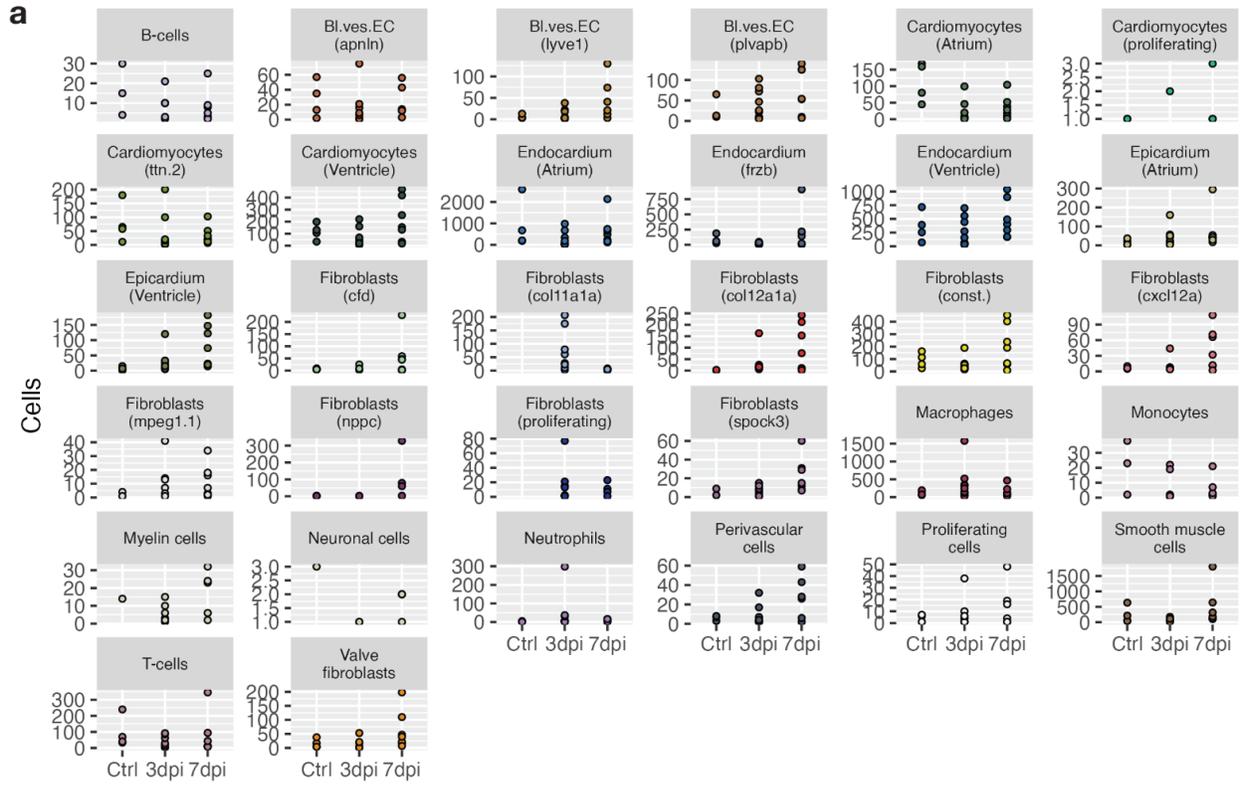
**Fig. S14 | Predicted driver TFs for Ciliated and Non-ciliated neurons, Glia and excretory cells.**

**a,b.** UMAPs of the ABxp lineage, colored by moslin/CellRank 2-computed fate probabilities (**a**) and the expression of predicted driver TFs (**b**). Each TF shown here is among the 20 most correlated TFs with the corresponding fate probabilities and has previously been reported to be important for the corresponding developmental trajectory (Methods and Additional file 3: Table S1).



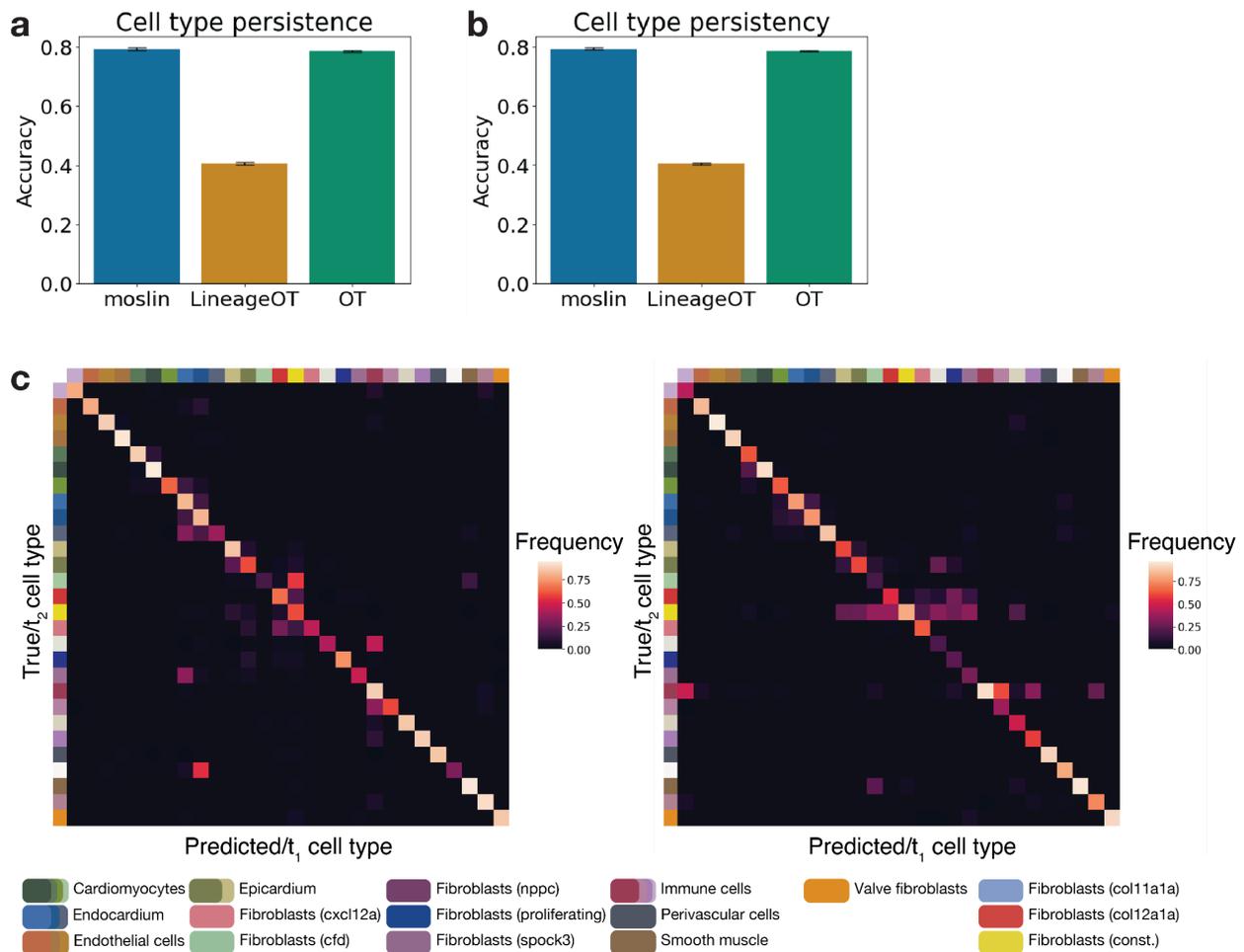
**Fig. S15 | Complete heatmap of smoothed gene expression along the non-ciliated neuron trajectory.**

As in Fig. 3f, with all gene names included.



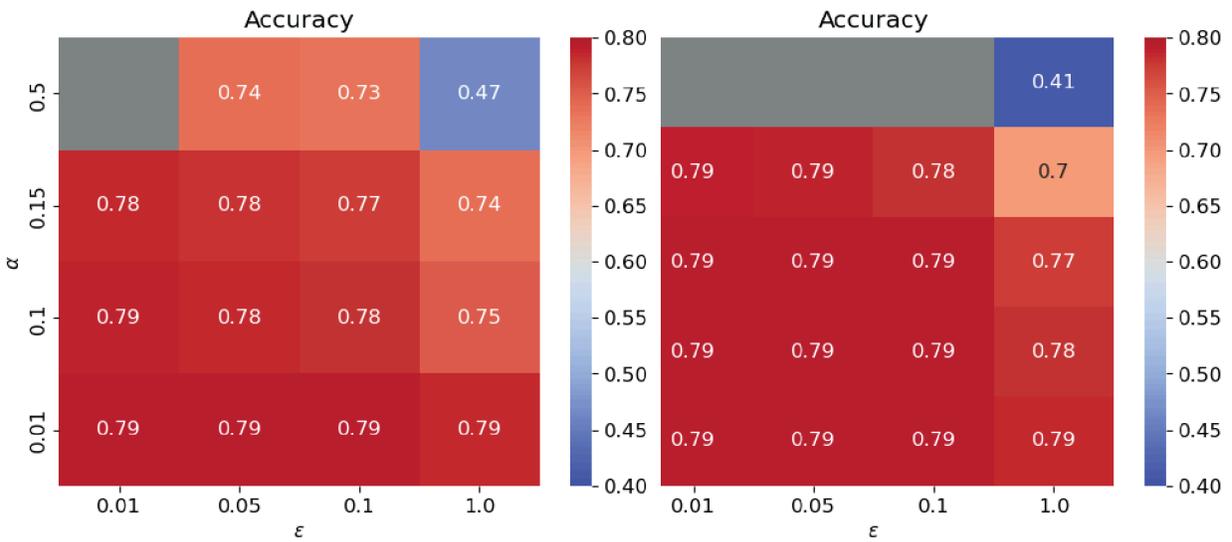
**Fig. S16 | Overview of the zebrafish dataset.**

**a.** Cell type frequencies per dataset. **b.** Annotated exemplary lineage tree. This lineage tree represents cell divisions during early stage development as measured by LINNAEUS. Every node represents a cell and the pie chart coloring indicates the eventual cell types that originate from this cell. Since not all divisions are measured, the tree is not necessarily binary. Dashed edges indicate that the temporal placement of the cell (relative to the other cells of the same generation) is unsure. The lineage tree branches show that cells from early divisions give rise to distinct adult cell types.



**Fig. S17 | Cell type persistence test**

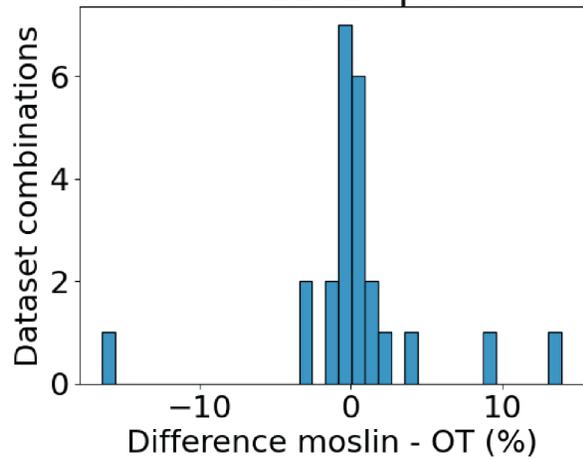
**a.** Persistence test accuracy with 95% confidence intervals (Methods). **b.** An approach of subsampling  $t_2$  cells (both at 3dpi and 7dpi) yields 95% confidence rates on accuracy that are very similar to those calculated through variability between dataset combinations ( $<0.012$  for subsampling,  $<0.008$  for variability between combinations, Methods). **c.** Cell type ancestors confusion matrix that underlies the persistence test, normalized by ground truth/ $t_2$  cell type frequency (left) and prediction/ $t_1$  cell type frequency (right).



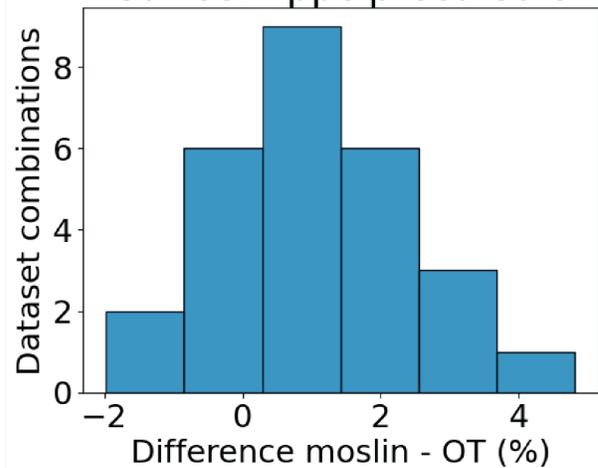
**Fig. S18 | Moslin performance on zebrafish heart regeneration data is robust to variations in hyperparameter values.**

Cell type persistence accuracy over varying values of hyperparameters  $\alpha$ ,  $\epsilon$  and  $\tau_a$ , controlling the weight given to lineage vs. gene expression information, the amount of entropic regularization, and the level of unbalancedness at the source marginal, respectively (Methods). Hyperparameters  $\tau_a = 0.4$  and  $\alpha = 0.01$  were kept fixed at their optimal values in the left and right plots, respectively. Gray tiles indicate non-convergence.

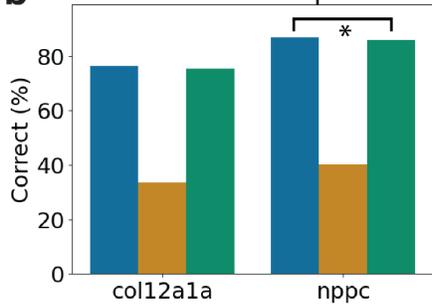
**a** Correct col12a1a precursors



Correct nppc precursors

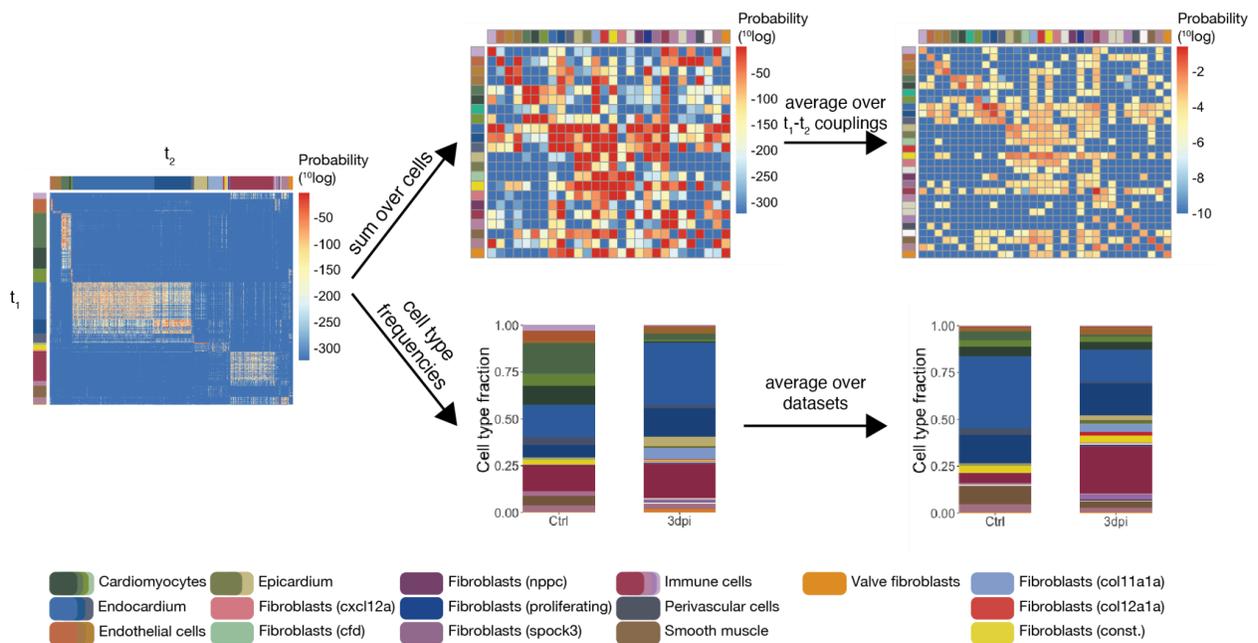


**b** Transient fibroblast precursors



**Fig. S19 | Moslin increases optimal transport performance by a small but significant amount on transient nppc fibroblasts.**

**a.** Histograms of differences between percentage of transient fibroblasts (col12a1a, left, nppc, right) from ground truth precursors between moslin and a pure Wasserstein term ( $W$ ;  $\alpha = 0$ , only gene expression information) over all dataset combinations. **b.** Percentage of transient fibroblasts from ground truth ancestor cell types. Colors indicate methods as in Additional file 1: Fig. S17. On nppc fibroblasts, moslin outperforms  $W$  by 1.2% ( $p=0.00027$ , Methods).



**Fig. S20 | Calculation of cell type transitions in zebrafish heart regeneration.**

The moslin-calculated cell-cell coupling matrix for every combination of  $t_1$  and  $t_2$  datasets is aggregated to obtain cell type couplings; we also calculate cell type fractions. Both of these are averaged - the cell type fractions over all datasets and the couplings over all  $t_1$ - $t_2$  dataset combinations, and the averages are used to calculate cellular flows. Coupling color-scales shown are log-10 of the coupling values.