

# Contents

<b>1</b>	<b>Cell type annotation</b>	<b>1</b>
1.1	Scatter plots . . . . .	1
1.2	Bar plots . . . . .	2
1.3	Correlation plots . . . . .	5
<b>2</b>	<b>Data integration</b>	<b>5</b>
2.1	Linear optimal transport . . . . .	5
2.1.1	Scatter plots . . . . .	5
2.1.2	Correlation plots . . . . .	10
2.2	Supervised optimal transport . . . . .	11
<b>3</b>	<b>HSC score</b>	<b>12</b>
<b>4</b>	<b>Trajectory inference</b>	<b>12</b>
4.1	X plots Erythroid . . . . .	13
<b>5</b>	<b>Early populations</b>	<b>13</b>

## 1 Cell type annotation

### 1.1 Scatter plots

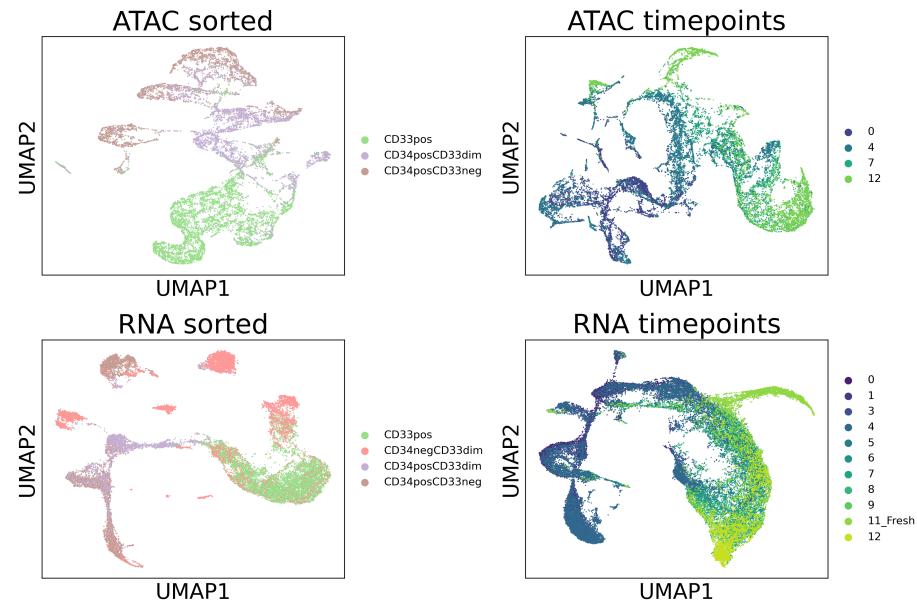


Figure 1: Scatter: Time points (days) and sorted assays (CD) annotation

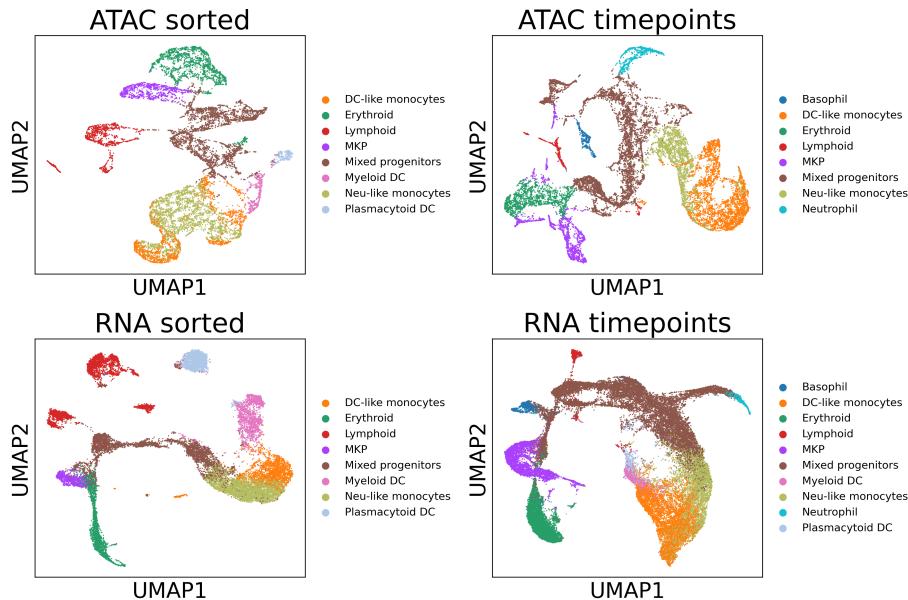


Figure 2: Scatter: Leiden identity simple annotation

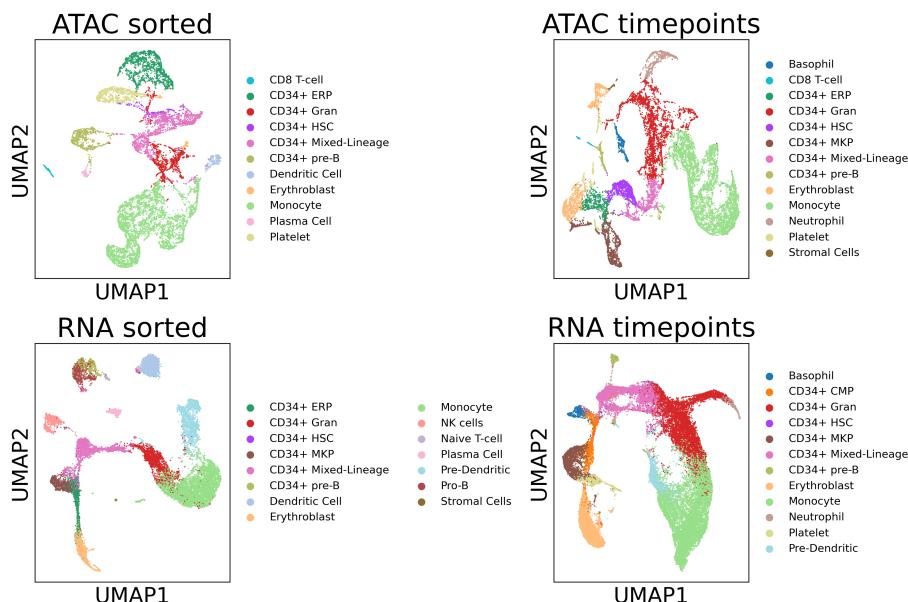


Figure 3: Scatter: Leiden identity annotation

## 1.2 Bar plots

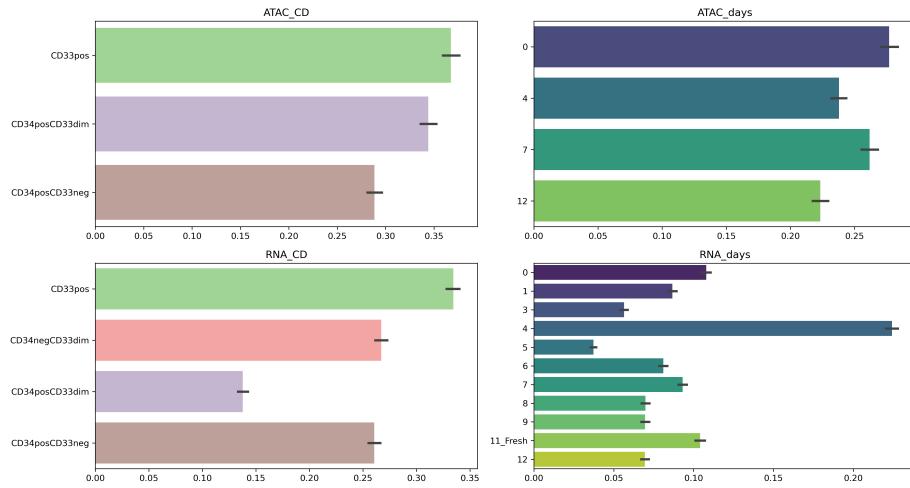


Figure 4: Proportions: Leiden identity simple annotation

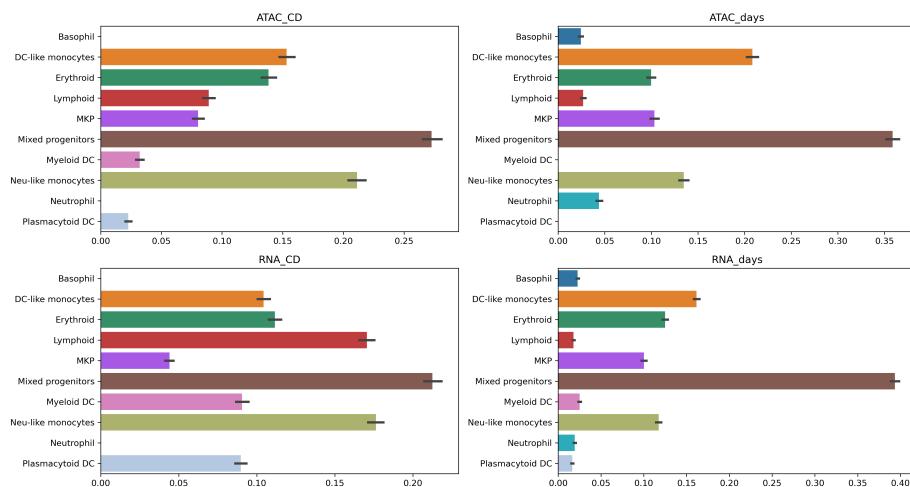


Figure 5: Proportions: Leiden identity simple annotation

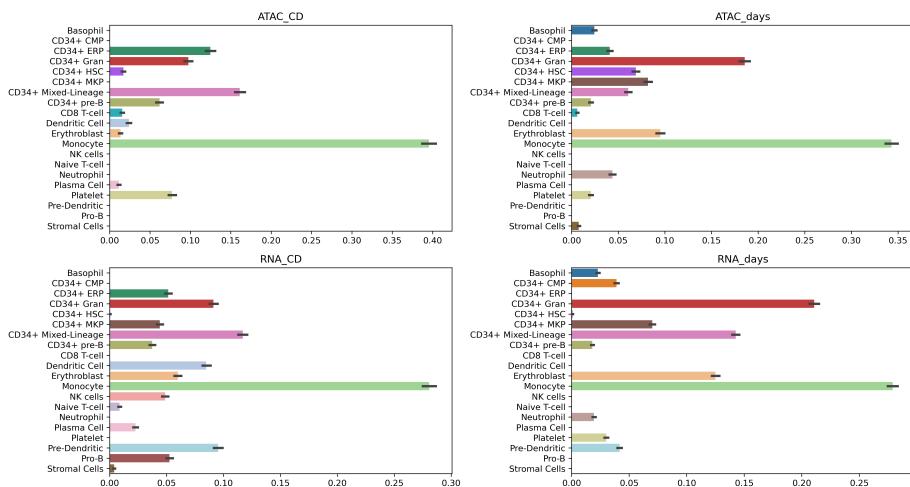


Figure 6: Proportions: Leiden identity annotation

### 1.3 Correlation plots

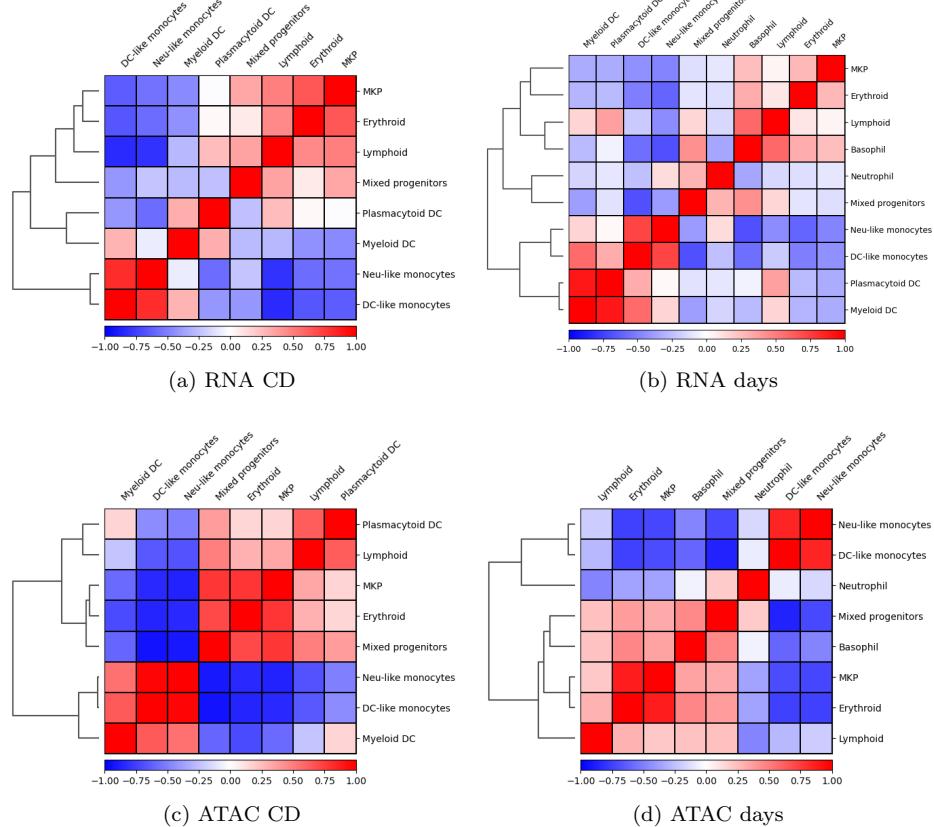


Figure 7: Correlation between cell type averages (obtained using top PCs)

## 2 Data integration

### 2.1 Linear optimal transport

#### 2.1.1 Scatter plots

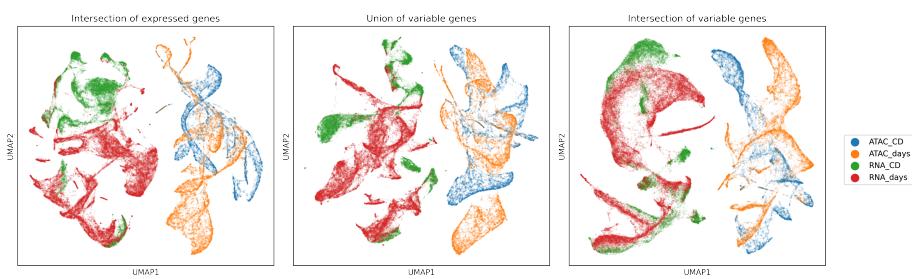


Figure 8: Assays in RNA days PCA subspace

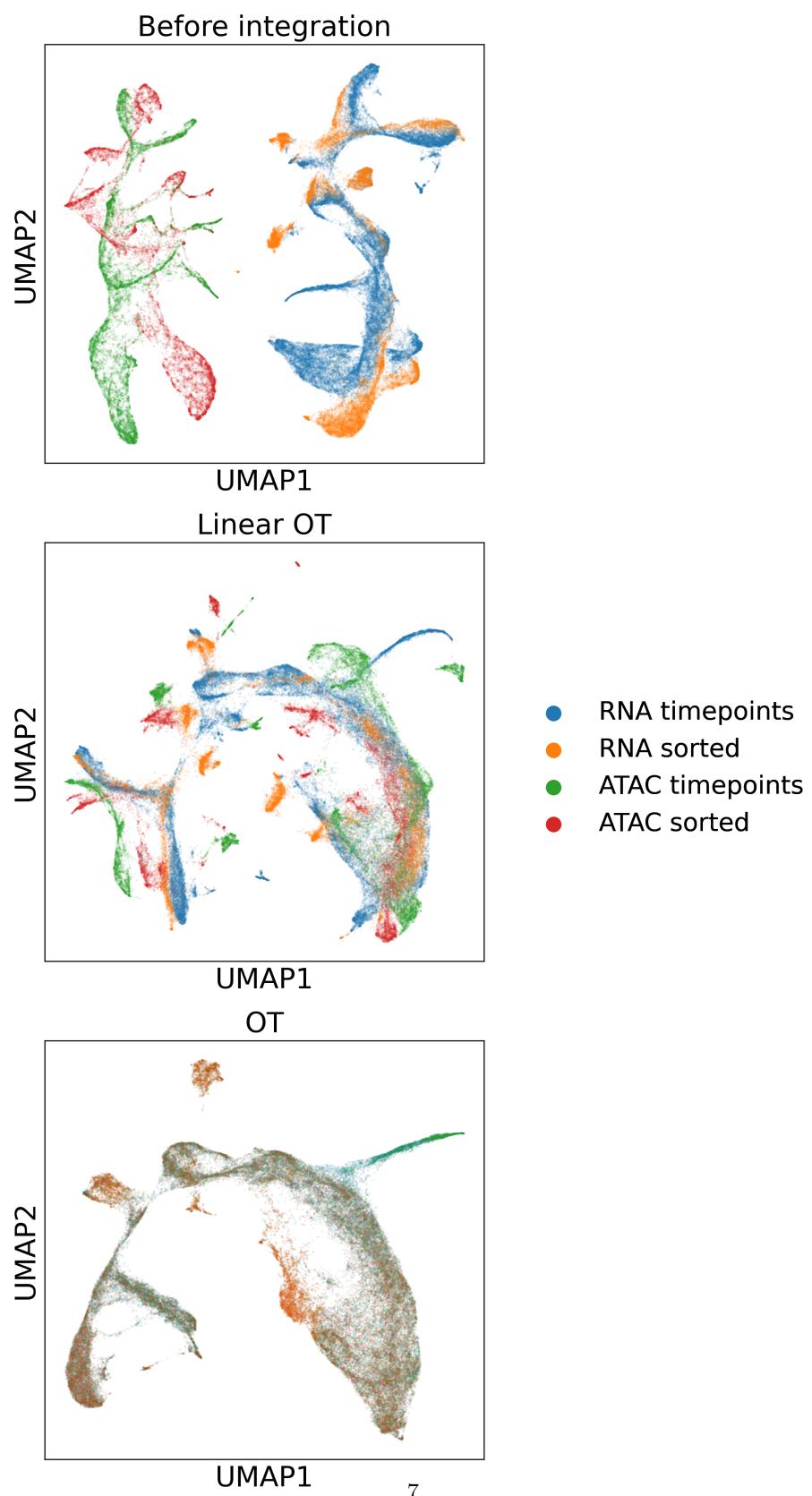


Figure 9: Assays in RNA days PCA subspace, after alignment with linear optimal transport

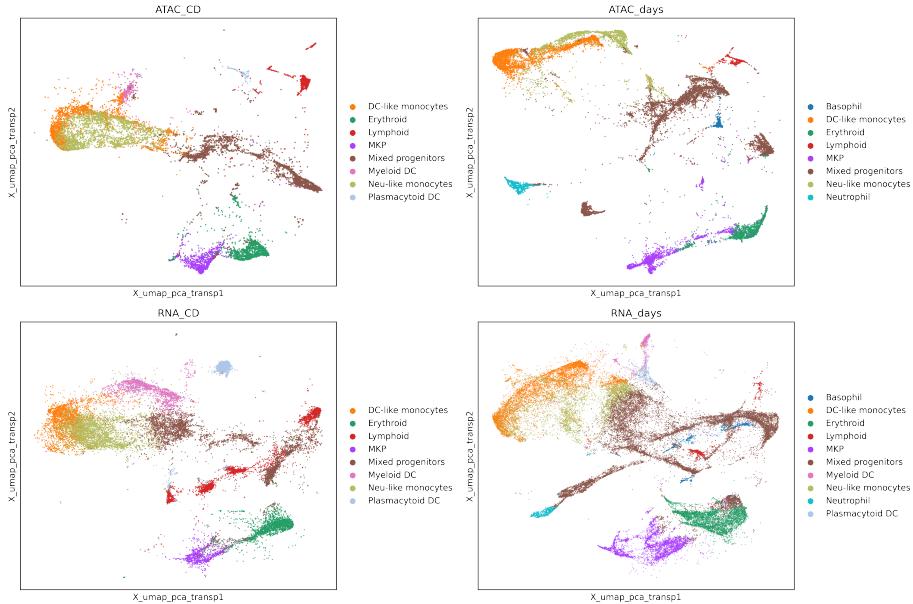


Figure 10: Assays in RNA days PCA subspace obtained using the intersection of expressed genes, after alignment with linear optimal transport.

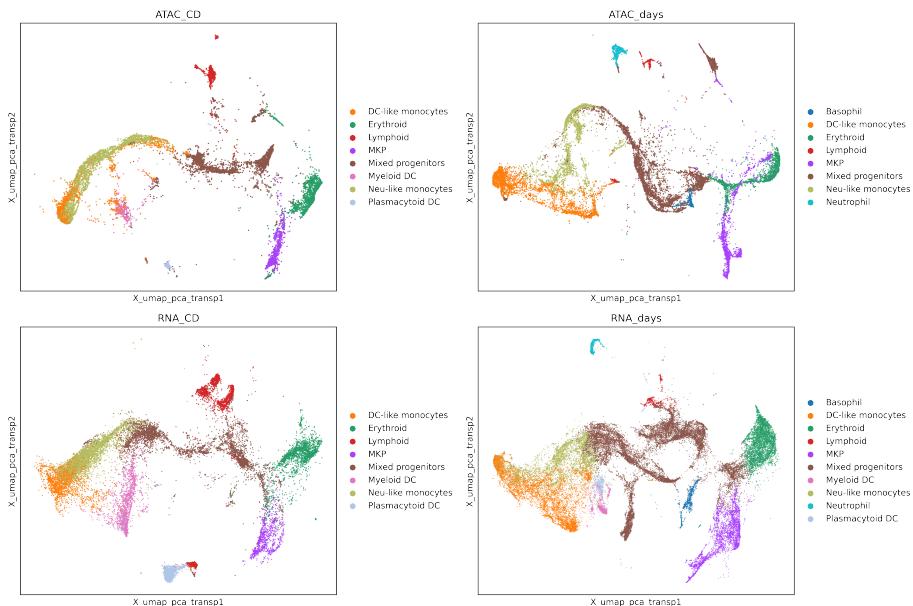


Figure 11: Assays in RNA days PCA subspace obtained using the union of variable genes, after alignment with linear optimal transport.

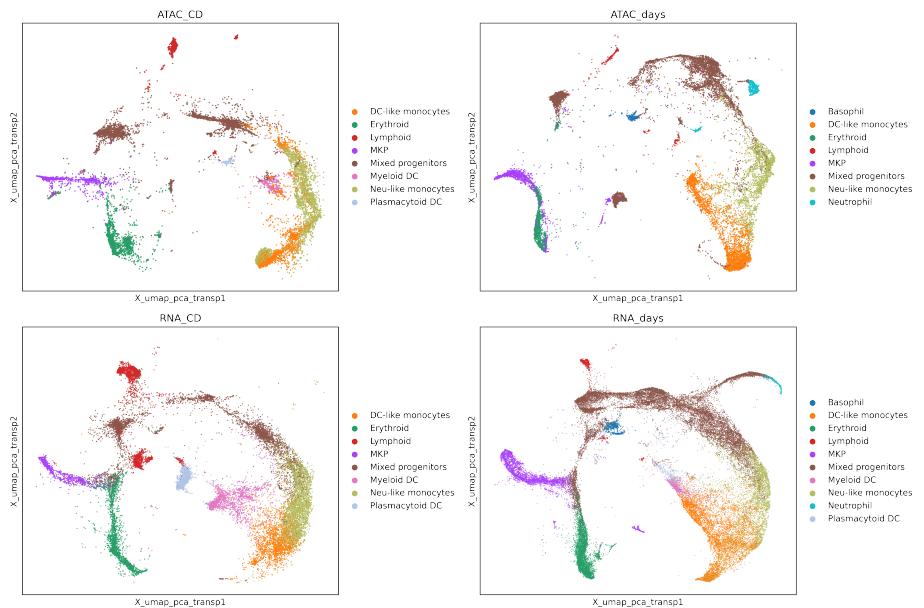


Figure 12: Assays in RNA days PCA subspace obtained using the intersection of variable genes, after alignment with linear optimal transport.

### 2.1.2 Correlation plots

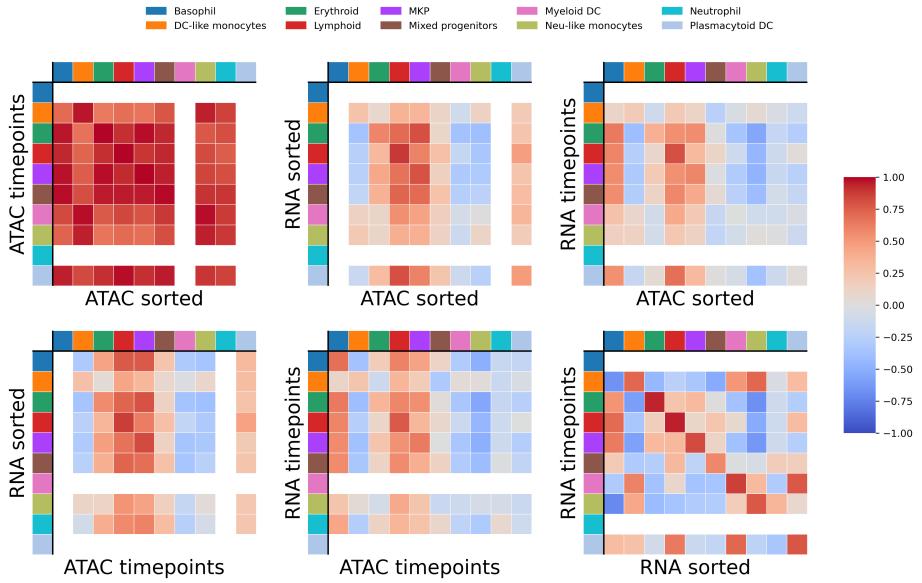


Figure 13: Correlation between assays' cell type averages (obtained using top PCs of assays, each in their own PCA subspace)

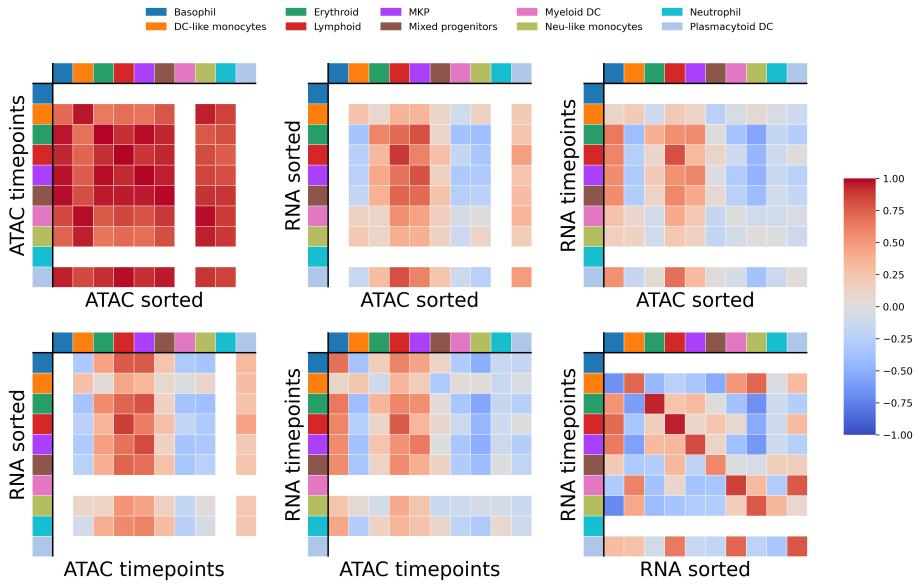


Figure 14: Correlation between assays' cell type averages (obtained using top PCs of assays in RNA days PCA subspace)

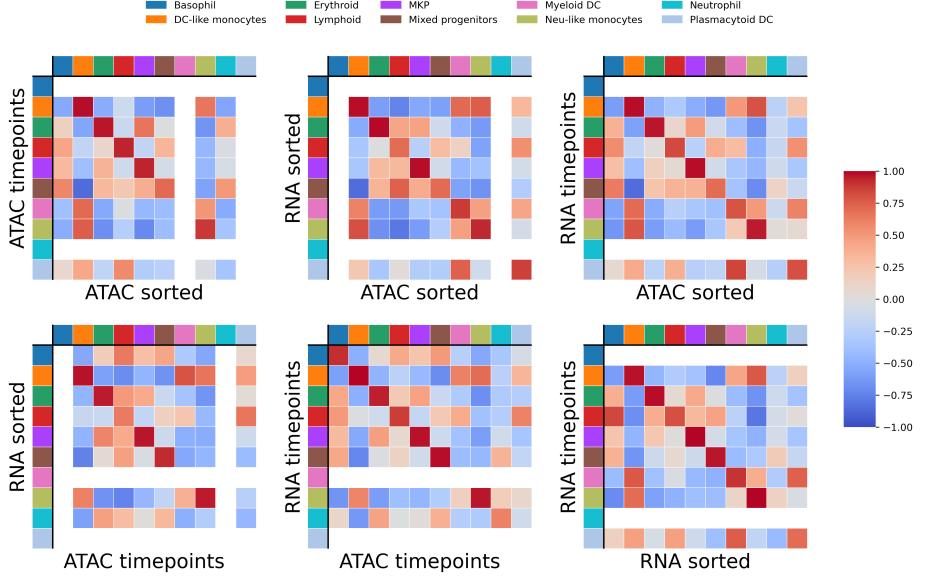


Figure 15: Correlation between assays’ cell type averages (obtained using top PCs of assays in RNA days PCA subspace, after alignment with linear optimal transport)

## 2.2 Supervised optimal transport

### Grouping of categories to go from full to simple annotation

- ‘Early-ERP’, ‘Erythroblast’, ‘CD34+ ERP’ = ‘Erythroid’
- ‘Platelet’, ‘CD34+ MKP’ = ‘MKP’
- ‘Pre-Dendritic’, ‘Dendritic Cell’ = ‘Dendritic’
- ‘CD34+ CLP’, ‘CD34+ pre-B’, ‘Pro-B’, ‘Plasma Cell’, ‘NK cells’, ‘Naive T-cell’, ‘CD8 T-cell’ = ‘Lymphoid’
- ‘CD34+ Mixed-Lineage’, ‘CD34+ HSC’, ‘CD34+ CMP’, ‘CD34+ Gran’, ‘Eosinophil’, ‘Stromal Cells’ = ‘Mixed-Lineage’
- change unlikely ‘Erythroid’ annotation (outlier, mixed cluster) for late ATAC days (day 7 and 12) to ‘Mixed-Lineage’

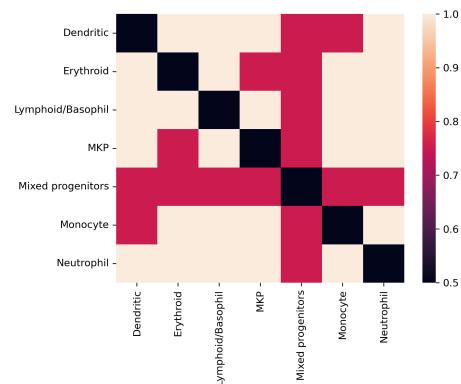


Figure 16: Supplementary: supervision of OT cost matrix

### 3 HSC score

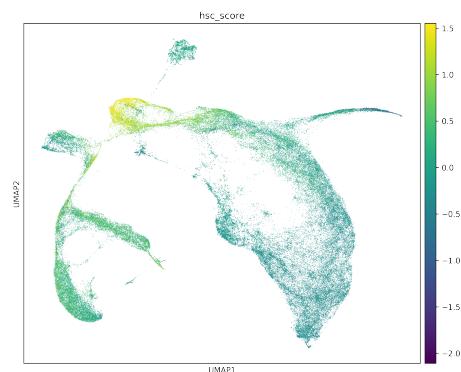


Figure 17: Supplementary: hematopoietic stem cell score

### 4 Trajectory inference

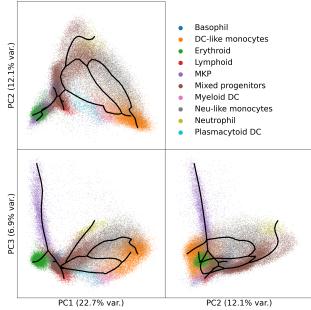


Figure 18: Supplementary: Trajectory visualized in PCA space

#### 4.1 X plots Erythroid

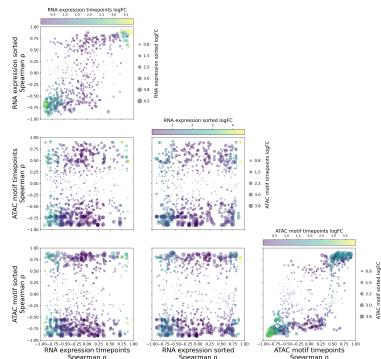


Figure 19: Supplementary: X-plot for the Erythroid branch for all pairs of assays, with MAGIC smoothing for RNA and chromVAR motifs

## 5 Early populations

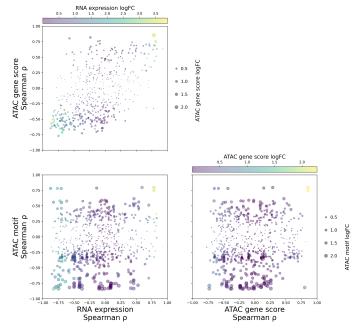


Figure 20: Supplementary: X-plot for the Erythroid branch, with MAGIC smoothing, grouping RNA assays together and chromVAR motifs together

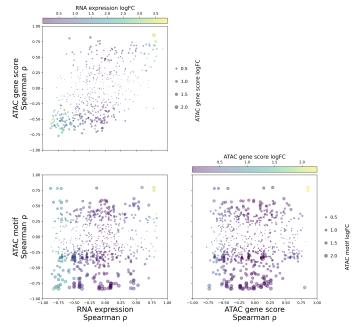


Figure 21: Supplementary: X-plot for the Erythroid branch, with MAGIC smoothing, grouping RNA assays together and chromVAR motifs together

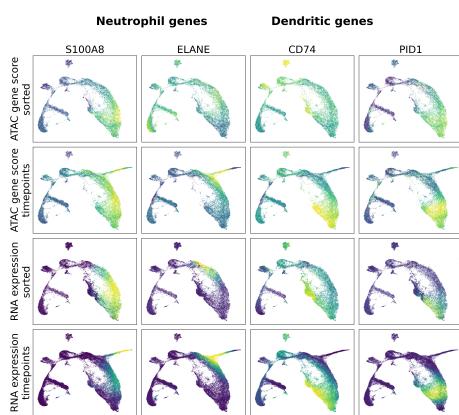


Figure 22: Supplementary: Canonical marker genes distinguish early monocyte populations: dendritic-like and neutrophil-like