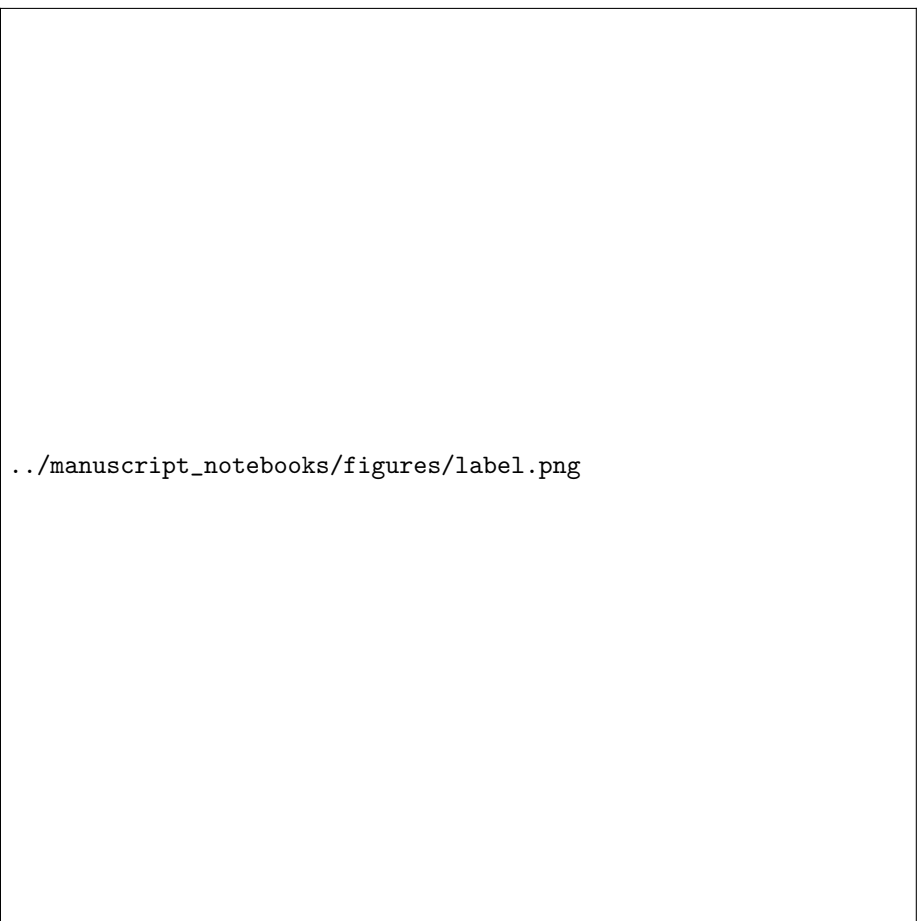


Contents

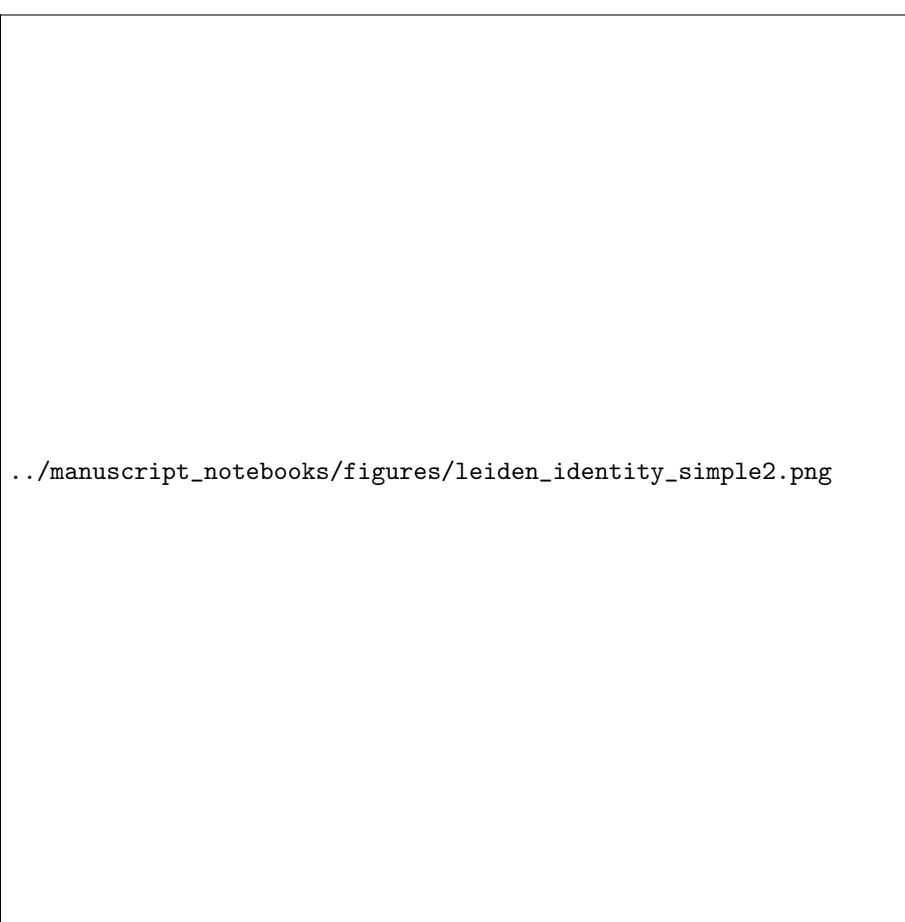
1 Cell type annotation

1.1 Scatter plots




`../manuscript_notebooks/figures/label.png`

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



`../manuscript_notebooks/figures/leiden_identity_simple2.png`

Figure 2: Scatter: Leiden identity simple annotation



`../manuscript_notebooks/figures/leiden_identity.png`

Figure 3: Scatter: Leiden identity annotation

1.2 Bar plots



Figure 4: Proportions: Leiden identity simple annotation



Figure 5: Proportions: Leiden identity simple annotation

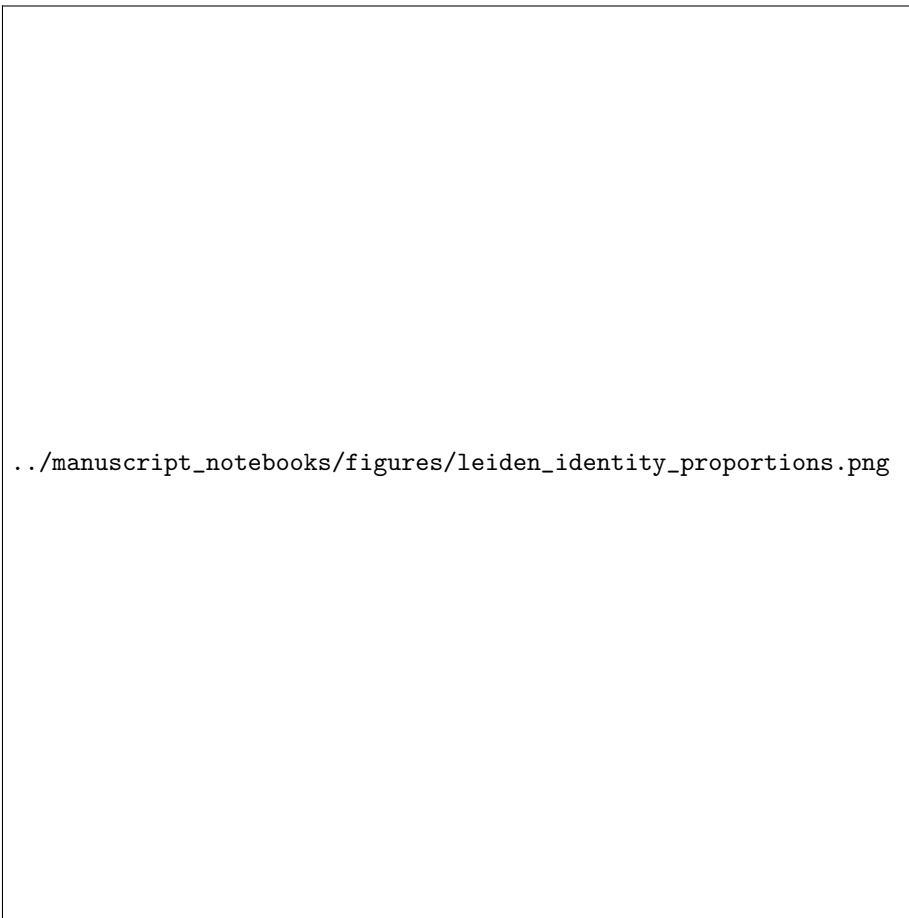


Figure 6: Proportions: Leiden identity annotation

1.3 Correlation plots

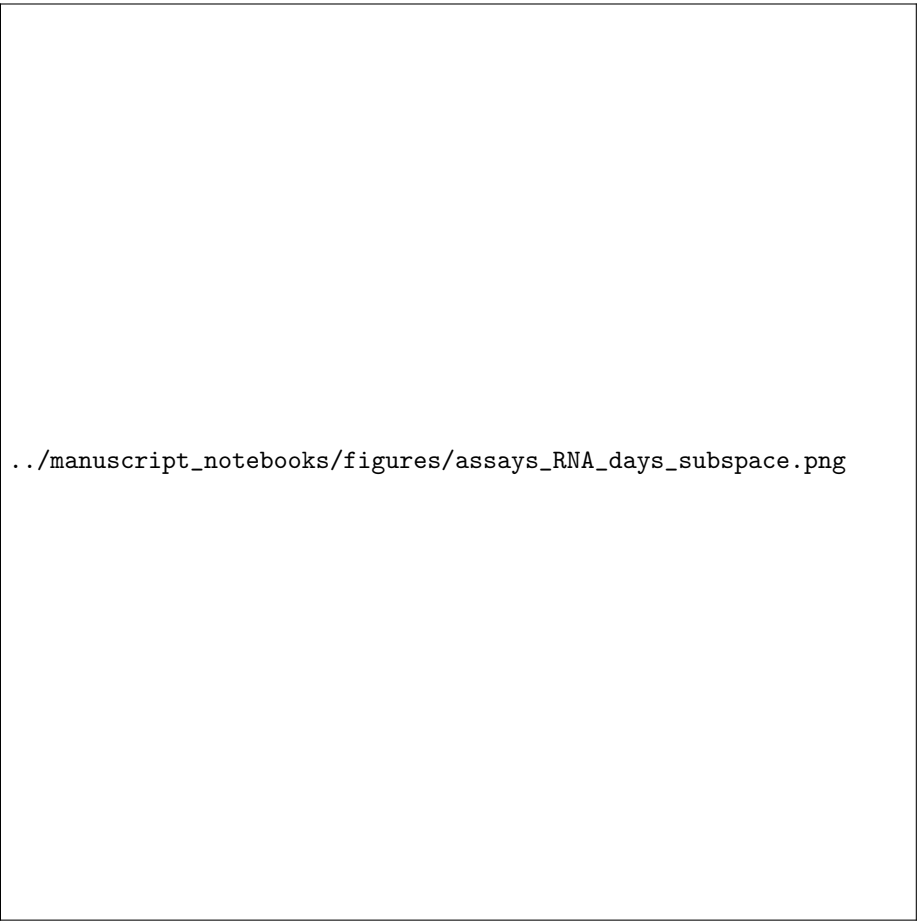


Figure 7: Correlation between cell type averages (obtained using top PCs)

2 Data integration

2.1 Linear optimal transport

2.1.1 Scatter plots



`../manuscript_notebooks/figures/assays_RNA_days_subspace.png`

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.



../manuscript_notebooks/figures/assays_RNA_days_subspace_aligned_Intersection_of_variable

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'

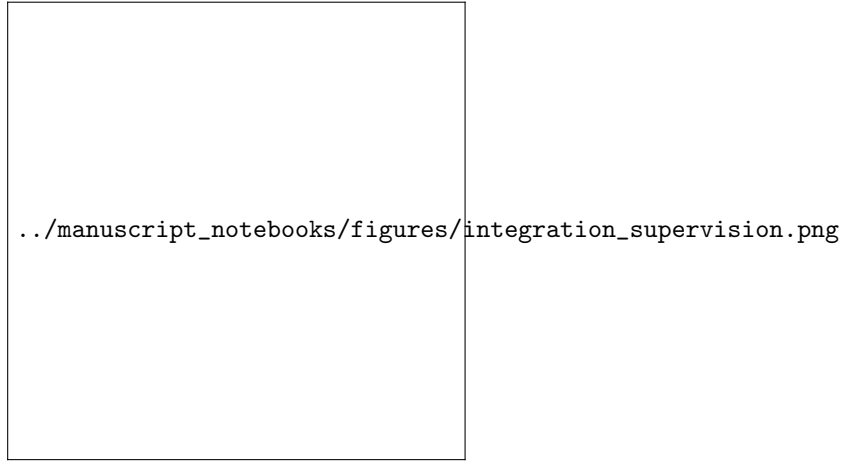


Figure 16: Supplementary: supervision of OT cost matrix

3 Trajectory inference

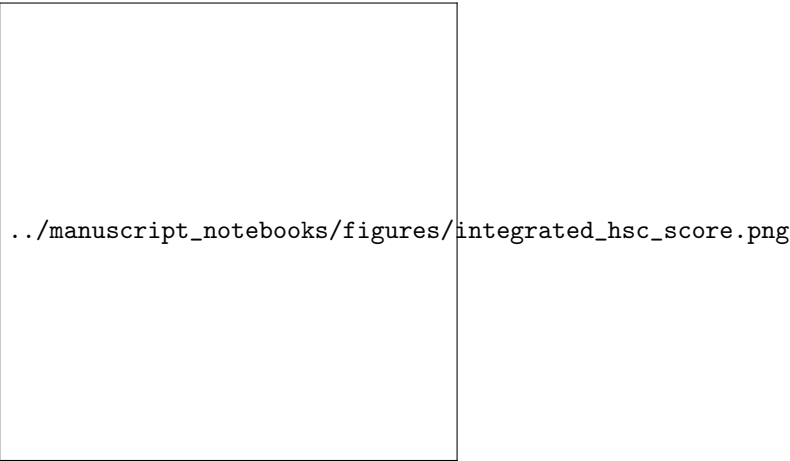


Figure 17: Supplementary: hematopoeitic stem cell score



Figure 18: Top markers of the Erythroid branch based on log fold change