

Analysis of Blood Differentiation Dataset (15K+ Cells)

Load and Explore with Scanpy

```
In [4]: import scanpy as sc
from matplotlib import rcParams
import pandas as pd

sc.settings.set_figure_params(dpi=200, frameon=False, figsize=(10, 10), facecolor='white') # low dpi (dots per inch) yields small inline figures

datadir = "../Data/"

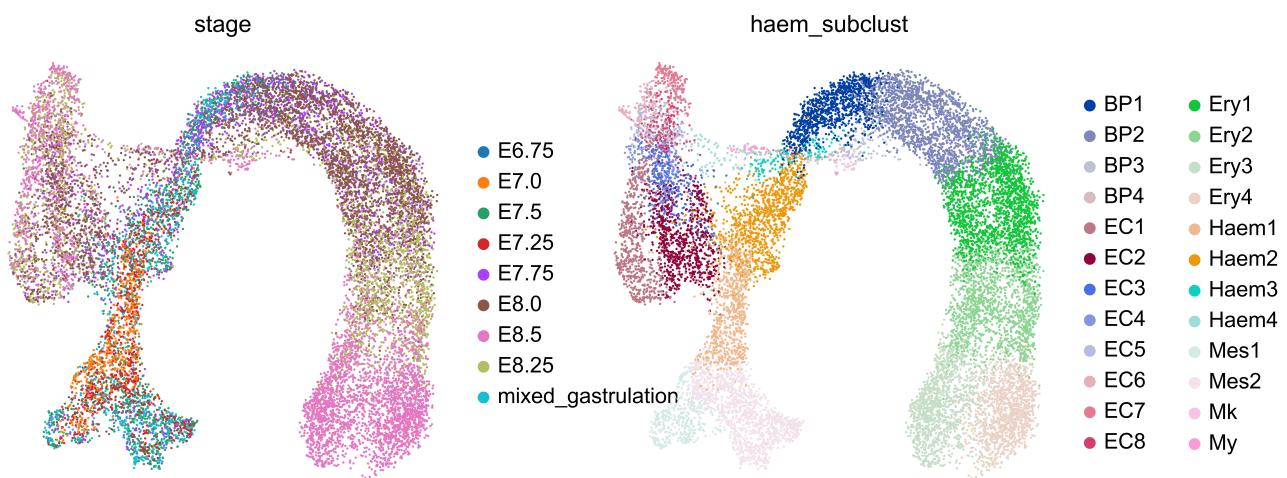
adata = sc.read_csv(datadir+'haem_coord.csv')
meta = pd.read_csv(datadir+'haem_meta.csv')
adata.obs = meta
adata
```

```
Out[4]: AnnData object with n_obs × n_vars = 15875 × 50
         obs: 'Unnamed: 0', 'cell', 'barcode', 'sample', 'stage', 'sequencing.batch', 'theiler', 'doub.density', 'doublet', 'cluster', 'cluster.sub', 'cluster.stage', 'cluster.theiler', 'stripped', 'celltype', 'colour', 'umapX', 'umapY', 'haem_gephiX', 'haem_gephiY', 'haem_subclust', 'endo_gephiX', 'endo_gephiY', 'endo_trajectoryName', 'endo_trajectoryDPT', 'endo_gutX', 'endo_gutY', 'endo_gutDPT', 'endo_gutCluster'
```

```
In [5]: sc.pp.neighbors(adata, n_neighbors=10)
sc.tl.umap(adata)
```

```
In [6]: import matplotlib.pyplot as plt
sc.settings.set_figure_params(dpi=500, frameon=False, figsize=(5, 5), facecolor='white') # low dpi (dots per inch) yields small inline figures
fig = sc.pl.umap(adata,color = ['stage','haem_subclust'],return_fig = True,save = 'haem_umap.pdf')
```

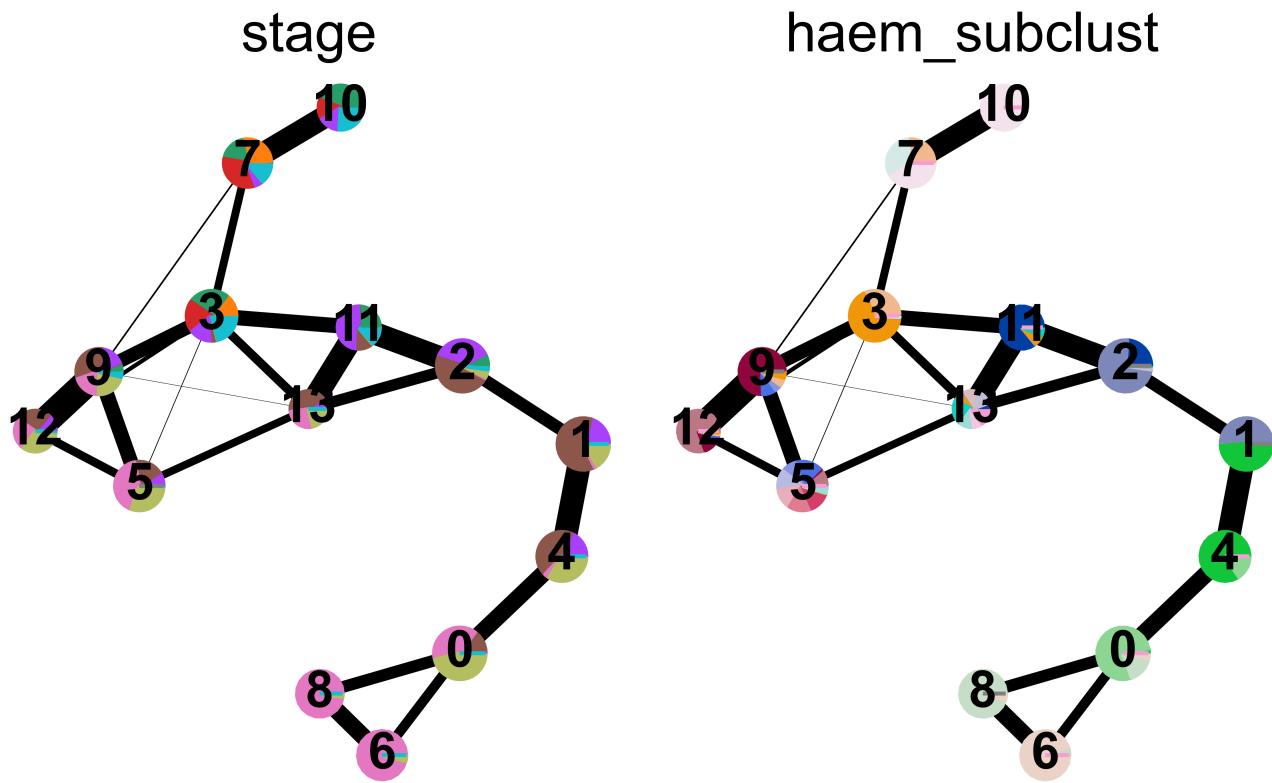
```
... storing 'barcode' as categorical
... storing 'stage' as categorical
... storing 'theiler' as categorical
... storing 'celltype' as categorical
... storing 'colour' as categorical
... storing 'haem_subclust' as categorical
```



```
In [7]: sc.pp.neighbors(adata, n_neighbors=10)
sc.tl.leiden(adata, resolution=1.0)
sc.tl.paga(adata, groups='leiden')
```

```
In [57]: sc.settings.set_figure_params(dpi=500, frameon=False, figsize=(3, 3), facecolor='white') # low dpi (dots per inch) yields small inline figures  
sc.pl.paga(adata, color=['stage', 'haem_subclust'], save = 'haem_paga.pdf')
```

WARNING: saving figure to file figures/pagahaem_paga.pdf

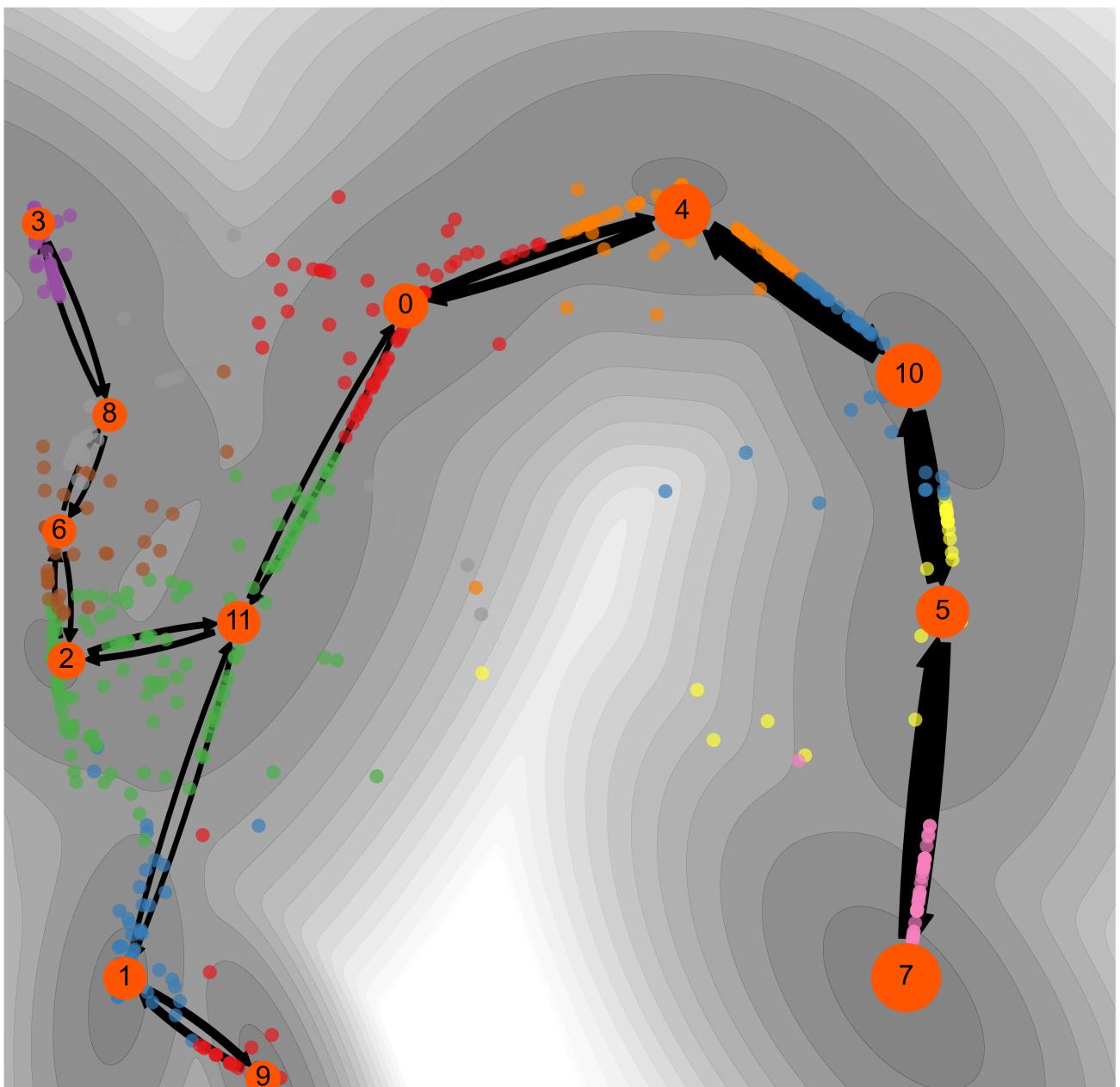


MuTrans Analysis

```
In [15]: import pyMuTrans as pm  
par = {"choice_distance": "cosine", "K_cluster": 12.0, "trials": 100, "weight_scale": True,  
"initial": "pca", "reduction_coord": "umap", "reduce_large_scale": True, "reduce_num_meta_c  
ell": 1500.0}  
adata = pm.dynamical_analysis(adata, par)
```

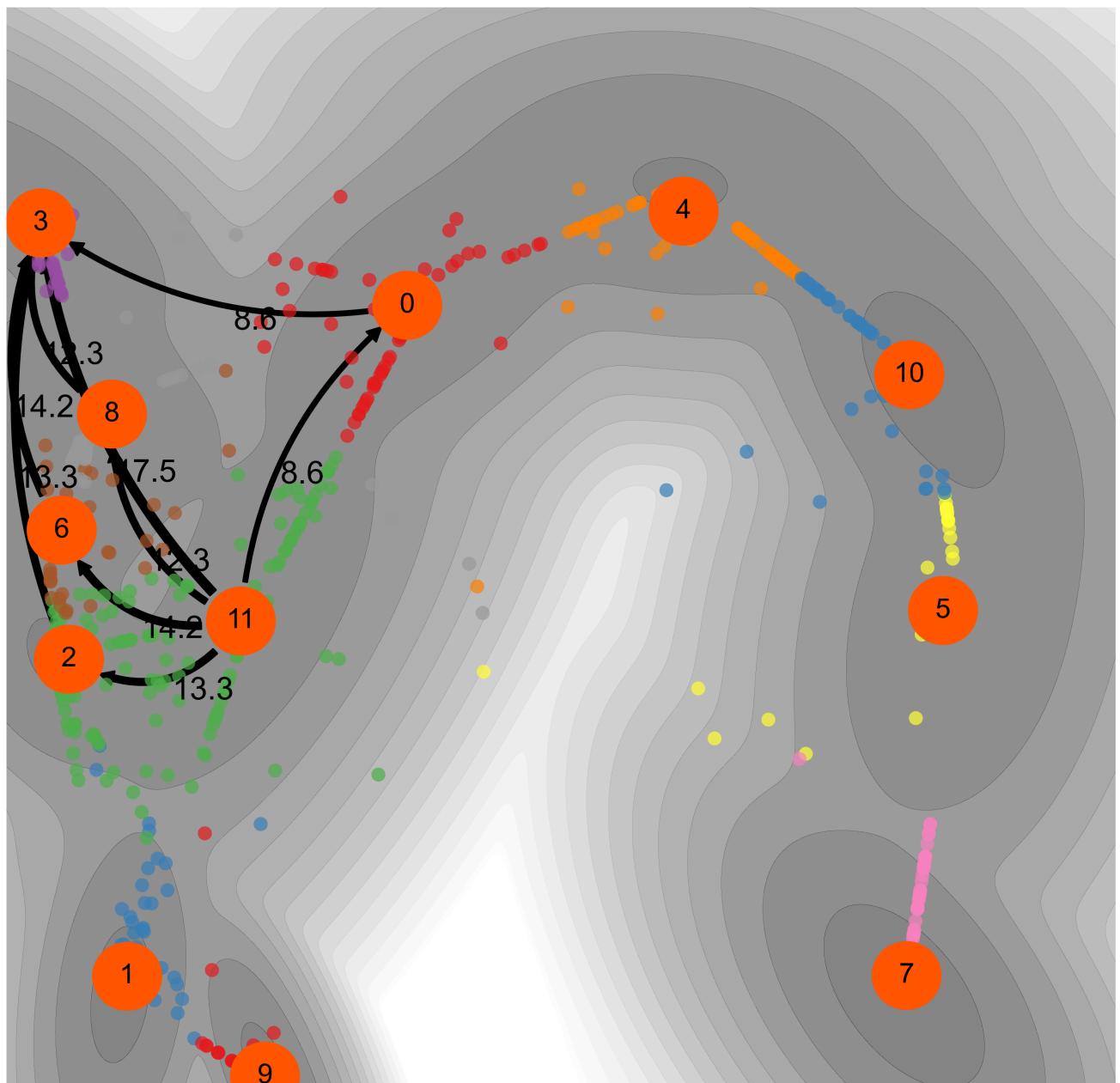
/Users/cliffzhou/opt/anaconda3/lib/python3.7/site-packages/pyemma/__init__.py:92: UserWarning: You are not using the latest release of PyEMMA. Latest is 2.5.7, you have 2.5.6.
.format(latest=latest, current=current), category=UserWarning)

```
In [21]: import matplotlib.pyplot as plt  
fig = plt.figure(figsize = (10,10))  
pm.infer_lineage(adata,si=0,sf = 2,method = "MPFT",size_state = 0.2,size_text=16,alph  
a_point = 0.7)
```



```
In [22]: fig.savefig('haem_mpft.pdf')
```

```
In [34]: fig = plt.figure(figsize = (10,10))
pm.infer_lineage(adata,si=11,sf = 3,method = "MPPT",flux_fraction = 0.6,size_state = 0.2,size_text=16,alpha_point = 0.7)
fig.savefig('haem_mppt_ec.pdf')
```



```
In [35]: fig = plt.figure(figsize = (10,10))
pm.infer_lineage(adata,si=11,sf = 7,method = "MPPT",flux_fraction = 0.35,size_state =
0.2,size_text=16,alpha_point = 0.7)
fig.savefig('haem_mppt_eryth.pdf')
```

