

cell_topic_analysis

June 15, 2022

```
[ ]: import experiment_cell_topic
from analysis import _topics, _umap_viz, _gsea
import pandas as pd

spruce = experiment_cell_topic.get_experiment()
print(spruce.model_id)
```

```
[ ]: dfh = spruce.cell_topic.h
df_umap= pd.DataFrame()
df_umap['cell'] = dfh['cell']
```

```
[ ]: import umap

umap_2d = umap.UMAP(n_components=2, init='random', random_state=0,min_dist=0.
    ↪0,metric='cosine')
proj_2d = umap_2d.fit(dfh.iloc[:,1:])
df_umap['umap1'] = umap_2d.embedding[:,0]
df_umap['umap2'] = umap_2d.embedding[:,1]
df_umap['topic'] = dfh.iloc[:,1:].idxmax(axis=1)
```

```
[ ]: df_umap.head()
```

```
[ ]:
```

	cell	umap1	umap2	topic
0	CGCGTTTTCTCAACC-1_GSM5022602_D4_GSE164898	5.794563	16.481131	h2
1	CID4495_AGCTCTCAGGAGCGTT_GSE176078	7.931469	3.237960	h11
2	CID4471_GTTTCTAAGCGTAATA_GSE176078	-2.594879	13.459086	h15
3	CID4398_CCGGGATGTCATGCAT_GSE176078	3.537929	-3.046828	h23
4	TTTGGAGAGACGCCAA-1_GSM5022606_D11_GSE164898	-6.983099	-2.546422	h22

```
[ ]: import matplotlib.pyplot as plt
plt.rcParams['figure.figsize'] = [12.50, 10.50]
plt.rcParams['figure.autolayout'] = True
import colorcet as cc
import seaborn as sns

cp = sns.color_palette(cc.glasbey, n_colors=len(df_umap['topic'].unique()))
sns.scatterplot(data=df_umap, x='umap1', y='umap2', hue='topic',s=1,palette=cp)
```

```
[ ]: <AxesSubplot:xlabel='umap1', ylabel='umap2'>
```

