

In [1]:

```
1 from matplotlib import pyplot as plt
2 %matplotlib inline
3 import seaborn as sns; sns.set()
4 import pandas as pd
```

In [2]:

```
1 url = "https://raw.githubusercontent.com/\
2 Stemanz/python_tutorials/master/datasets/gene_expression.csv"
3 data = pd.read_csv(url, index_col="gene symbol")
```

In [3]:

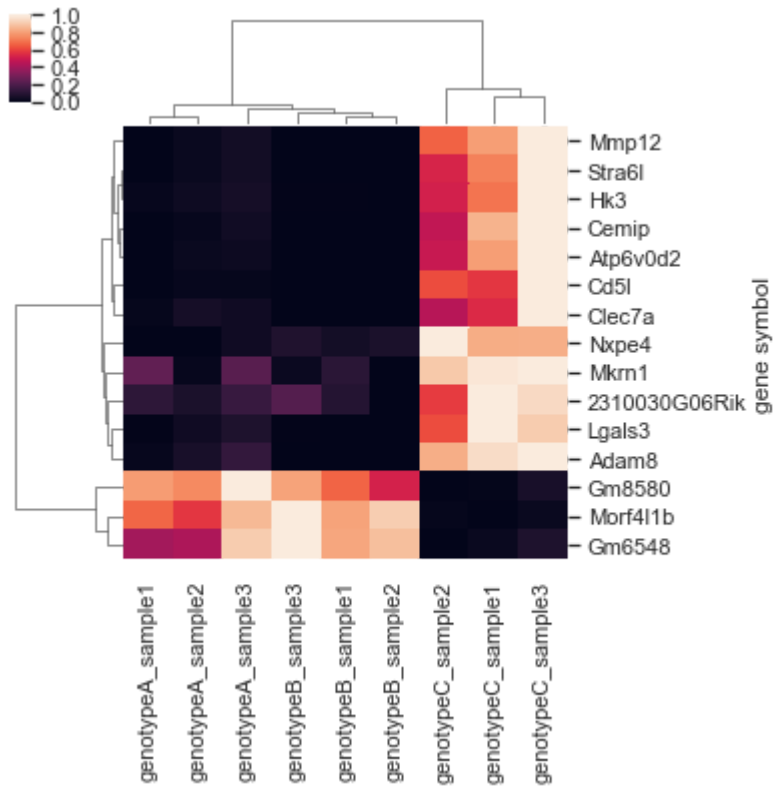
```
1 data.head()
```

Out[3]:

| | genotypeA_sample1 | genotypeA_sample2 | genotypeA_sample3 | genotypeB_sample1 | gene symbol |
|----------|-------------------|-------------------|-------------------|-------------------|-------------|
| Gm8580 | 5.767900 | 5.455137 | 7.157003 | 4.900520 | |
| Morf4l1b | 109.484714 | 97.039848 | 134.512688 | 127.100854 | |
| Gm6548 | 39.844572 | 40.872696 | 64.365084 | 59.735055 | |
| Cemip | 0.083976 | 1.620770 | 4.182101 | 0.114739 | |
| Stra6l | 0.041988 | 1.246746 | 2.193053 | 0.057370 | |

In [4]:

```
1 res = sns.clustermap(data, standard_scale=0, figsize=(5,5))
```



In [5]:

```
1 # from the doc:
2 # To access the reordered row indices, use: clustergrid.dendrogram_row.reordered_ind
3 # Column indices, use: clustergrid.dendrogram_col.reordered_ind
4 print(res.dendrogram_row.reordered_ind)
```

```
[7, 4, 10, 3, 6, 5, 8, 13, 12, 14, 9, 11, 0, 1, 2]
```

In [6]:

```
1 def extract_clustered_table(res, data):
2     """
3     input
4     =====
5     res:      <sns.matrix.ClusterGrid>  the clustermap object
6     data:      <pd.DataFrame>           input table
7
8     output
9     =====
10    returns: <pd.DataFrame>              reordered input table
11    """
12
13    # if sns.clustermap is run with row_cluster=False:
14    if res.dendrogram_row.reordered_ind is None:
15        print("Apparently, rows were not clustered.")
16        return -1
17
18    if res.dendrogram_col.reordered_ind is not None:
19        # reordering index and columns
20        new_cols = data.columns[res.dendrogram_col.reordered_ind]
21        new_ind = data.index[res.dendrogram_row.reordered_ind]
22
23        return data.loc[new_ind, new_cols]
24
25    else:
26        # reordering the index
27        new_ind = data.index[res.dendrogram_row.reordered_ind]
28
29        return data.loc[new_ind,:]
```

In [7]:

```
1 ordered = extract_clustered_table(res, data)
```

In [8]:

1

ordered

Out[8]:

| | genotypeA_sample1 | genotypeA_sample2 | genotypeA_sample3 | genotypeB_sample1 |
|---------------|-------------------|-------------------|-------------------|-------------------|
| gene symbol | | | | |
| Mmp12 | 4.030854 | 27.366074 | 51.868253 | 2.077976 |
| Stra6l | 0.041988 | 1.246746 | 2.193053 | 0.000000 |
| Hk3 | 7.179959 | 11.033702 | 15.198367 | 5.194930 |
| Cemip | 0.083976 | 1.620770 | 4.182101 | 0.074210 |
| Atp6v0d2 | 1.007714 | 9.038908 | 12.240296 | 1.038980 |
| Cd5l | 3.568985 | 41.267291 | 34.833841 | 1.706900 |
| Clec7a | 6.088269 | 15.646662 | 12.750308 | 3.562240 |
| Nxpe4 | 1.040044 | 1.088409 | 1.530037 | 2.077976 |
| Mktn1 | 65.249450 | 51.989306 | 63.904543 | 52.691460 |
| 2310030G06Rik | 1.175666 | 0.997397 | 1.275031 | 1.558480 |
| Lgals3 | 24.479041 | 63.958067 | 112.151708 | 21.892930 |
| Adam8 | 8.691529 | 19.324562 | 34.884842 | 4.749650 |
| Gm8580 | 5.767900 | 5.455137 | 7.157003 | 5.869530 |
| Morf4l1b | 109.484714 | 97.039848 | 134.512688 | 151.387780 |
| Gm6548 | 39.844572 | 40.872696 | 64.365084 | 68.343800 |

In []:

1