# main

July 11, 2021

## 1 Metrics summary for splicing

```python
[1]: import numpy as np
     import pandas as pd
```

```python
[2]: df = pd.read_csv('../../_m/cluster_ds_results_annotated.txt', sep='\t')
     df['chr'] = df.coord.str.split(":", expand=True)[0]
     df.head(2)
```

```
[2]:    clusterID   N                  coord  gene  annotation          FDR   chr
     0  clu_739_-   3  chrX:53217966-53220839  KDM5C   annotated  7.290000e-35  chrX
     1  clu_729_-  12  chrX:53176622-53193437  KDM5C     cryptic  2.620000e-25  chrX
```

```python
[3]: print("There are %d unique DS!" % len(df.clusterID.unique()))
     print("There are %d unique DS (gene name)!" % len(df.gene.unique()))
```

```
There are 31 unique DS!
There are 27 unique DS (gene name)!
```

```python
[4]: dfx = df[(df['chr']=='chrX')].copy()
     print("There are %d unique DS for allosomes!" % len(dfx.clusterID.unique()))
     print("There are %d unique DS (gene name) for allosomes!" % len(dfx.gene.
      ↪unique()))
```

```
There are 7 unique DS for allosomes!
There are 5 unique DS (gene name) for allosomes!
```

```python
[5]: print("There are %d unique DS for autosomes!" % (len(df.clusterID.unique()) -␣
      ↪len(dfx.clusterID.unique())))
     print("There are %d unique DS (gene name) for autosomes!" % (len(df.gene.
      ↪unique()) - len(dfx.gene.unique())))
```

```
There are 24 unique DS for autosomes!
There are 22 unique DS (gene name) for autosomes!
```

```python
[ ]:
```