fig_color_18S

July 30, 2020

0.0.1 Fig color 18S (Fig 2)

- Read in the structural positions of the probe target sites and write a file for Ribovision in order to make Fig. 2

```
[1]: #Imports
     import sys
     import pandas as pd
     import matplotlib.pyplot as plt
     import os
     import gffutils
     import seaborn as sns
     import numpy as np
     import scipy.stats

     sys.path.append('../scripts/')
     from plot_helpers import *

     %matplotlib inline
     %load_ext autoreload
     %autoreload 2
```

```
[2]: #Make outdir and load the data
     outdir = '../figures/F2/'
     os.makedirs(outdir, exist_ok = True)
```

```
[3]: #Assign colors to the structure based on no probe, one probe, overlapping probes
     len_target = 1995
     target_vals = np.array([0]*len_target)

     data_file = '../figures/F1/TableS1_18S_candidate_properties.csv'
     data_df = pd.read_csv(data_file, index_col = 'probe_num')
     short_probe_df = data_df.loc[1:30].copy()
     short_probe_df['py_struc_start'] = short_probe_df['structure_start'] - 1
     short_probe_df.sort_values('structure_start', inplace = True)

     #Try the one ahead and the one behind together
     l = short_probe_df[['py_struc_start', 'structure_end']].values
```

```python
probe_vals = [range(*i) for i in l]

#Give even probes score of 1 and even probes a score of 10.
#Overlapping regions will therefore have a score of 11.
for p in range(0, len(probe_vals)):
    if p%2 == 0:
        target_vals[probe_vals[p]] += 1
    else:
        target_vals[probe_vals[p]] += 10
```

```python
#Map the regions to specified colors and write the ribovision outfile
#http://apollo.chemistry.gatech.edu/RiboVision/
#http://apollo.chemistry.gatech.edu/RiboVision/Documentation/UserDataTutorial1.
 ↪html
#colors can take hex codes except for the following
#"#000000" or "#858585"
#0 - gray
#1 - wine
#2 - indigo
#11 - olive
color_dict = {0: color_dict['grey'], 1: color_dict['wine'],
              10: color_dict['indigo'], 11: color_dict['olive']}

colors = [color_dict[i] for i in target_vals]
residues = ['18S:%s' % i for i in range(1, len_target + 1)]
df = pd.DataFrame({'ColorCol': colors, 'resNum': residues})
df.to_csv(os.path.join(outdir,'ribovision_positions.csv'))
```

```python
```