

main

August 31, 2021

## 1 Summary of features for interaction analysis

```
[1]: import pandas as pd
```

### 1.1 Function

```
[2]: def get_summary(feature):
    fn = "../../_m/%s/significant_geneSNP_pairs_3tissues.tsv" % feature
    df = pd.read_csv(fn, sep='\t')
    cd = df[(df["N_Regions_Shared"] == 2) & (df["Hippocampus"] == 0)].shape[0]
    ch = df[(df["N_Regions_Shared"] == 2) & (df["DLPFC"] == 0)].shape[0]
    dh = df[(df["N_Regions_Shared"] == 2) & (df["Caudate"] == 0)].shape[0]
    cc = df[(df["N_Regions_Shared"] == 1) & (df["Caudate"] == 1)].shape[0]
    dd = df[(df["N_Regions_Shared"] == 1) & (df["DLPFC"] == 1)].shape[0]
    hh = df[(df["N_Regions_Shared"] == 1) & (df["Hippocampus"] == 1)].shape[0]
    print(feature.upper())
    #print(df.groupby("N_Regions_Shared").size())
    print("There are {} ({:.1%}) tissue specific eFeatures!".format(df.
    ↳groupby("N_Regions_Shared").size()[1],
                                                    df.
    ↳groupby("N_Regions_Shared").size()[1] / df.shape[0]))
    print("There are %d caudate specific eFeatures!" % cc)
    print("There are %d DLPFC specific eFeatures!" % dd)
    print("There are %d hippocampus specific eFeatures!\n" % hh)
    print("There are {} ({:.1%}) eFeatures shared between two brain regions!".
    ↳format(df.groupby("N_Regions_Shared").size()[2],
                                                    df.
    ↳groupby("N_Regions_Shared").size()[2] / df.shape[0]))
    print("There are %d shared only between caudate and DLPFC!" % cd)
    print("There are %d shared only between caudate and hippocampus!" % ch)
    print("There are %d shared only between DLPFC and hippocampus!\n" % dh)
    print("There are {} ({:.1%}) eFeatures shared across brain regions!\n".
    ↳format(df.groupby("N_Regions_Shared").size()[3],
                                                    df.
    ↳groupby("N_Regions_Shared").size()[3] / df.shape[0]))
```

## 1.2 Summary

```
[3]: for feature in ["genes", "transcripts", "exons", "junctions"]:  
      get_summary(feature)
```

### GENES

There are 1998 (12.7%) tissue specific eFeatures!

There are 1932 caudate specific eFeatures!

There are 66 DLPFC specific eFeatures!

There are 0 hippocampus specific eFeatures!

There are 1377 (8.8%) eFeatures shared between two brain regions!

There are 955 shared only between caudate and DLPFC!

There are 422 shared only between caudate and hippocampus!

There are 0 shared only between DLPFC and hippocampus!

There are 12327 (78.5%) eFeatures shared across brain regions!

### TRANSCRIPTS

There are 6976 (22.1%) tissue specific eFeatures!

There are 6951 caudate specific eFeatures!

There are 25 DLPFC specific eFeatures!

There are 0 hippocampus specific eFeatures!

There are 2166 (6.9%) eFeatures shared between two brain regions!

There are 1388 shared only between caudate and DLPFC!

There are 778 shared only between caudate and hippocampus!

There are 0 shared only between DLPFC and hippocampus!

There are 22433 (71.0%) eFeatures shared across brain regions!

### EXONS

There are 6559 (13.7%) tissue specific eFeatures!

There are 6422 caudate specific eFeatures!

There are 137 DLPFC specific eFeatures!

There are 0 hippocampus specific eFeatures!

There are 4292 (9.0%) eFeatures shared between two brain regions!

There are 2830 shared only between caudate and DLPFC!

There are 1461 shared only between caudate and hippocampus!

There are 1 shared only between DLPFC and hippocampus!

There are 36873 (77.3%) eFeatures shared across brain regions!

### JUNCTIONS

There are 10635 (18.1%) tissue specific eFeatures!

There are 10369 caudate specific eFeatures!

There are 266 DLPFC specific eFeatures!

There are 0 hippocampus specific eFeatures!

There are 4750 (8.1%) eFeatures shared between two brain regions!

There are 3354 shared only between caudate and DLPFC!

There are 1396 shared only between caudate and hippocampus!

There are 0 shared only between DLPFC and hippocampus!

There are 43250 (73.8%) eFeatures shared across brain regions!

[ ]: