

main

September 27, 2021

1 Summary of features for interaction analysis

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

1.1 Function

```
[2]: def get_summary(feature):  
    fn = "../_m/%s/brainseq_ancestry_4tissues_mashr.tsv" % feature  
    df = pd.read_csv(fn, sep='\t')  
    cc = df[(df["Caudate"] == 1)].copy()  
    gg = df[(df["Dentate Gyrus"] == 1)].copy()  
    dd = df[(df["DLPFC"] == 1)].copy()  
    hh = df[(df["Hippocampus"] == 1)].copy()  
    sig = df[(df["N_Regions_Shared"] != 0)].copy()  
    print(feature.upper())  
    #print(df.groupby("N_Regions_Shared").size())  
    print("There are {} ({:.1%}) Caudate specific eFeatures!".  
→format(cc[(cc["N_Regions_Shared"] == 1)].shape[0],  
    cc[(cc["N_Regions_Shared"] == 1)].shape[0]/cc.shape[0]))  
    print("There are {} ({:.1%}) Dentate Gyrus specific eFeatures!".  
→format(gg[(gg["N_Regions_Shared"] == 1)].shape[0],  
    gg[(gg["N_Regions_Shared"] == 1)].shape[0]/gg.shape[0]))  
    print("There are {} ({:.1%}) DLPFC specific eFeatures!".  
→format(dd[(dd["N_Regions_Shared"] == 1)].shape[0],  
    dd[(dd["N_Regions_Shared"] == 1)].shape[0]/dd.shape[0]))  
    print("There are {} ({:.1%}) Hippocampus specific eFeatures!\n".  
→format(hh[(hh["N_Regions_Shared"] == 1)].shape[0],  
    hh[(hh["N_Regions_Shared"] == 1)].shape[0]/hh.shape[0]))  
    print("There are {} ({:.1%}) DE {} shared between two brain regions!".  
→format(df.groupby("N_Regions_Shared").size()[2],  
    df.groupby("N_Regions_Shared").size()[2]/sig.shape[0],
```

```

→ feature))
    print("There are {} ({:.1%}) DE {} shared across three brain regions!".
→format(df.groupby("N_Regions_Shared").size()[3],

→      df.groupby("N_Regions_Shared").size()[3]/sig.shape[0],

→      feature))
    print("There are {} ({:.1%}) DE {} shared across four brain regions!\n".
→format(df.groupby("N_Regions_Shared").size()[4],

→      df.groupby("N_Regions_Shared").size()[4]/sig.shape[0],

→      feature))

```

1.2 Summary

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

GENES

There are 930 (18.7%) Caudate specific eFeatures!

There are 907 (21.8%) Dentate Gyrus specific eFeatures!

There are 790 (15.9%) DLPFC specific eFeatures!

There are 808 (16.1%) Hippocampus specific eFeatures!

There are 2387 (26.6%) DE genes shared between two brain regions!

There are 1731 (19.3%) DE genes shared across three brain regions!

There are 1434 (16.0%) DE genes shared across four brain regions!

TRANSCRIPTS

There are 2162 (24.2%) Caudate specific eFeatures!

There are 3549 (33.6%) Dentate Gyrus specific eFeatures!

There are 2402 (23.5%) DLPFC specific eFeatures!

There are 2732 (25.4%) Hippocampus specific eFeatures!

There are 6055 (27.3%) DE transcripts shared between two brain regions!

There are 3702 (16.7%) DE transcripts shared across three brain regions!

There are 1614 (7.3%) DE transcripts shared across four brain regions!

EXONS

There are 10388 (27.3%) Caudate specific eFeatures!

There are 11647 (37.5%) Dentate Gyrus specific eFeatures!

There are 11385 (29.0%) DLPFC specific eFeatures!

There are 10077 (26.4%) Hippocampus specific eFeatures!

There are 17001 (21.1%) DE exons shared between two brain regions!

There are 11513 (14.3%) DE exons shared across three brain regions!
There are 8647 (10.7%) DE exons shared across four brain regions!

JUNCTIONS

There are 3214 (27.1%) Caudate specific eFeatures!
There are 4107 (39.4%) Dentate Gyrus specific eFeatures!
There are 3917 (31.6%) DLPFC specific eFeatures!
There are 3187 (26.7%) Hippocampus specific eFeatures!

There are 5323 (20.5%) DE junctions shared between two brain regions!
There are 3429 (13.2%) DE junctions shared across three brain regions!
There are 2819 (10.8%) DE junctions shared across four brain regions!

[]: