

main

September 22, 2021

1 Examine extreme partials

```
[1]: import os, errno
import functools
import pandas as pd
import collections as cx
from pybiomart import Dataset
# GO analysis
from goatools.base import download_go_basic_obo
from goatools.base import download_ncbi_associations
from goatools.obo_parser import GODag
from goatools.anno.genetogo_reader import Gene2GoReader
from goatools.goea.go_enrichment_ns import GOEnrichmentStudyNS
```

1.1 Functions

1.1.1 Cached functions

```
[2]: @functools.lru_cache()
def get_database():
    dataset = Dataset(name="hsapiens_gene_ensembl",
                      host="http://www.ensembl.org",
                      use_cache=True)
    db = dataset.query(attributes=["ensembl_gene_id",
                                  "external_gene_name",
                                  "entrezgene_id"],
                       use_attr_names=True).dropna(subset=['entrezgene_id'])
    return db
```

1.1.2 Simple functions

```
[3]: def mkdir_p(directory):
    """
    Make a directory if it does not already exist.

    Input: Directory name
    """
    try:
```

```

        os.makedirs(directory)
    except OSError as e:
        if e.errno != errno.EEXIST:
            raise

def extract_top_bottom(tissue, ml_df, percent=0.05):
    df = ml_df[(ml_df["Tissue"] == tissue)].sort_values("Partial_R2",
    ↪ascending=False)
    df["ensemblID"] = df.Geneid.str.replace("\\.*", "", regex=True)
    n = round(df.shape[0] * percent)
    top = df.head(n)
    bottom = df.tail(n)
    return top, bottom

def extract_extremes(tissue, ml_df, val1, val2):
    df = ml_df[(ml_df["Tissue"] == tissue)].sort_values("Partial_R2",
    ↪ascending=False)
    df["ensemblID"] = df.Geneid.str.replace("\\.*", "", regex=True)
    return df[(df["Partial_R2"] > val1)], df[(df["Partial_R2"] < val2)]

def old_convert2entrez(tissue, ml_df, percent):
    top, bottom = extract_top_bottom(tissue, ml_df, percent)
    df1 = top.merge(get_database(), left_on='ensemblID',
                     right_on='ensembl_gene_id')
    df2 = bottom.merge(get_database(), left_on="ensemblID",
                       right_on="ensembl_gene_id")
    return df1, df2

def convert2entrez(top, bottom):
    df1 = top.merge(get_database(), left_on='ensemblID',
                     right_on='ensembl_gene_id')
    df2 = bottom.merge(get_database(), left_on="ensemblID",
                       right_on="ensembl_gene_id")
    return df1, df2

def obo_annotation(alpha=0.05):
    # database annotation
    fn_obo = download_go_basic_obo()
    fn_gene2go = download_ncbi_associations() # must be gunzip to work
    obodag = GODag(fn_obo) # downloads most up-to-date
    anno_hs = Gene2GoReader(fn_gene2go, taxids=[9606])
    # get associations

```

```

ns2assoc = anno_hs.get_ns2assoc()
for nspc, id2gos in ns2assoc.items():
    print("{NS} {N:,} annotated human genes".format(NS=nspc, N=len(id2gos)))
goeaobj = GSEnrichmentStudyNS(
    get_database()['entrezgene_id'], # List of human genes with entrez IDs
    ns2assoc, # geneid/GO associations
    obodag, # Ontologies
    propagate_counts = False,
    alpha = alpha, # default significance cut-off
    methods = ['fdr_bh'])
return goeaobj

def run_goea(tissue, top, bottom, dname, label=''):
    df1, df2 = convert2entrez(top, bottom)
    t_name = tissue.lower().replace(" ", "_")
    d = {"Top": df1, "Bottom": df2}
    for study in ["Top", "Bottom"]:
        print(study)
        df = d[study]
        geneids_study = {z[0]:z[1] for z in zip(df['entrezgene_id'],
↳df['external_gene_name'])}
        goeaobj = obo_annotation()
        goea_results_all = goeaobj.run_study(geneids_study)
        goea_results_sig = [r for r in goea_results_all if r.p_fdr_bh < 0.05]
        ctr = cx.Counter([r.NS for r in goea_results_sig])
        print('Significant results[{TOTAL}] = {BP} BP + {MF} MF + {CC} CC'.
↳format(
            TOTAL=len(goea_results_sig),
            BP=ctr['BP'], MF=ctr['MF'], CC=ctr['CC']))
        if label == '':
            goeaobj.wr_xlsx("%s/%s_GO_analysis_%s.xlsx" % (dname, t_name,
↳study),
                                goea_results_sig)
            goeaobj.wr_txt("%s/%s_GO_analysis_%s.txt" % (dname, t_name, study),
                                goea_results_sig)
        else:
            goeaobj.wr_xlsx("%s/%s_GO_analysis_%s_%s.xlsx" % (dname, t_name,
↳study, label),
                                goea_results_sig)
            goeaobj.wr_txt("%s/%s_GO_analysis_%s_%s.txt" % (dname, t_name,
↳study, label),
                                goea_results_sig)

```

1.2 Extract by tissue by Pst

1.2.1 Elastic net estimated Pst

```
[4]: model = "enet"  
      mkdir_p(model)
```

```
[5]: enet = pd.read_csv("../..//partial_r2/enet_partial_r2_metrics.tsv", sep='\t')  
      print(enet.shape)  
      enet.groupby("Tissue").size()
```

(9324, 6)

```
[5]: Tissue  
      Caudate          2929  
      DLPFC            2711  
      Dentate Gyrus    773  
      Hippocampus      2911  
      dtype: int64
```

```
[6]: enet[(enet["Partial_R2"] > 0.25)].groupby("Tissue").size()
```

```
[6]: Tissue  
      Caudate          452  
      DLPFC            446  
      Dentate Gyrus    172  
      Hippocampus      366  
      dtype: int64
```

```
[7]: enet[(enet["Partial_R2"] < 0.025)].groupby("Tissue").size()
```

```
[7]: Tissue  
      Caudate          741  
      DLPFC            658  
      Dentate Gyrus    111  
      Hippocampus      854  
      dtype: int64
```

Enrichment and extraction

```
[8]: percent = 0.05; val1 = 0.25; val2 = 0.025  
      top_df = pd.DataFrame()  
      bottom_df = pd.DataFrame()  
      for tissue in ["Caudate", "DLPFC", "Dentate Gyrus", "Hippocampus"]:  
          top, bottom = extract_extremes(tissue, enet, val1, val2)  
          run_goea(tissue, top, bottom, model)  
          top_df = pd.concat([top_df, top], axis=0)  
          bottom_df = pd.concat([bottom_df, bottom], axis=0)
```

Top

```
requests.get(http://purl.obolibrary.org/obo/go/go-basic.obo, stream=True)
```

WROTE: go-basic.obo

```
FTP RETR ftp.ncbi.nlm.nih.gov gene/DATA gene2go.gz -> gene2go.gz
gunzip gene2go.gz
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.430023 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes
```

```
Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association
```

```
Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association
```

```
Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association
```

```
Run BP Gene Ontology Analysis: current study set of 333 IDs ... 68%    227 of
333 study items found in association
100%    333 of    333 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
882 GO terms are associated with    227 of    333 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
```

```
Run CC Gene Ontology Analysis: current study set of 333 IDs ... 74%    247 of
333 study items found in association
100%    333 of    333 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
254 GO terms are associated with    247 of    333 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
```

```
Run MF Gene Ontology Analysis: current study set of 333 IDs ... 72%    239 of
```

```

333 study items found in association
100% 333 of 333 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
372 GO terms are associated with 239 of 333 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
0 items. NOT WRITING enet/caudate_GO_analysis_Top.xlsx
0 GOEA results. NOT WRITING enet/caudate_GO_analysis_Top.txt
Bottom
EXISTS: go-basic.obo
EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.966386 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 587 IDs ... 78% 459 of
587 study items found in association
100% 587 of 587 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
2,074 GO terms are associated with 459 of 587 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 587 IDs ... 85% 498 of

```

```

587 study items found in association
100% 587 of 587 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
450 GO terms are associated with 498 of 587 study items
METHOD fdr_bh:
5 GO terms found significant (< 0.05=alpha) ( 5 enriched + 0
purified): statsmodels fdr_bh
313 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 587 IDs ... 84% 494 of
587 study items found in association
100% 587 of 587 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
667 GO terms are associated with 494 of 587 study items
METHOD fdr_bh:
2 GO terms found significant (< 0.05=alpha) ( 2 enriched + 0
purified): statsmodels fdr_bh
363 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)
Significant results[7] = 0 BP + 2 MF + 5 CC
7 items WROTE: enet/caudate_GO_analysis_Bottom.xlsx
7 GOEA results for 437 study items. WROTE:
enet/caudate_GO_analysis_Bottom.txt
Top
EXISTS: go-basic.obo
EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.270911 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

```

```

Run BP Gene Ontology Analysis: current study set of 337 IDs ... 68%    230 of
337 study items found in association
100%    337 of    337 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    907 GO terms are associated with    230 of    337 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) (  1 enriched +  0
purified): statsmodels fdr_bh
    6 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 337 IDs ... 73%    246 of
337 study items found in association
100%    337 of    337 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    274 GO terms are associated with    246 of    337 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) (  1 enriched +  0
purified): statsmodels fdr_bh
    6 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 337 IDs ... 71%    239 of
337 study items found in association
100%    337 of    337 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    359 GO terms are associated with    239 of    337 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (  0 enriched +  0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[2] = 1 BP + 0 MF + 1 CC
    2 items WROTE: enet/dlpfc_GO_analysis_Top.xlsx
    2 GOEA results for    11 study items. WROTE:
enet/dlpfc_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.755375 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

```



```

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 519 IDs ... 85%    442 of
519 study items found in association
100%    519 of    519 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
2,078 GO terms are associated with    442 of    519 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) ( 1 enriched + 0
purified): statsmodels fdr_bh
    25 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 519 IDs ... 89%    460 of
519 study items found in association
100%    519 of    519 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
447 GO terms are associated with    460 of    519 study items
METHOD fdr_bh:
    12 GO terms found significant (< 0.05=alpha) ( 12 enriched + 0
purified): statsmodels fdr_bh
    350 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 519 IDs ... 86%    448 of
519 study items found in association
100%    519 of    519 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
671 GO terms are associated with    448 of    519 study items
METHOD fdr_bh:
    6 GO terms found significant (< 0.05=alpha) ( 6 enriched + 0
purified): statsmodels fdr_bh
    326 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

```

```

Significant results[19] = 1 BP + 6 MF + 12 CC
    19 items WROTE: enet/dlpfc_GO_analysis_Bottom.xlsx
    19 GOEA results for 420 study items. WROTE:
enet/dlpfc_GO_analysis_Bottom.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.504341 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 135 IDs ... 77% 104 of
135 study items found in association
100% 135 of 135 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
656 GO terms are associated with 104 of 135 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 135 IDs ... 81% 109 of
135 study items found in association
100% 135 of 135 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
168 GO terms are associated with 109 of 135 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)

```

```

    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 135 IDs ... 77%    104 of
135 study items found in association
100%    135 of    135 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    232 GO terms are associated with    104 of    135 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/dentate_gyrus_GO_analysis_Top.xlsx
    0 GOEA results. NOT WRITING enet/dentate_gyrus_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.650448 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 91 IDs ... 77%    70 of
91 study items found in association
100%    91 of    91 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    582 GO terms are associated with    70 of    91 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)

```

```

    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 91 IDs ... 90%      82 of
91 study items found in association
100%      91 of      91 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    123 GO terms are associated with      82 of      91 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 91 IDs ... 85%      77 of
91 study items found in association
100%      91 of      91 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    188 GO terms are associated with      77 of      91 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/dentate_gyrus_GO_analysis_Bottom.xlsx
    0 GOEA results. NOT WRITING enet/dentate_gyrus_GO_analysis_Bottom.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.520065 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact

```

```

70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 277 IDs ... 66%    183 of
277 study items found in association
100%    277 of    277 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
        746 GO terms are associated with    183 of    277 study items
    METHOD fdr_bh:
        0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
        0 study items associated with significant GO IDs (enriched)
        0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 277 IDs ... 73%    202 of
277 study items found in association
100%    277 of    277 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
        224 GO terms are associated with    202 of    277 study items
    METHOD fdr_bh:
        0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
        0 study items associated with significant GO IDs (enriched)
        0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 277 IDs ... 69%    190 of
277 study items found in association
100%    277 of    277 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
        312 GO terms are associated with    190 of    277 study items
    METHOD fdr_bh:
        0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
        0 study items associated with significant GO IDs (enriched)
        0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/hippocampus_GO_analysis_Top.xlsx
    0 GOEA results. NOT WRITING enet/hippocampus_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.992221 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes

```

MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 676 IDs ... 84% 568 of
676 study items found in association

100% 676 of 676 study items found in population(29107)

Calculating 12,416 uncorrected p-values using fisher_scipy_stats

12,416 GO terms are associated with 17,843 of 29,107 population items

2,382 GO terms are associated with 568 of 676 study items

METHOD fdr_bh:

1 GO terms found significant (< 0.05=alpha) (1 enriched + 0
purified): statsmodels fdr_bh

11 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 676 IDs ... 88% 593 of
676 study items found in association

100% 676 of 676 study items found in population(29107)

Calculating 1,752 uncorrected p-values using fisher_scipy_stats

1,752 GO terms are associated with 18,721 of 29,107 population items

523 GO terms are associated with 593 of 676 study items

METHOD fdr_bh:

16 GO terms found significant (< 0.05=alpha) (16 enriched + 0
purified): statsmodels fdr_bh

533 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 676 IDs ... 85% 577 of
676 study items found in association

100% 676 of 676 study items found in population(29107)

Calculating 4,418 uncorrected p-values using fisher_scipy_stats

4,418 GO terms are associated with 17,841 of 29,107 population items

785 GO terms are associated with 577 of 676 study items

METHOD fdr_bh:

1 GO terms found significant (< 0.05=alpha) (1 enriched + 0
purified): statsmodels fdr_bh

442 study items associated with significant GO IDs (enriched)

```

    0 study items associated with significant GO IDs (purified)
Significant results[18] = 1 BP + 1 MF + 16 CC
    18 items Wrote: enet/hippocampus_GO_analysis_Bottom.xlsx
    18 GOEA results for 577 study items. Wrote:
enet/hippocampus_GO_analysis_Bottom.txt

```

Save extremes

```

[9]: top_df["Variation_Explained"] = "High"
    bottom_df["Variation_Explained"] = "Low"
    dt = pd.concat([top_df, bottom_df], axis=0)
    dt.to_csv("%s/extremes_partial_r2_enet.tsv" % model, sep='\t', index=False)
    dt.head()

```

```

[9]:
      Geneid  N_Features  Partial_R2  Full_R2  Reduced_R2  \
1313  ENSG00000166435.15         31    0.954305   10.019849   219.274557
2193  ENSG00000228906.1         20    0.909336   20.355269   224.512782
 54    ENSG00000042317.16         11    0.867759   29.948981   226.472712
2150  ENSG00000226752.8         40    0.866798   25.970472   194.970508
2773  ENSG00000272865.1         31    0.865386   29.756291   221.048565

```

```

      Tissue      ensemblID  Variation_Explained
1313  Caudate  ENSG00000166435             High
2193  Caudate  ENSG00000228906             High
 54    Caudate  ENSG00000042317             High
2150  Caudate  ENSG00000226752             High
2773  Caudate  ENSG00000272865             High

```

1.2.2 Random forest estimated Pst

```

[10]: model = "rf"
    mkdir_p(model)

```

```

[11]: rf = pd.read_csv("../..//partial_r2/rf_partial_r2_metrics.tsv", sep='\t')
    print(rf.shape)
    rf.groupby("Tissue").size()

```

```

(9295, 6)

```

```

[11]: Tissue
      Caudate      2925
      DLPFC      2691
Dentate Gyrus      773
Hippocampus      2906
dtype: int64

```

```

[12]: rf[(rf["Partial_R2"] > 0.25)].groupby("Tissue").size()

```

```
[12]: Tissue
      Caudate          461
      DLPFC            449
      Dentate Gyrus    253
      Hippocampus      363
      dtype: int64
```

```
[13]: rf[(rf["Partial_R2"] < 0.01)].groupby("Tissue").size()
```

```
[13]: Tissue
      Caudate          541
      DLPFC            429
      Dentate Gyrus     47
      Hippocampus      681
      dtype: int64
```

Enrichment and extraction

```
[14]: percent = 0.05; val1 = 0.25; val2 = 0.01
      top_df = pd.DataFrame()
      bottom_df = pd.DataFrame()
      for tissue in ["Caudate", "DLPFC", "Dentate Gyrus", "Hippocampus"]:
          top, bottom = extract_extremes(tissue, rf, val1, val2)
          run_goea(tissue, top, bottom, model)
          top_df = pd.concat([top_df, top], axis=0)
          bottom_df = pd.concat([bottom_df, bottom], axis=0)
```

Top

```
EXISTS: go-basic.obo
EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.563487 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes
```

Load BP Gene Ontology Analysis ...

```
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association
```

Load CC Gene Ontology Analysis ...

```
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association
```

Load MF Gene Ontology Analysis ...

```
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association
```



```

Run BP Gene Ontology Analysis: current study set of 348 IDs ... 72%    251 of
348 study items found in association
100%    348 of    348 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    1,174 GO terms are associated with    251 of    348 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 348 IDs ... 77%    268 of
348 study items found in association
100%    348 of    348 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    292 GO terms are associated with    268 of    348 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 348 IDs ... 75%    261 of
348 study items found in association
100%    348 of    348 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    428 GO terms are associated with    261 of    348 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/caudate_GO_analysis_Top.xlsx
    0 GOEA results. NOT WRITING rf/caudate_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:05.017316 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

```

```

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 412 IDs ... 76%    312 of
412 study items found in association
100%    412 of    412 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
1,449 GO terms are associated with    312 of    412 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 412 IDs ... 82%    338 of
412 study items found in association
100%    412 of    412 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
373 GO terms are associated with    338 of    412 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) ( 1 enriched + 0
purified): statsmodels fdr_bh
    104 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 412 IDs ... 82%    337 of
412 study items found in association
100%    412 of    412 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
479 GO terms are associated with    337 of    412 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) ( 1 enriched + 0
purified): statsmodels fdr_bh
    238 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[2] = 0 BP + 1 MF + 1 CC

```

```

    2 items WROTE: rf/caudate_GO_analysis_Bottom.xlsx
    2 GOEA results for 269 study items. WROTE:
rf/caudate_GO_analysis_Bottom.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.253940 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 358 IDs ... 69% 247 of
358 study items found in association
100% 358 of 358 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
984 GO terms are associated with 247 of 358 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 358 IDs ... 74% 264 of
358 study items found in association
100% 358 of 358 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
276 GO terms are associated with 264 of 358 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

```

```

Run MF Gene Ontology Analysis: current study set of 358 IDs ... 71%    255 of
358 study items found in association
100%    358 of    358 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    358 GO terms are associated with    255 of    358 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/dlpfc_GO_analysis_Top.xlsx
    0 GOEA results. NOT WRITING rf/dlpfc_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.500557 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 333 IDs ... 81%    270 of
333 study items found in association
100%    333 of    333 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    1,201 GO terms are associated with    270 of    333 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

```

```

Run CC Gene Ontology Analysis: current study set of 333 IDs ... 88%    293 of
333 study items found in association
100%    333 of    333 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    323 GO terms are associated with    293 of    333 study items
METHOD fdr_bh:
    5 GO terms found significant (< 0.05=alpha) (  5 enriched +  0
purified): statsmodels fdr_bh
    180 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 333 IDs ... 86%    287 of
333 study items found in association
100%    333 of    333 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    477 GO terms are associated with    287 of    333 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) (  1 enriched +  0
purified): statsmodels fdr_bh
    207 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[6] = 0 BP + 1 MF + 5 CC
    6 items WROTE: rf/dlpfc_GO_analysis_Bottom.xlsx
    6 GOEA results for    257 study items. WROTE:
rf/dlpfc_GO_analysis_Bottom.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:05.334015 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact

```

```

70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 189 IDs ... 78%    148 of
189 study items found in association
100%    189 of    189 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    900 GO terms are associated with    148 of    189 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 189 IDs ... 84%    159 of
189 study items found in association
100%    189 of    189 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    195 GO terms are associated with    159 of    189 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 189 IDs ... 83%    157 of
189 study items found in association
100%    189 of    189 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    282 GO terms are associated with    157 of    189 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/dentate_gyrus_GO_analysis_Top.xlsx
    0 GOEA results. NOT WRITING rf/dentate_gyrus_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:05.019642 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes

```

MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 38 IDs ... 71% 27 of
38 study items found in association

100% 38 of 38 study items found in population(29107)

Calculating 12,416 uncorrected p-values using fisher_scipy_stats

12,416 GO terms are associated with 17,843 of 29,107 population items

171 GO terms are associated with 27 of 38 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 38 IDs ... 79% 30 of
38 study items found in association

100% 38 of 38 study items found in population(29107)

Calculating 1,752 uncorrected p-values using fisher_scipy_stats

1,752 GO terms are associated with 18,721 of 29,107 population items

62 GO terms are associated with 30 of 38 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 38 IDs ... 79% 30 of
38 study items found in association

100% 38 of 38 study items found in population(29107)

Calculating 4,418 uncorrected p-values using fisher_scipy_stats

4,418 GO terms are associated with 17,841 of 29,107 population items

73 GO terms are associated with 30 of 38 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

```

    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/dentate_gyrus_GO_analysis_Bottom.xlsx
    0 GOEA results. NOT WRITING rf/dentate_gyrus_GO_analysis_Bottom.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.531564 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 276 IDs ... 67%    186 of
276 study items found in association
100%    276 of    276 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    791 GO terms are associated with    186 of    276 study items
    METHOD fdr_bh:
        0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
        0 study items associated with significant GO IDs (enriched)
        0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 276 IDs ... 74%    205 of
276 study items found in association
100%    276 of    276 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    254 GO terms are associated with    205 of    276 study items
    METHOD fdr_bh:
        0 GO terms found significant (< 0.05=alpha) (    0 enriched +    0
purified): statsmodels fdr_bh
        0 study items associated with significant GO IDs (enriched)

```



```

    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 276 IDs ... 67%    186 of
276 study items found in association
100%    276 of    276 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    342 GO terms are associated with    186 of    276 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/hippocampus_GO_analysis_Top.xlsx
    0 GOEA results. NOT WRITING rf/hippocampus_GO_analysis_Top.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.710429 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 526 IDs ... 83%    439 of
526 study items found in association
100%    526 of    526 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    2,006 GO terms are associated with    439 of    526 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)

```

```

0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 526 IDs ... 86%    454 of
526 study items found in association
100%    526 of    526 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    450 GO terms are associated with    454 of    526 study items
METHOD fdr_bh:
    7 GO terms found significant (< 0.05=alpha) ( 7 enriched + 0
purified): statsmodels fdr_bh
    329 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 526 IDs ... 84%    441 of
526 study items found in association
100%    526 of    526 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    645 GO terms are associated with    441 of    526 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) ( 1 enriched + 0
purified): statsmodels fdr_bh
    329 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[8] = 0 BP + 1 MF + 7 CC
    8 items Wrote: rf/hippocampus_GO_analysis_Bottom.xlsx
    8 GOEA results for 408 study items. Wrote:
rf/hippocampus_GO_analysis_Bottom.txt

```

Save extremes

```

[15]: top_df["Variation_Explained"] = "High"
      bottom_df["Variation_Explained"] = "Low"
      dt = pd.concat([top_df, bottom_df], axis=0)
      dt.to_csv("%s/extremes_partial_r2_rf.tsv" % model, sep='\t', index=False)
      dt.head()

```

```

[15]:
      Geneid  N_Features  Partial_R2  Full_R2  Reduced_R2  \
1311  ENSG00000166435.15      28    0.888735  24.397504  219.274557
292    ENSG00000101255.10     278    0.878123  26.944644  221.080224
1082  ENSG00000152669.8      68    0.829583  36.868430  216.342463
2581  ENSG00000260565.6     135    0.813190  38.828169  207.848307
2282  ENSG00000233927.4      13    0.809802  43.630645  229.396302

      Tissue  ensemblID  Variation_Explained
1311  Caudate  ENSG00000166435              High
292    Caudate  ENSG00000101255              High

```

1082	Caudate	ENSG00000152669	High
2581	Caudate	ENSG00000260565	High
2282	Caudate	ENSG00000233927	High

1.3 Enrichment of top and bottom 100

```
[16]: model = "enet"
for tissue in ["Caudate", "DLPFC", "Dentate Gyrus", "Hippocampus"]:
    df = enet[(enet["Tissue"] == tissue)].sort_values("Partial_R2",
    ↪ascending=False)
    df["ensemblID"] = df.Geneid.str.replace("\\.*", "", regex=True)
    top = df.head(100)
    bottom = df.tail(100)
    run_goea(tissue, top, bottom, model, 'n100')
```

Top

```
EXISTS: go-basic.obo
EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.232312 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes
```

Load BP Gene Ontology Analysis ...

```
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association
```

Load CC Gene Ontology Analysis ...

```
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association
```

Load MF Gene Ontology Analysis ...

```
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association
```

```
Run BP Gene Ontology Analysis: current study set of 71 IDs ... 65%      46 of
71 study items found in association
```

```
100%      71 of      71 study items found in population(29107)
```

```
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
```

```
12,416 GO terms are associated with 17,843 of 29,107 population items
```

```
221 GO terms are associated with      46 of      71 study items
```

```
METHOD fdr_bh:
```

```
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
```

```
0 study items associated with significant GO IDs (enriched)
```

```
0 study items associated with significant GO IDs (purified)
```

```

Run CC Gene Ontology Analysis: current study set of 71 IDs ... 66%      47 of
71 study items found in association
100%      71 of      71 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    95 GO terms are associated with      47 of      71 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 71 IDs ... 68%      48 of
71 study items found in association
100%      71 of      71 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    90 GO terms are associated with      48 of      71 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/caudate_GO_analysis_Top_n100.xlsx
    0 GOEA results. NOT WRITING enet/caudate_GO_analysis_Top_n100.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.399722 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

```

```

Run BP Gene Ontology Analysis: current study set of 80 IDs ... 68%      54 of
80 study items found in association
100%      80 of      80 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
324 GO terms are associated with      54 of      80 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 80 IDs ... 81%      65 of
80 study items found in association
100%      80 of      80 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
90 GO terms are associated with      65 of      80 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 80 IDs ... 82%      66 of
80 study items found in association
100%      80 of      80 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
115 GO terms are associated with      66 of      80 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
0 items. NOT WRITING enet/caudate_GO_analysis_Bottom_n100.xlsx
0 GOEA results. NOT WRITING enet/caudate_GO_analysis_Bottom_n100.txt
Top
EXISTS: go-basic.obo
EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:05.296941 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

```

```

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 72 IDs ... 60%      43 of
72 study items found in association
100%      72 of      72 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
178 GO terms are associated with      43 of      72 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) ( 1 enriched + 0
purified): statsmodels fdr_bh
    5 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 72 IDs ... 60%      43 of
72 study items found in association
100%      72 of      72 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
72 GO terms are associated with      43 of      72 study items
METHOD fdr_bh:
    4 GO terms found significant (< 0.05=alpha) ( 4 enriched + 0
purified): statsmodels fdr_bh
    4 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 72 IDs ... 58%      42 of
72 study items found in association
100%      72 of      72 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
82 GO terms are associated with      42 of      72 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

```

Significant results[5] = 1 BP + 0 MF + 4 CC
 5 items WROTE: enet/dlpfc_GO_analysis_Top_n100.xlsx
 5 GOEA results for 5 study items. WROTE:
 enet/dlpfc_GO_analysis_Top_n100.txt
 Bottom
 EXISTS: go-basic.obo
 EXISTS: gene2go
 go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
 HMS:0:00:04.871583 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
 gene2go
 BP 18,501 annotated human genes
 CC 19,433 annotated human genes
 MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
 fisher module not installed. Falling back on scipy.stats.fisher_exact
 70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
 fisher module not installed. Falling back on scipy.stats.fisher_exact
 74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
 fisher module not installed. Falling back on scipy.stats.fisher_exact
 70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 77 IDs ... 87% 67 of
 77 study items found in association
 100% 77 of 77 study items found in population(29107)
 Calculating 12,416 uncorrected p-values using fisher_scipy_stats
 12,416 GO terms are associated with 17,843 of 29,107 population items
 520 GO terms are associated with 67 of 77 study items
 METHOD fdr_bh:
 0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
 purified): statsmodels fdr_bh
 0 study items associated with significant GO IDs (enriched)
 0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 77 IDs ... 87% 67 of
 77 study items found in association
 100% 77 of 77 study items found in population(29107)
 Calculating 1,752 uncorrected p-values using fisher_scipy_stats
 1,752 GO terms are associated with 18,721 of 29,107 population items
 150 GO terms are associated with 67 of 77 study items
 METHOD fdr_bh:
 0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
 purified): statsmodels fdr_bh
 0 study items associated with significant GO IDs (enriched)

```

    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 77 IDs ... 88%      68 of
77 study items found in association
100%      77 of      77 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    159 GO terms are associated with      68 of      77 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/dlpfc_GO_analysis_Bottom_n100.xlsx
    0 GOEA results. NOT WRITING enet/dlpfc_GO_analysis_Bottom_n100.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.194186 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 78 IDs ... 76%      59 of
78 study items found in association
100%      78 of      78 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    326 GO terms are associated with      59 of      78 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)

```



```

    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 78 IDs ... 79%      62 of
78 study items found in association
100%      78 of      78 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    106 GO terms are associated with      62 of      78 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 78 IDs ... 76%      59 of
78 study items found in association
100%      78 of      78 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    130 GO terms are associated with      59 of      78 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/dentate_gyrus_GO_analysis_Top_n100.xlsx
    0 GOEA results. NOT WRITING enet/dentate_gyrus_GO_analysis_Top_n100.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.976882 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact

```

```

70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 84 IDs ... 76%      64 of
84 study items found in association
100%      84 of      84 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
  12,416 GO terms are associated with 17,843 of 29,107 population items
    543 GO terms are associated with      64 of      84 study items
  METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 84 IDs ... 89%      75 of
84 study items found in association
100%      84 of      84 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
  1,752 GO terms are associated with 18,721 of 29,107 population items
    117 GO terms are associated with      75 of      84 study items
  METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 84 IDs ... 83%      70 of
84 study items found in association
100%      84 of      84 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
  4,418 GO terms are associated with 17,841 of 29,107 population items
    169 GO terms are associated with      70 of      84 study items
  METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
  0 items. NOT WRITING enet/dentate_gyrus_GO_analysis_Bottom_n100.xlsx
  0 GOEA results. NOT WRITING enet/dentate_gyrus_GO_analysis_Bottom_n100.txt
Top
  EXISTS: go-basic.obo
  EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.303162 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes

```

MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 73 IDs ... 60% 44 of
73 study items found in association

100% 73 of 73 study items found in population(29107)

Calculating 12,416 uncorrected p-values using fisher_scipy_stats

12,416 GO terms are associated with 17,843 of 29,107 population items

203 GO terms are associated with 44 of 73 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 73 IDs ... 70% 51 of
73 study items found in association

100% 73 of 73 study items found in population(29107)

Calculating 1,752 uncorrected p-values using fisher_scipy_stats

1,752 GO terms are associated with 18,721 of 29,107 population items

83 GO terms are associated with 51 of 73 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 73 IDs ... 64% 47 of
73 study items found in association

100% 73 of 73 study items found in population(29107)

Calculating 4,418 uncorrected p-values using fisher_scipy_stats

4,418 GO terms are associated with 17,841 of 29,107 population items

95 GO terms are associated with 47 of 73 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

```

    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/hippocampus_GO_analysis_Top_n100.xlsx
    0 GOEA results. NOT WRITING enet/hippocampus_GO_analysis_Top_n100.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.747972 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 73 IDs ... 73%      53 of
73 study items found in association
100%      73 of      73 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    305 GO terms are associated with      53 of      73 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 73 IDs ... 82%      60 of
73 study items found in association
100%      73 of      73 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    126 GO terms are associated with      60 of      73 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)

```

```

    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 73 IDs ... 84%      61 of
73 study items found in association
100%      73 of      73 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    124 GO terms are associated with      61 of      73 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING enet/hippocampus_GO_analysis_Bottom_n100.xlsx
    0 GOEA results. NOT WRITING enet/hippocampus_GO_analysis_Bottom_n100.txt

```

1.3.1 Random forest

```

[17]: model = "rf"
      for tissue in ["Caudate", "DLPFC", "Dentate Gyrus", "Hippocampus"]:
          df = rf[(rf["Tissue"] == tissue)].sort_values("Partial_R2", ascending=False)
          df["ensemblID"] = df.Geneid.str.replace("\\.*", "", regex=True)
          top = df.head(100)
          bottom = df.tail(100)
          run_goea(tissue, top, bottom, model, 'n100')

```

Top

```

    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.549471 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

```

Load BP Gene Ontology Analysis ...

```

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

```

Load CC Gene Ontology Analysis ...

```

fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

```

Load MF Gene Ontology Analysis ...

```

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

```

```

Run BP Gene Ontology Analysis: current study set of 78 IDs ... 72%      56 of
78 study items found in association
100%      78 of      78 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
294 GO terms are associated with      56 of      78 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 78 IDs ... 76%      59 of
78 study items found in association
100%      78 of      78 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
107 GO terms are associated with      59 of      78 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 78 IDs ... 74%      58 of
78 study items found in association
100%      78 of      78 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
133 GO terms are associated with      58 of      78 study items
METHOD fdr_bh:
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
0 items. NOT WRITING rf/caudate_GO_analysis_Top_n100.xlsx
0 GOEA results. NOT WRITING rf/caudate_GO_analysis_Top_n100.txt
Bottom
EXISTS: go-basic.obo
EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.441260 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

```

```

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 62 IDs ... 71%      44 of
62 study items found in association
100%      62 of      62 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
12,416 GO terms are associated with 17,843 of 29,107 population items
186 GO terms are associated with      44 of      62 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 62 IDs ... 77%      48 of
62 study items found in association
100%      62 of      62 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
1,752 GO terms are associated with 18,721 of 29,107 population items
87 GO terms are associated with      48 of      62 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 62 IDs ... 79%      49 of
62 study items found in association
100%      62 of      62 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
4,418 GO terms are associated with 17,841 of 29,107 population items
76 GO terms are associated with      49 of      62 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

```

```

Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/caudate_GO_analysis_Bottom_n100.xlsx
    0 GOEA results. NOT WRITING rf/caudate_GO_analysis_Bottom_n100.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:05.327410 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 79 IDs ... 61%      48 of
79 study items found in association
100%      79 of      79 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    202 GO terms are associated with      48 of      79 study items
    METHOD fdr_bh:
        0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
        0 study items associated with significant GO IDs (enriched)
        0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 79 IDs ... 65%      51 of
79 study items found in association
100%      79 of      79 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    107 GO terms are associated with      51 of      79 study items
    METHOD fdr_bh:
        1 GO terms found significant (< 0.05=alpha) ( 1 enriched + 0
purified): statsmodels fdr_bh
        2 study items associated with significant GO IDs (enriched)
        0 study items associated with significant GO IDs (purified)

```



```

Run MF Gene Ontology Analysis: current study set of 79 IDs ... 66%      52 of
79 study items found in association
100%      79 of      79 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    91 GO terms are associated with      52 of      79 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[1] = 0 BP + 0 MF + 1 CC
    1 items WROTE: rf/dlpfc_GO_analysis_Top_n100.xlsx
    1 GOEA results for      2 study items. WROTE:
rf/dlpfc_GO_analysis_Top_n100.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.965924 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 66 IDs ... 76%      50 of
66 study items found in association
100%      66 of      66 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    248 GO terms are associated with      50 of      66 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)

```

```

    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 66 IDs ... 83%      55 of
66 study items found in association
100%      66 of      66 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    108 GO terms are associated with      55 of      66 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 66 IDs ... 83%      55 of
66 study items found in association
100%      66 of      66 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    129 GO terms are associated with      55 of      66 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/dlpfc_GO_analysis_Bottom_n100.xlsx
    0 GOEA results. NOT WRITING rf/dlpfc_GO_analysis_Bottom_n100.txt

Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.558997 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact

```

```

70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 77 IDs ... 84%      65 of
77 study items found in association
100%      77 of      77 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
  12,416 GO terms are associated with 17,843 of 29,107 population items
    443 GO terms are associated with      65 of      77 study items
  METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 77 IDs ... 86%      66 of
77 study items found in association
100%      77 of      77 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
  1,752 GO terms are associated with 18,721 of 29,107 population items
    102 GO terms are associated with      66 of      77 study items
  METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 77 IDs ... 83%      64 of
77 study items found in association
100%      77 of      77 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
  4,418 GO terms are associated with 17,841 of 29,107 population items
    156 GO terms are associated with      64 of      77 study items
  METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
  0 items. NOT WRITING rf/dentate_gyrus_GO_analysis_Top_n100.xlsx
  0 GOEA results. NOT WRITING rf/dentate_gyrus_GO_analysis_Top_n100.txt
Bottom
  EXISTS: go-basic.obo
  EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.919144 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes

```

MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...

fisher module not installed. Falling back on scipy.stats.fisher_exact
70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 77 IDs ... 75% 58 of
77 study items found in association

100% 77 of 77 study items found in population(29107)

Calculating 12,416 uncorrected p-values using fisher_scipy_stats

12,416 GO terms are associated with 17,843 of 29,107 population items

548 GO terms are associated with 58 of 77 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 77 IDs ... 83% 64 of
77 study items found in association

100% 77 of 77 study items found in population(29107)

Calculating 1,752 uncorrected p-values using fisher_scipy_stats

1,752 GO terms are associated with 18,721 of 29,107 population items

135 GO terms are associated with 64 of 77 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 77 IDs ... 81% 62 of
77 study items found in association

100% 77 of 77 study items found in population(29107)

Calculating 4,418 uncorrected p-values using fisher_scipy_stats

4,418 GO terms are associated with 17,841 of 29,107 population items

155 GO terms are associated with 62 of 77 study items

METHOD fdr_bh:

0 GO terms found significant (< 0.05=alpha) (0 enriched + 0
purified): statsmodels fdr_bh

0 study items associated with significant GO IDs (enriched)

```

    0 study items associated with significant GO IDs (purified)
Significant results[0] = 0 BP + 0 MF + 0 CC
    0 items. NOT WRITING rf/dentate_gyrus_GO_analysis_Bottom_n100.xlsx
    0 GOEA results. NOT WRITING rf/dentate_gyrus_GO_analysis_Bottom_n100.txt
Top
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.478971 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 71 IDs ... 66%      47 of
71 study items found in association
100%      71 of      71 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    294 GO terms are associated with      47 of      71 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (  0 enriched +  0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)

Run CC Gene Ontology Analysis: current study set of 71 IDs ... 75%      53 of
71 study items found in association
100%      71 of      71 study items found in population(29107)
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
    1,752 GO terms are associated with 18,721 of 29,107 population items
    109 GO terms are associated with      53 of      71 study items
METHOD fdr_bh:
    1 GO terms found significant (< 0.05=alpha) (  1 enriched +  0
purified): statsmodels fdr_bh
    2 study items associated with significant GO IDs (enriched)

```

```

    0 study items associated with significant GO IDs (purified)

Run MF Gene Ontology Analysis: current study set of 71 IDs ... 72%      51 of
71 study items found in association
100%      71 of      71 study items found in population(29107)
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
    4,418 GO terms are associated with 17,841 of 29,107 population items
    117 GO terms are associated with      51 of      71 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (  0 enriched +  0
purified): statsmodels fdr_bh
    0 study items associated with significant GO IDs (enriched)
    0 study items associated with significant GO IDs (purified)
Significant results[1] = 0 BP + 0 MF + 1 CC
    1 items WROTE: rf/hippocampus_GO_analysis_Top_n100.xlsx
    1 GOEA results for      2 study items. WROTE:
rf/hippocampus_GO_analysis_Top_n100.txt
Bottom
    EXISTS: go-basic.obo
    EXISTS: gene2go
go-basic.obo: fmt(1.2) rel(2021-09-01) 47,191 GO Terms
HMS:0:00:04.868952 330,404 annotations, 20,688 genes, 18,642 GOs, 1 taxids READ:
gene2go
BP 18,501 annotated human genes
CC 19,433 annotated human genes
MF 18,194 annotated human genes

Load BP Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,231 of 29,107 population items found in association

Load CC Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    74% 21,438 of 29,107 population items found in association

Load MF Gene Ontology Analysis ...
fisher module not installed. Falling back on scipy.stats.fisher_exact
    70% 20,357 of 29,107 population items found in association

Run BP Gene Ontology Analysis: current study set of 74 IDs ... 80%      59 of
74 study items found in association
100%      74 of      74 study items found in population(29107)
Calculating 12,416 uncorrected p-values using fisher_scipy_stats
    12,416 GO terms are associated with 17,843 of 29,107 population items
    344 GO terms are associated with      59 of      74 study items
METHOD fdr_bh:
    0 GO terms found significant (< 0.05=alpha) (  0 enriched +  0
purified): statsmodels fdr_bh

```

```
0 study items associated with significant GO IDs (enriched)
0 study items associated with significant GO IDs (purified)
```

```
Run CC Gene Ontology Analysis: current study set of 74 IDs ... 86%      64 of
74 study items found in association
```

```
100%      74 of      74 study items found in population(29107)
```

```
Calculating 1,752 uncorrected p-values using fisher_scipy_stats
```

```
1,752 GO terms are associated with 18,721 of 29,107 population items
```

```
148 GO terms are associated with      64 of      74 study items
```

```
METHOD fdr_bh:
```

```
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
```

```
0 study items associated with significant GO IDs (enriched)
```

```
0 study items associated with significant GO IDs (purified)
```

```
Run MF Gene Ontology Analysis: current study set of 74 IDs ... 85%      63 of
74 study items found in association
```

```
100%      74 of      74 study items found in population(29107)
```

```
Calculating 4,418 uncorrected p-values using fisher_scipy_stats
```

```
4,418 GO terms are associated with 17,841 of 29,107 population items
```

```
131 GO terms are associated with      63 of      74 study items
```

```
METHOD fdr_bh:
```

```
0 GO terms found significant (< 0.05=alpha) ( 0 enriched + 0
purified): statsmodels fdr_bh
```

```
0 study items associated with significant GO IDs (enriched)
```

```
0 study items associated with significant GO IDs (purified)
```

```
Significant results[0] = 0 BP + 0 MF + 0 CC
```

```
0 items. NOT WRITING rf/hippocampus_GO_analysis_Bottom_n100.xlsx
```

```
0 GOEA results. NOT WRITING rf/hippocampus_GO_analysis_Bottom_n100.txt
```

```
[ ]:
```