SAGD

Sex-Associated Gene Database

Welcome to SAGD

Sex-Associated Gene Database

Large amount of sex-associated genes (SAG), which have varied expression between females and males, are responsible for the remarkable sex differences in morphological, physiological and behavioral phenotypes of animals. SAGD systematically incorporating all public transcriptomic datasets with paired female and male samples, as well as related publications to analyze SAG. It contains 91,836 SAGs in 21 species by curating data from 2,828 samples of 38 projects.

In SAGD, users can:

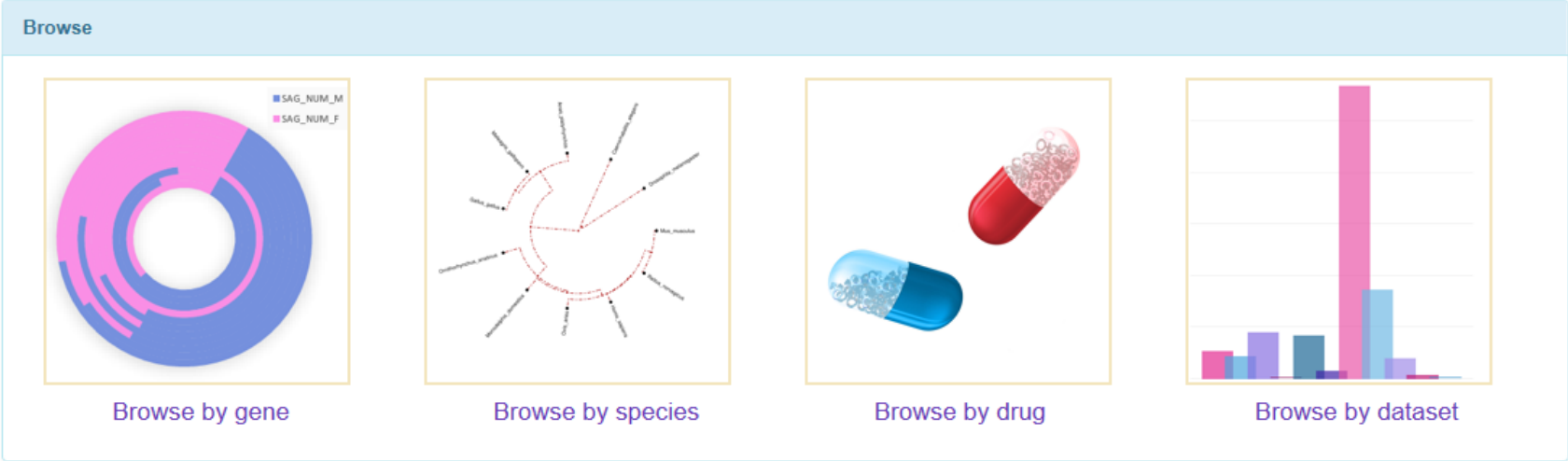
- Browse or search genes in different species and filter by their sex bias and significance in certain development stage and tissue;
- Browse or search drugs and find out the sex bias of their target genes;
- Browse or search high-throughput transcriptomic data with high quality paired female and male samples.

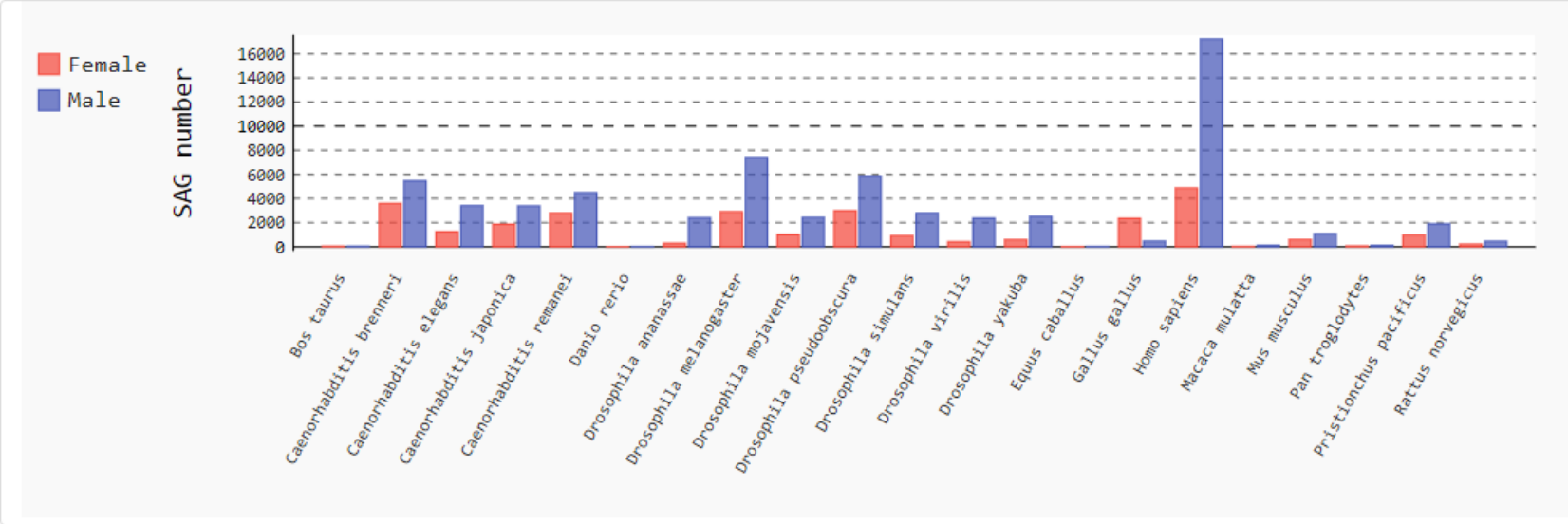
Citation:

SAGD: a comprehensive sex-associated gene database from transcriptomes

Meng-Wei Shi, Na-An Zhang, Chuan-Ping Shi, Chun-Jie Liu, Zhi-Hui Luo, Dan-Yang Wang, An-Yuan Guo*, Zhen-Xia Chen*

Nucleic Acids Research 2019; https://doi.org/10.1093/nar/gky1040

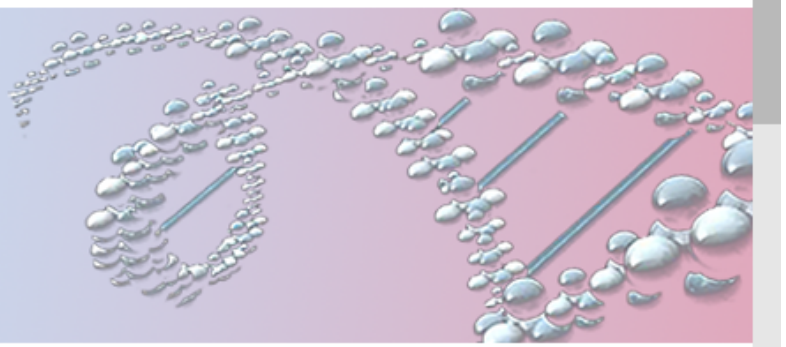




Website：http://bioinfo.life.hust.edu.cn/SAGD#!/

SAGD

Sex-Associated Gene Database



Drosophila melanogaster(Fruitfly)

SAGD Group	Project	Tissue	Stage	SAG number	Male-biased number	Female-biased number
SAGD_00134	SRP045429	whole body	adult	6046	5046	1000
SAGD_00161	SRP069203	Adult whole organisms without gonad	adult	1921	1725	196
SAGD_00183	SRP108530	abdomen without digestive or reproductive system	adult	322	131	191
SAGD_00184	SRP108530	digestive plus excretory system	adult	109	59	50
SAGD_00185	SRP108530	genitalia	adult	833	563	270
SAGD_00186	SRP108530	gonad	adult	7384	5987	1397
SAGD_00187	SRP108530	head	adult	89	36	53
SAGD_00188	SRP108530	reproductive system without gonad and genitalia	adult	2168	1216	952
SAGD_00189	SRP108530	thorax without digestive system	adult	127	41	86

clusterProfiler

platforms all | rank 52 / 1649 | posts 13 / 0.6 / 1 / 0 | in Bioc 8 years
build warnings | updated since release

DOI: 10.18129/B9.bioc.clusterProfiler

statistical analysis and visualization of functional profiles for genes and gene clusters

Bioconductor version: Release (3.8)

This package implements methods to analyze and visualize functional profiles (GO and KEGG) of gene and gene clusters.

Author: Guangchuang Yu [aut, cre, cph] (<https://orcid.org/0000-0002-6485-8781>), Li-Gen Wang [ctb], Giovanni Dall'Olio [ctb] (formula interface of compareCluster)

Maintainer: Guangchuang Yu <guangchuangyu at gmail.com>

Citation (from within R, enter `citation("clusterProfiler")`):

- 加载包

```
source("http://bioconductor.org/biocLite.R")
biocLite('clusterProfiler')
library("clusterProfiler")
#依赖于 bioconductor中心的 OrgDb 系列包
#http://www.bioconductor.org/packages/release/BiocViews.html#___OrgDb
biocLite("org.Dm.eg.db")
library(org.Dm.eg.db)
keytypes(org.Dm.eg.db)
biocLite("DOSE")
library("DOSE")
```

加载并处理数据

	X	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj	FPKM_M	FPKM_F
12026	FBgn0034350	0.00000000	NA	NA	NA	NA	NA	0.000000000	0.000000000
12027	FBgn0034363	1.03015288	-0.859551102	1.3037012	-0.659316010	0.50969286	NA	0.044386181	0.076567401
12028	FBgn0034364	0.00000000	NA	NA	NA	NA	NA	0.000000000	0.000000000
12029	FBgn0034372	0.00000000	NA	NA	NA	NA	NA	0.000000000	0.000000000
12030	FBgn0034377	1.34085698	-0.326891741	1.0478526	-0.311963482	0.75506828	NA	0.215662013	0.244506950

```
DEG_fruitfly<-read.csv("/home/wangdy/course/head_fruity.csv")
DEG_fruitfly<-na.omit(DEG_fruitfly)
ncol=ncol(DEG_fruitfly)
nrow=nrow(DEG_fruitfly)
DEG_fruitfly$sexbias <- as.factor(ifelse(DEG_fruitfly$padj<0.05&abs(DEG_fruitfly$log2FoldChange)>1,ifelse(DEG_fruitfly$log2FoldChange>1,'male','female'),'unbias'))
```

功能富集

	X	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj	FPKM_M	FPKM_F	sexbias
123	FBgn0004959	39.686765	1.8015788	0.23216003	7.760073	8.488046e-15	7.443947e-13	3.7683924	9.946213e-01	male
124	FBgn0037137	350.710132	-0.7085992	0.09157118	-7.738234	1.008073e-14	8.769422e-13	11.0046346	1.680423e+01	unbias
125	FBgn0014031	1067.770933	-1.4028960	0.18192007	-7.711606	1.242438e-14	1.072174e-12	55.4578763	1.372488e+02	female
126	FBgn0030616	7940.747693	-0.7099687	0.09233256	-7.689256	1.479936e-14	1.266990e-12	620.4873507	9.497304e+02	unbias

```
transID<-bitr(DEG_fruitfly$X, fromType="ENSEMBL", toType="ENTREZID", OrgDb="org.Dm.eg.db")
go_table<-merge(DEG_fruitfly,transID,by.x="X",by.y="ENSEMBL")
male<-subset(go_table,DEG_fruitfly$sexbias=="male")
female<-subset(go_table,DEG_fruitfly$sexbias=="female")

go_female <- enrichGO(female$ENTREZID, OrgDb = org.Dm.eg.db, ont='ALL',pAdjustMethod = 'BH',pvalueCutoff = 0.05,
                      qvalueCutoff = 0.2,keyType = 'ENTREZID')
go_male <- enrichGO(male$ENTREZID, OrgDb = org.Dm.eg.db, ont='ALL',pAdjustMethod = 'BH',pvalueCutoff = 0.05,
                      qvalueCutoff = 0.2,keyType = 'ENTREZID')
```

查看结果

```
> dim(go_male)
[1] 199  10
> dim(go_male[go_male$ONTOLOGY=='BP',])
[1] 158  10
> dim(go_male[go_male$ONTOLOGY=='CC',])
[1] 28 10
> dim(go_male[go_male$ONTOLOGY=='MF',])
[1] 13 10
```

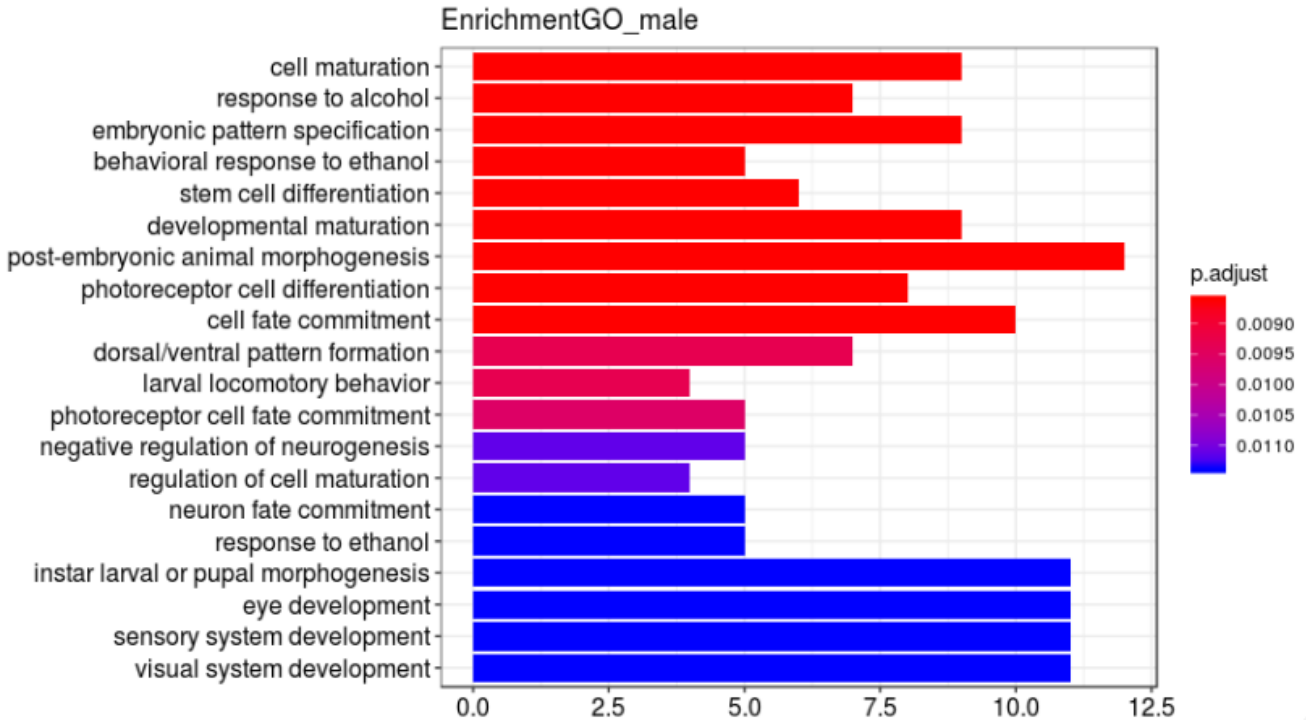
```
> head(go_male,2)
           ONTOLOGY         ID         Description GeneRatio   BgRatio       pvalue    p.adjust      qvalue                                                geneID Count
GO:0048469       BP GO:0048469     cell maturation      9/80 211/12621 7.140797e-06 0.008614196 0.006586697 42313/35540/37757/38389/36810/43604/33974/43901/42379     9
GO:0097305       BP GO:0097305 response to alcohol      7/80 131/12621 1.870644e-05 0.008614196 0.006586697                 44811/35540/33281/31718/36699/37734/34176     7
```

- enrichGO参数

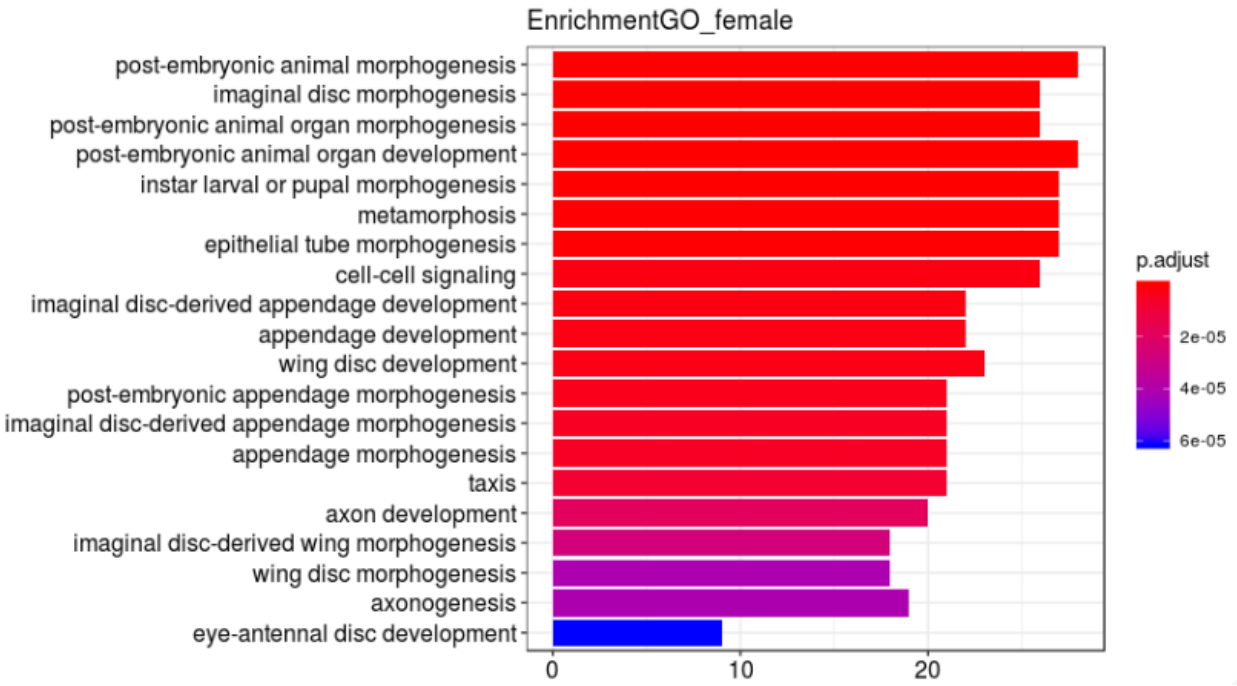
```
enrichGO     GO Enrichment Analysis of a gene set. Given a vector of genes, this
function will return the enrichment GO categories after FDR control.
Usage:
  enrichGO(gene, OrgDb, keyType = "ENTREZID", ont = "MF", pvalueCutoff = 0.05,
           pAdjustMethod = "BH", universe, qvalueCutoff = 0.2, minGSSize = 10,
           maxGSSize = 500, readable = FALSE, pool = FALSE)
Arguments:
  gene                  a vector of entrez gene id.
OrgDb                 OrgDb
keyType               keytype of input gene
ont                   One of "MF", "BP", and "CC" subontologies or 'ALL'.
pvalueCutoff          Cutoff value of pvalue.
pAdjustMethod         one of "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"
universe              background genes
qvalueCutoff          qvalue cutoff
minGSSize             minimal size of genes annotated by Ontology term for testing.
maxGSSize             maximal size of genes annotated for testing
readable              whether mapping gene ID to gene Name
pool                  If ont='ALL', whether pool 3 GO sub-ontologies
```

作图

```
barplot(go_female,showCategory=20,drop=T)
barplot(go_male,showCategory=20,drop=T)
dotplot(go_female,showCategory=30)
dotplot(go_male,showCategory=30)
```

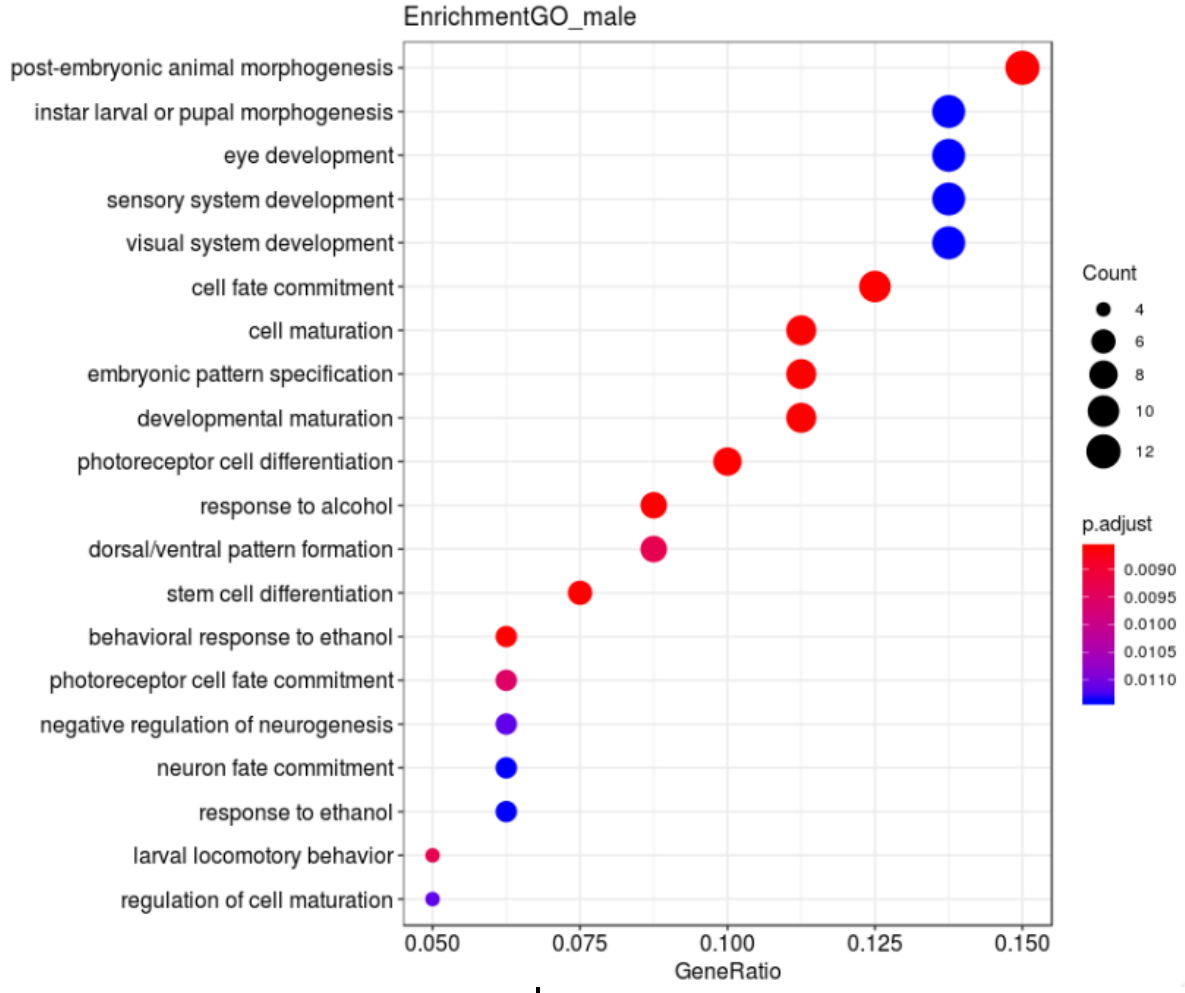


male

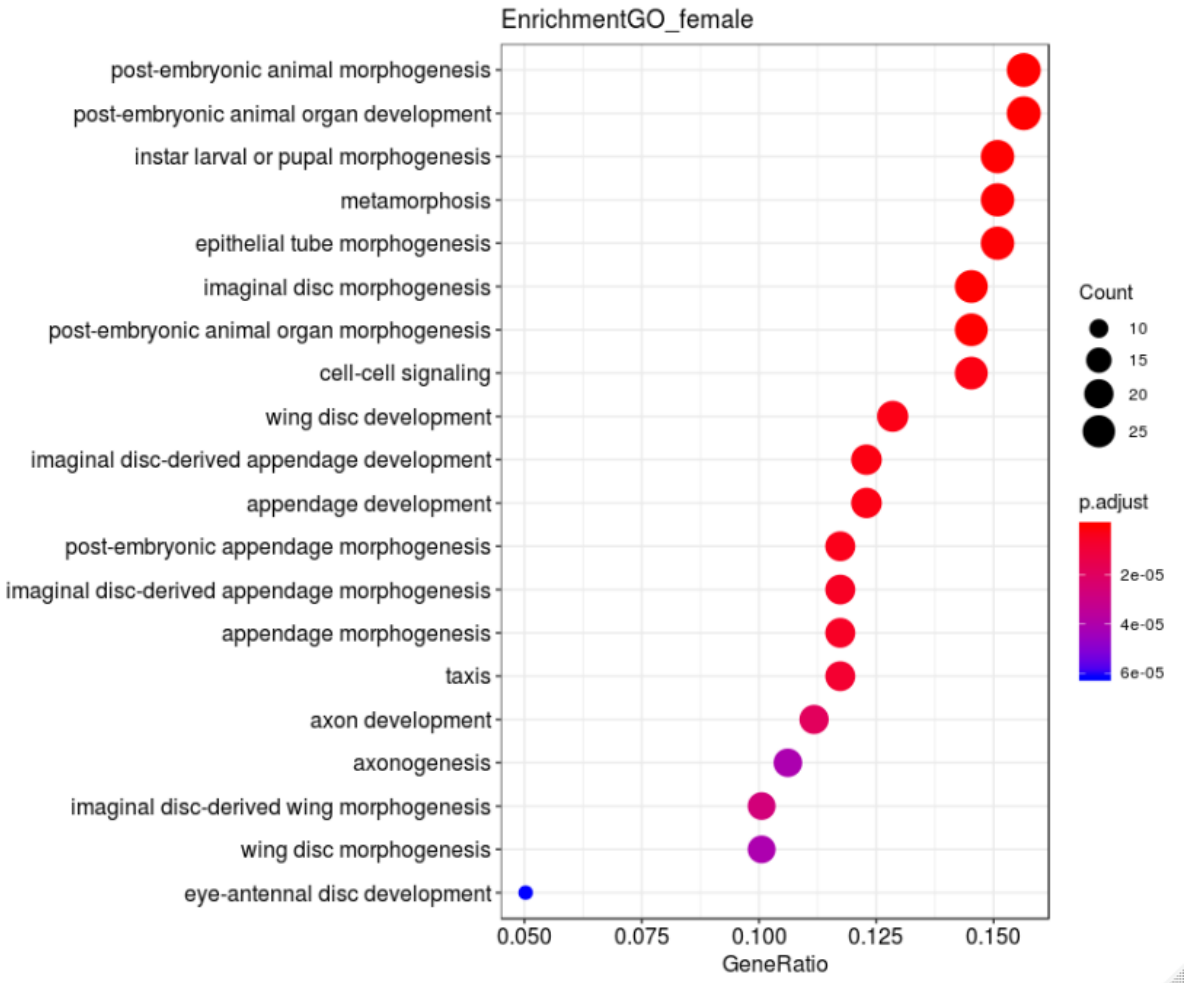


female

```
barplot(go_female,showCategory=20,drop=T)
barplot(go_male,showCategory=20,drop=T)
dotplot(go_female,showCategory=30)
dotplot(go_male,showCategory=30)
```

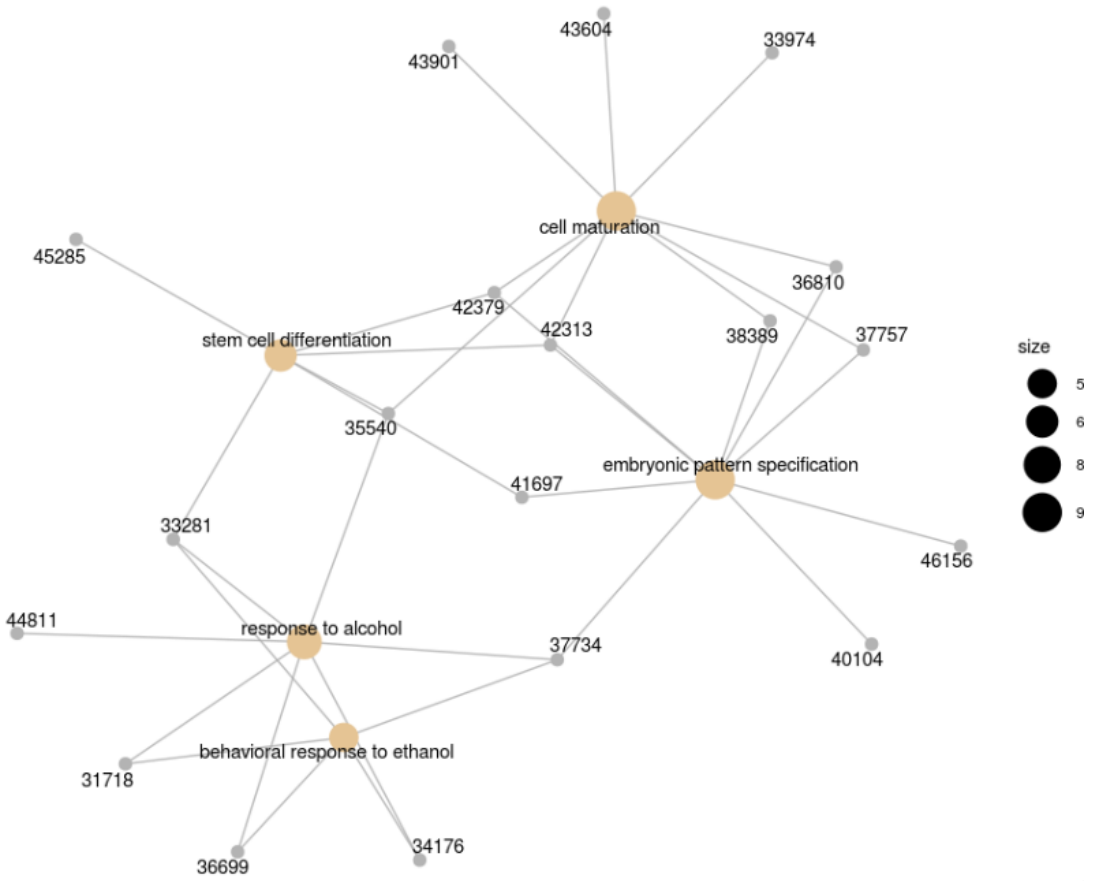


male

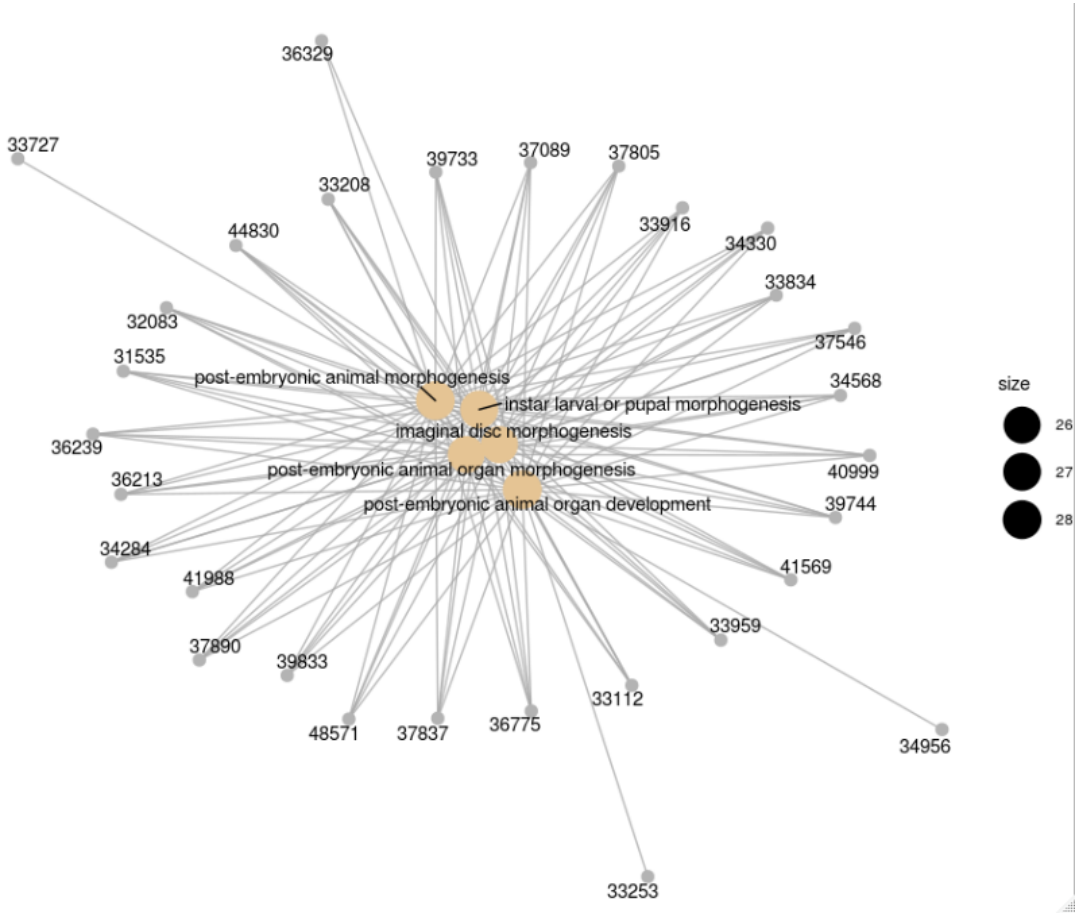


female

```
cnetplot(go_male)
cnetplot(go_female)
```



male



female