

# 代谢组 Kegg 富集分析

阎俊安

## 目录

1 第一章数据下载	1
1.1 KEGG 代谢通路数据下载	1
1.2 富集分析	1

## 1 第一章数据下载

### 1.1 KEGG 代谢通路数据下载

```
wget -c http://rest.kegg.jp/list/pathway

cat pathway |cut -d ":" -f 2|awk '{print $1}'|\
while read id; do wget -c http://rest.kegg.jp/link/compound/$id; done

cat map* > map.txt
```

### 1.2 富集分析

做富集分析总共需要三个文件，kegg 数据库注释文件，代谢通路背景文件，差异代谢物编号

```
library(tidyverse)
library(magrittr)

# kegg 数据库文件
keggannotation <- read_tsv("pathway",col_names = F) %>%
  left_join(.,read_tsv('map.txt',col_names = F),by="X1") %>%
  select(-1) %>% set_colnames(c("pathway","ID")) %>%
  mutate(across("ID",str_replace,"cpd:","")) %>% select(2,1) %>%
  arrange(ID)
```

```
# 背景文件
```

```
allkeggid <- read_tsv("meta_intensity_neg.anno.xls") %>% select(KEGG) %>%  
  bind_rows(.,read_tsv("meta_intensity_pos.anno.xls") %>% select(KEGG)) %>%  
  arrange() %>% filter(KEGG != "_") %>% set_colnames(c("ID"))
```

```
# 差异代谢物
```

```
diffkeggID <- read_tsv("diff.xls") %>% select(KEGG) %>%  
  arrange() %>% filter(KEGG != "_") %>% set_colnames(c("ID"))
```

```
total <- right_join(keggannotation,allkeggid,by="ID") %>% select(2,1)
```

```
# 富集分析
```

```
x <- clusterProfiler::enricher(gene = diffkeggID$ID,TERM2GENE = total,  
                               minGSSize = 1,pvalueCutoff = 1,qvalueCutoff = 1)
```

```
# 导出文件
```

```
write.csv(as.data.frame(x@result) %>% select(-1,-2),file="KEGG_enrichment_result.csv")
```