

main

August 16, 2021

1 Functional enrichment analysis with g:Profiler

```
[1]: library(tidyverse)
library(gprofiler2)
```

Attaching packages tidyverse
1.3.1

ggplot2	3.3.5	purrr	0.3.4
tibble	3.1.2	dplyr	1.0.7
tidyr	1.1.3	stringr	1.4.0
readr	1.4.0	forcats	0.5.1

Conflicts
tidyverse_conflicts()
dplyr::filter() masks stats::filter()
dplyr::lag() masks stats::lag()

1.1 Load DEG results

```
[2]: deg <- data.table::fread('../..../metrics_summary/_m/male_specific_DE_4features.
  ↳txt') %>%
  select(gencodeID, ensemblID, Symbol, logFC, "adj.P.Val")
deg %>% head(2)
```

	gencodeID	ensemblID	Symbol	logFC	adj.P.Val
	<chr>	<chr>	<chr>	<dbl>	<dbl>
A data.table: 2 × 5	ENSG00000157303.10	ENSG00000157303	SUSD3	-0.6828182	0.0006471114
	ENSG00000166428.12	ENSG00000166428	PLD4	-0.6095456	0.0014670510

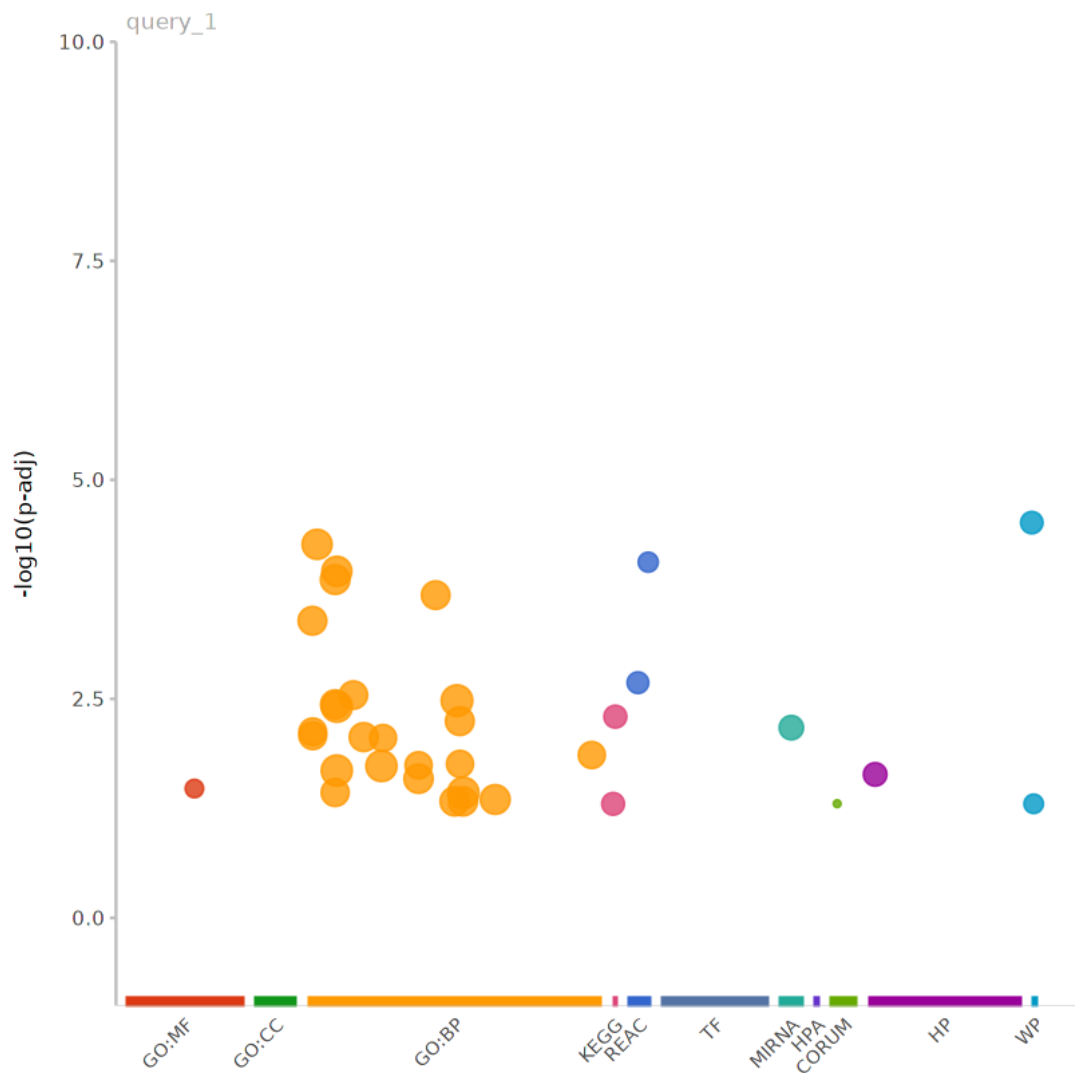
1.2 Calculated enrichment and visual plot

```
[3]: save_ggplots <- function(fn, p, w, h){
  for(ext in c('.pdf', '.png', '.svg')){
    ggsave(paste0(fn, ext), plot=p, width=w, height=h)
  }
}
```

```
[4]: gostres <- gost(query=deg$ensemblID, organism="hsapiens",
                    correction_method="g_SCS", significant=TRUE)
gostres$result %>%
  data.table::fwrite(file = "DEGs_functional_enrichment.tsv", sep="\t")
gostres$result %>% head(2)
```

	query <chr>	significant <lg1>	p_value <dbl>	term_size <int>	query_size <int>	intersection_size <int>	prec <dbl>
A data.frame: 2 × 14	1 query_1	TRUE	4.97118e-02	3	18	2	0.11
	2 query_1	TRUE	5.44763e-05	3438	90	40	0.44

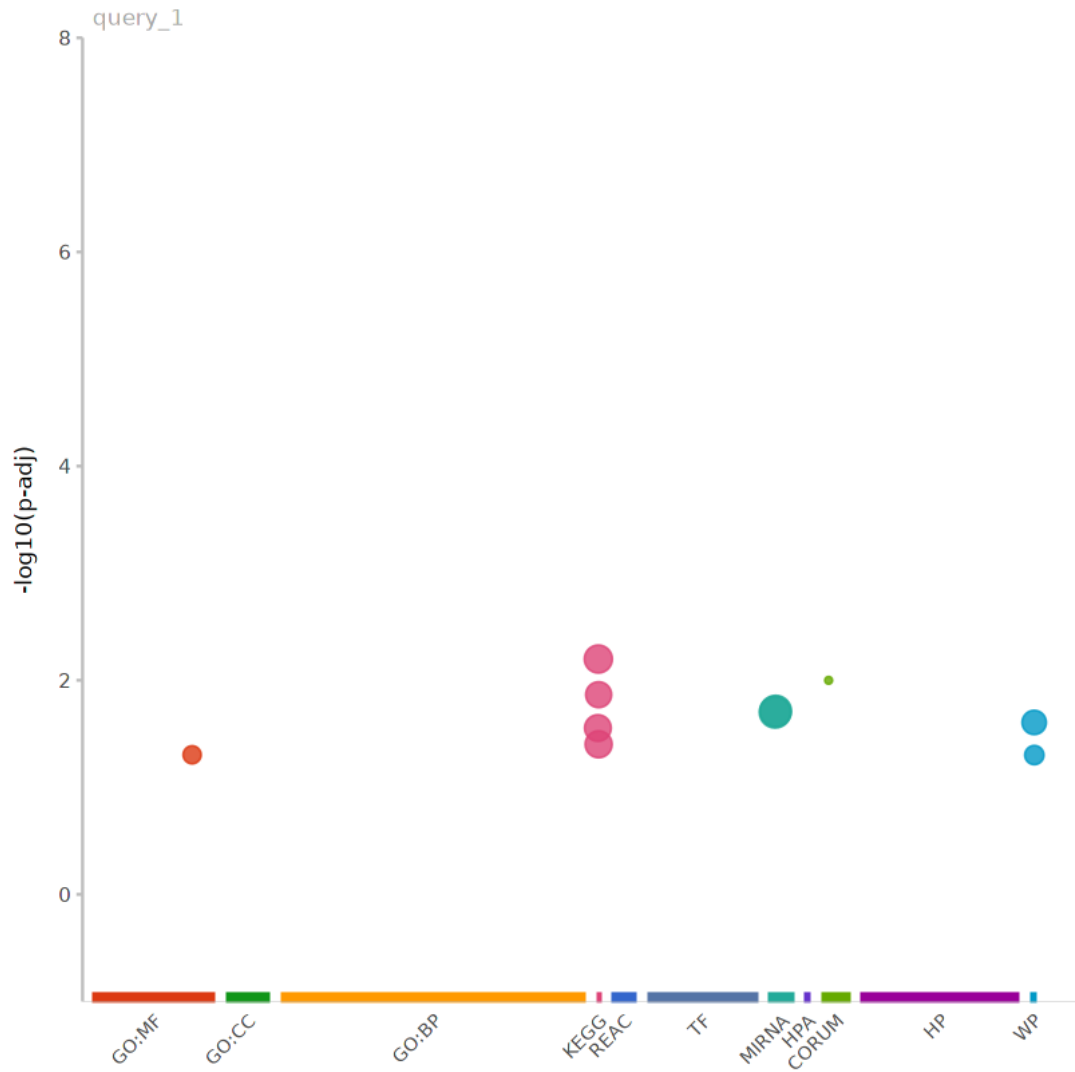
```
[5]: p <- gostplot(gostres, capped = FALSE, interactive = FALSE)
print(p)
save_ggplots("DEGs_manhattan", p, 9, 5)
```



```
[6]: deg_sz = deg %>% filter(logFC > 0)
      gostres <- gost(query=deg_sz$ensemblID, organism="hsapiens",
                      correction_method="g_SCS", significant=TRUE)
      gostres$result %>%
        data.table::fwrite(file = "upreg_DEGs_functional_enrichment.tsv", sep="\t")
      gostres$result %>% head(5)
```

		query	significant	p_value	term_size	query_size	intersection_size	pre
		<chr>	<lgl>	<dbl>	<int>	<int>	<int>	<dbl>
A data.frame: 5 × 14	1	query_1	TRUE	0.010018418	3	12	2	0.16
	2	query_1	TRUE	0.049673291	8	40	2	0.05
	3	query_1	TRUE	0.006333676	131	16	4	0.25
	4	query_1	TRUE	0.013638446	62	16	3	0.18
	5	query_1	TRUE	0.027931048	79	16	3	0.18

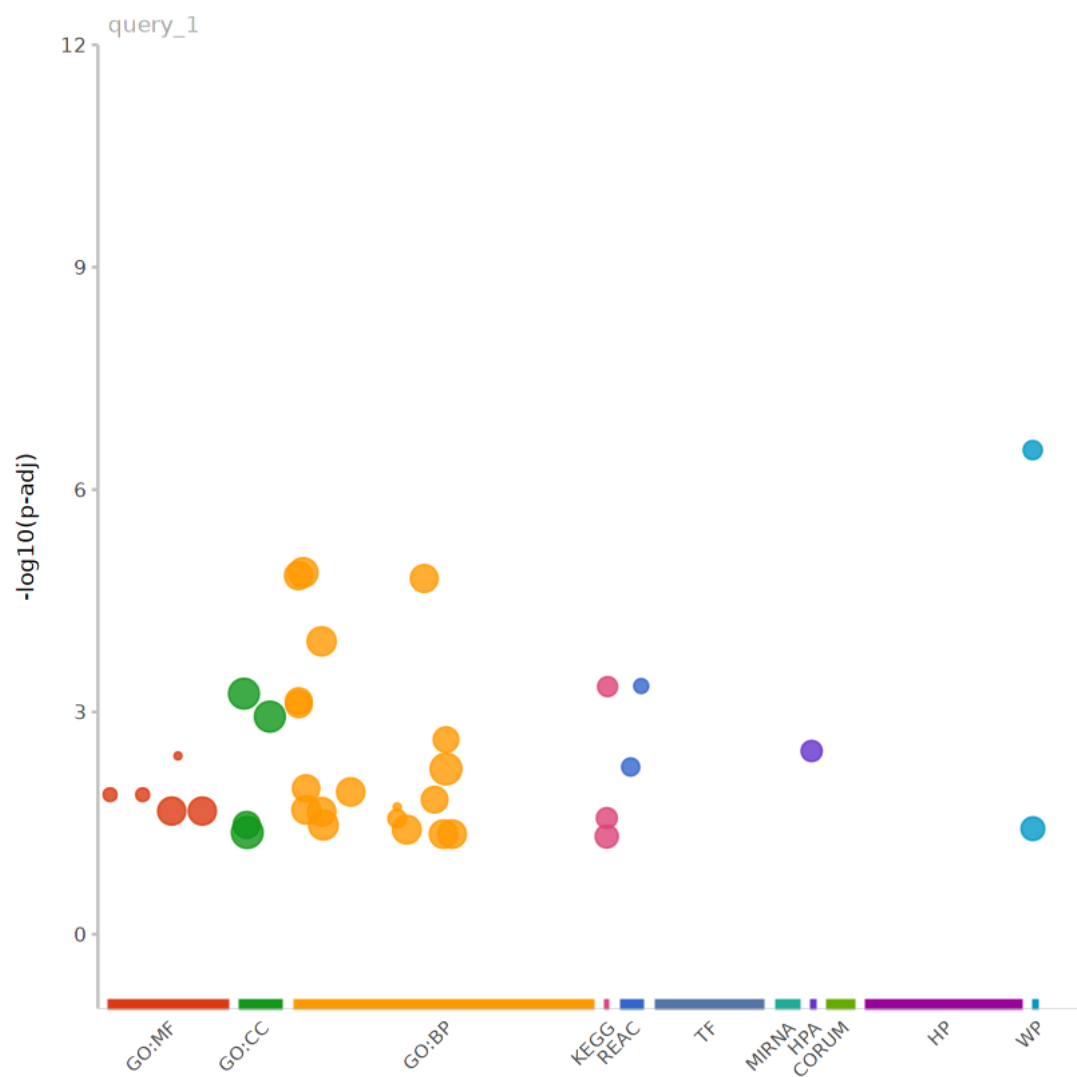
```
[7]: p <- gostplot(gostres, capped = FALSE, interactive = FALSE)
      print(p)
      save_ggplots("upreg_DEGs_manhattan", p, 9, 5)
```



```
[8]: deg_ctl = deg %>% filter(logFC < 0)
gostres <- gost(query=deg_ctl$ensemblID, organism="hsapiens",
               correction_method="g_SCS", significant=TRUE)
gostres$result %>%
  data.table::fwrite(file = "downreg_DEGs_functional_enrichment.tsv",
    ↪sep="\t")
gostres$result %>% head(2)
```

A data.frame: 2 × 14							
	query <chr>	significant <lgl>	p_value <dbl>	term_size <int>	query_size <int>	intersection_size <int>	pre <d
1	query_1	TRUE	1.315792e-05	3438	50	28	0.5
2	query_1	TRUE	1.449021e-05	1507	50	19	0.3

```
[9]: p <- gostplot(gostres, capped = FALSE, interactive = FALSE)
print(p)
save_ggplots("downreg_DEGs_manhattan", p, 9, 5)
```



1.3 Reproducibility Information

```
[10]: Sys.time()
proc.time()
options(width = 120)
sessioninfo::session_info()
```

```
[1] "2021-08-16 19:20:43 EDT"
```

```
user system elapsed
8.688 0.294 14.980
```

Session info

```
setting value
version R version 4.0.3 (2020-10-10)
os      Arch Linux
system  x86_64, linux-gnu
ui      X11
language (EN)
collate en_US.UTF-8
ctype   en_US.UTF-8
tz      America/New_York
date    2021-08-16
```

Packages

package	* version	date	lib	source
assertthat	0.2.1	2019-03-21	[1]	CRAN (R 4.0.2)
backports	1.2.1	2020-12-09	[1]	CRAN (R 4.0.2)
base64enc	0.1-3	2015-07-28	[1]	CRAN (R 4.0.2)
bitops	1.0-7	2021-04-24	[1]	CRAN (R 4.0.3)
broom	0.7.8	2021-06-24	[1]	CRAN (R 4.0.3)
Cairo	1.5-12.2	2020-07-07	[1]	CRAN (R 4.0.2)
cellranger	1.1.0	2016-07-27	[1]	CRAN (R 4.0.2)
cli	3.0.0	2021-06-30	[1]	CRAN (R 4.0.3)
colorspace	2.0-2	2021-06-24	[1]	CRAN (R 4.0.3)
crayon	1.4.1	2021-02-08	[1]	CRAN (R 4.0.3)
data.table	1.14.0	2021-02-21	[1]	CRAN (R 4.0.3)
DBI	1.1.1	2021-01-15	[1]	CRAN (R 4.0.2)
dbplyr	2.1.1	2021-04-06	[1]	CRAN (R 4.0.3)
digest	0.6.27	2020-10-24	[1]	CRAN (R 4.0.2)
dplyr	* 1.0.7	2021-06-18	[1]	CRAN (R 4.0.3)
ellipsis	0.3.2	2021-04-29	[1]	CRAN (R 4.0.3)
evaluate	0.14	2019-05-28	[1]	CRAN (R 4.0.2)
fansi	0.5.0	2021-05-25	[1]	CRAN (R 4.0.3)
farver	2.1.0	2021-02-28	[1]	CRAN (R 4.0.3)
forcats	* 0.5.1	2021-01-27	[1]	CRAN (R 4.0.2)
fs	1.5.0	2020-07-31	[1]	CRAN (R 4.0.2)
generics	0.1.0	2020-10-31	[1]	CRAN (R 4.0.2)
ggplot2	* 3.3.5	2021-06-25	[1]	CRAN (R 4.0.3)
glue	1.4.2	2020-08-27	[1]	CRAN (R 4.0.2)
gprofiler2	* 0.2.0	2020-08-27	[1]	CRAN (R 4.0.3)
gtable	0.3.0	2019-03-25	[1]	CRAN (R 4.0.2)
haven	2.4.1	2021-04-23	[1]	CRAN (R 4.0.3)
hms	1.1.0	2021-05-17	[1]	CRAN (R 4.0.3)
htmltools	0.5.1.1	2021-01-22	[1]	CRAN (R 4.0.2)
htmlwidgets	1.5.3	2020-12-10	[1]	CRAN (R 4.0.2)
httr	1.4.2	2020-07-20	[1]	CRAN (R 4.0.2)

IRdisplay	1.0	2021-01-20	[1]	CRAN	(R 4.0.2)
IRkernel	1.2	2021-05-11	[1]	CRAN	(R 4.0.3)
jsonlite	1.7.2	2020-12-09	[1]	CRAN	(R 4.0.2)
labeling	0.4.2	2020-10-20	[1]	CRAN	(R 4.0.2)
lazyeval	0.2.2	2019-03-15	[1]	CRAN	(R 4.0.2)
lifecycle	1.0.0	2021-02-15	[1]	CRAN	(R 4.0.3)
lubridate	1.7.10	2021-02-26	[1]	CRAN	(R 4.0.3)
magrittr	2.0.1	2020-11-17	[1]	CRAN	(R 4.0.2)
modelr	0.1.8	2020-05-19	[1]	CRAN	(R 4.0.2)
munsell	0.5.0	2018-06-12	[1]	CRAN	(R 4.0.2)
pbdZMQ	0.3-5	2021-02-10	[1]	CRAN	(R 4.0.3)
pillar	1.6.1	2021-05-16	[1]	CRAN	(R 4.0.3)
pkgconfig	2.0.3	2019-09-22	[1]	CRAN	(R 4.0.2)
plotly	4.9.4.1	2021-06-18	[1]	CRAN	(R 4.0.3)
purrr	* 0.3.4	2020-04-17	[1]	CRAN	(R 4.0.2)
R6	2.5.0	2020-10-28	[1]	CRAN	(R 4.0.2)
Rcpp	1.0.7	2021-07-07	[1]	CRAN	(R 4.0.3)
RCurl	1.98-1.3	2021-03-16	[1]	CRAN	(R 4.0.3)
readr	* 1.4.0	2020-10-05	[1]	CRAN	(R 4.0.2)
readxl	1.3.1	2019-03-13	[1]	CRAN	(R 4.0.2)
repr	1.1.3	2021-01-21	[1]	CRAN	(R 4.0.2)
reprex	2.0.0	2021-04-02	[1]	CRAN	(R 4.0.3)
rlang	0.4.11	2021-04-30	[1]	CRAN	(R 4.0.3)
rstudioapi	0.13	2020-11-12	[1]	CRAN	(R 4.0.2)
rvest	1.0.0	2021-03-09	[1]	CRAN	(R 4.0.3)
scales	1.1.1	2020-05-11	[1]	CRAN	(R 4.0.2)
sessioninfo	1.1.1	2018-11-05	[1]	CRAN	(R 4.0.2)
stringi	1.7.3	2021-07-16	[1]	CRAN	(R 4.0.3)
stringr	* 1.4.0	2019-02-10	[1]	CRAN	(R 4.0.2)
svglite	2.0.0	2021-02-20	[1]	CRAN	(R 4.0.3)
systemfonts	1.0.2	2021-05-11	[1]	CRAN	(R 4.0.3)
tibble	* 3.1.2	2021-05-16	[1]	CRAN	(R 4.0.3)
tidyr	* 1.1.3	2021-03-03	[1]	CRAN	(R 4.0.3)
tidyselect	1.1.1	2021-04-30	[1]	CRAN	(R 4.0.3)
tidyverse	* 1.3.1	2021-04-15	[1]	CRAN	(R 4.0.3)
utf8	1.2.1	2021-03-12	[1]	CRAN	(R 4.0.3)
uuid	0.1-4	2020-02-26	[1]	CRAN	(R 4.0.2)
vctrs	0.3.8	2021-04-29	[1]	CRAN	(R 4.0.3)
viridisLite	0.4.0	2021-04-13	[1]	CRAN	(R 4.0.3)
withr	2.4.2	2021-04-18	[1]	CRAN	(R 4.0.3)
xml2	1.3.2	2020-04-23	[1]	CRAN	(R 4.0.2)

[1] /home/jbenja13/R/x86_64-pc-linux-gnu-library/4.0

[2] /usr/lib/R/library