

BB512/BB612 - Week VIII

ORA

```
suppressPackageStartupMessages(library(enrichR))

selected_genes <- c("PPARA", "FABP4", "PLTP", "PPARG", "RXRA", "ATRX", "TP53", "KRAS", "MYC")

dbs <- listEnrichrDbs()
head(dbs)
```

	geneCoverage	genesPerTerm	libraryName	link	numTerms
## 1	13362	275	Genome_Browser_PWMs		615
## 2	27884	1284	TRANSFAC_and_JASPAR_PWMs		326
## 3	6002	77	Transcription_Factor_PPIs		290
## 4	47172	1370	ChEA_2013		353
## 5	47107	509	Drug_Perturbations_from_GEO_2014		701
## 6	21493	3713	ENCODE_TF_ChIP-seq_2014		498

```
##
## 1 http://hgdownload.cse.ucsc.edu/goldenPath/hg18/database/
## 2 http://jaspar.genereg.net/html/DOWNLOAD/
## 3
## 4 http://amp.pharm.mssm.edu/lib/cheadownload.jsp
## 5 http://www.ncbi.nlm.nih.gov/geo/
## 6 http://genome.ucsc.edu/ENCODE/downloads.html
##
## appyter categoryId
## 1 ea115789fcbf12797fd692cec6df0ab4dbc79c6a 1
## 2 7d42eb43a64a4e3b20d721fc7148f685b53b6b30 1
## 3 849f222220618e2599d925b6b51868cf1dab3763 1
## 4 7ebe772afb55b63b41b79dd8d06ea0fdd9fa2630 7
## 5 ad270a6876534b7cb063e004289dcd4d3164f342 7
## 6 497787ebc418d308045efb63b8586f10c526af51 7
```

```
dbs[grepl("KEGG", dbs$libraryName), ]
```

	geneCoverage	genesPerTerm	libraryName	link	numTerms	appyter	categoryId
## 11	4128	48	KEGG_2013	http://www.kegg.jp/kegg/download/			
## 60	3800	48	KEGG_2015	http://www.kegg.jp/kegg/download/			
## 94	7010	87	KEGG_2016	http://www.kegg.jp/kegg/download/			
## 148	7802	92	KEGG_2019_Human	https://www.kegg.jp/			
## 149	8551	98	KEGG_2019_Mouse	https://www.kegg.jp/			
## 173	8078	102	KEGG_2021_Human	https://www.kegg.jp/			

```
##
## numTerms
## 11 200 eb26f55d3904cb0ea471998b6a932a9bf65d8e50 7
## 60 179 e80d25c56de53c704791ddfdc6ab5eec28ae7243 7
## 94 293 43f56da7540195ba3c94eb6e34c522a699b36da9 7
## 148 308 3477bc578c4ea5d851dcb934fe2a41e9fd789bb4 7
## 149 303 187eb44b2d6fa154ebf628eba1f18537f64e797c 7
## 173 320 2
```

```
dbs[grepl("GO", dbs$libraryName), ]
```

##	geneCoverage	genesPerTerm	libraryName	
## 15	12753	57	GO_Molecular_Function_2015	
## 20	13236	82	GO_Cellular_Component_2015	
## 21	14264	58	GO_Biological_Process_2015	
## 46	7682	78	GO_Biological_Process_2013	
## 47	7324	172	GO_Cellular_Component_2013	
## 48	8469	122	GO_Molecular_Function_2013	
## 112	10427	38	GO_Cellular_Component_2017	
## 113	10601	25	GO_Molecular_Function_2017	
## 114	13822	21	GO_Biological_Process_2017	
## 115	8002	143	GO_Cellular_Component_2017b	
## 116	10089	45	GO_Molecular_Function_2017b	
## 117	13247	49	GO_Biological_Process_2017b	
## 130	14433	36	GO_Biological_Process_2018	
## 131	8655	61	GO_Cellular_Component_2018	
## 132	11459	39	GO_Molecular_Function_2018	
## 176	14937	33	GO_Biological_Process_2021	
## 177	11497	80	GO_Cellular_Component_2021	
## 178	11936	34	GO_Molecular_Function_2021	
##			link	numTerms
## 15	http://www.geneontology.org/GO.downloads.annotations.shtml			1136
## 20	http://www.geneontology.org/GO.downloads.annotations.shtml			641
## 21	http://www.geneontology.org/GO.downloads.annotations.shtml			5192
## 46	http://www.geneontology.org/GO.downloads.annotations.shtml			941
## 47	http://www.geneontology.org/GO.downloads.annotations.shtml			205
## 48	http://www.geneontology.org/GO.downloads.annotations.shtml			402
## 112	http://www.geneontology.org/			636
## 113	http://www.geneontology.org/			972
## 114	http://www.geneontology.org/			3166
## 115	http://www.geneontology.org/			816
## 116	http://www.geneontology.org/			3271
## 117	http://www.geneontology.org/			10125
## 130	http://www.geneontology.org/			5103
## 131	http://www.geneontology.org/			446
## 132	http://www.geneontology.org/			1151
## 176	http://www.geneontology.org/			6036
## 177	http://www.geneontology.org/			511
## 178	http://www.geneontology.org/			1274
##	appyter	categoryId		
## 15	f531ac2b6acdf7587a54b79b465a5f4aab8f00f9	7		
## 20	e1d004d5797cbd2363ef54b1c3b361adb68795c6	7		
## 21	bf120b6e11242b1a64c80910d8e89f87e618e235	7		
## 46	5216d1ade194ffa5a6c00f105e2b1899f64f45fe	7		
## 47	fd1332a42395e0bc1dba82868b39be7983a48cc5	7		
## 48	7e3e99e5aae02437f80b0697b197113ce3209ab0	7		
## 112	8fed21d22dfcc3015c05b31d942fdcf851cc8e04	7		
## 113	b4018906e0a8b4e81a1b1afc51e0a2e7655403eb	7		
## 114	d9da4dba4a3eb84d4a28a3835c06dfbbe5811f92	7		
## 115	ecf39c41fa5bc7deb625a2b5761a708676e9db7c	7		
## 116	8d8340361dd36a458f1f0a401f1a3141de1f3200	7		
## 117	6404c38bffc2b3732de4e3fbe417b5043009fe34	7		
## 130	b5b77681c46ac58cd050e60bcd4ad5041a9ab0a9	7		

```
## 131 e9ebe46188efacbe1056d82987ff1c70218fa7ae      7
## 132 79ff80ae9a69dd00796e52569e41422466fa0bee      7
## 176                                              3
## 177                                              3
## 178                                              3
```

```
res <- enrichr(selected_genes, databases = c("KEGG_2021_Human", "GO_Biological_Process_2017"))
```

```
## Uploading data to Enrichr... Done.
## Querying KEGG_2021_Human... Done.
## Querying GO_Biological_Process_2017... Done.
## Parsing results... Done.
```

```
head(res$KEGG_2021_Human)
```

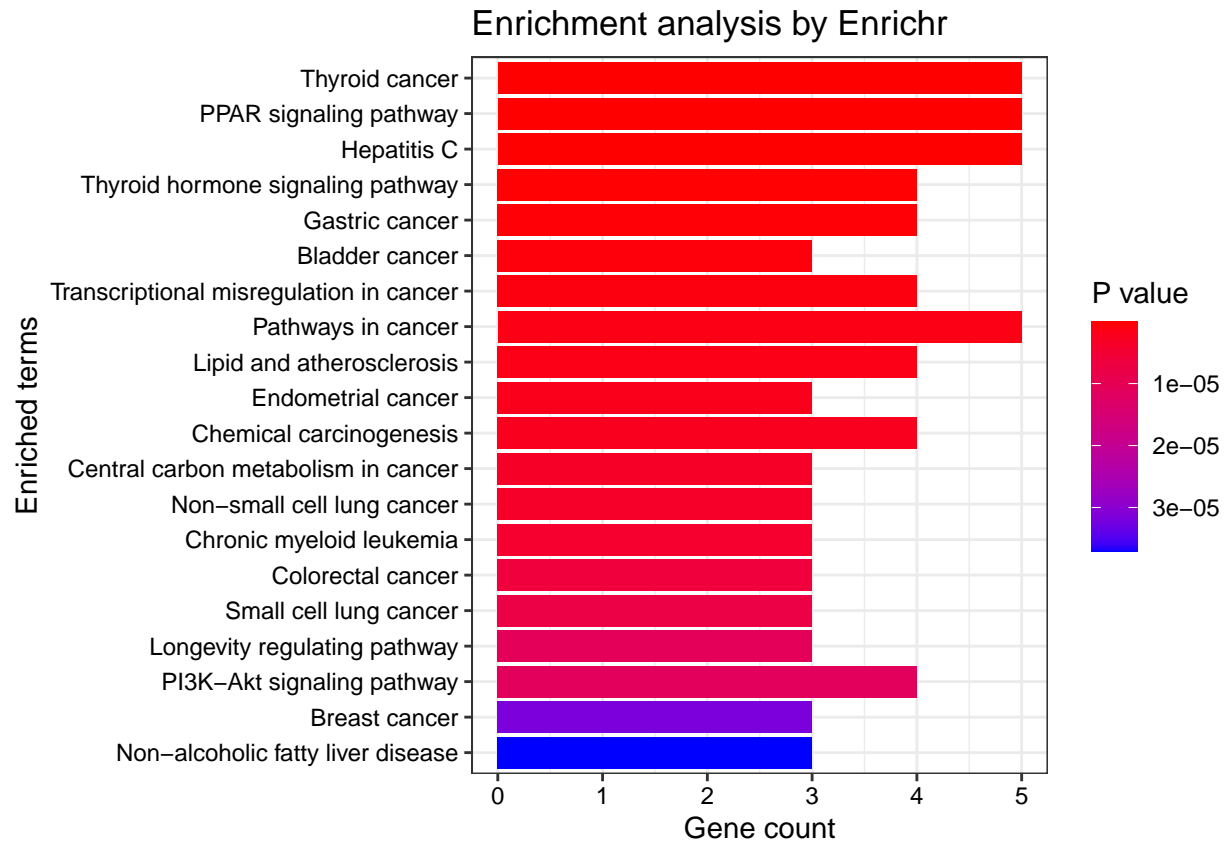
```
##              Term Overlap   P.value Adjusted.P.value
## 1          Thyroid cancer   5/37 2.0490e-12      2.2948e-10
## 2      PPAR signaling pathway   5/74 7.5260e-11      4.2146e-09
## 3          Hepatitis C     5/157 3.4347e-09      1.2823e-07
## 4 Thyroid hormone signaling pathway 4/121 1.5687e-07      4.3924e-06
## 5          Gastric cancer   4/149 3.6210e-07      8.1110e-06
## 6          Bladder cancer    3/41 6.6589e-07      1.2430e-05
## Old.P.value Old.Adjusted.P.value Odds.Ratio Combined.Score
## 1           0           0       779.65      20983.2
## 2           0           0       360.91      8412.7
## 3           0           0       163.15      3179.7
## 4           0           0       135.89      2129.1
## 5           0           0       109.50      1624.0
## 6           0           0       262.54      3733.9
##              Genes
## 1  RXRA;MYC;PPARG;KRAS;TP53
## 2 FABP4;RXRA;PPARG;PPARA;PLTP
## 3  RXRA;MYC;KRAS;PPARA;TP53
## 4  RXRA;MYC;KRAS;TP53
## 5  RXRA;MYC;KRAS;TP53
## 6    MYC;KRAS;TP53
```

```
head(res$GO_Biological_Process_2017)
```

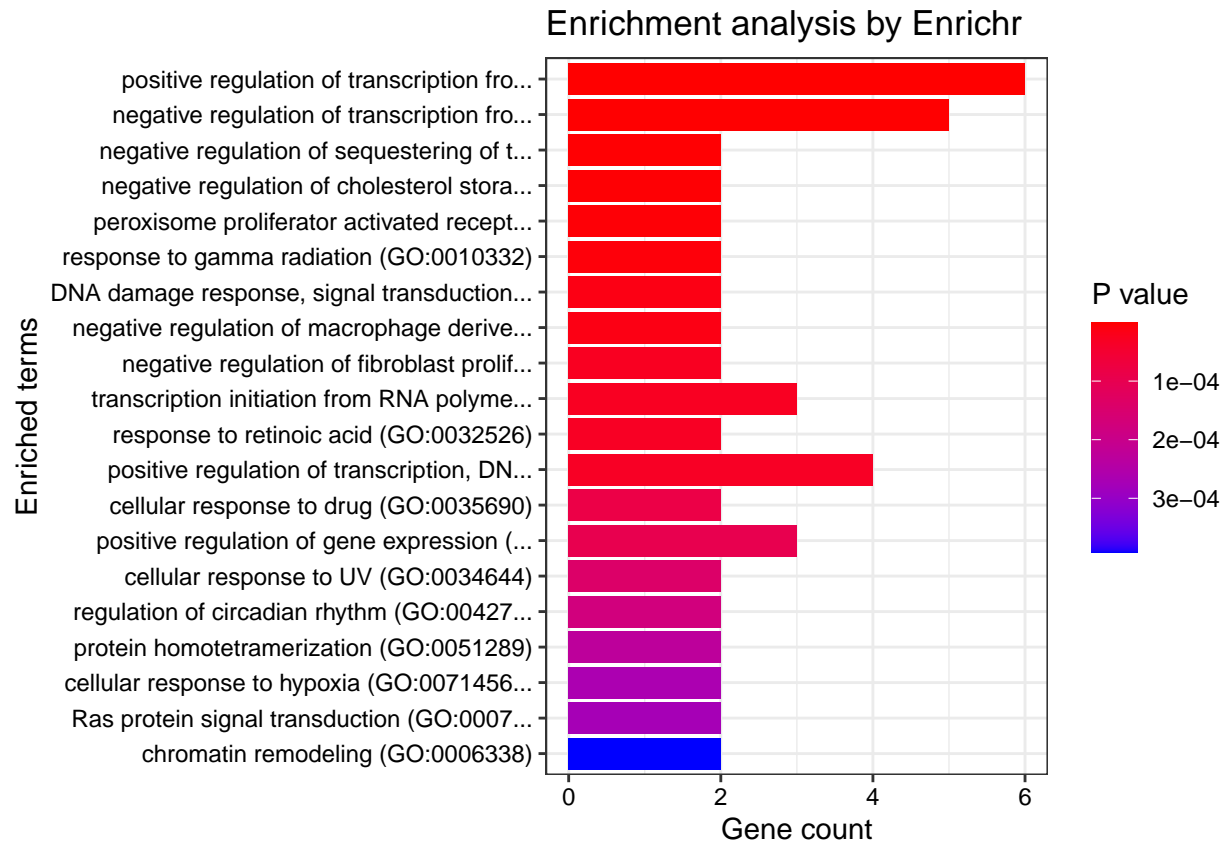
```
##              Term
## 1 positive regulation of transcription from RNA polymerase II promoter (GO:0045944)
## 2 negative regulation of transcription from RNA polymerase II promoter (GO:0000122)
## 3          negative regulation of sequestering of triglyceride (GO:0010891)
## 4          negative regulation of cholesterol storage (GO:0010887)
## 5      peroxisome proliferator activated receptor signaling pathway (GO:0035357)
## 6          response to gamma radiation (GO:0010332)
## Overlap   P.value Adjusted.P.value Old.P.value Old.Adjusted.P.value
## 1  6/712 1.5278e-07      2.3070e-05           0           0
## 2  5/473 8.4392e-07      6.3716e-05           0           0
## 3   2/5 1.7987e-06      9.0534e-05           0           0
## 4   2/6 2.6974e-06      1.0183e-04           0           0
## 5   2/7 3.7755e-06      1.1402e-04           0           0
## 6   2/9 6.4693e-06      1.6281e-04           0           0
## Odds.Ratio Combined.Score              Genes
## 1    54.632      857.41 RXRA;MYC;ATRX;PPARG;PPARA;TP53
## 2    52.145      729.26  RXRA;MYC;PPARG;PPARA;TP53
```

##	3	1903.619	25181.94	PPARG;PPARA
##	4	1427.643	18306.98	PPARG;PPARA
##	5	1142.057	14260.84	RXRA;PPARG
##	6	815.673	9746.03	MYC;TP53

```
plotEnrich(res$KEGG_2021_Human)
```



```
plotEnrich(res$GO_Biological_Process_2017)
```



FCS

```
suppressPackageStartupMessages(library(fgsea))
data(examplePathways)
data(exampleRanks)
set.seed(42)
```

```
head(examplePathways, 2)
```

```
## $`1221633_Meiotic_Synapsis`
## [1] "12189" "13006" "15077" "15078" "15270" "15512"
## [7] "16905" "16906" "19357" "20842" "20843" "20957"
## [13] "20962" "21749" "21750" "22196" "23856" "24061"
## [19] "28113" "50878" "56739" "57321" "64009" "66654"
## [25] "69386" "71846" "74075" "77053" "94244" "97114"
## [31] "97122" "97908" "101185" "140557" "223697" "260423"
## [37] "319148" "319149" "319150" "319151" "319152" "319153"
## [43] "319154" "319155" "319156" "319157" "319158" "319159"
## [49] "319160" "319161" "319565" "320332" "320558" "326619"
## [55] "326620" "360198" "497652" "544973" "625328" "667250"
## [61] "100041230" "102641229" "102641751" "102642045"
##
## $`1368092_Rora_activates_gene_expression`
## [1] "11865" "12753" "12894" "18143" "19017" "19883" "20787" "217166"
## [9] "328572"
```

```
head(exampleRanks, 2)
```

```
## 170942 109711
## -63.337 -49.748
```

```
fgseaRes <- fgsea(pathways = examplePathways,
                 stats      = exampleRanks,
                 minSize    = 15,
                 maxSize     = 500)
```

```
## Warning in fgseaMultilevel(...): For some pathways, in reality P-values are less
## than 1e-10. You can set the `eps` argument to zero for better estimation.
```

```
head(fgseaRes[order(pval), ])
```

```
##           pathway pval      padj log2err      ES      NES size
## 1: 5990979_Cell_Cycle,_Mitotic 1e-10 3.9067e-09      NA 0.55948 2.7691 317
## 2:           5990980_Cell_Cycle 1e-10 3.9067e-09      NA 0.53885 2.7059 369
## 3:           5990981_DNA_Replication 1e-10 3.9067e-09      NA 0.64400 2.6394 82
## 4:           5990987_Synthesis_of_DNA 1e-10 3.9067e-09      NA 0.64786 2.6429 78
## 5:           5990988_S_Phase 1e-10 3.9067e-09      NA 0.60131 2.5296 98
## 6:           5990990_G1_S_Transition 1e-10 3.9067e-09      NA 0.62329 2.5738 84
##           leadingEdge
## 1: 66336,66977,12442,107995,66442,12571,...
## 2: 66336,66977,12442,107995,66442,19361,...
## 3: 57441,17219,69270,12575,69263,17215,...
## 4: 17219,69270,12575,69263,17215,68240,...
## 5: 67849,17219,69270,12575,69263,71988,...
## 6: 20135,13555,17219,12575,12448,17215,...
```

PT

```
suppressPackageStartupMessages(library(pathfindR))
head(RA_input)
```

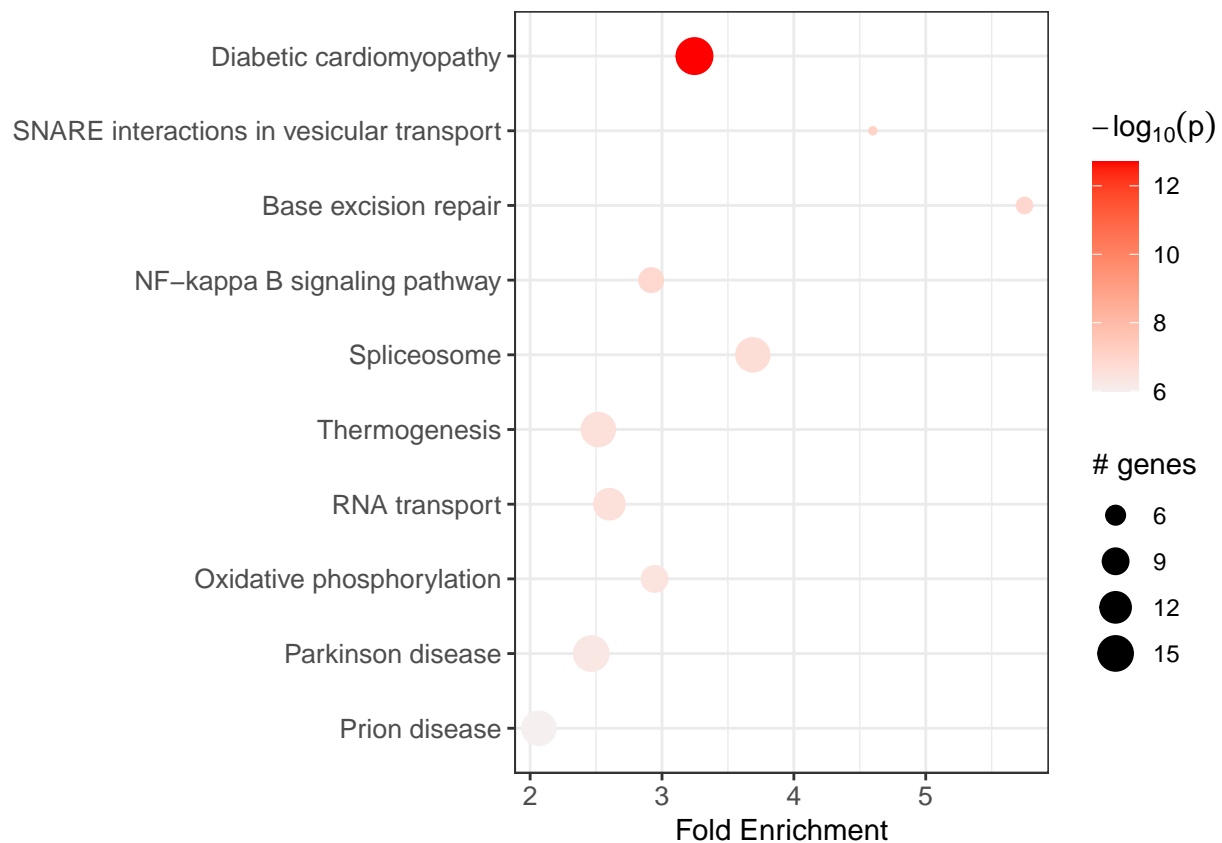
```
## Gene.symbol logFC adj.P.Val
## 1 FAM110A -0.69394 3.4087e-06
## 2 RNASE2 1.35350 1.0085e-05
## 3 S100A8 1.54483 3.4664e-05
## 4 S100A9 1.02809 2.2626e-04
## 5 TEX261 -0.32360 2.2626e-04
## 6 ARHGAP17 -0.69193 2.7081e-04
```

```
# RA_out_df <- run_pathfindR(RA_input, output_dir = tmp_dir)
RA_out_df <- RA_output
head(RA_out_df)
```

```
## ID Term_Description Fold_Enrichment occurrence
## 1 hsa05415 Diabetic cardiomyopathy 3.2464 10
## 2 hsa04130 SNARE interactions in vesicular transport 4.5990 10
## 3 hsa03410 Base excision repair 5.7488 1
## 4 hsa04064 NF-kappa B signaling pathway 2.9186 10
## 5 hsa03040 Spliceosome 3.6888 10
## 6 hsa04714 Thermogenesis 2.5175 10
## support lowest_p highest_p
```

```
## 1 0.0752907 1.9512e-13 1.9512e-13
## 2 0.0118360 8.0071e-08 8.0071e-08
## 3 0.0053191 1.2365e-07 1.2365e-07
## 4 0.0502592 1.2586e-07 1.2586e-07
## 5 0.0477457 2.2308e-07 4.4562e-05
## 6 0.0441982 2.5435e-07 2.5435e-07
##
## Up_regulated
## 1 NCF4, MMP9, NDUFA1, NDUFB3, UQCQRQ, COX6A1, COX7A2, COX7C, GAPDH
## 2 STX6
## 3 POLE4
## 4 LY96
## 5 SF3B6, LSM3, BUD31
## 6 NDUFA1, NDUFB3, UQCQRQ, COX6A1, COX7A2, COX7C
##
## Down_regulated
## 1 ATP2A2, MTOR, PDHA1, PDHB, VDAC1, SLC25A5, PARP1
## 2 STX2, BET1L, SNAP23
## 3 MUTYH, APEX2, POLD2, PARP1
## 4 PRKCQ, CARD11, TICAM1, IKBKB, PARP1, UBE2I, CSNK2A2
## 5 SNRPB, SF3B2, U2AF2, PUF60, SNU13, DDX23, EIF4A3, HNRNPA1, PCBP1, SRSF8, SRSF5
## 6 ADCY7, CREB1, KDM1A, SMARCA4, ACTG1, ACTB, ARID1A, MTOR
```

```
enrichment_chart(RA_out_df)
```



```
RA_clu_df <- cluster_enriched_terms(RA_out_df, method = "fuzzy")
```



