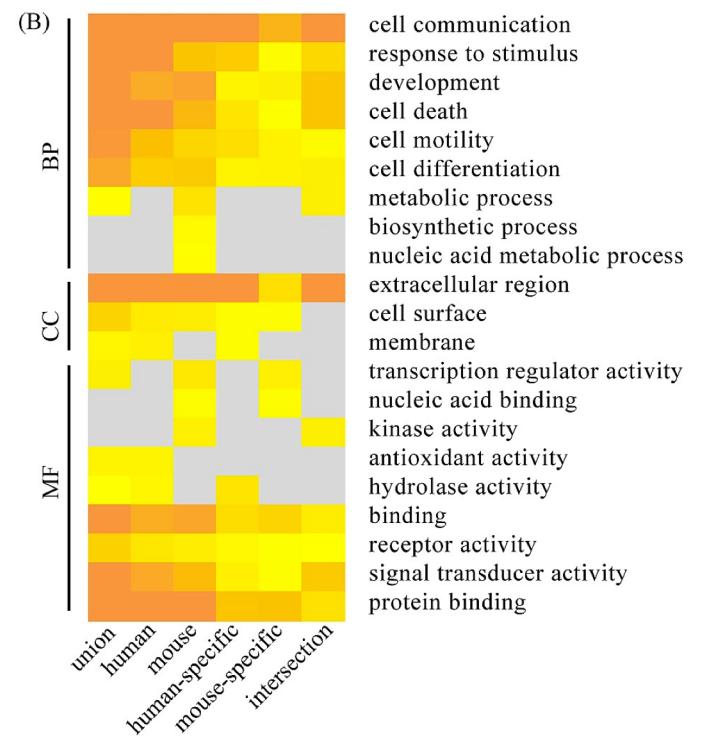
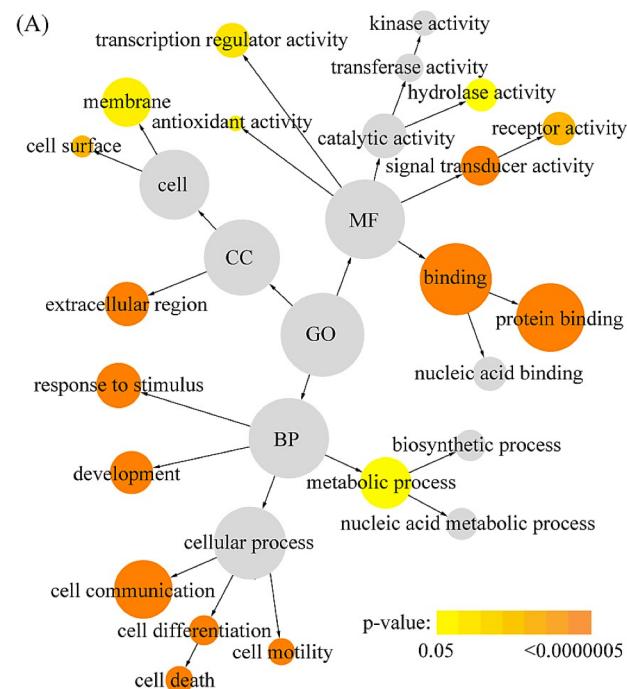
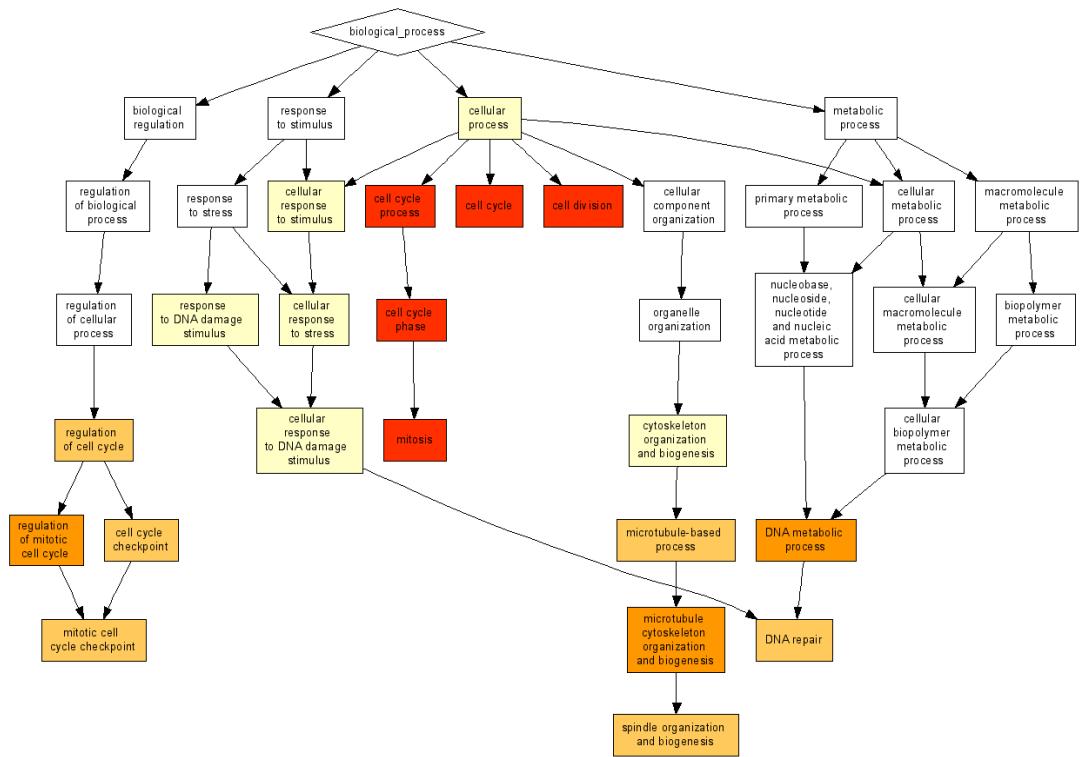


GO enrichment analysis

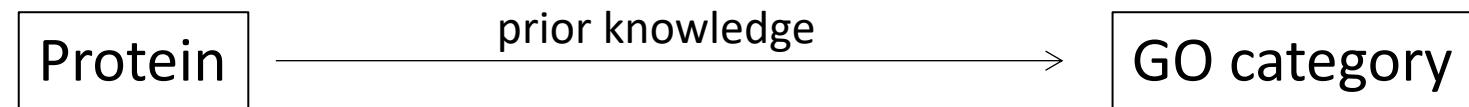


What do the genes do?

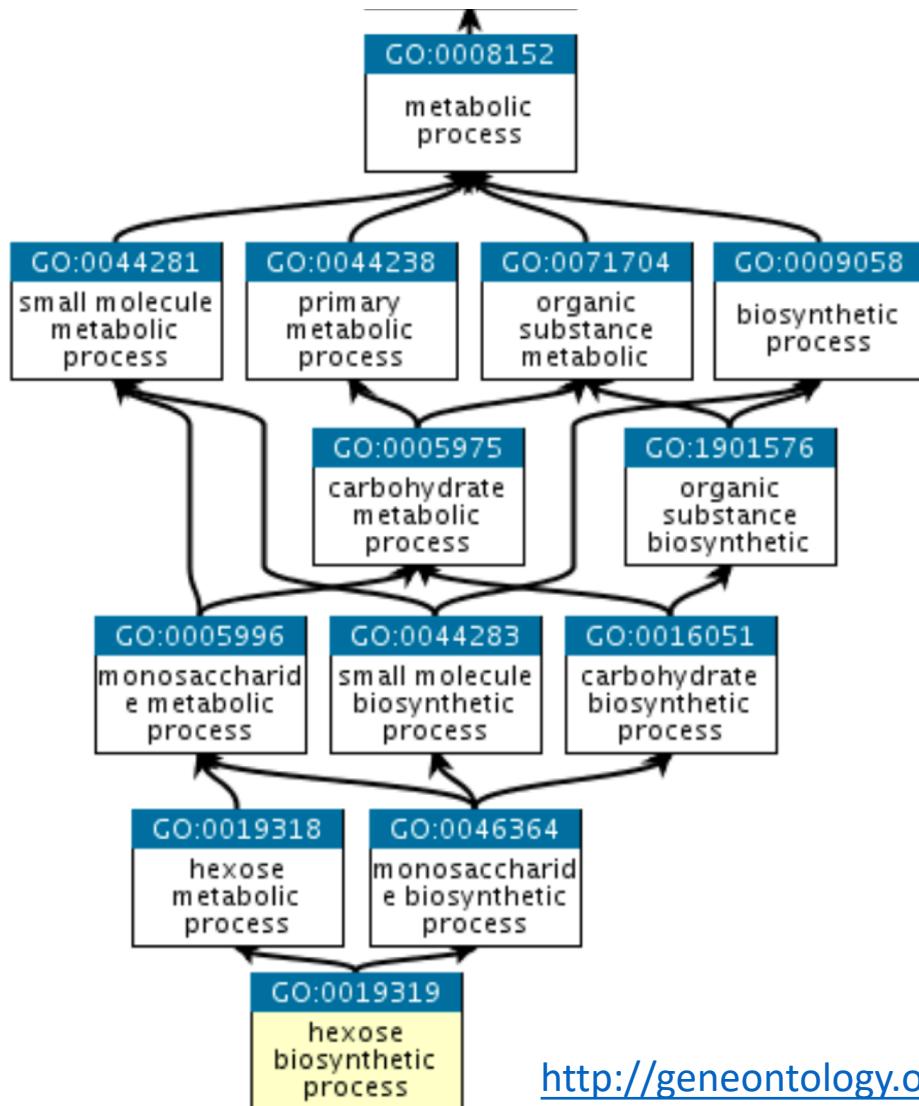
- Prupe.1G012200 -----> AT2G32120 - heat shock protein
 - 50 peach genes
 - 1,000 genes -----> gene ontology terms
- BLAST
- <https://www.arabidopsis.org/servlets/TairObject?name=AT2G32120&type=locus>

What is GO?

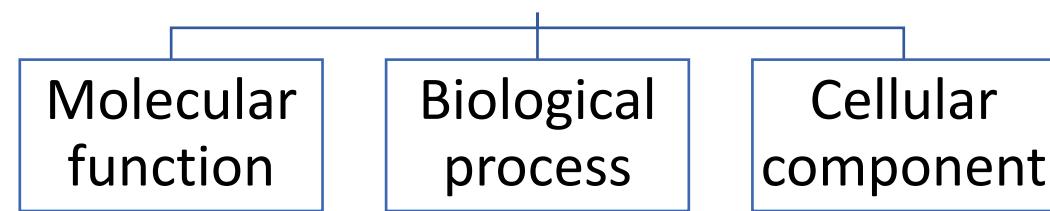
- The Gene Ontology (GO) describes our knowledge of the biological domain.



Gene ontology structure



Three independent root ontology term



GO is loosely hierarchical, with 'child' terms being more specialized than their 'parent' terms

GO term

- AT2G32120

| category | keyword |
|-----------------------|--|
| GO Biological Process | cellular response to heat , cellular response to unfolded protein , chaperone cofactor-dependent protein refolding , protein folding , protein refolding , response to heat , response to unfolded protein |
| GO Cellular Component | cytoplasm |
| GO Molecular Function | ATP binding , ATPase activity , ATPase activity, coupled , heat shock protein binding , misfolded protein binding , protein folding chaperone , unfolded protein binding |

- Example

| locusName | GO term |
|----------------|--|
| Prupe.1G001400 | GO:0016747, GO:0008610, GO:0016020, GO:0006633 |
| Prupe.1G006600 | GO:0046983 |
| Prupe.1G012200 | |
| Prupe.1G012900 | |
| Prupe.1G031300 | GO:0003676 |
| Prupe.1G031800 | GO:0006468, GO:0005524, GO:0004672 |
| Prupe.1G033800 | GO:0006468, GO:0005524, GO:0004672 |
| Prupe.1G034000 | GO:0006468, GO:0005524, GO:0004672 |
| Prupe.1G035500 | GO:0055114, GO:0016491 |

GO enrichment

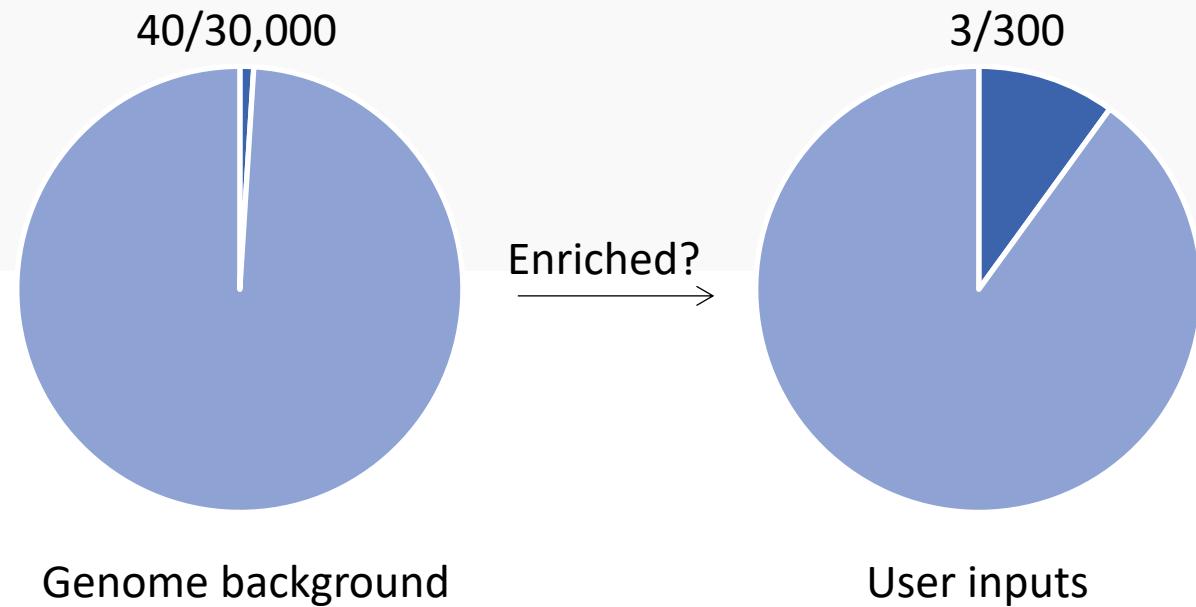


A Hypothetical Example:

In human genome background (30,000 gene total), 40 genes are involved in p53 signaling pathway. A given gene list has found that 3 out of 300 belong to p53 signaling pathway. Then we ask the question if 3/300 is more than random chance comparing to the human background of 40/30000.

A 2x2 contingency table is built on above numbers:

| | User Genes | Genome |
|----------------|------------|--------|
| In Pathway | 3-1 | 40 |
| Not In Pathway | 297 | 29960 |



https://david.ncifcrf.gov/content.jsp?file=functional_annotation.html

Other ontology databases

- KEGG - Kyoto Encyclopedia of Genes and Genomes
- COG – Clusters of Orthologous Groups
- KOG – Eukaryote Orthologous Groups

<https://www.genome.jp/kegg/>

<https://www.ncbi.nlm.nih.gov/COG/>

GO enrichment tools

- [AgriGO](#) – plant genomes
- [BiNGO](#) – model organisms, Cytoscape plugin
- [Gorilla](#) – model organisms
- [DAVID](#) – human and animal genomes

AgriGO enrichment output

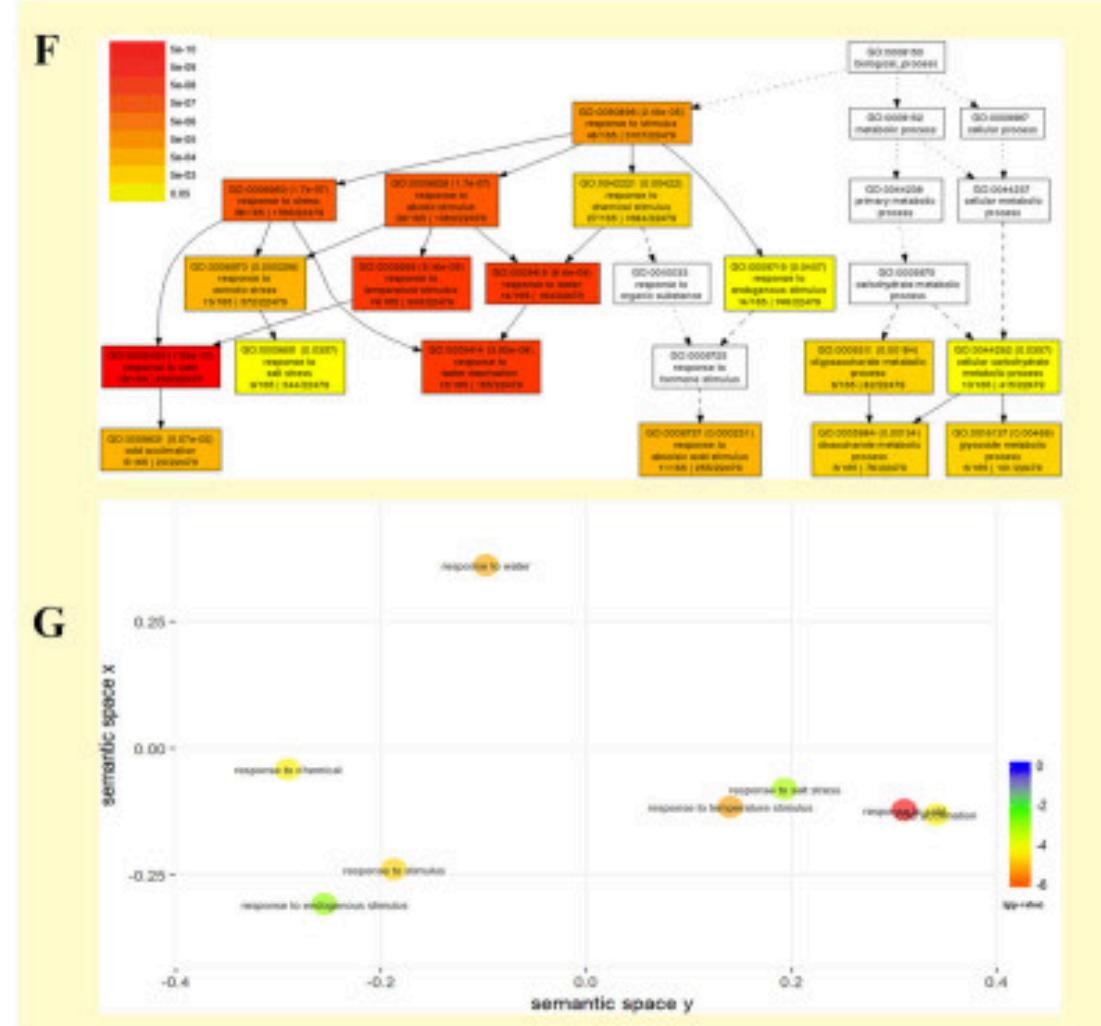
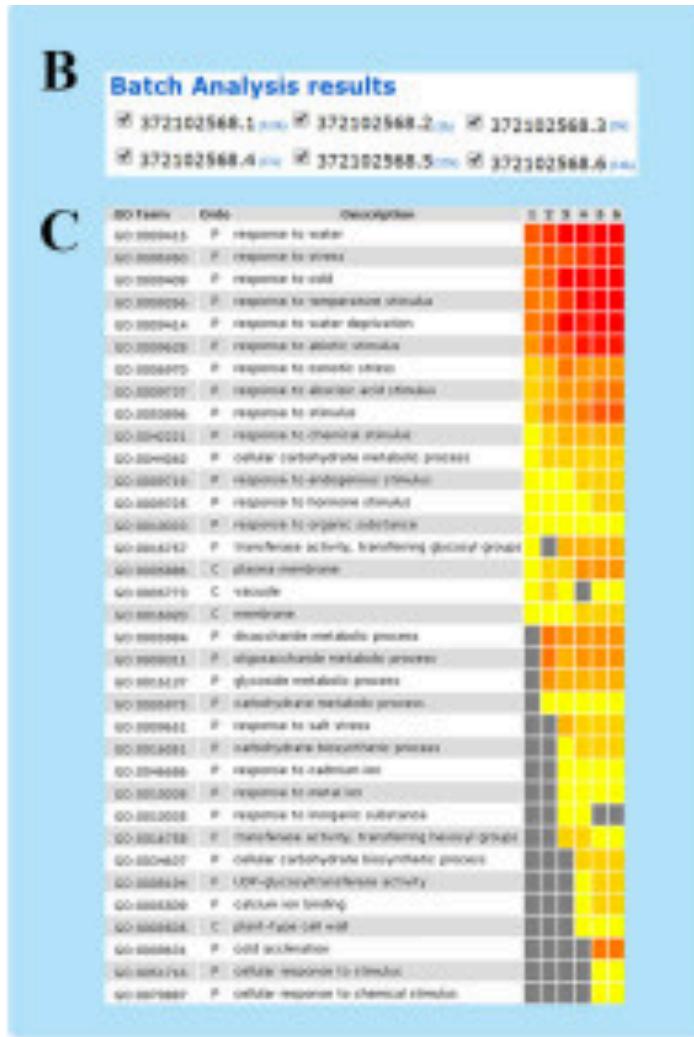
Detail information

You can [ Browse in tree traversing mode] [ Browse all GO terms] [ Download]

Or select from following significant terms to [Draw graphical results] [Create bar chart] [Scatter Plots analysis]

| <input type="checkbox"/> GO term | Ontology | Description | Number in input list | Number in BG/Ref | p-value | FDR |
|-------------------------------------|----------|---|----------------------|------------------|---------|---------|
| <input type="checkbox"/> GO:0005985 | P | sucrose metabolic process | 9 | 44 | 4.1e-08 | 4.9e-05 |
| <input type="checkbox"/> GO:0043038 | P | amino acid activation | 13 | 140 | 7.8e-07 | 0.00023 |
| <input type="checkbox"/> GO:0043039 | P | tRNA aminoacylation | 13 | 140 | 7.8e-07 | 0.00023 |
| <input type="checkbox"/> GO:0006418 | P | tRNA aminoacylation for protein translation | 13 | 138 | 6.6e-07 | 0.00023 |
| <input type="checkbox"/> GO:0055067 | P | monovalent inorganic cation homeostasis | 7 | 35 | 1.6e-06 | 0.00031 |
| <input type="checkbox"/> GO:0005984 | P | disaccharide metabolic process | 10 | 84 | 1.5e-06 | 0.00031 |
| <input type="checkbox"/> GO:0045490 | P | pectin catabolic process | 6 | 26 | 3.7e-06 | 0.00062 |
| <input type="checkbox"/> GO:0000272 | P | polysaccharide catabolic process | 8 | 69 | 2e-05 | 0.003 |
| <input type="checkbox"/> GO:0009311 | P | oligosaccharide metabolic process | 10 | 117 | 2.9e-05 | 0.0038 |
| <input type="checkbox"/> GO:0006771 | P | riboflavin metabolic process | 5 | 24 | 4.1e-05 | 0.0041 |

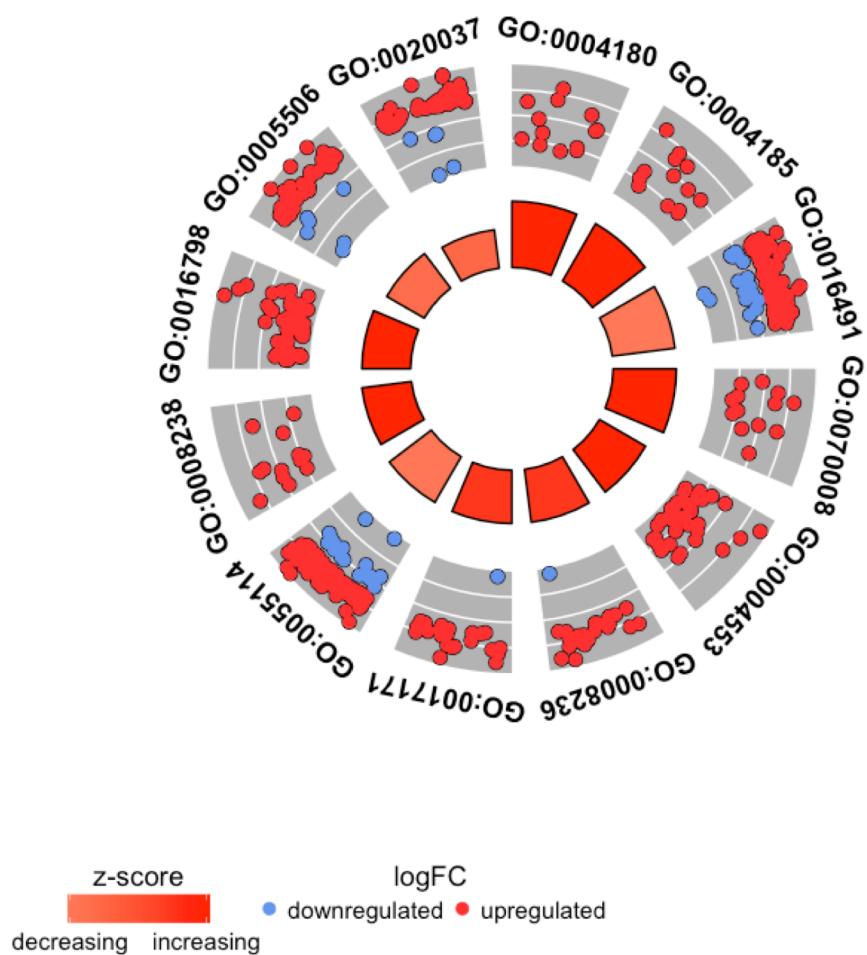
GO enrichment visualization



AgriGO demo

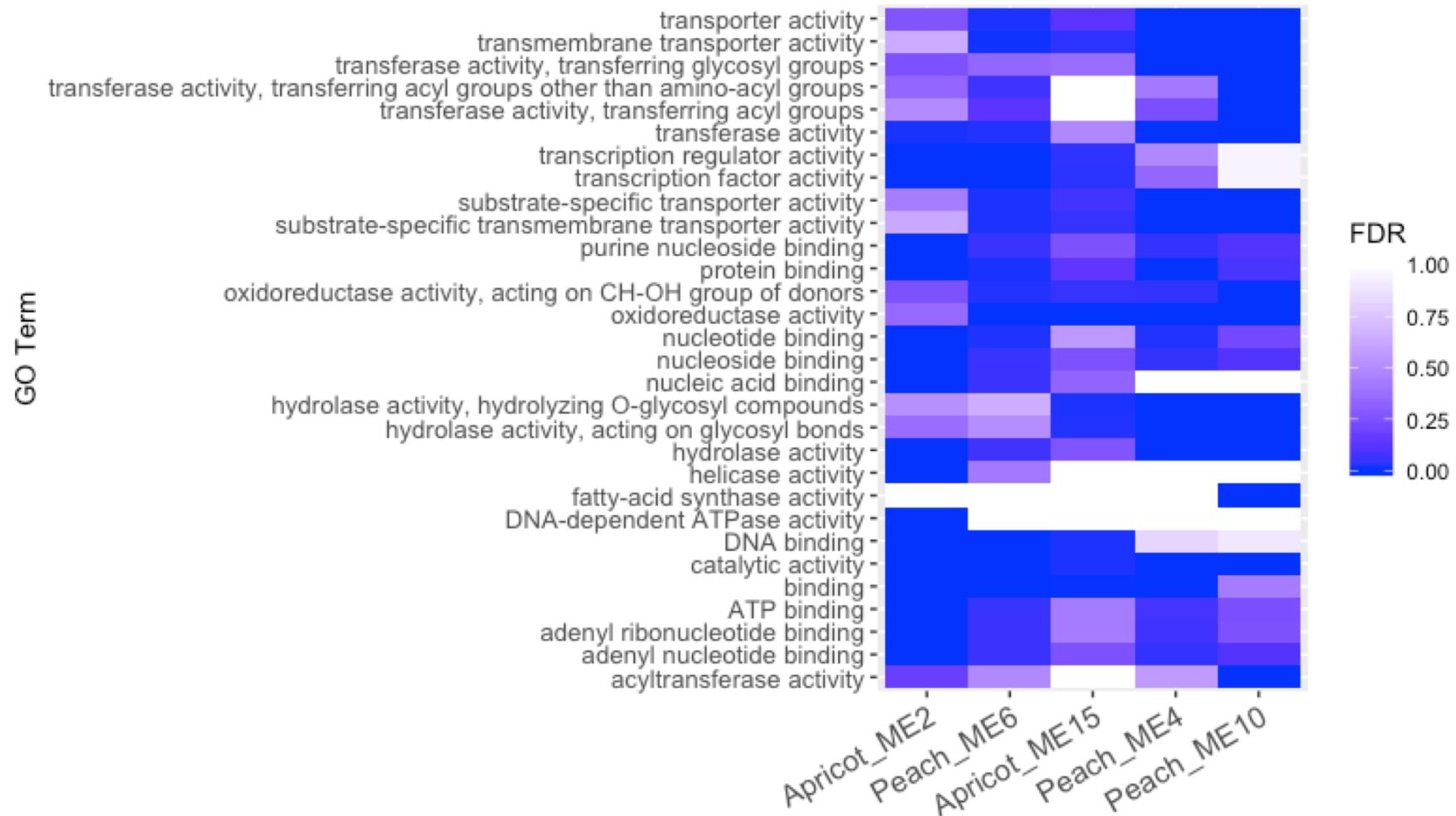
- <http://systemsbiology.cau.edu.cn/agriGOv2/>

Make your own figures with the results

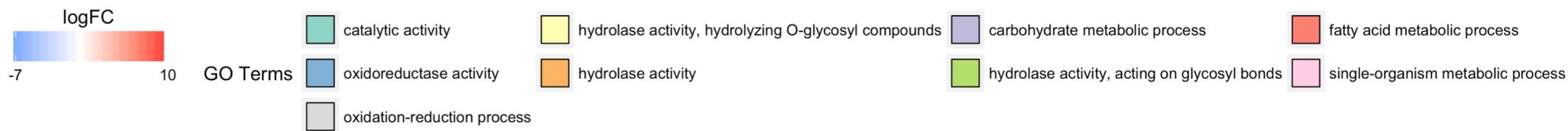


| ID | Description |
|------------|--|
| GO:0004180 | carboxypeptidase activity |
| GO:0004185 | serine-type carboxypeptidase activity |
| GO:0016491 | oxidoreductase activity |
| GO:0070008 | serine-type exopeptidase activity |
| GO:0004553 | hydrolase activity, hydrolyzing O-glycosyl compounds |
| GO:0008236 | serine-type peptidase activity |
| GO:0017171 | serine hydrolase activity |
| GO:0055114 | oxidation reduction |
| GO:0008238 | exopeptidase activity |
| GO:0016798 | hydrolase activity, acting on glycosyl bonds |
| GO:0005506 | iron ion binding |
| GO:0020037 | heme binding |

Make your own figures with the results



Make your own figures with the results



Reference

- <http://geneontology.org/>
- <https://david.ncifcrf.gov/>
- Tian T, Liu Y, Yan H, et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 2017;45(W1):W122–W129. doi:10.1093/nar/gkx382