Brief overview on how to start exploring functionality

Christian Werner

(Quantitative geneticist and biostatistician) **EiB, CIMMYT**, Texcoco (Mexico)

Filippo Biscarini

HerrFalloppio

(Biostatistician, bioinformatician, and quantitative geneticist) **CNR-IBBA**, Milan (Italy)

Oscar González-Recio

OscarGenomics

(Computer biologist, and quantitative geneticist) **INIA-UPM**, Madrid (Spain)

Functional analysis

- Examining the function of genes and their role in biological processes.
  - by looking at how changes in a gene's sequence affect its function
  - by studying how different genes interact with each other to carry out specific functions in the cell.
  - using bioinformatics tools to analyze large datasets of genomic information, such as the sequence of an entire genome or the expression levels of thousands of genes in different tissues or under different conditions. This can help identify patterns and connections between genes that may be involved in specific functions or processes, such as disease development or response to environmental stimuli.

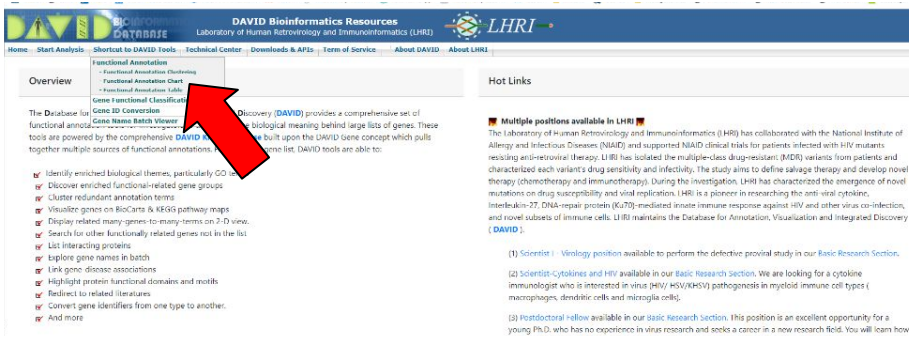
Overall, a genomic functional analysis can provide valuable insights into the roles that genes play in biological processes, and can help researchers better understand the underlying causes of diseases and other biological phenomena

Work Flow in FUMA

- Download bkg genes: https://www.ensembl.org/info/data/ftp/index.html
- Results overview in FUMA
  - https://fuma.ctglab.nl/snp2gene
  - Input file: GWASresults.txt
- Variant Effect Prediction in Ensembl
  - https://www.ensembl.org/Multi/Tools/VEP
  - Input file (significant SNPs): Map.selected.rs
  - Output file (Functional Info): Select 'Gene' column
- Enrichment analysis
  - https://fuma.ctglab.nl/gene2func
  - Input files: GENES from significant SNPs ('Gene' column from VEP)
  - -Background genes from the specie (Canis_familiaris.bkg_genes from here

Work Flow in DAVID

- Visit the DAVID website and create an account, if you don't already have one.
  - https://david.ncifcrf.gov/
- Log in to your account and go to the "Functional Annotation" section of the platform.
- Choose the appropriate algorithms and methods for your analysis. DAVID offers a variety of options for functional analysis, including clustering, enrichment analysis, and pathway mapping.

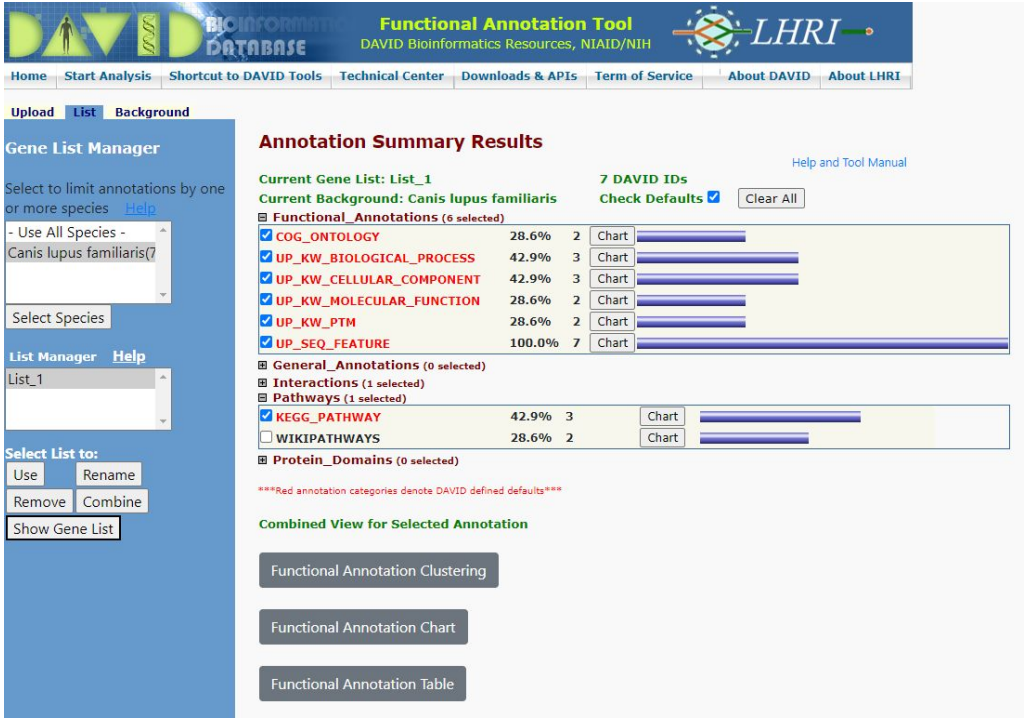


Work Flow in DAVID

- Paste the list of genes (output from the getGenes.R)
- Select OFFICIAL_GENE_SYMBOL
- Type the species
- Submit analysis

Use the results of your functional analysis to gain insights into the biological processes and functions that are represented in your data. You can use the information from DAVID to help identify potential pathways and mechanisms involved in the development of diseases, for example, or to investigate how different genes or proteins interact to perform specific functions in the cell.

Work Flow in DAVID



Go again to the Shortcut to ***David Tools*** and select ***'Functional Annotation'***

You can select the different options to gain insights into the biological processes and functions that are represented in your data. You can use the information from DAVID to help identify potential pathways and mechanisms involved in the expression of the phenotype. For example, to investigate how different genes or proteins interact to perform specific functions in the cell.

Limitations

- FA relies on algorithms and methods that make assumptions and simplifications about the data being analyzed. These assumptions may not always hold true in all cases, and they can affect the accuracy and reliability of the results.
- FA are often based on large datasets that may not be representative of all possible scenarios. This can lead to bias in the results, and can make it difficult to generalize the findings to other situations or organisms.
- FA can provide valuable insights, but it is important to carefully consider the limitations and uncertainties of the methods and algorithms used, and to interpret the results with caution.