

Brief overview on how to start exploring **functionality**

Christian Werner

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

Filippo Biscarini

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



HerrFaloppio

Oscar González-Recio

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



OscarGenomics



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



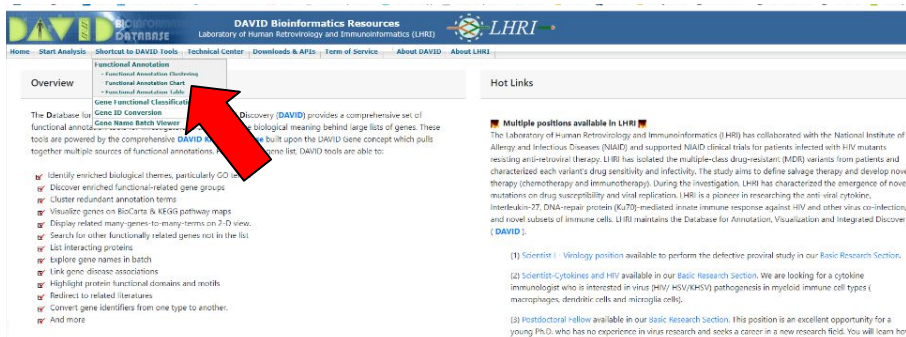
Work Flow in FUMA

- 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.



DAVID Bioinformatics Resources
Laboratory of Human Retrovirology and Immunoinformatics (LHRI)

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service About DAVID About LHRI

Functional Annotation
Functional Annotation (Database)
Functional Annotation Chart
Functional Annotation Table
Gene Functional Classification

Overview

The Database for Annotation, Visualization and Integrated Discovery (DAVID) provides a comprehensive set of functional annotation tools to help researchers understand the biological meaning behind large lists of genes. These tools are powered by the comprehensive DAVID Gene Ontology (GO) and KEGG pathway maps. DAVID tools are able to:

- Identify enriched biological themes, particularly GO terms
- Discover enriched functional-related gene groups
- Cluster redundant annotation terms
- Visualize genes on BioCarta & KEGG pathway maps
- Display related many-genes-to-many-terms on 2-D view
- Search for other functionally related genes not in the list
- List Interacting proteins
- Explore gene names in batch
- Link gene-disease associations
- Highlight protein functional domains and motifs
- Redirect to related literatures
- Convert gene identifiers from one type to another
- And more

Hot Links

Multiple positions available in LHRI

The Laboratory of Human Retrovirology and Immunoinformatics (LHRI) has collaborated with the National Institute of Allergy and Infectious Diseases (NIAID) and supported NIAID clinical trials for patients infected with HIV mutants residing anti-retroviral therapy. LHRI has isolated the multiple-class drug-resistant (MDR) variants from patients and characterized each variant's drug sensitivity and infectivity. The study aims to define salvage therapy and develop novel therapy (chemotherapy and immunotherapy). During the investigation, Linki has characterized the emergence of novel mutations on drug susceptibility and viral replication. LHRI is a pioneer in researching the anti-viral cytokines, Interleukin-27, DNA-repair protein (Ku70)-mediated innate immune response against HIV and other virus co-infection, and novel subsets of immune cells. LHRI maintains the Database for Annotation, Visualization and Integrated Discovery (DAVID).

- (1) Scientist - Virology position available to perform the defective proviral study in our Basic Research Section.
- (2) Scientist-Cytokines and HIV available in our Basic Research Section. We are looking for a cytokine immunologist who is interested in virus (HIV/HSV/KHSV) pathogenesis in myeloid immune cell types (macrophages, dendritic cells and microglia cells).
- (3) Postdoctoral Fellow available in our Basic Research Section. This position is an excellent opportunity for a young Ph.D. who has no experience in virus research and seeks a career in a new research field. You will learn how

Work Flow in DAVID



Gene Name Batch Viewer
DAVID Bioinformatics Resources, NIAID/NIH

Home | Start Analysis | Shortcut to DAVID Tools | Technical Center | Downloads & APIs | Term of Service | About DAVID | About LHRI

Upload | List | Background

Gene Name Batch Viewer

Submit your gene list to start!

Tell us how you like the tool
Read technical notes of the tool
Contact us for questions

What does this tool do?

- Quickly translate given gene IDs to corresponding gene names in a batch way
- Provide links for each gene to DAVID Gene Report for in-depth information
- Search functionally related genes within user's input gene list or genome

Key Concepts of "Search Related Genes"

Any given gene is associating with a set of annotation terms. If genes share similar set of those terms (annotation profile), they are most likely involved in similar biological mechanisms. The algorithm adopts kappa statistics to quantitatively measure the degree of the agreement how genes share the ~75,000 annotation terms collected by DAVID knowledgebase. For any given gene(s), the tool instantly searches and lists the related genes passed kappa similarity measurement threshold. The searching scope could be within user's input gene list, selected genome or all genomes (~1.2 million genes) as user's choice.

Find Related Genes Tool is very different and complementary to the common gene clustering methods, such as homologous genes based on sequence similarity; protein families based on one common by measuring the similarity of their global annotation profile, which facilitates new understanding of the biological network. [More](#)

Step 1: Enter Gene List
A: Paste a list

PDZRN4
CNTN1
LRRK2
SLC2A13

Clear

Or
B: Choose From a File
Seleccionar archivo Ninguno archivo seleccionado
Multi-List File

Step 2: Select Identifier
OFFICIAL_GENE_SYMBOL

Step 2a: Select species
Canis lupus familiaris

Step 3: List Type
Gene List
Background

Step 4: Submit List
Submit List

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



The screenshot displays the DAVID Functional Annotation Tool interface. The top navigation bar includes 'Home', 'Start Analysis', 'Shortcut to DAVID Tools', 'Technical Center', 'Downloads & APIs', 'Term of Service', 'About DAVID', and 'About LHR'. The main content area is titled 'Annotation Summary Results' and shows the following data:

Annotation Category	Percentage	Count	Chart
COG_ONTOLOGY	28.6%	2	Chart
UP_KW_BIOLOGICAL_PROCESS	42.9%	3	Chart
UP_KW_CELLULAR_COMPONENT	42.9%	3	Chart
UP_KW_MOLECULAR_FUNCTION	28.6%	2	Chart
UP_KW_PTM	28.6%	2	Chart
UP_SEQ_FEATURE	100.0%	7	Chart

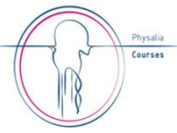
Below the table, there are sections for 'General_Annotations (0 selected)', 'Interactions (1 selected)', and 'Pathways (1 selected)'. The 'Pathways' section shows 'KEGG_PATHWAY' (42.9%, 3) and 'WIKIPATHWAYS' (28.6%, 2). At the bottom, there are three buttons: 'Functional Annotation Clustering', 'Functional Annotation Chart', and 'Functional Annotation Table'.

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.

