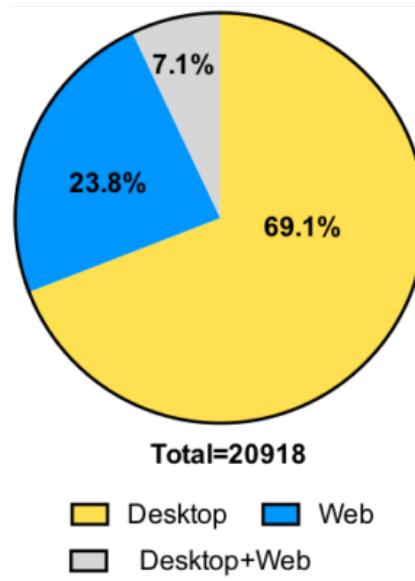


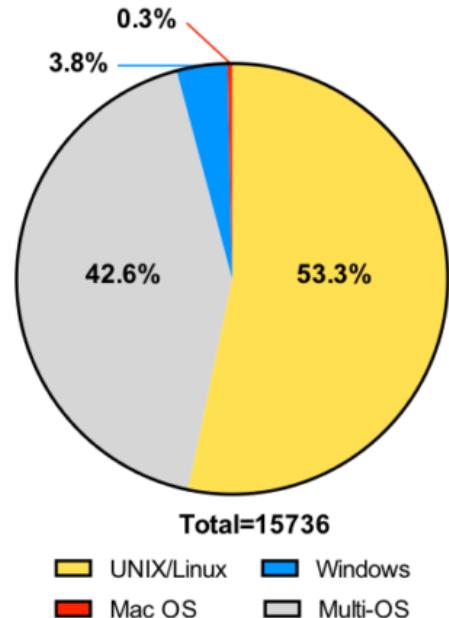


## Más de la mitad de herramientas bioinformáticas con uso conocido se ejecutan en UNIX/Linux

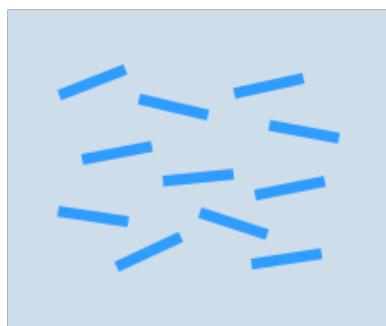
Tools specifications



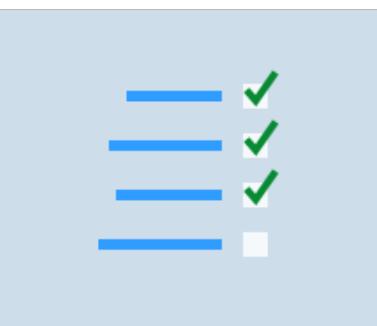
Tools operating systems



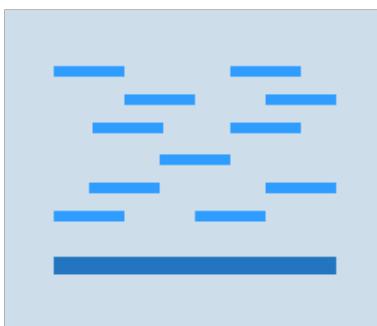
[https://omictools.com/blog/dataontools-4-evolution-tool-specifications?utm\\_source=USER+NEWSLETTER&utm\\_campaign=0f0931de67-MAINTAINER\\_COPY\\_01&utm\\_medium=email&utm\\_term=0\\_ac8d2cb0a2-0f0931de67-225798421](https://omictools.com/blog/dataontools-4-evolution-tool-specifications?utm_source=USER+NEWSLETTER&utm_campaign=0f0931de67-MAINTAINER_COPY_01&utm_medium=email&utm_term=0_ac8d2cb0a2-0f0931de67-225798421)



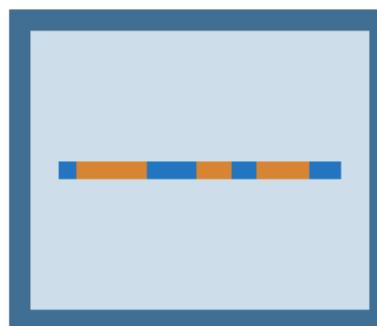
Sequencing



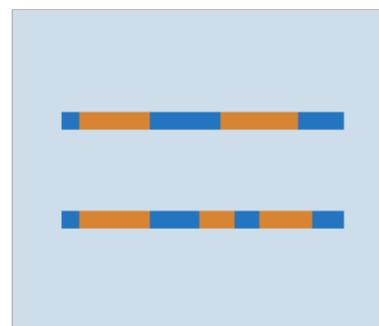
Quality control



Assembly



Annotation



Comparison



# IX. Ontología Genética (Gene Ontology)

La ontología genética proporciona un **vocabulario controlado y estructurado** que cubre varios dominios de biología celular y molecular, es libremente disponible para realizar **anotaciones de genes, productos de genes y secuencias** y aplica a todos los organismos.

# Estructura

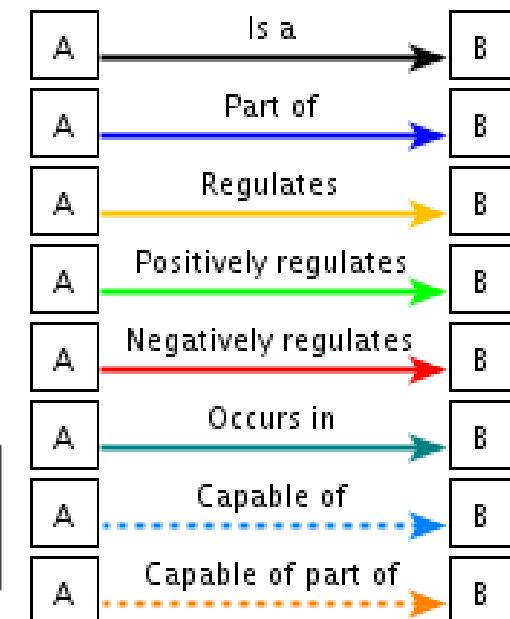
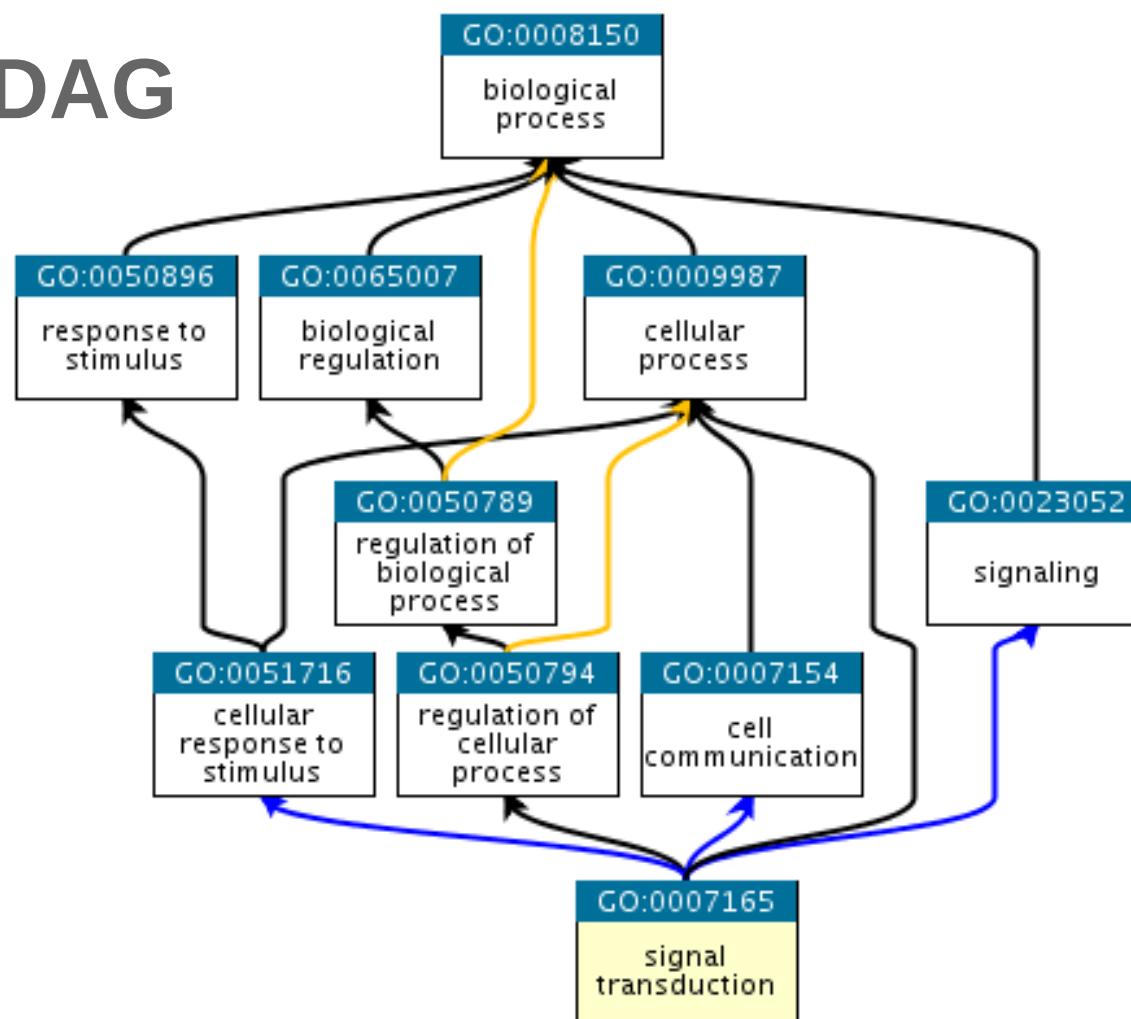
**Directed Acyclic Graph:** es como un gráfico acíclico donde cada término tiene relaciones definidas con uno o más términos en el mismo dominio

Padre

Especificidad

Hijo

DAG



# Anotación Funcional

¿Qué es?

Para “Gene Ontology” una anotación es...

... una declaración donde el producto de un gen

tiene una **función molecular** particular

o está implicado en un **proceso biológico** particular

o está localizado dentro de cierto **componente celular**

# Expansion of the Gene Ontology knowledgebase and resources

The Gene Ontology Consortium\*

List of authors of the Gene Ontology Consortium is provided in the Appendix.

Received October 25, 2016; Editorial Decision October 26, 2016; Accepted November 16, 2016

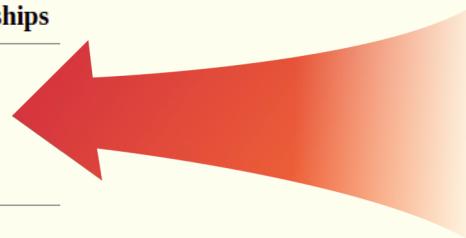
## ABSTRACT

The Gene Ontology (GO) source of computable knowledge about genes and gene products is expanding rapidly.

Table 1.

### Number of terms and relationships in the three aspects of the Gene Ontology, as of October 2016

Aspect	Terms (classes)	Relationships
Molecular function (MF)	10 417	14 039
Cellular component (CC)	4022	7854
Biological process (BP)	29 146	71 372



# Septiembre 2018

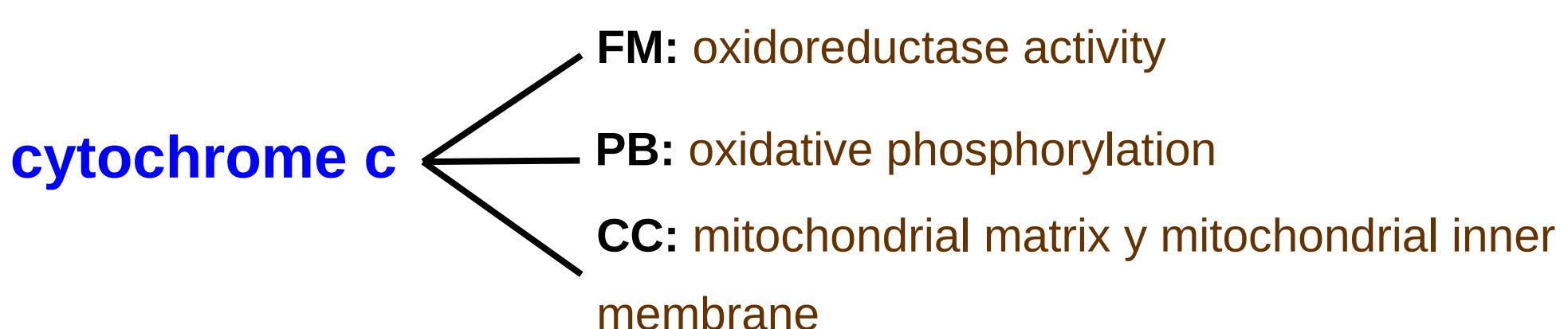
Proceso biológico	30782	GO terms
Función molecular	12110	GO terms
Componente celular	4386	GO terms

# Aspectos

Función Molecular: actividad elemental o tarea  
nuclease, DNA binding, catalytic activity

Proceso Biológico: objetivo específico  
mitosis, signal transduction, metabolism

Componente celular: localización o complejo  
nucleus, ribosome, membrane



# **Elementos**

**Validación de estructura de genes**

**Búsqueda de literatura**

**Búsqueda de homología/dominios**

**Asignación de funciones**

**Mantenimiento de los datos**

# Fuentes de anotación



Anotación manual

**UniprotKB**

Anotación electrónica

**Uniprot-GOA**

ARUK-UCL	<a href="#">Alzheimers Research Gene Ontology Initiative</a>
AgBase	<a href="#">AgBase resource for functional analysis of agricultural plant and animal gene products</a>
Alzheimers_University_of_Toronto	<a href="#">Alzheimers Project at University of Toronto</a>
AspGD	<a href="#">Aspergillus Genome Database</a>
BHF-UCL	<a href="#">Cardiovascular Gene Ontology Annotation Initiative</a>
CACAO	<a href="#">Community Assessment of Community Annotation with Ontologies</a>
CAFA	<a href="#">Critical Assessment of Protein Function Annotation</a>
CGD	<a href="#">Candida Genome Database</a>
CollecTF	<a href="#">Database of transcription factor binding sites (TFBS) in the Bacteria domain</a>
ComplexPortal	<a href="#">Complex Portal database of macromolecular complexes</a>
DFLAT	<a href="#">Developmental Functional Annotation at Tufts</a>
EcoCyc	<a href="#">Encyclopedia of E. coli metabolism</a>
EcoliWiki	<a href="#">EcoliWiki from EcoliHub</a>
Ensembl	<a href="#">Ensembl database of automatically annotated genomic data</a>
EnsemblFungi	<a href="#">Ensembl Fungi, the Ensembl database for accessing genome-scale data from fungi.</a>
EnsemblMetazoa	<a href="#">Ensembl Metazoa, the Ensembl database for accessing genome-scale data from non-vertebrate metazoa.</a>
EnsemblPlants	<a href="#">Ensembl Plants, the Ensembl database for accessing genome-scale data from plants.</a>
EnsemblProtists	<a href="#">Ensembl Protists, the Ensembl database for accessing genome-scale data from protists.</a>
FlyBase	<a href="#">FlyBase</a>
GDB	<a href="#">Human Genome Database</a>
GOC	<a href="#">Gene Ontology Consortium</a>
GOC-OWL	<a href="#">Gene Ontology Consortium - Logical inferences</a>
GO_Central	<a href="#">GO Central</a>
GR	<a href="#">Gramene</a>
GeneDB	<a href="#">GeneDB</a>
HGNC	<a href="#">HUGO Gene Nomenclature Committee</a>
HPA	<a href="#">Human Protein Atlas tissue profile information</a>
IntAct	<a href="#">IntAct protein interaction database</a>

InterPro	InterPro database of protein domains and motifs
JCVI	J. Craig Venter Institute
LIFEdb	LifeDB
MENGO	Microbial ENergy processes Gene Ontology Project
MGI	Mouse Genome Informatics
MTBBASE	Collection and Refinement of Physiological Data on <i>Mycobacterium tuberculosis</i>
NTNU_SB	Norwegian University of Science and Technology, Systems Biology team
PAMGO_GAT	Genome Annotation Tool ( <i>Agrobacterium tumefaciens</i> C58); PAMGO Interest Group
PAMGO_MGG	<i>Magnaporthe grisea</i> database
PINC	Proteome Inc.
ParkinsonsUK-UCL	Parkinsons Disease Gene Ontology Initiative
PomBase	PomBase
PseudoCAP	<i>Pseudomonas</i> Genome Project
RGD	Rat Genome Database
RNAcentral	RNAcentral
Reactome	Reactome - a curated knowledgebase of biological pathways
Roslin_Institute	Roslin Institute
SGD	<i>Saccharomyces</i> Genome Database
SYSCILIA_CCNET	Syscilia
SynGO	The Synapse Gene Ontology and Annotation Initiative
SynGO-UCL	The Synapse Gene Ontology and Annotation Initiative at UCL
TAIR	The Arabidopsis Information Resource
TIGR	J. Craig Venter Institute
UniProt	Universal Protein Knowledgebase
WormBase	WormBase database of nematode biology
YuBioLab	CITGeneDB
ZFIN	Zebrafish Information Network
dictyBase	dictyBase

# Evidence codes used in GO curation



**IEA** Inferred from Electronic Annotation

**IDA** Inferred from Direct Assay →

**IMP** Inferred from Mutant Phenotype

**IPI** Inferred from Protein Interaction

**IEP** Inferred from Expression Pattern

**IGI** Inferred from Genetic Interaction

**ISS** Inferred from Sequence or Structural Similarity

**IGC** Inferred from Genomic Context

**RCA** Reviewed Computational Analysis

**TAS** Traceable Author Statement →

**NAS** Non-traceable Author Statement

**IC** Inferred from Curator judgement

**ND** No Data available

**IDA:**

- Enzyme assays
- *In vitro* reconstitution (transcription)
- Immunofluorescence
- Cell fractionation

**TAS:**

- In the literature source the original experiments are referenced.



<http://www.geneontology.org/GO.evidence.shtml>

EMBL-EBI



# GO slim

Es una herramienta útil **para resumir** los resultados de **anotaciones** de genomas, microarreglos, transcriptomas, proteomas, etc., cuando se requiere una amplia clasificación funcional de productos de genes.

## Herramientas para crear GO slims

**OBO-Edit** (Desarrollado por el Consorcio GO)

**AgBase**

**CGD Gene Ontology Slim Mapper**

**Map2slim**

Cada una con sus ventajas y limitaciones

# GO slims

- Generic GO subset Developed by GO Consortium
- Plant subset Developed by [The Arabidopsis Information Resource](#)
- *Candida albicans* Developed by [Candida Genome Database](#)
- Protein Information Resource subset Developed by Darren Natale, [PIR](#)
- *Schizosaccharomyces pombe* subset Developed by Val Wood, [PomBase](#)
- Yeast subset Developed by [Saccharomyces Genome Database](#)
- *Aspergillus* subset Developed by [Aspergillus Genome Data](#)
- Metagenomics subset Developed by Jane Lomax and the [InterPro](#) group
- Virus subset Developed by Jane Lomax and Rebecca Foulger.
- ChEMBL Drug Target subset developed by Prudence Mutowo and Jane Lomax

# Herramientas para hacer anotaciones

## Metagenomics

<http://www.metagenomics.wiki/>

→ Prokka

Annotation tool for bacterial, archaeal, and viral genomes

→ RAST server

for bacterial and archaeal genomes

→ Maker

eukaryotic and prokaryotic genomes

→ GeneMark

- eukaryotics; prokaryotics; viruses, plasmids and plasmids  
- part of genome annotation pipelines at NCBI

→ MetaGeneMark

gene identification in metagenomic sequences

AmiGO

CateGORizer

g:Profiler

GeneMerge

Generic GO Term

Mapper

go-moose

go-perl

GOSlimViewer

QuickGO

ToppGene Suite

Blast2GO

Cluster Enrichment

GO Slim Mapper

GOHyperGAll

GOToolBox

# FunctionAnnotator

[Home](#) | [Analysis](#) | [Tutorial](#) | [Demo & Benchmark](#)

Job ID	15356438649821
Fasta file	710_trasnscritos_Calanus_finmarchicus.fasta
File size	1,011,586 bytes
Number of Entries	0 entries
Uploaded on	Thu, 30 Aug 18 23:44:24 +0800

Filtered:

[Basic information](#) [Download](#)

Job **15356438649821** is now in queue.  
There are currently **11** jobs before this job.  
You may need to wait for a while.  
Please **DO NOT** submit the same file again.



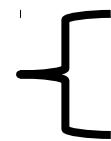
Bioinformatics core, Chang Gung University  
259 Wen-Hwa 1st Road, Kwei-Shan, Tao-Yuan, 333, Taiwan  
TEL: +886-3-2118800 ext.5136 FAX: +886-3-2118122  
If you have any question, please feel free to [e-mail](#) us.

# Aplicación de GO slims

Genomas

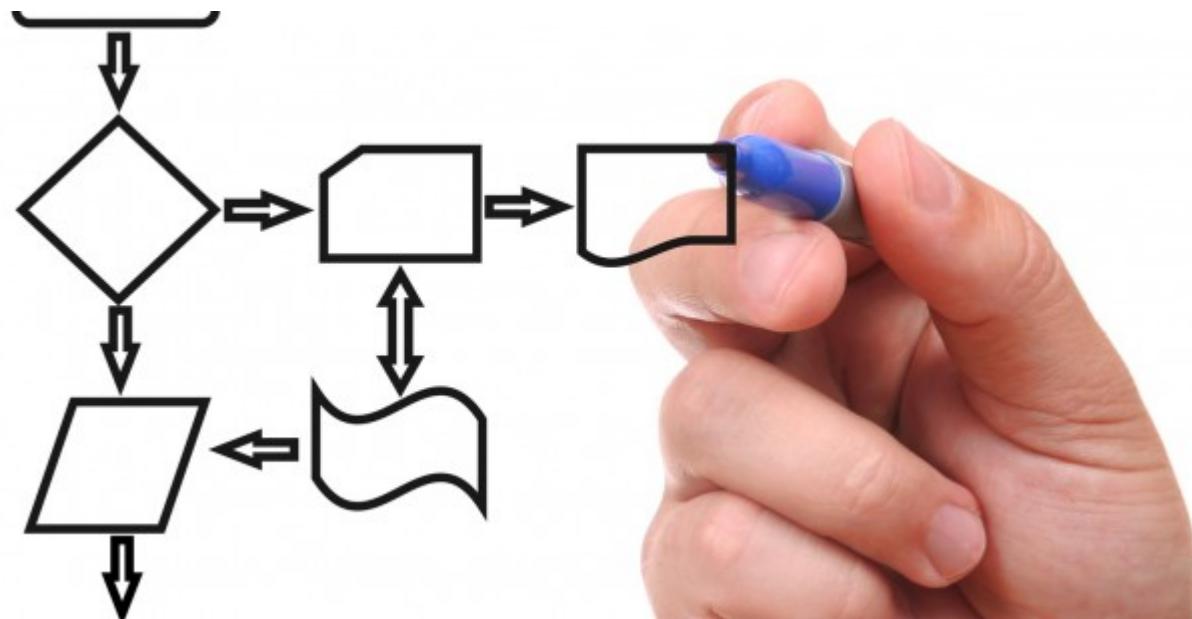
Transcriptomas

Proteomas



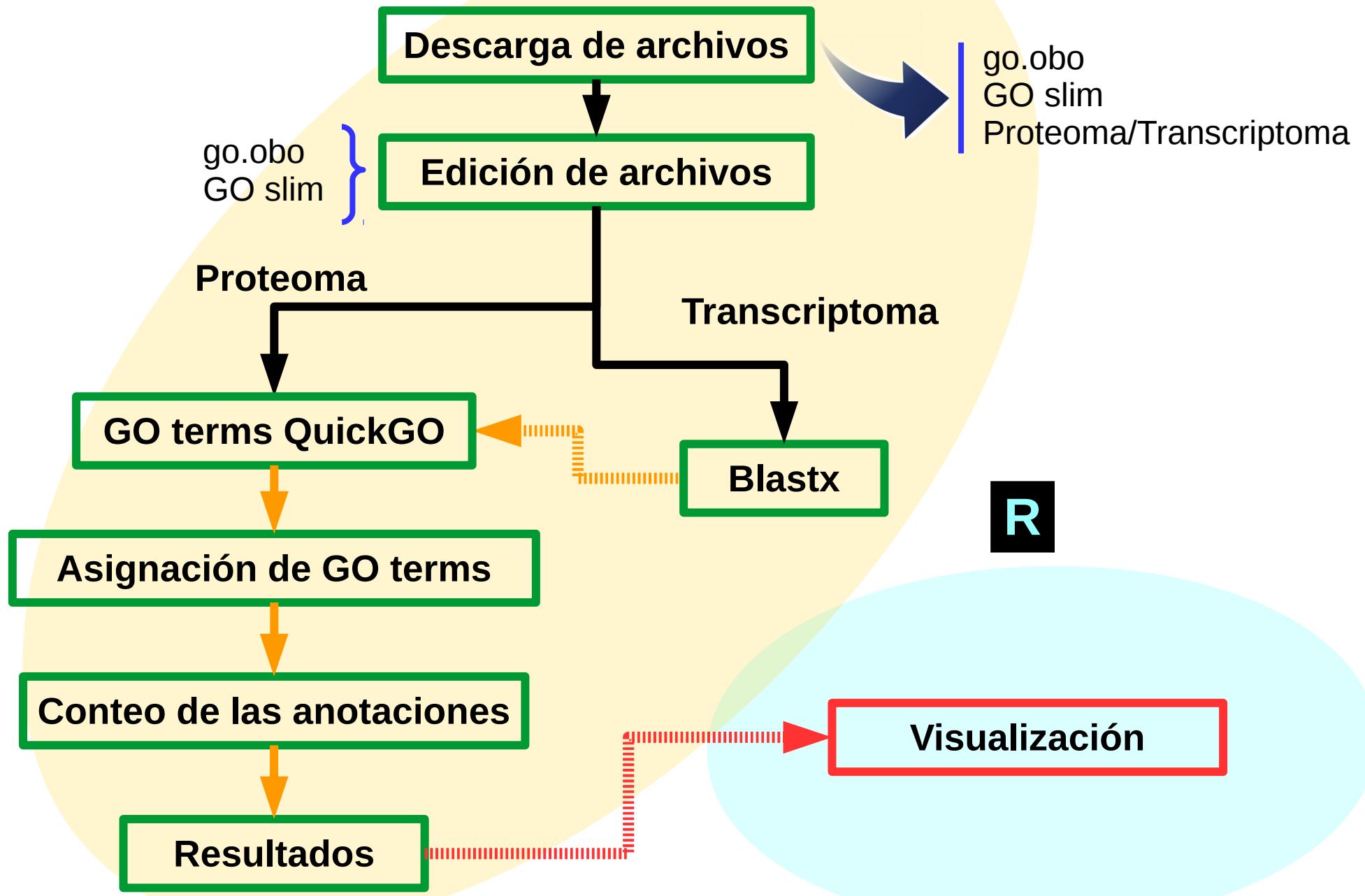
Gene Ontology

# Cómo se hace una anotación



# Anotación Funcional

Python/Linux



# 2 Ejemplos

- Tenemos un conjunto de 300 proteínas obtenidas a partir de un acetiloma en *Zea mays*.
- Tenemos un conjunto de transcritos obtenidos a partir de un transcriptoma.

