



# 2021 International Conference on Human Genetics and Genomics

1-2 December Yazd University



With the participation of University of Algarve, Portugal  
and Sechenov University, Russia

## Towards pathway-specific model organisms with orthology relationships



Sina Majidian

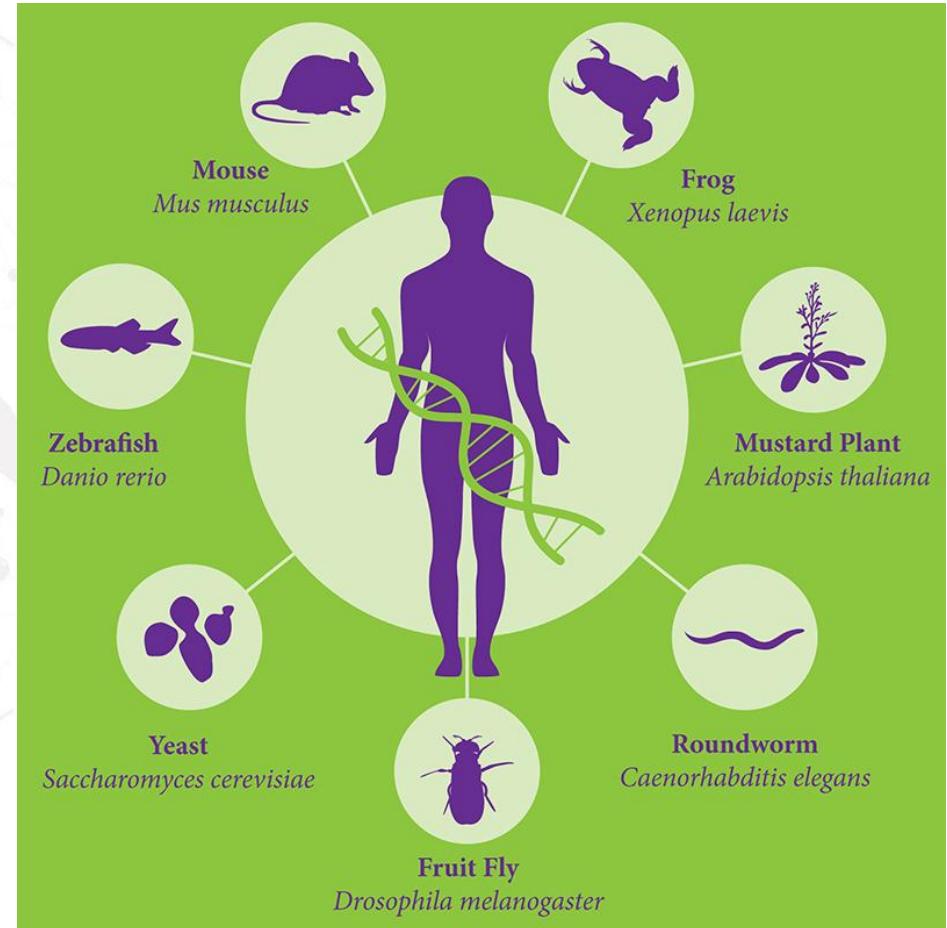


Swiss Institute of  
Bioinformatics

Unil  
University of Lausanne

# Introduction: Model Organisms

- A model organism is a non-human species that is used by scientists to understand biological processes in human cells.



<https://www.yeastgenome.org/blog/support-model-organism-databases>

# Available tools on model organism selection

- 6-9 model organisms
- Use of the models provided raises ethical concerns
- Focus on disease gene/rare gene variant discovery



**ALLIANCE**  
of GENOME RESOURCES



**MARRVEL**<sup>2</sup>



Here is a precomputed example using 6 worm genes modulating dauer inductions, [show](#)

You may run MORPHIN with your choice of pathway/disease database only (check your choice below).

GOBP  KEGG  GWASCAT  DO  OMIM  GAD  HPO

Model Species

Select Your Species ▾

Inparalog score threshold (bigger than or equal to)

\* Default threshold value is recommended.

0 (Very sensitive, Default) ▾

Input a set of model organism genes for a function/pathway/phenotype. (Max=500).

Each gene name must be separated by comma, tab, white space or new line.

Expected calculation time : 10 ~ 20 minutes.

For a test run, you may submit a toy example of each species by selecting 'Examples' below.

\*\*Examples\*\* ▾

# MARRVEL<sup>2</sup>

**i** MARRVEL version 1 is still available at [v1.marrvel.org](https://v1.marrvel.org)

- Human gene  Human protein  Model organism gene  Upload VCF file  Multiple human genes

Human Gene Symbol (official HGNC symbol)

Please use official HGNC Gene Symbol

Human Variant (hg19)

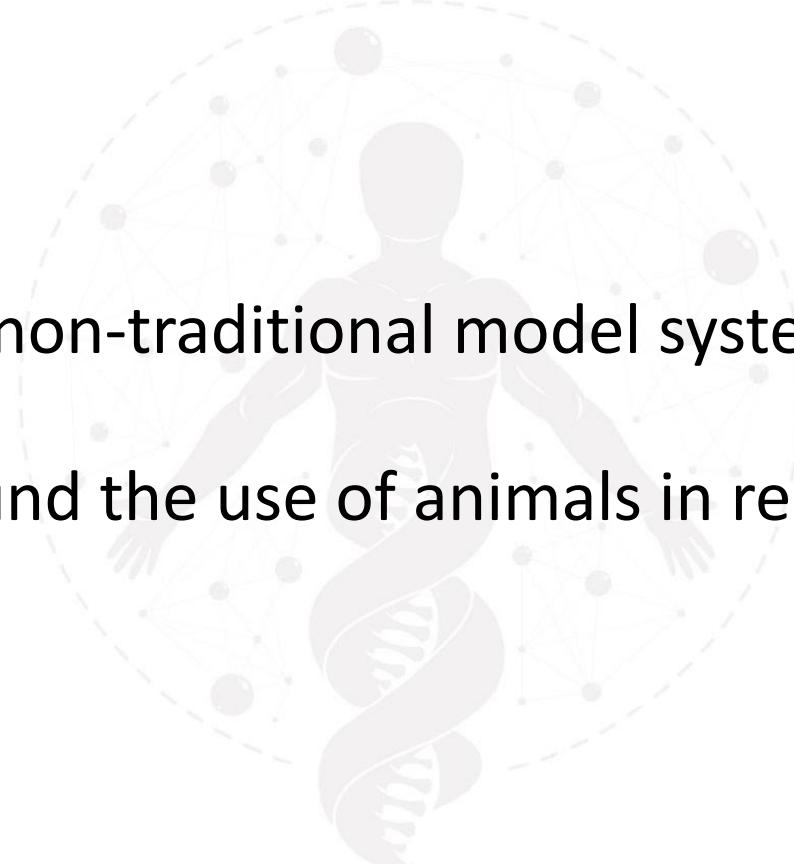
Example: [6:99365567 T>C / FBXL4](#) or [6:99365567 T>C](#) or [FBXL4](#) or [NM\\_012160.3:c.541A>G](#)

Search

Tutorial video

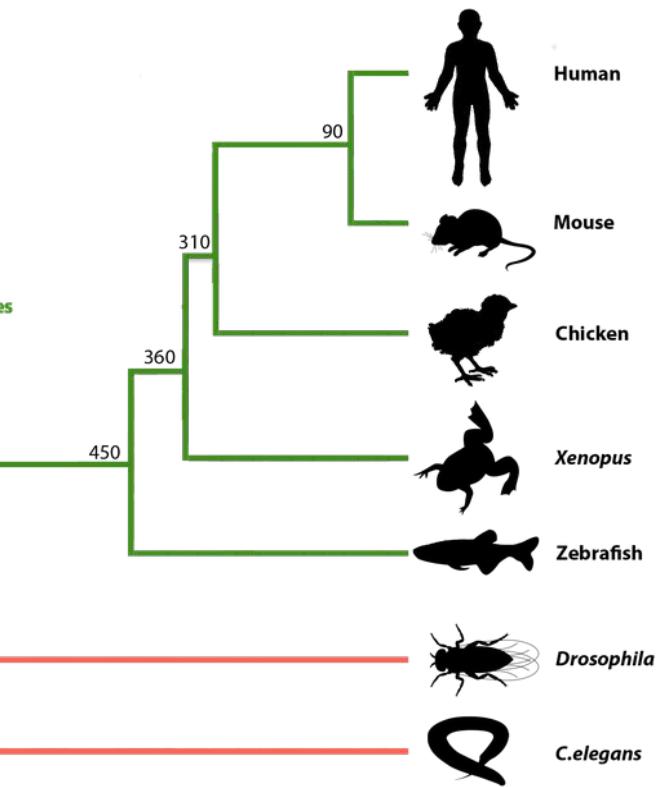
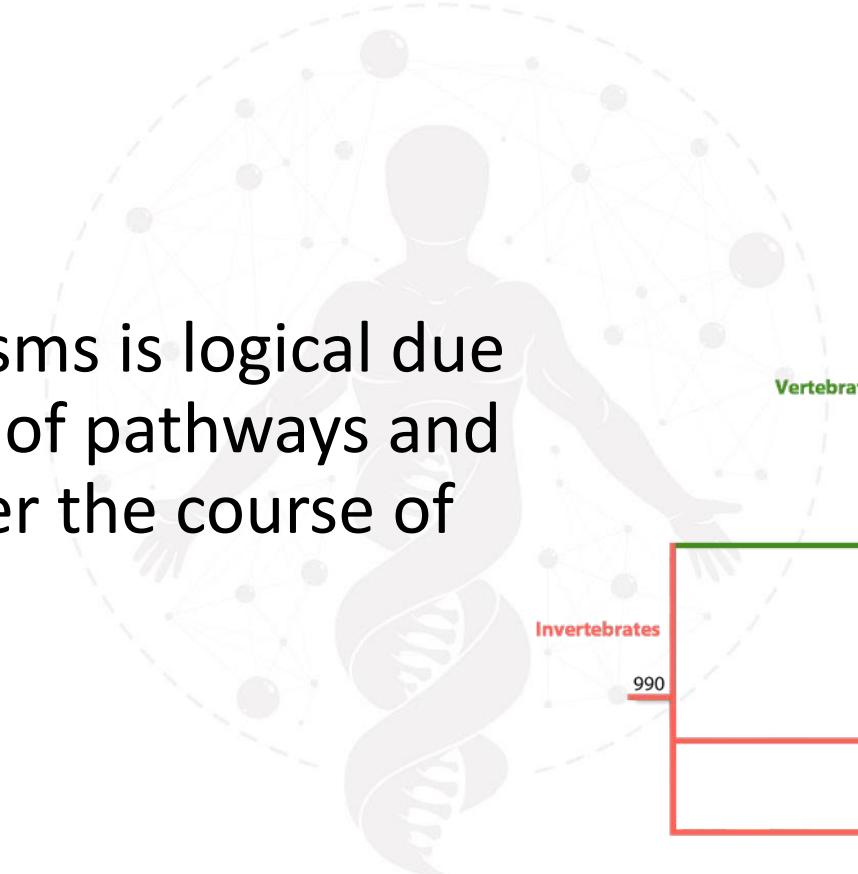
Gene	OMIM			ClinVar			Geno2MP			gnomAD						DGV		GTEX	GOs	
	Gene	MIM Number	#Pheno	#Var	P	LP	LB	R	Hom	Het	HPO	Syn_Z	Syn.o/e	Mis_Z	Mis.o/e	LoF.pLI	LoF.o/e	Total Gain	Total Loss	
<b>OXR1</b>	605609	1	4	36	1	1	3	6	841	511	-0.84	1.09 (0.96 - 1.24)	1.02	0.86 (0.79 - 0.94)	0.84	0.20 (0.12 - 0.94)	88	212	GTEX	<ul style="list-style-type: none"> <li>F molecular_function</li> <li>C nucleoplasm</li> <li>C nucleolus</li> <li>C mitochondrion</li> <li>P response to oxidative stress</li> <li>P adult walking behavior</li> <li>F oxidoreductase activity</li> <li>C intracellular membrane-bound organelle</li> <li>P negative regulation of neuron apoptotic process</li> <li>P neuron apoptotic process</li> <li>P oxidation-reduction process</li> <li>P cellular response to hydroperoxide</li> <li>P negative regulation of peptidyl-cysteine S-nitrosylation</li> <li>P negative regulation of oxidative stress-induced neuron death</li> </ul>
<b>IQSEC1</b>	610166	1	2	21	0	6	13	61	1415	1074	-0.11	1.01 (0.92 - 1.11)	2.28	0.75 (0.70 - 0.81)	1.00	0.10 (0.05 - 0.81)	19	25	GTEX	<ul style="list-style-type: none"> <li>F ARF guanyl-nucleotide exchange factor activity</li> <li>F protein binding</li> <li>C nucleolus</li> <li>C cytoplasm</li> <li>F lipid binding</li> <li>C membrane</li> <li>P actin cytoskeleton organization</li> <li>P regulation of ARF protein signal transduction</li> <li>C intracellular membrane-bound organelle</li> <li>P positive regulation of GTPase activity</li> <li>P positive regulation of keratinocyte migration</li> <li>P positive regulation of adherens junction organization</li> </ul>
<b>WDR27</b>	NA	NA	NA	59	4	2	7	33	2132	1470	-0.04	1.00 (0.89 - 1.13)	-0.59	1.07 (1.00 - 1.16)	0.00	0.67 (0.51 - 1.16)	7	628	GTEX	<ul style="list-style-type: none"> <li>F protein binding</li> <li>C nucleoplasm</li> </ul>
<b>ANKLE2</b>	616062	1	3	25	3	47	73	121	1132	718	-1.18	1.10 (0.99 - 1.22)	0.34	0.96 (0.89 - 1.03)	0.00	0.59 (0.42 - 1.03)	14	130	GTEX	<ul style="list-style-type: none"> <li>F protein binding</li> <li>C endoplasmic reticulum</li> <li>C endoplasmic reticulum membrane</li> <li>P mitotic nuclear envelope reassembly</li> <li>P central nervous system development</li> <li>C membrane</li> <li>F protein phosphatase regulator activity</li> <li>C integral component of endoplasmic reticulum membrane</li> <li>P positive regulation of protein dephosphorylation</li> <li>P negative regulation of phosphorylation</li> <li>P negative regulation of apoptotic process</li> </ul>

# Issues and limitations

- 
1. Lack of databases of non-traditional model systems for research
  2. Ethical concerns around the use of animals in research

# Introduction: Model Organisms

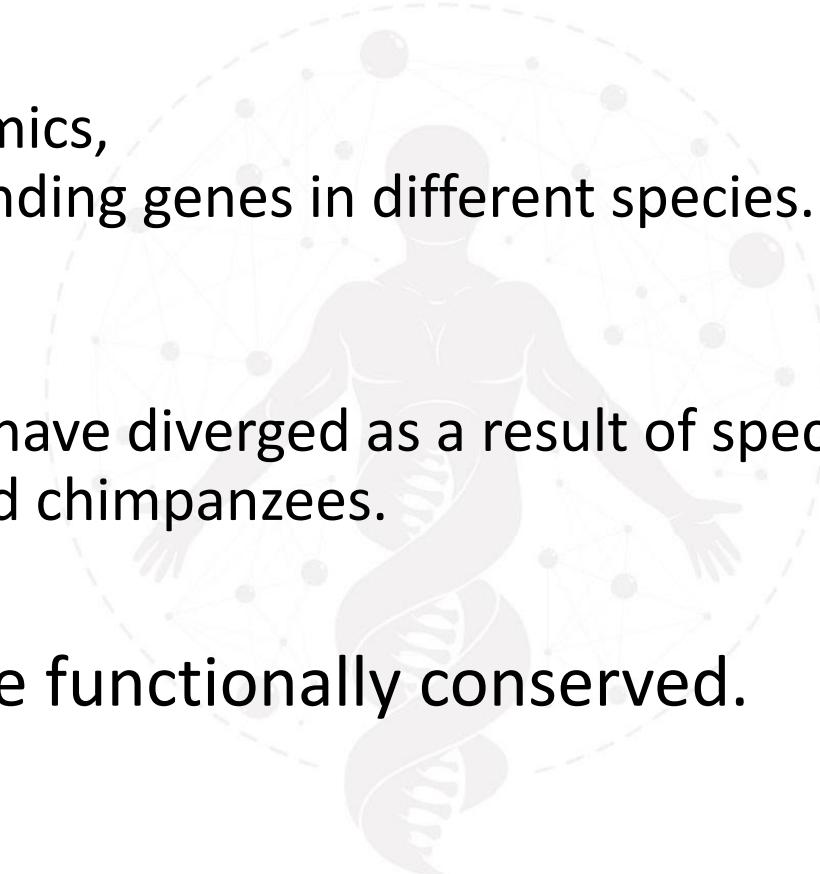
- Using model organisms is logical due to the conservation of pathways and genetic material over the course of evolution.



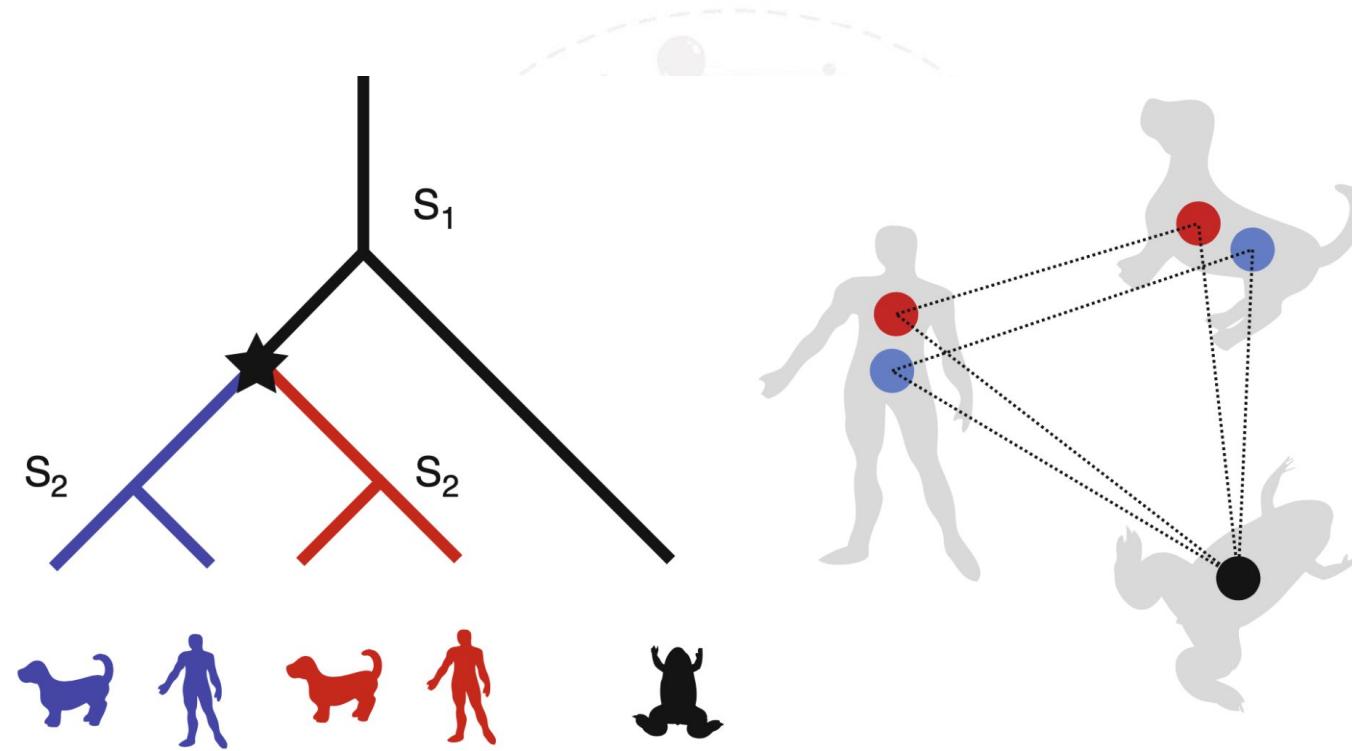
Wheeler & Brändli 2009 Dev Dyn 238:1287-1308.

# Introduction: orthology

- In comparative genomics,
  - relating corresponding genes in different species.
- Orthologs
  - pairs of genes that have diverged as a result of speciation.
  - $\beta$ -Hb in humans and chimpanzees.
- Orthologs tend to be functionally conserved.

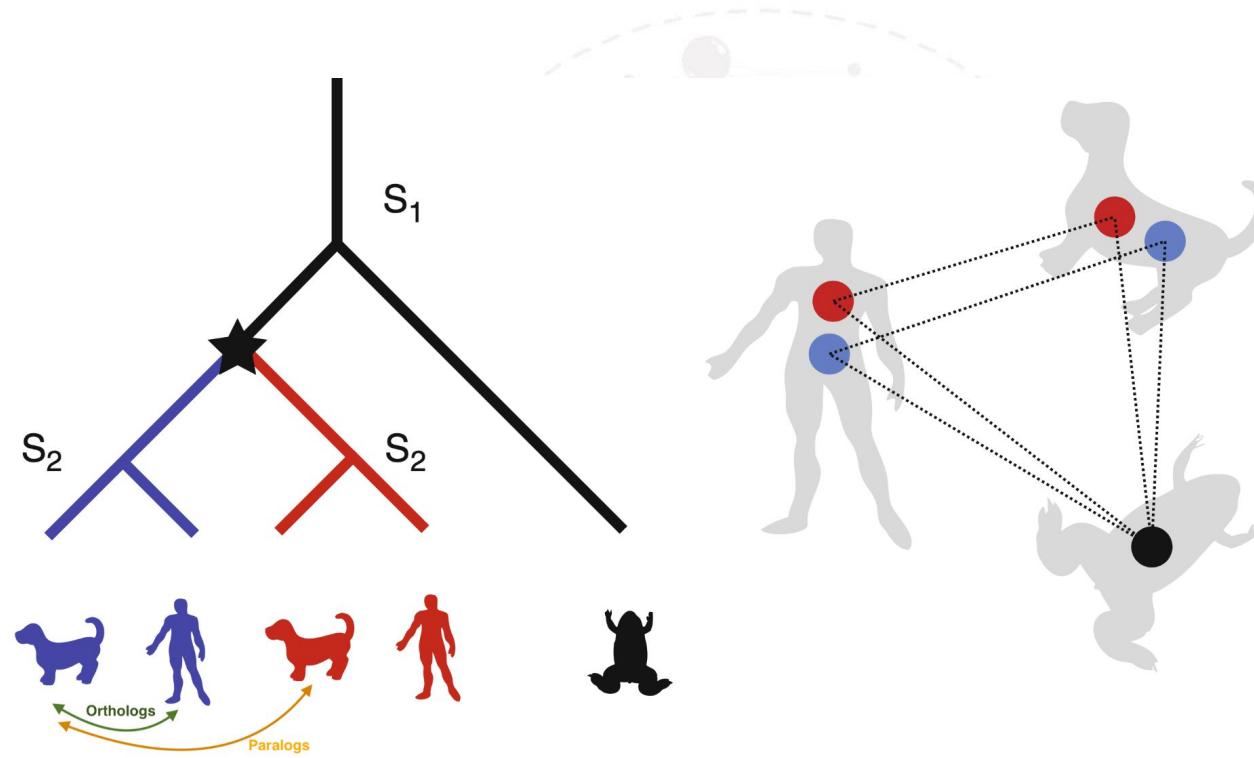


# An evolutionary scenario of a gene family



speciation events (S<sub>1</sub> and S<sub>2</sub>) and a duplication event (star)

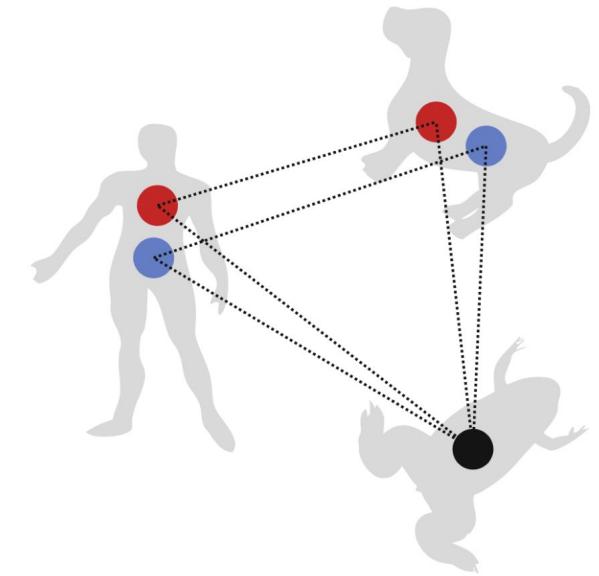
# An evolutionary scenario of a gene family



speciation events (S1 and S2) and a duplication event (star)

# Introduction: orthology

- Orthologs are inferred based on sequence similarity.
- all-against-all alignment between all the genes in all the genomes
- reciprocal best hit approach:
  - orthology relationship between proteins A and B
  - if A is the genome-wide closest relative of B and vice versa



EggNOG 5.0.0

Search protein or OG

Navigation

- Home
- Sequence search
  - eggNOG-mapper v2  
(Batch Functional Annotation)
- Downloads
- API
- Methods

EggNOG v5.0

A database of orthology relationships, functional annotation, and gene evolutionary histories.

Organisms | Viruses | Orthologous Groups | Tree & Algs

5,090 | 2,502 | 4.4M | 4.4M

Search

Citation

**eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses.**

*Jaime Huerta-Cepas, Damian Szklarczyk, Davide Heller, Ana Hernández-Plaza, Sofia K Forslund, Helen Cook, Daniel R Mende, Ivica Letunic, Thomas Rattei, Lars J Jensen, Christian von Mering, Peer Bork*  
Nucleic Acids Res. 2019 Jan 8; 47(Database issue): D309–D314. doi: 10.1093/nar/gky1085

Previous versions

News Contact

Tweets by @eggnogetDb

EggNOG database Retweeted Professor Lesley Hoyles @BugsInYourGuts Very pleased eggNOG-mapper now includes CAZy and BiGG model information in its outputs.

Jun 25, 2020

EggNOG database Retweeted Xavier Grau-Bové @xgraubove Replying to @zolotarg I agree: hmmer-based eggNOG mapper works great to get comprehensive annotations for nonmodel species. If appropriate, you can also choose a taxonomically focused eggNOG reference dataset.

**Oma** | orthologous matrix  
browser

Search all

P53\_RAT | "auxin response factor" | Fungi | Insulin Human



Expand (e.g. to paste a long amino-acid sequence)

**Examples:** Entry HUMAN22169 - Entry P53\_RAT - Search for "photoreceptor" - "Drosophila" species

SCROLL TO DISCOVER MORE



2,424  
Full genomes

17,080,590  
Proteins

1,043,597  
OMA groups

792,770  
Deepest HOGs

All.Apr2021  
Release



## First time here?

Introduction to OMA: See OMA in a nutshell ([Introduction](#)) in OMA doc.

Familiarise yourself with [orthology basics](#): Swiss Orthology , Orthology benchmark.

Familiarise yourself with the [different types of homologs](#) in the OMA doc.

Get ideas about [different types of analyses you could do with OMA](#).

[Start to read about OMA](#)



## OMA tools



### Corona OMA Browser

[Online tool](#)

[Software](#)

### OMA StandAlone

### pyHam

[Python library](#)

### Functional prediction

[Online tool](#)

### Synteny dotplot

[Online tool](#)

### Combined Orthology prediction from OMA & orthoDB

[Ressource](#)

[Discover more tools](#)



## Download options

**Current release:** All data is available for download in a variety of formats.

**API:** The OMA Browser now has a REST API !

**Export All-All:** Export a set of genomes and their precomputed all-against-all comparisons for OMA standalone.

[See more download options](#)



## Latest news

New blog post: OMA Browser Apr2021 is out! 27th release with 2,424 genomes <https://ift.tt/3f9v4Kh>

#newOMAfeature: We identify and use the best-conserved isoform—not necessarily the longest one. Currently, the best...

<https://twitter.com/i/web/status/1331165581043752962>

## Stay in touch



**Oma** | orthologous matrix  
browser

Search all

HBA\_

**HBA\_HUMAN**

HUMAN25352

*Homo sapiens***HBA\_HUMAN**

HUMAN25349

*Homo sapiens***HBA\_LAMPA**

VICPA09091

*Vicugna pacos***HBA\_LATCH**

LATCH02689

*Latimeria chalumnae*



# HOG:A0514501.1a.2a.3a.3a with 7 members (hemoglobin subunit)

Euteleostomi / Sarcopterygii / Tetrapoda / Amniota / Mammalia / Theria / Eutheria / ... / Homininae / [Lower Level ▶](#)

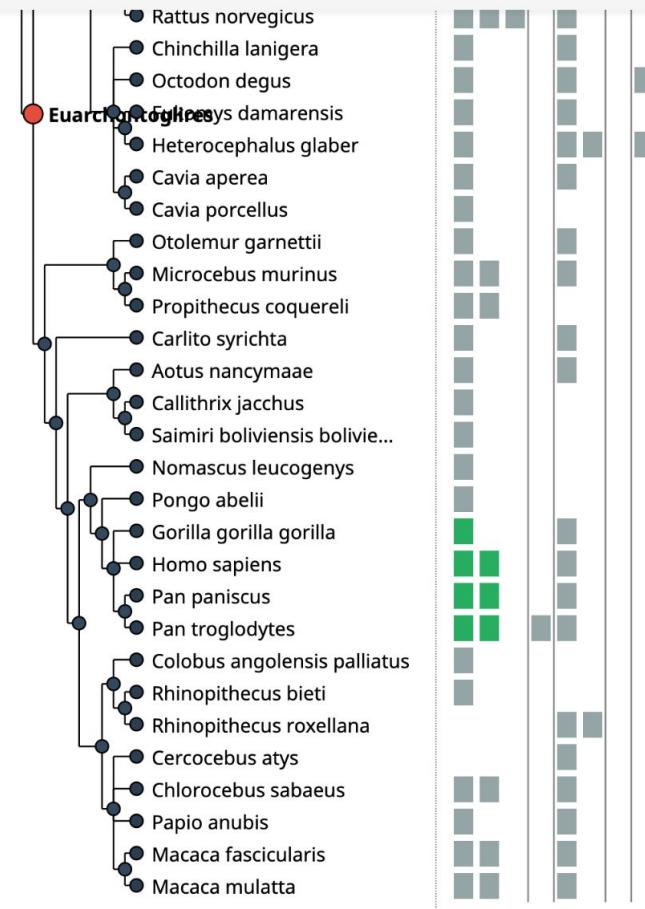
## Graphical viewer

Members

Alignment

Ancestral synteny

Similar HOGs



# Gene HUMAN34043 (BRCA1\_HUMAN)

 Homo sapiens | Breast cancer type 1 susceptibility protein [BRCA1]

Orthologs 78

Paralogs 3

Gene information

GO Annotations

Sequences/Isoforms 32

Local synteny

	-5	-4	-3	-2	-1	0	1	2
HUMAN	33995 >	33997 >	34004 >	< 34005	34014 >	< 34017		
GORG	15983 >	15982 >	15981 >	< 15980	15979 >	< 15978	15977 >	15976 >
PANPA	16570 >	16569 >	16568 >	< 16566	16565 >	< 16559	16556 >	16553 >
PANTR	13813 >	13812 >	13810 >	< 13808	13807 >	< 13802	13796 >	13795 >
PONAB	< 05865	05864 >	05863 >	< 05862	05861 >	< 05860	05859 >	05858 >
NOMLE	16445 >	16446 >	16447 >	< 16448	16449 >	< 16450	16451 >	16452 >
CERAT	29747 >	29748 >	29749 >	< 29750	29752 >	< 29753	29762 >	29767 >
MACFA	14277 >	14278 >	14279 >	< 14281	14283 >	< 14287	14293 >	14294 >
MACMU	09983 >	09984 >	09985 >	< 09986	09987 >	< 09988	09990 >	09991 >

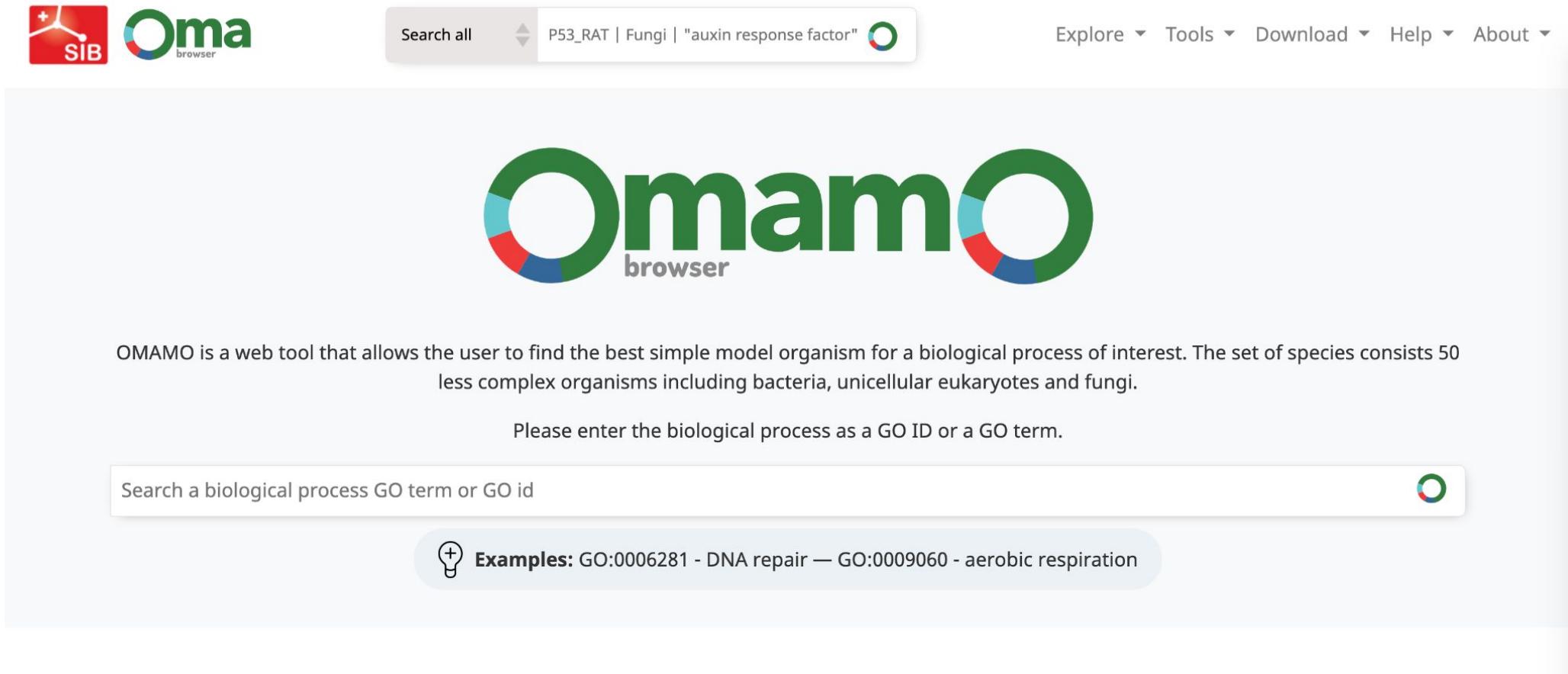
# Goal of project

Create a database and software

that allows to query a **pathway**

and provide the user with  
the best-suited simple **model organism**.

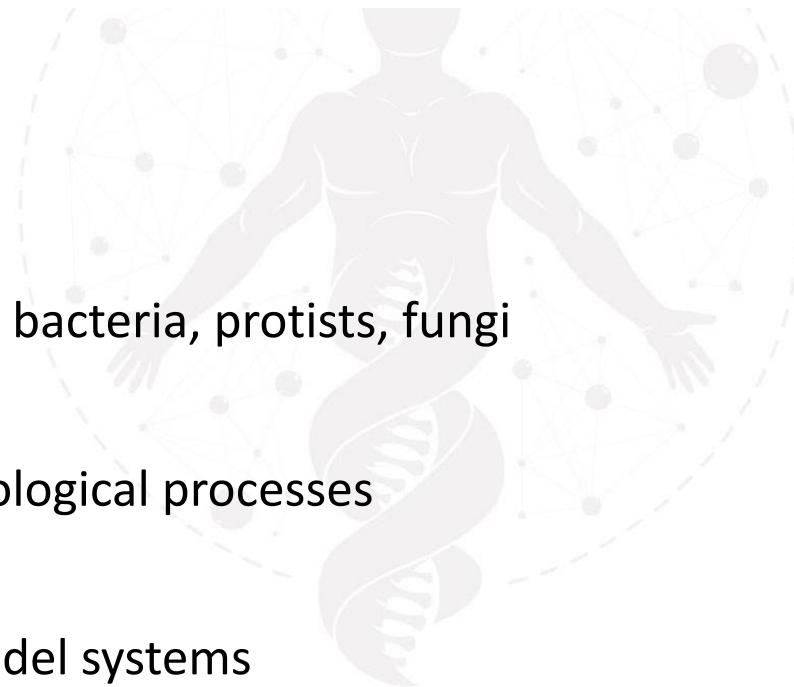
# omabrowser.org/omamo



The screenshot shows the homepage of the OmamO browser. At the top left is the SIB logo (a red square with a white 'S' and 'IB' inside). Next to it is the Oma browser logo, which consists of a green 'O' with a red 'M' inside, followed by the word 'browser'. In the center is a search bar with the text "Search all" and a dropdown menu showing "P53\_RAT | Fungi | 'auxin response factor'". To the right of the search bar are navigation links: "Explore", "Tools", "Download", "Help", and "About". Below the header is a large, stylized "OmamO" logo where the 'O's are filled with a circular gradient of red, blue, and green. Underneath the logo, the word "browser" is written in a smaller, gray sans-serif font. A descriptive text block below the logo states: "OMAMO is a web tool that allows the user to find the best simple model organism for a biological process of interest. The set of species consists 50 less complex organisms including bacteria, unicellular eukaryotes and fungi." Below this text is a placeholder message: "Please enter the biological process as a GO ID or a GO term." At the bottom of the search bar is a text input field containing "Search a biological process GO term or GO id" and a small green search icon. To the left of the input field is a lightbulb icon with a plus sign, followed by the text "Examples: GO:0006281 - DNA repair — GO:0009060 - aerobic respiration".



- 50 species
- All organisms are simple, e.g. bacteria, protists, fungi
- Versatile - can query 4620 biological processes
- Potential to discover new model systems



Screenshot of a GitHub repository page for 'DessimozLab / omamo'.

**Header:** Search or jump to... / Pull requests / Issues / Marketplace / Explore / Unwatch (6) / Unstar (1) / Fork (1)

**Repository Information:** DessimozLab / omamo (Public)

**Navigation:** Code (selected), Issues, Pull requests, Actions, Projects, Wiki, Security, Insights, Settings

**Code Overview:** master branch, 1 branch, 0 tags

**Commit History:**

Commit Message	Author	Date	Commits
sina add link to preprint	bc9312a	7 days ago	9 commits
data	first commit		last month
README.md	add link to preprint		7 days ago
logo-omamo.jpg	first commit		last month
omamo_base.py	first commit		last month
omamo_dataframe.py	first commit		last month

**README.md Content:**

OMAMO: orthology-based model organism selection



**About:** OMAMO: orthology-based model organism selection

**Readme:**

**Releases:** No releases published. Create a new release

**Packages:** No packages published. Publish your first package

**Languages:** Python 100.0%



# OmamO browser

OMAMO is a web tool that allows the user to find the best simple model organism for a biological process of interest. The set of species consists 50 less complex organisms including bacteria, unicellular eukaryotes and fungi.

Please enter the biological process as a GO ID or a GO term.

DNA rep



- GO:0006274 - DNA replication termination
- GO:0006275 - regulation of DNA replication
- GO:0006279 - premeiotic DNA replication
- GO:0006281 - DNA repair
- GO:0006282 - regulation of DNA repair
- GO:0006307 - DNA dealkylation involved in DNA repair
- GO:0006335 - DNA replication-dependent nucleosome assembly
- GO:0006336 - DNA replication-independent nucleosome assembly
- GO:0008156 - negative regulation of DNA replication
- GO:0010213 - non-photoreactive DNA repair


**Examples:** GO:0006281 - DNA repair — GO:0009060 - aerobic respiration

## GO:0006281 - DNA repair



	Domains	Species	Taxon	No. of orthologs	Average func. similarity	Score
+	E	SCHPO	<b><i>Schizosaccharomyces pombe</i></b> (strain 972 / ATCC 24843)	131	0.4884 ± 0.2033	63.98
+	E	SCHJY	<b><i>Schizosaccharomyces japonicus</i></b> (strain yFS275 / FY16936)	126	0.4941 ± 0.2211	62.26
+	E	SCHCR	<b><i>Schizosaccharomyces cryophilus</i></b> (strain OY26 / ATCC MYA-4695 / CBS 11777 / NBRC 106824 / NRRL Y48691)	123	0.4945 ± 0.2124	60.82
+	E	SCHOY	<b><i>Schizosaccharomyces octosporus</i></b> (strain yFS286)	121	0.5013 ± 0.2086	60.65
+	E	DICPU	<b><i>Dictyostelium purpureum</i></b>	112	0.5252 ± 0.2549	58.82



# SCHPO - *Schizosaccharomyces pombe* (strain 972 / ATCC 24843)

Proteome version: Ensembl Fungi 22; ASM294v2; 17-MAR-2014 with 5087 proteins.

Eukaryota / Opisthokonta / Fungi / Dikarya / Ascomycota / Taphrinomycotina / Schizosaccharomycetes / Schizosaccharomycetales / Schizosaccharomycetaceae / **Schizosaccharomyces**

## Genome information

List genes

Related Species >

Synteny

Common name: Fission yeast

UniProt identifier: SCHPO

NCBI Taxonomy ID: 284812

DB Release: Ensembl Fungi 22; ASM294v2; 17-MAR-2014

Information [http://fungi.ensembl.org/schizosaccharomyces\\_pombe/Info/Index](http://fungi.ensembl.org/schizosaccharomyces_pombe/Info/Index)

Number of genes: 5087

Number of proteins (all isoforms) 5087

Proteins in OMA groups: 4675

Proteins in HOGs: 4785



## GO:0006281 - DNA repair

grid
down

	Domains	Species	Taxon	No. of orthologs	Average func. similarity	Score
-	E	SCHPO	<b>Schizosaccharomyces pombe</b> (strain 972 / ATCC 24843)	131	0.4884 ± 0.2033	63.98

**Human genes:** RAD1 NSMCE4A RPA4 NAP1L4 XRCC6 ENDOV PNKP ERCC5 CHEK2 GTF2H2C CUL4B MSH3 RAD51B NTHL1 RPA2 ERCC4 NSMCE1 UPF1 GINS2 DDX11 GTF2H4 DNNT POLE CUL4A ASCC3 RAD50 APEX2 ERCC8 VCP RPA1 RFC5 TDP1 GLTSCR2 GTF2H3 HUS1B POLD1 MCM3 PCNA SEM1 HUS1 ERCC2 RAD54B DNA2 RAD54L PSMD14 KAT5 ACTR5 SMC5 SUPT16H PDS5B RAD17 EXO5 CDC45 POLR2I ACTL6A MMS19 POLR3A ERCC1 UBE2B MCM5 MCM4 FEN1 GINS4 MCM7 ACTR8 CSNK1E UBE2A RUVBL2 MCM2 INO80 LIG1 UBE2V2 MORF4L1 SMC1A ATR PMS2 RFC2 XRCC5 DDB1 SSRP1 MUS81 RUVBL1 FAN1 POLE2 MSH6 MCM6 TRRAP SMC6 MRE11 ERCC3 SLX1A UBE2V1 XAB2 RAD23A DMC1 MUTYH PRPF19 PARK7 POLA1 POLD2 XPA MSH2 FH LIG4 RAD51 H2AX RAD23B WDHD1 RFC1 RFC4 JUNB UNG SMC3 RAD9B UBE2N SPO11 POLM POLG GTF2H2 RFC3 RAD9A MLH1 PDS5A H2AW UBE2NL EID3

**SCHPO genes:** pku80 tfb4 rpb9 hus1 TYDP1\_SCHPO nse4 pcu4 rad1 rhp14 rfc3 cds1 mug81 pob3 fum1 alp13 hta2 rad50 YGNB\_SCHPO hsp3106 mcl1 mip1 dmc1 psm1 dpb2 spm2 mcm4 rad32 rhp57 upf1 mcm7 mcm6 mus81 ddb1 msh6 dna2 arp5 rad15 rfc4 rhp23 pms1 ung1 pol2 rad3 mcm2 ubc13 rad13 ARP8\_SCHPO rec12 cdc17 pol4 rrp16 cwf3 rvb1 mst1 rpn11 hhp1 rpc1 pcn1 nth1 YDA6\_SCHPO smc6 apn2 tra1 rfc2 ssb1 mms19 pds5 swi10 rpn15 sld5 spt16 rnh201 prp19 sna41 msh2 ptr8 nse1 chl1 pol1 pol3 psm3 slx1 tfb2 rad2 cdc1 mcm5 alp5 msh3 pku70 hta1 rhp6 rhp54 mcm3 YAMB\_SCHPO rad17 lig4 ino80 smc5 ssb2 mlh1 rdh54 rhp51 rvb2 ssl1 psf2 cdc48 pnk1 rad9 FAN1\_SCHPO myh1 rfc1 rad16 exo5 ckn1 rfc5

# Gene HUMAN100076 (A6NMQ1)

 Homo sapiens | DNA polymerase [POLA1]

Groups ▾

Genome

## General Information

Orthologs 520

Paralogs 3

### Gene information

GO Annotations

Sequences/Isoforms 4

Local synteny

Description DNA polymerase alpha 1, catalytic subunit [Source:HGNC Symbol;Acc:HGNC:9173]; transcript\_id=ENST00000379068.8

Organism HUMAN - Homo sapiens

Locus [X]: 24693962 ... 24995950

Number of exons 37

Exons join(24693962..24694004, 24699425..24699549, 24703251..24703347, 24704389..24704469, 24714554..24714669, 24715141..24715203, 24716362..24716454, 24716884..24716971, 24717290..24717491, 24717580..24717758, 24723155..24723267, 24724335..24724451, 24725981..24726055, 24726933..24727071, 24727782..24727936, 24732370..24732454, 24733755..24733816, 24735399..24735488, 24737625..24737741, 24739375..24739550, 24741375..24741504, 24742002..24742121, 24743230..24743329, 24745418..24745542, 24748311..24748460, 24748870..24748992, 24809898..24809930, 24810708..24810800, 24812658..24812863, 24814979..24815111, 24821452..24821583, 24826427..24826601, 24841652..24841830, 24843546..24843677, 24888006..24888122, 24930453..24930549, 24995805..24995950)

## IDs and Cross-references

UniProtKB/TrEMBL ★ A6NMQ1

Ensembl Protein ENSP00000368358.3

Ensembl Gene ENSG00000101868.13

Ensembl Transcript ENST00000379068 ENST00000379068.8

RefSeq NP\_001317289 NP\_001365232 XP\_005274609 XP\_011543842 XP\_016885083 XP\_016885084 XP\_016885085

EntrezGene 5422

# Gene HUMAN100076 (A6NMQ1)

 Homo sapiens | DNA polymerase [POLA1]

Groups ▾

Genome

Protein sequences and cDNA of the selected isoforms are displayed below the tables

Search



Orthologs 520

Paralogs 3

Gene information

GO Annotations

Sequences/Isoforms 4

Local synteny

	Protein ID	Sequence length	Exons	Domain Architectures	Number exons	Cross reference
<input type="radio"/>	HUMAN100075	96		n/a	4	 A0A5F9ZHW2
<input type="radio"/>	HUMAN100077	563			15	 ENSG00000101868.13
<input type="radio"/>	HUMAN100078	1462			37	 DPOLA_HUMAN
<input checked="" type="radio"/>	HUMAN100076 reference isoform	1468			37	 A6NMQ1

## Protein Sequence

Copy to clipboard

Download:  Fasta

```

MAPVHGDDCE IGASALSDSG SFVSSRARRE KKSKKGRQEA LERLKKAKAG EKYKYEVEDF 60
TGVYEEVDEE QYSKLVQARQ DDDWIVDDDG IGYVEDGREI FDDDLEDDAL DADEKGKDGGK 120
ARNKDKRNVK KLAVTKPNNI KSMFIACAGK KTADKAVIDLS KDGLLGDIHQ DLNTETPQIT 180
PPPVMILKKK RSIGASPNNP SVHTATAVPS GKIASPVSRK EPPLTPVPLK RAEFAGDDVQ 240
VESTEEEQES GAMEFEDGDF DEPMEVVEVD LEPMAAKAWD KESEPAEEVK QEADSGKTV 300
SYLGSFLPDV SCWDIDQEGD SSFSVQEYQV DSSHLPVLKG ADEEQVFHFY WLDAYEDQYN 360

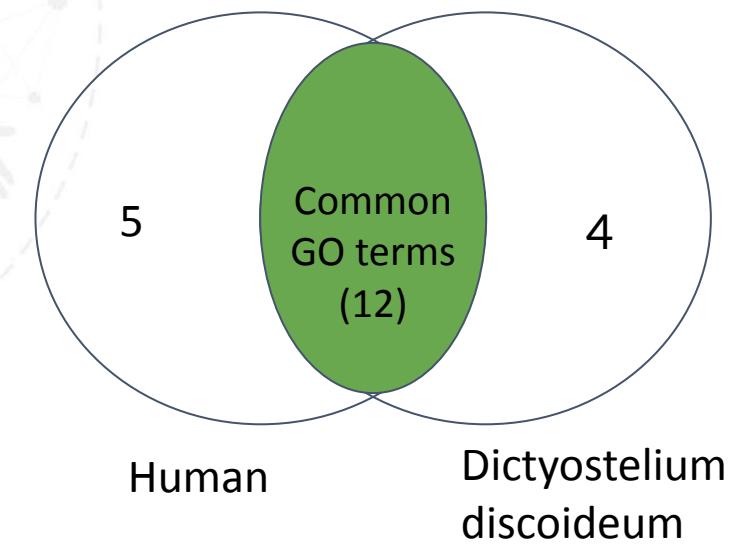
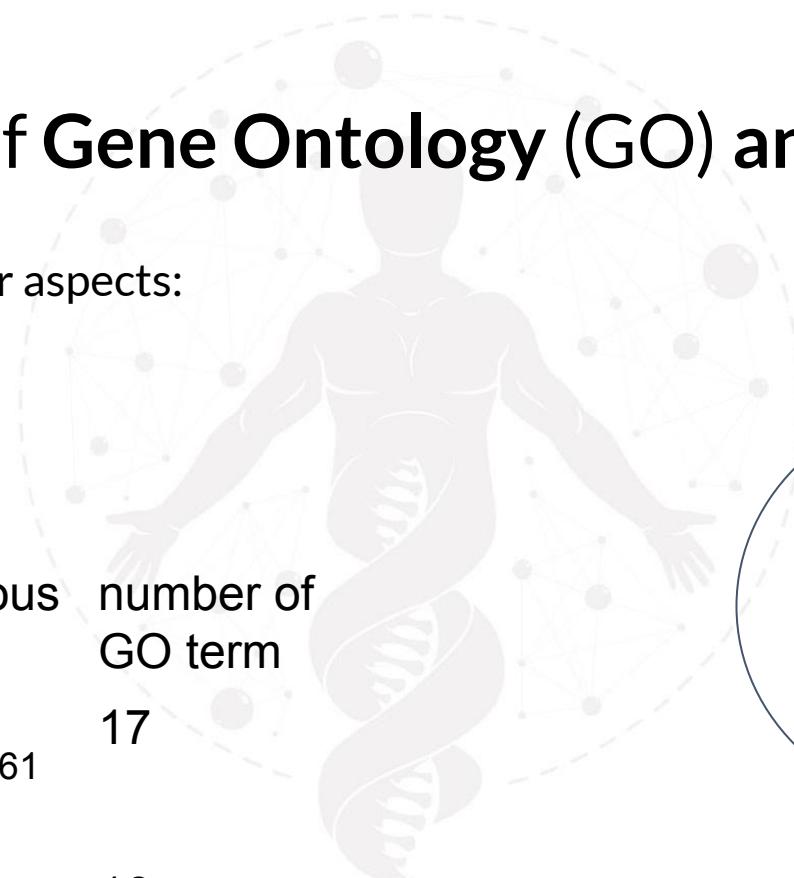
```

# Methods

Each gene has a set of Gene Ontology (GO) annotations

- They are split into three major aspects:
  - Cellular component
  - Biological process
  - Molecular function

Species	orthologous proteins	number of GO term
Human	Transport protein Sec61	17
Dictyostelium discoideum	Transport protein Sec61	16



# Information content-based similarity measure

$$\text{Similarity} = \frac{\text{Intersection}}{\text{Union}}$$

Intersection = IC stored by **overlapping GO terms**

Union = IC stored by all GO terms in both genes

# Discussion

'Dictyostelium is a powerful model for investigating neutrophil **chemotaxis** and the molecular mechanisms that regulate this process'

*Williams et al., Trends in molecular medicine, Cell Press, 2006.*

	Species	No. of ogs	Average % func. similarity	Score
<b>1</b>	DICDI	21	25.05 +/- 13.27	5.26
<b>2</b>	BACTN	1	53.39 +/- 0.00	0.53
<b>3</b>	CAUVC	1	53.39 +/- 0.00	0.53
<b>4</b>	CAUVN	1	53.39 +/- 0.00	0.53

# Discussion

*'Paramecium* has a prominent secretory pathway that makes it a suitable model to study regulated secretion ..exocytosis.'

*Beisson et al., Cold Spring Harbor Protocols, 2010.*

	Species	No. of OGs	Average % func. similarity ± st. dev.	Score
1	PARTE	118	$28.83 \pm 17.67$	34.02
2	DICDI	51	$37.53 \pm 20.03$	19.14
3	EMEND	37	$42.88 \pm 19.27$	15.87
4	EMENI	33	$47.17 \pm 21.37$	15.57

# Future steps

- Considering several more species in our database
- improving the scoring system
  - sequence similarity
  - conservation of protein structure
  - reproduction time
- providing models for plants and animals
- Collaborations for validating results
- <https://lab.dessimoz.org/research>

Thanks for your attention !