

BIOINFORMÁTICA BANCOS DE DADOS

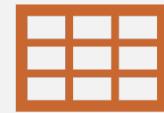
Prof. Leandro Martins de Freitas, PhD
IMS/UFBA



BANCO DE DADOS BIOLÓGICOS



Grande quantidade de dados



Dados organizados

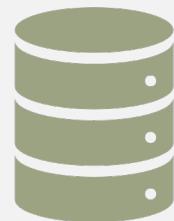
BANCO DE DADOS BIOLÓGICOS



Update



Pesquisa



Recuperação de
dados

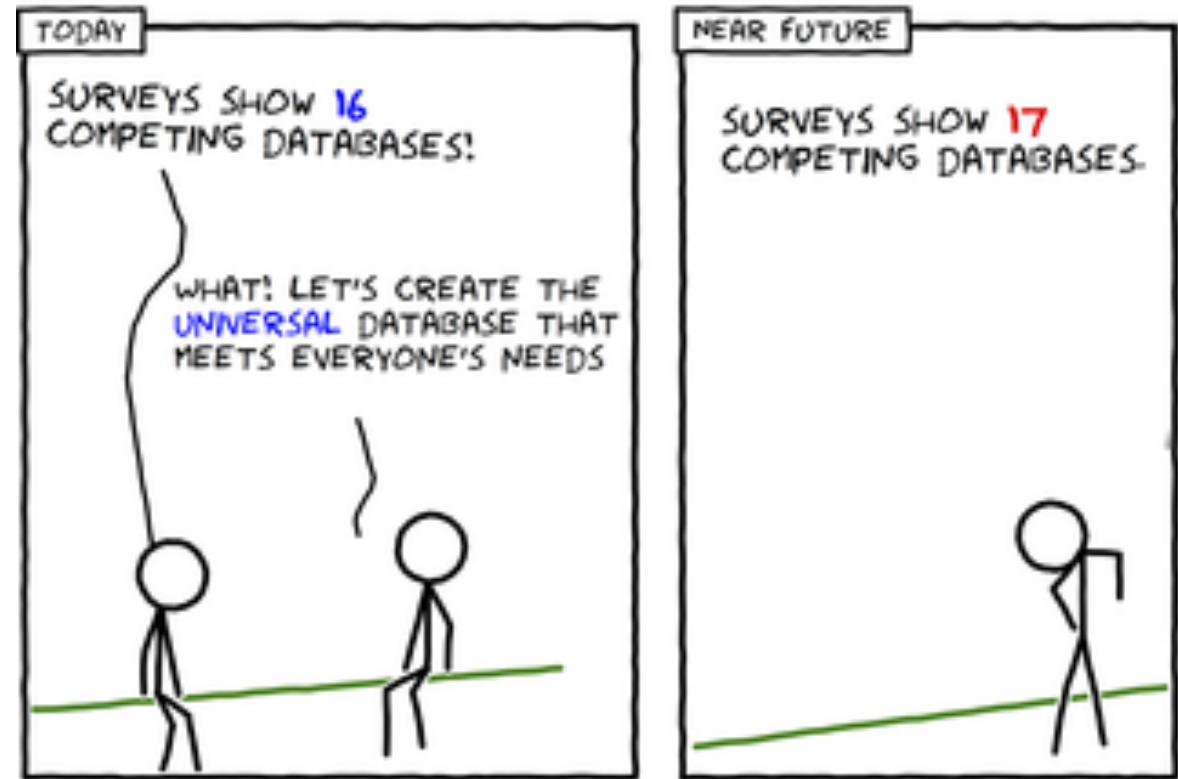
BANCO DE DADOS





CRESCIMENTO DOS BANCOS DE DADOS

HOW BIOLOGICAL DATABASES PROLIFERATE



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

EDITORIAL

Ten Simple Rules for Developing Public Biological Databases

Mohamed Helmy, Alexander Crits-Christoph, Gary D. Bader

Published: November 10, 2016 • <https://doi.org/10.1371/journal.pcbi.1005128>

Activities Table BANCO DE DADOS – NÃO RELACIONAL

Student	Activity1	Cost1	Activity2	Cost2
John Smith	Tennis	\$36	Swimming	\$17
Jane Bloggs	Squash	\$40	Swimming	\$17
John Smith	Tennis	\$36		
Mark Antony	Swimming	\$15	Golf	\$47

ORACLE®



BANCO DE DADOS – RELACIONAL

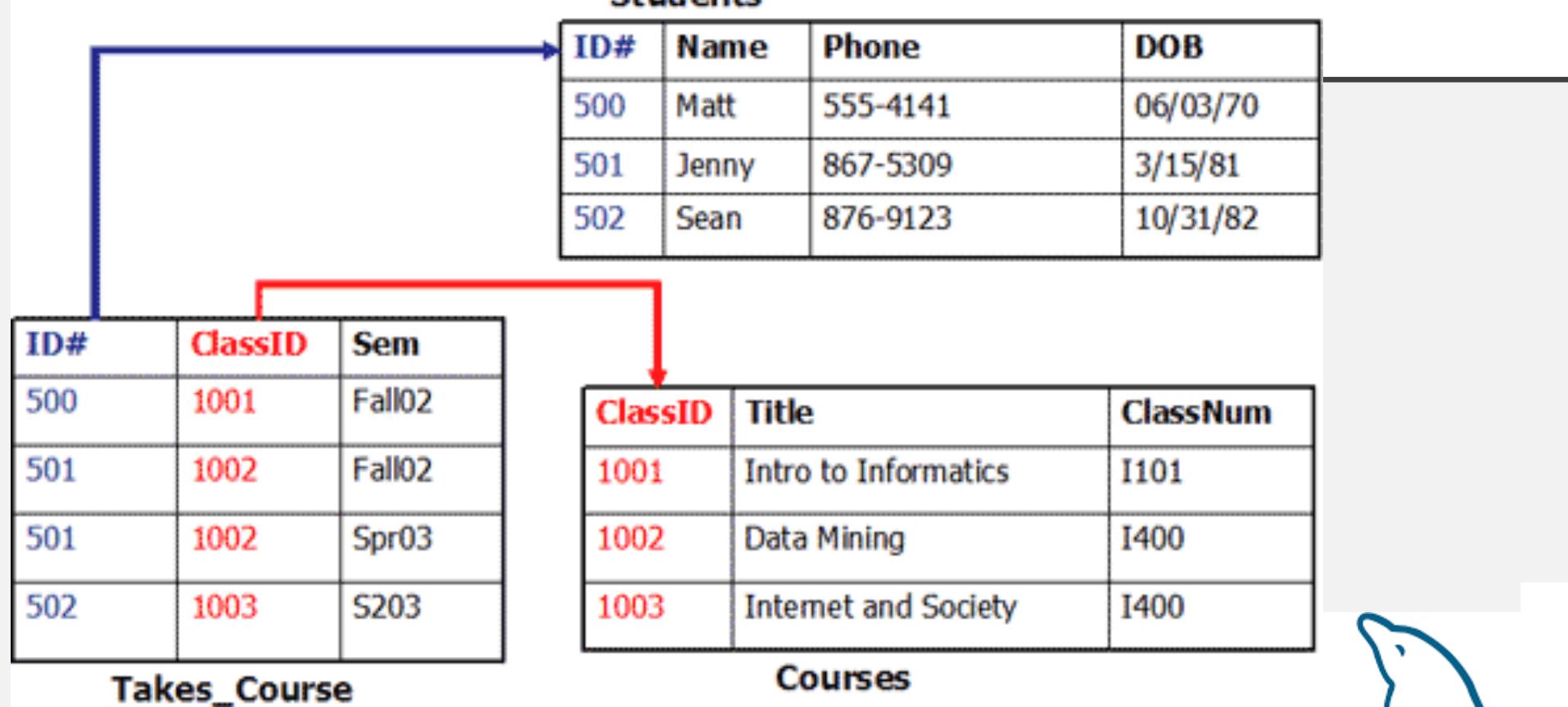
- Sem esquema fixo.
- Capacidade para lidar com grandes dados melhor.
- Menor custo do servidor.

The Oracle logo, featuring the word "ORACLE" in a bold, red, sans-serif font with a registered trademark symbol.

ORACLE®



DANOS DE DADOS RELACIONAL



ORACLE® MySQL®

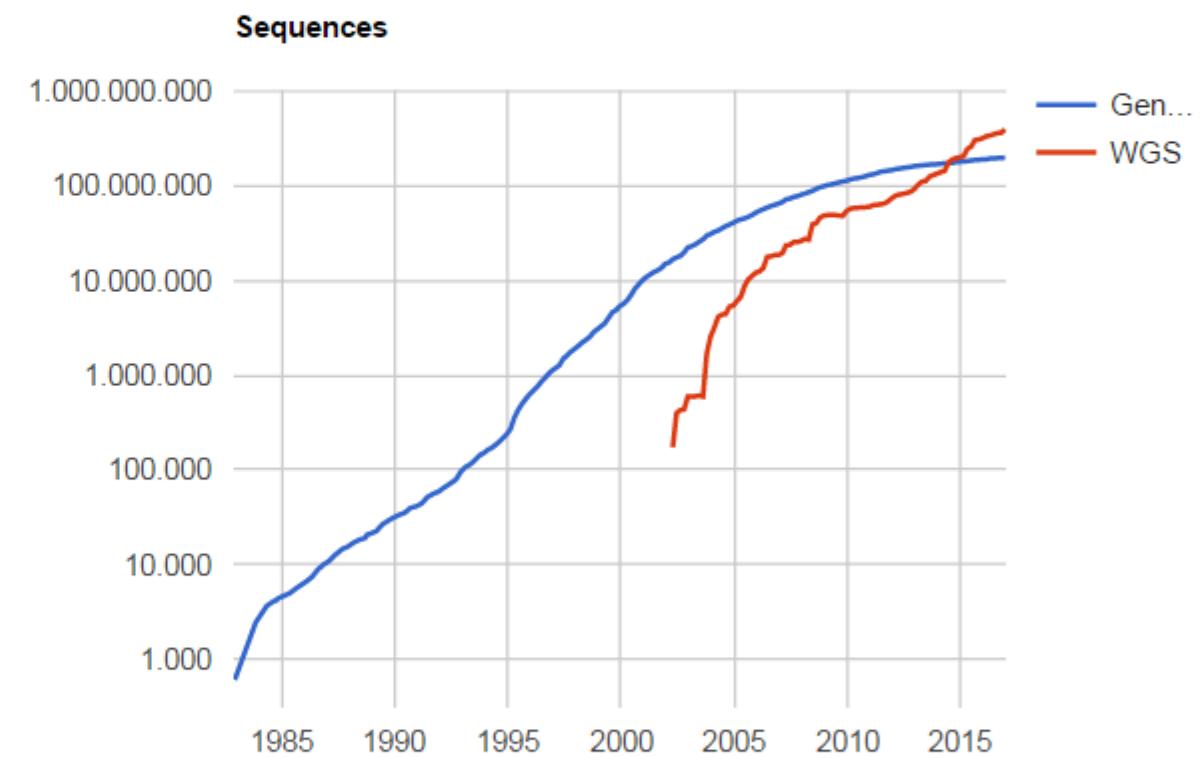
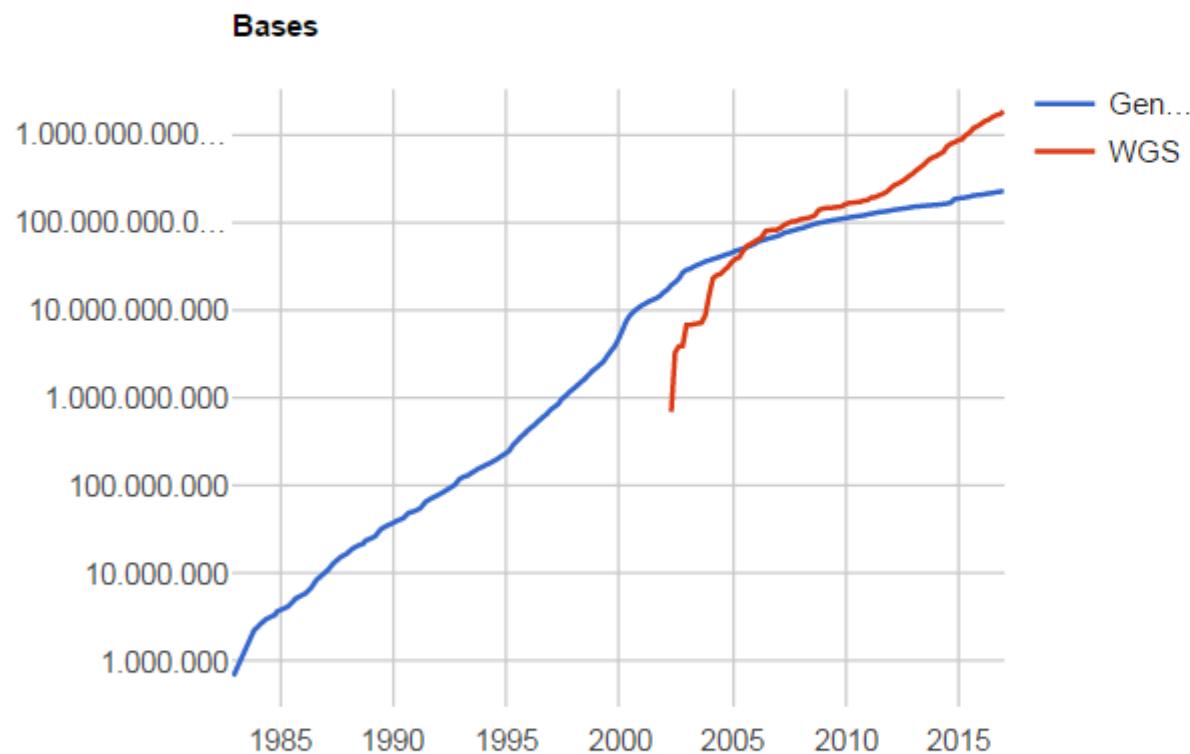


BANCO DE DADOS - RELACIONAL

- Evita a duplicação de dados.
- Evita registros inconsistentes.
- Mais fácil de mudar dados.
- Mais fácil de alterar o formato de dados.
- Os dados podem ser adicionados e removidos facilmente.

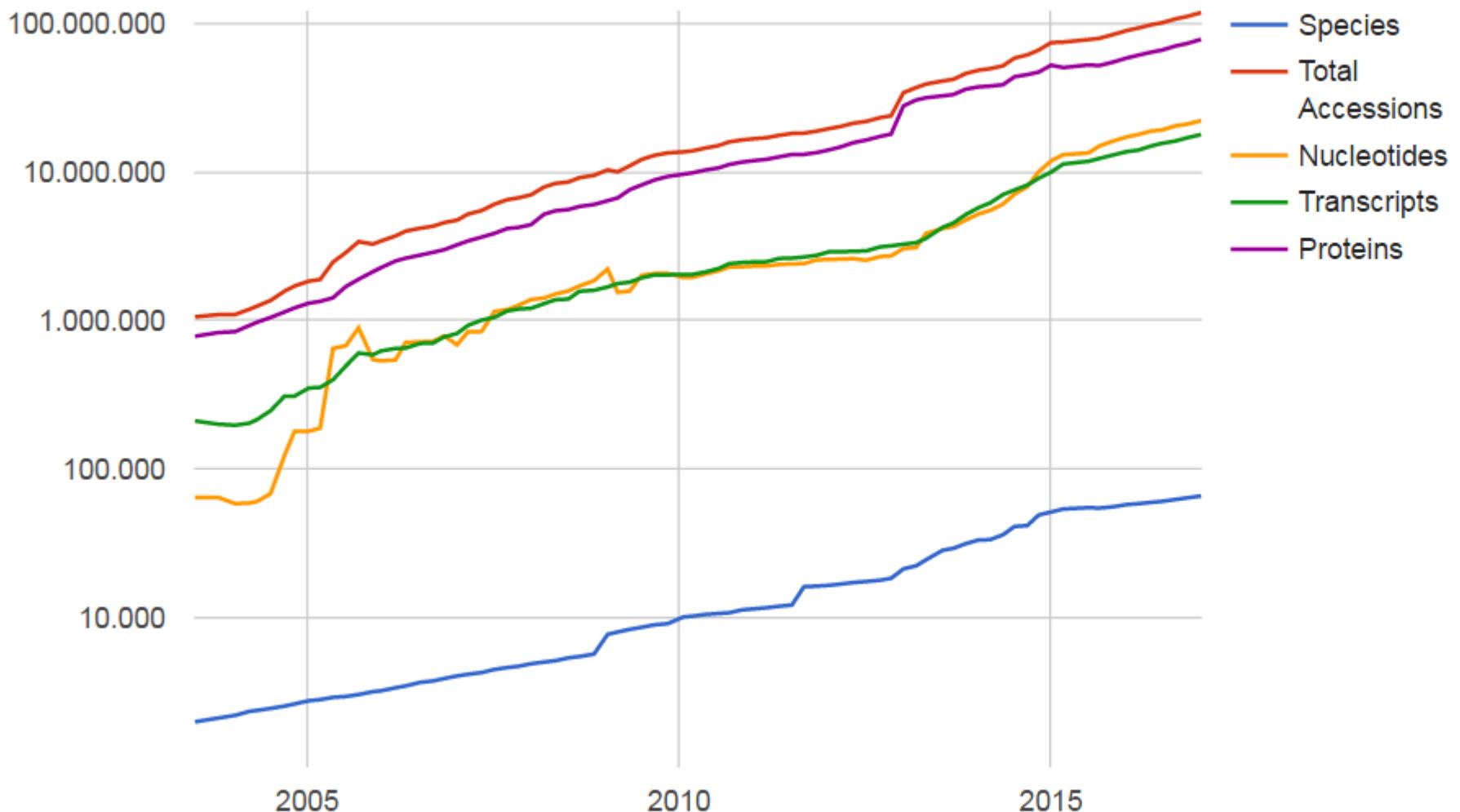
The Oracle logo, featuring the word "ORACLE" in a bold, red, sans-serif font with a registered trademark symbol.

GenBank and WGS Statistics

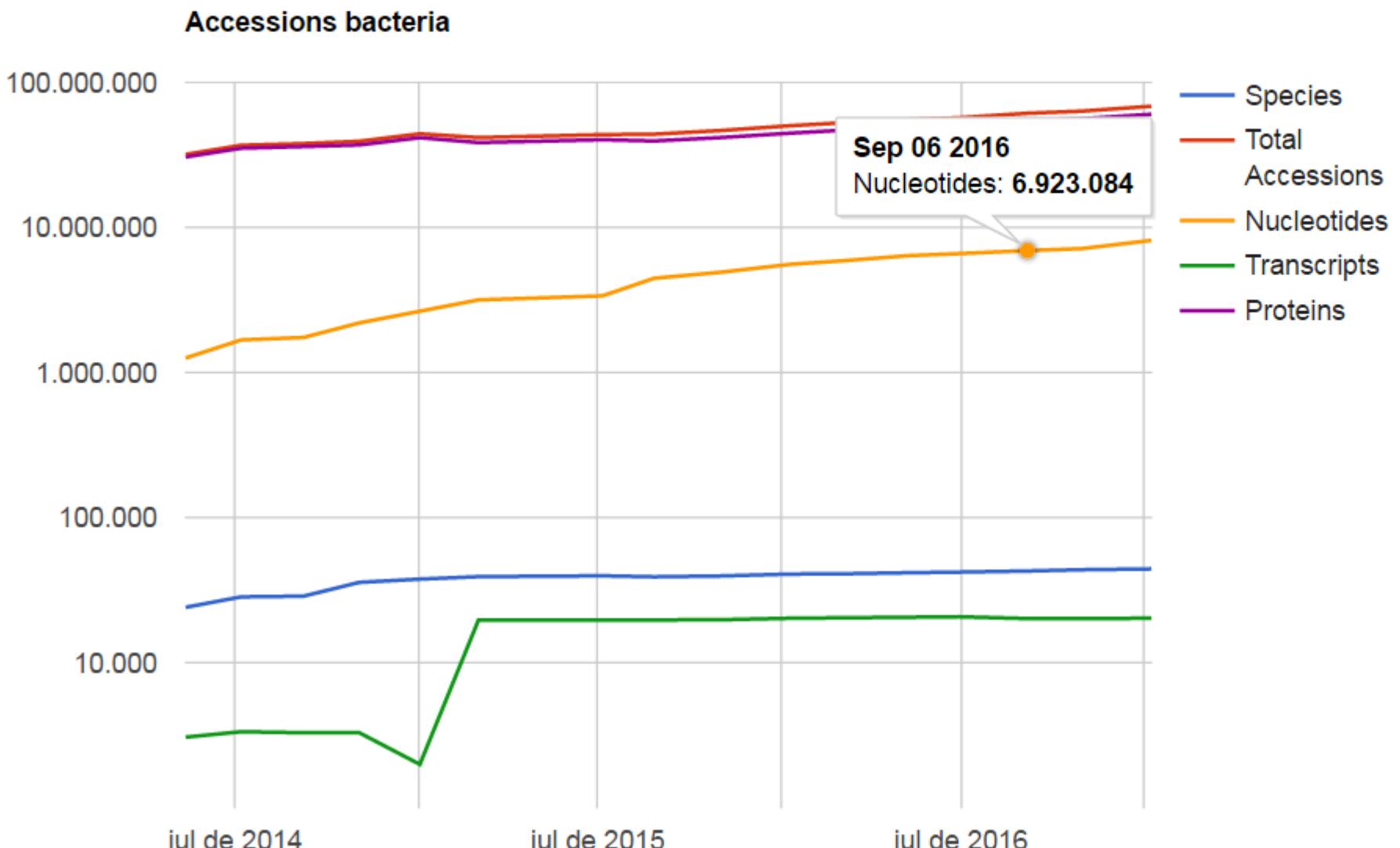


ESTATÍSTICAS DO REFSEQ/NCBI

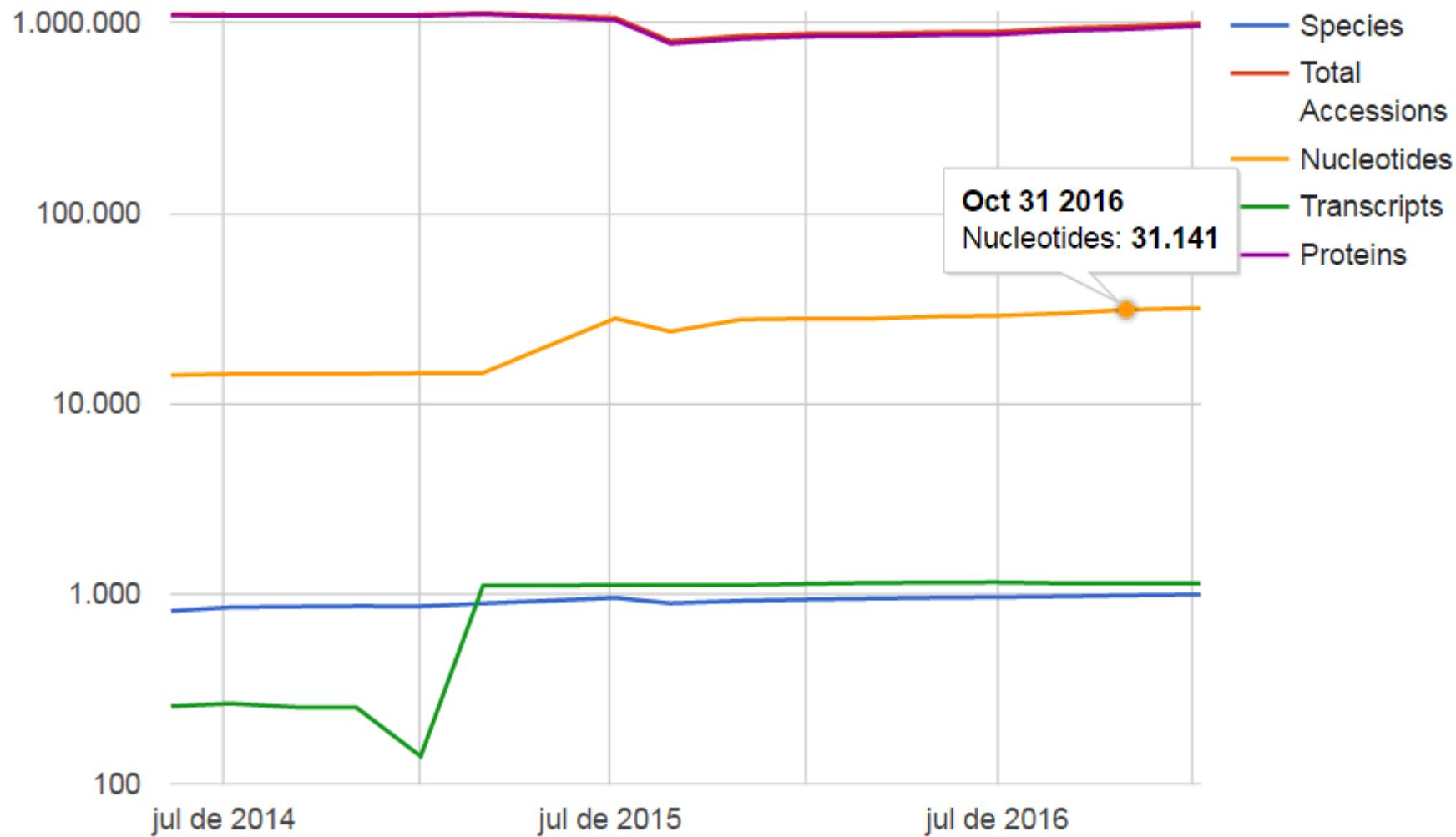
Accessions complete



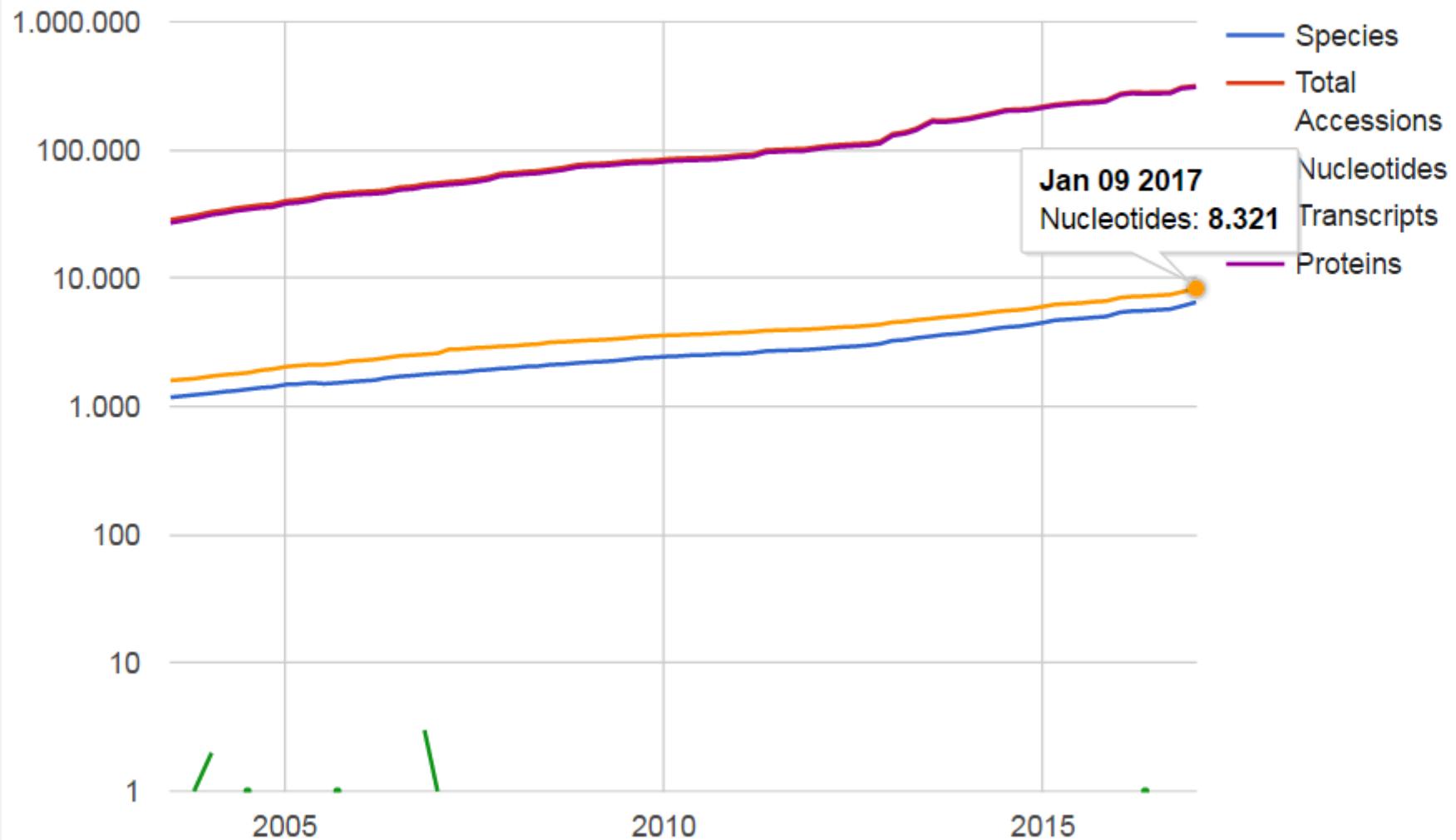
ESTATÍSTICAS DO REFSEQ/NCBI



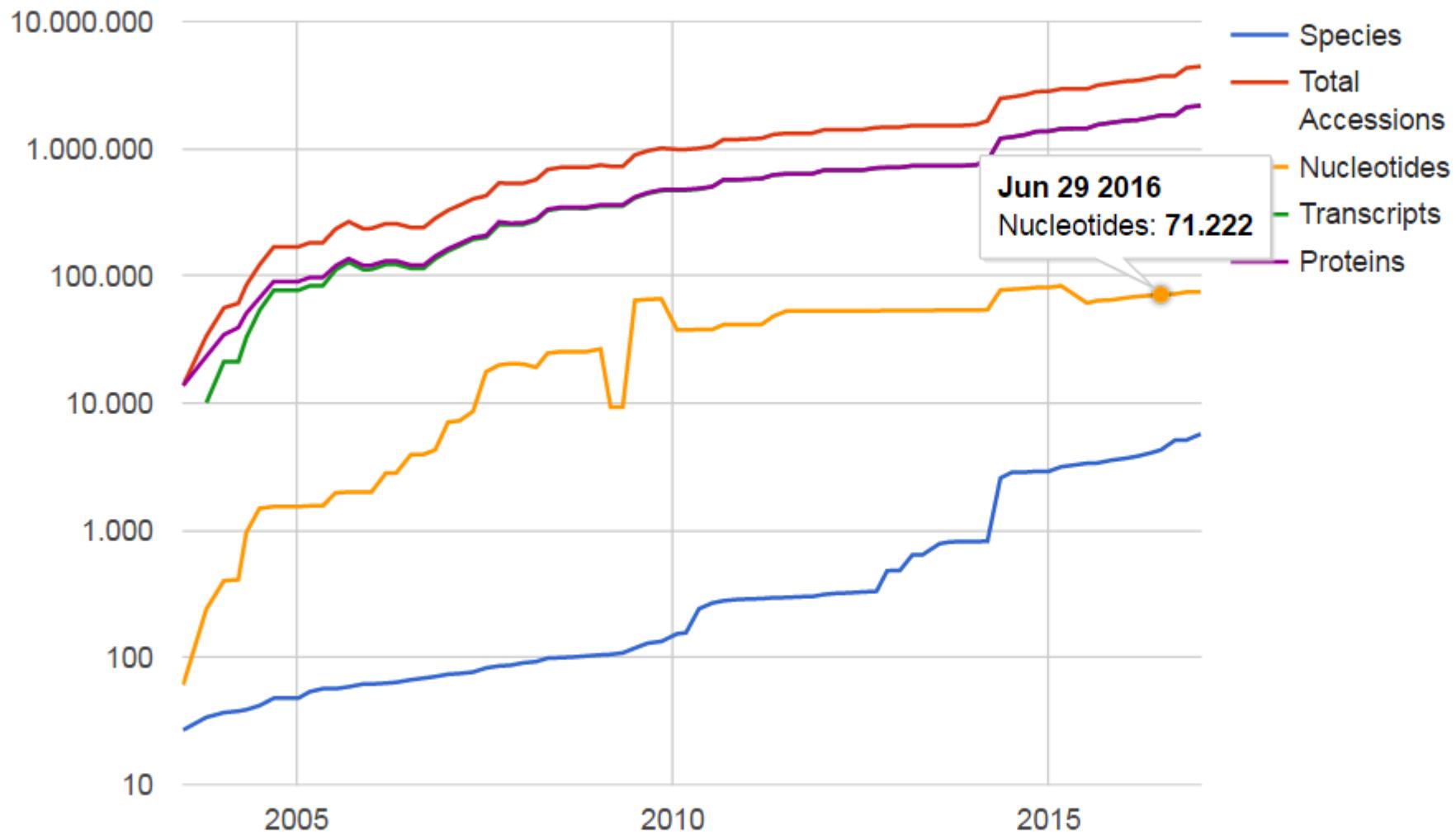
Accessions archaea



Accessions viral



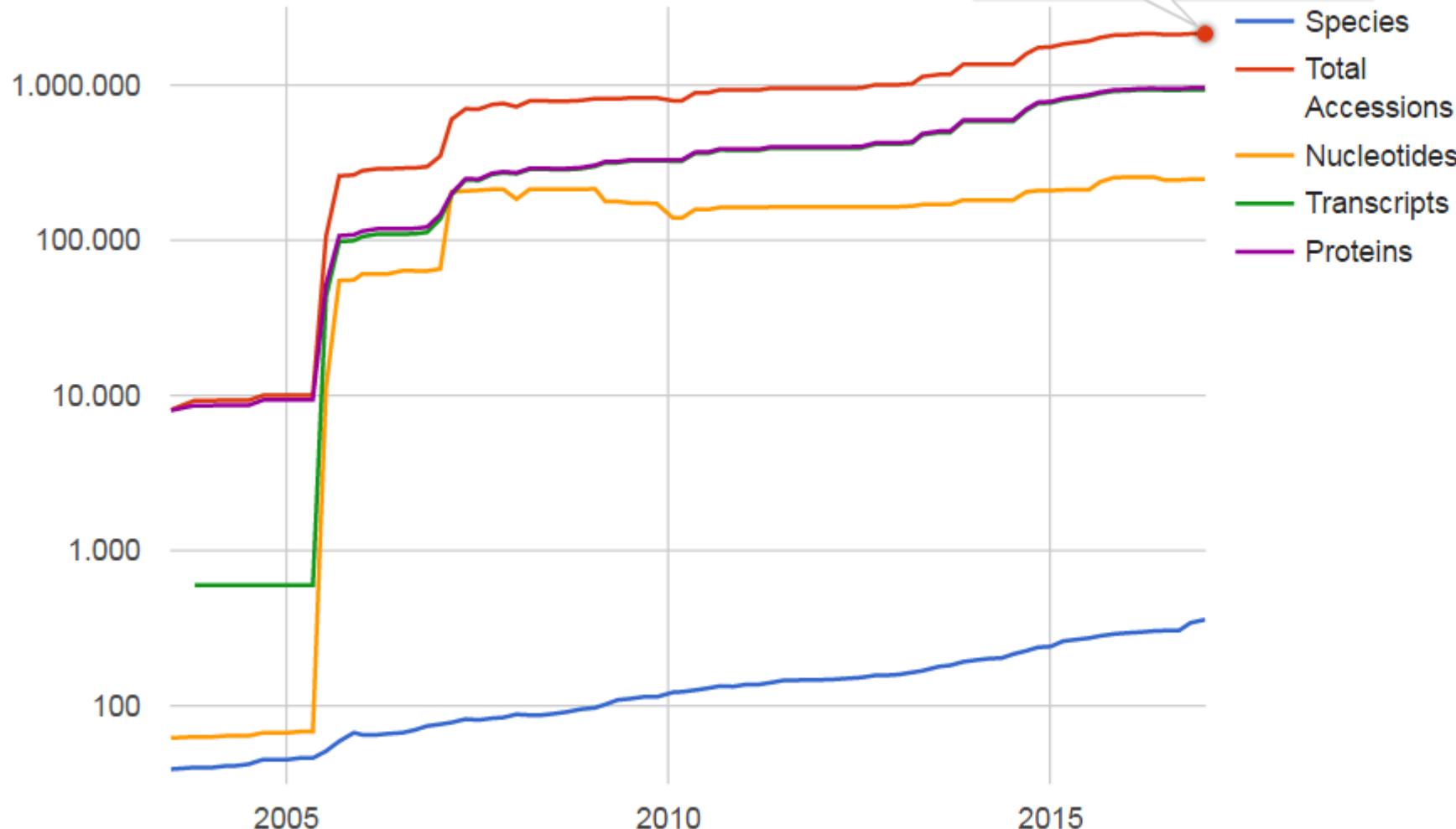
Accessions fungi



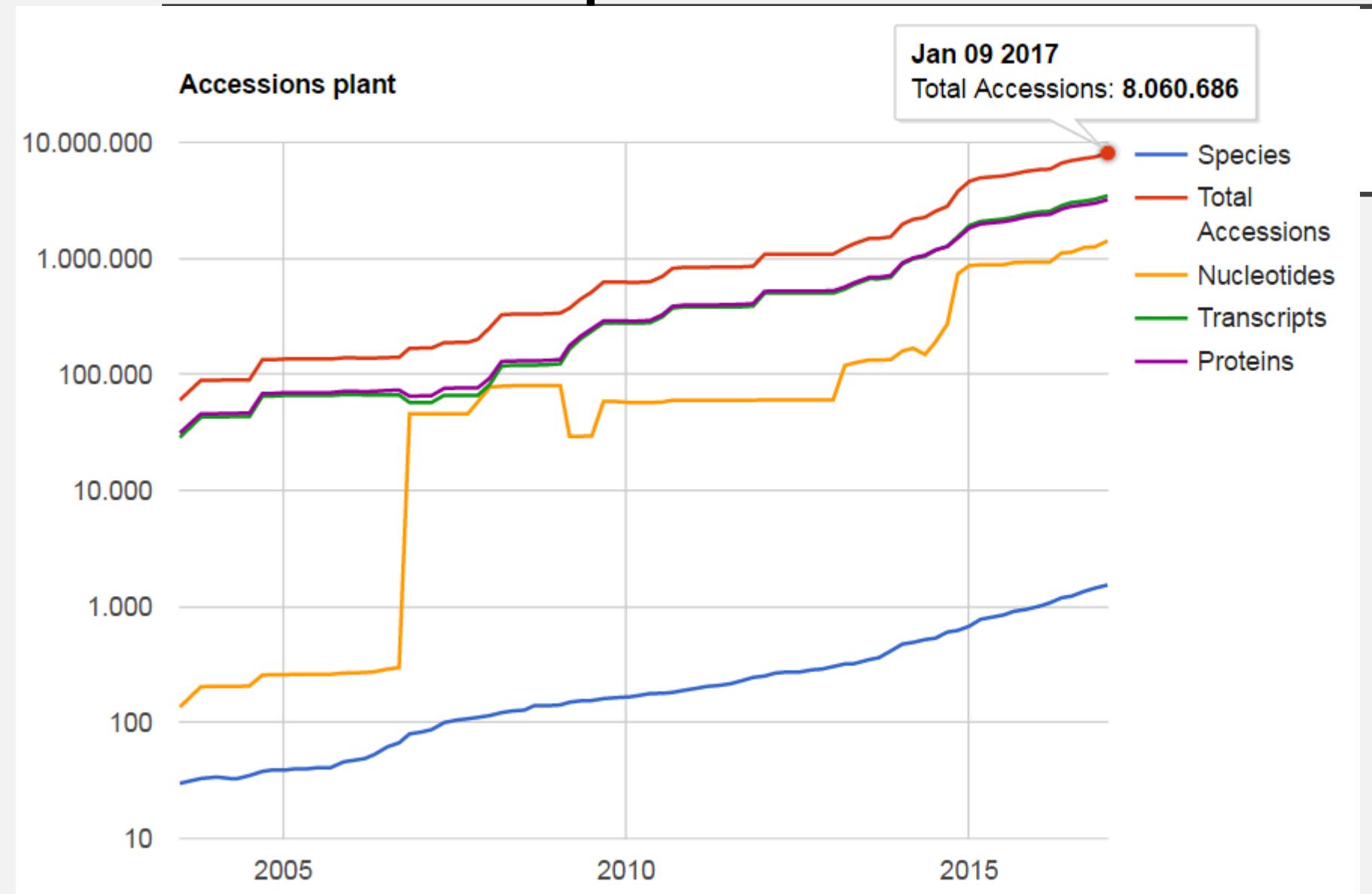
Accessions protozoa

Jan 09 2017

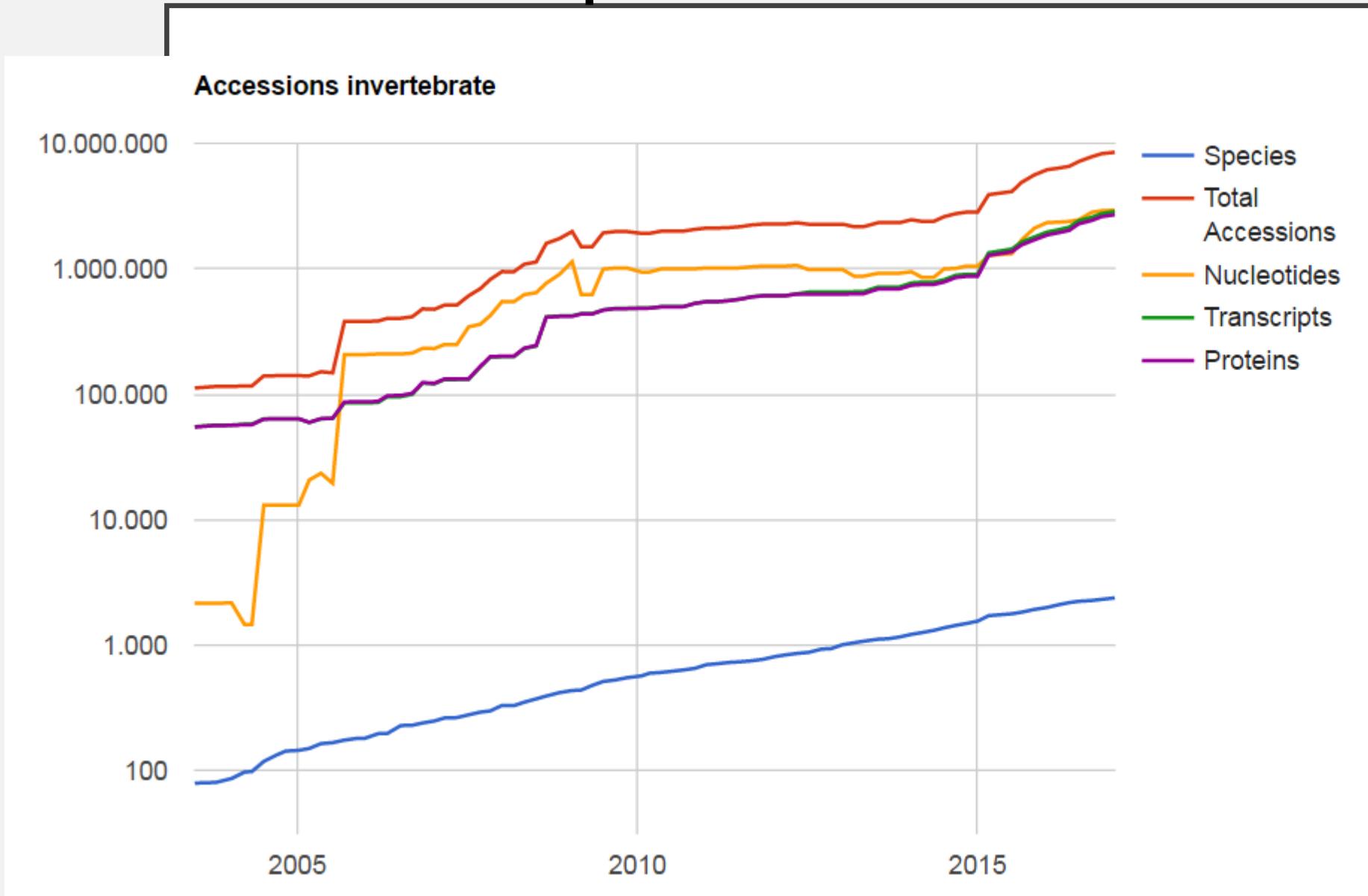
Total Accessions: 2.140.466



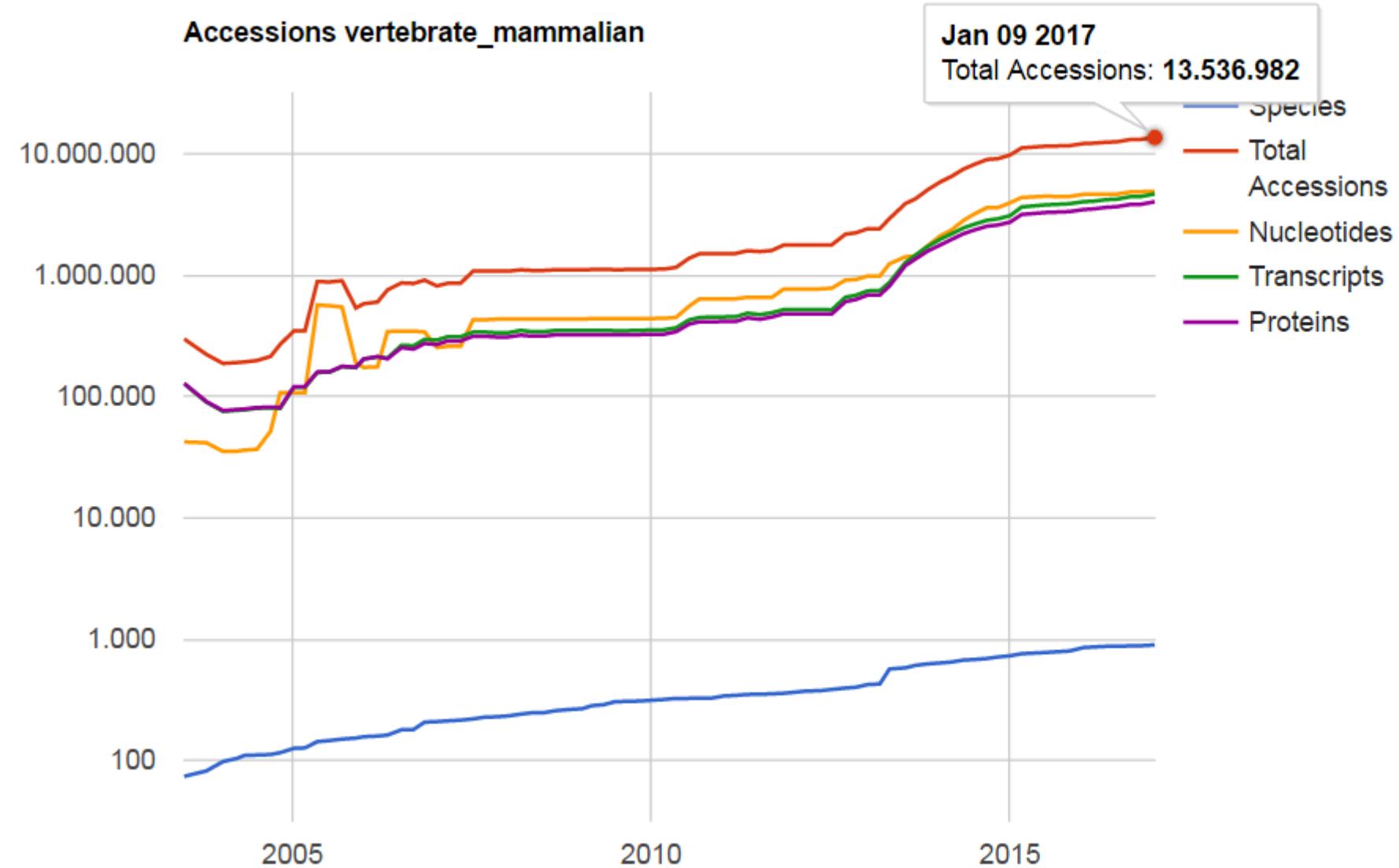
Estatísticas do RefSeq/NCBI



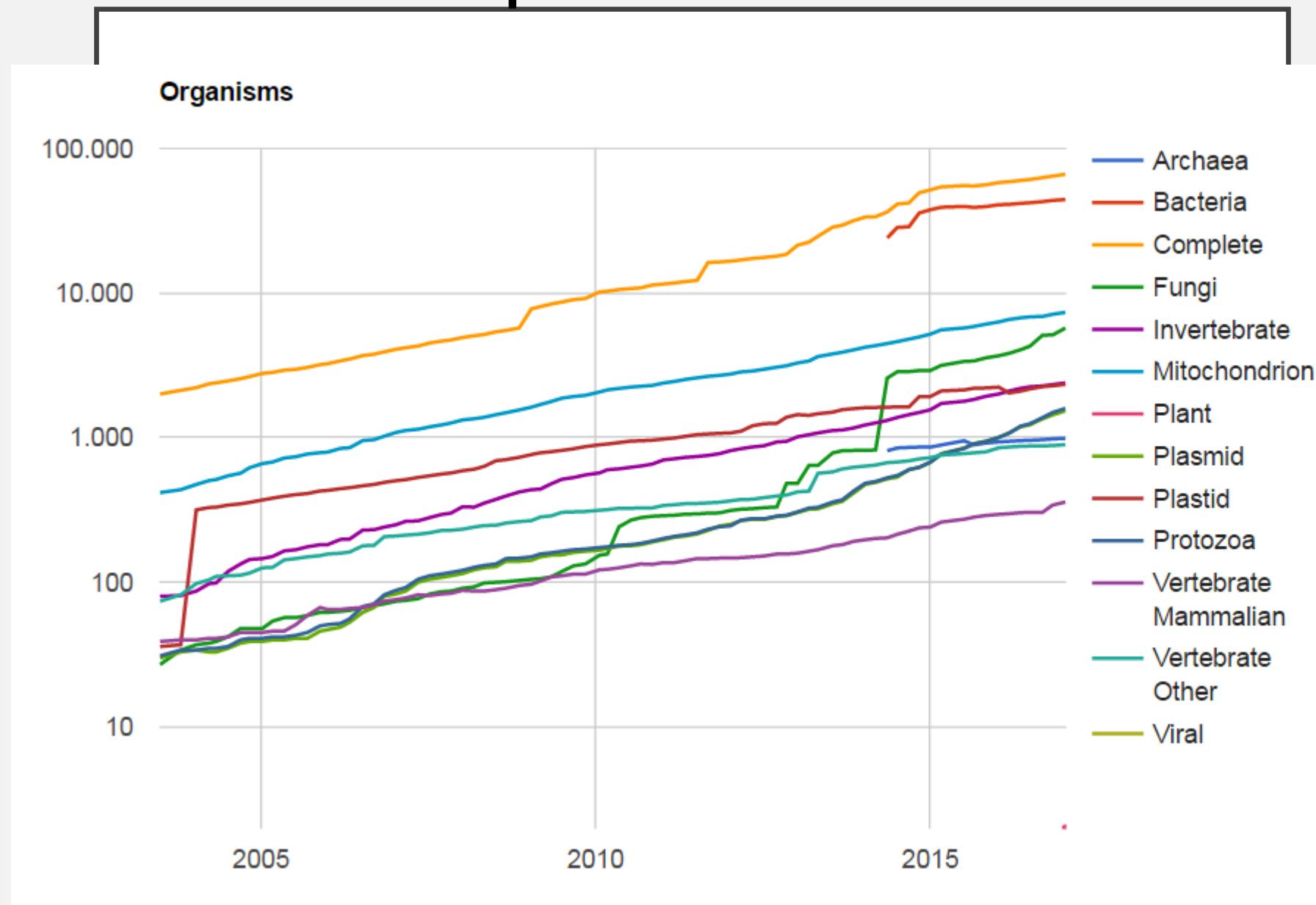
Estatísticas do RefSeq/NCBI

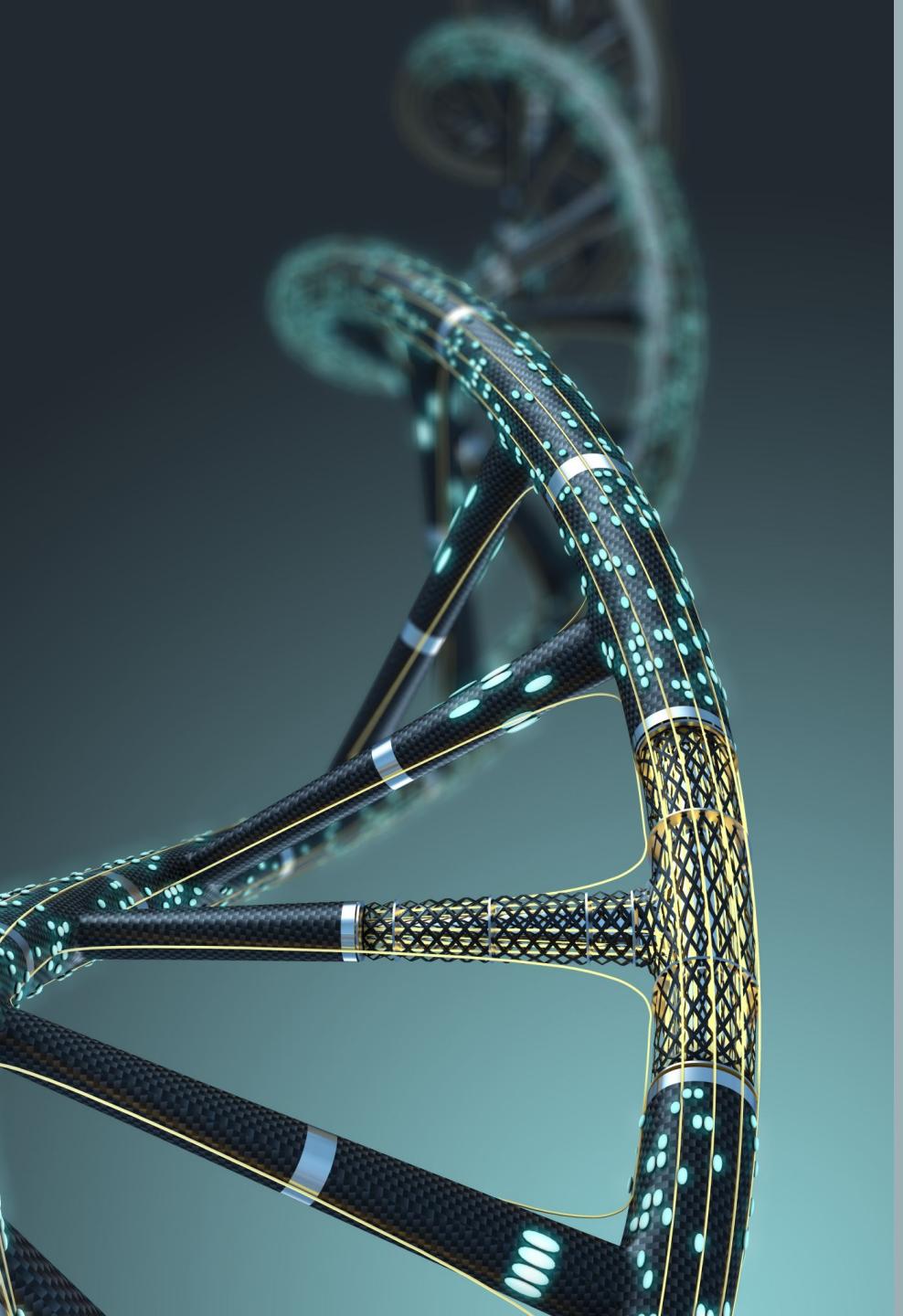


Estatísticas do RefSeq/NCBI



Estatísticas do RefSeq/NCBI





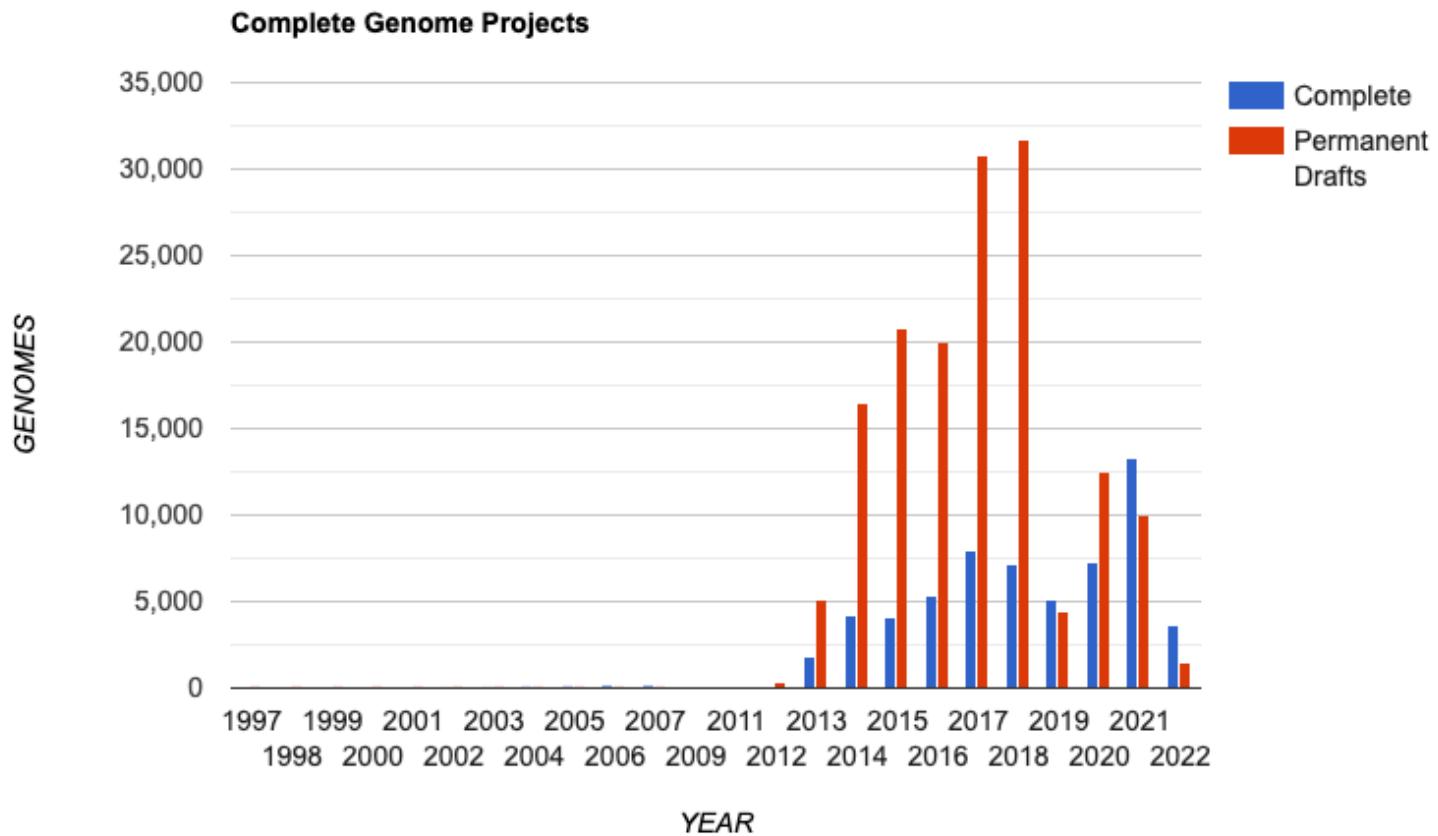
GOLD –GENOMES ONLINE DATABASE

<https://gold.jgi.doe.gov>

CRESCIMENTO DO GOLD

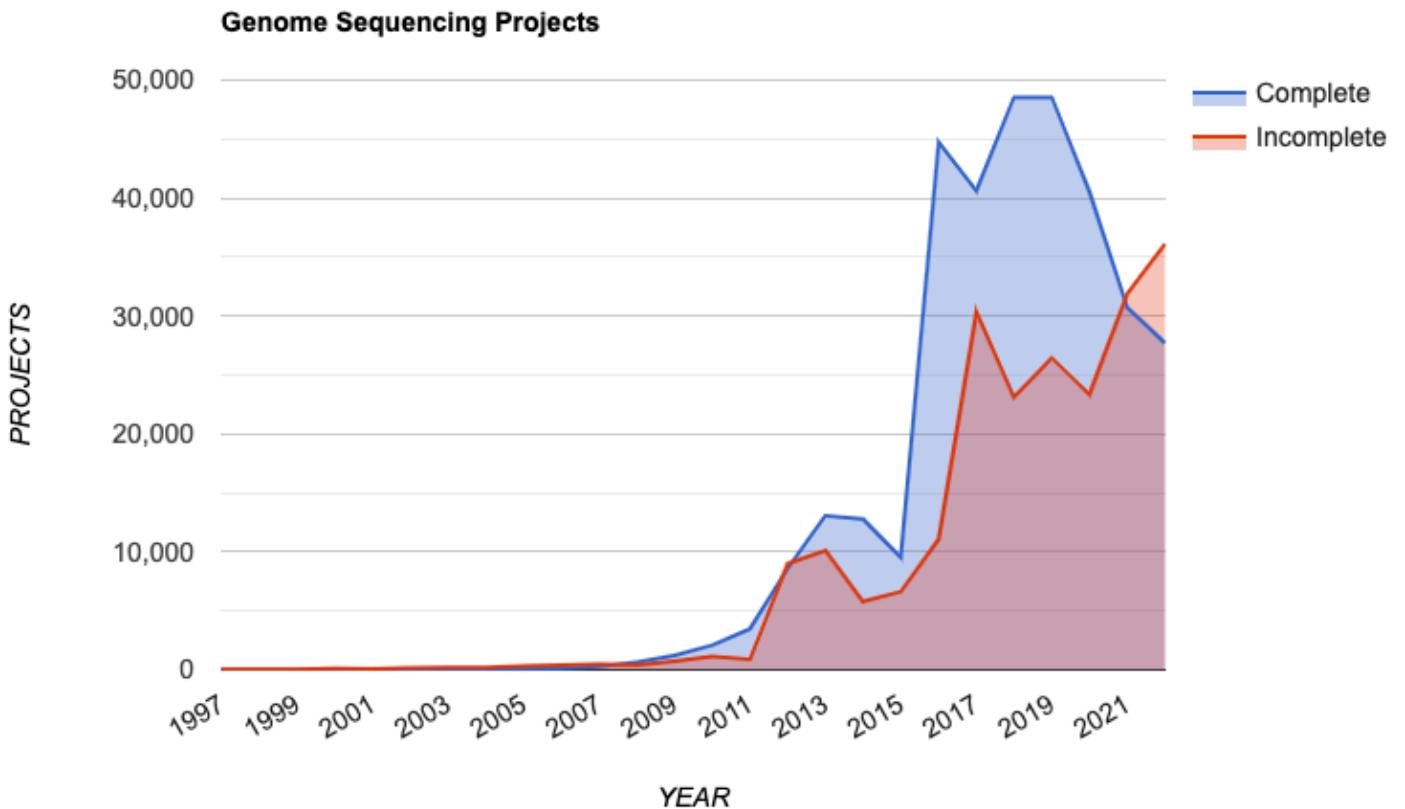
Statistics

Complete and Permanent Draft Genome Totals in GOLD (by year and status)



CRESCIMENTO DO GOLD

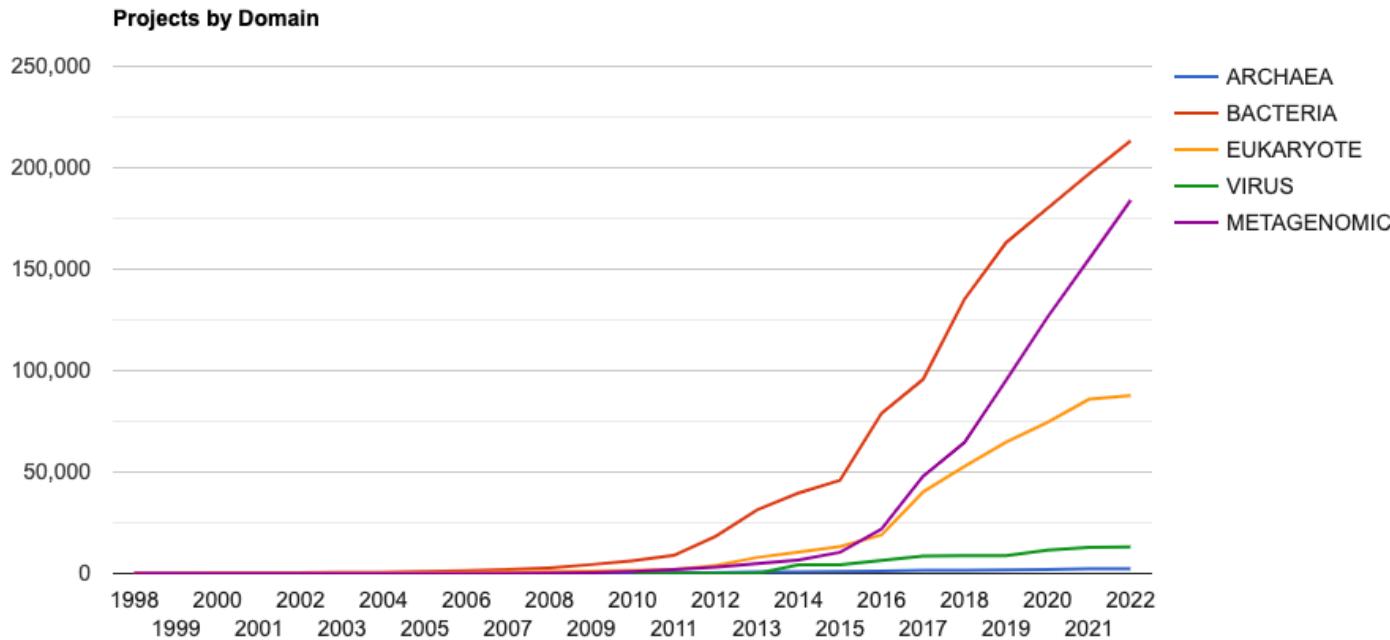
Genome Totals in GOLD (by year and status)



CRESCIMENTO DO GOLD

- Grupos taxonômicos

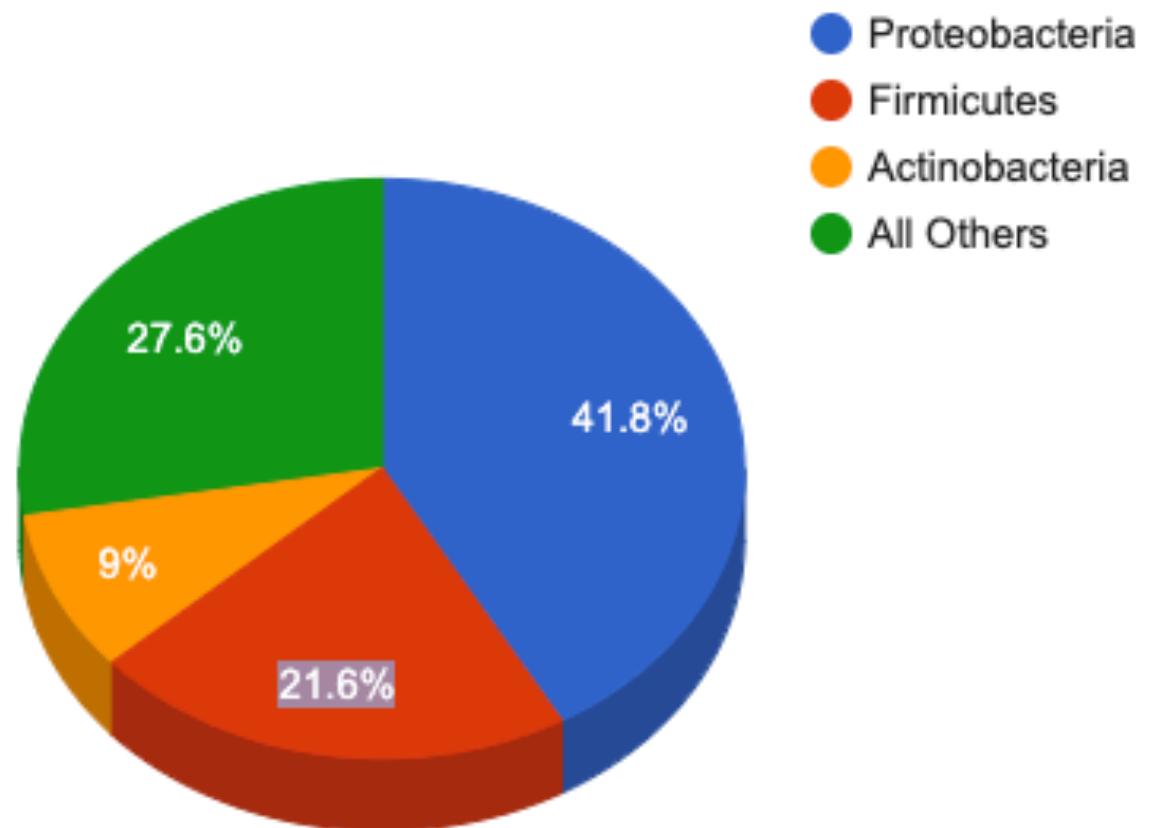
Project Totals in GOLD (by year and Domain Group)



GOLD – DIVISÃO DOS GRUPOS

Phylogenetic distribution of Bacterial Genome Projects

Projects by Bacterial Phylogenetic Group

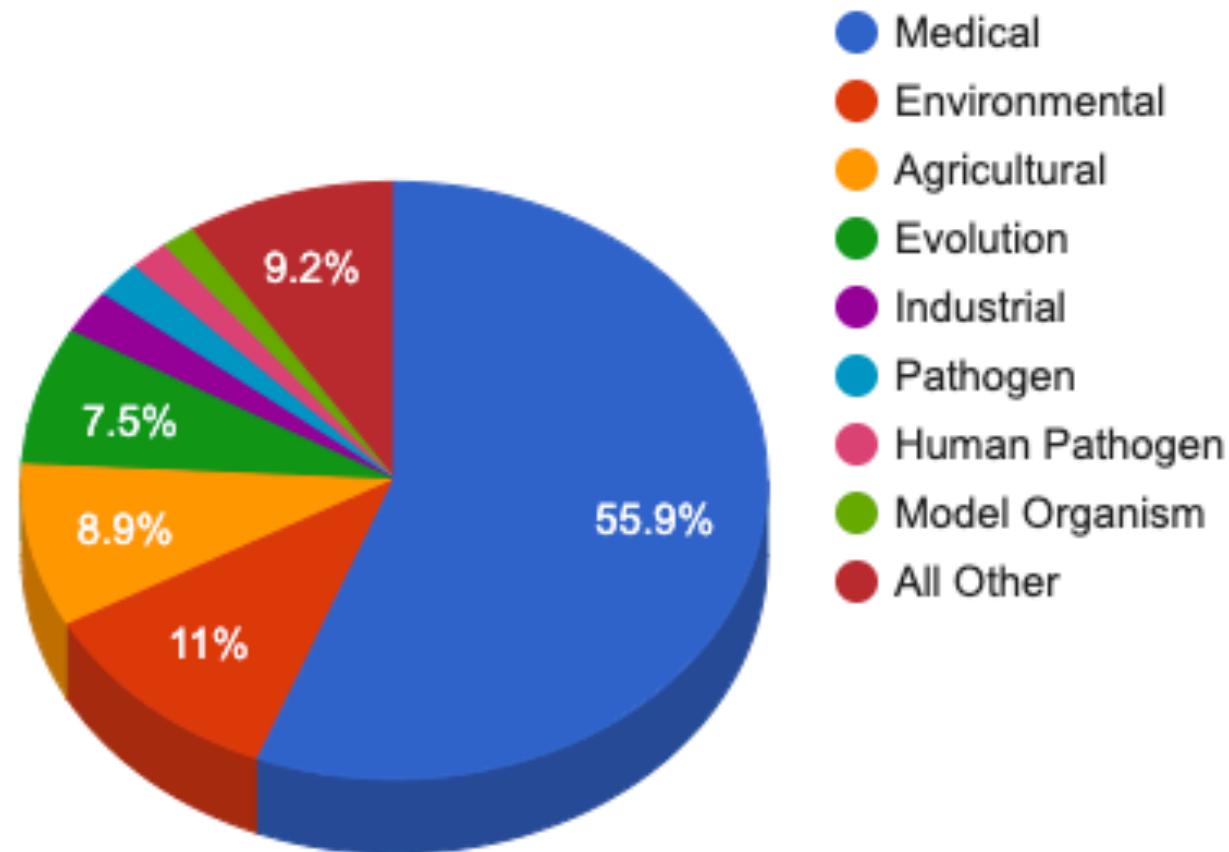


GOLD – DIVISÃO DOS GRUPOS

- Relevância dos genomas

Project Relevance of Bacterial Genome Projects

Project Relevance of Bacterial Projects

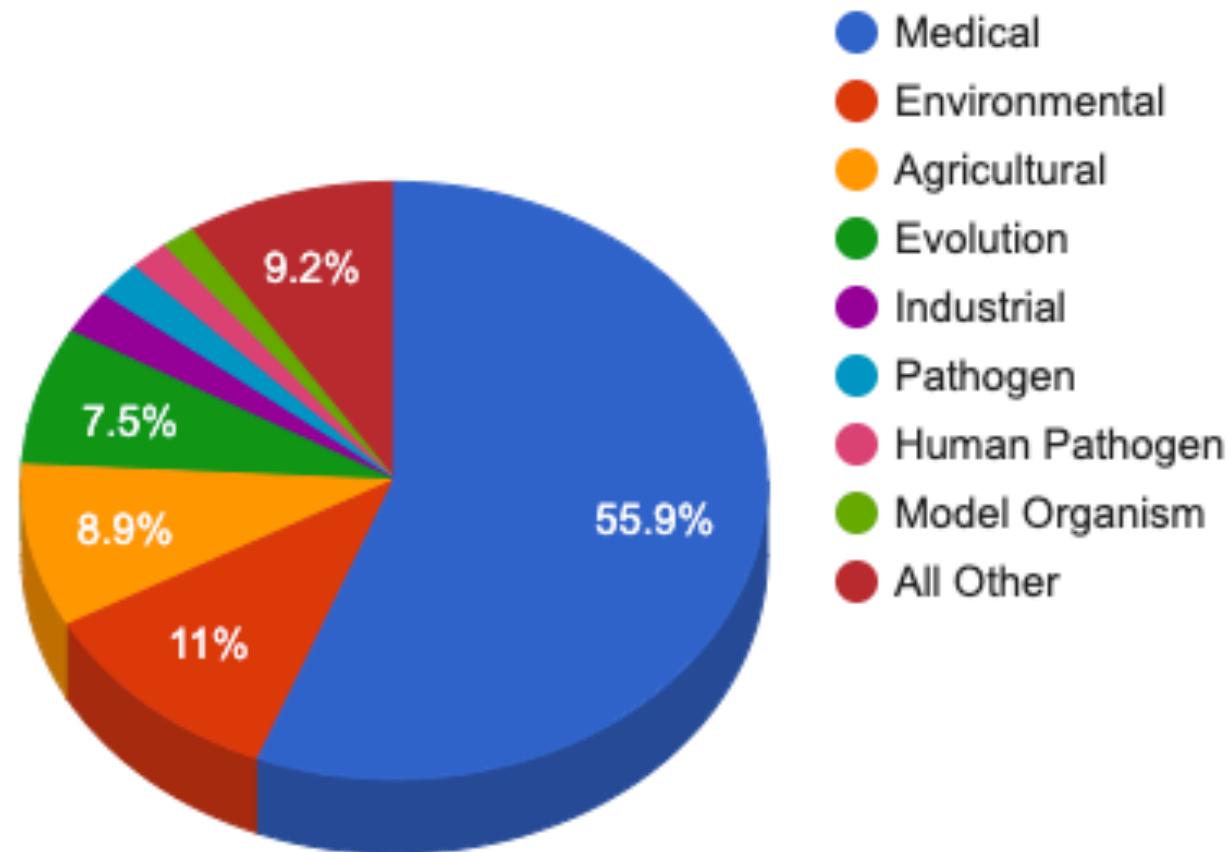


GOLD – DIVISÃO DOS GRUPOS

- Relevância dos genomas

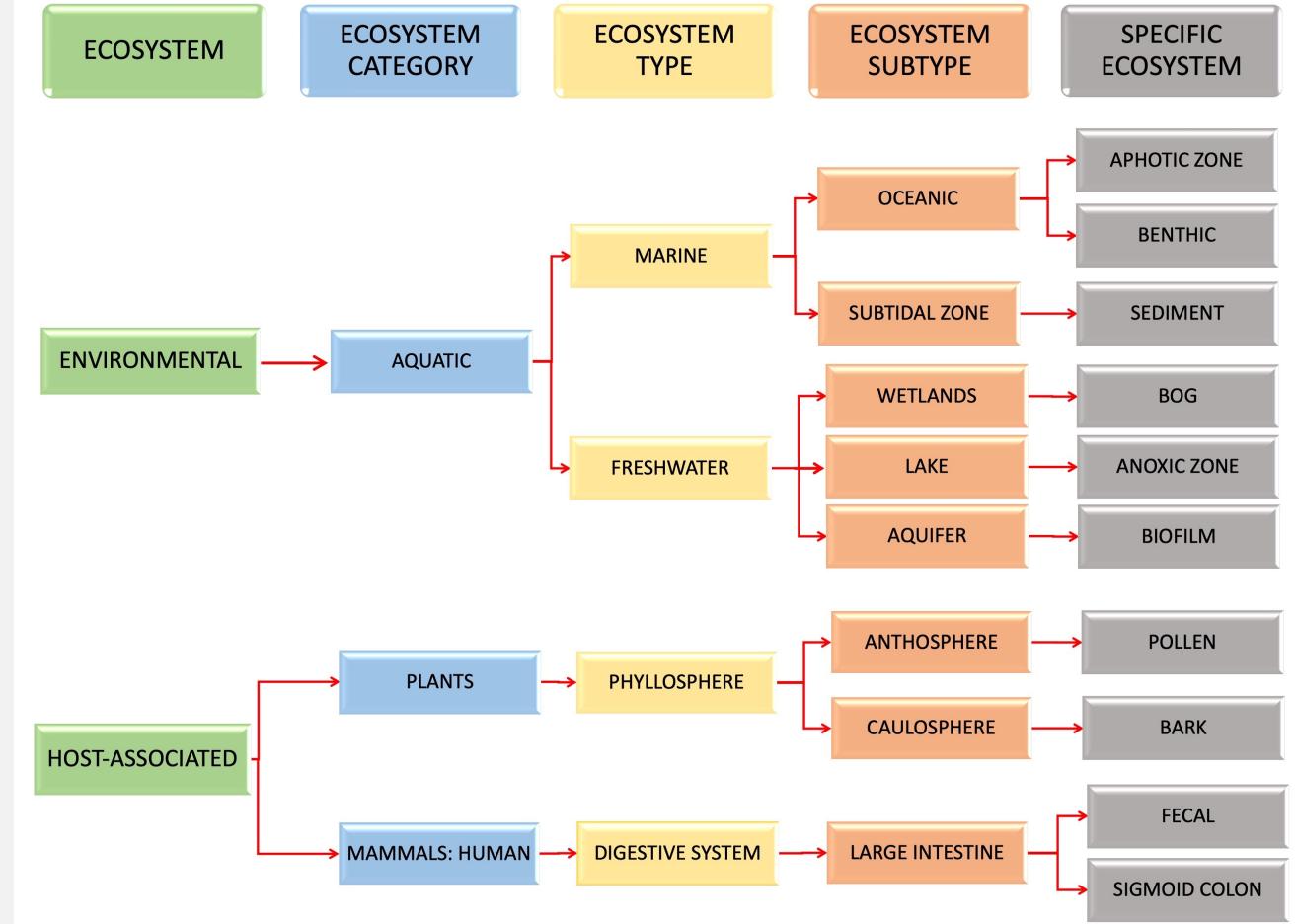
Project Relevance of Bacterial Genome Projects

Project Relevance of Bacterial Projects



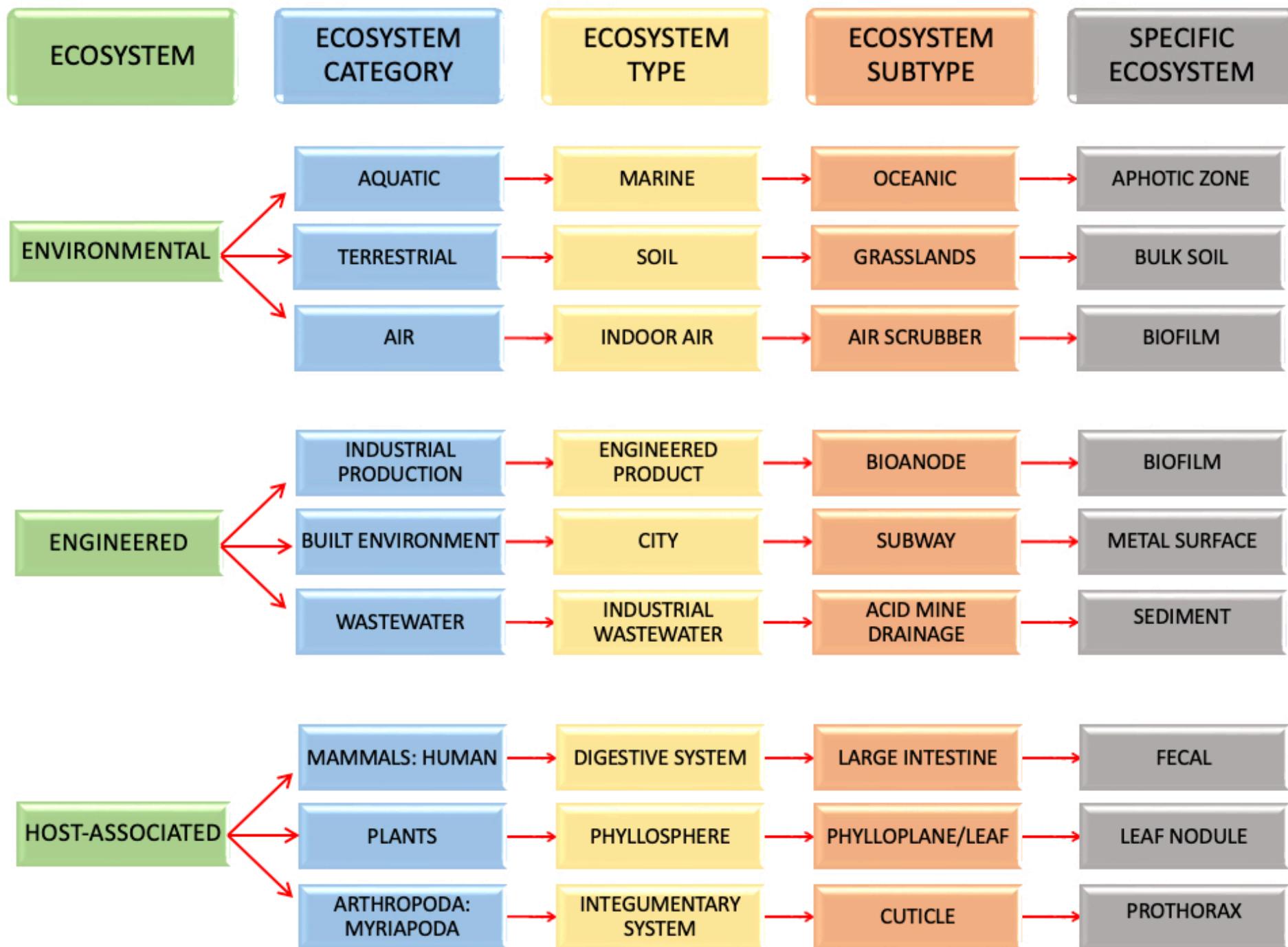
GOLD – ECOSSISTEMA

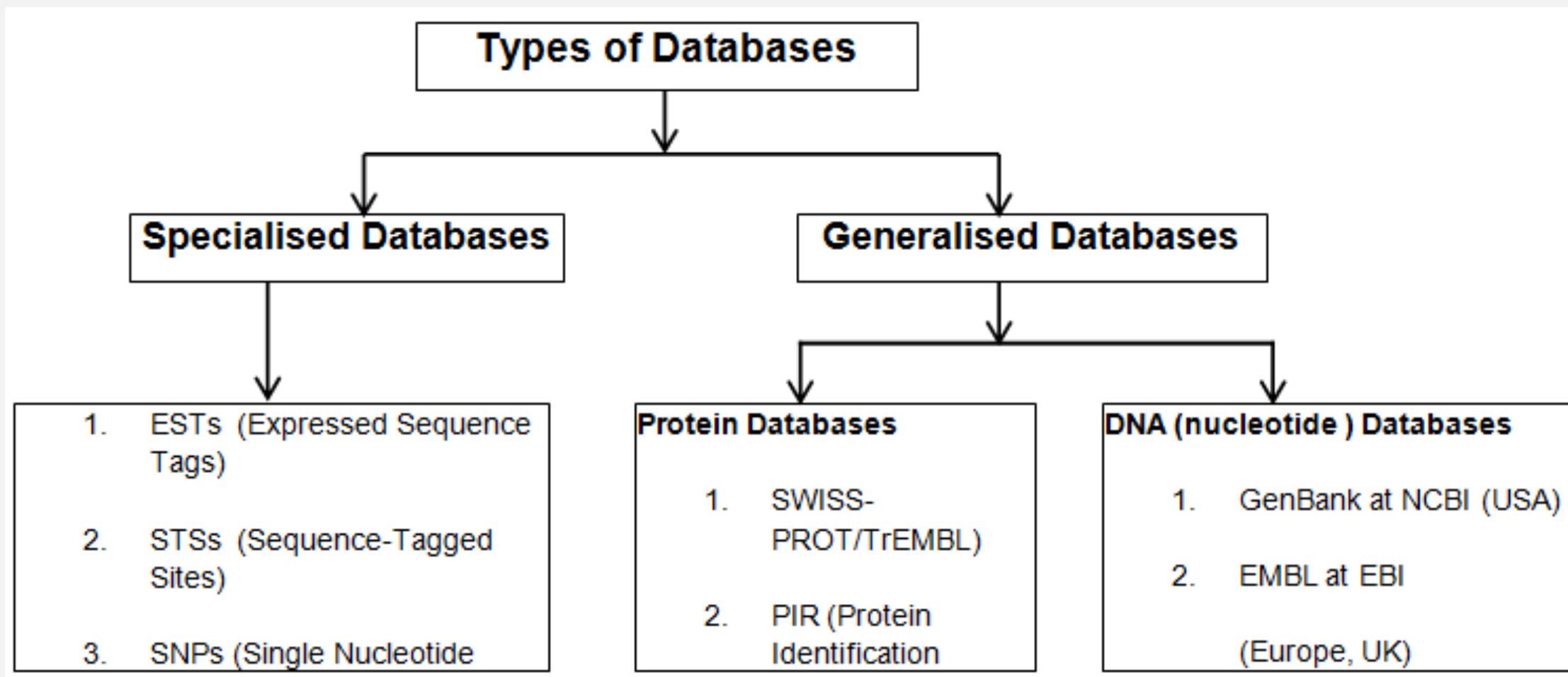
- Amostras

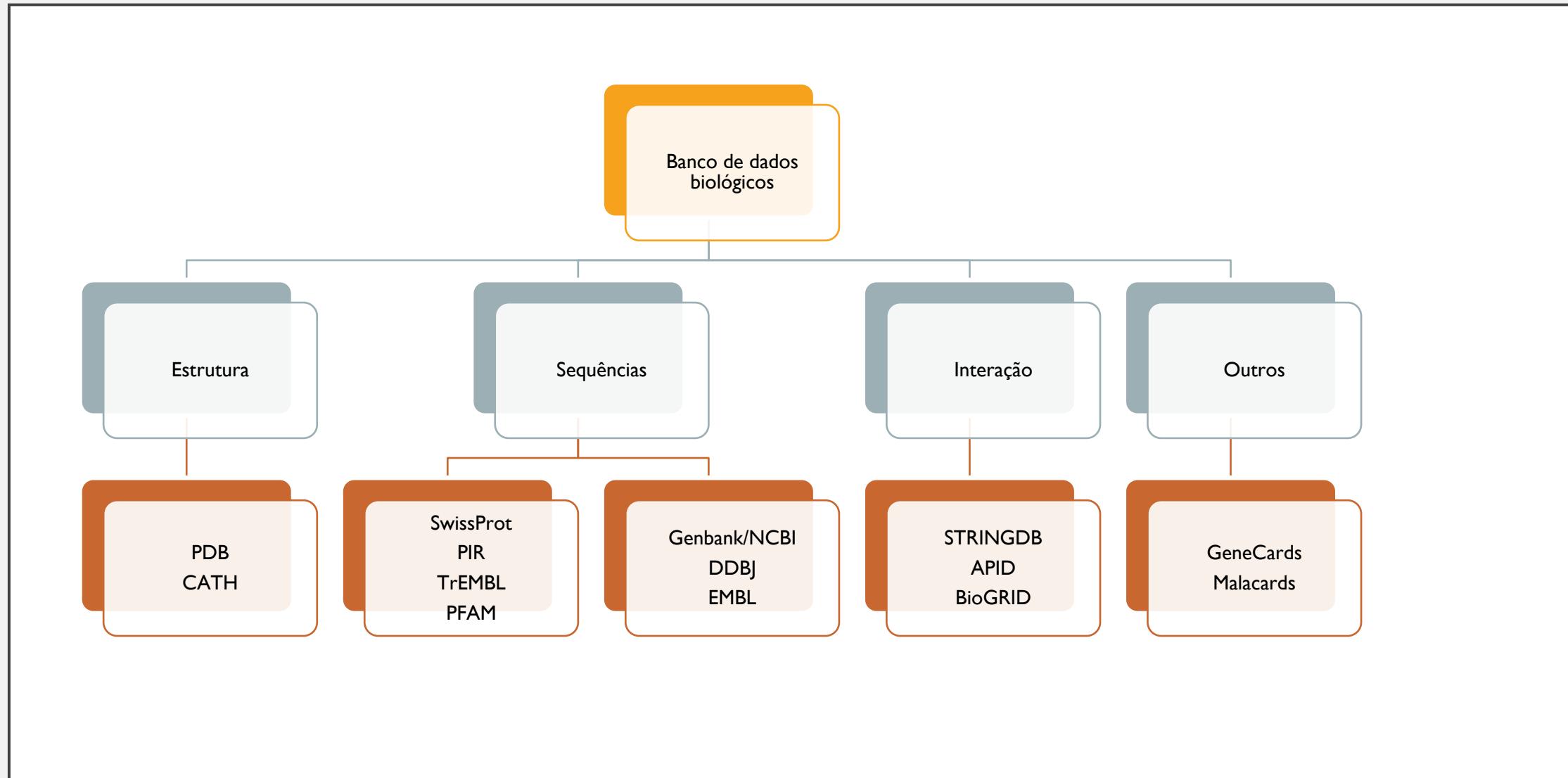


E

Am

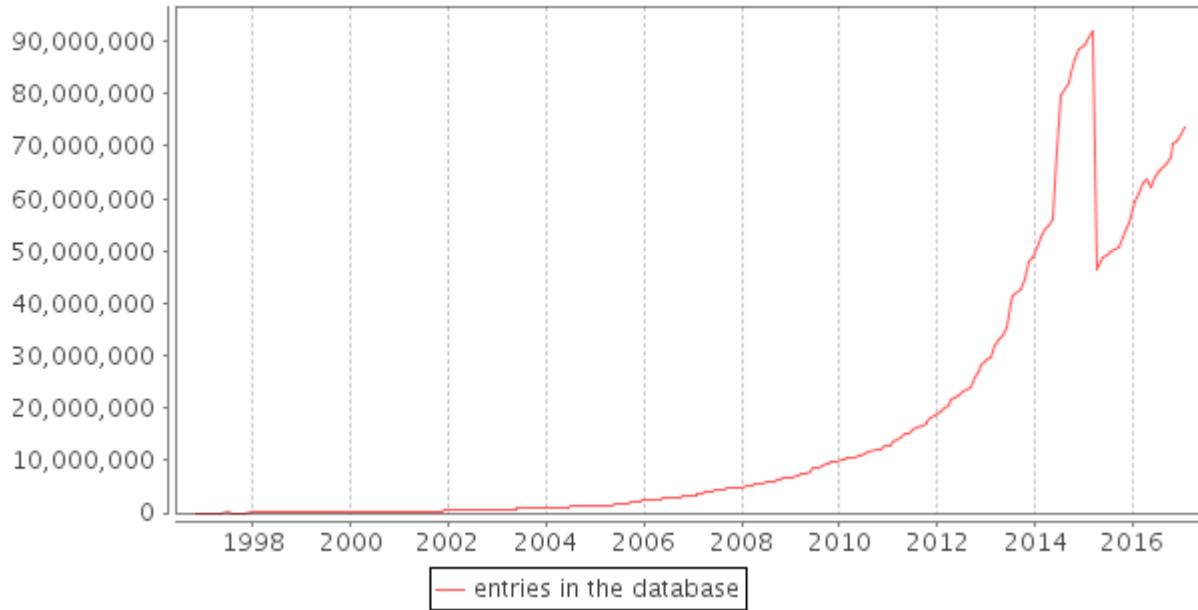




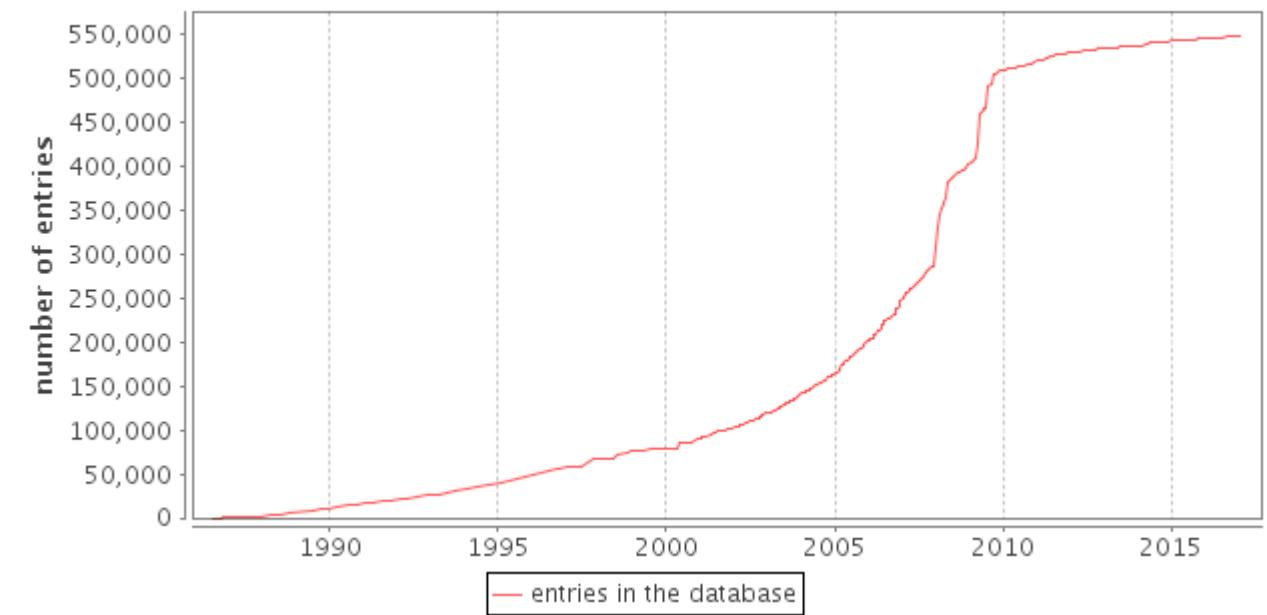




Number of entries in UniProtKB/TrEMBL over time



Number of entries in UniProtKB/Swiss-Prot over time



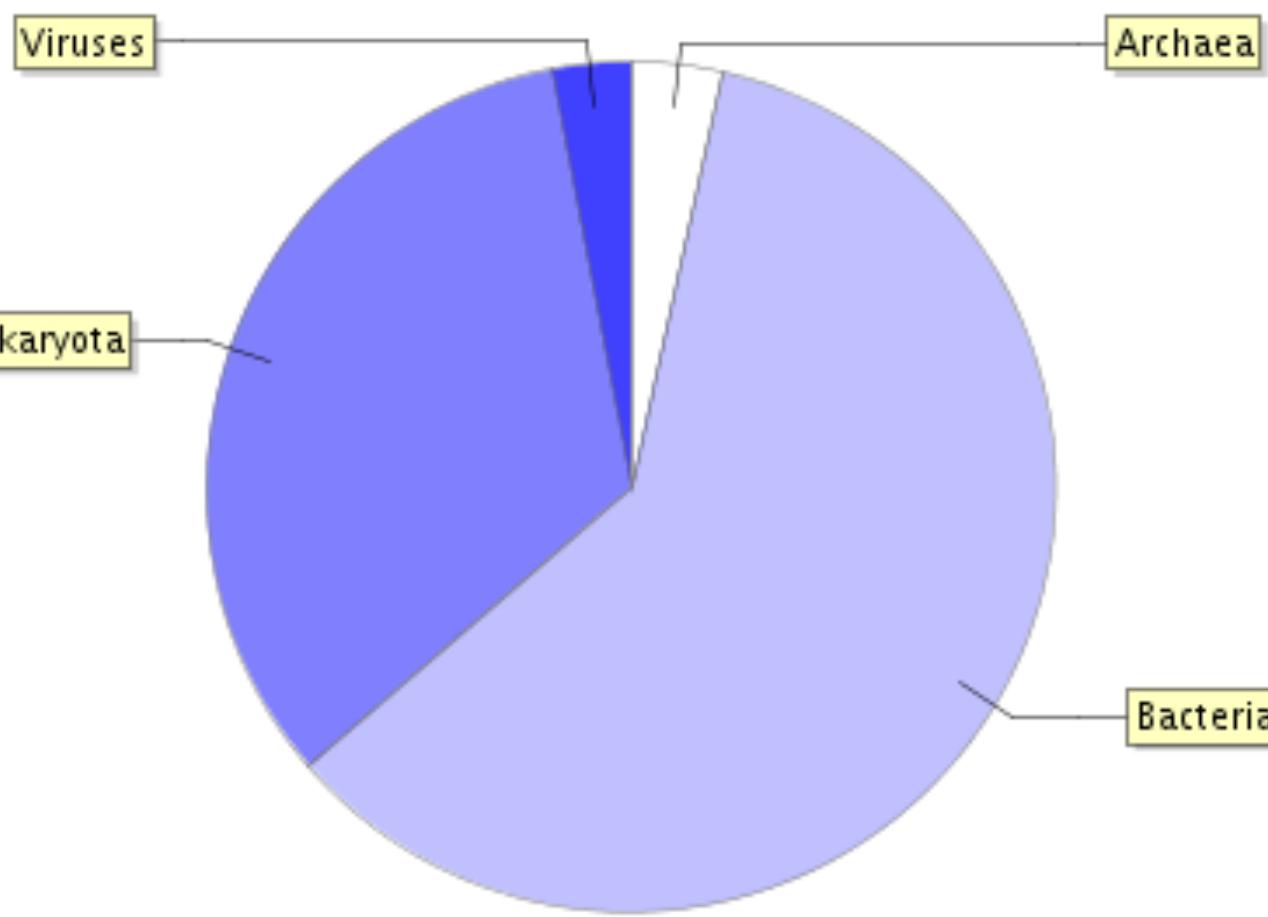
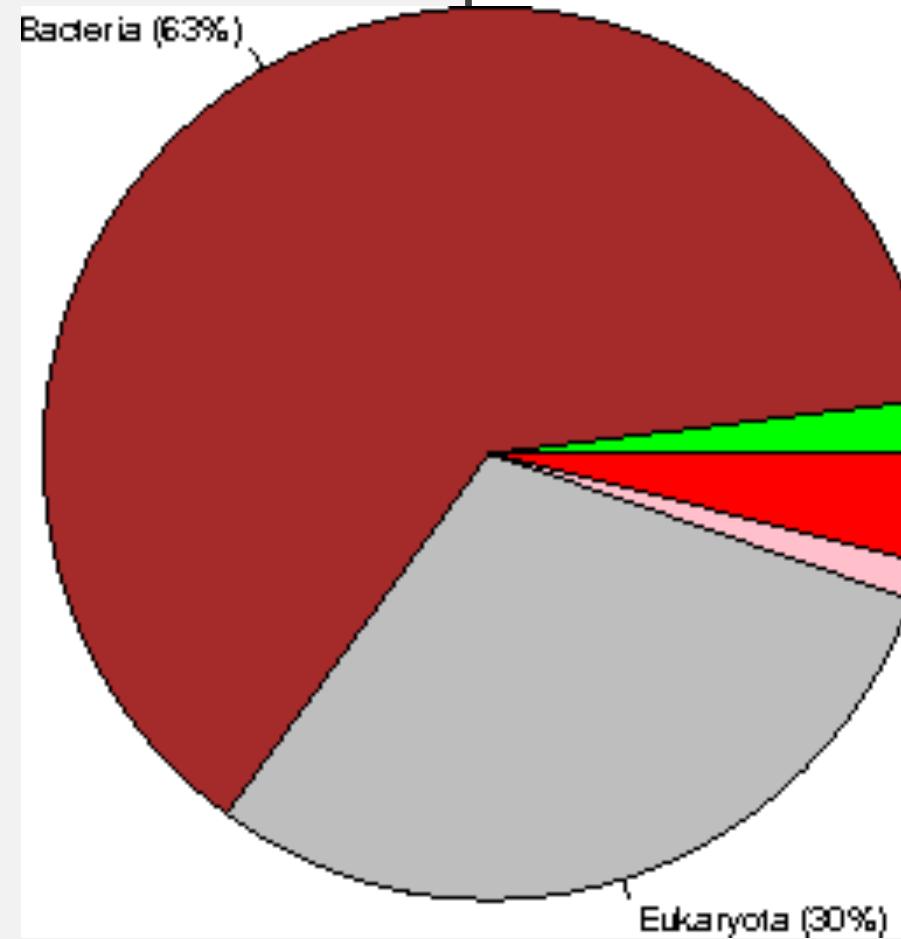
UNIPROTKB/TREMBL

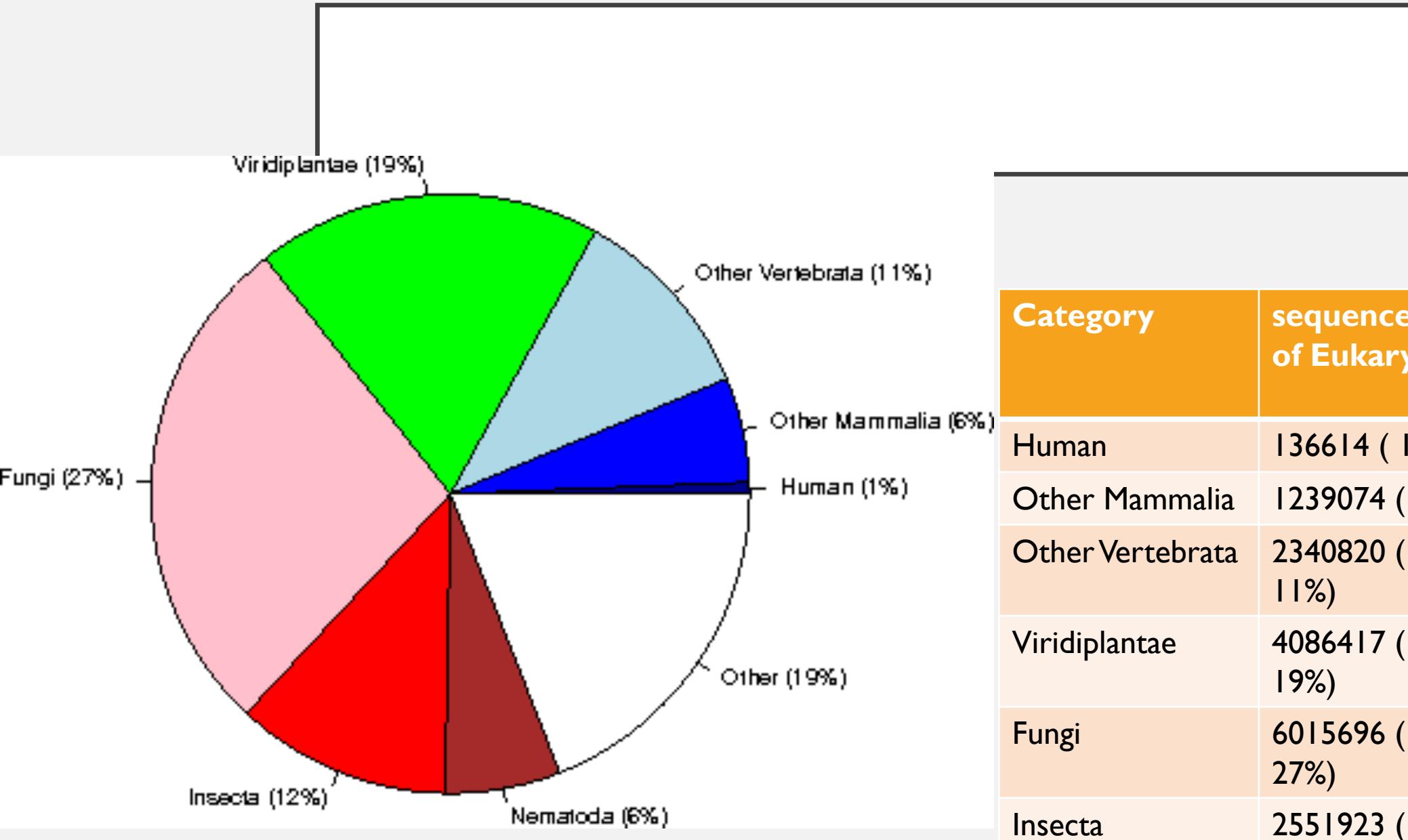
Protein Existence (PE)	Number of entries
Evidence at protein level	125,544
Evidence at transcript level	1,063,419
Inferred from homology	17,236,726
Predicted	55,286,192



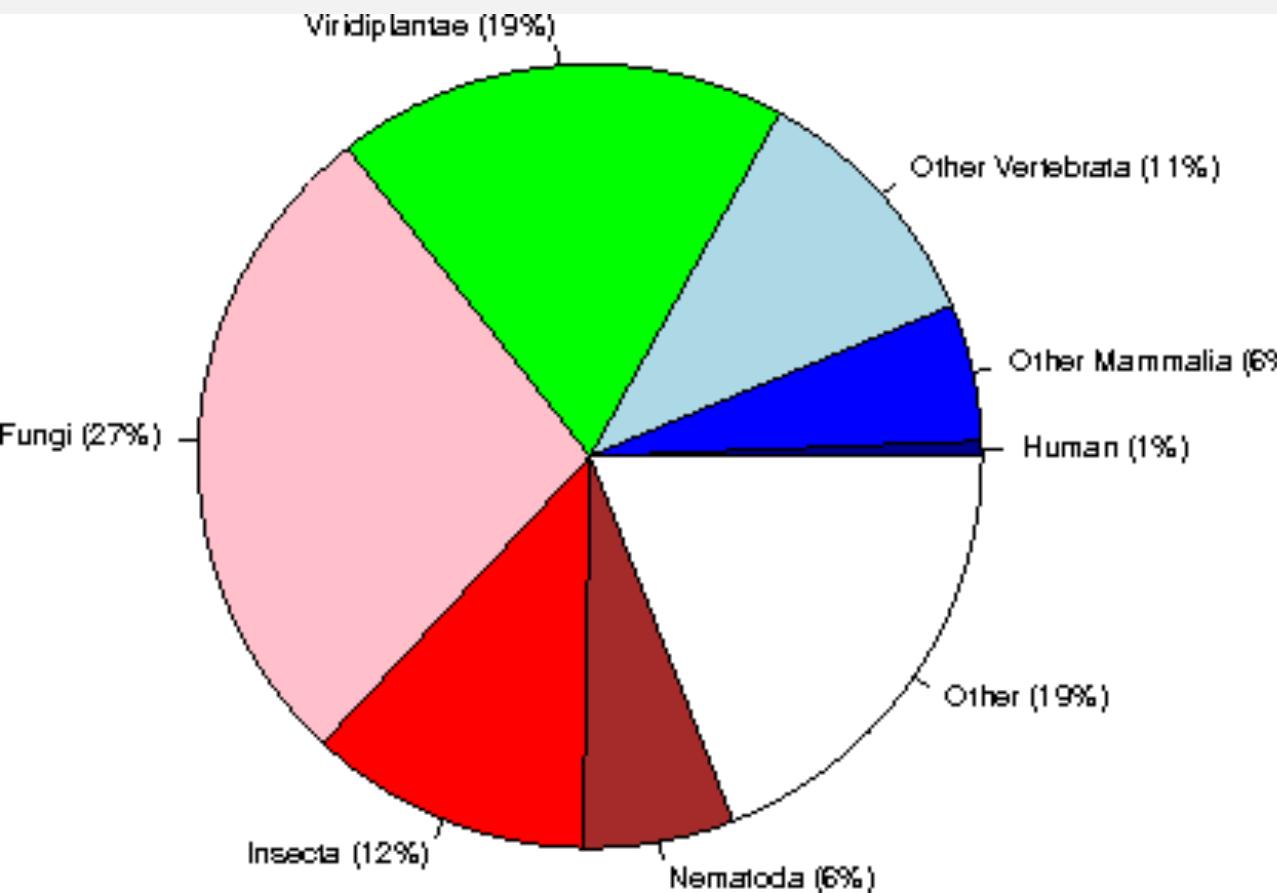
Number	Frequency	Species
1	743808	Human immunodeficiency virus I
2	668601	marine sediment metagenome
3	506807	<i>Daphnia magna</i>
4	259928	Arundo donax (Giant reed) (<i>Donax arundinaceus</i>)
5	201351	uncultured bacterium
6	191907	<i>Escherichia coli</i>
7	179540	<i>Bacillus cereus</i>
8	165724	Fundulus heteroclitus (Killifish) (Mummichog)
9	156305	<i>Streptococcus pneumoniae</i>
10	146583	<i>Pseudomonas fluorescens</i>
11	145614	<i>Triticum aestivum</i> (Wheat)
12	139437	<i>Zea mays</i> (Maize)
13	136539	<i>Homo sapiens</i> (Human)

Swiss-Prot entries per taxonomic group

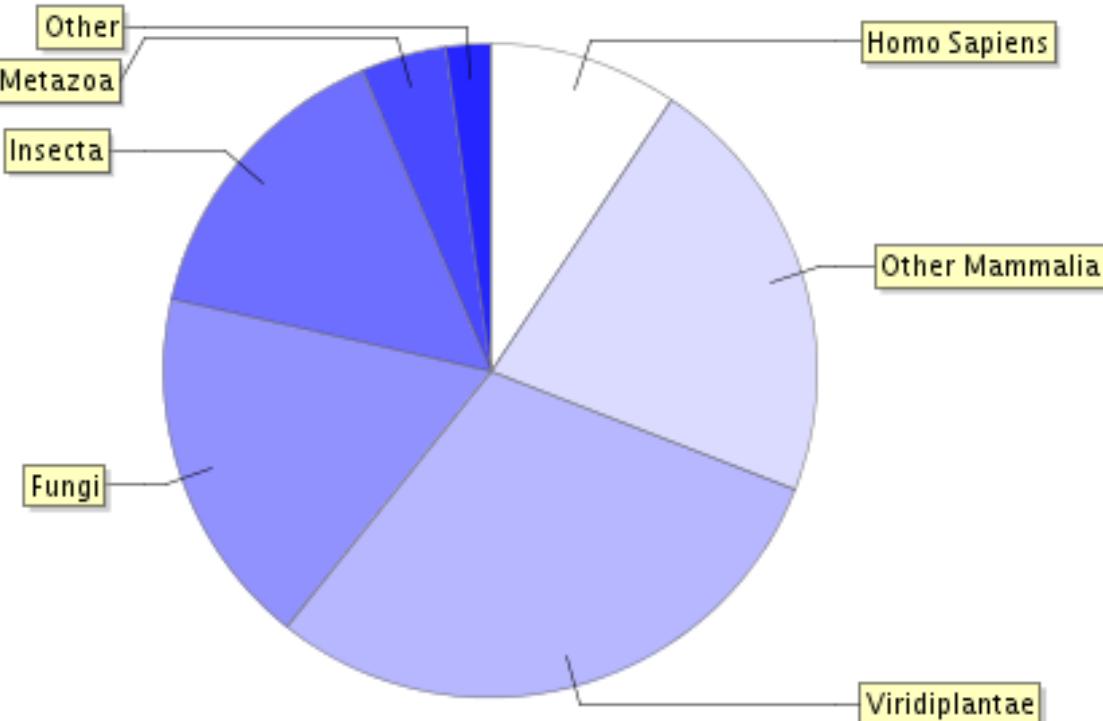




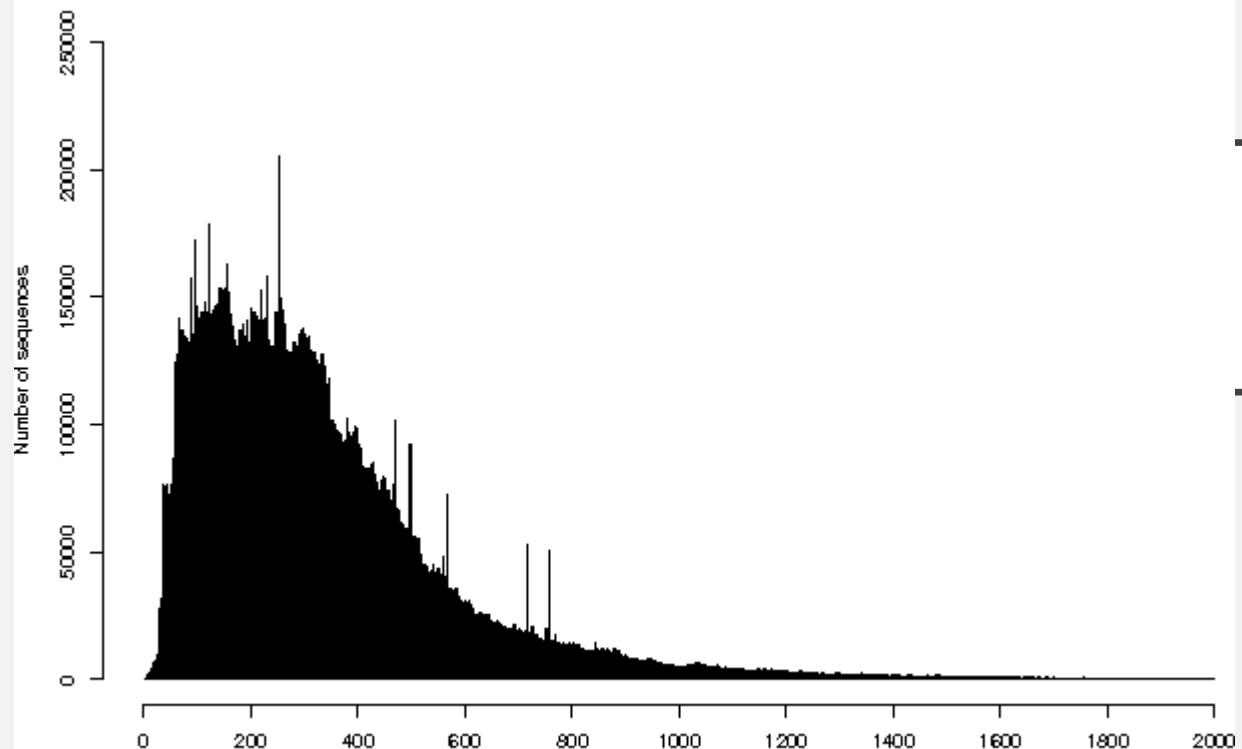
Category	sequences (% of Eukaryota)	(% of the complete database)
Human	136614 (1%)	(0%)
Other Mammalia	1239074 (6%)	(2%)
Other Vertebrata	2340820 (11%)	(3%)
Viridiplantae	4086417 (19%)	(6%)
Fungi	6015696 (27%)	(8%)
Insecta	2551923 (12%)	(3%)



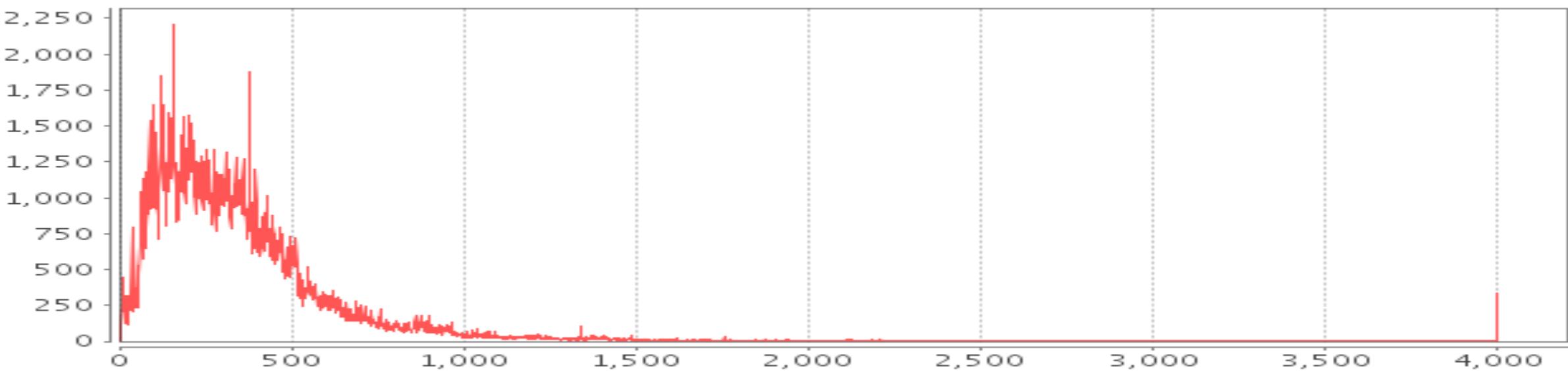
Swiss-Prot entries in Eukaryota

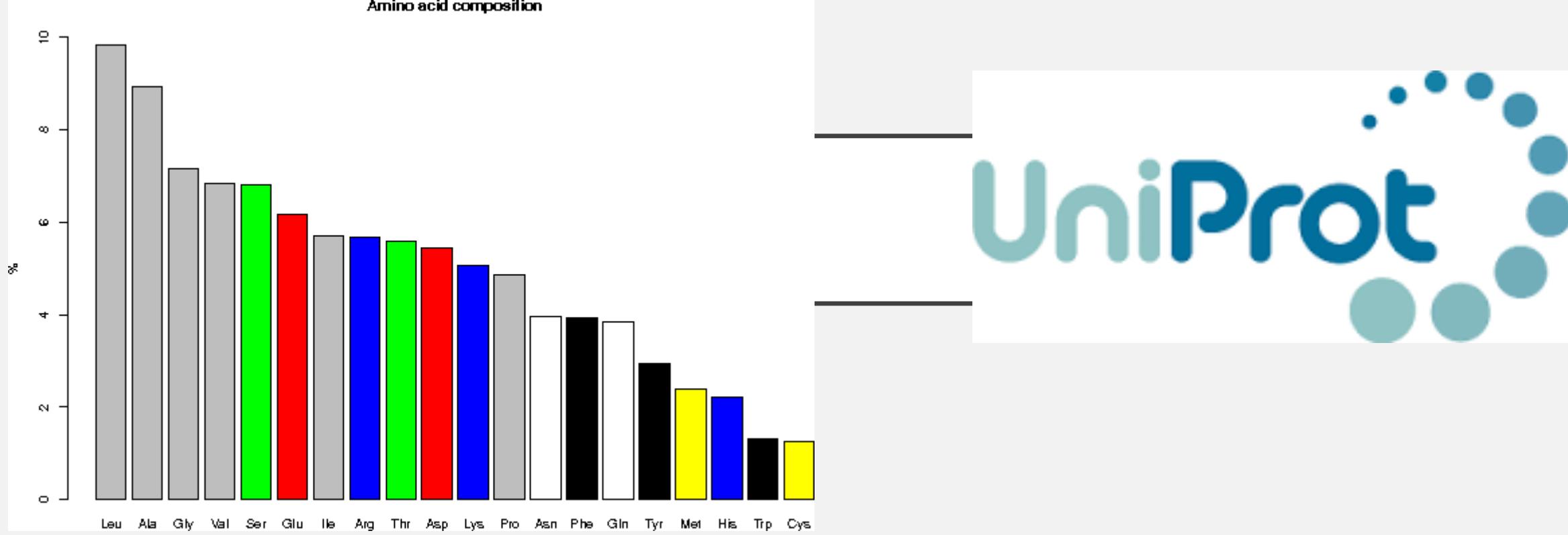


Length distribution of the sequences

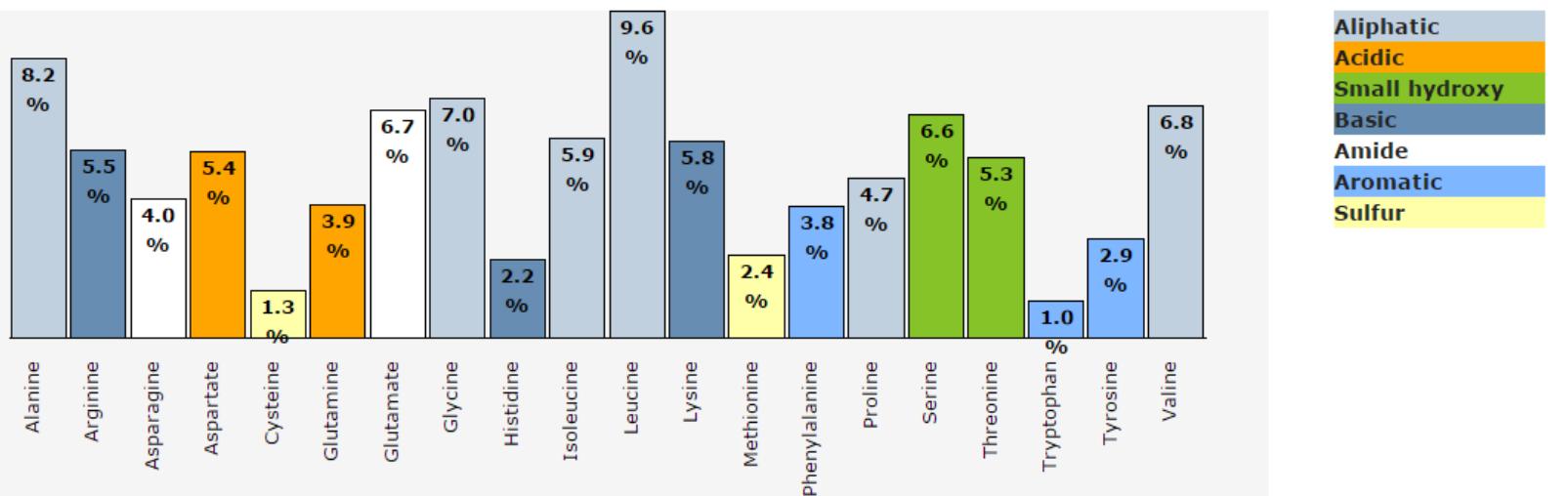


Sequence length distribution in UniProtKB/Swiss-Prot





Amino acid distribution statistics



UniProt



HUGO GENE



The resource for approved human gene nomenclature

Search symbols, keywords or IDs



Last updated: 2021-04-08

HUGO GENE

HGNC

Gene data ▾ Tools ▾ Downloads ▾ VGNC ▾ Contact us ▾ More ▾ Request symbol ⚡

Search symbols, keywords or IDs

Symbol report for PTEN stable symbol ?

Report HCOP homology predictions

HGNC data for PTEN

Approved symbol ? PTEN

Approved name ? phosphatase and tensin homolog

Locus type ? gene with protein product

HGNC ID ? HGNC:9588

Symbol status ? Approved

Previous symbols ? BZS; MHAM

Alias symbols ? MMAC1; TEP1; PTEN1

Alias names ? " mutated in multiple advanced cancers 1 "

Chromosomal location ? 10q23.31

Gene groups ? PTEN protein phosphatases
C2 tensin-type domain containing
Phosphoinositide phosphatases

Gene resources for PTEN

Ensembl ENSG00000171862  Curated

Ensembl region in detail ,
Ensembl gene sequence 

UCSC uc001kfb.4 

NCBI Gene 5728  Curated

Alliance of Genome
Resources HGNC:9588 

Nucleotide resources for PTEN

INSDC U92436  Curated

ENA , GenBank , DDBJ 

CCDS CCDS31238  Curated

RefSeq NM_000314  Curated

NCBI sequence viewer 

Protein resources for PTEN

UniProt/Swiss-Prot P60484 

InterPro , PDBe , Reactome 

+ Enzyme Commission

3.1.3.16  Curated

Expasy , BRENDA , KEGG , IntEnz , IUBMB Enzyme Nomenclature 

Orthologs from selected species for PTEN

Bos taurus PTEN (VGNC:33496 ) 

Canis familiaris PTEN (VGNC:45138 ) 

Equus caballus PTEN (VGNC:21984 ) 

Felis catus PTEN (VGNC:69138 ) 

Mus musculus Pten (MGI:109583 )  Curated

Pan troglodytes PTEN (VGNC:12442 ) 

Rattus norvegicus Pten (RGD:61995 )



 NCBI  Entrez, The Life Sciences Search Engine.

HOME | SEARCH | SITE MAP | PubMed | Entrez | Human Genome | GenBank | Map Viewer | BLAST

Search across databases Help

Welcome to the new Entrez cross-database search page

 PubMed: biomedical literature citations and abstracts	 Books: online books
 PubMed Central: free, full text journal articles	 OMIM: online Mendelian Inheritance in Man
 Journals: detailed information about journals in Entrez	 Site Search: NCBI web and FTP sites
 Nucleotide: sequence database (GenBank)	 UniGene: gene-oriented clusters of transcript sequences
 Protein: sequence database	 CDD: conserved protein domain database
 Genome: whole genome sequences	 3D Domains: domains from Entrez Structure
 Structure: three-dimensional macromolecular structures	 UniSTS: markers and mapping data
 Taxonomy: organisms in GenBank	 PopSet: population study data sets
 SNP: single nucleotide polymorphism	 GEO: expression and molecular abundance profiles
	 GEO DataSets: experimental sets of GEO data

Enter terms and click 'GO' to run the search against ALL the databases, OR
Click Database Name or Icon to go directly to the Search Page for that database, OR
Click Question Mark for a short explanation of that database.

Figure 1 - Entrez Cross-database Search Page

NCBI/GENE

- <https://www.ncbi.nlm.nih.gov/gene/>

NCBI Resources How To

Gene Advanced Search Help

Gene

Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.

Using Gene

- Gene Quick Start
- FAQ
- Download/FTP
- RefSeq Mailing List
- Gene News
- Factsheet

Gene Tools

- Submit GeneRIFs
- Submit Correction
- Statistics
- BLAST
- Genome Workbench
- Splign

Other Resources

- OMIM
- RefSeq
- RefSeqGene
- Protein Clusters

NCBI/GENE

GENE

Was this helpful?



PTEN – phosphatase and tensin homolog

Homo sapiens (human)

Also known as: 10q23del, BZS, CWS1, DEC, GLM2, MHAM, MMAC1, PTEN1, PTENbeta, TEP1

Gene ID: 5728

[RefSeq transcripts](#) (3) [RefSeq proteins](#) (3) [RefSeqGene](#) (1) [PubMed](#) (2,413)

Orthologs

Genome Browser

BLAST

Download

RefSeq Sequences



Search results

Items: 1 to 20 of 6561

<< First < Prev Page of 329 Next > Last >>

[See also 58 discontinued or replaced items.](#)

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> PTEN ID: 5728	phosphatase and tensin homolog [<i>Homo sapiens</i> (human)]	Chromosome 10, NC_000010.11 (87863625..87971930)	10q23del, BZS, CWS1, DEC, GLM2, MHAM, MMAC11, PTENbeta, TEP1, PTEN	601728
<input type="checkbox"/> Pten ID: 19211	phosphatase and tensin homolog [<i>Mus musculus</i> (house mouse)]	Chromosome 19, NC_000085.7 (32734977..32803560)	2310035O07Rik, A130070J02Rik, AI463227, B430203M17Rik, MMAC, MMAC1, PTENbeta, TEP, TEP1	
<input type="checkbox"/> Pten ID: 50557	phosphatase and tensin homolog [<i>Rattus norvegicus</i> (Norway rat)]	Chromosome 1, NC_051336.1 (230631303..230696754)	MMAC1, Mmac, TEP1	

NCBI/GENOME

Genome Genome homo sapiens[orgn]
Create alert Limits Advanced

Homo sapiens (human)
Reference genome: **Homo sapiens (assembly GRCh38.p13)**
Download sequences in FASTA format for **genome, transcript, protein**
Download genome annotation in **GFF, GenBank or tabular format**
BLAST against **Homo sapiens genome, transcript, protein**

All 908 genomes for species:
[Browse the list](#)
Download sequence and annotation from [RefSeq](#) or [GenBank](#)
NEW Try [NCBI Datasets](#) - a new way to download genome sequence and annotation we're testing in NCBI Labs

Display Settings: Overview Send to: ID: 51

[Organism Overview](#) ; [Genome Assembly and Annotation report \[908\]](#) ; [Organelle Annotation Report \[25\]](#)

 **Homo sapiens (human)**
Human genome projects have generated an unprecedented amount of knowledge about human genetics and health.

Lineage: [Eukaryota](#)[6800]; [Metazoa](#)[2884]; [Chordata](#)[1712]; [Craniata](#)[1691]; [Vertebrata](#)[1691]; [Euteleostomi](#)[1675]; [Mammalia](#)[454]; [Eutheria](#)[430]; [Euarchontoglires](#)[163]; [Primates](#)[60]; [Haplorrhini](#)[41]; [Catarrhini](#)[29]; [Hominidae](#)[6]; [Homo](#)[1]; [Homo sapiens](#)[1]

Study of the human condition such as genetic and infectious disease, the intersection between genetics and the environment, and population variation is supported by a wealth of genome-scale data. These data sets include: a) numerous sequenced genomes including several which have been assembled; b) studies that examine transcript and protein existence, [More...](#)

 **Summary**

Sequence data: genome assemblies: 908; sequence reads: 606 (See [Genome Assembly and Annotation report](#))
Statistics: median total length (Mb): 2859.97
median protein count: 122962
median GC%: 40.4
NCBI Annotation Release: 109

NCBI/GENOME

Representative (genome information for reference and representative genomes)

Reference genome:

- *Homo sapiens GRCh38.p13*

Submitter: Genome Reference Consortium

Loc	Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Other RNA	Gene	Pseudogene	
	Chr	1	NC_000001.11	CM000663.2	248.96	42.3	11,288	17	90	4,612	5,091	1,416	
	Chr	2	NC_000002.12	CM000664.2	242.19	40.3	8,519	-	7	3,834	3,879	1,207	
	Chr	3	NC_000003.12	CM000665.2	198.3	39.7	7,566	-	4	2,869	2,988	914	
	Chr	4	NC_000004.12	CM000666.2	190.22	38.3	4,763	-	1	2,258	2,443	807	
	Chr	5	NC_000005.10	CM000667.2	181.54	39.5	4,892	-	17	2,263	2,603	792	
	Chr	6	NC_000006.12	CM000668.2	170.81	39.6	5,756	-	138	2,588	3,020	893	
	Chr	7	NC_000007.14	CM000669.2	159.35	40.7	5,359	-	22	2,482	2,782	913	
	Chr	8	NC_000008.11	CM000670.2	145.14	40.2	4,157	-	4	2,018	2,172	678	
	Chr	9	NC_000009.12	CM000671.2	138.4	42.3	4,764	-	3	2,284	2,275	728	
	Chr	10	NC_000010.11	CM000672.2	133.8	41.6	5,518	-	3	2,231	2,181	643	
	Chr	11	NC_000011.10	CM000673.2	135.09	41.6	6,800	-	13	2,543	2,920	842	
	Chr	12	NC_000012.12	CM000674.2	133.28	40.8	6,190	-	9	2,575	2,540	701	
	Chr	13	NC_000013.11	CM000675.2	114.36	40.2	2,116	-	4	1,290	1,384	475	
	Chr	14	NC_000014.9	CM000676.2	107.04	42.2	3,590	-	18	1,746	2,061	587	
	Chr	15	NC_000015.10	CM000677.2	101.99	43.4	3,632	-	9	1,838	1,827	573	
	Chr	16	NC_000016.10	CM000678.2	90.34	45.1	4,697	-	27	1,823	1,955	486	
	Chr	17	NC_000017.11	CM000679.2	83.26	45.3	6,269	-	33	2,298	2,461	579	
	Chr	18	NC_000018.10	CM000680.2	80.37	39.8	2,052	-	1	1,001	985	301	
	Chr	19	NC_000019.10	CM000681.2	58.62	47.9	6,944	-	6	2,001	2,494	530	
	Chr	20	NC_000020.11	CM000682.2	64.44	43.9	2,871	-	-	1,360	1,359	344	
	Chr	21	NC_000021.9	CM000683.2	46.71	42.2	1,291	12	1	715	778	208	
	Chr	22	NC_000022.11	CM000684.2	50.82	47.7	2,525	-	-	1,034	1,186	360	
	Chr	X	NC_000023.11	CM000685.2	156.04	39.6	3,830	-	4	1,297	2,220	916	
	Chr	Y	NC_000024.10	CM000686.2	57.23	45.4	335	-	-	321	583	398	
		MT	NC_012920.1	J01415.2	0.02	44.4	13	2	22	-	37	-	
		Un	-	-	-	183.8	44.3	7,225	33	161	4,057	7,014	2,076



GEO DataSets

This database stores curated gene expression DataSets, as well as original Series and Platform records in the Gene Expression Omnibus (GEO) repository. Enter search terms to locate experiments of interest. DataSet records contain additional resources including cluster tools and differential expression queries.

Getting Started

[GEO Documentation](#)[GEO FAQ](#)[About GEO DataSets](#)[Construct a Query](#)[Download Options](#)

GEO Tools

[Submit to GEO](#)[Advanced Search](#)[DataSet Browser](#)[Programmatic Access](#)[GEO2R](#)

More Resources

[GEO Home](#)[GEO Profiles](#)[SRA](#)

NCBI/GEO

NCBI/GEO

GEO DataSets

GEO DataSets

(PDAC) AND "Homo sapiens"[porgn:txid9606]

Create alert Advanced

Entry type clear Summary ▾ 20 per page ▾ Sort by Default order ▾

Send to: ▾ F

Entry type

DataSets (0)

✓ Series (269)

Samples (0)

Platforms (0)

Organism

Customize ...

Study type

Expression profiling by array

Methylation profiling by array

Customize ...

Author

Customize ...

Attribute name

tissue (103)

strain (4)

Customize ...

Publication dates

30 days

1 year

Custom range...

[Clear all](#)

[Show additional filters](#)

Search results

Items: 1 to 20 of 269

<< First < Prev Page 1 of 14 Next > Last >>

Filters activated: Series. [Clear all](#) to show 4201 items.

[Identification and Functional Analysis of Novel Oncogenes in Pancreatic Ductal Adenocarcinoma](#)

1. (Submitter supplied) For identifying the specific expressed genes in PDAC, we constructed sequencing libraries from polyadenylated-RNA extracted from 6 PDAC specimens, 3 non-tumor adjacent tissues and 3 pancreas tissues from normal individuals.

Organism: **Homo sapiens**

Type: Expression profiling by high throughput sequencing

Platform: GPL11154 6 Samples

Download data: CSV

Series Accession: GSE171485 ID: 200171485

[SRA Run Selector](#)

[The muscle and adipose transcriptome data from patients with PDAC cachexia](#)

2. (Submitter supplied) The aim of the study is to identify genes and pathways associated with muscle and adipose wasting in PDAC cachexia. Muscle and adipose were collected from same individuals to study the concurrent muscle and adipose wasting.

Organism: **Homo sapiens**

Type: Expression profiling by high throughput sequencing

Platform: GPL17303 68 Samples

Download data: TXT

Series Accession: GSE133979 ID: 200133979

[SRA Run Selector](#)

[Pancreatic cancer prognosis is predicted by chromatin accessibility](#)

3. (Submitter supplied) This SuperSeries is composed of the SubSeries listed below.

Organism: **Homo sapiens**

PDB - RCSB

- <https://www.rcsb.org>

The screenshot shows the RCSB PDB homepage. At the top left is the RCSB PDB logo. To its right is a search bar with the placeholder "Enter search term(s)" and a magnifying glass icon. Below the search bar are links for "Advanced Search" and "Browse Annotations". A "Help" link is also present. The top right features a banner for "Celebrating 25 YEARS OF Protein Data Bank" with social media icons for Facebook, Twitter, YouTube, and LinkedIn. On the far left, a vertical sidebar menu lists "Welcome", "Deposit", "Search", "Visualize", "Analyze", "Download", and "Learn", each with a corresponding icon. The main content area has a header "A Structural View of Biology" with a subtext about the archive's purpose. It includes two images: one for "COVID-19 CORONAVIRUS Resources" showing a 3D model of a virus, and another for "PDB50: A special symposium" featuring a molecular structure. To the right is a large image of the SARS-CoV-2 Spike protein bound to antibodies, with the caption "April Molecule of the Month".



The image shows the RCSB Protein Data Bank (PDB) website interface. At the top left is the RCSB PDB logo with the text "176528 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education". To the right is a search bar with the placeholder "Enter search term(s)" and a magnifying glass icon. Below the search bar are links for "Advanced Search" and "Browse Annotations".

A banner at the top right celebrates "Celebrating 50 YEARS OF Protein Data Bank" with icons for social media (Facebook, Twitter, YouTube) and a map of the world.

The main content area features a navigation bar with tabs: Structure Summary, 3D View, Annotations, Experiment, Sequence, Genome, and Versions. The "3D View" tab is currently selected, showing a ribbon diagram of a protein structure labeled "Biological Assembly 1". Below the diagram are links for "3D View: Structure" and "Electron Density".

The protein structure itself is composed of multiple colored domains (red, green, blue, orange) forming a complex fold. To the right of the structure, the identifier "1XNI" is displayed in large letters, followed by the title "Tandem Tudor Domain of 53BP1".

Below the title are details about the structure: DOI: 10.2210/pdb1XNI/pdb, Classification: CELL CYCLE, Organism(s): Homo sapiens, Expression System: Escherichia coli BL21, and Mutation(s): No.

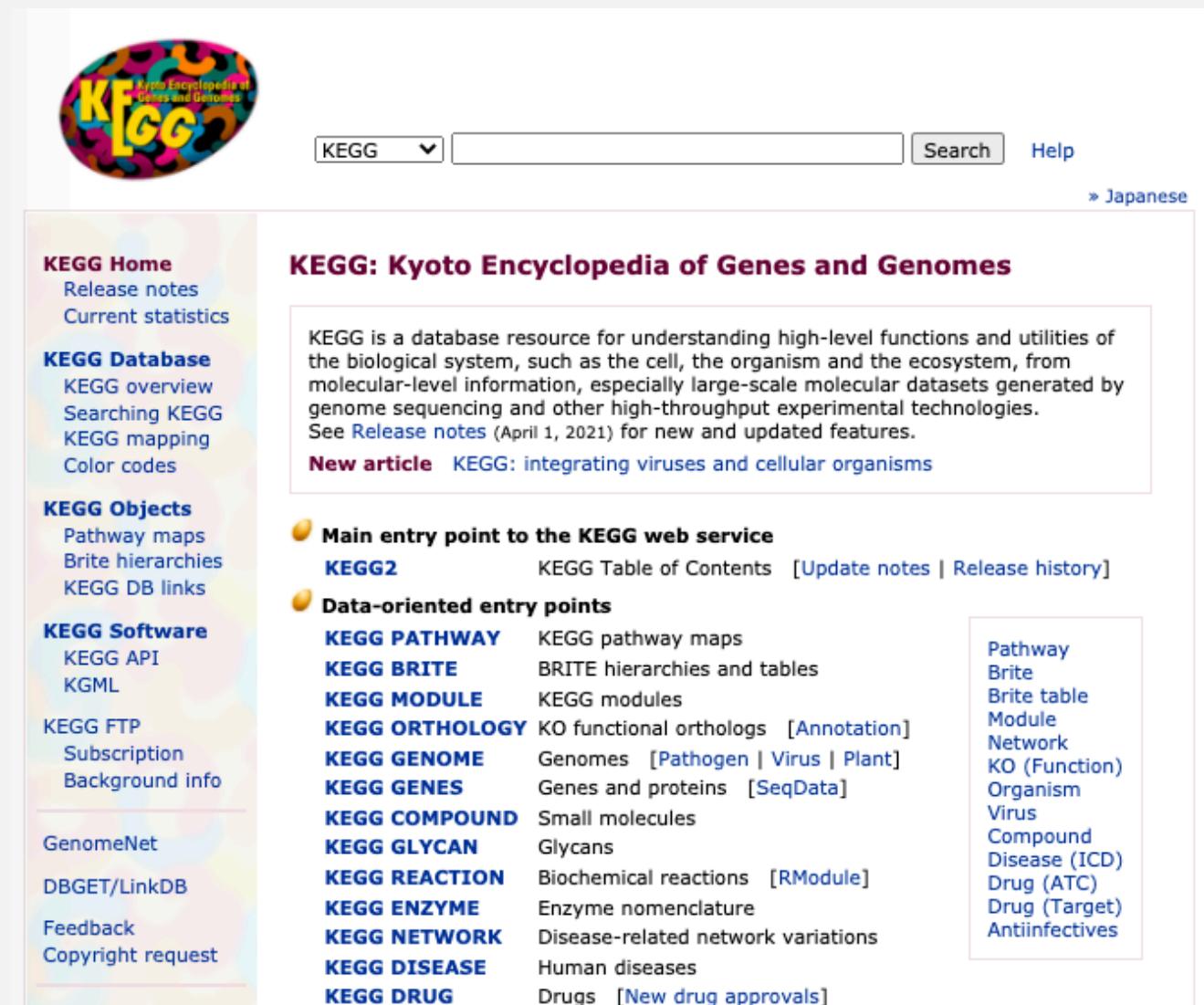
Information about deposition and release is provided: Deposited: 2004-10-05, Released: 2004-11-30, and Deposition Author(s): Huyen, Y., Zgheib, O., DiTullio Jr., R.A., Gorgoulis, V.G., Zacharatos, P., Petty, T.J., Sheston, E.A., Mellert, H.S., Stavridi, E.S., Halazonetis, T.D.

The "Experimental Data Snapshot" section lists the method as X-RAY DIFFRACTION, resolution as 2.80 Å, and R-Value Free as 0.269. The "wwPDB Validation" section includes a table with validation metrics and percentile ranks:

Metric	Percentile Ranks	Value
Rfree	27	0.267
Clashscore	27	0.9%
Ramachandran outliers	27	0.9%

KEGG: KYOTO ENCYCLOPEDIA OF GENES AND GENOMES

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



The screenshot shows the KEGG homepage with a dark header and sidebar. The header features the KEGG logo and navigation links for "KEGG", "Search", and "Help". A "Japanese" link is also present. The sidebar contains links for "KEGG Home", "Release notes", "Current statistics", "KEGG Database", "KEGG Objects", "KEGG Software", and "KEGG FTP". The main content area is titled "KEGG: Kyoto Encyclopedia of Genes and Genomes" and describes the database's purpose in understanding biological systems. It includes a "New article" section and a "Main entry point to the KEGG web service" section listing various databases like KEGG PATHWAY, KEGG BRITE, and KEGG DRUG, each with a brief description and links to more information. A sidebar on the right lists categories such as Pathway, Brite, Module, Network, KO (Function), Organism, Virus, Compound, Disease (ICD), Drug (ATC), Drug (Target), and Antiinfectives.

KEGG Home
Release notes
Current statistics

KEGG Database
KEGG overview
Searching KEGG
KEGG mapping
Color codes

KEGG Objects
Pathway maps
Brite hierarchies
KEGG DB links

KEGG Software
KEGG API
KGML

KEGG FTP
Subscription
Background info

GenomeNet

DBGET/LinkDB

Feedback
Copyright request

KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies. See [Release notes](#) (April 1, 2021) for new and updated features.

New article KEGG: integrating viruses and cellular organisms

Main entry point to the KEGG web service

KEGG2 KEGG Table of Contents [Update notes | Release history]

Data-oriented entry points

KEGG PATHWAY	KEGG pathway maps
KEGG BRITE	BRITE hierarchies and tables
KEGG MODULE	KEGG modules
KEGG ORTHOLOGY	KO functional orthologs [Annotation]
KEGG GENOME	Genomes [Pathogen Virus Plant]
KEGG GENES	Genes and proteins [SeqData]
KEGG COMPOUND	Small molecules
KEGG GLYCAN	Glycans
KEGG REACTION	Biochemical reactions [RModule]
KEGG ENZYME	Enzyme nomenclature
KEGG NETWORK	Disease-related network variations
KEGG DISEASE	Human diseases
KEGG DRUG	Drugs [New drug approvals]

Pathway
Brite
Brite table
Module
Network
KO (Function)
Organism
Virus
Compound
Disease (ICD)
Drug (ATC)
Drug (Target)
Antiinfectives

KEGG: KYOTO ENCYCLOPEDIA OF GENES AND GENOMES

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

