

Taller 1: Uso de herramientas puntuales de genómica comparativa



Profesores: Mauricio Quimbaya-Ph.D. Y Fabián Tobar Tossé-Ph.D.

Pontificia Universidad Javeriana Cali

Proyecto 1: Epigenómica, Redes Genéticas, Genómica Funcional y Estructural de Cultivos

www.omicas.co



El futuro
es de todos

Gobierno
de Colombia



Un problema de genómica comparativa

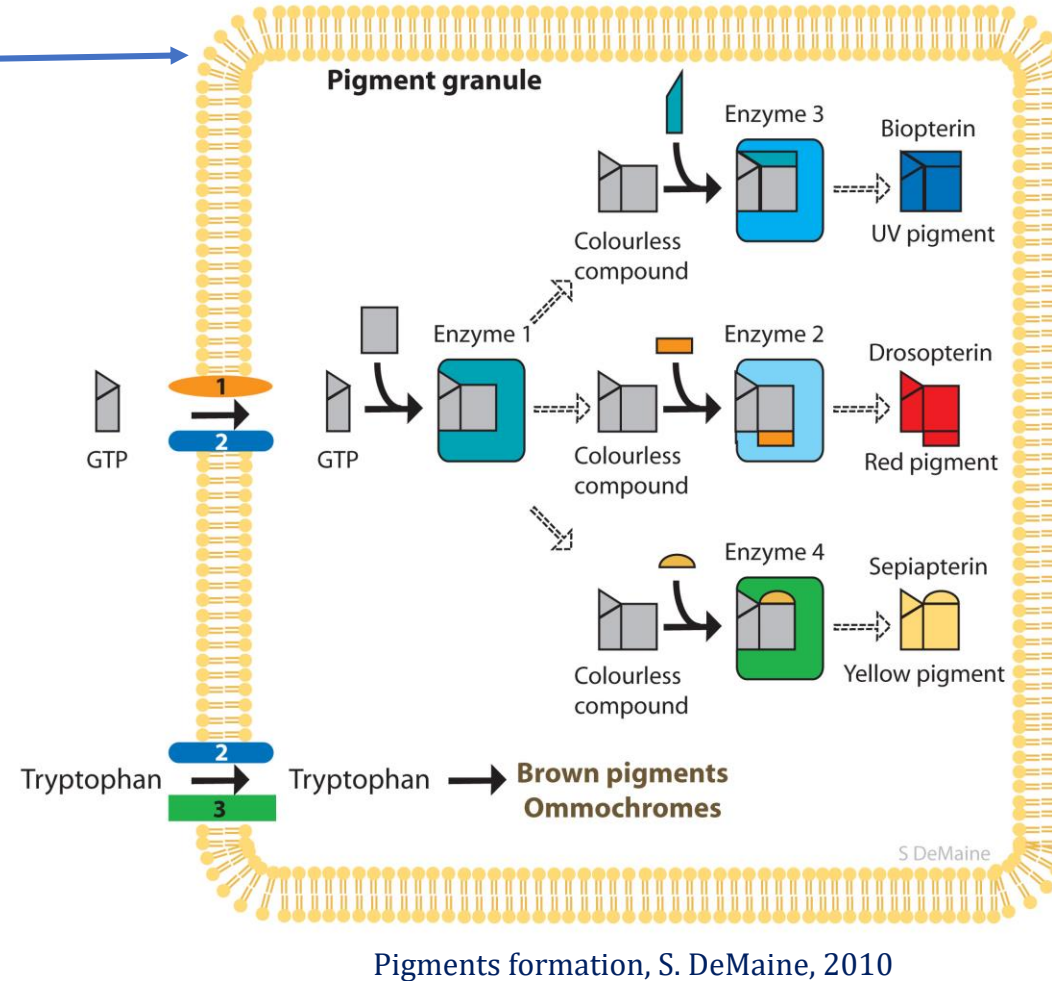
El fenotipo está asociado a procesos celulares puntuales

Existe un fenotipo



Ruta de síntesis de las drosópteras y xantópteras

El proceso está determinado por la acción de genes específicos



Existe un fenotipo equivalente

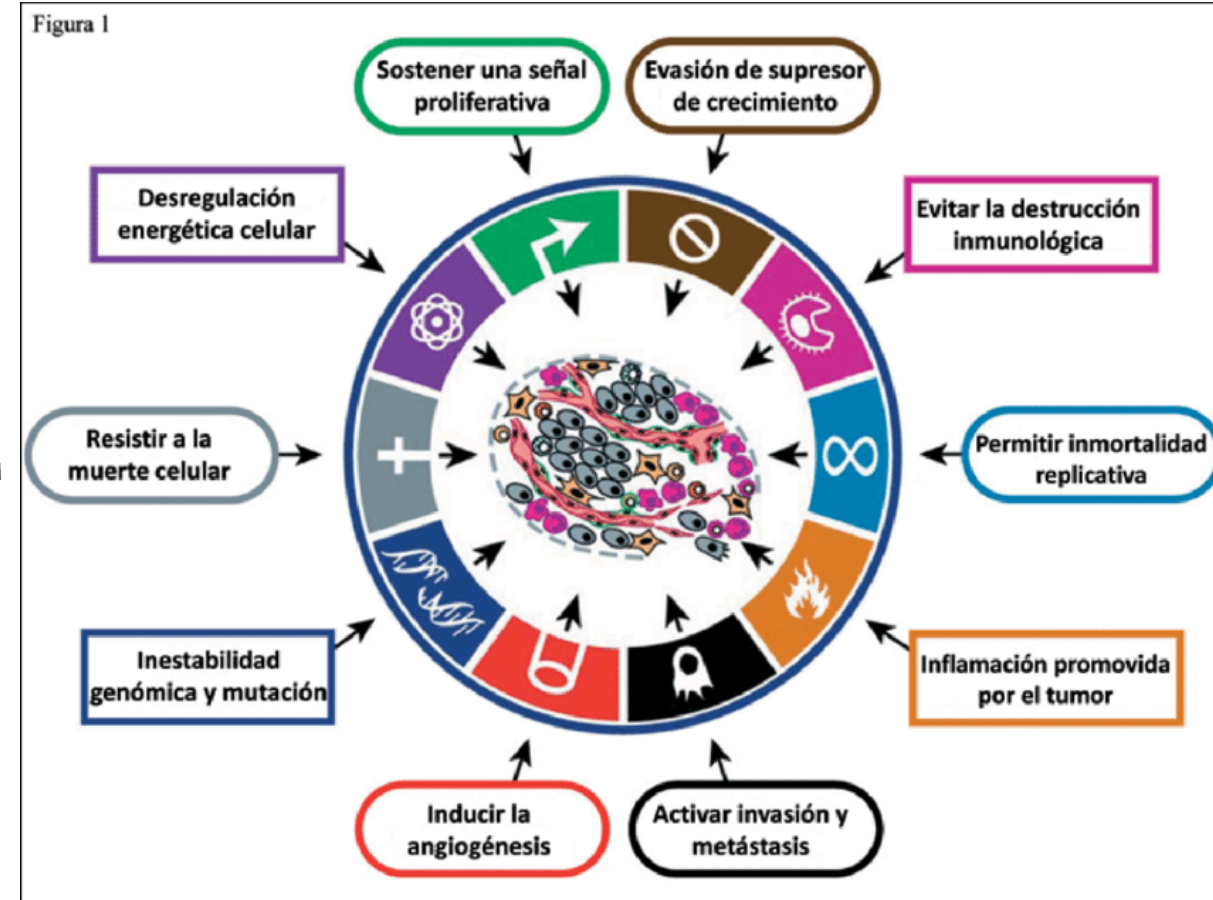
¿Está este nuevo fenotipo soportado por las mismas rutas y los mismos genes?

Un problema de genómica comparativa



¿Son los mismos procesos celulares los que determinan la altura en plantas y humanos?

Hallmarks of Cancer: Next Generation. Douglas Hanahan & Robert Weinberg. 2011



¿Existen procesos celulares equivalentes a las marcas distintivas del cáncer en plantas?

Fenólogos = Relaciones de ortología entre genes

Systematic discovery of nonobvious human disease models through orthologous phenotypes

Kriston L. McGary^{a,1}, Tae Joo Park^{a,b,1}, John O. Woods^a, Hye Ji Cha^a, John B. Wallingford^{a,b}, and Edward M. Marcotte^{a,c,2}

^aCenter for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, ^bThe Howard Hughes Medical Institute and Department of Molecular Cell and Developmental Biology, and ^cDepartment of Chemistry and Biochemistry, University of Texas, Austin, TX 78712

Edited* by William H. Press, University of Texas, Austin, TX, and approved February 26, 2010 (received for review September 6, 2009)

¿Qué es el genoma?

Es el código que define y estructura las instrucciones necesarias para el funcionamiento celular.

Está constituido por el **total de la información genética** que portan las células de cualquier entidad biológica.

La información contenida en el genoma es extensa. Si usted es capaz de leer **600 palabras por minuto**, le tomaría **20 años** el poder leer la información que está escrita en su genoma.



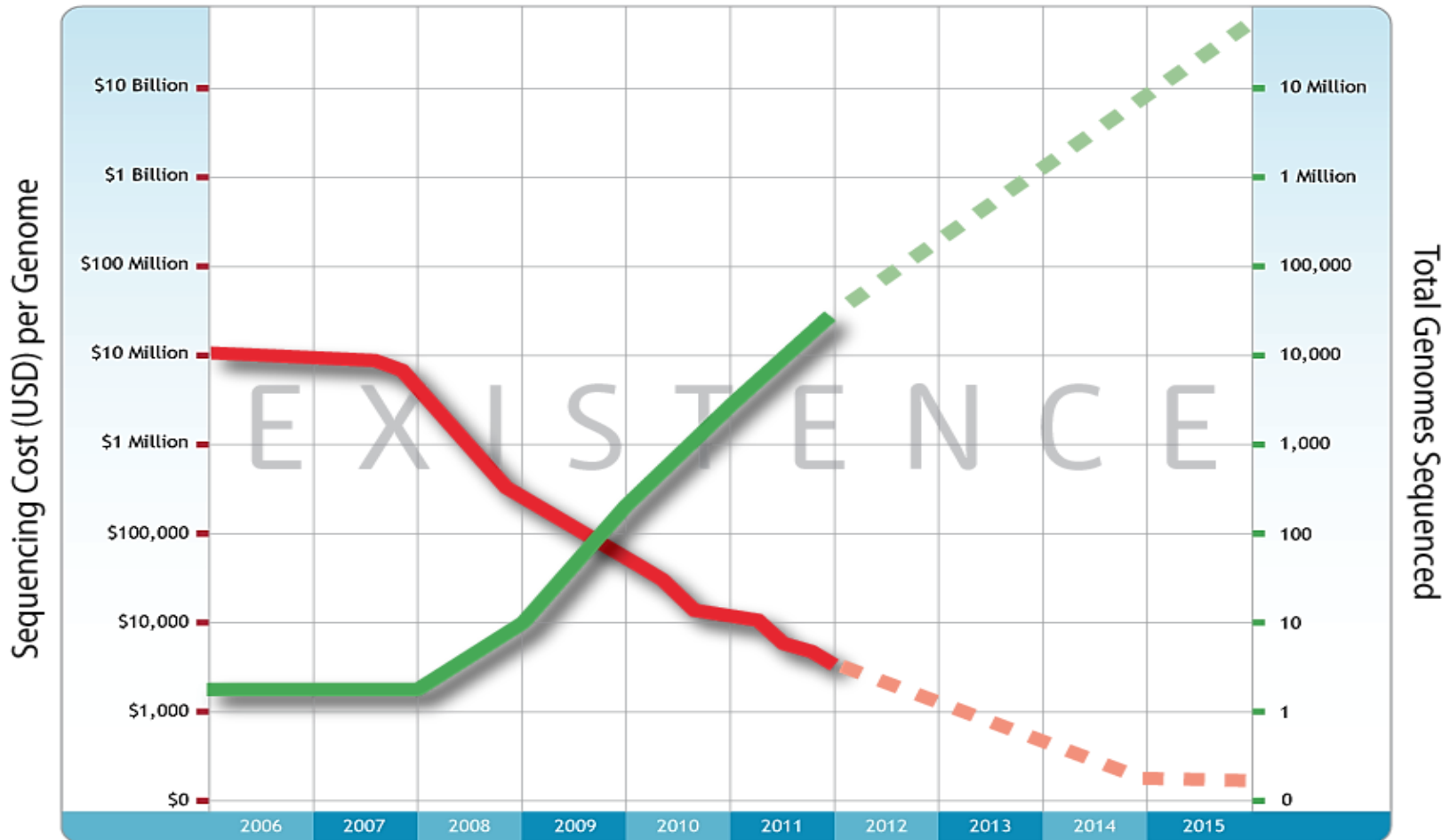
Tomado de <https://www.engr.wisc.edu>

<https://www.youtube.com/watch?v=s6rJLXq1Re0>

¿Qué es el genoma?

Full Genome Sequencing & The Genetic Revolution

Cost per Human Genome vs Total Number of Genomes Sequenced



Sequencing Projects

- Complete Projects 19.157
- Permanent Drafts 201.055
- Incomplete Projects 100.358
- Targeted Projects 1.090

Genómica



El futuro
es de todos

Gobierno
de Colombia



¿Qué es la genómica comparativa?

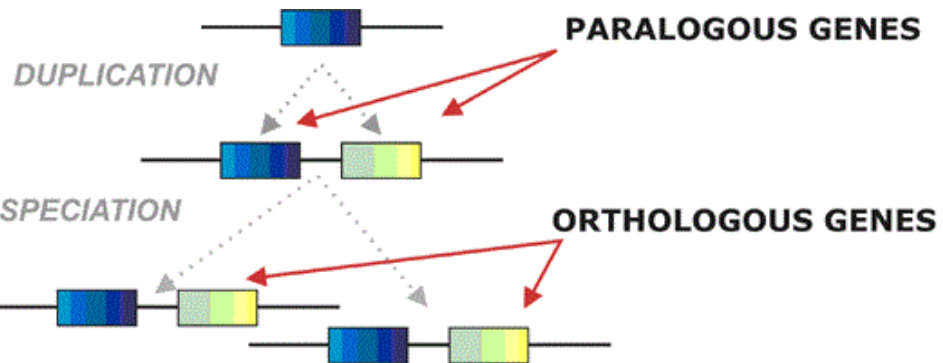


La genómica comparativa es una estrategia para el entendimiento de la composición y evolución de los genomas a través de la comparación sistemática de los distintos elementos estructurales que lo componen.

- Es útil para saber qué genes, regiones reguladoras o circuitos génicos se han mantenido constantes o invariables en la historia de la vida, identificando los elementos fundamentales para el sostenimiento de procesos biológicos vitales.
- Sirve para determinar cómo ha variando la estructura genómica, determinada por la composición y orden de los genes y sus elementos reguladores, a través del proceso evolutivo.

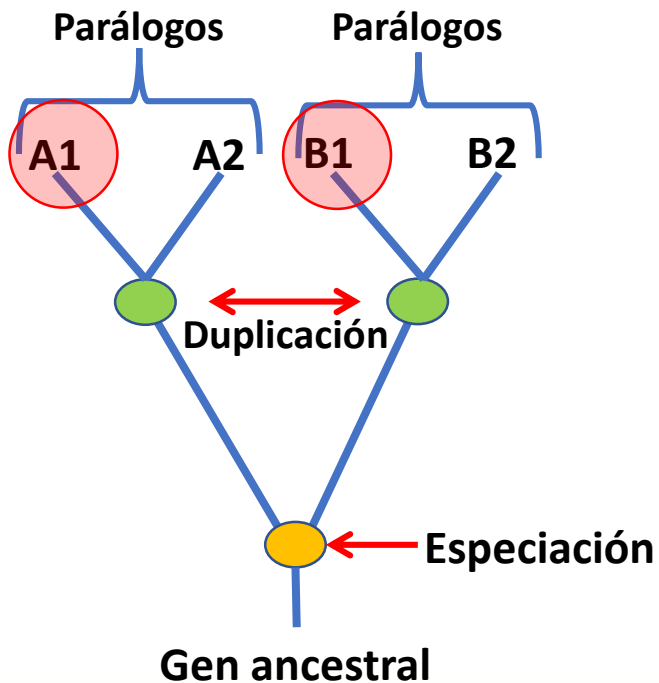
Fenólogos = Relaciones de ortología entre genes

Elementos clave en la genómica comparativa: Ortólogos



Ortólogos: Misma función en distintos genomas.

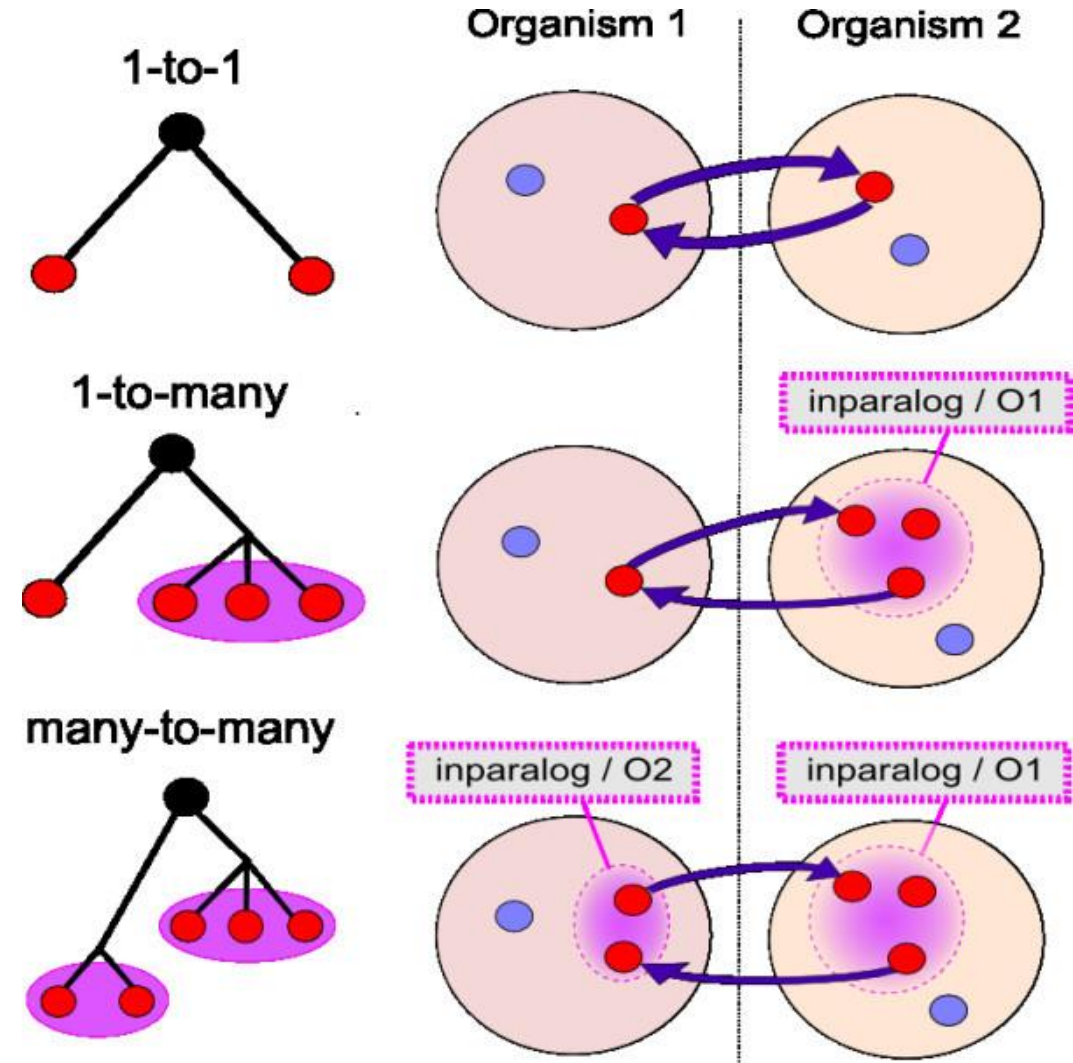
Parálogos: Funciones similares o diferentes en el mismo genoma.



¿Cómo se detectan ortólogos?

Elementos clave en la genómica comparativa: Ortólogos

BLAST: Reciprocal Best Hits



OrthoInspector: comprehensive orthology analysis
and visual exploration por Benjamin Linard, Julie D Thompson, Olivier Poch, Odile Lecompte

Detección de Ortólogos con ENSEMBL

 [BLAST/BLAT](#) | [VEP](#) | [Tools](#) | [BioMart](#) | [Downloads](#) | [Help & Docs](#) | [Blog](#)

Tools
[All tools](#)

BioMart >
Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT >
Search our genomes for your DNA or protein sequence

Variant Effect Predictor >
Analyse your own variants and predict the functional consequences of known and unknown variants

Search

for

e.g. [BRCA2](#) or [rat 5:62797383-63627669](#) or [rs699](#) or [coronary heart disease](#)

All genomes

- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

Favourite genomes

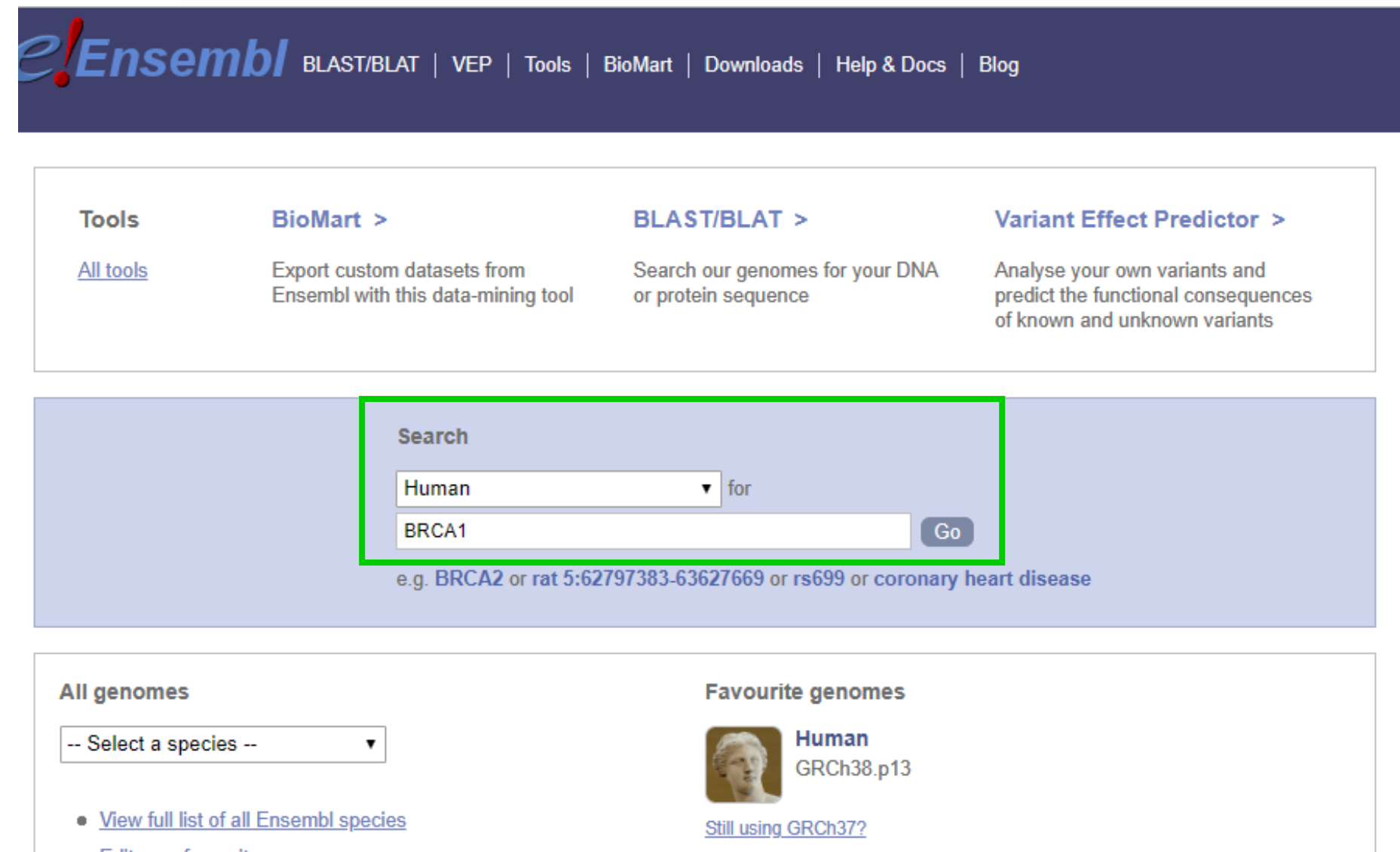
**Human**
GRCh38.p13
[Still using GRCh37?](#)

**Mouse**
GRCm38.p6

ENSEMBL es un repositorio de información genética y genómica, que explora en su mayoría genomas de vertebrados (en este momento ENSEMBL posee información de 150 genomas). Su proceso de curaduría es meticuloso y posee un gran número de herramientas que permiten extraer información específica de un genoma, un gen o una secuencia específica.

<http://www.ensembl.org/info/website/tutorials/index.html>

Detección de Ortólogos con ENSEMBL



The screenshot displays the Ensembl website's navigation bar and main content area. The navigation bar includes links for BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog. Below this, there are four main tool categories: Tools (with a link to All tools), BioMart (for exporting custom datasets), BLAST/BLAT (for searching genomes), and Variant Effect Predictor (for analyzing variants). The BLAST/BLAT section is highlighted with a green box, showing a search interface with a dropdown menu set to 'Human', a text input field containing 'BRCA1', and a 'Go' button. Below the search bar, an example search string is provided: 'e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease'. At the bottom, there are sections for 'All genomes' (with a species selection dropdown) and 'Favourite genomes' (showing 'Human GRCh38.p13' with a link to 'Still using GRCh37?').

Realizaremos un análisis de ortología para el gen **BRCA1** presente en humanos

Detección de Ortólogos con ENSEMBL

 [BLAST/BLAT](#) | [VEP](#) | [Tools](#) | [BioMart](#) | [Downloads](#) | [Help & Docs](#) | [Blog](#)

[New Search](#) [Jobs ▼](#)

Current selection:

< all Species

Only searching Human

Restrict category to:

Gene	20
Transcript	145
Variant	42
Phenotype	6
GeneTree	1
Clones & Regions	1
Protein Domain	12
Protein Family	11

Only searching Human ▼ **BRCA1** 🔍

238 results match BRCA1 when restricted to species: Human ✕

[Did you mean... ▼](#)

[BRCA1 \(Human Gene\)](#)
ENSG00000012048 17:43044295-43170245:-1
BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]

LRG_292 (LRG display in Ensembl gene record; description: Locus Reference Genomic record for **BRCA1**,) is an external reference matched to Gene ENSG00000012048
[Variant table](#) • [Phenotypes](#) • [Location](#) • [External Refs.](#) • [Regulation](#) • [Orthologues](#) • [Gene tree](#)

[BRCA1-226 \(Human Transcript\)](#)
ENST00000591534 17:43045563-43125329:-1
BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100].
[Location](#) • [External Refs.](#) • [cDNA seq.](#) • [Exons](#) • [Variant table](#) • [Protein seq.](#) • [Population](#) • [Protein summary](#)

[BRCA1-227 \(Human Transcript\)](#)
ENST00000591849 17:43045563-43125329:-1

Detección de Ortólogos con ENSEMBL



BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Search Human...

Human (GRCh38.p13) ▼

Location: 17:43,044,295-43,170,245

Gene: BRCA1

Jobs ▼

Gene-based displays

Summary

- Splice variants
- Transcript comparison
- Gene alleles

Sequence

- Secondary structure

Comparative Genomics

- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues

Ensembl protein families

Ontologies

- GO: Molecular function
- GO: Biological process
- GO: Cellular component

Phenotypes

Genetic Variation

- Variant table
- Variant image
- Structural variants

Gene expression

Pathway

Regulation

External references

Supporting evidence

ID History

- Gene history

Gene: BRCA1 ENSG00000012048

Description

BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:[HGNC:1100](#)]

Gene Synonyms

BRCC1, FANCS, PPP1R53, RNF53

Location

[Chromosome 17: 43,044,295-43,170,245](#) reverse strand.
GRCh38:CM000679.2

About this gene

This gene has 34 transcripts ([splice variants](#)), [200 orthologues](#), is a member of [1 Ensembl protein family](#) and is associated with [88 phenotypes](#).

Transcripts

Show transcript table

Summary ?

Name

[BRCA1](#) (HGNC Symbol)

CCDS

This gene is a member of the Human CCDS set: [CCDS11453.1](#), [CCDS11454.2](#), [CCDS11455.2](#), [CCDS11456.2](#), [CCDS11459.2](#)

UniProtKB

This gene has proteins that correspond to the following UniProtKB identifiers: [P38398](#)

RefSeq

This Ensembl/Gencode gene does not contain any transcripts for which we have [selected identical model\(s\) in RefSeq](#). If there are other RefSeq transcripts available they will be in the [External references](#) table

LRG

[LRG_292](#) provides a stable genomic reference framework for describing sequence variants for this gene

Ensembl version

ENSG00000012048.22

Other assemblies

This gene maps to [41,196,312-41,322,262](#) in GRCh37 coordinates.
View this locus in the GRCh37 archive: [ENSG00000012048](#)

Gene type

Protein coding

Genómica



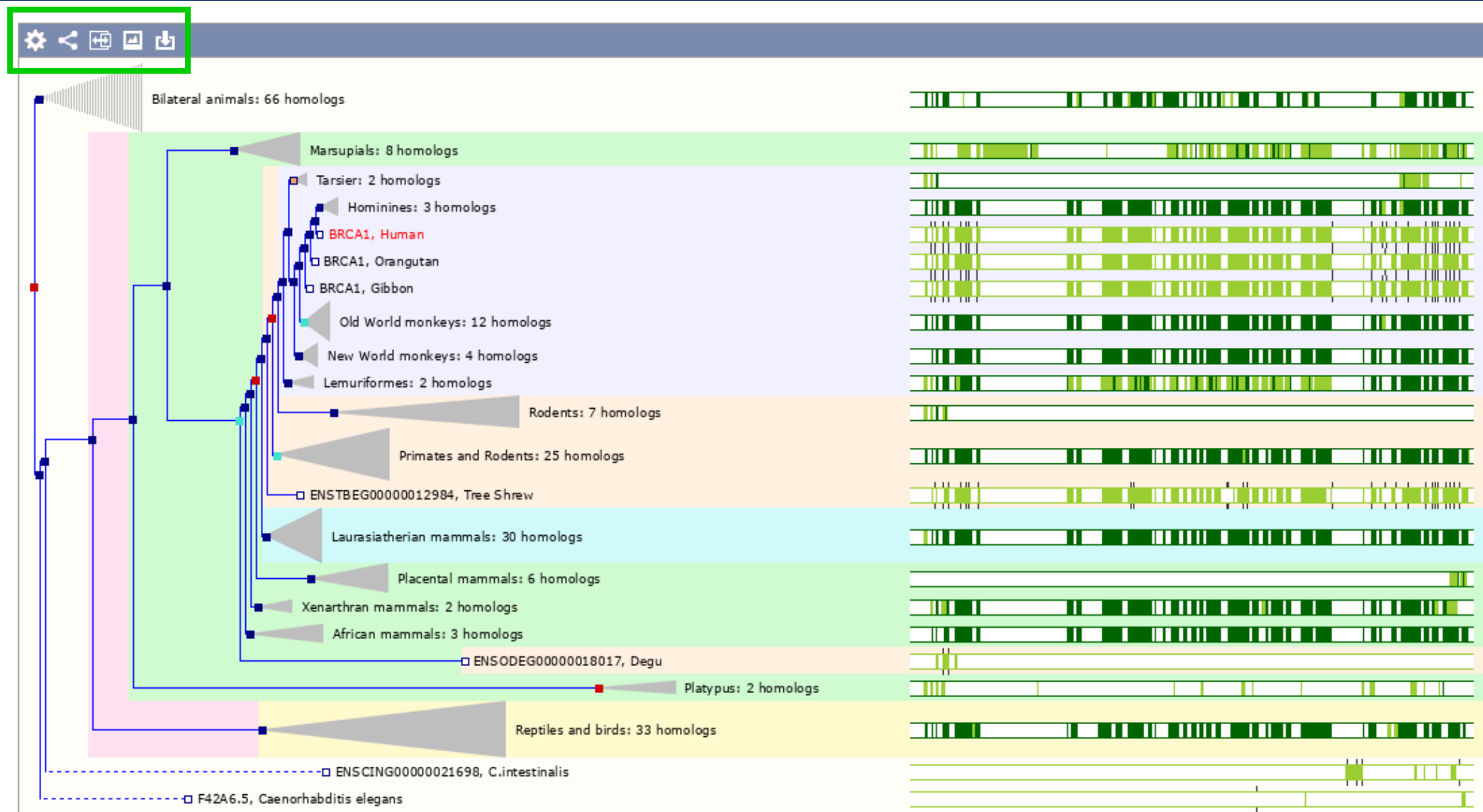
El futuro
es de todos

Gobierno
de Colombia





Detección de Ortólogos con ENSEMBL



Detección de Ortólogos con ENSEMBL

File name:

BRCA1_gene_tree

File format:

FASTA

Preview

Download

Download Compressed

Guide to file formats

CLUSTALW

CLUSTAL W(1.81) multiple sequence

```
homo_sapiens/1-465588 CCTCAGGAC
pan_troglodytes/1-465588 CCCCAGGAC
** *****
```

```
homo_sapiens/1-465588 CCCAGTGCC
pan_troglodytes/1-465588 CCCAGTGCC
*****
```

FASTA

```
>homo_sapiens/1-464300
CCTCAGGACCGACGGCAAACCAACCAGAT
CCCAGTGCCTTCGACTGCCTCCTGGGCC
TGGGACAGAGAGAGAACACAGCTGGCTC
AGGGGCTGGTTGGGGGTTAGATCAAA
CCGACGTGGATCCTGATATTGGCCACCTC
CCAGGCTCTGTGCAAAAAGGTTGCTGTG
AGGAAGACCGTGGTGCCTCTGCCTGTGT
AAAGATGGGGTGGGTGTGGATTCTCTT
GGGAGAGGGAAGAAGAAAGGGCCCTGGG
CAGCCCTGGGGCTCTAGCCCACTGGGC
```

Mega

```
#mega
!Title: ProjectedMultiAlign;
!Format datatype=dna identical=.

#homo_sap CCTCAGGACC GACGGCAAAC
#pan_trog ..C.....

#homo_sap CCCAGTGCCT TCGACTGCCT
#pan_trog .....

#homo_sap TGGGACAGAG AGAGAACCAC
```

MSF

```
ProjectedMultiAlign MSF: 2 Type:

Name: homo_sapiens/1-465588 Le
Name: pan_troglodytes/1-465588 Le

//

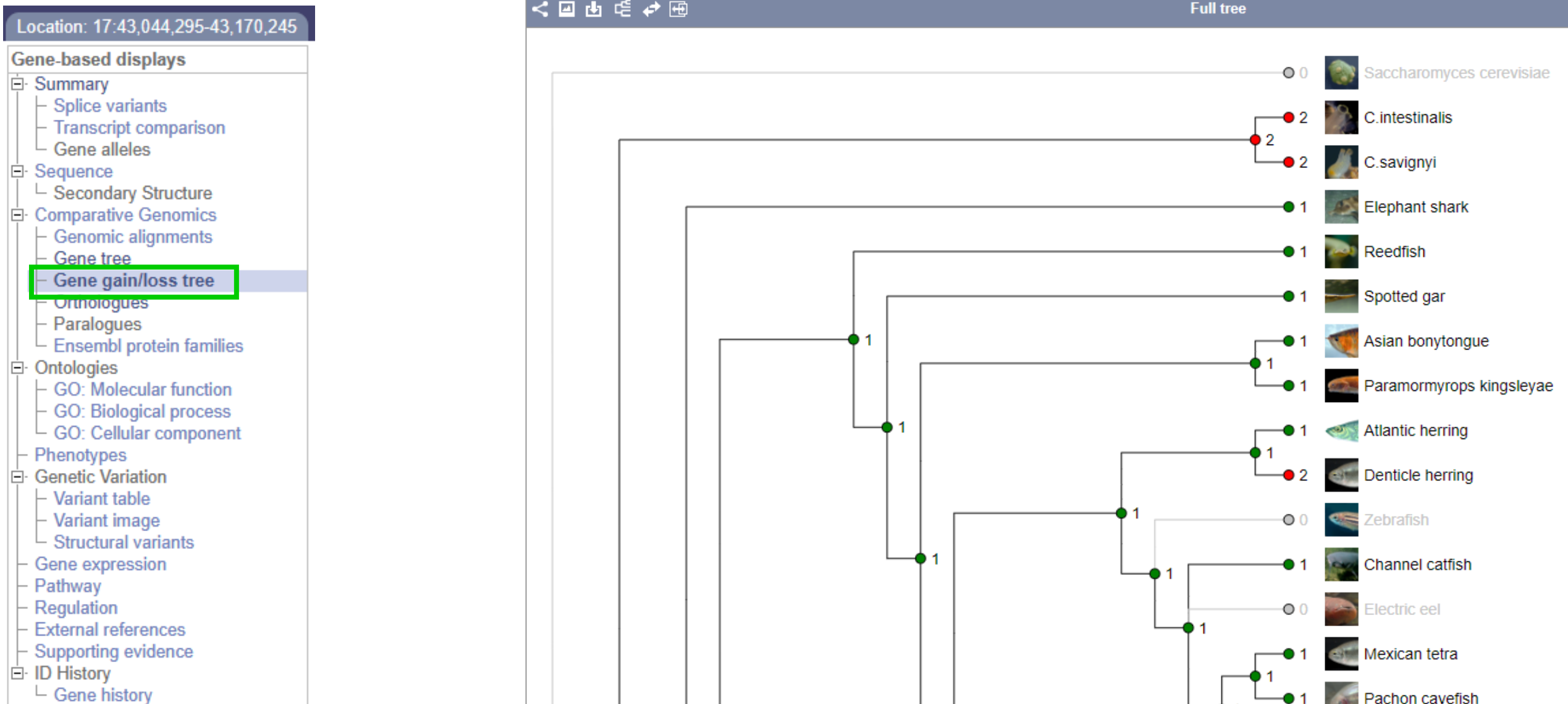
homo_sapiens/1-465588 CCTCAGGACC G
pan_troglodytes/1-465588 CCCCAGGAC

homo_sapiens/1-465588 GGGTCAACAC C
```

Newick

```
(((((ENSTRUP00000015030_Trub_:0.07
ENSTNIP00000002435_Tnig_:0.103149):0.
ENSGACP00000015199_Gacu_:0.161942):0.
((ENSPPOP00000001575_Pfor_:0.042925,
ENSXMAP00000006983_Xmac_:0.044847):0.
ENSONIP00000006940_Onil_:0.29611):0.0
ENSORLP00000004773_Olat_:0.550113):0.1
ENSGMOP00000010385_Gmor_:0.36066):0.1
(ENSAMXP00000013440_Amex_:0.474771,
ENSADAP000000099674_Drez_:0.588918):0.
ENSLOCPO00000009962_Locu_:0.219888):0.
(((((((ENSGALP000000027524_Ggal_:0.03
ENSMGAP00000015990_Mgal_:0.045749):0.
ENSAPLP00000007411_Apla_:0.122877):0.
```

Detección de Ortólogos con ENSEMBL



Detección de Ortólogos con ENSEMBL

Location: 17:43,044,295-43,170,245

 Download orthologues

Gene-based displays

- [-] Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
- [-] Sequence
 - Secondary Structure
- [-] Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree**
 - Orthologues
 - Paralogues
 - Ensembl protein families
- [-] Ontologies
 - GO: Molecular function
 - GO: Biological process
 - GO: Cellular component
- Phenotypes
- [-] Genetic Variation
 - Variant table
 - Variant image
 - Structural variants
- Gene expression
- Pathway
- Regulation
- External references
- Supporting evidence
- [-] ID History
 - Gene history

Summary of orthologues of this gene [Hide](#) ⊖

Click on 'Show details' to display the orthologues for one or more groups of species. Alternatively, click on 'Configure this page' to choose a custom list of species.

Species set	Show details	With 1:1 orthologues	With 1:many orthologues	With many:many orthologues	Without orthologues
Primates (26 species) Humans and other primates	☐	23	2	0	1
Rodents and related species (31 species) Rodents, lagomorphs and tree shrews	☐	28	1	0	2
Laurasiatheria (27 species) Carnivores, ungulates and insectivores	☐	25	2	0	0
Placental Mammals (89 species) All placental mammals	☐	81	5	0	3
Sauropsida (34 species) Birds and Reptiles	☐	33	0	0	1
Fish (60 species) Ray-finned fishes	☐	55	1	0	4
All (199 species) All species, including invertebrates	☒	178	10	0	11

Detección de Ortólogos con ENSEMBL

Location: 17:43,044,295-43,170,245

Gene-based displays

- Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
- Sequence
 - Secondary Structure
- Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree**
 - Orthologues
 - Paralogues
 - Ensembl protein families
- Ontologies
 - GO: Molecular function
 - GO: Biological process
 - GO: Cellular component
- Phenotypes
- Genetic Variation
 - Variant table
 - Variant image
 - Structural variants
- Gene expression
- Pathway
- Regulation
- External references
- Supporting evidence
- ID History
 - Gene history

File name:

File format:

Settings

Sequence to export:

- ☒ Alignments - DNA
- ☐ Alignments - amino acids
- ☐ Unaligned sequences - CDS
- ☐ Unaligned sequences - proteins

Guide to file formats

CLUSTALW

```
CLUSTAL W(1.81) multiple sequence
homo_sapiens/1-465588 CCTCAGGAC
pan_troglodytes/1-465588 CCCCAGGAC
** *****
```

FASTA

```
Text sequence(s):
DNA and/or amino acids
>11 dna:chromosome chromosome:GRCh38:11:10:
CAGCGCGAAGCCACAGGCGCATCCCTAGTAGGGCTACTTGC
TCTGGCCCTCAGACAAGAACTCTCCCCACATTTCAGTTGGC
CCCAAGTATGGAGCAGGCTCAGGCGTGACGCCGGTTGTAGT
TTCTAAATCCCTGTAGACTTACCCTCCCGCCCGCTGGAC
```

Mega

```
#mega
!Title: ProjectedMultiAlign:
!Format datatype=dna identical=.
#homo_sap CCTCAGGACC GACGGCAAAC
#pan_trog ..C.....
```

MSF

```
ProjectedMultiAlign MSF: 2 Type:
Name: homo_sapiens/1-465588 Le
Name: pan_troglodytes/1-465588 Le
//
homo_sapiens/1-465588 CCTCAGGACC G
```

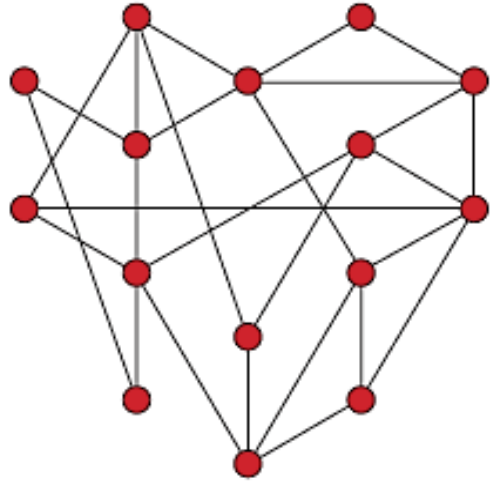
Nexus

```
#NEXUS
[TITLE: ProjectedMultiAlign]
begin data;
dimensions ntax=2 nchar=465588;
format interleave datatype=dna gap=
matrix
homo_sapiens CCTCAGGACC
pan_troglodytes CCCCAGGACC
```

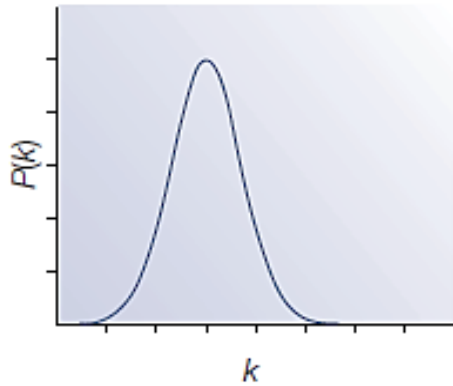
Detección de Ortólogos con ENSEMBL

A Random network

Aa

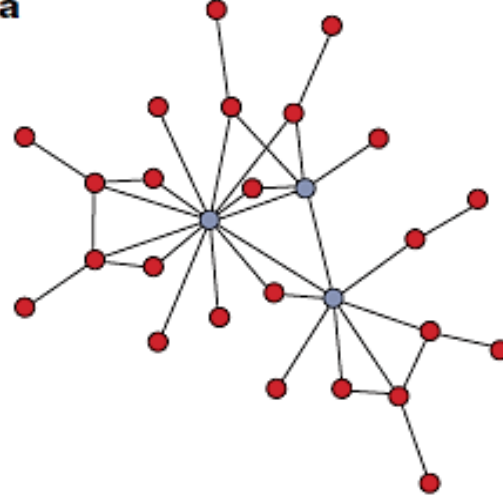


Ab

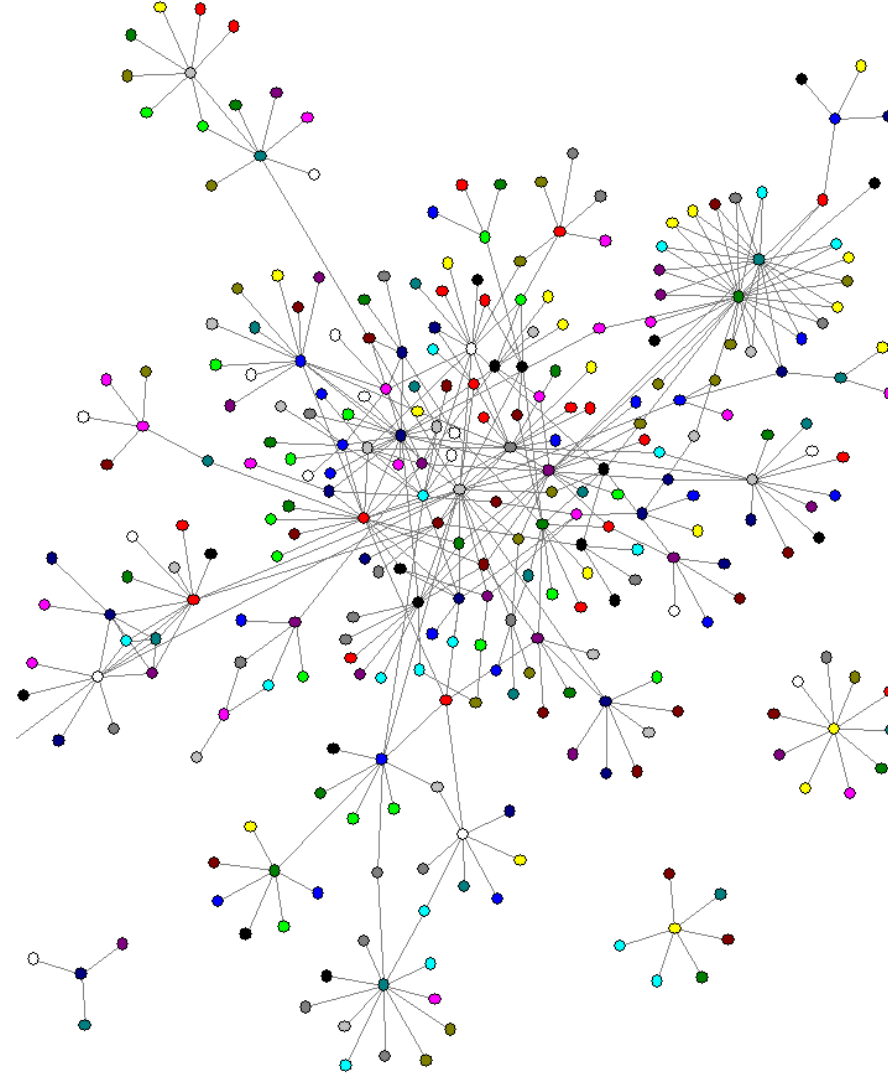
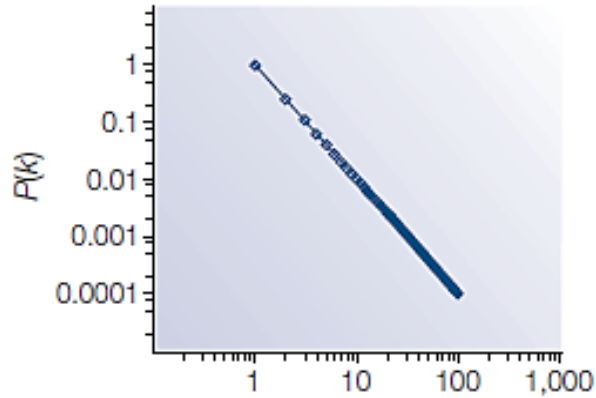


B Scale-free network

Ba



Bb



Nodos
altamente
interconectados
tienden a ser de
copia única

Detección de Ortólogos con ENSEMBL

Ensembl protein families

Ontologies

- GO: Molecular function
- GO: Biological process**
- GO: Cellular component

Phenotypes

Genetic Variation

- Variant table
- Variant image
- Structural variants

Gene expression

Pathway

Regulation

External references

Supporting evidence

ID History

- Gene history

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

GO: Biological process ?

Show All entries		Show/hide columns (1 hidden)		Filter	
Accession	Term	Evidence	Annotation source	Transcript IDs	
GO:0000724	double-strand break repair via homologous recombination	IDA	HGNC	ENST00000493795 ENST00000491747 ENST00000461221 ENST00000471181 ENST00000461798 ENST00000357654 ENST00000468300 ENST00000352993	Search BioMart View on karyotype
GO:0000729	DNA double-strand break processing	TAS	Reactome	ENST00000471181 ENST00000461798 ENST00000357654 ENST00000461221 ENST00000491747 ENST00000493795 ENST00000352993 ENST00000468300	Search BioMart View on karyotype
GO:0006260	DNA replication	TAS	Reactome	ENST00000468300 ENST00000352993 ENST00000461221 ENST00000491747 ENST00000493795 ENST00000461798 ENST00000357654 ENST00000471181	Search BioMart View on karyotype
GO:0006281	DNA repair	IEA	InterPro	ENST00000412061	Search BioMart

Detección de Ortólogos Usando PLAZA

Available PLAZA instances

News



Latest news items

- [Preliminary release PLAZA Monocots 4.5](#) 2019-07-10
- [Best instance selection](#) 2019-01-07
- [Data Warehouse API](#) 2018-10-13
- [PLAZA 4.0 release notes](#) 2017-09-01

[View all news items](#)

Monocots PLAZA 4.5 (2019)



Summary

This intermediate update of the PLAZA platform, with a focus on the Monocots, became necessary due to substantial annotation updates to species present in the 4.0 PLAZA instances. For example, the latest wheat release (IWGSCv1.1) is a big step up from the version in Monocots 4.0 (TGACv1.0). Furthermore, the genomes for some other important crop species have also become available in the meantime, necessitating an update to the PLAZA platform.

Under construction! This version can be accessed, but may still contain errors or unreachable pages!

[Go to Monocots PLAZA 4.5](#)

[Locate most appropriate PLAZA instance](#)

Dicots PLAZA 4.0 (2017)



Summary

Version 4.0 of the PLAZA platform. Focused on dicot species, but with the necessary outgroup organisms as reference species.

[PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics.](#)

New release!

[Go to Dicots PLAZA 4.0](#)

PLAZA es una plataforma para la genómica comparativa entre especies de plantas.

Genómica

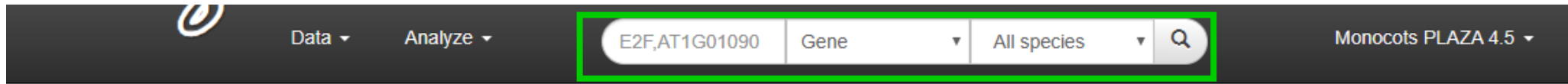


El futuro
es de todos

Gobierno
de Colombia



Detección de Ortólogos Usando PLAZA



Introduction

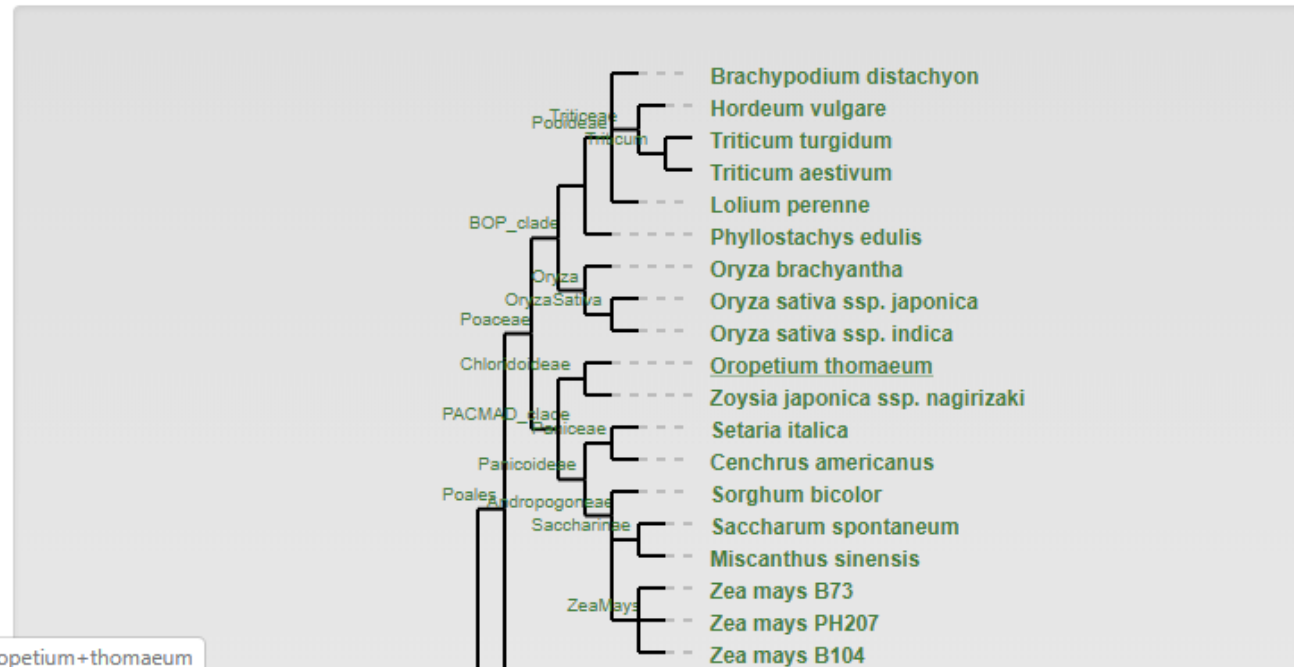
PLAZA is an access point for plant comparative genomics centralizing genomic data produced by different genome sequencing initiatives. It integrates plant sequence data and comparative genomics methods and provides an online platform to perform evolutionary analyses and data mining within the green plant lineage (*Viridiplantae*).

Please use the search functions in the menu bar to search for your favorite gene(s) or use the different Analyze tools to explore gene families and genomic homology. For help and details, please check the [Documentation](#) and/or [Tutorials](#).

News

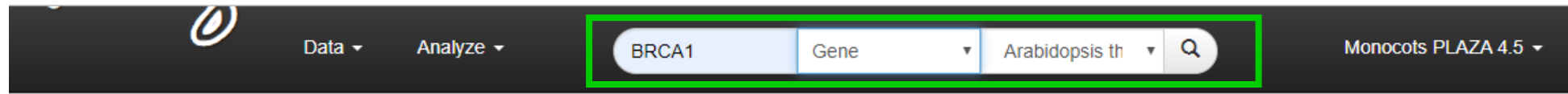
Monocots PLAZA 4.5 summary

- Integration of structural and functional annotation of 39 species.
- Includes 1,563,555 genes, of which 95.8% are protein coding. These protein coding genes are clustered in 38,489 multi-gene gene families (53.7% multi-species gene families).
- More information on the data content can be found in the [data overview](#)



ics.psb.ugent.be/plaza/versions/plaza_v4_5_monocots/organism/view/Oropetium+thomaeum

DetECCIÓN de Ortólogos Usando PLAZA



Introduction

PLAZA is an access point for plant comparative genomics centralizing genomic data produced by different genome sequencing initiatives. It integrates plant sequence data and comparative genomics methods and provides an online platform to perform evolutionary analyses and data mining within the green plant lineage (*Viridiplantae*).

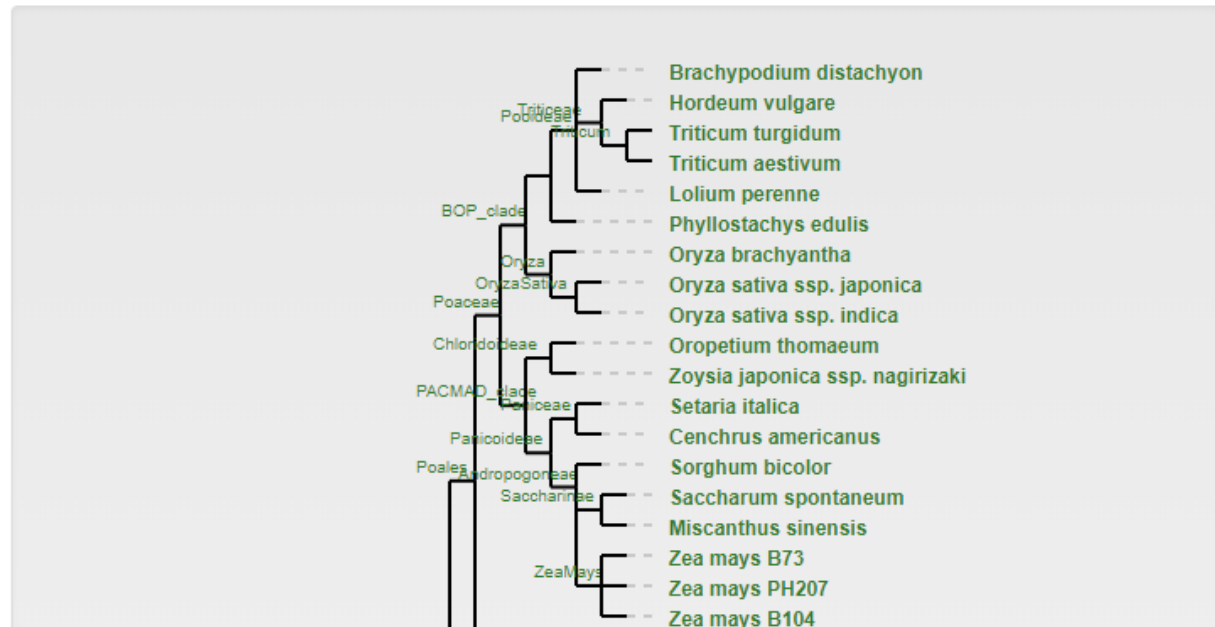
Please use the search functions in the menu bar to search for your favorite gene(s) or use the different Analyze tools to explore gene families and genomic homology. For help and details, please check the [Documentation](#) and/or [Tutorials](#).

News

- [Preliminary release PLAZA Monocots 4.5](#) 2019-07-10

Monocots PLAZA 4.5 summary

- Integration of structural and functional annotation of 39 species.
- Includes 1,563,555 genes, of which 95.8% are protein coding. These protein coding genes are clustered in 38,489 multi-gene gene families (53.7% multi-species gene families).
- More information on the data content can be found in the [data overview](#)



Detección de Ortólogos Usando PLAZA



Data ▾

Analyze ▾

E2F,AT1G01090

Gene ▾

All species ▾



Monocots PLAZA 4.5 ▾

Result page

The query returned 15 entries in 1 species for the search keyword *BRCA1*

[Arabidopsis thaliana \(ath\)](#)

15 entries

AT4G21070 ATBRCA1,ARABIDOPSIS THALIANA BREAST CANCER SUSCEPTIBILITY1 [HOM04x5M001607](#) [ORTHO04x5M001519](#)

breast cancer susceptibility1 (BRCA1); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: DNA repair, response to gamma radiation; LOCATED IN: intracellular, nucleus; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type, conserved site (InterPro:IPR017907), BRCA1 (InterPro:IPR011364), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, C3HC4 RING-type (InterPro:IPR018957), BRCT (InterPro:IPR001357); BEST Arabidopsis thaliana protein match is: breast cancer associated RING 1 (TAIR:AT1G04020.2); Has 8999 Blast hits to 8407 proteins in 1134 species: Archae - 0; Bacteria - 191; Metazoa - 6557; Fungi - 520; Plants - 721; Viruses - 32; Other Eukaryotes - 978 (source: NCBI BLink).

Encodes AtBRCA1, an ortholog of the human breast cancer susceptibility gene 1. Contains one N-terminal RING finger, two C-terminal BRCT and the p300/CBP interacting domain. Strongly induced by gamma rays, consistent with a putative role in DNA repair and in cell cycle control.

BRCA1

AT1G80210 BRCA1/BRCA2-containing complex 36 homolog A [HOM04x5M004977](#) [ORTHO04x5M005998](#)

Genómica



El futuro
es de todos

Gobierno
de Colombia



Detección de Ortólogos Usando PLAZA



Data ▾

Analyze ▾

E2F,AT1G01090

Gene ▾

All species ▾



Monocots PLAZA 4.5 ▾

Gene: **AT4G21070** (*Arabidopsis thaliana*)

Overview

Gene Identifier AT4G21070
Transcript Identifier AT4G21070.1
Gene Type Coding gene
Location Chr4 : 11248174-11252633 :
positive

Family Information

Homologous gene family HOM04x5M001607
(146 genes in 38 species)

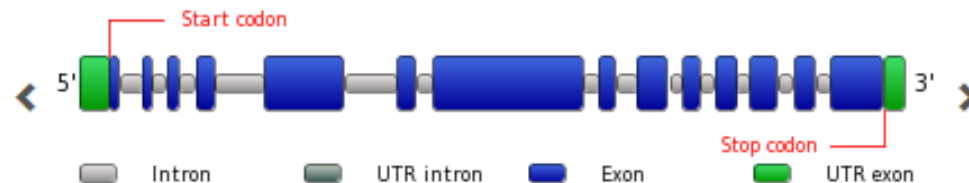
Viridiplantae specific family

Orthologous gene family ORTHO04x5M001519
(146 genes in 38 species)

Viridiplantae specific family

Identifiers

Type	Value
Alias	ATBRCA1,ARABIDOPSIS THALIANA BREA SUSCEPTIBILITY1
Dbxref	PMID:15644464,PMID:16517762,gene:3439
symbol	BRCA1
tid	AT4G21070.1



Genómica



El futuro
es de todos

Gobierno
de Colombia



Detección de Ortólogos Usando PLAZA

Orthologous gene family: **ORTHO04x5M001519** ⓘ

Sequences

[Download DNA sequences](#)
[Download protein sequences](#)

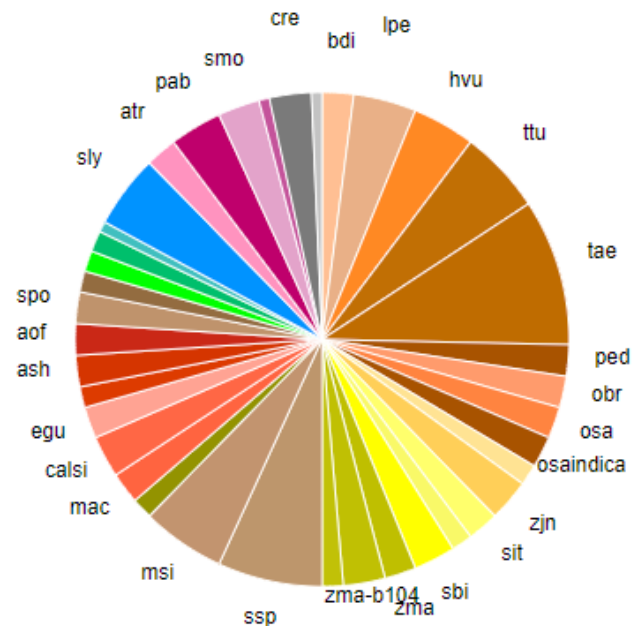
Parent family

HOM04x5M001607
(146 genes in 38 species)

Keywords

Gene	Biological	DNA repair
Ontology	Process	
InterPro	BRCA1-associated ; Extended PHD (ePHD) domain ; BRCT domain	

Overview of gene family content



The gene family *ORTHO04x5M001519* has 146 genes in 38 species (7 outlier genes).
Smallest encompassing phylogenetic clade for this gene family : *Viridiplantae*
Tandem/block/both gene duplicate percentage within this gene family : 9%, 23%, 3%

The order of species in the graph is based on the phylogenetic species tree.

Toolbox

Explore

- ...the local gene organization for homologous genes.
- ...the phylogenetic trees of this gene family.
- ...the expansion/depletion of species in this gene family.

Create

- ...a custom phylogenetic tree using this gene family as seed.

View

- ...the multiple sequence alignment of this gene family with BioJS
- ...the similarity heatmap of this gene family.
- ...the genome wide organization of this gene family.

Detección de Ortólogos Usando PLAZA

Gene list

Index of genes with 1 constraints:

- gene family : [ORTHO04x5M001519](#)

Page 1 of 8, showing 20 records out of 146 total, starting on record 1, ending on 20

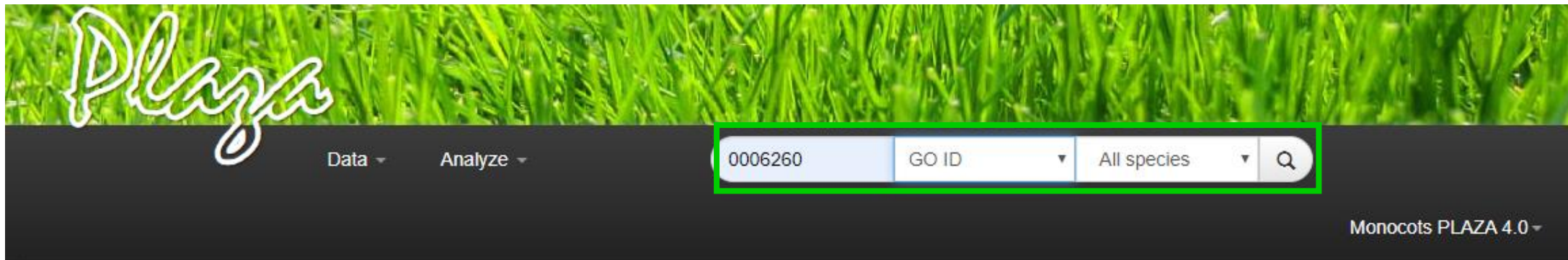
gene_id	species	strand	chr
Aco008966	aco	+	LG09
Aco012603	aco	-	LG13
ASH_gene15189	ash	-	KZ452009.1
ASH_gene20855	ash	-	KZ454132.1
ASH_gene4657	ash	-	KZ451916.1
AT1G04020	ath	+	Chr1
AT4G21070	ath	+	Chr4
ATR0736G050	atr	+	AmTr_v1.0_scaffold00067
ATR0736G145	atr	+	AmTr_v1.0_scaffold00067
ATR1098G120	atr	+	AmTr_v1.0_scaffold00092
Bradi2g20260	bdi	-	Bd2
Bradi2g21960	bdi	+	Bd2
Bradi5g15500	bdi	-	Bd5
CALS1_Maker00007680	calsi	-	scaffold1361

Download genes (select columns to include)

- ☒ gene_id
- ☒ species
- ☒ strand
- ☒ chr
- ☒ outlier

Download

Detección de Ortólogos Usando PLAZA



Introduction

PLAZA is an access point for plant comparative genomics centralizing genomic data produced by different genome sequencing initiatives. It integrates plant sequence data and comparative genomics methods and provides an online platform to perform evolutionary analyses and data mining within the green plant lineage (*Viridiplantae*).

Please use the search functions in the menu bar to search for your favorite gene(s) or use the different Analyze tools to explore gene families and genomic homology. For help and details, please check the [Documentation](#) and/or [Tutorials](#).

News

- [Move from http to https](#) September 2017

PLAZA 4.0 Monocots summary

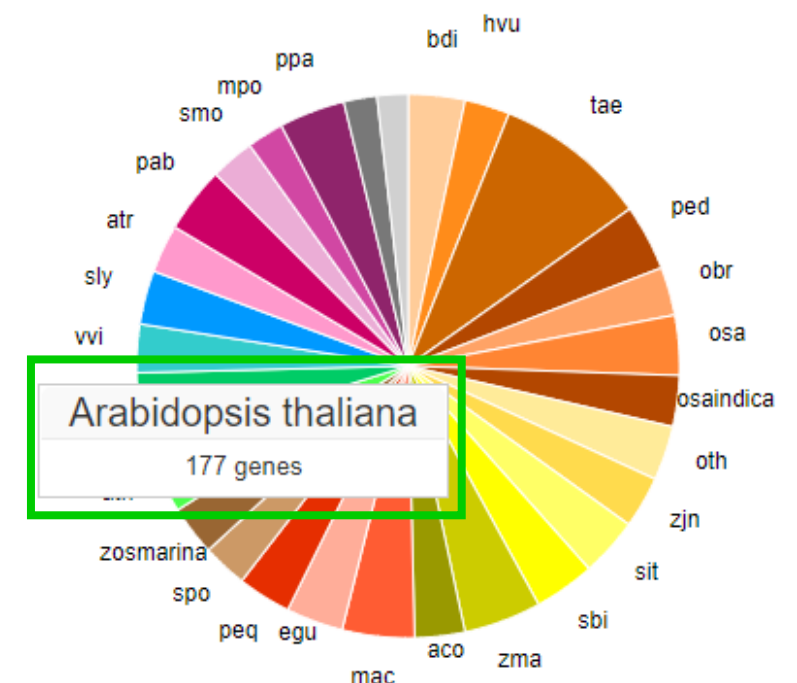
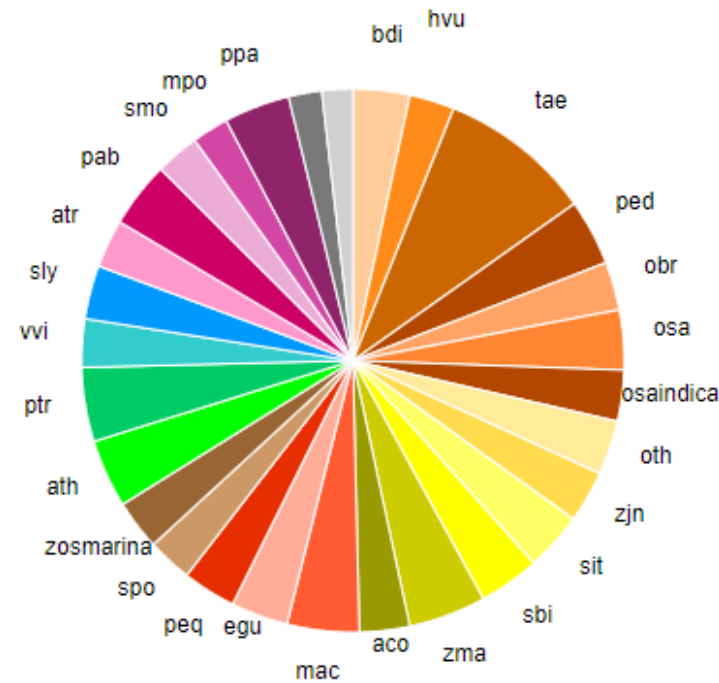
- Integration of structural and functional annotation of 29 species.



Detección de Ortólogos Usando PLAZA

Description	DNA replication	GO depth	5
Type	Biological Process	Number of genes	4355 (all) / 2321 (primary)

GO distribution



Genómica



El futuro
es de todos

Gobierno
de Colombia



Detección de Ortólogos Usando PLAZA

Index of genes with 2 constraints:

- GO term : [GO-0006260](#)
- species : [Arabidopsis thaliana](#)

GO term parent-child relationships are used in this table.

[View genome wide organization](#)

Page 1 of 9, showing 20 records out of 180 total, starting on record 1, ending on 20

gene_id	species	go	evidence	go_source	provider
AT5G27740	ath	GO:0006260	IEA	primary	interproscan
AT4G19130	ath	GO:0006260	IEA	primary	interproscan
AT5G22010	ath	GO:0006260	IEA	primary	interproscan
AT4G24790	ath	GO:0006260	IEA	primary	interproscan
AT4G29910	ath	GO:0006260	IEA	primary	interproscan
AT3G52630	ath	GO:0006260	IEA	primary	interproscan
AT5G16690	ath	GO:0006260	IEA	primary	interproscan
AT1G21690	ath	GO:0006260	IEA	primary	interproscan
AT2G07690	ath	GO:0006260	IEA	primary	interproscan
AT1G50840	ath	GO:0006260	IEA	primary	interproscan
AT1G14460	ath	GO:0006260	IEA	primary	interproscan
AT4G18820	ath	GO:0006260	IEA	primary	interproscan

Download genes (select columns to include)

- ☒ gene_id
- ☒ species
- ☒ evidence
- ☒ go_source

Download

Genómica



El futuro
es de todos

Gobierno
de Colombia



¡Gracias!



¡Gracias!