

BANCOS DE DADOS BIOLÓGICO

Aula 9

LGN0232 – Genética Molecular



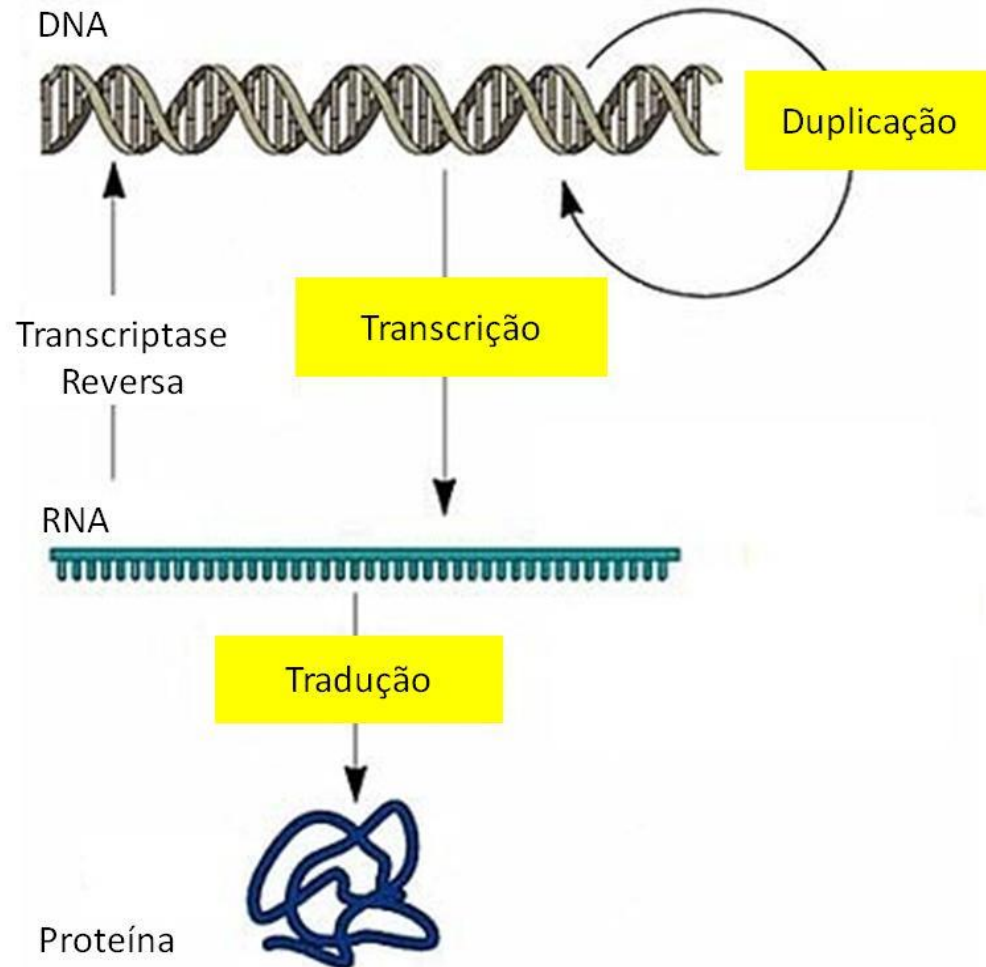
Departamento de Genética

SUMÁRIO:

- 1. DEFINIÇÃO DE BANCO DE DADOS BIOLÓGICOS;**
- 2. BIOINFORMÁTICA**
- 3. RECURSOS OFERECIDOS PELO NCBI**
- 4. UTILIZAÇÃO DO PROGRAMA BLAST**

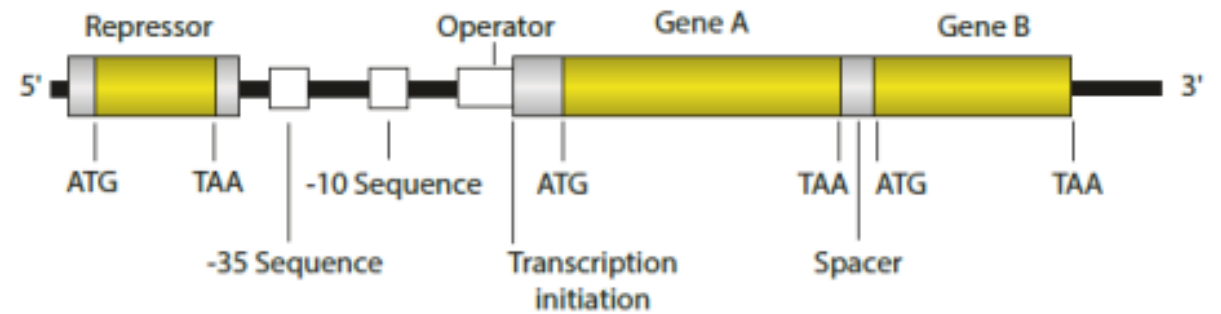
RELEMBRANDO:

✓ O FLUXO DA INFORMAÇÃO GENÉTICA

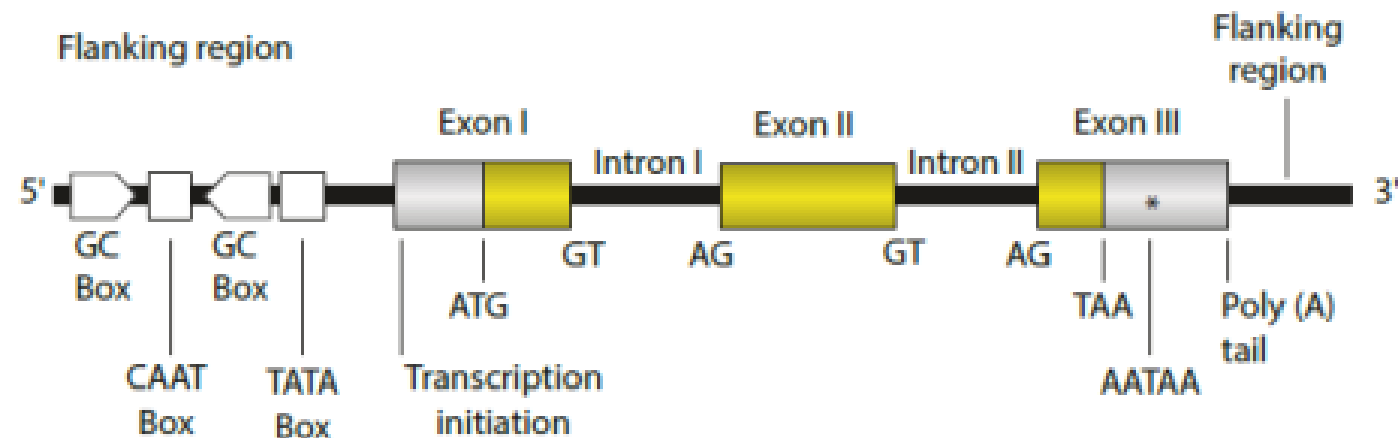


✓ ORGANIZAÇÃO DOS GENES

PROCARIOTOS:



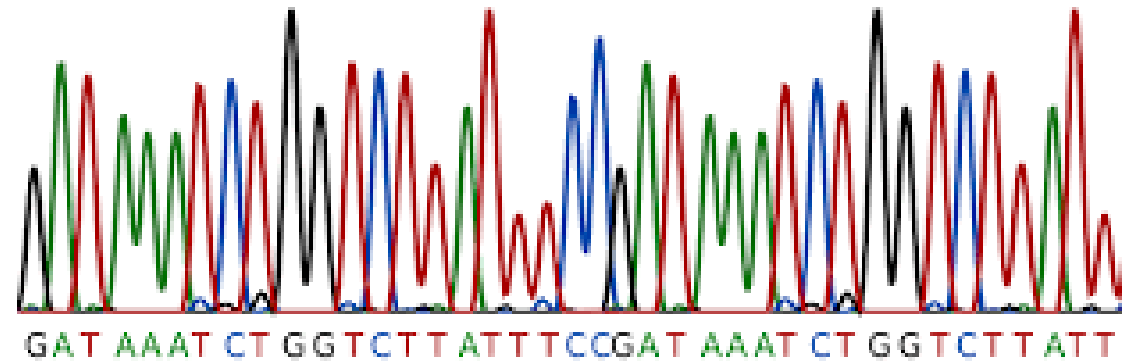
EUCARIOTOS:



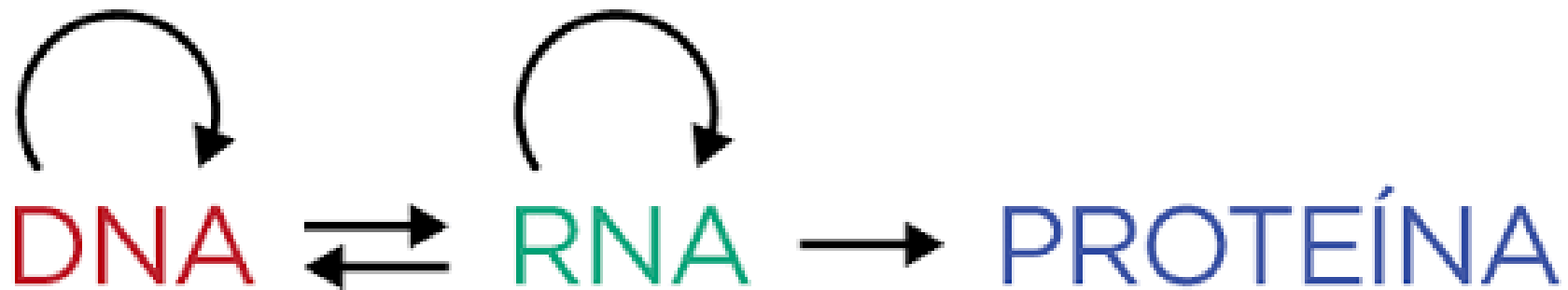
PROJETOS DE SEQUENCIAMENTO



SEQUENCIAMENTO DE DNA



PREDIÇÃO



PROJETOS DE SEQUENCIAMENTO

AUMENTO DO NÚMERO DE PROJETOS DE SEQUENCIAMENTO
(Novas tecnologias e redução de custos)

+

**COMPARTILHAMENTO DAS INFORMAÇÕES OBTIDAS COM A
COMUNIDADE CIENTÍFICA**



BANCO DE DADOS BIOLÓGICOS



BANCO DE DADOS BIOLÓGICOS (BDB) E BIOINFORMÁTICA

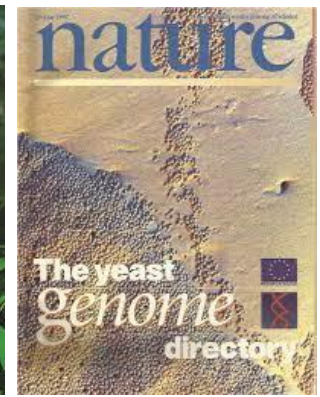
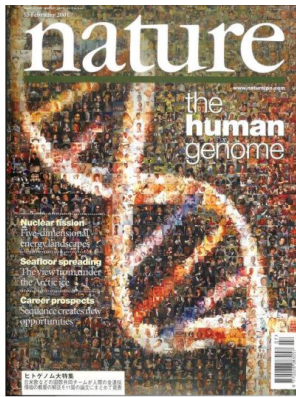
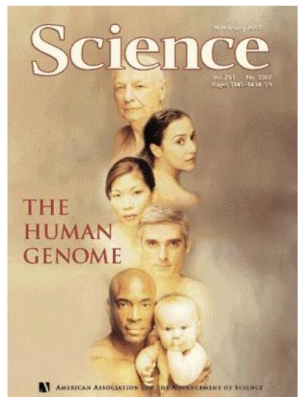
➤ O QUE SÃO?

São **repositórios** online que centralizam as informações genéticas a nível de DNA, RNA ou proteína, dentre outros.

➤ **OBJETIVOS DOS BDBS:** Centralizar os dados e permitir o acesso a informações geradas nas pesquisas.



Permite, por exemplo, comparar genes/genomas de espécies distintas.

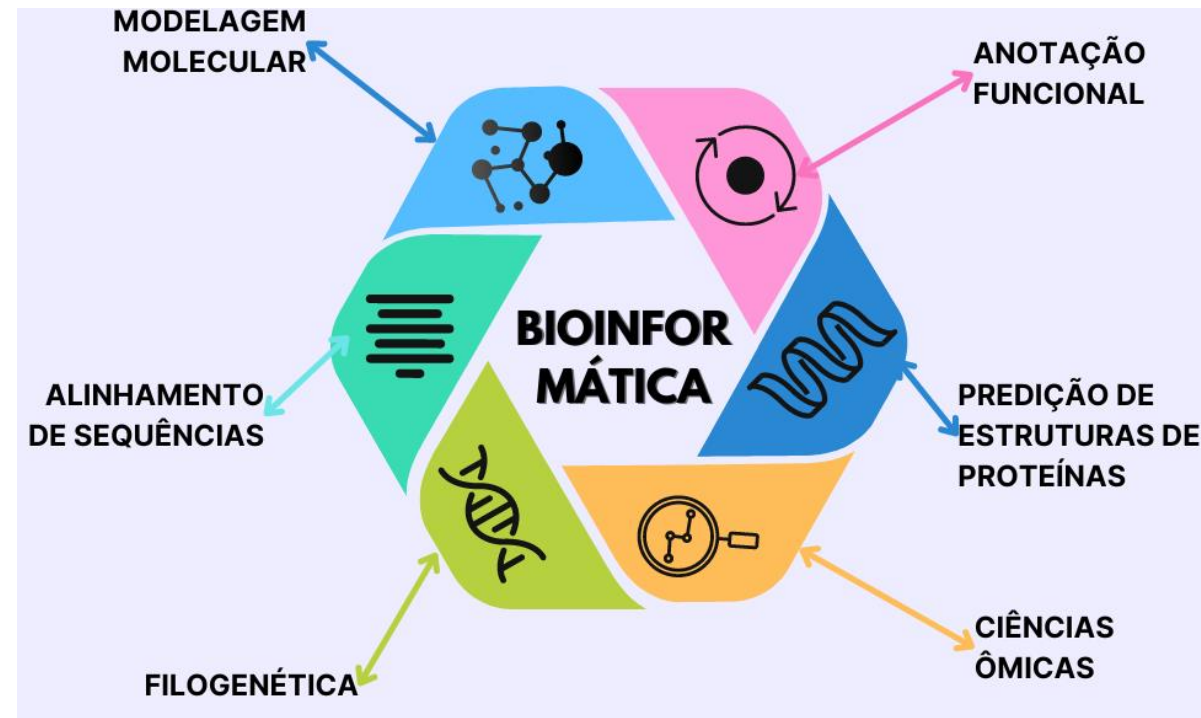
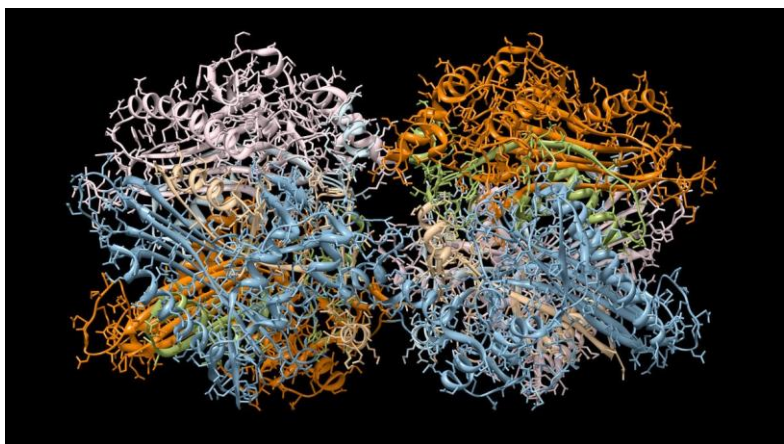


BANCO DE DADOS BIOLÓGICOS (BDB) E BIOINFORMÁTICA

➤ BIOINFORMÁTICA

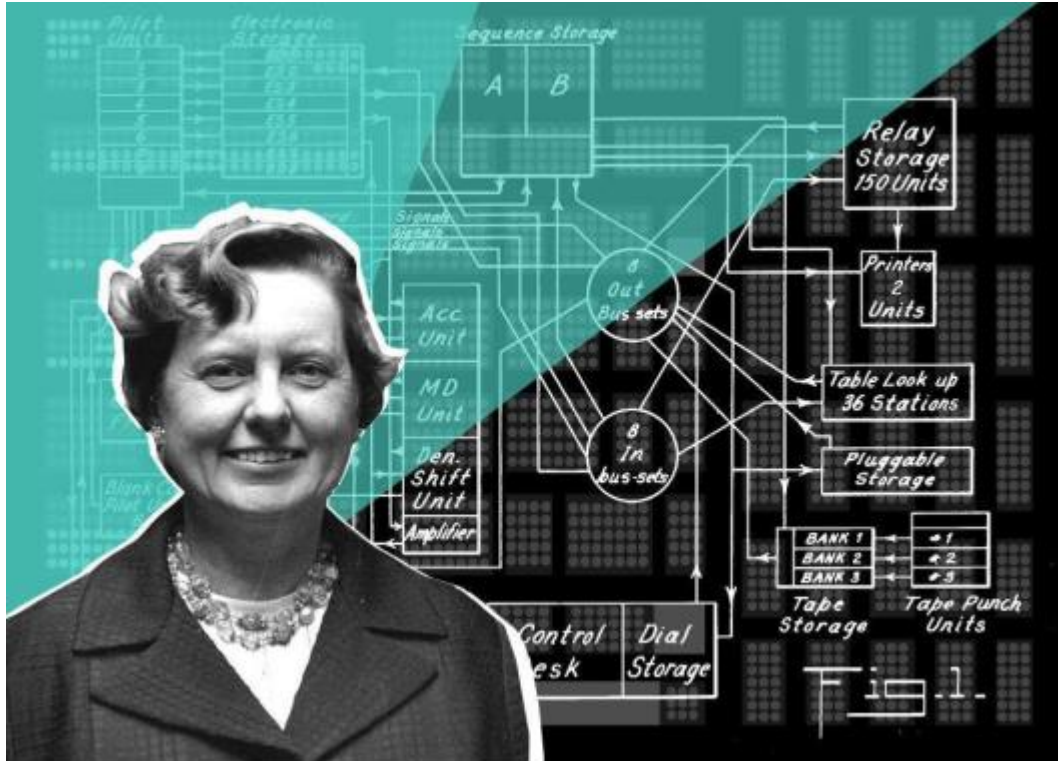
Ciência multidisciplinar que usa tecnologia e ferramentas da computação na genética para organizar, armazenar, interpretar, visualizar e disseminar dados biológicos.

É utilizada, por exemplo, para estudar a determinação de funções de genes e proteínas, estabelecer relações evolutivas e prever formas tridimensionais das proteínas.



BANCO DE DADOS BIOLÓGICOS (BDB) E BIOINFORMÁTICA

Dra. Margaret Belle (Oakley) Dayhoff

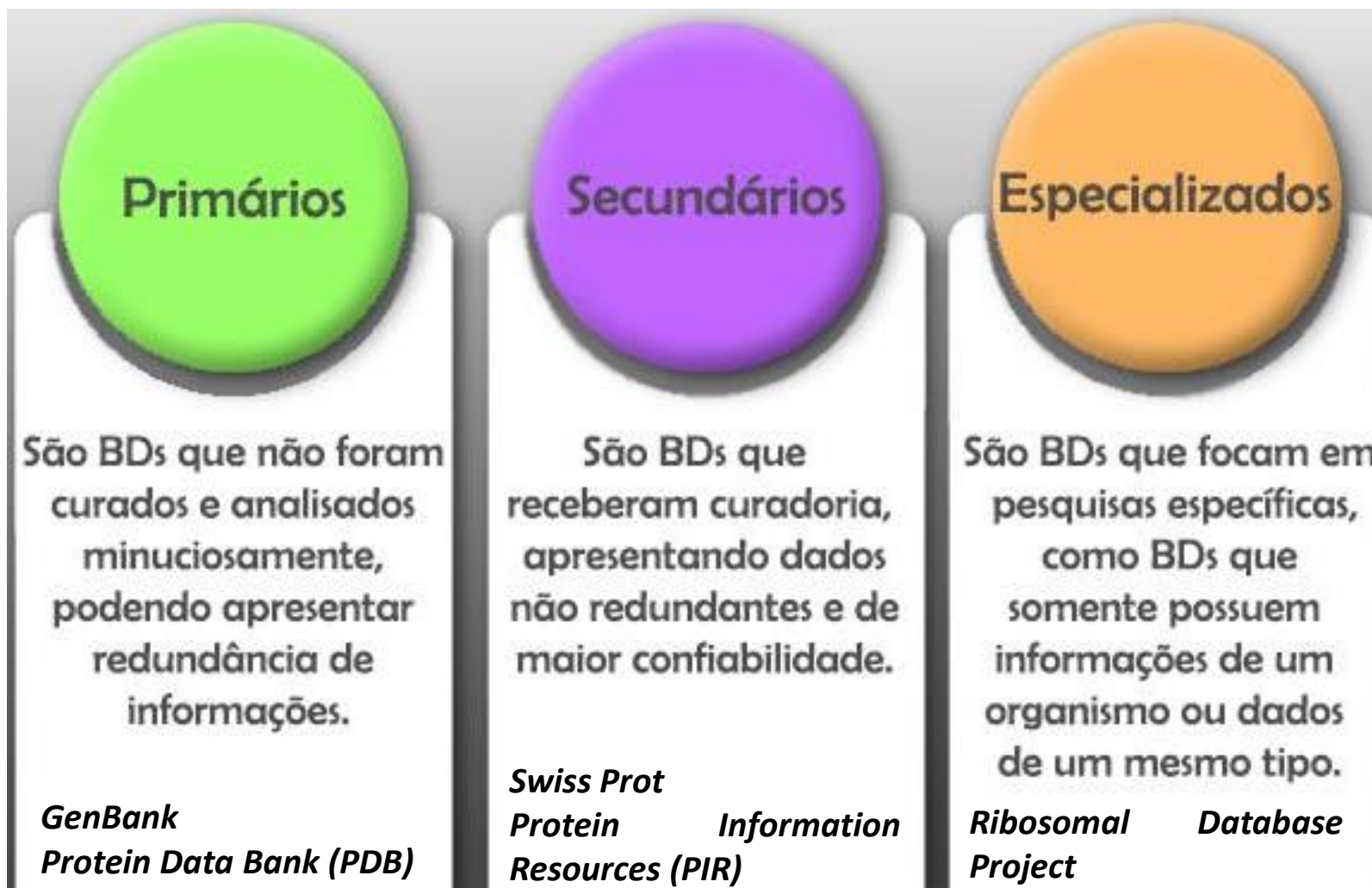


Pioneira na aplicação de métodos matemáticos e computacionais ao campo da bioquímica. Dedicou sua carreira à aplicação das tecnologias computacionais para promover avanços no campo da biologia e da medicina, em especial à criação de bases de dados de ácidos nucleicos e proteínas, além de ferramentas para utilizá-las.

A mãe da Bioinformática

BANCO DE DADOS BIOLÓGICOS (BDB) E BIOINFORMÁTICA

➤ TIPOS DE BANCOS DE DADOS BIOLÓGICOS



➤ O PRINCIPAL BANCO DE DADOS BIOLÓGICOS:

NCBI (*National Center for Biotechnology Information*) – fundado em 1988;

O website <https://www.ncbi.nlm.nih.gov/> foi criado em 1994.

- É UMA PLATAFORMA *ONLINE*, QUE REUNE DIVERSOS BANCOS DE DADOS E PERMITE QUE QUALQUER USUÁRIO ACESSE TAIS INFORMAÇÕES;
- UMA COLEÇÃO DE REGISTROS INTEGRADOS DE SEQUENCIAS (DNA, RNA, PROTEÍNAS);
- É UM REPOSITÓRIO PARA O DEPÓSITO DE SEQUENCIAS E GENOMAS.

➤ O PRINCIPAL BANCO DE DADOS BIOLÓGICOS:

NCBI (*National Center for Biotechnology Information*) – fundado em 1988;

O website <https://www.ncbi.nlm.nih.gov/> foi criado em 1994.

É de uso gratuito e acolhe diversos BDBs, separados por categorias:

Literature — Repositório de artigos científicos, livros, entre outros. Um dos bancos mais utilizados dessa categoria é o **PubMed Central**.

Genes — São encontradas sequências gênicas e **anotações** para estudo de estrutura de ortólogos, expressão e evolução.

Genomes — Possui bancos de sequências genômicas, dados de genômica funcional e origem de amostras biológicas. Um dos principais bancos dessa categoria é o **Nucleotide**, que tem o **GenBank** como um dos seus principais componentes.

Proteins — Apresenta dados como sequências proteicas, estruturas tridimensionais (3D) e domínios proteicos.

BLAST — É uma ferramenta que realiza consultas em diferentes bancos de dados, como **Nucleotide** e **Protein**.

PubChem — Repositório de informações químicas, rotas metabólicas e ferramentas para screening de atividade biológica.

BANCO DE DADOS BIOLÓGICOS (BDB)

➤ Formatos de arquivos de dados comuns nos BDBs e em bioinformática:

Formato FASTA

>AB537496.1 *Eucalyptus grandis* chloroplast rbcL gene for ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, complete cds

```
ATGTCACCACAAACAGAGACTAAAGCAAGTGTTGGATTCAAAGCTGGTGTTAAAGATTATAAACTGACTTATTATACTCCT
GACTATGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCTGGAGTTCCTCCTGAGGAAGCAGGG
GCTGCGGTAGCTGCTGAATCTTCTACTGGTACATGGACAACCTGTGTGGACCGATGGGCTTACCAGCCTTGATCGTTATAAA
GGAAGATGCTACCACATCGAGCCTGTTGCTGGAGAAGAAAATCAATATATATGTTATGTAGCTTACCCTTTAGACCTTTTTT
GAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGGTTCAAAGCCCTGCGCGCTCTACGTCTG
GAGGATCTGCGAATCCCTCCTTCCTATACGAAAACCTTCCAAGGCCCGCCTCATGGCATCCAAGTTGAGAGAGATAAATTA
ACAAATATGGGCGTCCCCTATTGGGATGTACTATTAAACCGAAATTGGGGTTATCCGCTAAGAACTACGGTAGAGCAGTTT
ATGAATGTCTTCGTGGTGGACTTGATTTTACGAAAGATGATGAGAACGTGAACTCACAACCATTTATGCGTTGGAGAGACC
GTTTCTTATTTTGTGCCGAAGCCATTTTTTAAATCACAGGCTGAAACAGGTGAAATCAAAGGGCATTACTTGAATGCTACTG
CAGGTACATGCGAAGAAATGATCAAAAGGGCTGTATTTGCCAGAGAATTGGGAGTTCC
```

Esse “tipo” de informação é gerada após o sequenciamento e processamento dos dados !!!

**ACOMPANHE ACESSANDO O SITE DO NCBI,
PELO CELULAR OU PC**

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

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

NCBI – National Center for Biotechnology Information

**National Library of Medicine**
National Center for Biotechnology Information

Log in

All Databases

Search

[NCBI Home](#)
[Resource List \(A-Z\)](#)
[All Resources](#)
[Chemicals & Bioassays](#)
[Data & Software](#)
[DNA & RNA](#)
[Domains & Structures](#)
[Genes & Expression](#)
[Genetics & Medicine](#)
[Genomes & Maps](#)
[Homology](#)
[Literature](#)
[Proteins](#)
[Sequence Analysis](#)
[Taxonomy](#)
[Training & Tutorials](#)
[Variation](#)

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit

Deposit data or manuscripts into NCBI databases



Download

Transfer NCBI data to your computer



Learn

Find help documents, attend a class or watch a tutorial



Develop

Use NCBI APIs and code libraries to build applications



Analyze

Identify an NCBI tool for your data analysis task



Research

Explore NCBI research and collaborative projects



Popular Resources

- [PubMed](#)
- [Bookshelf](#)
- [PubMed Central](#)
- [BLAST](#)
- [Nucleotide](#)
- [Genome](#)
- [SNP](#)
- [Gene](#)
- [Protein](#)
- [PubChem](#)

NCBI News & Blog

Comparing Yeast Species Used in Beer Brewing and Bread Making
29 Sep 2023

Using the NIH Comparative Genomics Resource (CGR) to gain knowledge

Join NCBI at ASHG 2023
27 Sep 2023

November 1-5 in Washington, D.C. We look forward to seeing you in person at

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

NCBI – National Center for Biotechnology Information

The screenshot displays the NCBI website interface. At the top, the NIH logo and "National Library of Medicine" text are visible, along with a "Log in" button. Below this is a search bar with a dropdown menu set to "Genome" and the search term "EUCALYPTUS GRANDIS". A search button is to the right of the search bar. On the left side, there is a vertical navigation menu with links such as "NCBI Home", "Resource List (A-Z)", "All Resources", "Chemicals & Bioassays", "Data & Software", "DNA & RNA", "Domains & Structures", "Genes & Expression", "Genetics & Medicine", "Genomes & Maps", "Homology", "Literature", "Proteins", "Sequence Analysis", "Taxonomy", and "Training & Tutorials". A dropdown menu is open from the "Genome" link, showing a list of database categories including "All Databases", "Assembly", "Biocollections", "BioProject", "BioSample", "Books", "ClinVar", "Conserved Domains", "dbGaP", "dbVar", "Gene", "GEO DataSets", "GEO Profiles", "GTR", "HomoloGene", "Identical Protein Groups", "MedGen", "MeSH", and "NLM Catalog". The "Gene" option is currently selected. The main content area features a large "NCBI" heading and a description of the center's mission. Below this, there are three columns of resources: "Submit" (for manuscripts and databases), "Download" (to transfer data to a computer), and "Learn" (to find help documents, attend a class, or watch a tutorial). At the bottom, there are three more columns: "Develop" (using NCBI APIs and code libraries), "Analyze" (identifying tools for data analysis), and "Research" (exploring research and collaborative projects). On the right side, there is a "Popular Resources" section with links to PubMed, Bookshelf, PubMed Central, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem. Below this is an "NCBI News & Blog" section with a link to "Comparing Yeast Species Used in Beer Brewing and Bread Making" dated 29 Sep 2023.

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

Genome EUCALYPTUS GRANDIS Search

NCBI

Center for Biotechnology Information advances science and health by providing access to genomic information.

Submit Download Learn

Transfer NCBI data to your computer

Find help documents, attend a class or watch a tutorial

Develop Analyze Research

Use NCBI APIs and code libraries to build applications

Identify an NCBI tool for your data analysis task

Explore NCBI research and collaborative projects

Popular Resources

PubMed
Bookshelf
PubMed Central
BLAST
Nucleotide
Genome
SNP
Gene
Protein
PubChem

NCBI News & Blog

Comparing Yeast Species Used in Beer Brewing and Bread Making
29 Sep 2023

Using the NIH Comparative Genomics

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

[NCBI Datasets](#) [Taxonomy](#) [Genome](#) [Gene](#) [Command-line tools](#) [Documentation](#)

[Eukaryota](#) / [Viridiplantae](#) / [Streptophyta](#) / [Magnoliopsida](#) / [Myrtales](#) / [Myrtaceae](#) / [Eucalyptus](#)

Eucalyptus grandis ☆

Rose gum (*Eucalyptus grandis*) is a species of eudicot in the family *Myrtaceae* (myrtle family).

[Browse taxonomy](#)

Current scientific name	<i>Eucalyptus grandis</i>
Common name	rose gum
Taxonomic rank	species
NCBI Taxonomy ID	71139

For more details see [NCBI Taxonomy](#)
View the legacy [Genome](#) page



Genome

[Browse all 3 genomes](#)

Reference genome

[ASM1654582v1](#)

Australian National University (2021). Isolate: ANBG69807.140.
RefSeq [GCF_016545825.1](#)

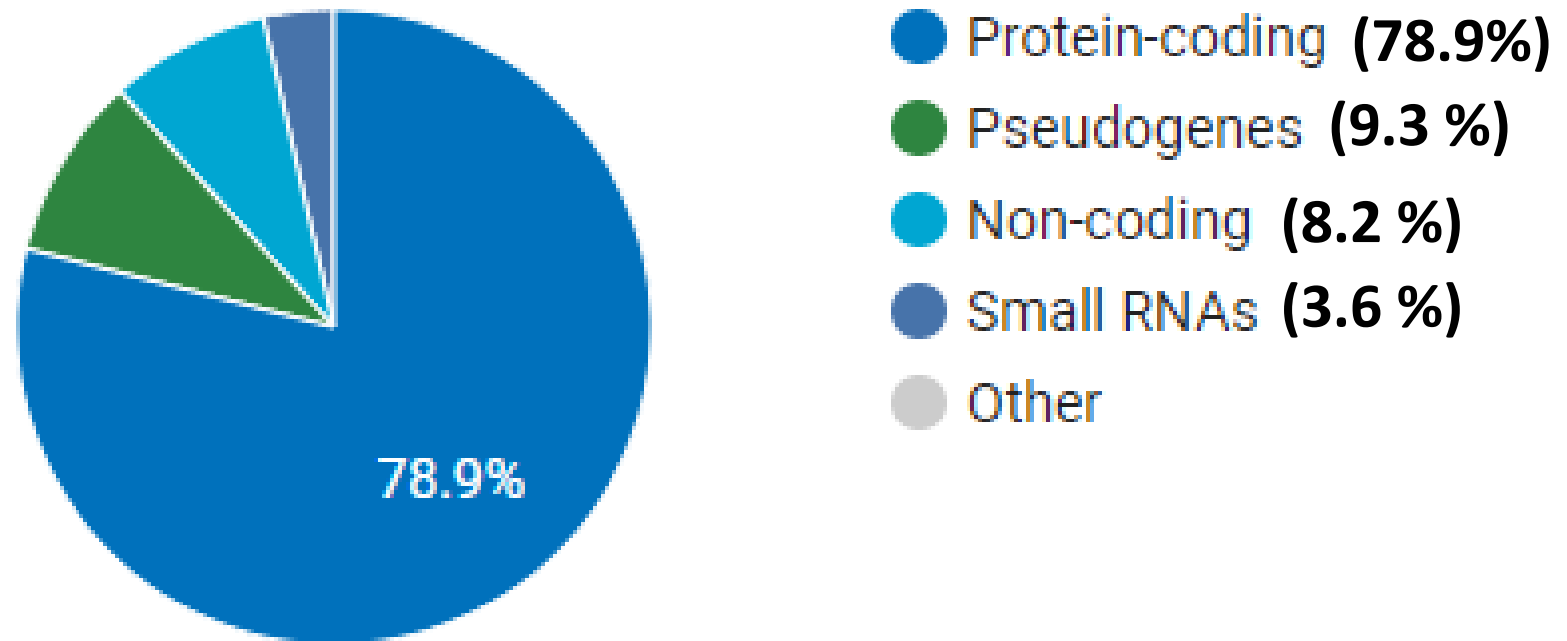
[Download](#)

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ ACESSANDO O GENOMA DE *EUCALYPTUS GRANDIS*

Gene

Current gene set



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ INFORMAÇÕES OBTIDAS COM UM GENOMA SEQUENCIADO

Sample details

BioSample ID	SAMN13016251
Description	Plant sample from Eucalyptus grandis
Owner name	Australian National University
Isolate	ANBG69807.140
Dev stage	tree
Geo loc name	Australia
Tissue	leaf

Assembly methods

Sequencing technology	Oxford Nanopore MinION
Assembly method	Canu v. 1.9

Annotation details

[See full annotation report](#)

	RefSeq
Provider	NCBI RefSeq
Name	NCBI Annotation Release 102
Date	Feb 2, 2021
Genes	42.619
Protein-coding	33.352
Non-coding	5.319
Software version	8.5

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ BUSCANDO POR GENES DE INTERESSE

The screenshot shows the NCBI homepage with a search bar at the top. The 'All Databases' dropdown menu is open, displaying a list of databases. Red arrows point to 'Gene', 'Nucleotide', and 'Protein' in the list. The background shows the main navigation menu on the left, the NCBI logo and mission statement in the center, and a 'Popular Resources' section on the right.

All Databases ▼

- Conserved Domains
- dbGaP
- dbVar
- Gene
- Genome
- GEO DataSets
- GEO Profiles
- GTR
- HomoloGene
- Identical Protein Groups
- MedGen
- MeSH
- NLM Catalog
- Nucleotide
- OMIM
- PMC
- PopSet
- Protein
- Protein Clusters
- Protein Family Models

NCBI

Center for Biotechnology Information advances science and health by providing access to genomic information.

[Home](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit
Submit your manuscripts and data to public databases

Download
Transfer NCBI data to your computer

Learn
Find help documents, attend a class or watch a tutorial

Develop
Use NCBI APIs and code libraries to build applications

Analyze
Identify an NCBI tool for your data analysis task

Research
Explore NCBI research and collaborative projects

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

Comparing Yeast Species Used in Beer Brewing and Bread Making
29 Sep 2023

Using the NIH Comparative Genomics Resource (CGR) to gain knowledge

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ BUSCA POR GENES DE INTERESSE – EX: RUBISCO EUCALYPTUS

Gene

Gene

rubisco eucalyptus

×

Search

Create RSS

Save search

Advanced

Help

Gene sources

Genomic

Organelles

Plastids

Categories

Alternatively spliced

Annotated genes

Protein-coding

Sequence content

RefSeq

Status

✓ Current

[Clear all](#)

[Show additional filters](#)

clear

Tabular ▾ 20 per page ▾ Sort by Relevance ▾

Send to: ▾

Hide sidebar >>

See [RBCS \(RUBISCO\) ribulose biphosphate carboxylase small chain, chloroplastic-like](#) in the Gene database

Search results

Items: 1 to 20 of 37

<< First < Prev Page 1 of 2 Next > Last >>

Showing Current items.

Name/Gene ID	Description	Location	Aliases
☐ LOC104415378 ID: 104415378	rubisco accumulation factor 1.1, chloroplastic [<i>Eucalyptus grandis</i>]	Chromosome 8, NC_052619.1 (59733730..59735591, complement)	EUGRSUZ_H042531
☐ LOC104432900 ID: 104432900	ruBisCO large subunit-binding protein subunit alpha [<i>Eucalyptus grandis</i>]	Chromosome 2, NC_052613.1 (41847086..41850549)	EUGRSUZ_B02532
☐ LOC104422107 ID: 104422107	ribulose biphosphate carboxylase small chain, chloroplastic [<i>Eucalyptus grandis</i>]	Chromosome 10, NC_052621.1 (18302801..18304215)	EUGRSUZ_J01502
☐ LOC104440598 ID: 104440598	protein PLASTID TRANSCRIPTIONALLY ACTIVE 14 [<i>Eucalyptus grandis</i>]	Chromosome 4, NC_052615.1 (210992..216873)	EUGRSUZ_D00007
☐ LOC104437383 ID: 104437383	N-lysine methyltransferase SETD6 [<i>Eucalyptus grandis</i>]	Chromosome 3, NC_052614.1 (39529715..39539186, complement)	EUGRSUZ_C02305

Filters: [Manage Filters](#)

Results by taxon

Top Organisms [\[Tree\]](#)

Eucalyptus grandis (25)

Eucalyptus globulus subsp. *globulus* (1)

Eucalyptus sieberi (1)

Eucalyptus elata (1)

Eucalyptus regnans (1)

All other taxa (8)

[More...](#)

Find related data

Database: [Select](#) ▾

[Find items](#)

Search details

(rubisco[All Fields] AND ("Eucalyptus" [Organism] OR eucalyptus[All Fields])) AND alive[prop]

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

BUSCA POR GENES DE INTERESSE

LOC104432900 ruBisCO large subunit-binding protein subunit alpha [*Eucalyptus grandis*]

[Download Datasets](#)

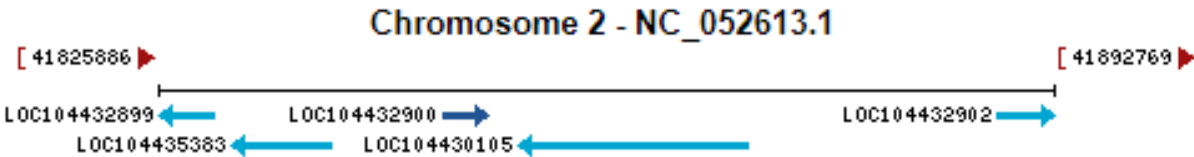
Gene ID: 104432900, updated on 18-Aug-2023

Summary

Genomic context

Location: chromosome: 2 See LOC104432900 in [Genome Data Viewer](#)
Exon count: 8

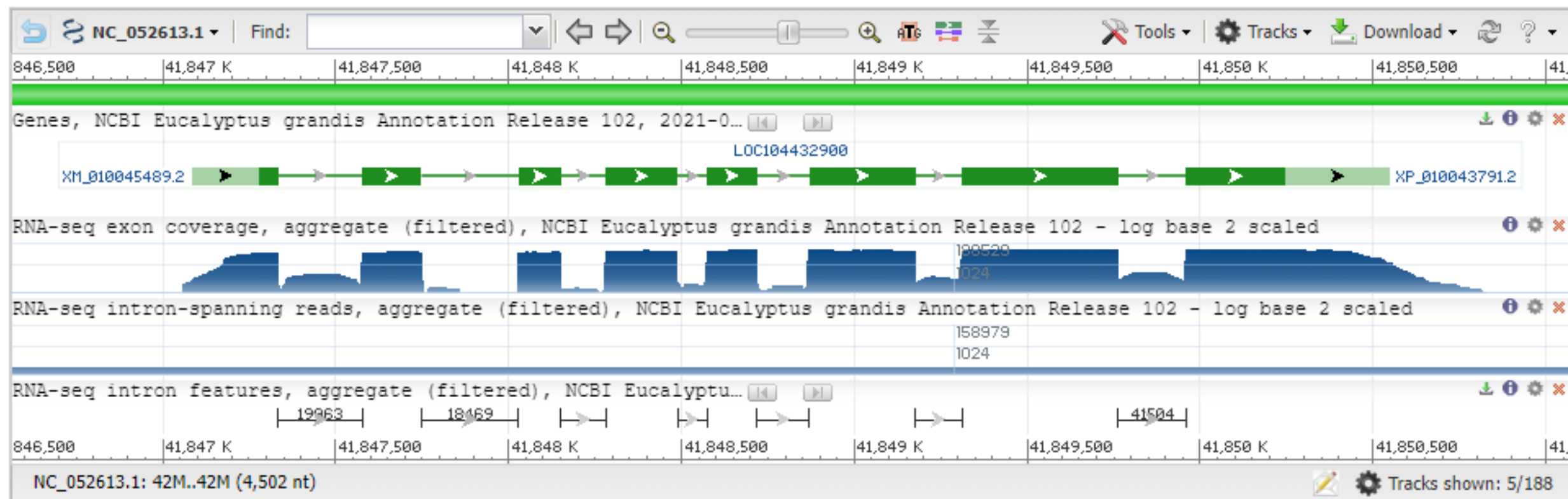
Annotation release	Status	Assembly	Chr	Location
102	current	ASM1654582v1 (GCF_016545825.1)	2	NC_052613.1 (41847086..41850549)
101	previous assembly	Egrandis1_0 (GCF_000612305.1)	Unplaced Scaffold	NW_010092439.1 (47408441..47411957)



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

Genomic Sequence:

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

Reference ASM1654582v1 Primary Assembly

Genomic

1. NC_052613.1 Reference ASM1654582v1 Primary Assembly

Range	41847086..41850549
Download	GenBank , FASTA , Sequence Viewer (Graphics)

1 ↑ 2 ↑

mRNA and Protein(s)

1. [XM_010045489.2](#) → [XP_010043791.2](#) ruBisCO large subunit-binding protein subunit alpha

UniProtKB/TrEMBL	A0A059D5S8		
Conserved Domains (1) summary			
	<table><tr><td>cd03344 Location:45 → 565</td><td>GroEL; GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They share a common general morphology, a double toroid of 2 stacked rings, each composed of 7-9 subunits. The symmetry of type I is seven-fold and they are found ...</td></tr></table>	cd03344 Location:45 → 565	GroEL; GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They share a common general morphology, a double toroid of 2 stacked rings, each composed of 7-9 subunits. The symmetry of type I is seven-fold and they are found ...
cd03344 Location:45 → 565	GroEL; GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They share a common general morphology, a double toroid of 2 stacked rings, each composed of 7-9 subunits. The symmetry of type I is seven-fold and they are found ...		

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ ACESSANDO O GENE DE INTERESSE

GenBank ▾

Eucalyptus grandis isolate ANBG69807.140 chromosome 8 whole genome shotgun sequence

NCBI Reference Sequence: NC_052619.1

[FASTA](#) [Graphics](#)

Go to: ☐

LOCUS NC_052619 1862 bp DNA linear CON 04-FEB-2021

DEFINITION Eucalyptus grandis isolate ANBG69807.140 chromosome 8, ASM1654582v1, whole genome shotgun sequence.

ACCESSION [NC_052619](#) REGION: complement(59733730..59735591)

VERSION NC_052619.1

DBLINK BioProject: [PRJNA698663](#)
BioSample: [SAMN13016251](#)
Assembly: [GCF_016545825.1](#)

KEYWORDS WGS; RefSeq.

SOURCE Eucalyptus grandis

ORGANISM [Eucalyptus grandis](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
Pentapetalae; rosids; malvids; Myrtales; Myrtaceae; Myrtoideae;
Eucalypteae; Eucalyptus.

COMMENT [REFSEQ INFORMATION](#): The reference sequence is identical to [CM028315.1](#).
Assembly name: ASM1654582v1
The genomic sequence for this RefSeq record is from the whole-genome assembly released by the Australian National University on 2021/01/06. The original whole-genome shotgun project has the accession JABKBJ000000000.1.

FEATURES
source

Location/Qualifiers
1..1862
/organism="Eucalyptus grandis"
/mol_type="genomic DNA"
/isolate="ANBG69807.140"
/db_xref="taxon:[71139](#)"
/chromosome="8"
/tissue_type="leaf"
/dev_stage="tree"
/country="Australia"

[gene](#)

1..1862
/gene="LOC104415378"
/note="Derived by automated computational analysis using gene prediction method: Gnomon."
/db_xref="GeneID:[104415378](#)"

[mRNA](#)

1..1862
/gene="LOC104415378"
/product="rubisco accumulation factor 1.1, chloroplastic"
/note="Derived by automated computational analysis using gene prediction method: Gnomon. Supporting evidence includes similarity to: 3 Proteins, and 100% coverage of the annotated genomic feature by RNAseq alignments"
/transcript_id="[XM_010026669.3](#)"
/db_xref="GeneID:[104415378](#)"

[CDS](#)

148..1542
/gene="LOC104415378"
/note="Derived by automated computational analysis using gene prediction method: Gnomon."
/codon_start=1
/product="rubisco accumulation factor 1.1, chloroplastic"
/protein_id="[XP_010024971.3](#)"
/db_xref="GeneID:[104415378](#)"

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ ACESSANDO O GENE DE INTERESSE - SEQUÊNCIA FASTA DO GENE

Eucalyptus grandis isolate ANBG69807.140 chromosome 2, ASM1654582v1, whole genome shotgun sequence

NCBI Reference Sequence: NC_052613.1

[GenBank](#) [Graphics](#)

```
>NC_052613.1:41847086-41850549 Eucalyptus grandis isolate ANBG69807.140 chromosome 2,
ASM1654582v1, whole genome shotgun sequence
AAAAAGAGCTCGAAGAATCTTCTGGAACCTCCCCTGTCCGTGCCTTAAGCCCTAAATCCAACACGGGCAG
AACCCTCTCTCCGTGCGAGCGCCCTCTCCTCCTTATCCATACATCCCGCATTGCAAAACCTCCTCCT
TCGTCTCCCGCACC GCCACAGCTCCGCCGCAAAACCCTAGCCAGGCTCCGGAAATGGCGTCGGCCAACG
CTATCTCCACGGCTCCATCCTCTGCTCCCCAAGCAGGTAACCGTTTCCCCCGCTCCGCGAACATCC
GCGCGACCCCGCGTCGCCGGACGTCGCTTCCGGCGGCGCTTCCGTCTCCATTTTCGGAGCGGATGAATG
CGAATCTTCGGCTCTCGGTCGCCGCGATGATTGAGAAGTTCTCGAGACTAGAATCGAGCCTCTCGCGGTT
GGTTGGGAAGCAGGTTGGTTGGTTTTACTTACCGTCGGTCGTTGTGTGTTTTGCGTCTGCCGTGATGTA
GGGGGGCTTGAGGAGGAAAGGGAACCAGAAGGTGAACCTCGGACAGTCCGGCAGGAGGTTCCGCGTGAGA
GCCTCCGCCAAGGAAATTGCGTTCGACCAGAAATCCAGGGCCGCTCTTCAGGCCGGCATCGACAAGCTCG
CGGACGCCGTCCGCCCTCAGCTCGGTCCGAGGGGTAAAGCTAGCTACGGTTCTCTTTTGATGTTGACTGG
ATGAGTTGTGTCGTCCGATGTGCCTATTCGAGTGGACGGTGAGGTCTGCAAGCTGTAGCGATTGGATTG
ACTGTAGTTCCCATTCATGTTTAGAACTCCTTGCAATTTATCTCTTCCATTATTGAACTCATAGGTTTA
TCTCTTTTGATATGGCTTTTTTTTTTTTTTTTTTCCGTAGCAATAGAGTGGCAGTGGTTATGCAAACTGT
ACTTATTGGGAAATTATTGCTCTGCTCTCTGCAGGGAGAAATGTCGTCTTGGATGAATTTGGCAGTCCAA
AAGTTGTAAATGACGGGGTCACAATTGCCCGAGCAATTGAGCTACCTGATGCTATGGAAAATGCTGGTGC
CGCTCTCATTAGAGAGGTGAGACAATCACATGGCAGTTTTGCACTCACTCTCAGTCAAGGACTTTCCATC
TTAAGAAAGTACCCTTTTCAGTTAGAATTGCTTCCAACATATGTTATTTTGCTGATAAATTTTGGTTCAA
TGCAGGTTGCAAGCAAACTAACGACTCTGCAGGTGATGGTACAACACTGTCATCAGTCCTTGCTCGTGA
AATAATTAAATTGGGATTGCTGAGTGTTACATCCGGTGCAAAATCCAGTTTCAGTGAAGAGAGGCATCGAT
AAACTGTGCAAGGATTGATACAGGAGCTGGAGAAGAGGGCTAGGCCGGTTGAAGGACGTGATGACATTA
AAGGTAATGTGCTCCCATGGTCCATGAAAGACGAAGCTTTGAGTGGAACAAGTCTGTTACACTGATACTC
TTTTGGGTTTCATTGTAGCCGTTGCGACAATTTCTGCTGGAAATGATGAGCTCATTGGAACAATGATTGC
GGATGCAATTGACAGGGTTGGCCCTGATGGTGTGTTGTCCATTGAATCATCATCGTTTGAGACAAC
GTTGATGTAGAGGAGGGAATGGAGGTCAGCATTTTCGTCAATTTTCCTTTGTGCGTCTTGTGTTTCATTAG
TTGGGGAAGAAAAGCATTCTGCCTTTTTCCCCTCCTGGTGGAGAAGGTGGTGGAAAGTTTTCACTCGTTCT
TATCTGTTCTTACAAATGTTTTTATTGGCTTGAGATAGACAGAGGATACATATCCCCTCAATTCGTTAC
```

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

Reference ASM1654582v1 Primary Assembly

Genomic

1. NC_052613.1 Reference ASM1654582v1 Primary Assembly

Range	41847086..41850549
Download	GenBank , FASTA , Sequence Viewer (Graphics)

mRNA and Protein(s)

1. [XM_010045489.2](#) → [XP_010043791.2](#) ruBisCO large subunit-binding protein subunit alpha

UniProtKB/TrEMBL	A0A059D5S8		
Conserved Domains (1) summary			
	<table><tr><td>cd03344 Location:45 → 565</td><td>GroEL; GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They share a common general morphology, a double toroid of 2 stacked rings, each composed of 7-9 subunits. The symmetry of type I is seven-fold and they are found ...</td></tr></table>	cd03344 Location:45 → 565	GroEL; GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They share a common general morphology, a double toroid of 2 stacked rings, each composed of 7-9 subunits. The symmetry of type I is seven-fold and they are found ...
cd03344 Location:45 → 565	GroEL; GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They share a common general morphology, a double toroid of 2 stacked rings, each composed of 7-9 subunits. The symmetry of type I is seven-fold and they are found ...		

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

PREDICTED: *Eucalyptus grandis* rubisco accumulation factor 1.1, chloroplastic (LOC104415378), mRNA

NCBI Reference Sequence: XM_010026669.3

[FASTA](#) [Graphics](#)

Go to: ☐

LOCUS XM_010026669 1862 bp mRNA linear PLN 04-FEB-2021

DEFINITION PREDICTED: *Eucalyptus grandis* rubisco accumulation factor 1.1, chloroplastic (LOC104415378), mRNA.

ACCESSION XM_010026669

VERSION XM_010026669.3

DBLINK BioProject: [PRJNA698663](#)

KEYWORDS RefSeq.

SOURCE *Eucalyptus grandis*

ORGANISM [Eucalyptus grandis](#)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae; Pentapetales; rosids; malvids; Myrtales; Myrtaceae; Myrtoideae; Eucalypteae; Eucalyptus.

COMMENT MODEL [REFSEQ](#): This record is predicted by automated computational analysis. This record is derived from a genomic sequence ([NC_052619.1](#)) annotated using gene prediction method: Gnomon. Also see:

[Documentation](#) of NCBI's Annotation Process

On Feb 4, 2021 this sequence version replaced [XM_010026669.2](#).

##Genome-Annotation-Data-START##

Annotation Provider :: NCBI

Annotation Status :: Full annotation

Annotation Name :: [Eucalyptus grandis Annotation Release 102](#)

Annotation Version :: 102

Annotation Pipeline :: NCBI eukaryotic genome annotation pipeline

Annotation Software Version :: [8.5](#)

Annotation Method :: Best-placed RefSeq; Gnomon

Features Annotated :: Gene; mRNA; CDS; ncRNA

##Genome-Annotation-Data-END##

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

FEATURES	Location/Qualifiers
source	1..2245 /organism="Eucalyptus grandis" /mol_type="mRNA" /isolate="ANBG69807.140" /db_xref="taxon: 71139 " /chromosome="2" /tissue_type="leaf" /dev_stage="tree" /country="Australia"
gene	1..2245 /gene="LOC104432900" /note="Derived by automated computational analysis using gene prediction method: Gnomon. Supporting evidence includes similarity to: 31 ESTs, 7 Proteins, and 100% coverage of the annotated genomic feature by RNAseq alignments, including 65 samples with support for all annotated introns" /db_xref="GeneID: 104432900 "
CDS	195..1943 /gene="LOC104432900" /codon_start=1 /product="ruBisCO large subunit-binding protein subunit alpha" /protein_id="XP_010043791.2" /db_xref="GeneID: 104432900 "



CDS
Coding Protein Sequence

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

```
/db_xref= GeneID: 104437910  
/translation="MASANAISTASILCSPKQGGLRRKGNQKVNFGQSGRRFAVRSA  
KEIAFDQKSRAALQAGIDKLADAVGLTLGPRGRNVVLDEFGSPKVVNDGVTIARAI  
PDAMENAGAALIREVASKTND SAGDGT TTSVLA REIIKLGLLSVTSGANPVSVKRG  
DKTVQGLIQELEKRARPVEGRDDIKAVATISAGNDELIGTMIADAIDRVGPDGVL  
SSSSFETTVDVEEGMEIDRGYISPQFVTNPEKLIVEFENARVLVTDQKISAIKDI  
PLEKTQQLRAPLLIIAEDVTGEALATLVNKL RGIINVA AIKAPGFGERRKAILQD  
IAILTGA EYQASDLG LLVENTSVEQLGIARKVTISKDSMTIIADAASKDELQAR  
IAQLKKE LSETDSVYDTEKLAERIAKLSGGVAVIK  
IVPGGGA AFVHLSACVPAIKDKLEDADEKLGADIIQKALV  
APASLIAQNAGIEGEVVVEKVKAREWEIGYNAMTDKYENLVEAGVIDPAKVTR  
CALQNSASVAGMVLTTQAIIVVEKAKPKAPAAASPQGLTI"
```

ORIGIN

```
1 atggcgtcgg ccaacgctat ctccacggcc tccat  
61 ttgaggagga aagggaacca gaaggtgaac ttcgg  
121 agagcctccg ccaaggaaat tgcgttcgac cagaa  
181 atcgacaagc tcgcggacgc cgtcggcctc acgct  
241 ttggatgaat ttggcagtc aaagattgta aatga  
301 gagctacctg atgctatgga aaatgctggg gccgc  
361 actaacgact ctgcagggtg tggatcaact actgc  
421 aaattgggat tgctgagtgt tacatccggg gcaaa  
481 gataaaactg tgcaaggatt gatacaggag ctgga  
541 cgtgatgaca ttaaagccgt tgcgacaatt tctgc  
601 atgattgcgg atgcaattga cagggttggc cctga  
661 tcatcgtttg agacaactgt tgatgtagag gaggg  
721 tcccctcaat tcgttactaa ccctgagaag ttgat  
781 ctggtgacag atcaaaagat ttcggcaata aaggatatta tccccttgct ggagaagaca  
841 actcagctga gagcgccttt gctcataatt gctgaagatg tgactggtga ggctttggct  
901 actcttgtcg tgaacaagtt gcggggtatc attaatgttg ctgctatcaa agcccctgga  
961 tttggtgaaa ggagaaaggc gatcctacaa gatattgcaa tcttgacagg agctgagtac  
1021 caagccagtg atcttggttt gcttggttag aacacatcag ttgagcagct tgggtagact  
1081 agaaaagtaa ccatctctaa ggattctatg accattattg ctgatgcggc atcaaaggat  
1141 gaactacaag ccaggattgc acaattgaaa aaagattgtg ctgaaactga ttctgtatat  
1201 gacacagaga aactcgtgta gaggattgcc aagctatccg ggggagttgc tgtgatcaaa  
1261 gttggtgcag ctactgagac tgagcttgag gataggaagc ttcggattga ggatgccaa  
1321 aatgcaactt ttgctgccat agaggaaggc atagtgcctg gcggtggtgc cgcatttgtt  
1381 cacctctctg cttgcgtgcc tgctatcaag gacaagcttg aggatgcaga tgaaaaattg
```

```
>XP_010043791.2 ruBisCO large subunit-binding protein subunit alpha [Euc  
MASANAISTASILCSPKQGGLRRKGNQKVNFGQSGRRFAVRSAKEIAFDQKSRAALQAGIDKLADAVGL  
TLGPRGRNVVLDEFGSPKVVNDGVTIARAIELPDAMENAGAALIREVASKTND SAGDGT TTSVLA REII  
KLGLLSVTSGANPVSVKRGIDKTVQGLIQELEKRARPVEGRDDIKAVATISAGNDELIGTMIADAIDRVG  
PDGVL SIESSSSFETTVDVEEGMEIDRGYISPQFVTNPEKLIVEFENARVLVTDQKISAIKDIIP LLEKT  
TQLRAPLLIIAEDVTGEALATLVNKL RGIINVA AIKAPGFGERRKAILQDIAIILTGA EYQASDLG LLVE  
NTSVEQLGIARKVTISKDSMTIIADAASKDELQARIAQLKKE LSETDSVYDTEKLAERIAKLSGGVAVIK  
VGAATETELED RKLRIEDAKNATFAAIEEGIVPGGGA AFVHLSACVPAIKDKLEDADEKLGADIIQKALV  
APASLIAQNAGIEGEVVVEKVKAREWEIGYNAMTDKYENLVEAGVIDPAKVTRCALQNSASVAGMVLTTQ  
AIIVVEKAKPKAPAAASPQGLTI
```

BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

➤ FERRAMENTA BLAST (*BASIC LOCAL ALIGNMENT SEARCH TOOL*)

Desenvolvida para realizar **buscas** e **comparar** sequências biológicas contra um banco de dados que contém uma grande quantidade de informação, retornando sequências mais **similares**, referentes à sequência pesquisada.

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

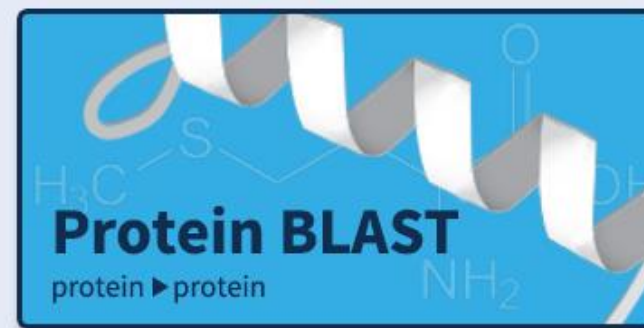
ClusteredNR database on BLAST+

The ClusteredNR database is now available for BLAST+

Thu, 24 Aug 2023

[More BLAST news...](#)

Web BLAST



BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

➤ COMPARANDO SEQUENCIAS - ALINHAMENTO

O alinhamento de sequencias consiste em comparar duas sequencias (de nucleotídeos ou aminoácidos) de forma a identificarmos **o grau de similaridade (identidade)** entre elas.

Identidade

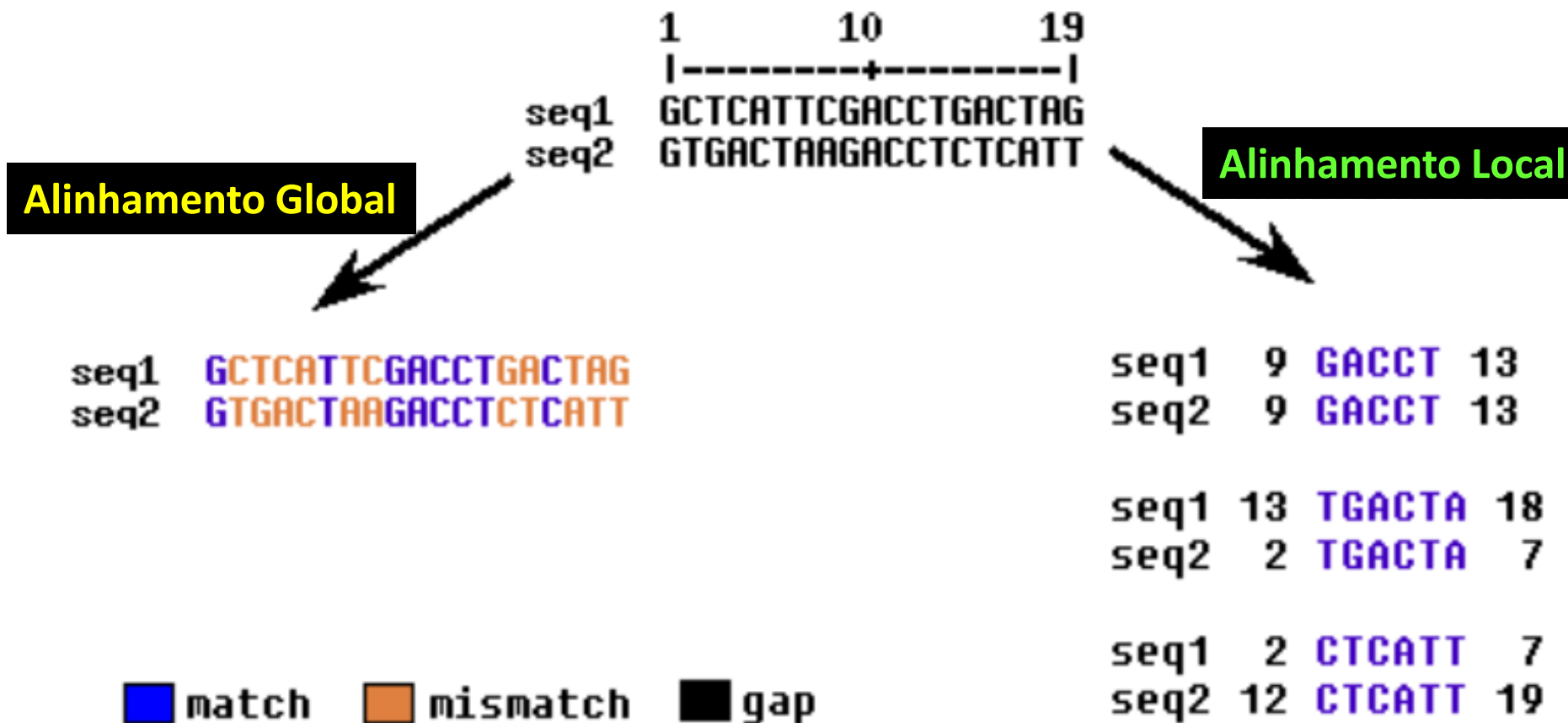
Número de posições invariáveis em duas sequências (nucleotídeos ou aminoácidos) alinhadas.

```
RBP:                26  RVKENFDKARFSGTWYMAKKDPEGLFLQDNIVA 59
                      + K++ + ++ GTW++MA + L + A
glicodelina: 23  QTKQDLELPKLAGGTWHMAMA-TNNISLMATLKA 55
```

BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

➤ COMPARANDO SEQUENCIAS - ALINHAMENTO

Pode ser utilizado para: identificação de genes e proteínas desconhecidas, comparação da ordem de genes em genomas de organismos proximamente relacionados, na montagem de genomas ...



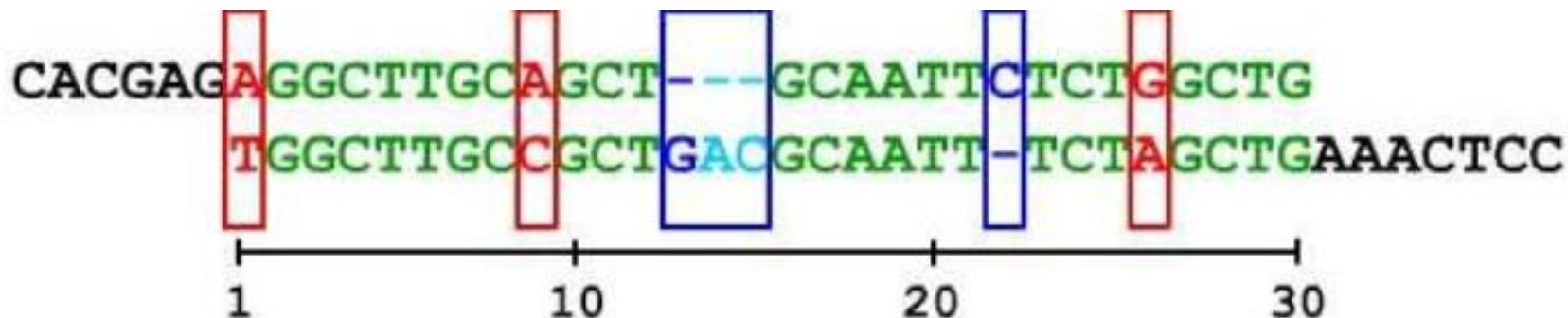
BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

➤ COMPARANDO SEQUENCIAS - ALINHAMENTO

Alinhamento Global: é feito quando comparamos uma sequência de aminoácidos ou nucleotídeos com outra, ao longo de toda sua extensão

```
seq1  GCTCATTGACCTGACTAG
seq2  GTGACTAAGACCTCTCATT
```

Alinhamento Local: acontece quando a comparação entre duas sequências não é feita ao longo de toda sua extensão, mas sim através de pequenas regiões destas. O *BLAST* é o principal programa para realizar o alinhamento local.



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

EX: **XM_010026669.3** **PREDICTED: Syzygium oleosum rubisco accumulation factor 1.1, chloroplastic (LOC115690120), i**
Sequence ID: [XM_030616373.2](#) Length: 1930 Number of Matches: 1

Range 1: 56 to 1928 [GenBank](#) [Graphics](#) [▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Gaps	Strand
2082 bits(1127)	0.0	1643/1886(87%)	60/1886(3%)	Plus/Plus
Query 24	aataaaaaaaTC-AAAGCCATCAGTGAACCTTCAGTTCAGATAAGGTTTCACTGAAAACCC	82		
Sbjct 56	AAGAAAAATATCAAAAGCCATCATTCAATTTTCAGTTCAGATAAGGTTTCACTGAAGACTC	115		
Query 83	ACAGTGACCTCCCTCCACTCCTCCCTGAACCACCAACTCCCGCCACCGCCACCGCC---T	139		
Sbjct 116	ACAGTGACCTCCTCCAC-CCTCACTGCACCACCGCCTCCCGCCGCCGCCACCGCCGCCG	174		
Query 140	CCG---CCACCATGCTCTCTCTAGCCCTCATCAAcaccacccac---cc-ccaccaccacc	193		
Sbjct 175	CCGCCTCCACCATGCTCTCTCTAGCCCTCATCAACACCACCCACACCCACCACCACCACC	234		
Query 194	aattcttcaatcccaccaccaccatcctcaaacccctcacttcctctccctccaccacc	253		
Sbjct 235	AATTCTTCAACCCCAACCACCA-CCAC-CA-CCCTTA-AGCCCCTCACTTCCCCTTTC	290		
Query 254	cttgccgccaccaccgccgccaccactccgccggcccaaggcccatctccgccGCCTCA	313		
Sbjct 291	C-T-CCACCACCCCAACCACCACTCGGCCGCCCGAGGCCCATCTCCGCCAGCCTCA	348		
Query 314	ACCCGAACCTCTAAtcccggcctccctccgccgcagcagcagcgccctaccagccct	373		
Sbjct 349	ACCCGAGCCCTCTAATCCCGGCCTCCCTCCGCCGCAGCAGCAGCGTGTCTACCAGCCCT	408		
Query 374	tccgccctccccc-tccccgct--cctcccagttccgctccctcGACACCGCCAGCCGCC	430		
Sbjct 409	TCCGCCCTCCCCCTCCCGCTCCCTCCAGTTCGCTCCCTCGACACCGCCGCCGCCGCC	468		

BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

Standard Nucleotide BLAST

blastn

blastp

blastx

tblastn

tblastx

BLASTN programs search nucleotide databases using a nucleotide query. more...

Reset page

Bookmark

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) ? Clear

Query subrange ?

From

To

Or, upload file

Escolher arquivo

Nenhum arquivo escolhido ?

Job Title

Enter a descriptive title for your BLAST search ?

☐ Align two or more sequences ?

Choose Search Set

Database

☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus

New

☐ Experimental databases

Try experimental taxonomic nt databases

For more info see [What are taxonomic nt databases?](#)

Download

Nucleotide collection (nr/nt) v ?

Organism

Optional

Enter organism name or id--completions will be suggested



exclude

Add organism

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown ?

Exclude

Optional

☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences

Limit to

Optional

☐ Sequences from type material

Feedback

BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

➤ FERRAMENTA BLAST (*BASIC LOCAL ALIGNMENT SEARCH TOOL*)

Desenvolvida para realizar **buscas** e **comparar** sequências biológicas contra um banco de dados que contém uma grande quantidade de informação, retornando sequências mais **similares**, referentes à sequência pesquisada.

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

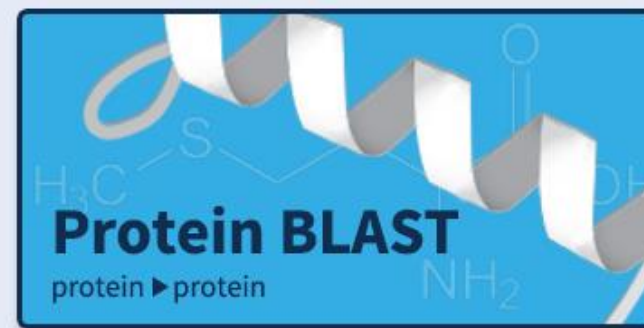
ClusteredNR database on BLAST+

The ClusteredNR database is now available for BLAST+

Thu, 24 Aug 2023

[More BLAST news...](#)

Web BLAST



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

- BLASTn**: pesquisa uma sequência de **nucleotídeo** em uma base de dados de **nucleotídeos**, o que é útil na identificação de sequências desconhecidas obtidas de sequenciamentos e PCR. Além disso o programa busca relacionar as sequências de forma evolutiva (busca por homologia) com a sequência de entrada e identificar genes.
- BLASTp**: realiza a comparação da **sequência proteína-proteína**. Neste tipo de busca é necessário ter uma sequência de aminoácidos para comparar com sequências já identificadas nos bancos de dados. O BLASTp também pode verificar a homologia entre sequências, conservação de aminoácidos e diferenças pontuais que interferem na atividade final das proteínas.
- BLASTx**: pesquisa **nucleotídeos** em um banco de dados de **proteínas**, traduzindo a sequência de interesse em tempo real. Esta é uma ferramenta muito importante para conhecer o produto gênico quando se tem apenas dados genômicos e não se tem ciência dos produtos que são codificados pelos genes representados na sequência de interesse.

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

•**tBLASTn**: pesquisa **proteínas** em um banco de dados de **nucleotídeos**, traduzindo o banco de dados em tempo real. Essa opção permite inferir quais genes ou porções do genoma estão relacionadas com uma determinada proteína, quando não se tem a sequência de nucleotídeos disponível.

tBLASTx: a sequência de **nucleotídeos** será convertida em **6 sequências de aminoácidos** (uma para cada fase de leitura), as quais serão comparadas com as 6 possíveis fases de leitura em um banco de dados de nucleotídeos. Isso é muito útil na identificação de proteínas provindas de diversos rounds de sequenciamento e novos genes.

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

FASTA

```
>XM_010026669.3 PREDICTED: Eucalyptus grandis rubisco accumulation factor 1.1,
chloroplastic (LOC104415378), mRNA
CTCCCCAAGACTTGTCCAAAAAATAAAAAATCAAAGCCATCAGTGAAC TTCAGTTCAGATAAGGTTT
CACTGAAAACCCACAGTGACCTCCCTCCACTCCTCCCTGAACCACCAACTCCCGCCACCGCCACCGCCTC
CGCCACCATGCTCTCTCTAGCCCTCATCAACACCACCCACCCCCACCACCACCAATTCTTCAATCCCACC
ACCACCATCCTCAAACCCCTCACTTCCTCTCCCTCCACCACCCTTGCCGCCACCACCGCCGCCACCACT
CCGCCGGCCCAAGGCCCATCTCCGCCAGCCTCAACCCGAACCCTCTAATCCCGGCCCTCCCTCCGCCGCA
GCAGCAGCGGCCCTACCAGCCCTTCCGCCCTCCCCCTCCCCGCTCCTCCAGTTCCTGCTCCCTCGACACC
GCCAGCCGCTCGACACCCTCTCCAACCGCCTCGGCCTCTGGTTCGAGTACGCCCCCTCATCCCCTCCC
TCATCCAGGAAGGCTTCACCCCTCCCTCCATCGAGGAGGTCACCGGCATCTCCGGCGTCGAGCAGAACCG
CCTCGTCGTCGCCGCCAAGGTCCGGGACTCCCTGATCCAATCCGACATCGACCCGGACGTCGTCGCCACC
TTCGACATCGGCGGCGCCGAGTCTCTACGAAATCCGGCTCCTGAGCGTCGCGCAGCGGTCCGCCGCGG
CGGAATTCATCGTCAGGAACAAGTTCGACGCGAGGGGCGCGCAGGACCTCGCCCGCGCCATGAAGGATTT
CCCCCGCCGGCGCGGGGATAGGGGATGGGAGAAATTTCAGTATAAATTGCCCGGAGACTGCTTGTCTTTC
ATGTATTATAGGCAGAGCAGAGAGCACAAGAACCCTCGGAGCAGAGGACGGCGGCGTTGGAGCAGGCGC
TGAGCGTGCGGAGACGGAGAAAGCGAAGGGCGTGATCCTGGAGGAATTGGAAGGGAAAGGTGGTGAGGG
AGATGGAAAAGAAGGGCAAGTGCAGGATGGGATTCGGGTGCCGGTGGTGAGGATGAAGACGGGGGAGGTG
GCGGAGGCGAGCACGGTGGCGGTGTTGCCGGTGTGCCGGTGGAGGAGAGGGAAAGAGAGGTCCTGGAGG
CTCCGTGGCGGTGCAAGAGCGAAGGGGAATTCGGGGTGGTGGTGCGGATAAAGGGTGGGCGAGGTGGGT
GGTTTTGCCGGGTGGGAGCCGGTGGTGAGACTGTCCGGCGGAGGAGTGGTGGTGTGCTTCCCGGACGCG
AGGGCCTTGCCGTGGAAGGTGAACAGGTGGTACAACGAGGAGGCCATCTTGGTTGTGGCCGACCGGGGGA
GGAAAGAGGTGGTCGAGGACGAAGGCTTTTATATGGTGGCGGTGGATGCCGACGATGACGCCAACGCCGC
GGGCGGCTACGGATTGAAGGTGGAGAGAGGGCGGAGTTTGAAGGAAATGGGGGTGGAGGAGAGCCTAGGG
ACGGTGGTGTGGTGGTGAGGCCGCCAAAAGAAGAGATTGAGGATCAATTGAGTGATGAAGATTGGGAGT
GAGAGGCAACAGCATGAACGTGAGGAGTTTATTGATGTGTACAATATTGATACAGATCAATATGCAACTT
TTTTGGATCTGAGACTATGGTGTGGAGACATCAACTTTTCAGGTTGAAAAGGGTTTTTATAAAATTACCC
ATTTTGGAGATTGGAAGGAAACCTCTTTTGATGTTTCTCTGTTGTTGTCTAAATACCAGACTTTGTTCA
```

blastn

blastp

blastx

tblastn

tblastx

BLASTN programs

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) ? Clear

XM_010026669.3

Or, upload file

Escolher arquivo Nenhum arquivo escolhido ?

Job Title

Enter a descriptive title for your BLAST search ?

☐ Align two or more sequences ?

Choose Search Set

Database

☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐

New

☐ Experimental databases

<

Try experiment

For more info see

Nucleotide collection (nr/nt)

Enter organism name or id--completions will be suggested

Enter organism common name, binomial, or tax id. Only 20 top taxa

Exclude

Optional

Limit to

Optional

☐ Models (XM/XP) ☐ Uncultured/environmental sample se

☐ Sequences from type material

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTn

BLAST® » blastn suite » results for RID-J2KSB1E701R

HomeRecent ResultsSaved StrategiesHelp

👤 Check out the ClusteredNR database on BLAST+

Learn more

Give us feedback

◀ Edit Search

Save Search

Search Summary ▼

🔗 How to read this report?

📺 BLAST Help Videos

↶ Back to Traditional Results Page

Job Title

ref|XM_010026669.3|

RID

[J2KSB1E701R](#) Search expires on 10-08 20:28 pm [Download All](#) ▼

Program

BLASTN ? [Citation](#) ▼

Database

nt [See details](#) ▼

Query ID

[XM_010026669.3](#)

Description

PREDICTED: Eucalyptus grandis rubisco accumulation fa ...

Molecule type

nucleic acid

Query Length

1862

Other reports

[Distance tree of results](#) [MSA viewer](#) ?

Filter Results

Organism only top 20 will appear ☐ exclude

[+ Add organism](#)

Percent Identity

to

E value

to

Query Coverage

to

Filter

Reset

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTn

Descriptions									
Sequences producing significant alignments									
Download Select columns Show 100 ?									
☒ select all 94 sequences selected									
GenBank Graphics Distance tree of results MSA Viewer									
	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒	PREDICTED: Eucalyptus grandis rubisco accumulation factor 1.1, chloroplastic (LOC104415378), mRNA	Eucalyptus gra...	3439	3439	100%	0.0	100.00%	1862	XM_010026669.3
☒	PREDICTED: Syzygium oleosum rubisco accumulation factor 1.1, chloroplastic (LOC115690120), mRNA	Syzygium oleos...	2082	2082	98%	0.0	87.12%	1930	XM_030616373.2
☒	PREDICTED: Rhodamnia argentea rubisco accumulation factor 1.1, chloroplastic-like (LOC115743939),...	Rhodamnia arg...	1844	1844	93%	0.0	86.17%	1781	XM_030678960.2
☒	PREDICTED: Camellia sinensis rubisco accumulation factor 1.1, chloroplastic-like (LOC114286420), m...	Camellia sinensis	640	640	63%	5e-178	76.79%	1818	XM_028229585.1
☒	PREDICTED: Camellia sinensis rubisco accumulation factor 1.1, chloroplastic-like (LOC114285100), m...	Camellia sinensis	640	640	63%	5e-178	76.79%	1816	XM_028228093.1
☒	Alnus glutinosa genome assembly, chromosome: 12	Alnus glutinosa	617	1179	63%	2e-171	76.63%	27214496	OY340909.1
☒	Theobroma cacao genome assembly, chromosome: 5	Theobroma cacao	115	115	5%	4e-20	86.54%	41542769	OY284442.1
☒	PREDICTED: Rhodamnia argentea uncharacterized LOC115743940 (LOC115743940), transcript varian...	Rhodamnia arg...	110	110	4%	2e-18	91.46%	1889	XM_048280600.1
☒	Microcos paniculata chromosome 2b	Microcos panic...	73.1	73.1	5%	2e-07	79.05%	49527071	CP131448.1

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTn

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
✓	PREDICTED: Eucalyptus grandis rubisco accumulation factor 1.1, chloroplastic (LOC104415378), mRNA	Eucalyptus gra...	3439	3439	100%	0.0	100.00%	1862	XM_010026669.3
✓	PREDICTED: Syzygium oleosum rubisco accumulation factor 1.1, chloroplastic (LOC115690120), mRNA	Syzygium oleos...	2082	2082	98%	0.0	87.12%	1930	XM_030616373.2
✓	PREDICTED: Rhodamnia argentea rubisco accumulation factor 1.1, chloroplastic-like (LOC115743939),...	Rhodamnia arg...	1844	1844	93%	0.0	86.17%	1781	XM_030678960.2
✓	PREDICTED: Camellia sinensis rubisco accumulation factor 1.1, chloroplastic-like (LOC114286420), m...	Camellia sinensis	640	640	63%	5e-178	76.79%	1818	XM_028229585.1

•O **max score** é o valor máximo que pode ser obtido naquele alinhamento, enquanto o **Total Score** é o valor obtido pelo alinhamento. Se o *total score* for igual ao *max score* significa que o alinhamento obteve o maior valor possível.

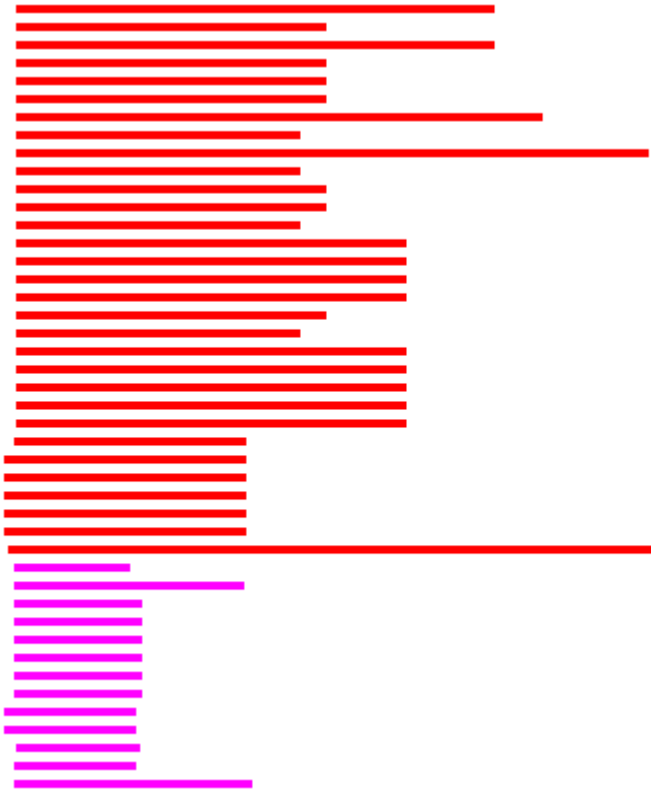
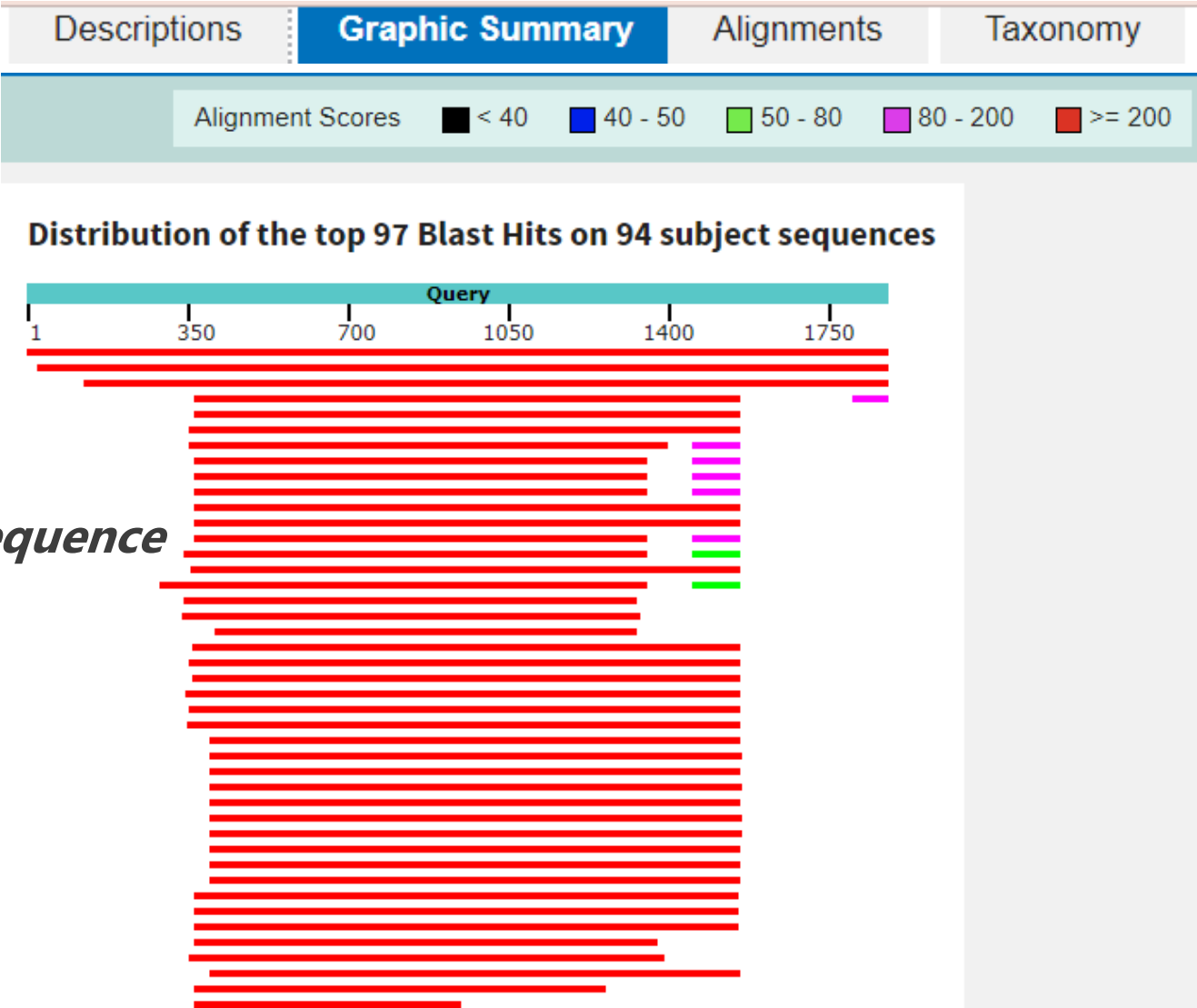
•Cobertura da **query** demonstra o quanto da sequência enviada conseguiu ser alinhada.

•**E-value** é um parâmetro muito importante, demonstra a possibilidade do alinhamento ter sido realizado ao acaso. Quanto mais próximo ao zero (0) o valor, mais confiabilidade pode se ter no alinhamento.

•A **Porcentagem de Identidade** está relacionada com a similaridade da sequência enviada pelo usuário com a sequência alinhada, levando-se em consideração a cobertura da sequência.

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTn



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTn - ALIGNMENTS

PREDICTED: Syzygium oleosum rubisco accumulation factor 1.1, chloroplastic (LOC115690120), i

Sequence ID: [XM_030616373.2](#) Length: 1930 Number of Matches: 1

Range 1: 56 to 1928 [GenBank](#) [Graphics](#)

▼ [Next Match](#) ▲

Score	Expect	Identities	Gaps	Strand
2082 bits(1127)	0.0	1643/1886(87%)	60/1886(3%)	Plus/Plus
Query 24	aataaaaaaaTC-AAAGCCATCAGTGAAC TTCAGTT CAGATAAGGTTTCACTGAAAACCC			82
Sbjct 56	AAGAAAAATATCAAAAGCCATCATTCAATTT CAGTT CAGATAAGGTTTCACTGAAGACTC			115
Query 83	ACAGTGACCTCCCTCCACTCCTCCCTGAACCACCAACTCCCGCCACCGCCACCGCC---T			139
Sbjct 116	ACAGTGACCTCCCTCCAC-CCTCACTGCACCACCGCCTCCCGCCGCCGCCACCGCCGCCG			174
Query 140	CCG---CCACCATGCTCTCTCTAGCCCTCATCAAcaccacccac---cc-ccaccaccacc			193
Sbjct 175	CCGCCTCCACCATGCTCTCTCTAGCCCTCATCAACACCACCCACACCCACCACCACCACC			234
Query 194	aattcttcaatcccaccaccaccatcctcaaaccctcacttcctctccctccaccacc			253
Sbjct 235	AATTCTTCAACCCCAACCACCA-CCAC-CA-CCCTTA-AGCCCTCACTTCCCCTTTC			290
Query 254	cttgccgccaccaccgccgccaccactccgccggcccaaggcccatctccgccGCCTCA			313
Sbjct 291	C-T-CCACCACCCCAACCACCACCTCGGCCGGCCCGAGGCCATCTCCGCCAGCCTCA			348
Query 314	ACCCGAACCTCTAAtcccgccctccctccgccgcagcagcagcgccctaccagccct			373
Sbjct 349	ACCCGAGCCCTCTAATCCCGGCCTCCCTCCGCCGAGCAGCAGCGTGTCTACCAGCCCT			408
Query 374	tccgccctccccc-tccccgct--cctcccagttccgctccctcGACACCGCCAGCCGCC			430
Sbjct 409	TCCGCCCTCCCCCTCCCGCTCCCTCCAGTTCCGCTCCCTCGACACCGCCGCCGCCG			468

Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Eucalyptus gra...	3439	3439	100%	0.0	100.00%	1862	XM_010026669.3
Syzygium oleos...	2082	2082	98%	0.0	87.12%	1930	XM_030616373.2

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

PREDICTED: Panicum virgatum rubisco accumulation factor 1, chloroplastic-like (LOC120688899)

Sequence ID: [XM_039971251.1](#) Length: 938 Number of Matches: 1

Range 1: 297 to 512 [GenBank](#) [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Gaps	Strand
180 bits(97)	1e-39	179/218(82%)	7/218(3%)	Plus/Plus
Query 360	gccctaccagcccttccgccctcc-ccctccccgct-cctcccag-ttccgctccctcGA 416			
Sbjct 297	GCTCTACCAGCCCTTCCACCCGCCGCGTCCCCGCTGCCGCCAGCTACCGCACCCCTCGA 356			
Query 417	C-ACCGCCAGCCGCCTCGACACCCTCTCCAACCGCCTCGGCCTCTGGTTCGAGTACGCCC 475			
Sbjct 357	CCTCACCAG-CGCCTCGAGGTCCTCCGCGACCGCATGGGCCGGTGGTACGAGTACGCGC 415			
Query 476	CCCTCATCCCCTCCCTCATCCAGGAA-GGCTTCACCCCTCCCTCCATCGAGGAGGTCACC 534			
Sbjct 416	CCCTCATCGCCGCGCTC-TCCCGCAACGGCTTCACCCCGTCCTCCATCCAGGAGGCCACC 474			
Query 535	GGCATCTCCGGCGTCGAGCAGAAACCGCCTCGTCGTCGC 572			
Sbjct 475	GGCATCTCCGGCGTCGAGCAGAACTGCCTCGTCGTCGC 512			

Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len
Panicum virgatum	180	180	11%	1e-39	82.11%	938

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

Standard Protein BLAST

blastn

blastp

blastx

tblastn

tblastx

BLASTP programs search protein databases using a protein query. [more...](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#)

XP_010043791.2

Query subrange [?](#)

From

To

Or, upload file

Escolher arquivo

Nenhum arquivo escolhido [?](#)

Job Title

XP_010043791:ruBisCO large subunit-binding...

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Databases

☒ Standard databases (nr etc.): **New** ☐ Experimental databases

Compare

☐ Select to compare standard and experimental database [?](#)



Try experimental clustered nr database

For more info see [What is clustered nr?](#)



Standard

Database

Non-redundant protein sequences (nr) [?](#)

Organism

Optional

☐

exclude

[Add organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTp

[← Edit Search](#)

[Save Search](#)

[Search Summary ▼](#)

[? How to read this report?](#)

[▶ BLAST Help Videos](#)

[↶ Back to Traditional Results Page](#)

Job Title	XP_010043791:ruBisCO large subunit-binding...
RID	J2T0E03G016 Search expires on 10-08 21:58 pm Download All ▼
Program	BLASTP ? Citation ▼
Database	nr See details ▼
Query ID	XP_010043791.2
Description	ruBisCO large subunit-binding protein subunit alpha [Euca ...
Molecule type	amino acid
Query Length	582
Other reports	Distance tree of results Multiple alignment MSA viewer ?

Filter Results

Organism only top 20 will appear

☐ exclude

[+ Add organism](#)

Percent Identity

to

E value

to

Query Coverage

to

[Filter](#)

[Reset](#)

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTp

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
✓	ruBisCO large subunit-binding protein subunit alpha [Eucalyptus grandis]	Eucalyptus gran...	1147	1147	100%	0.0	100.00%	582	XP_010043791.2
✓	hypothetical protein EUGRSUZ_B02532 [Eucalyptus grandis]	Eucalyptus gran...	1144	1144	100%	0.0	99.66%	656	KCW85781.1
✓	hypothetical protein EUGRSUZ_B02532 [Eucalyptus grandis]	Eucalyptus gran...	1139	1139	100%	0.0	99.49%	657	KCW85780.1
✓	hypothetical protein BT93_B2216 [Corymbia citriodora subsp. variegata]	Corymbia citriod...	1116	1116	100%	0.0	96.74%	582	KAF8039927.1
✓	hypothetical protein NL676_031747 [Syzygium grande]	Syzygium grande	1105	1105	100%	0.0	95.36%	582	KAI6685834.1
✓	ruBisCO large subunit-binding protein subunit alpha [Syzygium oleosum]	Syzygium oleosum	1098	1098	100%	0.0	94.85%	582	XP_030447891.1
✓	ruBisCO large subunit-binding protein subunit alpha [Rhodamnia argentea]	Rhodamnia arge...	1077	1077	100%	0.0	94.67%	582	XP_030552056.1
✓	PREDICTED: ruBisCO large subunit-binding protein subunit alpha [Populus euphratica]	Populus euphrat...	1071	1071	100%	0.0	91.64%	586	XP_011000529.1
✓	ruBisCO large subunit-binding protein subunit alpha [Populus trichocarpa]	Populus trichoca...	1070	1070	100%	0.0	91.30%	586	XP_002313525.1
✓	hypothetical protein NC652_022600 [Populus alba x Populus x berolinensis]	Populus alba x ...	1068	1068	100%	0.0	91.30%	586	KAJ6904624.1
✓	hypothetical protein POTOM_033871 [Populus tomentosa]	Populus toment...	1068	1068	100%	0.0	91.13%	586	KAG6760692.1
✓	ruBisCO large subunit-binding protein subunit alpha-like [Populus alba x Populus x berolinensis]	Populus alba x ...	1066	1066	100%	0.0	91.13%	586	KAJ6895201.1
✓	hypothetical protein SADUNF_Sadunf09G0005500 [Salix dunnii]	Salix dunnii	1066	1066	100%	0.0	91.30%	586	KAF9675181.1
✓	ruBisCO large-binding protein [Salix suchowensis]	Salix suchowensis	1066	1066	100%	0.0	91.13%	586	KAG5238052.1
✓	60 KDA HEAT SHOCK PROTEIN MITOCHONDRIAL [Salix koriyanagi]	Salix koriyanagi	1065	1065	100%	0.0	91.13%	586	KAJ6715735.1
✓	hypothetical protein OIU84_008426 [Salix udensis]	Salix udensis	1065	1065	100%	0.0	90.96%	586	KAJ6408726.1

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

hypothetical protein **EUGRSUZ_B02532** [*Eucalyptus grandis*]

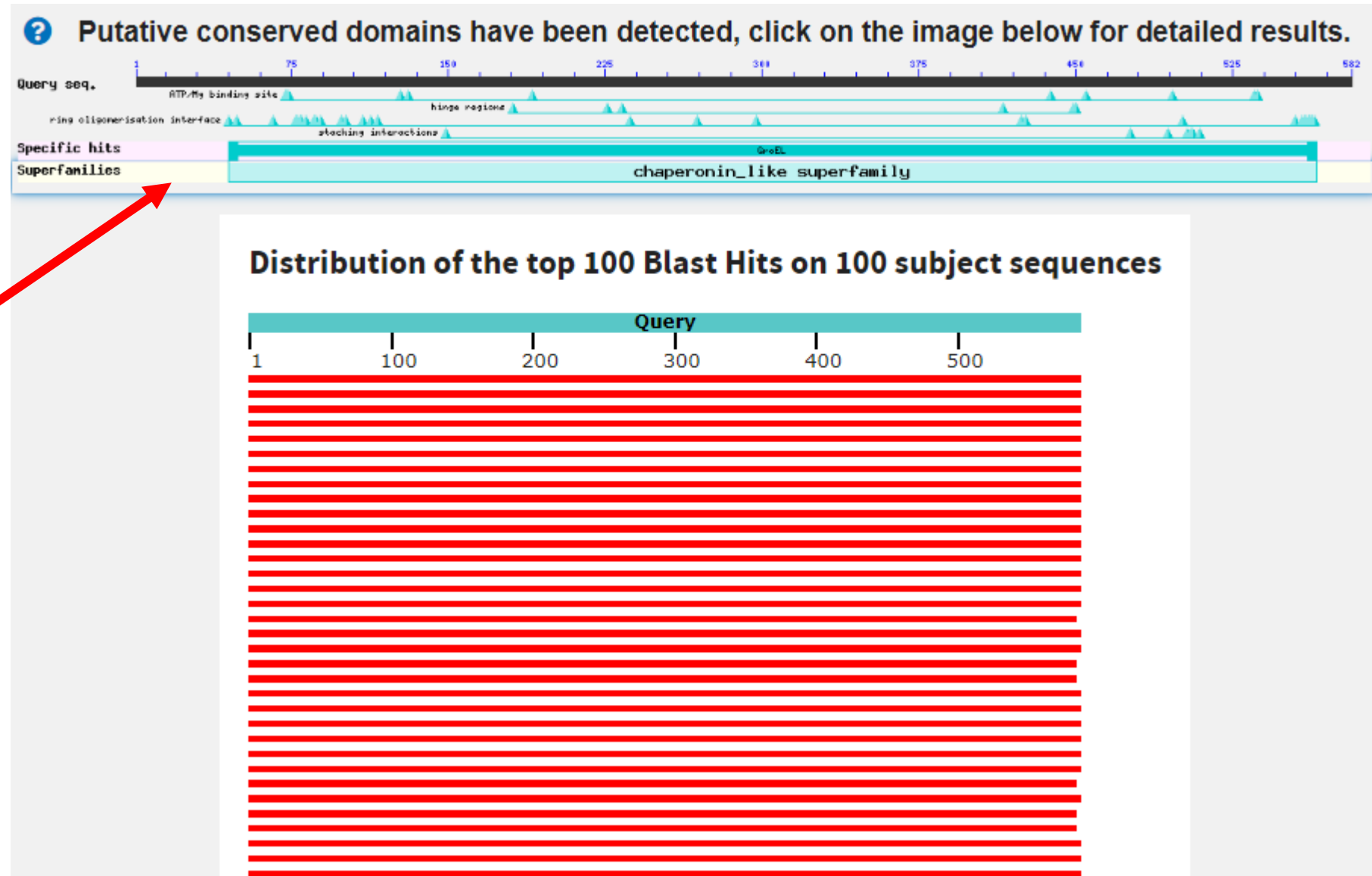
Sequence ID: [KCW85780.1](#) Length: **657** Number of Matches: **1**

Range 1: 75 to 657 [GenPept](#) [Graphics](#) [▼ Next Match](#) [▲ Pre](#)

Score	Expect	Method	Identities	Positives	Gaps
1139 bits(2946)	0.0	Compositional matrix adjust.	580/583(99%)	581/583(99%)	1/583(0%)
Query 1	MASANAISTASILCSPKQGGLRRKGNQKVNFGQSGRRFAVRASAKEIAFDQKSRAALQAG				60
Sbjct 75	MASANAISTASILCSPKQGGLRRKGNQKVNFGQSGRRFAVRASAKEIAFDQKSRAALQAG				134
Query 61	IDKLADAVGLTLGPRGRNVVLDEFGSPKVVNDGVTIARAIELPDAMENAGAALIREVASK				120
Sbjct 135	IDKLADAVGLTLGPRGRNVVLDEFGSPKVVNDGVTIARAIELPDAMENAGAALIREVASK				194
Query 121	TNSDAGDGTITASVLAREIIKLGLLSVTSGANPVSVKRGIDKTVQGLIQELEKRARPVEG				180
Sbjct 195	TNSDAGDGTITASVLAREIIKLGLLSVTSGANPVSVKRGIDKTVQGLIQELEKRARPVEG				254
Query 181	RDDIKAVATISAGNDELIGTMIADAIDRVGPDGVLSIESSSSFETTVDVEEGMEIDRGYI				240
Sbjct 255	RDDIKAVATISAGNDELIGTMIADAIDRVGPDGVLSIESSSSFETTVDVEEGMEIDRGYI				314
Query 241	SPQFVTNPEKLIVEFENARVLVTDQKISAIKDIIPLEKTTQLRAPLLIIAEDVTGEALA				300
Sbjct 315	SPQFVTNPEKLIVEFENARVLVTDQKISAIKDIIPLEKTTQLRAPLLIIAEDVTGEALA				374
Query 301	TLVVNKLKRGIIINVAAIKAPGFGERRKAILQDIAILTG-AEYQASDLGLLVENTSVEQLGI				359
Sbjct 375	TLVVNKLKRGIIINVAAIKAPGFGERRKAILQDIAILTGAEYQASDLGLLVENTSVEQLGI				434
Query 360	ARKVTISKDSMTIIADAASKDELQARIAQLKKELSETDSVYDTEKLAERIAKLSSGGVAVI				419
Sbjct 435	ARKVTISKDS TIIADAASKDELQARIAQLKKELSETDSVYDTEKLAERIAKLSSGGVAVI				494

BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTp



BANCO DE DADOS BIOLÓGICOS (BDB) - NCBI

➤ RESULTADO BLASTp

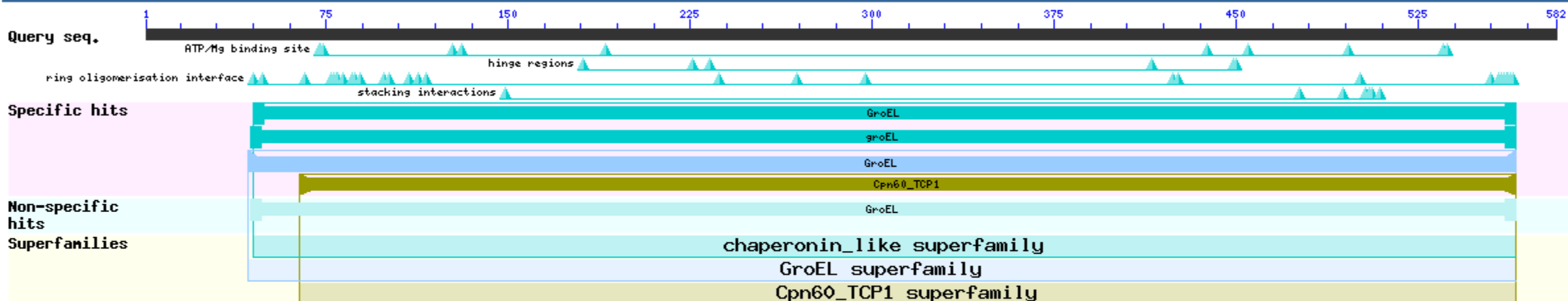
➤ DOMÍNIOS PROTEICOS

Conserved domains on [gi|1973640993|ref|XP_010043791|]

View Standard Results ?

ruBisCO large subunit-binding protein subunit alpha [Eucalyptus grandis]

Graphical summary ☐ Zoom to residue level show extra options ?



Search for similar domain architectures ?

Refine search ?

List of domain hits

+	Name	Accession	Description	Interval	E-value
[+]	GroEL	cd03344	GroEL_like type I chaperonin. Chaperonins are involved in productive folding of proteins. They ...	45-565	0e+00
[+]	groEL	PRK00013	chaperonin GroEL; Reviewed	44-565	0e+00
[+]	GroEL	TIGR02348	chaperonin GroL; This family consists of GroEL, the larger subunit of the GroEL/GroES ...	44-565	0e+00
[+]	GroEL	COG0459	Chaperonin GroEL (HSP60 family) [Posttranslational modification, protein turnover, chaperones]; ...	43-565	4.84e-150
[+]	Cpn60_TCP1	pfam00118	TCP-1/cpn60 chaperonin family; This family includes members from the HSP60 chaperone family ...	64-565	1.87e-83

► DOMÍNIOS PROTEICOS

+	Name	Accession	Description	Interval	E-value
[-]	GroEL	cd03344	GroEL like type I chaperonin. Chaperonins are involved in productive folding of proteins. They ...	45-565	0e+00

Pssm-ID: 239460 Cd Length: 520 Bit Score: 804.75 E-value: 0e+00

gi|1973640993 444 FAAIEEGIVPGGGAAFVHL SACPVAIKdKLEDADEKLGADIIQKALVAPASLIAQ NAGIEGEVVEKVKAREWEIGYNAM 523
Cdd:cd03344 401 RAAVEEGIVPGGGAALLRASPAIDKL--KALNGDEKLGIEIVRRAL EAPLROIAENAGVDGSEVVEKVLSPDGFGYDAA 478

BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

- E SE USARMOS A SEQUÊNCIA DE UM GENE PROCARIOTO PARA “BLASTAR” CONTRA O BANCO DE EUCARIOTOS ?

Bacillus thuringiensis strain T29 Cry1Ac gene, partial cds

GenBank: MK882923.1

[FASTA](#) [Graphics](#)

GenBank: MK882923.1

```
1 cctgggtcaa aaattgatat ttagtaaaat tagttgcact ttgtgcattt tttcataaga
61 tgagtcatat gttttaaaatt gtagtaatga aaaacagtat tatatcataa tgaattggta
121 tcttaataaaa agagatggag gtaacttatg gataacaatc cgaacatcaa tgaatgcatt
181 ccttataaatt gtttaagtaa ccctgaagta gaagtattag gtggagaaaag aatagaaact
241 ggttacaccc caatcgatat ttccttgtcg ctaacgcaat ttcttttgag tgaatttggt
301 cccggtgctg gatttgtggt aggactagtt gatataatat ggggaatttt tgggccctct
361 caatgggacg catttcttgt acaaattgaa cagttaatta accaaagaat agaagaattc
421 gctaggaacc aagccatttc tagattagaa ggactaagca atctttatca aatttacgca
481 gaatctttta gagagtggga agcagatcct actaatccag cattaagaga agagatgcgt
541 attcaattca atgacatgaa cagtgccctt acaaccgcta ttcctctttt tgcagttcaa
601 aattatcaag ttcctctttt atcagtatat gttcaagctg caaatttaca tttatcagtt
661 ttgagagatg tttcagtggt tggacaaaagg tggggatttg atgccgcgac tatcaatagt
721 cgttataatg atttaactag gcttattggc aactatacag attatgctgt acgctgggtac
781 aatacgggat tagaacgtgt atggggaccg gattctagag attgggtaag gtataatcaa
... ..
```

blastn

blastp

blastx

tblastn

tblastx

BLASTN programs search nucleotide databases using a nucleotide query. [more...](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#)

MK882923.1

Query subrange [?](#)

From

To

Or, upload file

Escolher arquivo

Nenhum arquivo escolhido [?](#)

Job Title

MK882923:Bacillus thuringiensis strain T29...

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database

☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus

New

☐ Experimental databases



Try experimental taxonomic nt databases

For more info see [What are taxonomic nt databases?](#)

[Download](#)

Nucleotide collection (nr/nt) [?](#)

Organism

Optional

Eucarya (taxid:2759)

☐ exclude

[Add organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

**O QUE VOCÊ ESPERARIA
COMO RESULTADO?**



BANCO DE DADOS BIOLÓGICOS (BDB) – NCBI - BLAST

➤ RESULTADO BLASTn

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

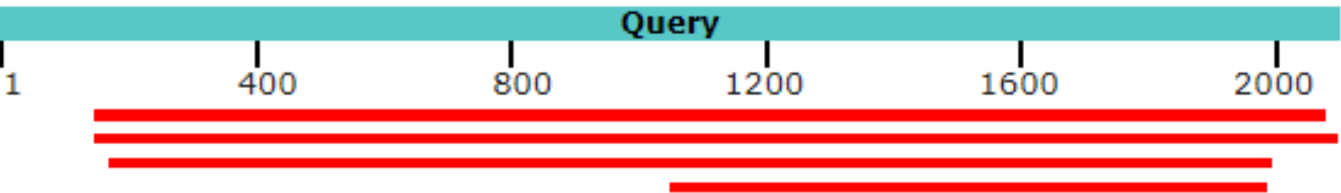
DownloadSelect columnsShow100?

☒ select all 4 sequences selected

[GenBank](#)[Graphics](#)[Distance tree of results](#)[MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒	Gossypium hirsutum transgenic cultivar GK-12 insect-resistant gene cassette, complete sequence	Gossypium hirs...	1537	1537	91%	0.0	81.07%	2471	GU583854.1
☒	Gossypium hirsutum transgenic cultivar Xinmian 33B insect-resistant gene cassette, complete sequence	Gossypium hirs...	1522	1522	92%	0.0	80.79%	4074	GU583853.1
☒	Oryza sativa Indica Group transgenic insect-resistant line TT51-1 integrated transgenes and flanking g...	Oryza sativa In...	1428	2857	86%	0.0	80.88%	10536	EU880444.1
☒	Oryza sativa Indica Group transgenic isolate KF6R1 cultivar Kefeng6 genomic sequence	Oryza sativa In...	944	944	44%	0.0	84.80%	1867	HQ161057.1

Distribution of the top 5 Blast Hits on 4 subject sequences



ESTUDO DIRIGIDO

1. O QUE SÃO BANCO DE DADOS BIOLÓGICOS?
2. A RELAÇÃO ENTRE BIOINFORMÁTICA E OS PROJETOS DE SEQUENCIAMENTO?
3. FUNÇÕES DO NCBI?
4. COMO A FERRAMENTA *BLAST* FUNCIONA?

