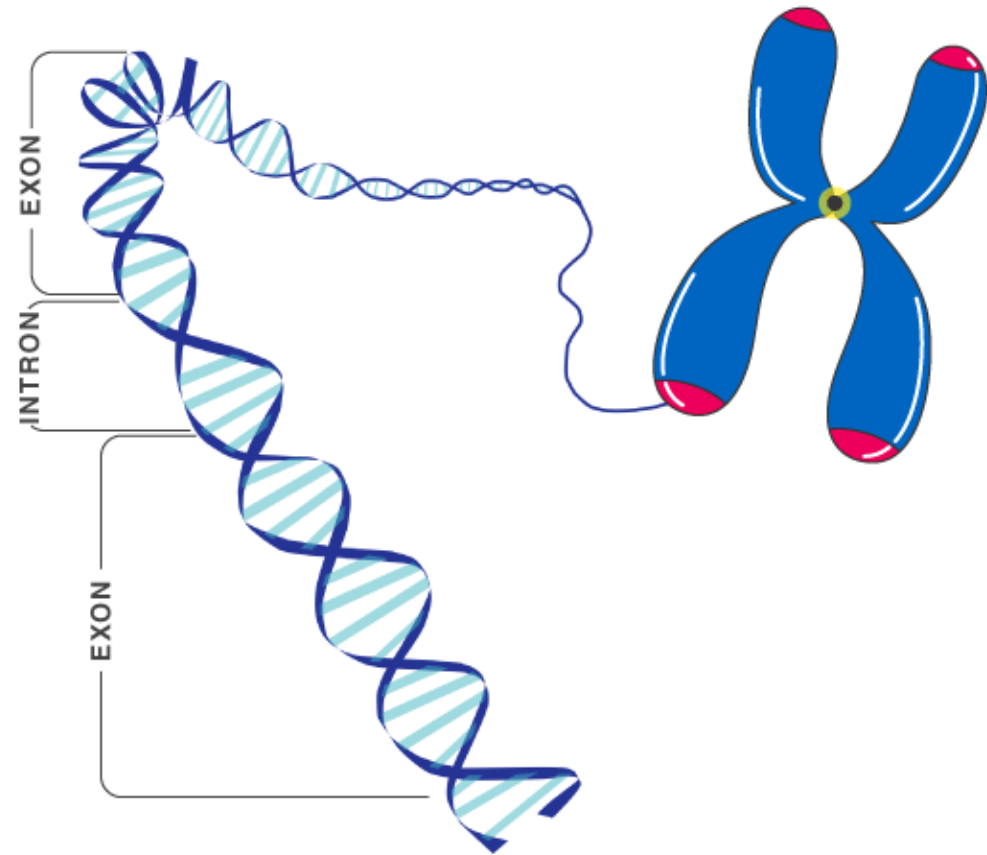
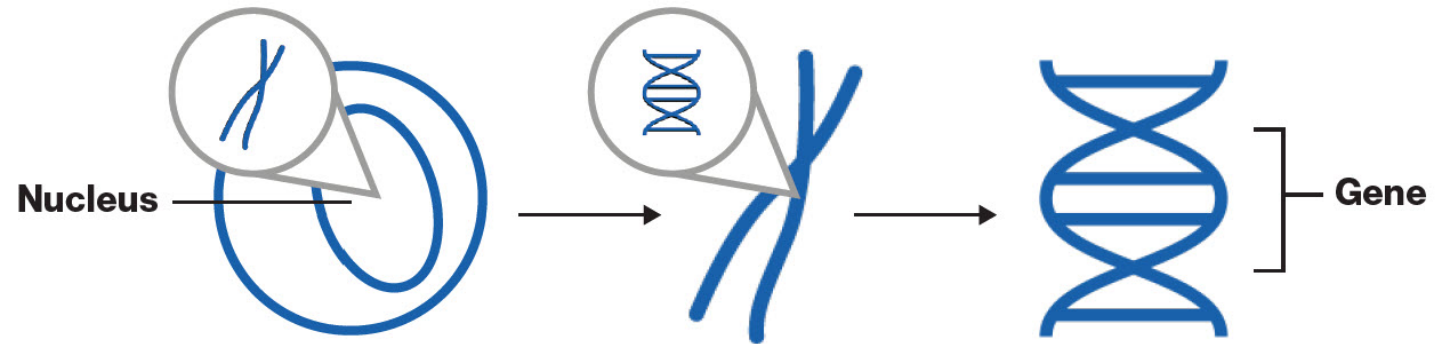


Gene Structure Visualization

Assoc. Prof. Dr. Ilker BUYUK



-
- The **gene** is the functional unit of Heredity.
 - Each gene is a segment of DNA that give rise to a protein product or RNA.
 - A gene may exist in alternative forms called **alleles**.
 - Chromosome in fact carry genes.
 - Each chromosome consists of a linear array of genes.



Cell

The nucleus controls the processes of the cell

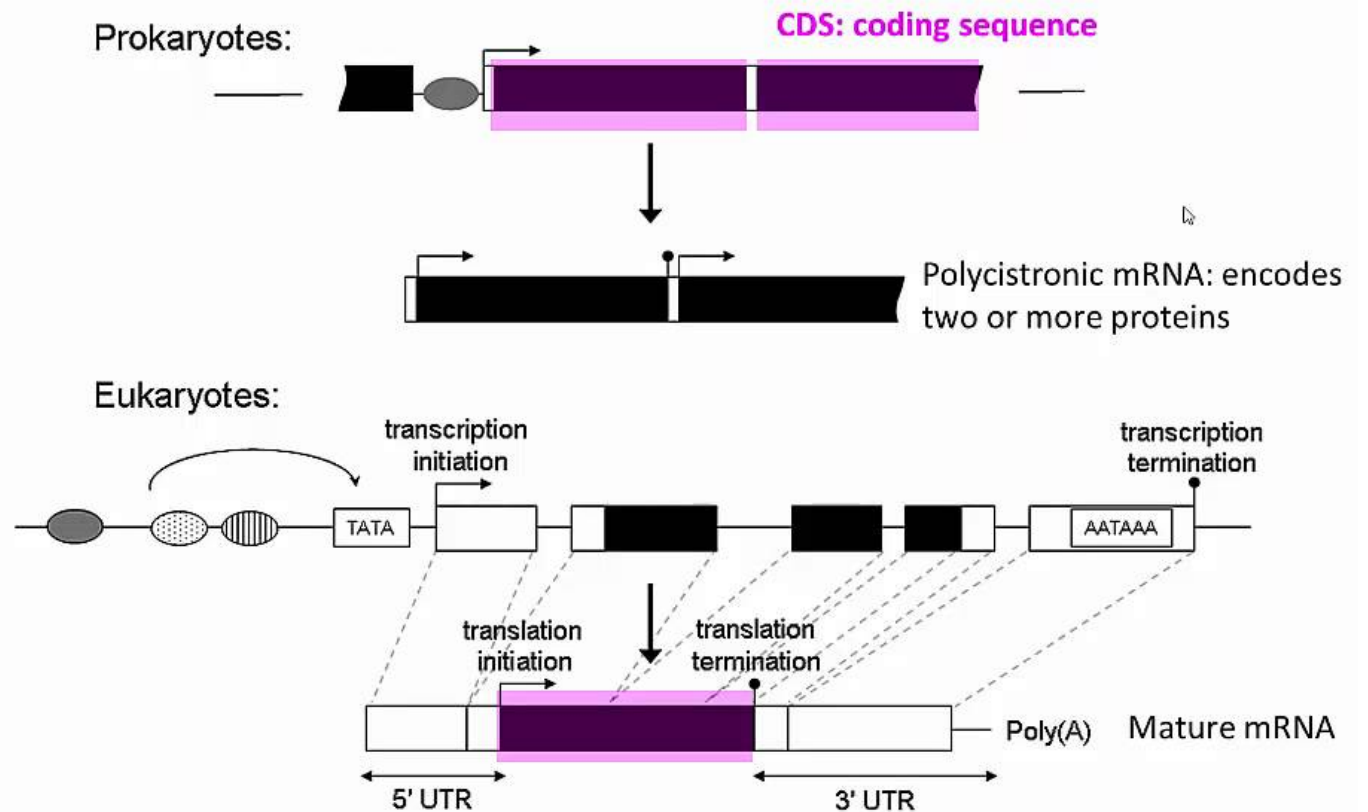
Chromosome

Chromosomes are thread-like structures made up of DNA tightly coiled many times around proteins called histones

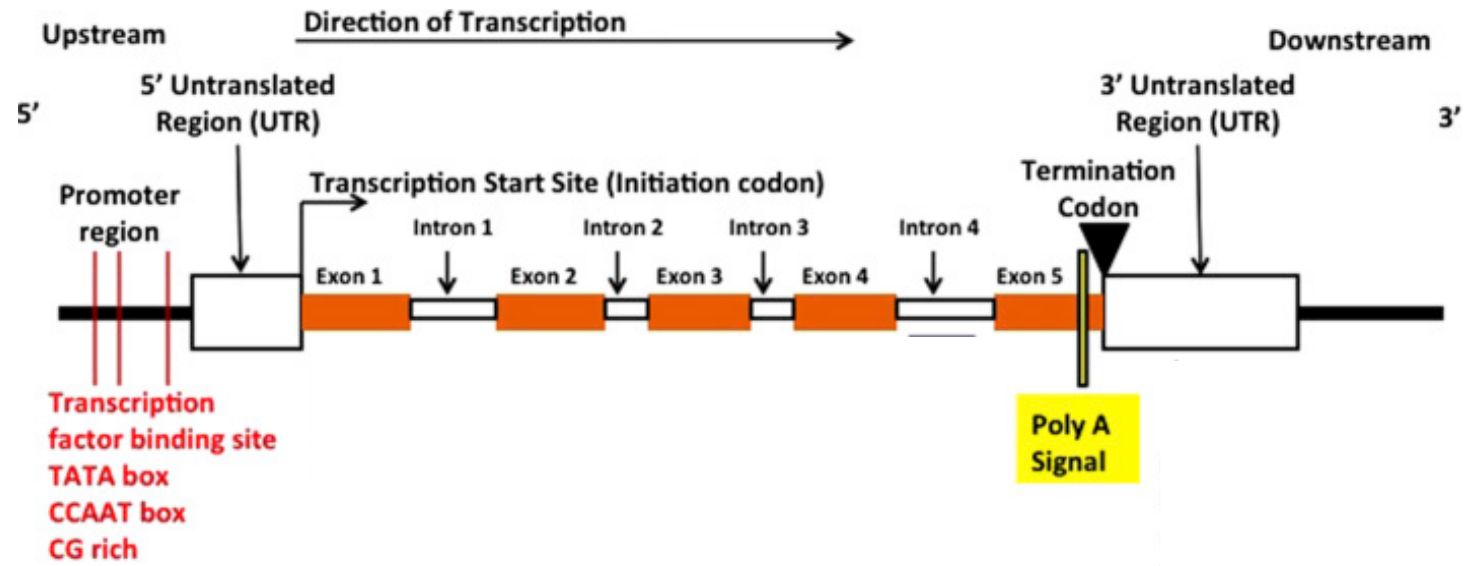
DNA

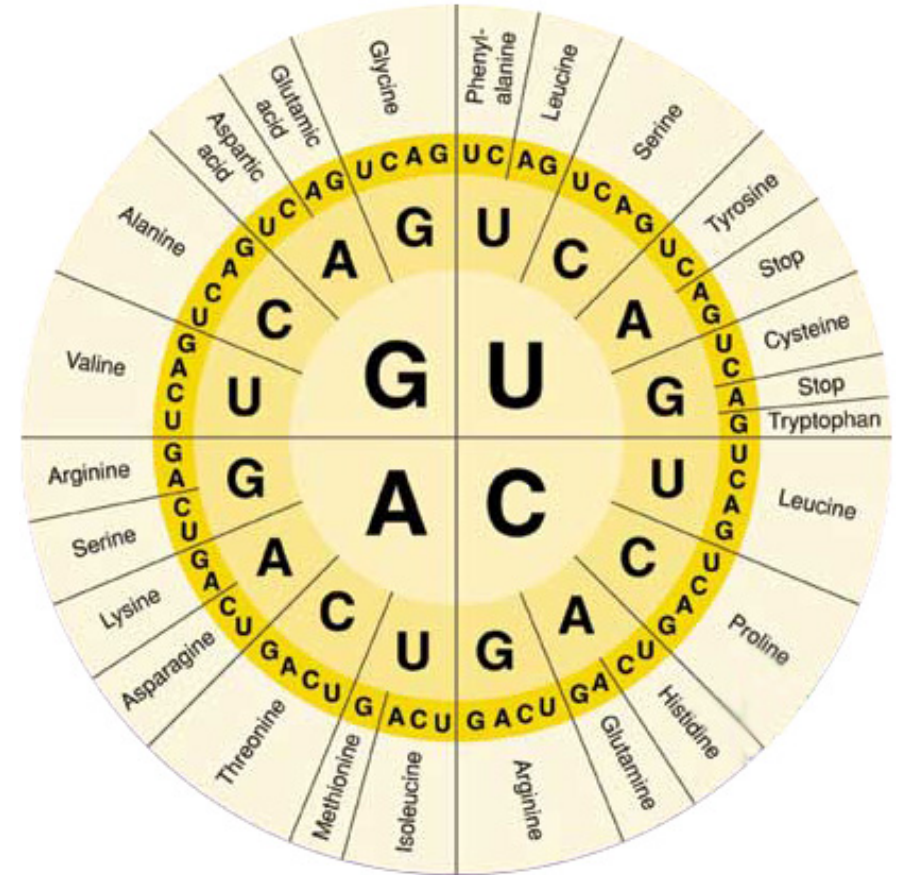
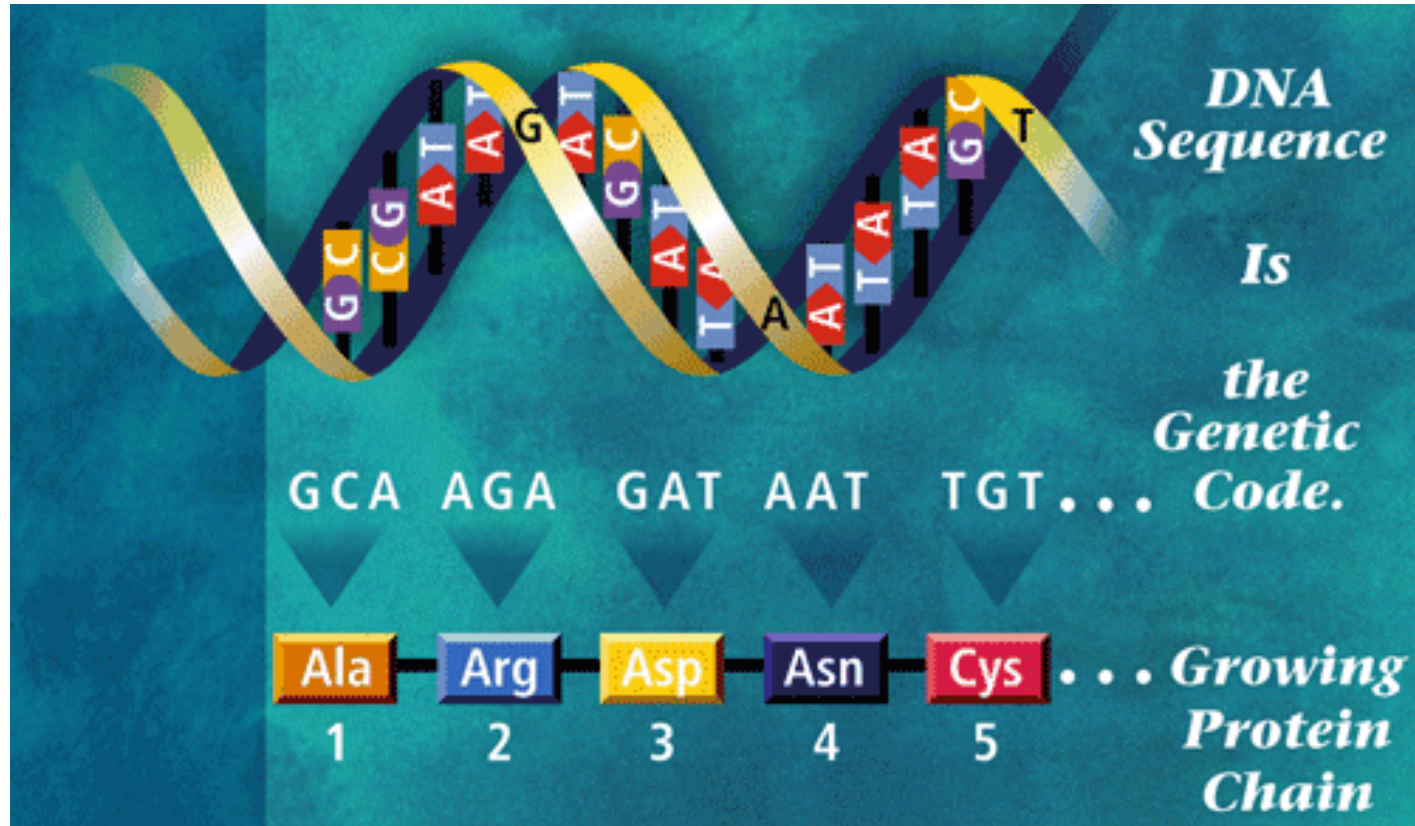
- A gene is a specific sequence of DNA containing genetic information required to make a specific protein
- Prokaryotic gene is **un-interrupted**.
- In Eukaryotic gene the coding sequences (exon) are separated by non-coding sequences called **introns**.
- In complex eukaryotes, introns account for more than 10 times as much DNA as exons.

Gene Structure

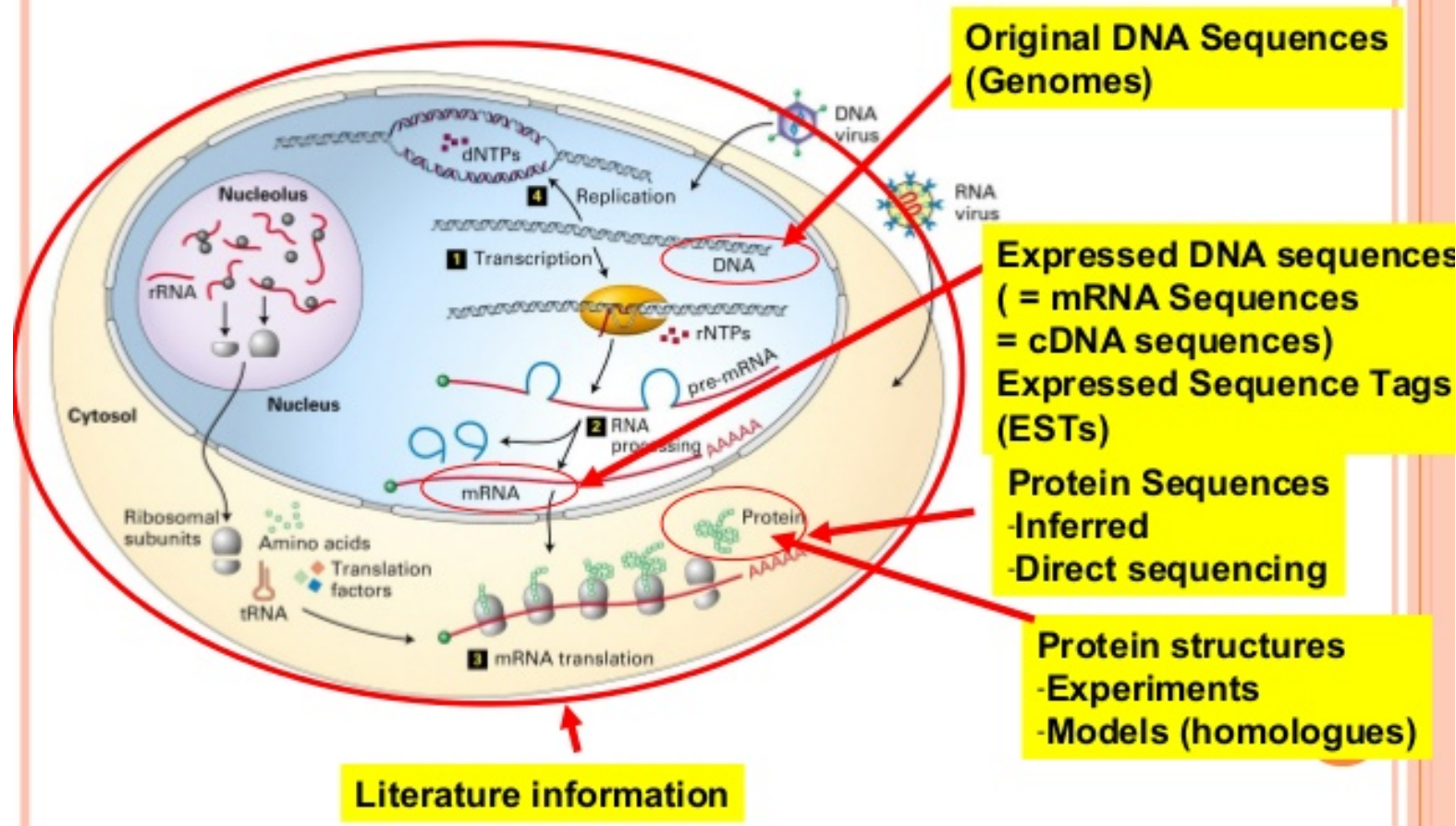


-
- **EXONS** coding sequence, transcribed and translated.
 - **INTRONS** coding sequences are separated by non-coding sequences called introns.





THE CENTRAL DOGMA & BIOLOGICAL DATA



FASTA

- In bioinformatics and biochemistry, the **FASTA format** is a text-based format for representing either nucleotide sequences or amino acid (protein) sequences, in which nucleotides or amino acids are represented using single-letter codes.
- The format also allows for sequence names and comments to precede the sequences.
- The format originates from the FASTA software package, but has now become a near universal standard in the field of bioinformatics.

Agave sisalana HSP70 (HSP70) mRNA, complete cds

GenBank: MH555341.1

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

```
>MH555341.1:55-2250 Agave sisalana HSP70 (HSP70) mRNA, complete cds
ATGGTCTCTCGAATCTGAAGGAGATTGCTGGGAAGAATTTGAACGCTGCGGTTGTTGATTGTCATCG
GGATTCCGGTGTACTTCACTGATTTGCAACGGCGAGCTGTGATGGATGCTGCGAGTATTGCGGGATTGCA
TCCGTTGAGGTTGTTTACGAGACCACAGCCACTGCTTTGGCTTATGGGATCTATAAGACCGATTTGCC
GAGAACGATCAGGTGAATGTGGCCTTTGTTGATGTTGGTCATGCTAGCATGCAGGTGTGCATTGCTGGGT
TCAAGAAGGGCCAGTTGAAGATTTTGGCCCATGCATTCGATCGGTCACTCGGTGGTTCGAGATTTTGTGA
AGCTCTGTTCAAGTATTTTGTGCTAAGTTTAAAGGAGGAGTACAAGATTGATGTGTACCAAATGCGAGG
GCTTGCCCTTAGGCTGCGTGCAGCTTGTGAGAAGTTGAAGAAGGTGCTGAGTGCAAACCTGAGGCACCGT
TGAATATAGAGTGTGGATGGATGAGAAAGATGTTAAGGGGTTTATCAAGAGGGAGGAGTTTGAATAT
TAGTGTGCCAATATTGAAAGAGTTAAAGGGCCTCTGGAGAAGGCCCTTATTGATGCAGGATTGAGTGTG
GAGAATATTCCTCTGTGGAGGTTGTTGGGTCTGGATCGAGGGTCCCGGCTATTATTAGGATATTGACGG
AGTTCTTTGGGAAGGAACCGAGAAGGACAATGAATGCGAGTGAATGTGTGGCTCGTGGGTGTGCTCTTCA
GTGCGCCATCCTCAGCCCACATTCAAAGTGCAGGATTTTCAAGTGCATGAAAGTTTCCCTTTCCAATT
GCCCTTTCATGGAAAGTTCTGCCCCAGACTCACAGAGTGCAGCTGAAAATCATCAGAGCTCAATTGTAT
TTCCCAAGGGAAATCCAATTCCAAGTGTAAAGCTCTGACATTCCTCAGATCTAGTACATTCTCAATTGA
TGTCATGTATGCTGATGCAAGCGAATTGCAAGTGTATCACCCAAGATTAGCACATATACGATTGGTCCC
TTTCAATCTGCTAAAGGTGATCGAGCAAAATTGAAAGTGAAGTCCGCTTGAATGTCCATGGAATTGTTT
CTGTTGAATCAGCAACTATGTTAGAGGAGGAAGAAGTTGAAGTTCAGTAACAGCTGGAAAGGAAGCACC
TAAAGAAGCTGCGAAAATGGAGACAGATGAAGCAAAAACCGACTCGTATGCGGCTGGAAGTGTGCAAC
ATGCAGGATGCTCCTGGTACCAGGGTTCGAGAATGGTGCAGCTGAGTCTGAAGATAAGCCTGTTTCAAGTGG
AAACTGATACCAAGGTTGAACCTTTGAAAAGGAAGGTCAAGAAGACTAATGTTCCAGTTGTTGAATTGGT
CTATGGAGGCCTTGCAGCTGCAGATTTGCAGAAAGCGGTGGAAAAGGAGTTTGAATGGCTCTTCAAGAC
```

FASTA alphabet

- Sequences are expected to be represented in the **standard amino acid** and **nucleic acid codes**, with these exceptions: lower-case letters are accepted and are mapped into upper-case; a single hyphen or dash can be used to represent a gap of indeterminate length; and in amino acid sequences, U and * are acceptable letters (see below). Before submitting a request, any numerical digits in the query sequence should either be removed or replaced by appropriate letter codes (e.g., N for unknown nucleic acid residue or X for unknown amino acid residue). The nucleic acid codes supported are:

A	adenosine	C	cytidine	G	guanine
T	thymidine	N	A/G/C/T (any)	U	uridine
<i>K</i>	<i>G/T (keto)</i>	<i>S</i>	<i>G/C (strong)</i>	<i>Y</i>	<i>T/C (pyrimidine)</i>
<i>M</i>	<i>A/C (amino)</i>	<i>W</i>	<i>A/T (weak)</i>	<i>R</i>	<i>G/A (purine)</i>
<i>B</i>	<i>G/T/C</i>	<i>D</i>	<i>G/A/T</i>	<i>H</i>	<i>A/C/T</i>
<i>V</i>	<i>G/C/A</i>	-	<i>gap of indeterminate length</i>		

Nucleotide

bZIP

Search



COVID-19 is an emerging, rapidly evolving situation.

Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS)

- 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

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



Develop

Download

Transfer NCBI data to your computer



Analyze

Learn

Find help documents, attend a class or watch a tutorial



Research

Popular Resources

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

NCBI News & Blog

Conserved Domain Database version 3.19 is available!

Items: 1 to 20 of 27051

<< First < Prev Page 1 of 1353 Next > Last >>

- 05) [Cicer arietinum bZip \(bZIP\) mRNA, complete cds](#)
1. 1,396 bp linear mRNA
Accession: KC464463.1 GI: 520993719
[Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- i1) [Oryza sativa bZIP \(bZIP\) mRNA, complete cds](#)
2. 2,575 bp linear mRNA
Accession: AF268596.3 GI: 13124870
[Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Chrysanthemum x morifolium bZIP \(bZIP\) mRNA, complete cds](#)
3. 867 bp linear mRNA
Accession: MK986728.1 GI: 1786006924
[Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Phaseolus acutifolius bZIP mRNA, complete cds](#)
4. 662 bp linear mRNA
Accession: AY026054.1 GI: 12829955
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Paeonia lactiflora cultivar Karl Rosenfield bZIP mRNA, complete cds](#)

Oryza sativa (279)
Neisseria meningitidis (624)
Arabidopsis thaliana (430)
Oryza sativa (279)
Zea mays (260)
All other taxa (24816)
More...

Find related data

Database:

Search details

bZIP[All Fields]

Recent activity

bZIP (27051)

Oryza sativa bZIP (bZIP) cds

GenBank

Oryza sativa bZIP (bZIP) mRNA, complete cds

GenBank: AF268596.3

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

Go to:

LOCUS AF268596 2575 bp mRNA linear PLN 26-FEB-2001
DEFINITION Oryza sativa bZIP (bZIP) mRNA, complete cds.
ACCESSION AF268596
VERSION AF268596.3
KEYWORDS .
SOURCE Oryza sativa Indica Group (long-grained rice)
ORGANISM [Oryza sativa Indica Group](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; BOP
clade; Oryzoideae; Oryzeae; Oryzinae; Oryza; Oryza sativa.
REFERENCE 1 (bases 1 to 2575)
AUTHORS Bi,X., Fu,B. and Song,Y.
TITLE Cloning and characterization of a novel rice bZIP-like protein cDNA
from ovule
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2575)
AUTHORS Bi,X., Fu,B. and Song,Y.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2000) College of Life Science, Developmental

Oryza sativa bZIP (bZIP) mRNA, complete cds

GenBank: [AF268596.3](#)

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

>AF268596.3 Oryza sativa bZIP (bZIP) mRNA, complete cds
ATGGGGTGCACGGCGTCAAGGTGGAGCAGGAGGACACGGTGCGGCGGT
TGAAGGAGGCGGTGGCGTCGCGGCAGCAGCTGGCGTCGGCGCACGCCGAC
CACCGCCCGCGCTCTCCCGCTTCGCGCAGGGCCACCCGTCGCTCGCCGTGTCGCACCACACCGCGCCG
GTGCTCCTCACCACGGCCGCGCCCGCGCTGGCGCCGACGCCGACGCCCGCCGCGCCGTCATCCACGGCGT
CGTCCTCGCTCCCGCCACCGACGCCGCTGCTCCCCAAGCACCAGCAGGCGCCGCGCCGCCACCGCCAC
GCAGTCGCATCAGCCGCCTCTCCCGTGGCGGTGAGGGCTCCCCGCGCGGGCCGCGTGCCTCAAGGTG
CCGCACATCCTGTCCGACTCCAGCGTCGCCAGCCCGCGCGGTCTGTTCCGGAAGCCGGTGGTGGGGA
CGCCGTGCTGTCGTGGCGTGGGACTGGGAGAATCTACCCGCGCTCGCCGCCGGACTCCGAGTTCTT
CGACCGCCGCAAGGCCGACCTCGAGGAGGCCAACCGCCTCCGCGAGCTCGAGGAGGAGGAGAAGGCCGG
GGCTACCTCCACCCCAACCTCAAGGAAGAGGACGAGGTCGACGACGACGACGACGAGAGGGAGGAGG
AGATGCATTGCGGCGGATGGGAGGACGACGACGACCACTACGCGTCGACGACCACTCGGAGACCAGATC
GGAGGAGGGGAAATGGGGAACAGATCGGAGTGGCGGCTTCGCGGCCAGATCGGAGTACGGCGGCACGGCG
CCGTCGGAGTACGCCGCCGCGCCGCTGCCACTGCCGCTGAGGAGGAGGGACGAGAGGTCGGAGGCCGGGG
ACTCCTCCTCCACGGTCACGGCGGCCGCGGAGATGCGGATGGTGATCCGCCACCGCACGCTGGCGGAGAT
CGTGGCCGCATCGAGGAGTACTTTGTCAAGGCGGCCGAGGCCGGCAATGGCGTCTCGGAGCTCCTGGAG
GCTAGCCGCGCGCAGCTGGACCCAACTCCGGCAGCTCAAAAAGACGGTGTACCACTCGAACAGCTTGC
TATCGTCGCTGTCGTGACATGGACTTCAAAGCCACCATGGCTGTTGCTACAAGTTGGACACCAATGC
GTTAGAGATGGAGTCAATGGAAGGGAAGAGCCATGGGTCGACACTGGAGCGTCTTTTGGCCTGGGAGAAA
AAGCTCTATCAGGAGGTCAAGGCTAGAGAGAGCGTTAAGATTGAGCACGAGAAGAAGCTTTTACTCTGC
AGAGCCTGGAGTACAGAGGGAGGGATAGTACCAAGCTGGATAAGACCAAGGCCCTCCATAAAACAAGCTGCA
ATCGTTGATCATCGTGACTTACAGGCCGCAACTACCACATCCTCAGCCATTGTCAGGGTGCCTGACAAT
GAGCTTGACACAGCTTGTGAGCTTTGCTTCGCGCTGTTGAGCATGTGGAGATCAATGAACCATTTC
ATGAGATCCAGAATGAAATTGTTACAGCAAGTCCGTGGTCTGGTAGACAATTCATGGCTGAGTCAACATC
TGATCTTACCCGGCTTGCCACCCGTGATCTTGAGGCTGCTGCTCAGCATGGCACTCAAACCTCAACCGT
CTCATCAAGTATCAACGTGATTATACGTGCCCTCTATGGCTGGCTGAAGCTCACACTCTTCCAAGTGG

Search NCBI

AF268596.3



Search

Results found in 2 databases

NUCLEOTIDE SEQUENCE

Oryza sativa bZIP (bZIP) mRNA, complete cds

Oryza sativa Indica Group

2,575 bp mRNA sequence

AF268596.3

FASTA

BLAST

Download

Literature

Bookshelf

0

MeSH

0

NLM Catalog

0

Genes

Gene

0

GEO DataSets

0

GEO Profiles

0

Proteins

Conserved Domains

0

Identical Protein Groups

1

Protein

0

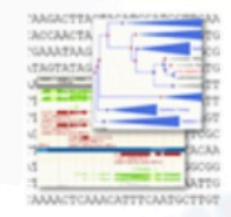


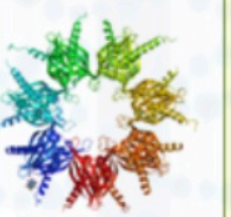



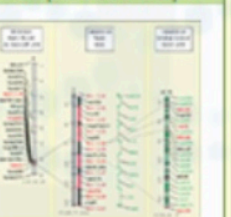
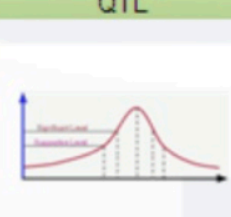





CROP DATABASES

1. Gramene

Explore Gramene

Quick Search

Genomes 	Genetic Diversity 	Pathways 	Proteins 
Genes 	Ontologies 	Markers 	Comparative Maps 
QTL 	BLAST 	Gramene Mart 	Species Pages 

<http://www.gramene.org/>

3. Phytozome

phytozome 10 [Species](#) [Tools](#) [Info](#) [Download](#) [Help](#) [Log In](#)  

Welcome to phytozome The JGI Comparative Plant Genomics Portal

Phytozome quick search (advanced)

Search a JGI flagship genome



Target: Query:

Search version 10 genomes and families

Search an Early Release genome

Help with Phytozome

Documentation

- View a tree representation of the species in Phytozome v10

Video tutorials

- Video tutorials will be available shortly

About Phytozome v10.0.4

Phytozome is the Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute. Families of related genes representing the modern descendants of ancestral genes are constructed at key phylogenetic nodes. These families allow easy access to clade-specific orthology/paralogy relationships as well as insights into clade-specific novelties and expansions. As of release v10, Phytozome provides access to forty-eight sequenced and annotated green plant genomes which have been clustered into gene families at 12 evolutionarily significant nodes. Each gene has been annotated with PFAH, KOG, KEGG, PANTHER and GO assignments, where possible. Query-based data access is provided by Phytozome's InterMine and BioMart instances, while bulk data sets can be accessed via the JGI's Genome Portal (registration required). JBrowse genome browsers are available for all genomes.

News (details...)

- (2014-11-13) **Improved Eucalyptus grandis assembly**
- (2014-09-11) **v10.0.4 update released**
- (2014-08-13) **v10.0.3 and willow genome released**

System Status (2014-12-10 02:03)

- ✓ Search
- ✓ BLAST
- ✓ BLAT
- ✓ InterMine
- ✓ Database

<http://www.phytozome.jgi.doe.gov/pz/>

19

4. GreenPhylDB

The screenshot shows the GreenPhyl v4 website interface. At the top, there is a search bar with the text "Enter your text to search here" and a "Search" button. Below the search bar is a navigation menu with links for "Home", "Search", "Gene Family lists", "Toolbox", "Documentation", "Statistics", and "My Space".

A note states: "Note: GreenPhyl v3 is still available at this URL: <http://www.greenphylogiv3/>".

The main content area contains a description of GreenPhylDB: "GreenPhylDB is a web resource designed for comparative and functional genomics in plants. The database contains a catalogue of gene families based on gene predictions of genomes, covering a broad taxonomy of green plants. Result of our automatic clustering is manually annotated and analyzed by a phylogenetic-based approach to predict homologous relationships. It supports evolution and functional studies to identify candidate gene affecting agronomic traits in crops."

To the right of the description is an "Information" box with the following links:

- [Release Notes of version 4](#)
- [Release Notes of version 3](#)
- [How to cite us](#)

Below the text is a phylogenetic tree showing relationships between various plant species. The tree is rooted on the left and branches out to the right. The species names are listed on the right side of the tree, with some names in different colors. The tree is labeled with "D1" and "D2" at various nodes.

The species names listed are:

- Cyanobacterium merolae
- Chlamydomonas reinhardtii
- Ostreococcus tauri
- Physocostis perfoliata
- Salvinella selaginoides
- Picea abies
- Arabidopsis thaliana
- Musa acuminata
- Musa balbisiana
- Eleusine indica
- Pharus dactyloides
- Oryza sativa
- Brachypodium distachyon
- Hordium vulgare
- Setaria italica
- Zea mays
- Sorghum bicolor

<http://www.greenphyl.org/cgi-bin/index.cgi/>

5. CoGE

Organisms: 16,962 Genomes: 23,442 Features: 457,831,491 Annotations: 643,045,640 Experiments: 3,339 (27G values)

New to CoGe?

[Get started](#) [Create an Account](#) [Tutorials](#) [Documentation](#) [FAQ](#)

Tools



OrganismView

Search for organisms, get an overview of their genomic make-up, and visualize them using a dynamic, interactive genome browser.
[Example](#)



CoGeBlast

Blast sequences against any number of organisms in CoGe.
[Example](#)



SynMap

Compare any two genomes to identify regions of synteny.
[Example](#)



SynFind

Search CoGe's annotation database for homologs.
[Example](#)



GEvo

Compare sequences and genomic regions to discover patterns of genome evolution.
[Example](#)

What do you want to do?

[Compare two genomes](#)

[Browse a genome](#)

[Load a new genome](#)

[Load experimental data](#)

Latest News

[Brassicaceae: Dogs of the Plant World](#)
December 5th 2014

[40 New Fish Genomes now Available](#)
November 21st 2014

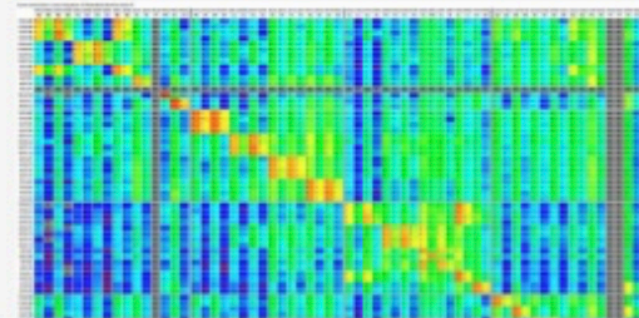
[Bug Fixes and Improved Stability](#)
November 12th 2014

[Tutorial for integrating genomes from JGI/Phytozome](#)
October 17th 2014

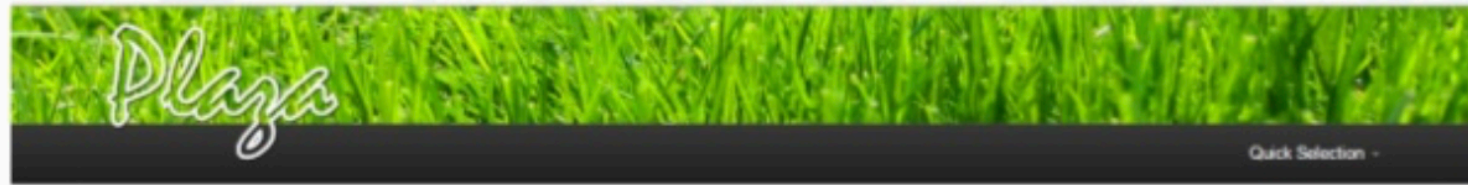
[Japanese eggplant genome now available](#)
October 17th 2014

[...more...](#)

Tutorials



6. Plaza



Available PLAZA versions *(select the most appropriate)*

PLAZA 3.0 Dicots (2014)



Summary

Latest iteration of the PLAZA platform. Focused on dicot species, but with some others as reference organisms.

This platform may still contain minor bugs. Do not hesitate to contact us if you find any!

PLAZA 3.0: an access point for plant comparative genomics

[Go to PLAZA 3.0 Dicots](#)

Species Included



PLAZA 3.0 Monocots (2014)



Summary

Latest iteration of the PLAZA platform. Focused on monocot species, but with some others as reference organisms.

This platform may still contain minor bugs. Do not hesitate to contact us if you find any!

PLAZA 3.0: an access point for plant comparative genomics

[Go to PLAZA 3.0 Monocots](#)

Species Included



Crop databases

Crop plant	Hyperlink
Rice	http://rgp.dna.affrc.go.jp/IRGSP/
Maize	http://www.maizegdb.org/
Sorghum	http://www.phytozome.net/sorghum.php
Wheat	http://www.wheatgenome.org/
Tomato	http://solgenomics.net/organism/Solanum_lycopersicum/genome
Potato	http://www.potatogenome.net/index.php/Main_Page
Rapeseed	http://www.brassica.info/info/reference/genome-sizes.php
Soybean	http://soybase.org/
Castor Bean	http://www.phytozome.net/ricinus.php
Flax	http://www.phytozome.net/soybean.php
Common Bean	http://www.phytozome.net/commonbean.php
Foxtail millet	http://foxtailmillet.genomics.org.cn/page/species/index.jsp
Cotton	http://www.cottontdb.org/wwwroot/cdbhome.php
Chick pea	http://www.icrisat.org/gt-bt/ICGGC/homepage.htm
Pigeon pea	http://gigadb.org/dataset/100028
Sunflower	https://www.sunflowergenome.org/

◀ Previous view

? Help with this page

Actions

↻ Revise query

🚀 Launch Jalview

👤 Find related ... ▾

+ Add to cart

🌿 Composite family

My Data (28)

🛒 View cart

+ Add to cart

📄 Upload user data

📄 Send to BioMart

📄 Send to PhytoMine

📄 Get from PhytoMine

Search for genes, families and sequences

1. Select a Target 1 species selected ✕

Target set: Phytozome 12.1 Pre-release species

Target type: Ancestor nodes Species

Phaseolus vulgaris v2.1

- *Boechera stricta* v1.2
- *Brassica oleracea capitata* v1.0
- *Brassica rapa* FPsc v1.3
- *Capsella grandiflora* v1.1
- *Capsella rubella* v1.0
- *Eutrema salsugineum* v1.0
- **Fabidae**
- *Cucumis sativus* v1.0
- *Fragaria vesca* v1.1
- *Glycine max* Wm82.a2.v1
- *Malus domestica* v1.0
- *Medicago truncatula* Mt4.0v1
- **Phaseolus vulgaris v2.1** ✓

2. Build your query GO

Search type: Keyword BLAST

Phvul.001G045400.1.p

⬆ ✕

Algorithm parameters

Add trailing wildcard:

Use family settings:

▼ Gene Info

Organism Phaseolus vulgaris**Transcript Name** Phvul.001G045400.1 (primary)**Location:** Chr01:3677580..3680026 reverse**Alias** Phvul.001G045400.v1.0 Phvul.001G045400.1.v1.0**Description** (1 of 16) PTHR27007:SF24 - L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IX.1-RELATED**Links** [B](#) [M](#)

Functional Annotation Genomic Sequences Protein Homologs Gene Ancestry Expression

Genomic sequence Transcript sequence CDS sequence Peptide sequence Show all

key: 5' UTR CDS 3' UTR

^ Genomic Sequence [2447]	BLAST this sequence at	Phytozome	NCBI
^ Transcript Sequence [2447]	BLAST this sequence at	Phytozome	NCBI
^ CDS Sequence [2004]	BLAST this sequence at	Phytozome	NCBI

Contact Disclaimer
Accessibility / Section 508 Statement

©1997-2017 The Regents of the University of California



Genomic sequence Transcript sequence CDS sequence Peptide sequence Show all

key: 5' UTR CDS 3' UTR

```
>Phvu1.001G045400.1
MLAISHLLLFIFFTTFSPIHSLFFNITNFDDPTSNISYOGDGRSTNGSIDLNKVSYYFRVGRALYSKPLRLWDPSSNVVTFD
VTRFTFSIDRVNSSETSYADGFAFYLAFLGYOIPPNSAGGTFALFNATTMSDLPONHVFAVEFDTF IGSTDPPMKHVGVDNSLTSVAFENFDIDNNLKG
MCHTLITYTASTOTL FVSWSFKGRPTTKDSNNSSLSYSIDLKILPEWVNI GFSASTGLYTEHNNVIYSWEFNSSLKDSSAENEGVKLNHGKSKLVLIVA
ILCPLVLLLVGASTFVVILIKRKRKDDCMLYDAGDDEIGPTSVKFDLDRGTIPRRFEYKELVDATNGFSDERRLGOGASGOVYKGVLSYLGRVVAIKRI
FADFENSERVF TNEVRIISRLIHKNLVOFIGWCHEEGEFLLIFEYMONGSLDTHLFGNKRMLEWHVRYKIALGVVTALHYLHEDAEOCVLHRDIKSANVL
```



Quick Start

[Search](#) | [Batch Query](#) | [BLAST](#) | [Download](#) | [Genome Browser](#) | [GO enrichment](#)



[Cucumber \(Chinese Long\) genome v3 | v2](#)

[Search](#) | [Batch Query](#) | [Genome Browser](#) | [Pathway](#) | [Download](#)

[Cucumber \(Gy14\) genome v2 | v1](#)

[Search](#) | [Batch Query](#) | [Genome Browser](#) | [Pathway](#) | [Download](#)

[Cucumber \(PI183967\) genome](#)

[Search](#) | [Batch Query](#) | [Genome Browser](#) | [Pathway](#) | [Download](#)

[Watermelon \(97103\) genome v2 | v1](#)

News

The wax gourd reference genome published in [Nature Communications](#) is now publicly available in the database. [Dec. 2019]

The manuscript describing the improved [watermelon '97103' genome assembly \(v2\)](#) and [genome resequencing of 414 accessions](#) representing the seven extant *Citrullus* species has been published in [Nature Genetics](#). [November, 2019]

SNPs called from GBS data of *Cucurbita* accessions, including *C. pepo*, *C. maxima* and *C. moschata* are made publicly available prior to publication under a [data release policy](#). [October, 2019]

SNPs called from GBS data of 1365 watermelon accessions are now freely available. The [paper](#) describing these data was published in [Plant Biotechnology Journal](#). [April, 2019]

[+ Gene sequence \(with intron\)](#)Legend: CDS exon

Hold the cursor over a type above to highlight its positions in the sequence below.

```
ATGGCGTTTCTTGGACTTCTCTTGGTGGGTCTTCTTCTATCTTTGGGATGTGATTCTGCTCTTGCCAACAACAATGGAGGTGGTTGGATTAGTGCCACGCCACAT
TCTACGGTGGGGGTGACGCCGCTGGTACAATGGGTATGTGAAAAGATTTATACTTGTGCATTTATTGAACGTACAACCGTAGCCTTTAACGACGGACTTTCTTTT
TGCAGGAGGGGCTTGTGGCTATGAAATCTGTACAGCCAGGGATATGGGACGAACACAGCAGCTCTGAGCACAGCTTTATTCAACAATGGACTGAGTTGCGGCCT
TGCTTTGAAATCAAATGTGTAATGACCCAAAATGGTGTCTTCAAAAGTCCATTGTGGTTACTGCCACAAACTTTTGTCCACCAACAATGCGCTCCCAACAACG
CCGGCGGCTGGTGTAAACCCTCCCAACACCATTTCGATCTCTCCAACCCGCTCTCGAACAAATCGCTGGTTACAAAGCCGGCATTGTCCCAGTGGCATAACAGAAG
GTACTGACTGATTTTGTAAAGTTAAAACACTATAACTGAAACAGAGACTGAGCAAATAACCAATGCAGGGTATCTTGTGAGAAGAAAGGAGGCATTAGATTACAA
TCAATGGTCATTCGACTTCAATTTAGTCCTTGAACCAACGTTGGTGGAGGTGGGGATGTTTCATGGAGTGTGGATTAAAGGGTCGAAAACAGGGTGGGAAGCAAT
GTCGAGGAATTGGGGACAGAATTGGCAGAGCAATTCATATCTGAATGGACAAAGCTTATCGTTCAAAGTGACCACCGGCGATGGCAGAACAGTGATATCCAACGAC
GTAGCGCCAGCTCAGTGGAGCTTTGGTCAAACCTTTAGTGGGTTGCAGTTCCGATGA
```

[+ mRNA sequence](#)[+ Coding sequence \(CDS\)](#)

```
ATGGCGTTTCTTGGACTTCTCTTGGTGGGTCTTCTTCTATCTTTGGGATGTGATTCTGCTCTTGCCAACAACAATGGAGGTGGTTGGATTAGTGCCACGCCACAT
TCTACGGTGGGGGTGACGCCGCTGGTACAATGGGAGGGGCTTGTGGCTATGAAATCTGTACAGCCAGGGATATGGGACGAACACAGCAGCTCTGAGCACAGCTTT
ATTCAACAATGGACTGAGTTGCGGCCTTGCTTTGAAATCAAATGTGTAATGACCCAAAATGGTGTCTTCAAAAGTCCATTGTGGTTACTGCCACAAACTTTTGT
CCACCAACAATGCGCTCCCAACAACGCCGGCGGCTGGTGTAAACCCTCCCAACACCATTTCGATCTCTCCAACCCGCTCTCGAACAAATCGCTGGTTACAAAG
CCGGCATTGTCCCAGTGGCATAACAGAAGGGTATCTTGTGAGAAGAAAGGAGGCATTAGATTACAAATCAATGGTCATTCGACTTCAATTTAGTCCTTGAACCA
CGTTGGTGGAGGTGGGGATGTTTCATGGAGTGTGGATTAAAGGGTCGAAAACAGGGTGGGAAGCAATGTCGAGGAATTGGGGACAGAATTGGCAGAGCAATTCAT
CTGAATGGACAAAGCTTATCGTTCAAAGTGACCACCGGCGATGGCAGAACAGTGATATCCAACGACGTAGCGCCAGCTCAGTGGAGCTTTGGTCAAACCTTTAGT
GGTTGCAGTTCCGATGA
```

[+ Protein sequence](#)

```
MAFLG L L L V G L L L S L G C D S A L A N N G G G W I S A H A T F Y G G G D A A G T M G G A C G Y G N L Y S Q Y G T N T A A L S T A L F N N G L S C G A C F E I K C V N D P K W C L Q K S I V V T A T N F C
P P N N A L P N N A G G W C N P P Q H H F D L S Q P V F E Q I A G Y K A G I V P V A Y R R V S C E K K G G I R F T I N G H S Y F N L V L V T N V G G G D V H G V W I K G S K T G W E A M S R N W G Q N W Q S N S Y
```

<http://gsds.gao-lab.org>

- The visualization of gene features such as composition and position of exons and introns for genes offers visual presentation for biologists to integrate annotation, and also helps them to produce high quality figures for publication.

GSDS_{2.0} Gene Structure Display Server

Home | Help | About | FAQ | Links: PlantRegMap

● Gene Features

Format: BED

Input: BED

Input data: Sequence(FASTA)

or upload file: Dosya seçilmedi

● Other Features to Display

● Output (Phylogenetic Tree/Order)

● Image Format: SVG

● Gene Features

Format

Please keep the sequence IDs consistent in the two fields.

➤ CDS sequence (FASTA)

Input data:

```
>Pvu1-SRS-1
ATGGCTGGGTTCTTCTCTCTAGGAGGAAGGCATAACAAAGCAGAACAAGAGGAGGATCAGAGAGA
GGACAACAACCCAACAACACCAACCACAACAACA
GTCAATTTTGTTCAGAAACGTTAACGAGGAGATCTACAACACCAACAAGGGCTTCGAGATATGG
CCGCAATCCTCGTACCACCACAACCTTCACCAACTA
CTACTCCTTCGGAGTGGGCCCTAGTCGCAGGAACAACGCCAACCAACAACAGCTCCTCCAACA
ACGTAAACGACGACGTCTCTGTCTCTTTCTCGGAT
```

Example

or upload file: Dosya seçilmedi

➤ Genomic sequence (FASTA)

Input data:

```
GGCATATTTCAGATGAGGAAAAAATGAAAGTGAATGATAGATTCGGCGTGGTGGATAGAAATGCT
CAGAACCTTTGACTGGGAGCACTTTATTAGTGTC
CAAATAAACACGCAAGTGACCTAGAAACAAATGAATTGATTAAGTATAGAGCTTGTCTTTTG
TGGTCGAAAAGTGG

>Pvu1-SRS-8
ATTTTCATGTTCTCCACGCACACATGACACACCACCACCCTCTCTCTCTCTCGCAGAAACCA
```

Example

● Gene Features

Format

Please keep the sequence IDs consistent in tl

➤ CDS sequence (FASTA)

Input data:

Example

or upload file: CDS.TXT

➤ Genomic sequence (FASTA)

Input data:

Example

or upload file: GENOMİK.TXT

● Other Features to Display

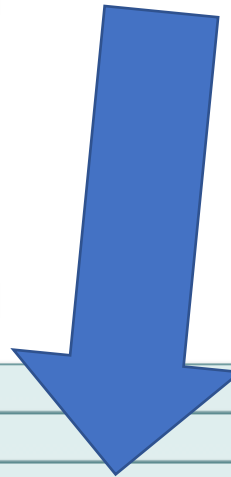
● Output (Phylogenetic Tree/Order)

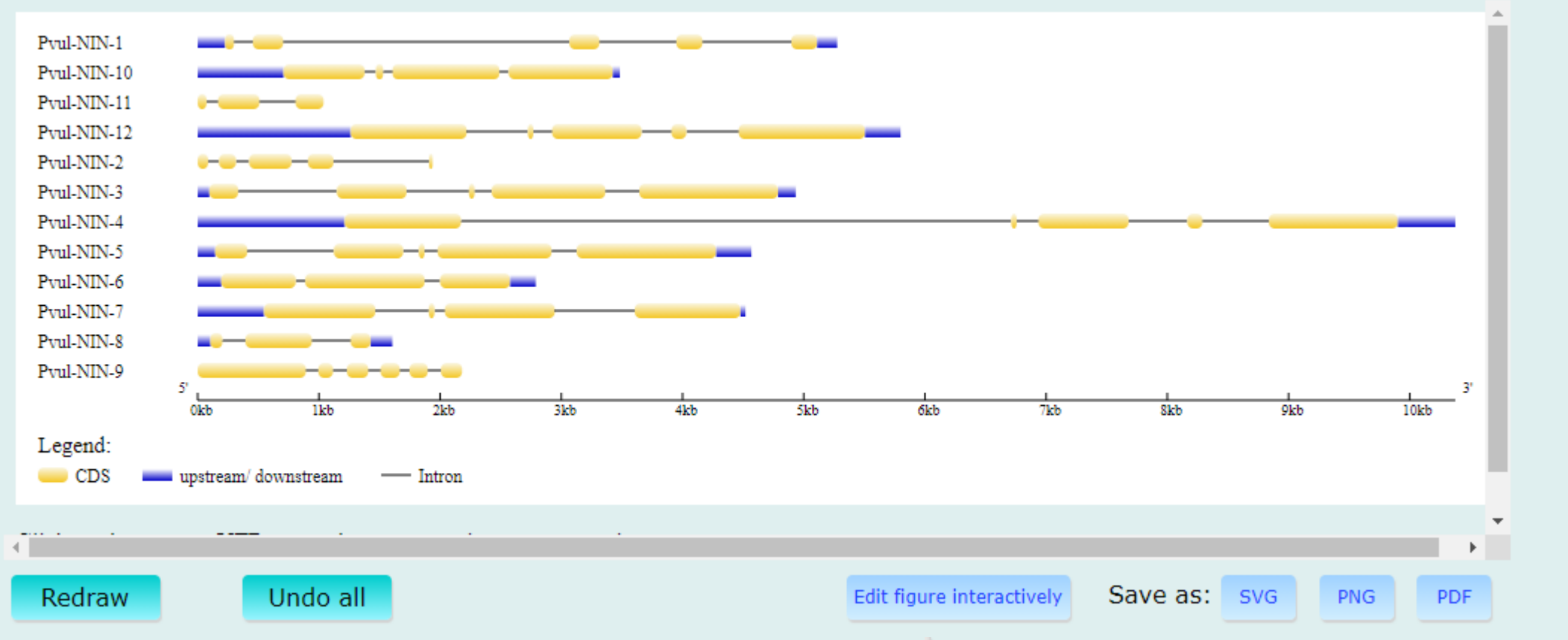
● Image Format

Reset

Submit

You job is running..., Please be patient...





Users can customize the size, shape and color of annotation features after their initial render and even fine-tune each element through an integrated visual editor.

Intron

Color Shape Line Line width px Intron rescale

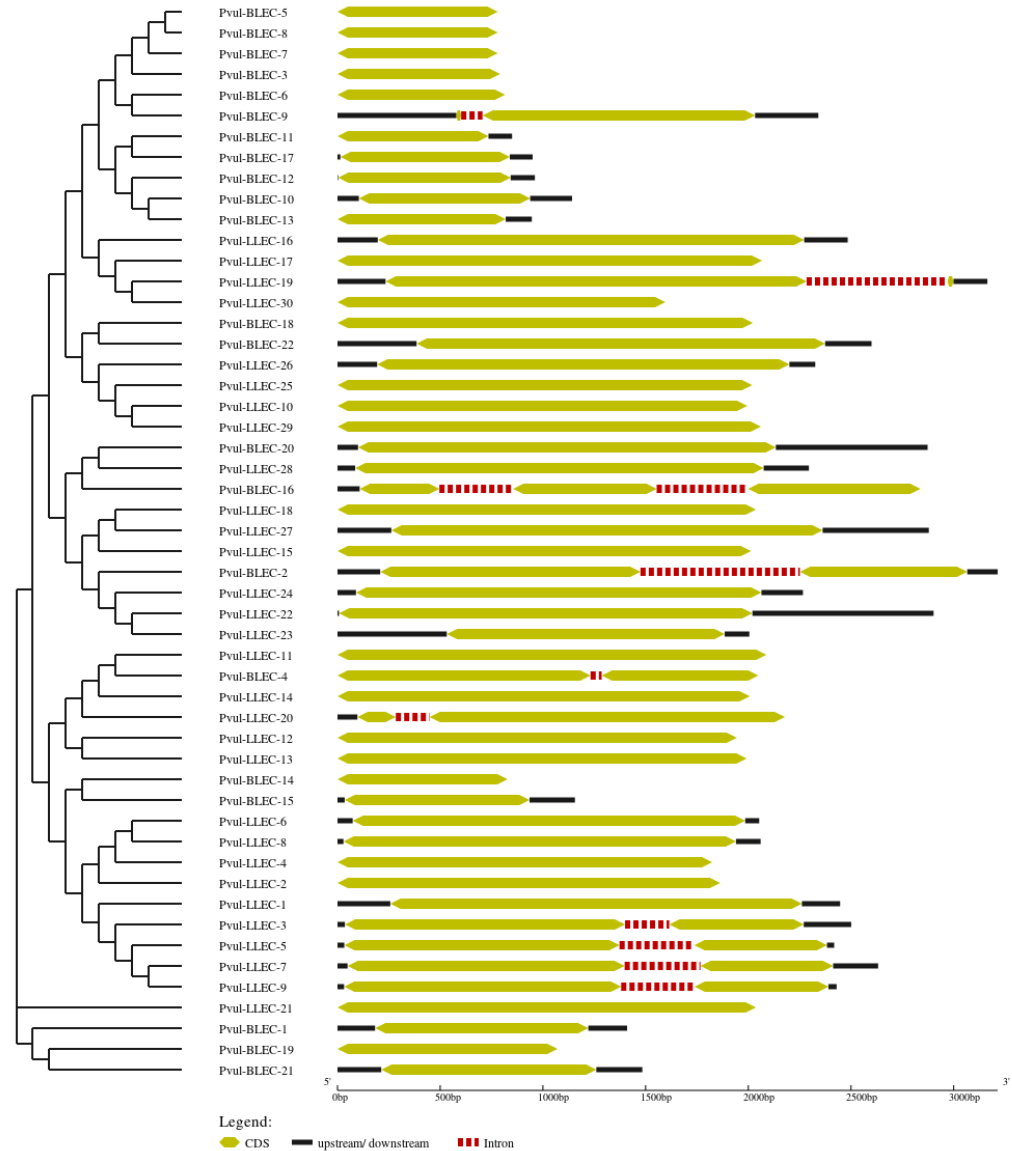
CDS

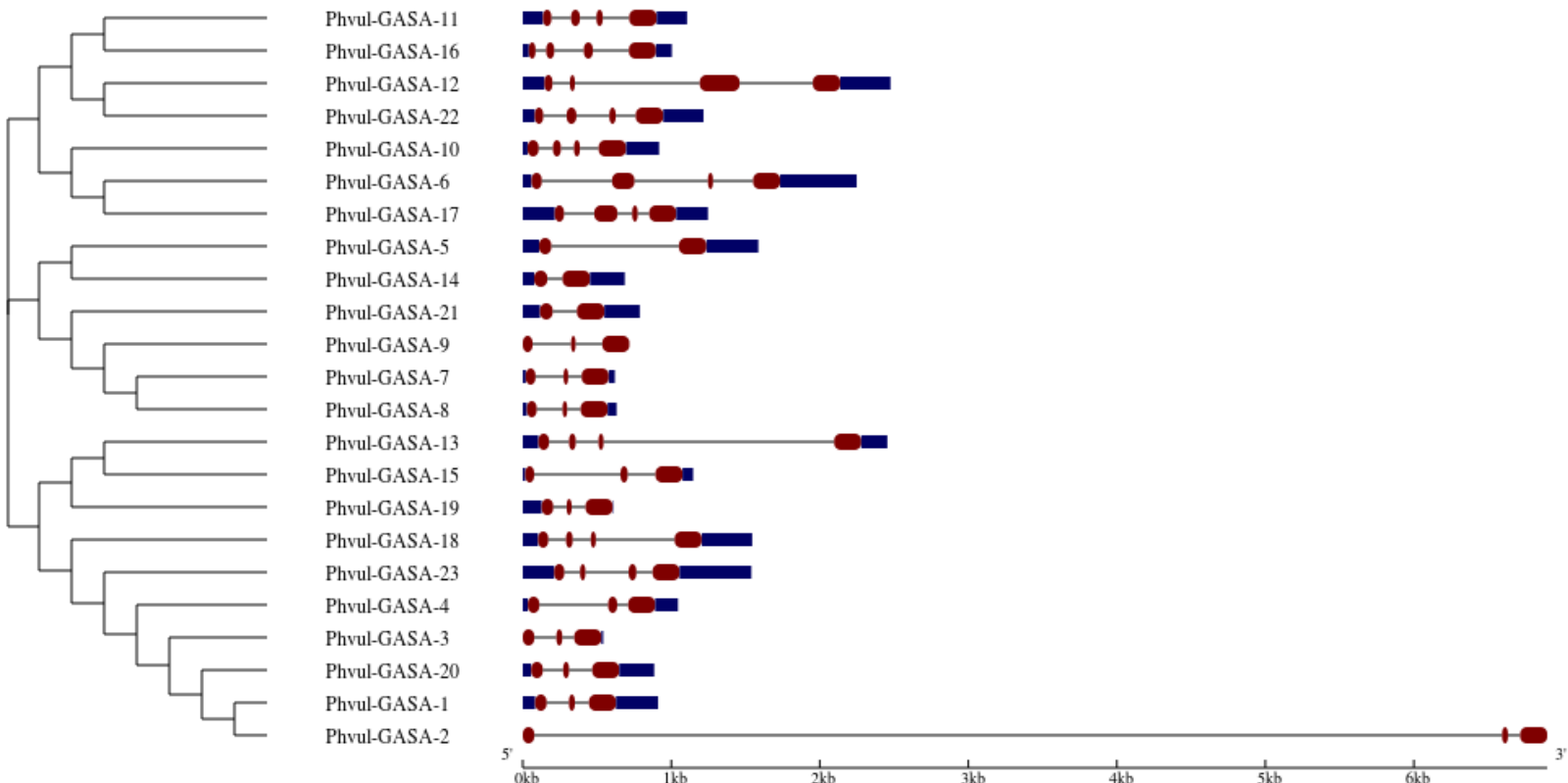
Color Shape Round-corner rect Filled Height px

UTR

Color Shape Rectangle Filled Height px

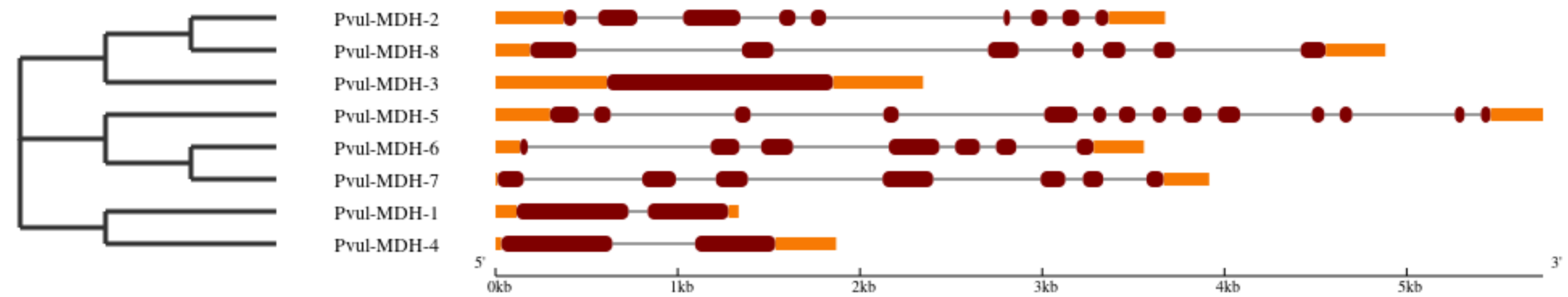
- To facilitate evolutionary analysis, a user-specified phylogenetic tree can be added to the figure. Finally, the generated figure can be exported as either vector graphic (in SVG and PDF format), or raster graphic (in PNG format).





Legend:

● CDS ■ upstream/ downstream — Intron



Legend:

● CDS ■ upstream/ downstream — Intron

Thanks for
listening

