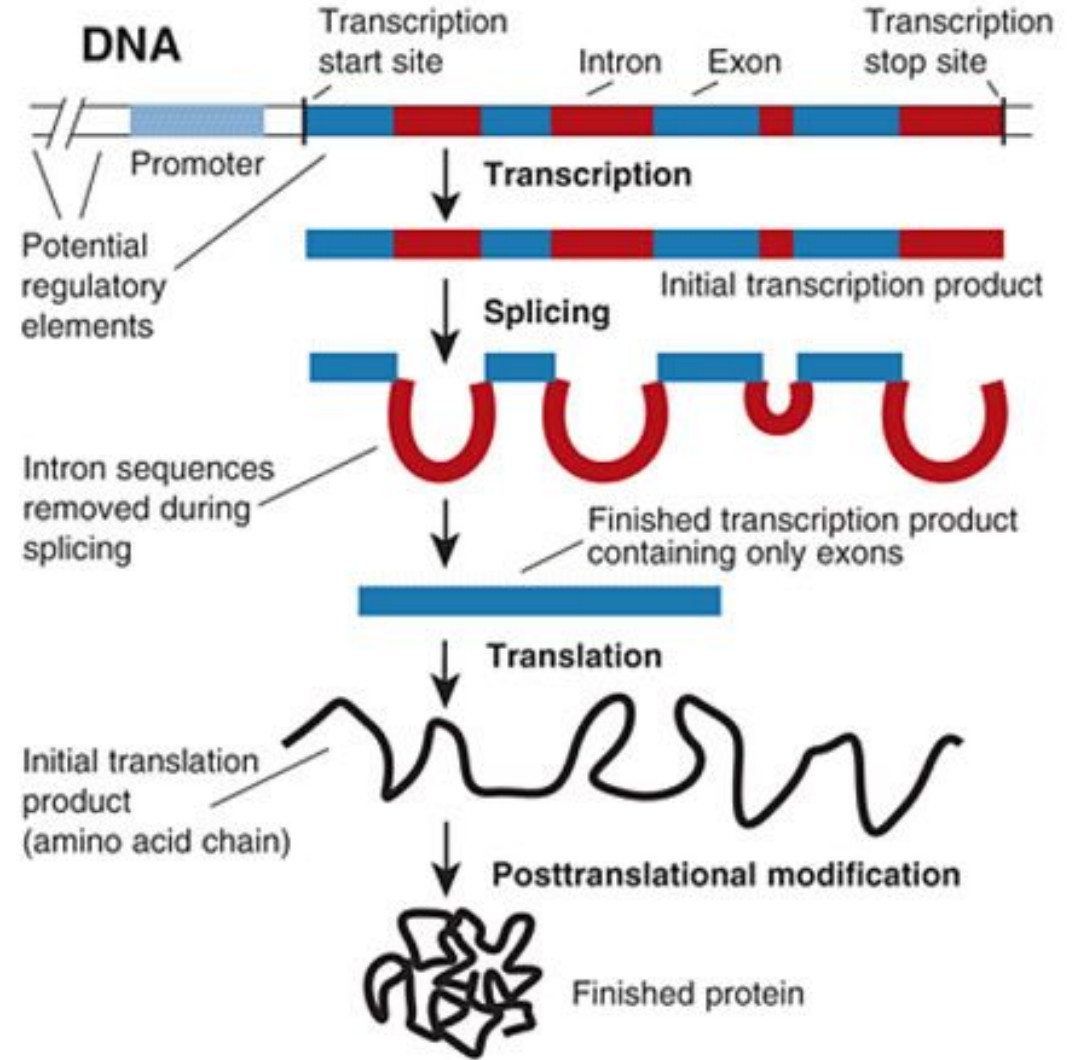
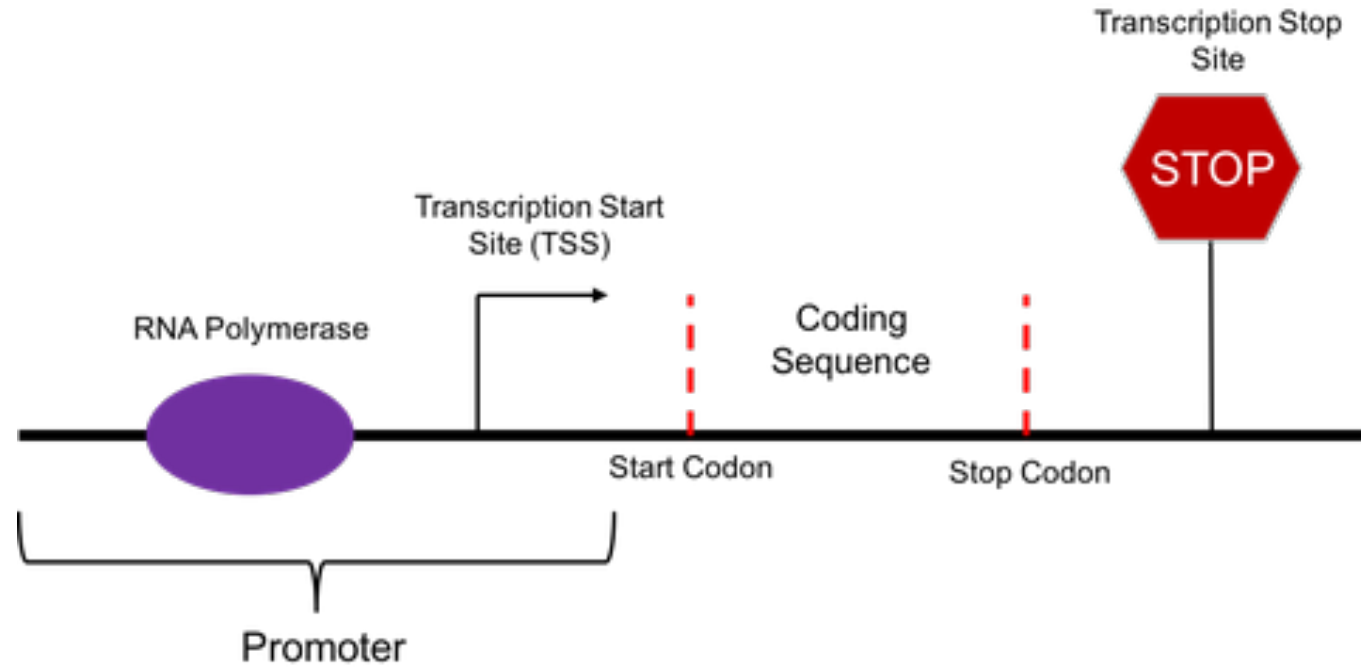


# Promoters-Cis Elements

Assoc. Prof. Dr. Ilker BUYUK





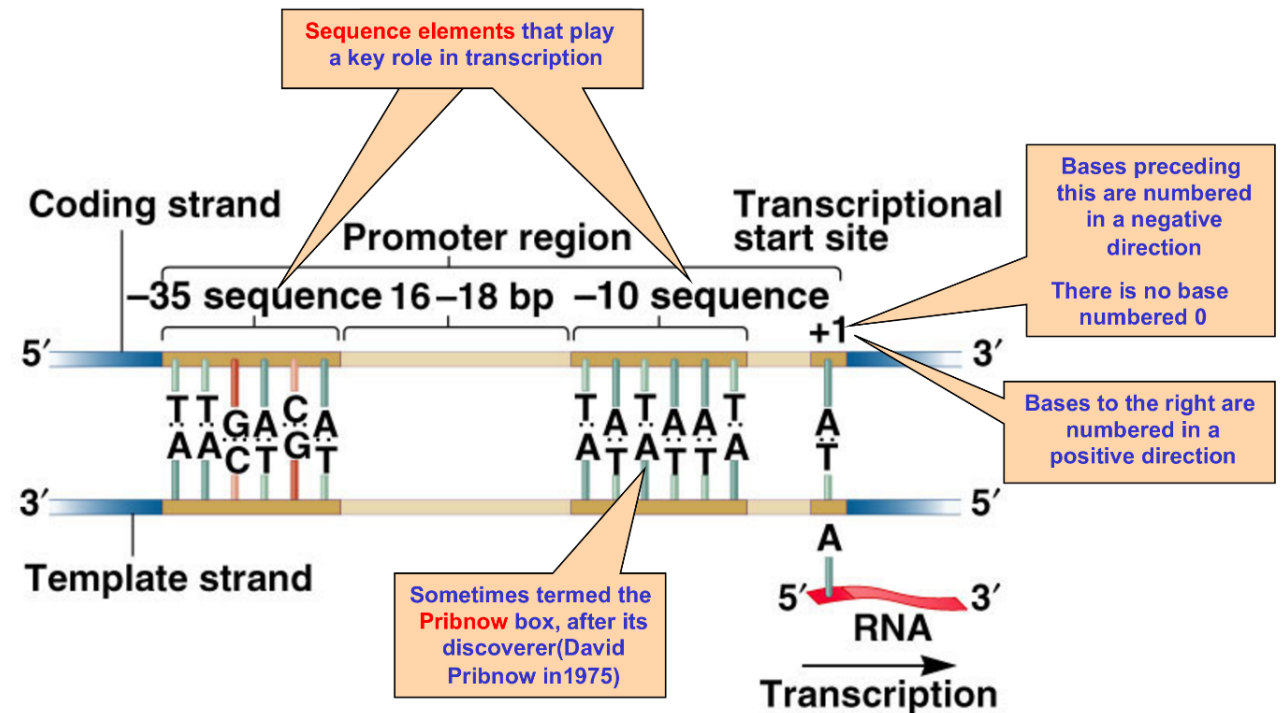
# PROMOTER

- A promoter is the main regulatory portion of a gene.
- The simplest analogy is that a promoter is a “switch” that turns a gene “on” or “off.” It is the portion of the gene where cellular machinery binds before transcribing the DNA blueprint into a useful RNA.

# PROMOTER

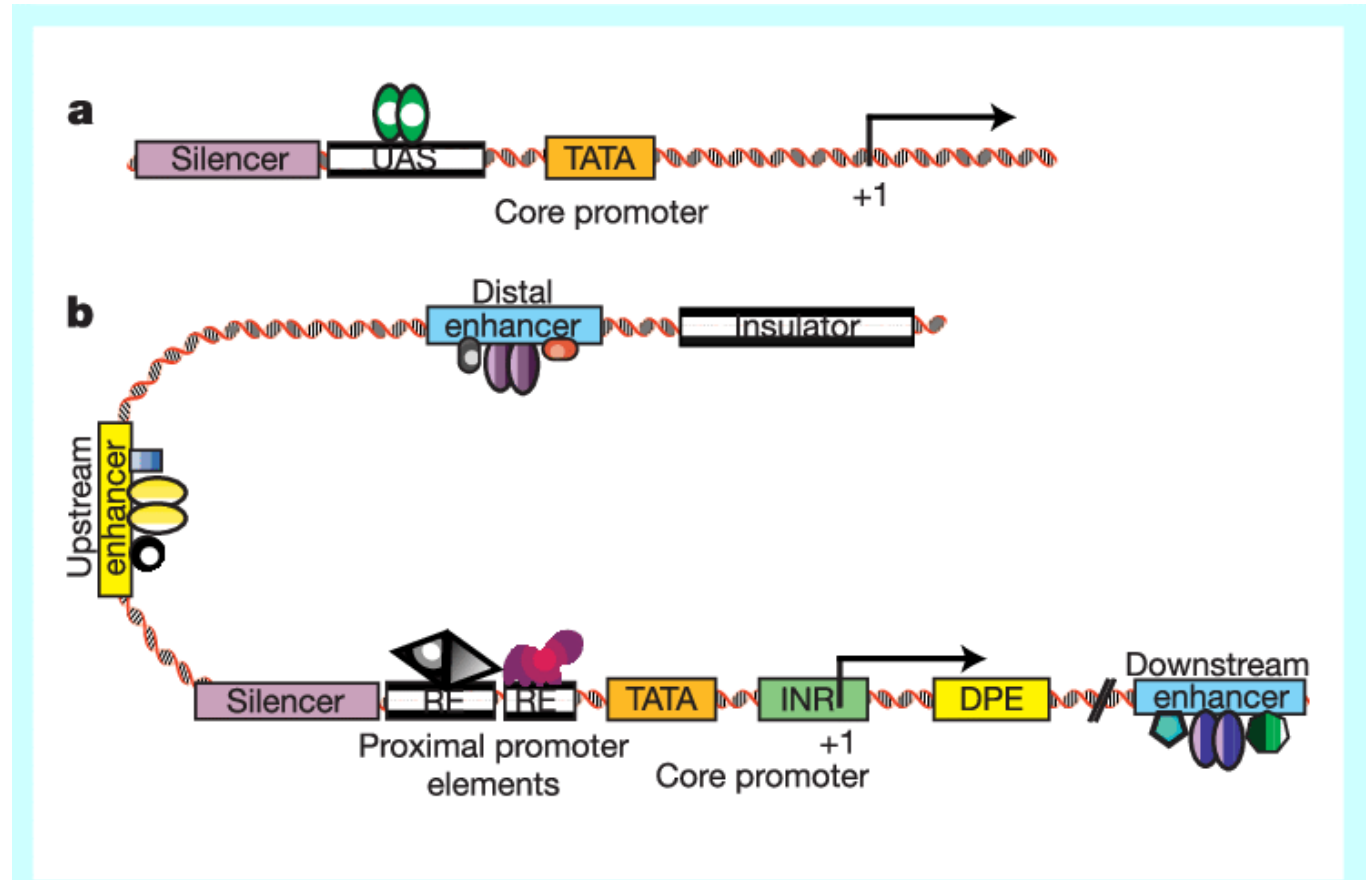
- In genetics, Promoters are DNA sequences located in the 5' region adjacent to the transcriptional start site. RNA polymerase and accessory proteins (transcription factors) bind to the promoter to initiate production of an mRNA transcript. Interactions of proteins at the promoter regulate gene activity by activating or repressing transcription

- The bases in a promoter sequence are numbered in relation to the transcription start site.



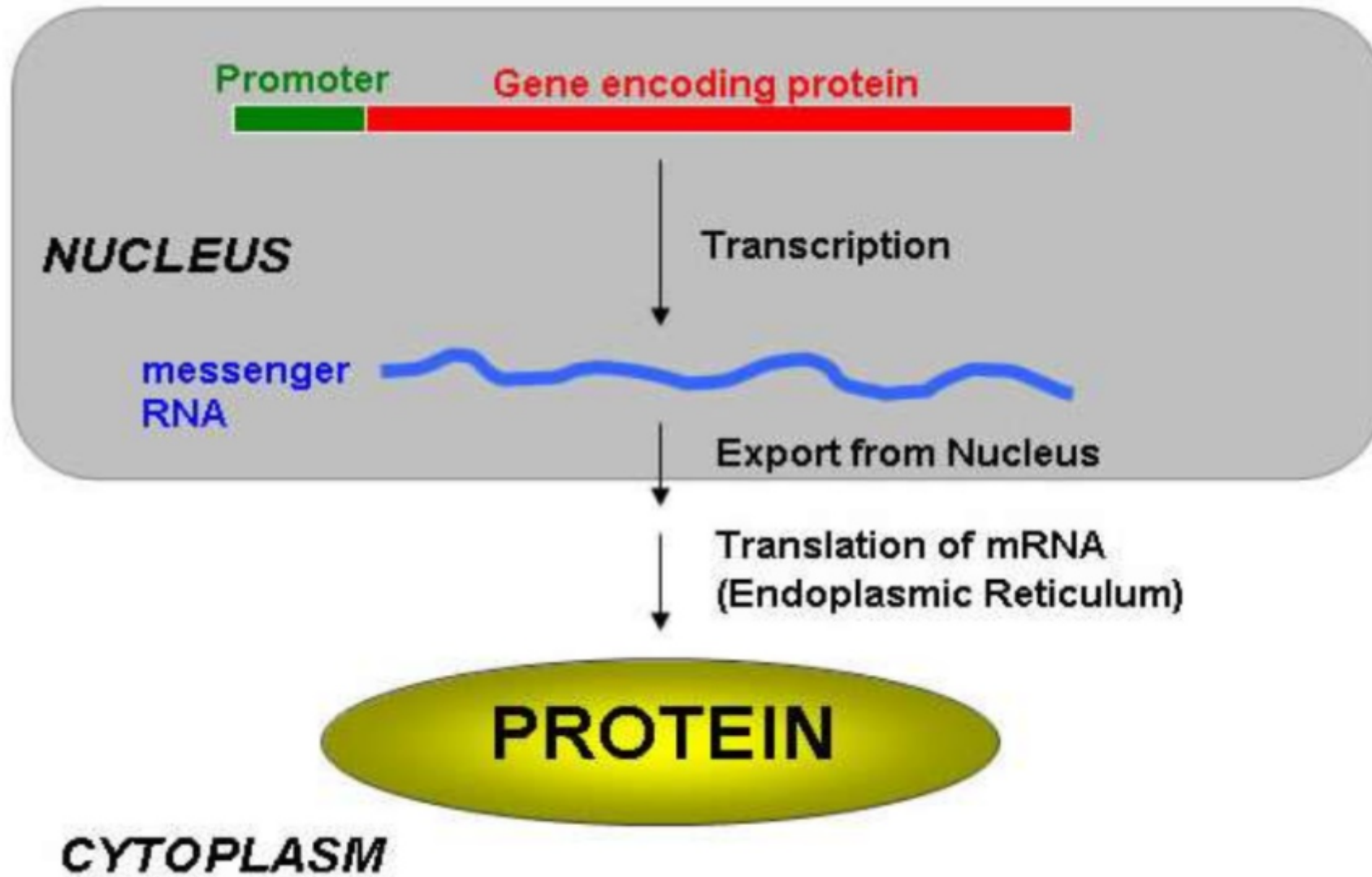
The conventional numbering system of promoters

- There is virtually an endless number of promoters, potentially as many as there are genes.
- Promoters, like genes, are made up of A's, G's, C's and T's all lined up in a certain order. Promoters can be about 100– 1000 base pairs long.
- The promoter region can be short or quite long; the longer the promoter is, the more available space for proteins to bind.



# Role in transcription

- The whole process of transcription starts when the RNA polymerase binds to the promoter. RNA Polymerase has 4 subunits i-e alpha , beta , beta' and sigma. Only the Sigma Factor is required for RNA polymerase to bind to the promoter. As the RNA polymerase binds to the promoter , DNA duplex becomes unwind, base pairs are broken down and transcription bubble is appeared. This is followed by elongation phase where the RNA grows and protrudes from the bubble.



# TYPES OF PROMOTERS USED FOR GENE EXPRESSION

## 1) Constitutive promoters:

- Induce the expression of the downstream located coding region in all tissues irrespective of environmental or developmental factors.
- **Example:** Plant pathogen promoters CaMV 35S promoter.

## 2) Tissue Specific Promoters:

- Operate in particular tissues and at certain developmental stages.
- Maybe induced by endogenous and exogenous factors.
- **Examples:** Tomato pz7 and pz10 gene promoters (for ovary gene expression)

# TYPES OF PROMOTERS USED FOR GENE EXPRESSION

## 3) Inducible promoters:

- Within this group, there are promoters modulated by presence or absence of biotic or abiotic factors such as light, oxygen levels, heat, cold and wounding. Inducible promoters are grouped as:
  - a ) Chemically regulated promoters
  - b) Physically regulated promoters

## 4) Synthetic promoters :

- Promoters made by bringing together the primary elements of a promoter region from diverse origins.
- **Example:**
  - Maize ubiquitin 1 gene ( Ubi 1) core promoter
  - Cytomegalovirus (CMV) promoter
  - CAG promoter



# PROMOTER REGION:

- There are three main portions that make up a promoter: core promoter, proximal promoter, and distal promoter.
- **CORE PROMOTER REGION:**
- The core promoter region is located most proximally and contains the RNA polymerase binding site, **TATA box**, and **transcription start site (TSS)**. RNA polymerase will bind to this core promoter region stably and transcription of the template strand will initiate.

# PROMOTER REGION:

- **PROXIMAL PROMOTER REGION:**

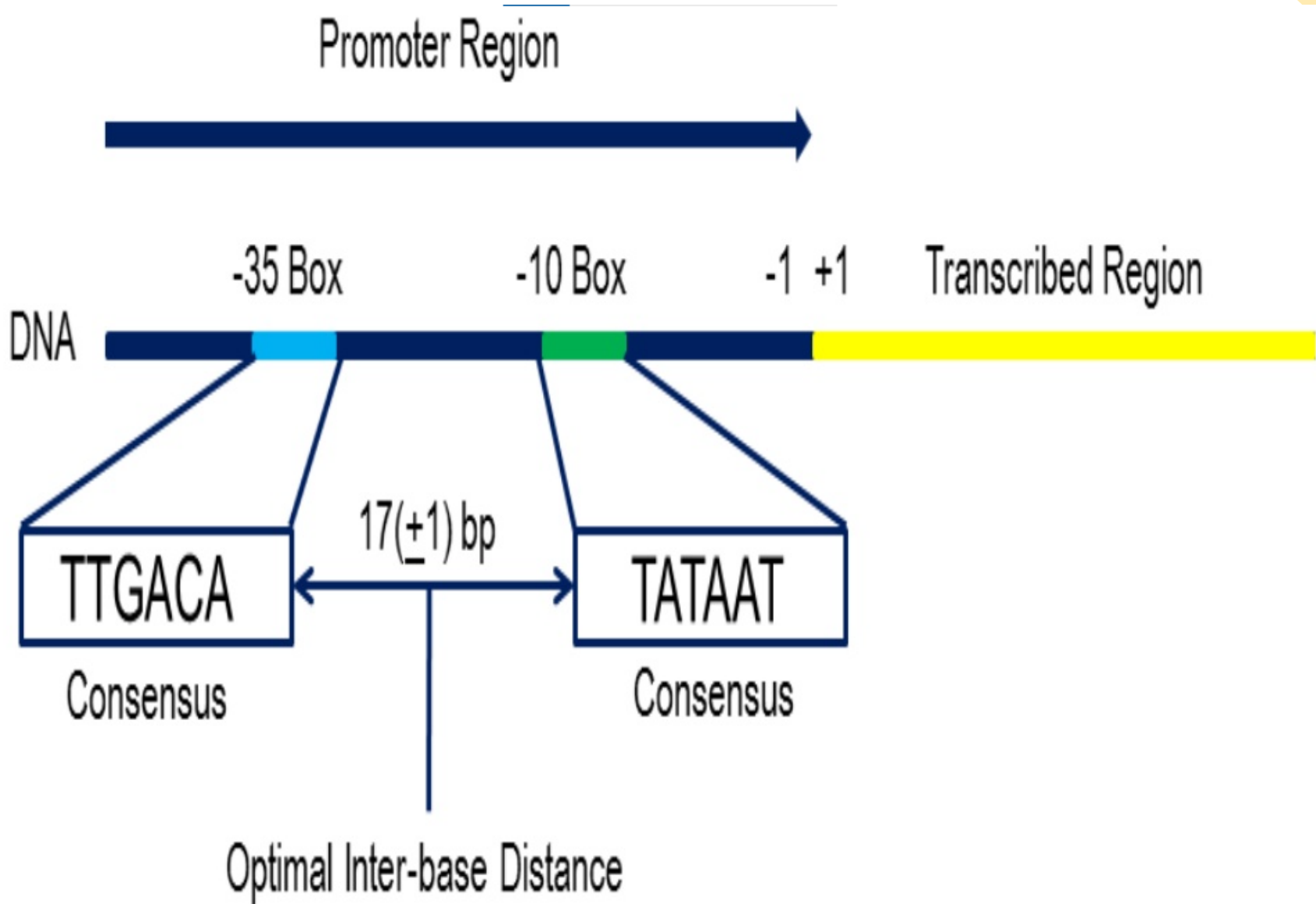
- Proximal promoters are upstream from the core promoter which contain many primary regulatory elements. The proximal promoter is found approximately 250 base pairs upstream from the TSS and it is the site where general transcription factors bind.

- **DISTAL PROMOTER REGION:**

- The final portion of the promoter region is called the distal promoter which is anything further upstream from the gene. The distal promoter also contains transcription factor binding sites, but mostly contains regulatory elements.

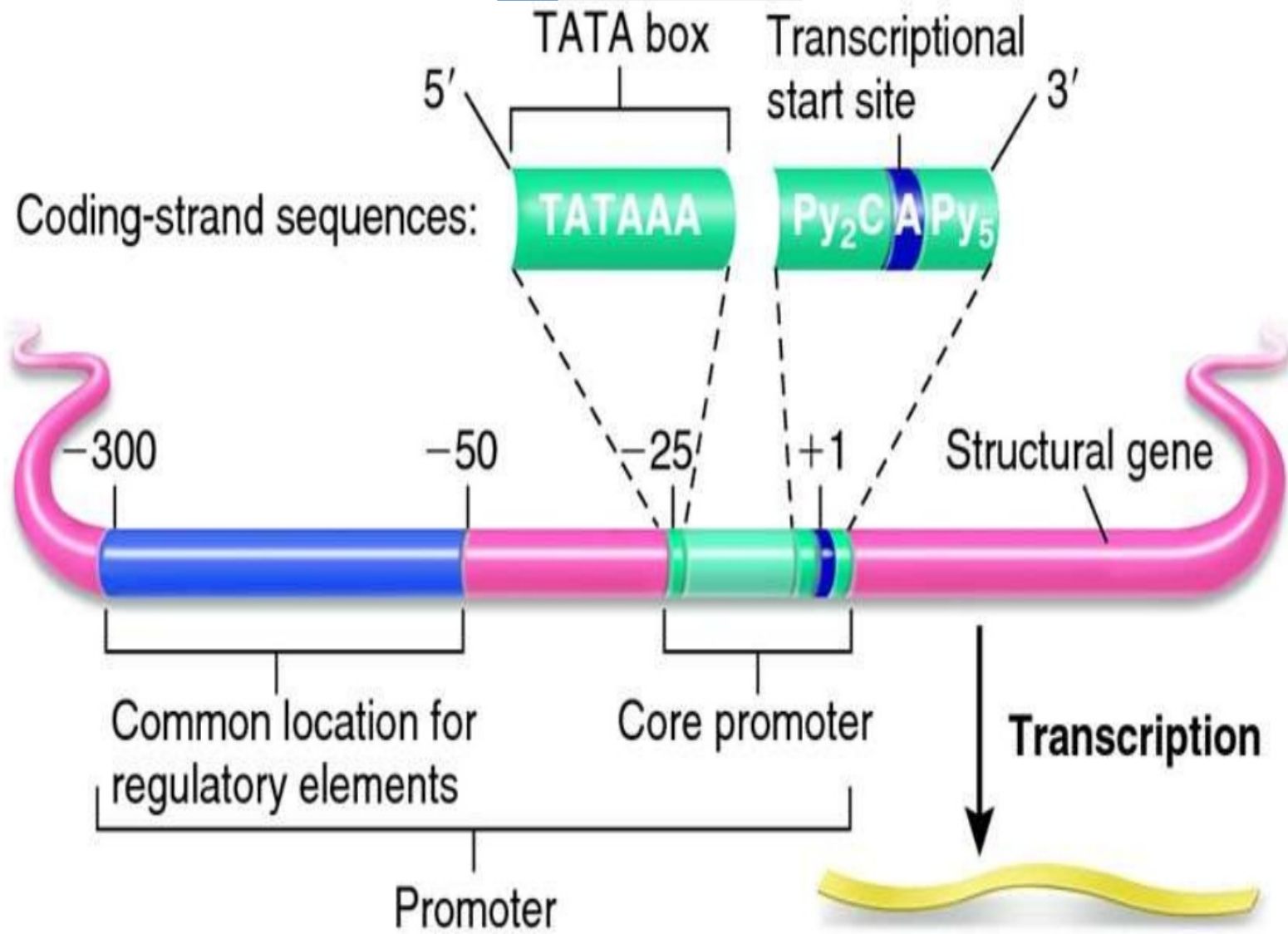
# PROKARYOTIC PROMOTERS

- Promoters in prokaryotic organisms are two short DNA sequences located at the -10 (10bp 5' or upstream) and -35 positions from the transcription start site (TSS).
- Their equivalent to the eukaryotic TATA box, the Pribnow box (TATAAT) is located at the -10 position and is essential for transcription initiation.
- The -35 position, simply titled the -35 element, typically consists of the sequence TTGACA and this element controls the rate of transcription.
- Prokaryotic cells contain sigma factors which assist the RNA polymerase in binding to the promoter region. Each sigma factor recognizes different core promoter sequences.



# EUKARYOTIC PROMOTERS

- Eukaryotic Promoters are much more complex and diverse than prokaryotic promoters. Eukaryotic promoters span a wide range of DNA sequence. There are two parts :
- The core promoter or the basal promoter
- Upstream promoter element
- Core promoter is constituted by TATA box and transcriptional start site. Initiation complex binds to the core promoter upstream elements are responsible of the regulation of transcription.



Conserved eukaryotic promoter elements	Consensus sequence
CAAT box	GGCCAATCT
TATA box	TATAA
GC box	GGGCGG
CAP site	TAC

# BIDIRECTIONAL PROMOTERS

- Pairs of genes control by same promoter but located on opposite strand and opposite direction. Their TSS are separated by less than 1,000 bp.
- In general, they are rich in CpG content .
- Function of genes represented in bidirectional class are often: DNA repair genes, chaperone protein, and mitochondrial genes.
- Genes control by bidirectional promoters are often co- express, but a minority of bidirectional genes have a mutual exclusive expression.
- No correlation between length of promoter and degree of expression



# BIDIRECTIONAL PROMOTERS

- Bi-directional promoters are a common feature of mammalian genomes About 11% of human genes are bi-directionally paired.
- Certain sequence characteristics have been observed in bidirectional promoters:
  - 1: A lack of TATA boxes
  - 2: An abundance of CpG islands
  - 3: Symmetry around the midpoint of dominant Cs and As on one side and Gs and Ts on the other.

- **Sub-genomic promoters**

- A sub-genomic promoter is a promoter added to a virus for a specific heterologous gene, resulting in the formation of mRNA for that gene alone.

- **Detection of promoters**

- A wide variety of algorithms have been developed to facilitate detection of promoters in genomic sequence, and promoter prediction is a common element of many gene prediction methods. A promoter region is located before the -35 and -10 Consensus sequences. The closer the promoter region is to the consensus sequences the more often transcription of that gene will take place.

# COMMONLY USED PROMOTERS

- The most commonly used promoters are:
- Lac (Lactose )
- Trp, ( Tryptophan )
- Tac, ( invitro combo of Lac and Trp)
- T 7 System
- All of these bind to the sigma-70 factor for transcription.

# Why the interest in promoters?

- The interest in promoters stems from the myriad opportunities for controlling gene expression.
- The study and understanding of the function of their multiple components and the factors associated with their performance have opened up the possibility of **modulation of the expression of genes** in homologous organisms as well as in heterologous organisms, where foreign promoters together with genes of interest are inserted.
- Promoters are regarded as **molecular biological tools crucial** for the regulation of the expression of genes of interest.
- As such, they have a huge influence in follow-on research and development in biotechnology.

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

All Databases Search

**COVID-19 Information**  
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

**UNITE**  
A new NIH initiative to end structural racism and achieve racial equity in the biomedical research enterprise.  
[LEARN MORE](#)

**Ending Structural Racism**  
nih.gov/ending-structural-racism

**NCBI Home**  
Resource List (A-Z)  
All Resources  
Chemicals & Bioassays  
Data & Software  
DNA & RNA  
Domains & Structures  
Genes & Expression  
Genetics & Medicine  
Genomes & Maps  
Homology

**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

**Popular Resources**  
PubMed  
Bookshelf  
PubMed Central  
BLAST  
Nucleotide  
Genome  
SNP  
Gene  
Protein  
PubChem

- The National Center for Biotechnology Information provides access to genomic information in order to advance science. The genome database currently comprises of 16,326 species and can be sorted by kingdom, group and subgroup.

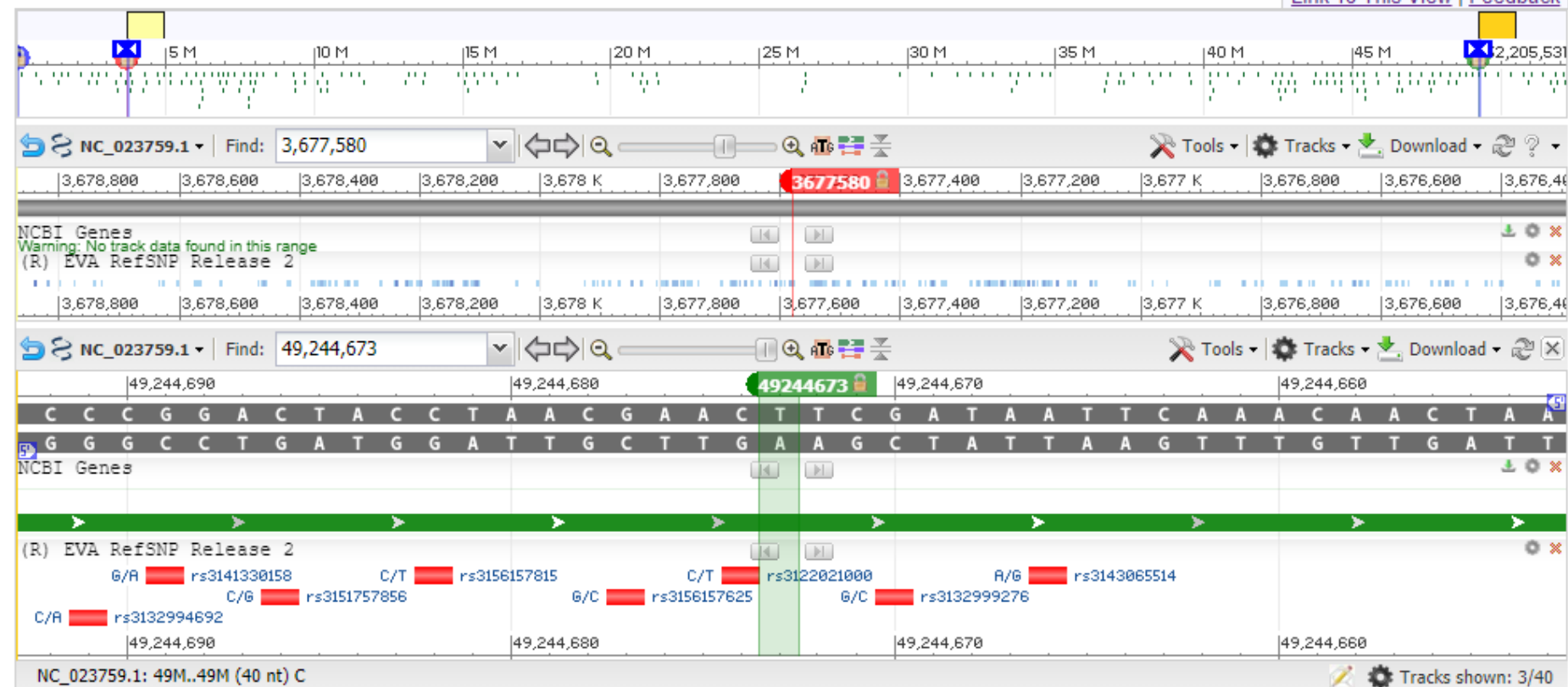
ID	Phaseolus vulgaris Genomic Database Identifier	Physical position on <i>P. vulgaris</i> genome			Protein length (aa)	pI	Molecular weight (Da)	Instability index	Aliphatic index	GRAVY	Stable or unstable	NCBI Accession No.
		Chr.	Start position (bp)	End Position (bp)								
<i>Pvul-LLEC-1</i>	Phvul.001G045400.1.p	1	3,677,580	3,680,026	667	6.04	75271.53	32.38	87.06	-0.176	stable	XP_007161136.1
<i>Pvul-BLEC-1</i>	Phvul.001G239200.1.p	1	49,244,673	49,246,082	345	5.28	37510.23	28.22	89.77	-0.126	stable	XP_007163496.1

## Phaseolus vulgaris cultivar G19833 chromosome 1, whole genome shotgun sequence

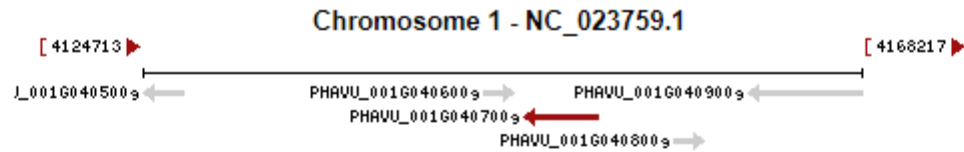
NCBI Reference Sequence: NC\_023759.1

[GenBank](#) [FASTA](#)

[Link To This View](#) | [Feedback](#)



Chromosome: 1; NC\_023759.1 (4147723..4152276, complement)



### regions, transcripts, and products

Sequence: NC\_023759.1

[Go to reference sequence details](#)

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

XP\_007161075.1

CDS: XP\_007161075.1  
Name: hypothetical protein  
Comment: encoded by transcript PHAVU\_001G0407001  
Location: complement(4,147,954..4,152,276)  
[Length]  
Span on NC\_023759.1: 4,323 nt  
CDS length: 2,154 nt  
Protein length: 717 aa  
[Positional Info]  
NC\_023759.1 position: 4,149,413  
Exon: 2 of 3  
CDS position: 910  
Protein position: 304  
Protein sequence: TGE LTER NLIHSWEF[S]STLNSSTASNDNPS

Download FASTA: [XP\\_007161075.1](#)

### Why

Gene References Into Functions

[GeneRIF?](#)

[View GeneRIF](#) [Correction](#)

### Gene information

#### Links & Tools

Phytozome: [Phvul.001G040700.1.p](#)

BLAST Protein: [XP\\_007161075.1](#)

BLAST nr: [NC\\_023759.1\(4,147,954..4,152,276\)](#)

FASTA record: [NC\\_023759.1\(4,147,954..4,152,276\)](#)

[XP\\_007161075.1](#)

GenBank record: [NC\\_023759.1\(4,147,954..4,152,276\)](#)

[XP\\_007161075.1](#)

Graphical View: [XP\\_007161075.1](#)

NC\_023759.1[4149498..4150299](-).fa - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

>ref|NC\_023759.1|:c4150299-4149498 Phaseolus vulgaris cultivar G19833 chromo  
shotgun sequence

```
AACAAATTCAGAACAAACTTATTCATTCTTGCAAACAAATCCACTTTTGTATTGTTCTTCTCTCAGTTC
TTAGCTCCGGCAAAGTTGAGAAAAAAACTGCCATTGAGAATGAGATTTGCAGACATGGTGACCATT
TTGTTCTATTCTTGCCATTCCTTCTCCTTTCCCTTAAACAGTGGAATCACTAAATTTCAACATTACAAA
CTTCGATGACCCTGAGAGTGCAAAAAACATGGCATAACATGGGTGACGGCAAACCAAAGAACGGTTCCATA
GACCTCAACATAGTCTCGTATTATTTCCGTGTTGGGAGAGCCTTGTATGCACAACCTTTGCACCTTTGGG
ATTCATCTTCAAGTGTGTAACAGACTTCACCACACGCTTCACCTTTTCTATTGAAAAAGGAAACGATGA
TACCGCTAGCTACGCTGATGGTTTCGCCTTCTACATAGCCCCTCACGGCTATCAGATTCCTCCAAATGGT
GCAGGTGGCACTTTTGGCCTTTTCAATGGCACTTTCAACAGCTTCCTTCCCCAAAATCGCATCCTTGCGG
TTGAGTTTGACACACTTAATGGCACCATTGACCCTCCCATGCAGCACGTTGGCATAGATGATAACTCCCT
CAAATCCGCGACTACTGGCAAGTTTGACATCGACGAAAATCTGGGAAAGAAATGTAACGCTTTGGTAACC
TACAATGCTTCCGACAAGACACTCTTTGTGGCTTGGTCCTTCAACGGAACCGCAACGATACACTCCAATT
CTTCACTTTCTTACAAGATCGACCTCTTGAGA
```



### Phytozome quick search (advanced)

[Flagships](#)
[Clustered Genomes and Families](#)
[Unclustered Genomes](#)

for

### Help with Phytozome

#### Documentation

- View a tree representation of the species in Phytozome v12.1
- View a Quick Start Guide to using Phytozome
- Check out the FAQs

### About Phytozome 12.1.6

**Note: all genome releases since May 2019 are available only in Phytozome v13, our newest version of Phytozome. Access it at Phytozome-Next.**

Phytozome, the Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute, provides JGI users and the broader plant science community a hub for accessing, visualizing and analyzing JGI-sequenced plant genomes, as well as selected genomes and datasets that have been sequenced elsewhere. As of release v12.1.6, Phytozome hosts 93 assembled and annotated genomes from 82 Viridiplantae species. [More](#)

### News (details...)

- (2020-07-14) **BULK DATA ACCESS is back!**
- (2020-07-02) **Limited DATA ACCESS Jul 10-14!**
- (2020-04-15)

### System Status (2021-05-02 05:55)

- ✓ Search
- ✓ BLAST
- ✓ Database
- ✓ PhytoMine
- ✓ PhytoMine-Web

- *Phytozome* is a plant comparative genomics portal wherein families of related genes representing the modern descendants of ancestral genes are constructed at key phylogenetic nodes. Currently *Phytozome* provides access to 58 sequenced and annotated green plant genomes.

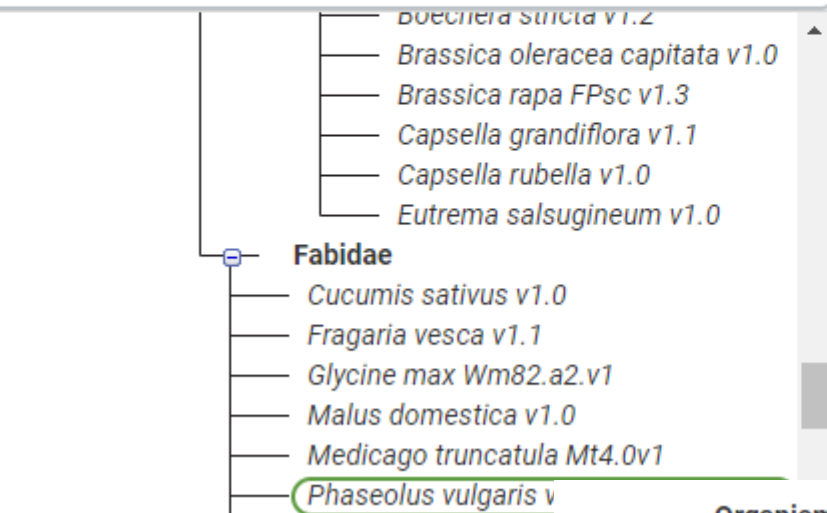
1. Select a Target 1 species selected

2. Build your query GO

Target set: Phytozome 12.1 Pre-release species

Target type: Ancestor nodes Species

Phaseolus vulgaris v2.1



Search type: Keyword BLAST

Phvul.001G040700.1.p

Algorithm parameters

Add trailing wildcard:

Use family settings:

Organism Phaseolus vulgaris

Transcript Name Phvul.001G040700.1 (primary)

Location: Chr01:4336106..4338606 forward

Alias Phvul.001G040700.v1.0 Phvul.001G040700.1.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 [2501] BLAST this sequence at Phytozome NCBI

Show flanking sequence: upstream: downstream: 0 Submit

```

>Phvul.001G040700 | Chr01:4336106..4338606 forward
ATTAAATGATCAAGACACCAATAAGAAAAAAGAAAAAAGTAGAGGAGATCAAAGACCAAACCTCCAATTGAGTCAAAGAAAAGATCTTAAATTAATCC
AAAACCTCCACCTTTAAAAATAATCTTATTCTATGACTCCATATCTCACTCACTAGTCATCGCAATCAAATAAAAAGCAGCAGAAAAACACTTCTAATCGC
AATCCCAGTCTATGTAATTACTCGTAGTGACAATCTTATTTTCCAACATCCTCCAAGCTGTAACCTAAGCTGAAGGCAAAACCTTAATCCTCCAAAAC

```

<https://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST>

TATTTCAAGAAAAAATCAATAAAAAGGATATCCACATAACATACTCTTACGTCATAATATTTTTTTATAATATGTGAGATAGGAGATAGACAAGTATG  
TCAAACAAATCTTACCGTGGAATCCAATTAACAATCTAAGAGAAATTATCATGGAATCATTCTTAGACCGAATCAATGTGTAATTAATATTATTAATT  
ATTTATTTAATTTGAAAATTATCATATGTGTCATTTGGAGGATGATAGAGACCATAAAACAATTTATTTAATAATTTTTTCCAAATCATTCTTGTT  
GATTCTACTTGCTCTTGAAGACTTGGTTCAACCATCACCATATATATTCCATCAAAATTAATTAATGACCTAACCGTCTCACAGTGAACCTTCGTCTT  
CATTACGCTTGACTACAACGTCAATAAAAACCTTTGAAAAAATAAAGAAAATTGCAGTACACCACATTGCAGTATTGGGACGCAACCTAGGAAGACCAA  
TTATACAGACAAAAAGGTTTAGTCTCAATCAGTAAAACCTTTTACCTGCTCTTCTTGTAGTTAATCTTTAATATAAGGATATACTAATATTGCTTATC  
ATACCATATTATTTAAAATTTTTGTTTTCAGAATTTATAAAACCTAATACACTTTTATAATTTTTCATTTATACTTAATATAAATTTTACAAAGTGAA  
AAGTAATTTTTATTCTTTGAATTCAAAACCTTGCATGATGGAATATAATATTAAGAAAATTACTAAAGACATTTGTTTAAATTGTTGAGTTAAAC  
TTTTATTTAAACAGCAACCTGCGTGAACAAATTCAGAACAACTTATTCACTCTTCAAACAAATCCACTTTTGTATTGTTCTTCTCAGTTCTTAGCT  
CCGGCAAAAGTTGAGAAAAAAAACCTGCCATTGAGAATGAGATTGCAGACATGGTGACCATTTTGTCTATTCTTGGCCATTCTCTCCTTTCCTTA

Genomic sequence

Transcript sequence

CDS sequence

Peptide sequence

Show all

key: 5' UTR CDS 3' UTR

ACGACCAAGGAGATTCTGTTGGTTTTGATTAATGCTTATGAAAGCTTCACACTCATCTCTTGGTTCACAAAAACCTTTGGCTGGGATATTAG  
GTATGTATGGTTAAATAATATATACTCTCTATCTTTTATTTATTTTATCATGAACAAAACCTTGGATACTTAGTGTTTAATTTAATTAGCATATA  
TAATTCCTTCAATCAGACTTTTATCAGAATCTATTTAATCACACTTTTATTGAGTTGATTACATCAGTTATGGAACCTCAATTACACACTTTGTCTTCT  
AAAAATTTCTAACAGGTACAAGGTAGCATTAGGCGTGGCTTTGGCACTTCGTTATCTTCATGAGGATGCCGAACAGAGTGTTCTTCATAGGGATATTAAG  
TCGGCTAATGTGTTGCTGGACACAGATTTTAGCACGAAGCTTGGAGATTTGGGATGGCCAAGTTGGTGGATCCAAGGTTGAAGACTCAGAGGACAGGAG  
TGGTGGGGACTTACGGGTACCTTGCCCCAGAATACATGAATGGAGGTAGGGCTAGCAAGGAATCAGACATGTATAGTTTTGGTGTGTGGCTCTTGAAT  
TGCATGTGGAAGAAGGACTTATCTGGATGGGGAGTTTCAATACCTCTGATGAACTGGGTGTGGCAACAGTATGTGGAAGGGAATGTGATGGATGTTGTT  
GATGAGAGATTGAACATGGAGTTTGTATGTGGATGAAATGAGAAGCTTGATCATTGTGGGTTGTGGTGCCTAACCTAACGACAAGGAAAGGCCAAAGG  
CAGCACAAGTAATAAAAGTTCTTGAGCTTGAAGCACCATTGCCAGAGCTTCCACTTGATATGCATGATCGTCTCTCTTCTCTAAGTACTTATAATCA  
TGCACAACCCACTCATAATTCCTTACAATCACTACCCTTCACTAACAGTTTGTCACTATTGGACGTTAATTTGCTAAGTCAGGCTCTGATCCCACACCA  
TGAAGTTATGCATAATTTTCGTTTCATTTATTTTATTTGATCTTGCCTTGTACCAAGTATTTAGTGACCAGCAACTAGTTGGAAGATGATTTCTTTTCTC

# Menu

Query CARE

Search for CARE

## Other Queries

[Classification...](#)

[Genes...](#)

[Name of Factor...](#)

[Name of Site...](#)

[Referencia...](#)

[Motif Sampler](#) NEW

[Clustering](#) NEW

[Enter new data](#) NEW

[General References...](#)

Help

About

Plant Systems Biology

Site maintained by  
[Stephane Rombauts](#)



## Search for CARE

email address to send the results back

reference name or ID for the sequence (optional)

Sequence to submit

please only submit **fasta formatted sequences** as pure/simple text files, no word documents or other binary things

currently file size limited to 100Kb

NEW

★ To use our output you will need a modern browser for the implemented DHTML features. Earlier version might work partially (printing will not work properly)

★ Before using scripts `dth=20 src='../html/GIF/star.gif'` because of the increased usage of the site, I have been forced to implement a scheduling and results will be returned via email.

Scripts should not be used anymore as one can upload a multi-fasta file, but if you do, please be gentle, and do introduce a sleep of 60sec between each request to submit a number of sequences through this page please contact [me](#)

(Those not respecting that will be blocked :-)

★ because of the increased usage of the site, I have implemented a scheduling and results will be returned via email.

This means that you might have to wait a little in order to get the results sent back to you, certainly when the load on the server is high. Please check you email spam folder if you haven't seen any email.

★ Please, use ASCII characters, roman alphabet, everything else will be trashed



## Search for CARE

email address to send the results back

reference name or ID for the sequence (optional)

Sequence to submit

please only submit **fasta formatted sequences** as pure simple text files, no word documents or other binary things

 Dosya seçilmedi

currently file size limited to 100Kb

**NEW**

★ To use our output you will need a modern browser for the implemented DHTML features. Earlier version might work partially (printing will not work properly)

>PlantCARE\_21713

```
+ GATAACAAAT TATTATAATT ATAAAAAATG AGCTGTAACA TTTATAATGA AATGATAAAA TAAAGATTCA
- CTATTGTTTA ATAATATTAA TATTTTTTAC TCGACATTGT AAATATTACT TTACTATTTT ATTTCTAAGT

+ TTTAAATTTT AAGAAATGAA TTGATTG TTC TAACAAGAAA ATATTTATTA TACACATTTT AACTAATGCA
- AAATTTAAAA TTCTTTACTT AACTAACAAG ATTGTTCTTT TATAAATAAT ATGTGTAAAA TTGATTACGT

+ ATTAAATAAA TAATACAGTA CGTTTTCCAA AATAGAAAAT CAGATATAAA TTTTGTGGT ATTTCAAGAA
- TAATTTATTT ATTATGTCAT GCAAAAGGTT TTATCTTTTA GTCTATATTT AAAACAAACA TAAAGTTCTT

+ AAAAAATCAA TAAAAGGATA TCCACATAAC ATACTCTTAC GTCATAATAT TTTTTTTATA ATATGTGAGA
- TTTTTTAGTT ATTTTCCTAT AGGTGTATTG TATGAGAATG CAGTATTATA AAAAAATAT TATACACTCT

+ TAGGAGATAG ACAAGTATGT CAAACAAATC TTACCGTGGA ATCCAATTAA CAATCTAAGA GAAATTATCA
- ATCCTCTATC TGTTCATAACA GTTTGTTTAG AATGGCACCT TAGGTTAATT GTTAGATTCT CTTTAATAGT

+ TGAATCATT TCTTAGACCG AATCAATGTG TAATTAATA TTATTAATTA TTTTATTTTA ATTTGAAAAT
- ACCTTAGTAA AGAATCTGGC TTAGTTACAC ATTAATTTAT AATAATTAAT AAAATAAAAT TAAACTTTTA

+ TATCATATGT GTCATTTGGA GGATGATAGA GACCATAAAA CAATTTATTT AATAATTTTT TTCCAAATCA
- ATAGTATACA CAGTAAACCT CCTACTATCT CTGGTATTTT GTTAAATAAA TTATTAAAAA AAGGTTTAGT
























+ TTCTTGTTG ATTCTACTTG CTCTTGAAGA CTGGTTCAA CCATCACCAT ATATATTCCA TCAAAATTAA
- AAGAACCAAC TAAGATGAAC GAGAACTTCT GAACCAAGTT GGTAGTGGTA TATATAAGGT AGTTTTAATT

+ ATTAATTGAC CTAACCGTCT CACAGTGAAC CTCGTCTTC ATTACGCTTG ACTACAACGT CAATAAAAAC
- TAATTAACTG GATTGGCAGA GTGTCACCTG GAAGCAGAAG TAATGCGAAC TGATGTTGCA GTTATTTTTG

+ TTTGAAAAAA TATAAAGAAA ATTGCAGTAC ACCACATTGC AGTATTGGGA CGCAACCTAG GAAGACCAAT
- AAATTTTTTT ATATTTCTTT TAACGTCATG TGGTGTAACG TCATAACCCT GCGTTGGATC CTTCTGGTTA

+ TATACAGACA AAAAGGTTTA GTCTCAATCA GTAAAACTTT TTACCTGCTC TTCTTGTTTA GTTAATCTTT
- ATATGTCTGT TTTTCCAAAT CAGAGTTAGT CATTTTGAAA AATGGACGAG AAGAACAAAT CAATTAGAAA
```

## Motifs Found

+		
+		ACE
+		AT~TATA-box
+		Box 4
+		CAAT-box
+		CGTCA-motif
+		ERE
+		GATA-motif
+		MRE
+		MYB
+		MYC
+		Myc
+		TATA
+		TATA-box
+		TCA
+		TCT-motif
+		TGA-box
+		TGACG-motif
+		Unnamed__1
+		Unnamed__4
+		W box
+		WUN-motif
+		as-1

+ AAAAAATCAA TAAAAGGATA TCCACATAAC ATACTCTTAC GTCATAATAT TTTTTTTATA ATATGTGAGA  
- TTTTTTAGTT ATTTTCCTAT AGGTGTATTG TATGAGAATG CAGTATTATA AAAAAAATAT TATACACTCT

+ TAGGAGATAG ACAAGTATGT CAAACAAATC TTACCGTGGA ATCCAATTAA CAATCTAAGA GAAATTATCA  
- ATCCTCTATC TG TTCATACA GTTTGTTTAG AATGGCACCT TAGGTTAATT GTTAGATTCT CTTAATAGT

+ TGGAATCATT TCTTAGACCG AATCAATGTG TAATTAAATA TTATTAATTA TTTTATTTTA ATTTGAAAAT  
- ACCTTAGTAA AGAATCTGGC TTAGTTACAC ATTAATTTAT AATAATTAAT AAAATAAAAT TAAACTTTTA

+ TATCATATGT GTCATTTGGA GGATGATAGA GACCATAAAA CAATTTATTT AATAATTTTT TTCCAAATCA  
- ATAGTATACA CAGTAAACCT CCTACTATCT CTGGTATTTT GTTAAATAAA TTATTAAAAA AAGGTTTAGT

+ TTCTTGGTTG ATTCTACTTG CTCTTGAAGA CTTGGTTCAA CCATCACCAT ATATATTCCA TCAAAATTAA  
- AAGAACCAAC TAAGATGAAC GAGA ACTTCT GAACCAAGTT GGTAGTGGTA TATATAAGGT AGTTTTAATT

+ ATTAATTGAC CTAACCGTCT CACAGTGAAC CTCGTCCTC ATTACGCTTG ACTACAACGT CAATAAAAAC  
- TAATTA ACTG GATTGGCAGA GTGTC ACTTG GAAGCAGAAG TAATGCGAAC TGATGTTGCA GTTATTTTTG

+ TTTGAAAAAA TATAAAGAAA ATTGCAGTAC ACCACATTGC AGTATTGGGA CGCAACCTAG GAAGACCAAT  
- AA ACTTTTTT ATATTTCTTT TAACGTCATG TGGTGTAACG TCATAACCCT GCGTTGGATC CTTCTGGTTA

+ TATACAGACA AAAAGGTTTA GTCTCAATCA GTAAAACTTT TTACCTGCTC TTCTTGTTTA GTTAATCTTT  
- ATATGTCTGT TTTTCCAAAT CAGAGTTAGT CATTTTGAAA AATGGACGAG AAGAACAAAT CAATTAGAAA

+ AATATAAGGA TATACTAATA TTGCTTATCA TACCATATTA TTTTAAAATT TTTGTTTTCA GAATTTTATA  
- TTATATTCCT ATATGATTAT AACGAATAGT ATGGTATAAT AAAATTTTAA AAACAAAAGT CTTAAAATAT

+ AAACCTAATA CACTTTTATA ATTTTTCATT TATACTTAAT ATAAATTTTA CAAAGTGAAA AGTAATTTTT  
- TTTGGATTAT GTGAAAATAT TAAAAAGTAA ATATGAATTA TATTTAAAAT GTTTC ACTTT TCATTA AAAA

+ TATTCTTTGA ATTCAAA ACT TTGCATGATG GAATATAATA TTAAAAAGAA AAATTACTAA AGACATTTGT  
- ATAAGAACT TAAGTTTTGA AACGTACTAC CTTATATTAT AATTTTCTT TTTAATGATT TCTGTAAACA

+ TTAAATTGTT GAGTTAACT TTATTTAAAC AGCAACCTGC GTGAACAAAT TCAGAACAAA CTTATTCATT  
- AATTTAACAA CTCAATTTGA AATAAATTTG TCGTTGGACG CACTTGTTTA AGTCTTGTTT GAATAAGTAA

+ TATA-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
<a href="#">TATA-box</a>	Brassica napus	12	+	6	ATTATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	61	+	8	TAAAGATT	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica juncea	40	-	7	TATAAAT	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Helianthus annuus	265	-	6	TATAAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	19	-	5	TATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	185	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	42	-	5	TATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	266	-	5	TATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	14	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	119	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Helianthus annuus	41	-	6	TATAAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Pisum sativum	264	-	7	TATAAAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	20	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica oleracea	184	+	6	ATATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	43	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	267	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	13	-	5	TATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica napus	18	+	6	ATTATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica napus	117	+	6	ATTATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica napus	539	+	6	ATATAT	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	118	-	5	TATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	540	+	6	TATATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica napus	541	+	6	ATATAT	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	542	+	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica oleracea	640	+	6	ATATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	641	-	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica napus	699	+	6	ATTATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	700	-	5	TATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	701	-	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	764	-	8	TAAAGATT	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Brassica oleracea	772	+	6	ATATAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	773	-	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	781	-	4	TATA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Pisum sativum	834	-	7	TATAAAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Helianthus annuus	835	-	6	TATAAA	core promoter element around -30 of transcription start
<a href="#">TATA-box</a>	Arabidopsis thaliana	836	-	5	TATAA	core promoter element around -30 of transcription start



