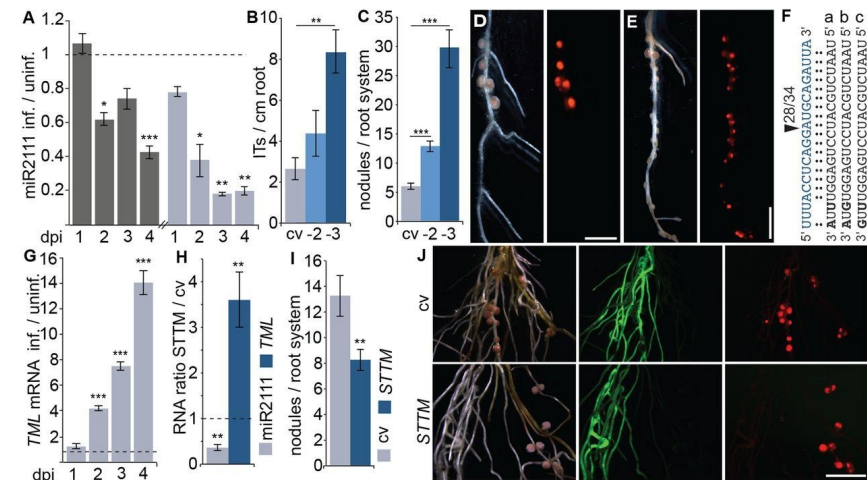


# Systemic control of legume susceptibility to rhizobial infection by a mobile microRNA

1

Nitrogen-fixing root nodules on legumes result from two developmental processes, bacterial infection and nodule organogenesis. To balance symbiosis and plant growth, legume hosts restrict nodule numbers through an inducible autoregulatory process. Here, we present a mechanism where repression of a negative regulator ensures symbiotic susceptibility of uninfected roots of the host *Lotus japonicus*. We show that microRNA miR2111 undergoes shoot-to-root translocation to control rhizobial infection through posttranscriptional regulation of the symbiosis suppressor TOO MUCH LOVE in roots. miR2111 maintains a susceptible default status in uninfected hosts and functions as an activator of symbiosis downstream of LOTUS HISTIDINE KINASE1-mediated cytokinin perception in roots and HYPERNODULATION ABERRANT ROOT FORMATION1, a shoot factor in autoregulation. The miR2111-*TML* node ensures activation of feedback regulation to balance infection and nodulation events.



**Fig. 1 miR2111 regulates *TML* posttranscriptionally.**

(A) miR2111 abundance in *Lotus* leaves (dark bars) and roots (light bars) at 1-4 days postinfection (dpi) with *M. loti*. (B) Infection thread (IT; 10 dpi) and (C) nodule numbers (21 dpi) in *pUBQ1:MIR2111-2* (-2) and *pUBQ1:MIR2111-3* (-3) compared to control (cv) roots. (D and E) Nodulation in control (D) and *pUBQ1:MIR2111-3* (E) roots (21 dpi). Right-hand panels visualize *M. loti* *DsRED* in nodules. Scale bars, 2 mm. (F) miR2111 directs *TML* cleavage. Bold font marks polymorphisms between miR2111 isoforms a-c (black). Numbers: degradome 5' ends at arrowhead/total within *TML* target region (blue). (G) *TML* mRNA in *M. loti*-infected roots (1-4 dpi). (H to J) *miR2111STTM* (*STTM*) expression reduced miR2111, increased *TML* (H), and reduced nodulation [(I) and (J)] compared to control roots (cv). (I)  $n = 23/26$  (*STTM*/cv). Green fluorescence [(J), center] shows co-transformation, red [(J), right] nodules with *M. loti* *DsRED*. Scale bar, 5 mm. Transgenic roots [(B) to (E) and (H) to (J)] were *A. rhizogenes*-induced. [(A), (G), and (H)] qRT-PCR analyses. RNA levels are relative to two reference genes. Error bars: SEM of at least three biological replicates. Student's *t* test *P* values: \* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ . (I) Results represent one biological replicate ( $P = 0.006$ ) and were similar in a second ( $P = 0.001$ ).

<http://science.sciencemag.org/content/early/2018/08/29/science.aat6907>

# Announcements

- Phytobacteriology in the News! Week 6 [LINK](#)
- Reviewing assignments, how is it going?
- Reading for next week (published last week) [LINK](#)

## **Phevamine A, a small molecule that suppresses plant immune responses**

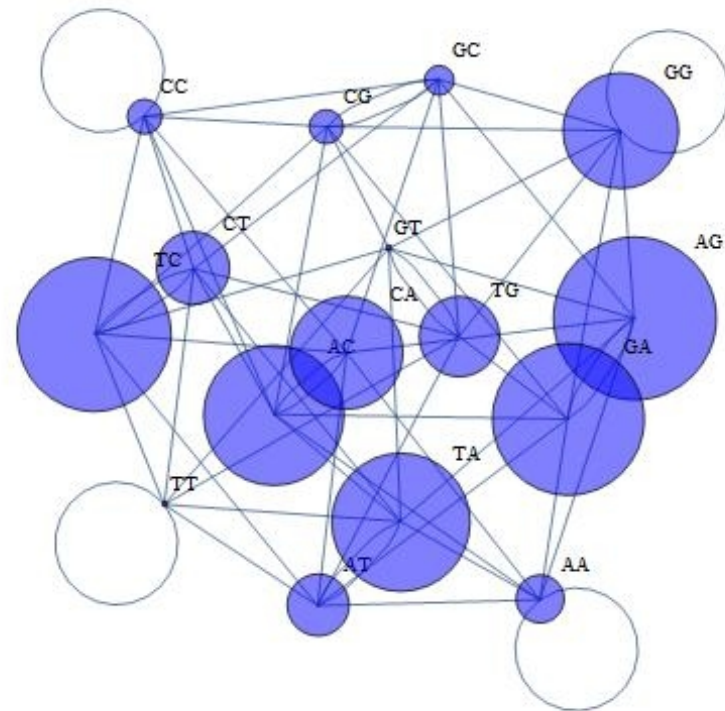
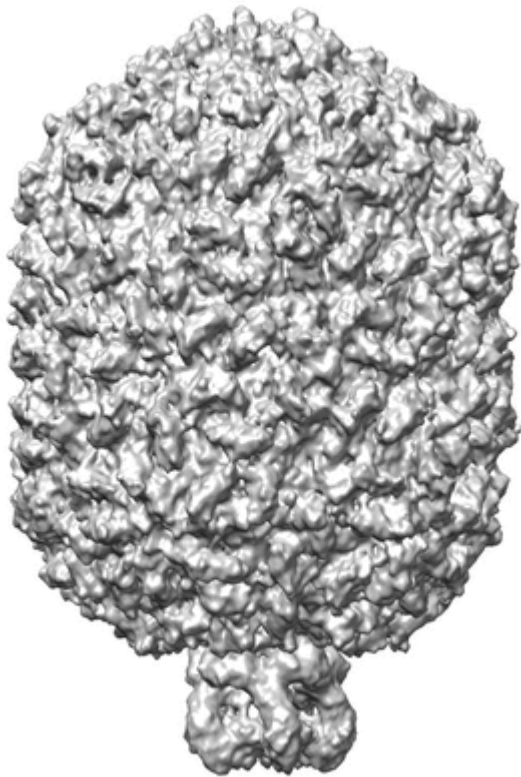


Erinn M. O'Neill, Tatiana S. Mucyn, Jon B. Patteson, Omri M. Finkel, Eui-Hwan Chung, Joshua A. Baccile, Elisabetta Massolo, Frank C. Schroeder, Jeffery L. Dangl, and Bo Li

PNAS published ahead of print September 20, 2018 <https://doi.org/10.1073/pnas.1803779115>

Edited by Sheng Yang He, Michigan State University, East Lansing, MI, and approved August 10, 2018 (received for review March 3, 2018)

[Article](#)[Figures & SI](#)[Info & Metrics](#)[PDF](#)



**Guest lecturer: Dr. Richard Allen White III**

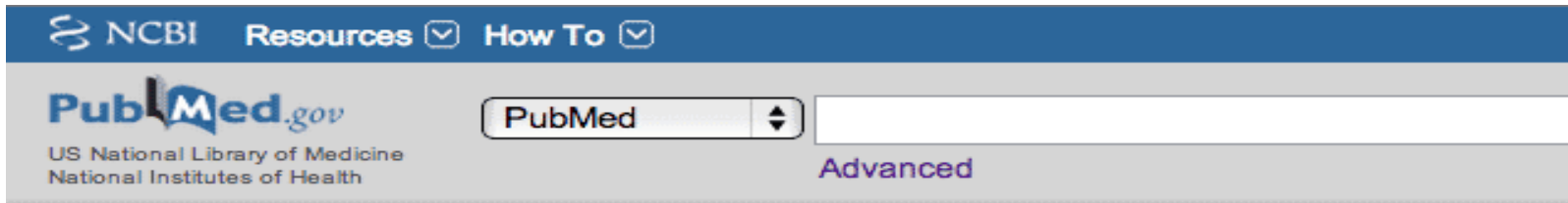
**Post-doc: Friesen lab**

# Learning Objectives

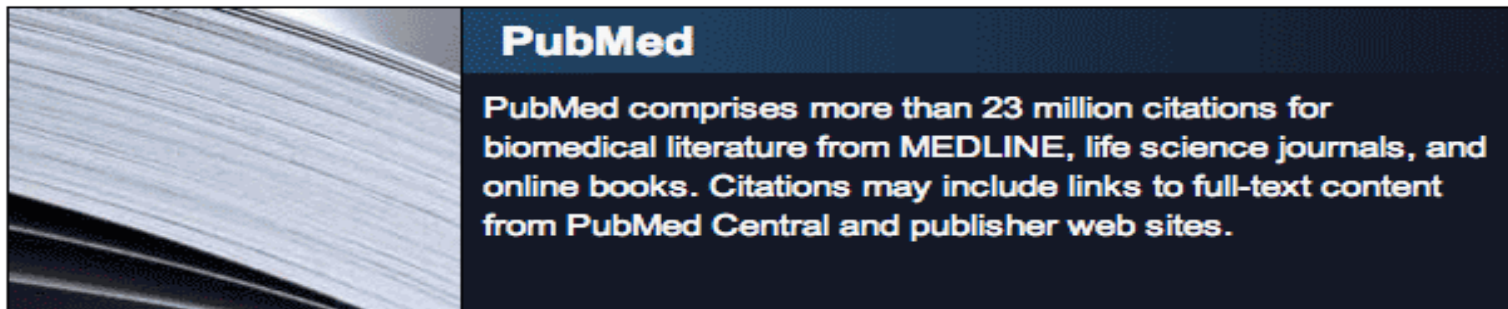
1. NCBI/pubmed walk through
2. BLAST tutorial
3. Global vs. local alignment
4. Multiple sequence alignment
5. Quick tree building in BLAST

# NCBI/Pubmed

Go to [www.pubmed.com](http://www.pubmed.com)



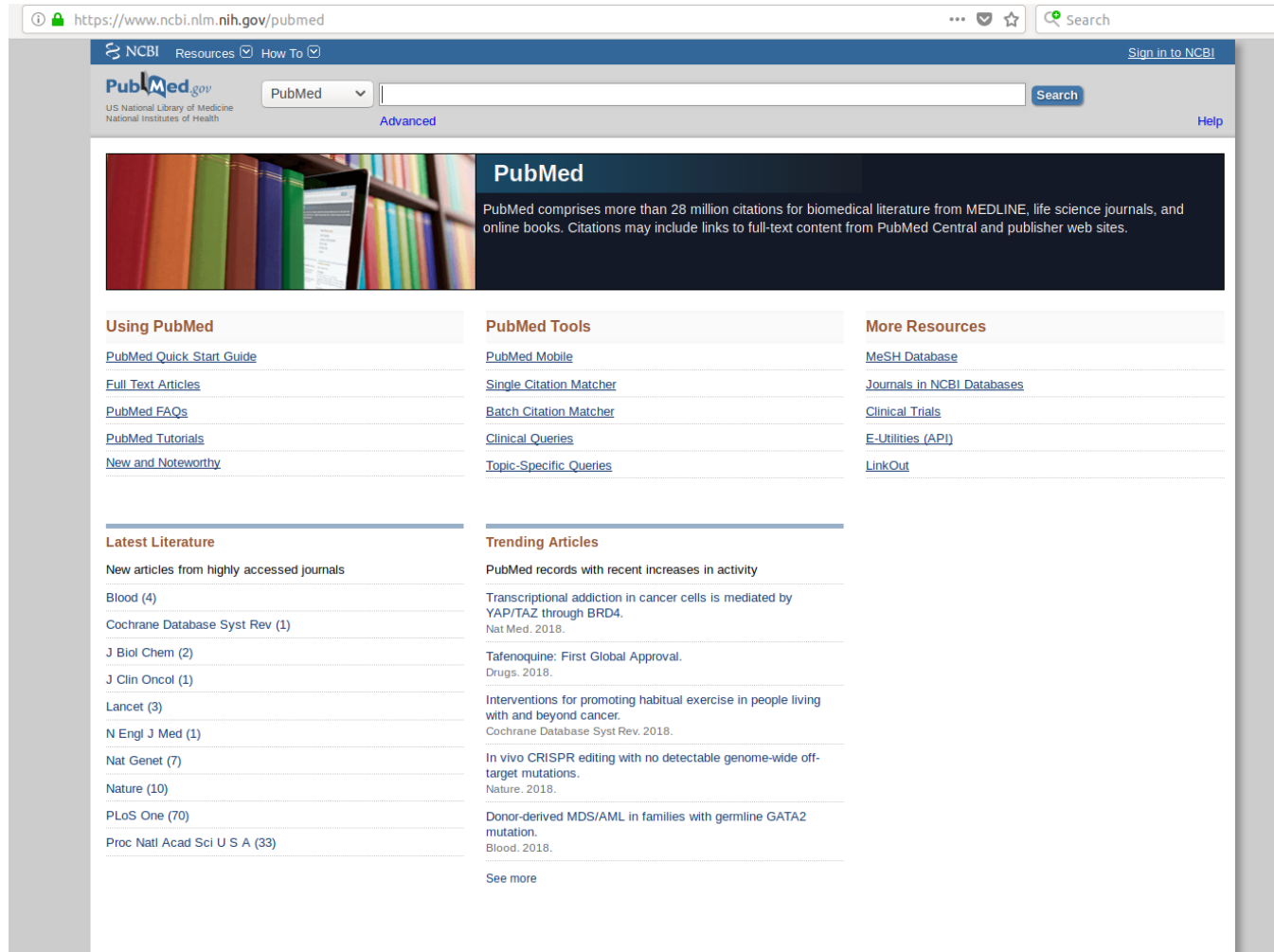
The screenshot shows the top navigation bar of the NCBI website. It features the NCBI logo on the left, followed by the text "Resources" and "How To", each with a dropdown arrow. Below this is the "PubMed.gov" logo, the text "US National Library of Medicine" and "National Institutes of Health", a search box containing the text "PubMed", and a link labeled "Advanced".



**PubMed**

PubMed comprises more than 23 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

# NCBI/Pubmed



The screenshot shows the NCBI PubMed website. At the top, there's a navigation bar with "NCBI", "Resources", and "How To" links. The main header features the "PubMed.gov" logo, a search bar with a "Search" button, and a "Sign in to NCBI" link. Below the header, a large banner image shows a bookshelf with a tablet displaying a PubMed search result. To the right of the image, text states: "PubMed comprises more than 28 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites."

The main content area is divided into three columns:

- Using PubMed:**
  - [PubMed Quick Start Guide](#)
  - [Full Text Articles](#)
  - [PubMed FAQs](#)
  - [PubMed Tutorials](#)
  - [New and Noteworthy](#)
- PubMed Tools:**
  - [PubMed Mobile](#)
  - [Single Citation Matcher](#)
  - [Batch Citation Matcher](#)
  - [Clinical Queries](#)
  - [Topic-Specific Queries](#)
- More Resources:**
  - [MeSH Database](#)
  - [Journals in NCBI Databases](#)
  - [Clinical Trials](#)
  - [E-Utilities \(API\)](#)
  - [LinkOut](#)

Below these columns, there are two sections:

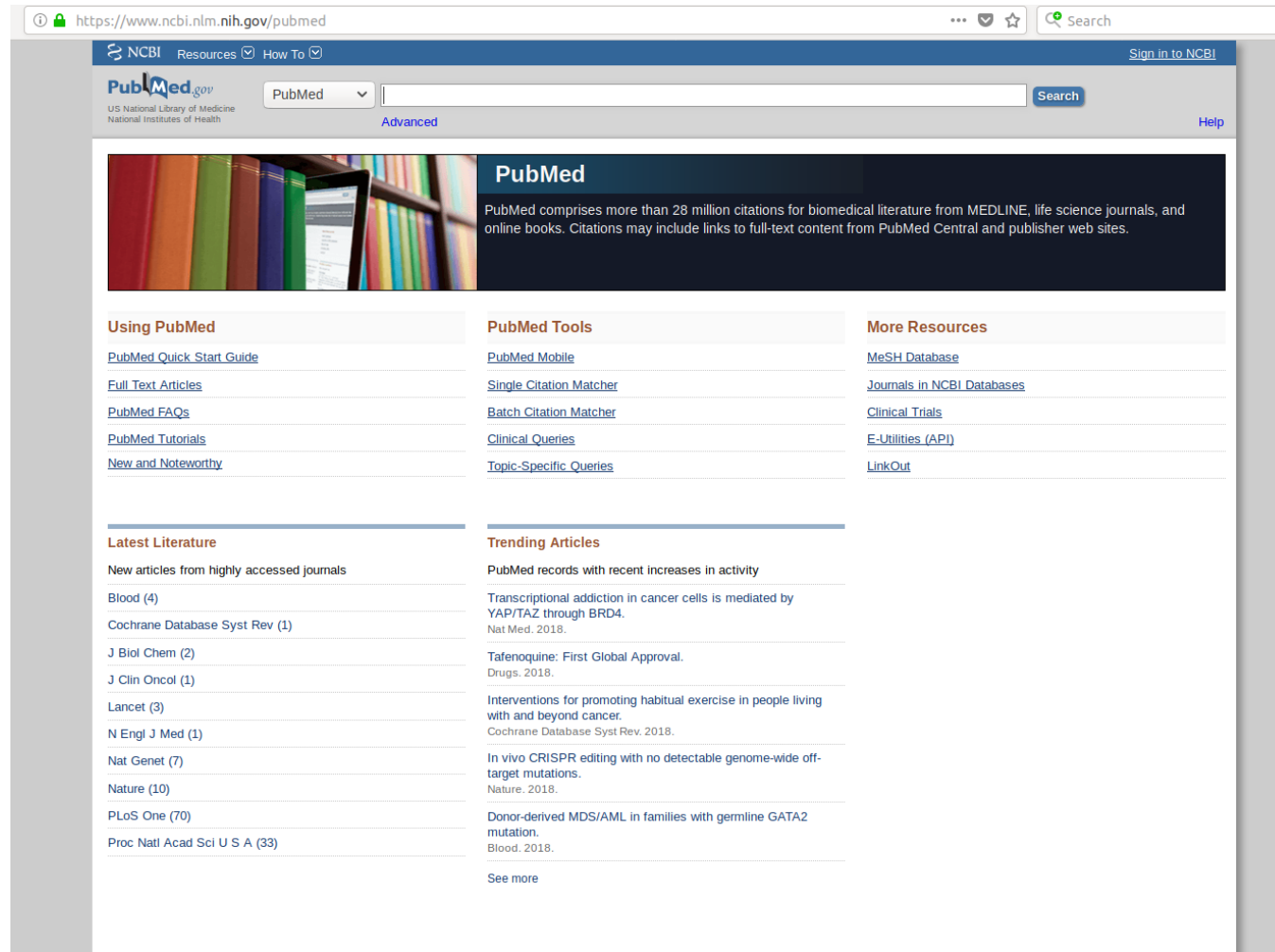
- Latest Literature:**
  - New articles from highly accessed journals
  - Blood (4)
  - Cochrane Database Syst Rev (1)
  - J Biol Chem (2)
  - J Clin Oncol (1)
  - Lancet (3)
  - N Engl J Med (1)
  - Nat Genet (7)
  - Nature (10)
  - PLoS One (70)
  - Proc Natl Acad Sci U S A (33)
- Trending Articles:**
  - PubMed records with recent increases in activity
  - Transcriptional addiction in cancer cells is mediated by YAP/TAZ through BRD4. Nat Med. 2018.
  - Tafenoquine: First Global Approval. Drugs. 2018.
  - Interventions for promoting habitual exercise in people living with and beyond cancer. Cochrane Database Syst Rev. 2018.
  - In vivo CRISPR editing with no detectable genome-wide off-target mutations. Nature. 2018.
  - Donor-derived MDS/AML in families with germline GATA2 mutation. Blood. 2018.
  - [See more](#)

# Type Phevamine A

Dr. Friesen - PLP514 - 2018-Sep-27 - L12

# NCBI/Pubmed

7



Type - *Pseudomonas syringae* pv. Tomato DC

# NCBI/Pubmed

8

The screenshot shows the NCBI Genome website. At the top, there's a navigation bar with 'NCBI', 'Resources', and 'How To'. A search bar is on the right. Below the navigation bar, the 'Genome' section is highlighted. A large banner image shows chromosomes. The main content area is divided into several columns: 'Using Genome' (with links like Help, Browse by Organism, Download / FTP, Download FAQ, Submit a genome), 'Custom resources' (with links like Human Genome, Microbes, Organelles, Viruses, Prokaryotic reference genomes), 'Other Resources' (with links like Assembly, BioProject, BioSample, Genome Data Viewer), 'Genome Tools' (with links like BLAST the Human Genome, Microbial Nucleotide BLAST), 'Genome Annotation and Analysis' (with links like Eukaryotic Genome Annotation, Prokaryotic Genome Annotation, PASC), and 'External Resources' (with links like GOLD, Bacteria Genomes at Sanger, Ensembl). At the bottom, there's a footer with 'You are here: NCBI > Genomes & Maps > Genome' and a 'Support Center' link. The footer also contains several sections: 'GETTING STARTED' (NCBI Education, NCBI Help Manual, NCBI Handbook, Training & Tutorials, Submit Data), 'RESOURCES' (Chemicals & Bioassays, Data & Software, DNA & RNA, Domains & Structures, Genes & Expression, Genetics & Medicine, Genomes & Maps, Homology), 'POPULAR' (PubMed, Bookshelf, PubMed Central, BLAST, Nucleotide, Genome, SNP, Gene), 'FEATURED' (Genetic Testing Registry, GenBank, Reference Sequences, Gene Expression Omnibus, Genome Data Viewer, Human Genome, Mouse Genome, Influenza Virus), and 'NCBI INFORMATION' (About NCBI, Research at NCBI, NCBI News & Blog, NCBI FTP Site, NCBI on Facebook, NCBI on Twitter, NCBI on YouTube, Privacy Policy).

Type - *Pseudomonas syringae* pv. Tomato DC



The screenshot shows the NCBI Genomes database interface. The search bar at the top contains "Pseudomonas syringae pv. tomato str. DC3000". The main content area displays the "Pseudomonas syringae group genomsp. 3" entry. It includes a "Reference genome" section with download options for FASTA, GFF, GenBank, and tabular formats. Below this is the "Organism Overview" section, which provides a summary of the genome assembly and annotation report, including the genome size (6.26051 Mb) and the number of genes (5505). The "Publications" section lists three recent articles related to the pathogen. The "Recent activity" section on the right shows a list of recent searches and genome assemblies.

**Pseudomonas syringae group genomsp. 3**  
**Reference genome: Pseudomonas syringae pv. tomato str. DC3000**  
 Download sequences in FASTA format for **genome**, **protein**  
 Download genome annotation in **GFF**, **GenBank** or **tabular** format  
 BLAST against Pseudomonas syringae group genomsp. 3 **genome**, **protein**  
**All 38 genomes for species:**  
 Browse the **list**  
 Download sequence and annotation from **RefSeq** or **GenBank**

Display Settings: Overview Send to: ▼

**Organism Overview** ; [Genome Assembly and Annotation report \[38\]](#) ; [Genome Tree report \[38\]](#) ; [Plasmid Annotation Report \[11\]](#) ID: 2253

**Pseudomonas syringae group genomsp. 3**  
 Pseudomonas syringae group genomsp. 3 overview

**Lineage:** Bacteria[22966]; Proteobacteria[7223]; Gammaproteobacteria[2706]; Pseudomonadales[356]; Pseudomonadaceae[232]; Pseudomonas[192]; Pseudomonas syringae group[19]; Pseudomonas syringae group genomsp. 3[1]

**Pseudomonas.** Bacteria belonging to the *Pseudomonas* group are common inhabitants of soil and water and can also be found on the surfaces of plants and animals. *Pseudomonas* bacteria are found in nature in a biofilm or in planktonic form. In the latter form the single cell can display an extremely high motility due to its polar [More...](#)

**Summary**

**Sequence data:** genome assemblies: 38; sequence reads: 2 (See [Genome Assembly and Annotation report](#))  
**Statistics:** median total length (Mb): 6.26051  
 median protein count: 5505  
 median GC%: 58.6

**Publications**

1. Adaptation of the pathogen, *Pseudomonas syringae*, during experimental evolution on a native vs. alternative host plant. Meaden S, et al. Mol Ecol 2017 Apr
2. Genome Sequences of Two *Pseudomonas syringae* pv. tomato Race 1 Strains, Isolated from Tomato Fields in California. Thapa SP, et al. Genome Announc 2016 Mar 10
3. Whole-Genome Sequencing of 10 *Pseudomonas syringae* Strains Representing Different Host Range Spectra. Bartoli C, et al. Genome Announc 2015 Apr 30

[More...](#)

**Representative** (genome information for reference and representative genomes)

**Reference genome:**

**Tools**  
 BLAST Genome

**Related information**  
 Assembly  
 BioProject  
 Gene  
 Components  
 Protein  
 PubMed  
 Taxonomy

**Search details**  
 txid223283[Organism:noexp]  
 Search See more...

**Recent activity**  
 Turn Off Clear

- Pseudomonas syringae group genomsp. 3  
Genome
- txid223283[Organism:noexp] (1)  
Genome
- pseudomonas syringae pv. tomato str. dc3000[orgn] (1)  
Taxonomy
- pseudomonas syringae pv. tomato dc3000[orgn] (1)  
Genome
- Pseudomonas syringae pv. tomato DC3000 (681)  
PubMed

Type - download the gff, contig fasta, and tell me the genome size and number of genes on the main chromosome

The screenshot shows the NCBI Genome database interface. The search bar at the top contains the query "monas syringae pv. tomato" and the results show "Pseudomonas syringae group genomosp. 3". The page is divided into several sections: "Tools" (BLAST Genome), "Related information" (Assembly, BioProject, Gene, Components, Protein, PubMed, Taxonomy), "Search details" (txid223283[Organism:noexp]), "Recent activity" (Pseudomonas syringae group genomosp. 3, txid223283[Organism:noexp], pseudomonas syringae pv. tomato str. dc3000[orgn] (1), pseudomonas syringae pv. tomato dc3000[orgn] (1)), and "Publications" (Adaptation of the pathogen, Pseudomonas syringae, during experimental evolution on a native vs. alternative host plant. Meaden S, et al. Mol Ecol 2017 Apr; Genome Sequences of Two Pseudomonas syringae pv. tomato Race 1 Strains, Isolated from Tomato Fields in California. Thapa SP, et al. Genome Announc 2016 Mar 10; Whole-Genome Sequencing of 10 Pseudomonas syringae Strains Representing Different Host Range Spectra. Bartoli C, et al. Genome Announc 2015 Apr 30). The "Summary" section provides sequence data: genome assemblies: 38; sequence reads: 2 (See Genome Assembly and Annotation report); median total length (Mb): 6.26051; median protein count: 5505; median GC%: 58.6. The "Organism Overview" section provides a brief description of the organism and its characteristics.

NCBI Resources How To Sign in to NCBI

Genome Genome Pseudomonas syringae pv. tomato str. DC3000 Search

Create alert Limits Advanced Help

**Pseudomonas syringae group genomosp. 3**  
Reference genome: **Pseudomonas syringae pv. tomato str. DC3000**  
Download sequences in FASTA format for **genome, protein**  
Download genome annotation in **GFF, GenBank** or **tabular** format  
BLAST against Pseudomonas syringae group genomosp. 3 **genome, protein**  
**All 38 genomes for species:**  
Browse the **list**  
Download sequence and annotation from **RefSeq** or **GenBank**

Display Settings: Overview Send to:

**Organism Overview** ; [Genome Assembly and Annotation report \[38\]](#) ; [Genome Tree report \[38\]](#) ; [Plasmid Annotation Report \[11\]](#) ID: 2253

**Pseudomonas syringae group genomosp. 3**  
Pseudomonas syringae group genomosp. 3 overview

**Lineage:** [Bacteria\[22966\]](#); [Proteobacteria\[7223\]](#); [Gammaproteobacteria\[2706\]](#); [Pseudomonadales\[356\]](#); [Pseudomonadaceae\[232\]](#); [Pseudomonas\[192\]](#); [Pseudomonas syringae group\[19\]](#); [Pseudomonas syringae group genomosp. 3\[1\]](#)

**Pseudomonas.** Bacteria belonging to the *Pseudomonas* group are common inhabitants of soil and water and can also be found on the surfaces of plants and animals. *Pseudomonas* bacteria are found in nature in a biofilm or in planktonic form. In the latter form the single cell can display an extremely high motility due to its polar [More...](#)

**Summary**

**Sequence data:** genome assemblies: 38; sequence reads: 2 (See [Genome Assembly and Annotation report](#))

**Statistics:** median total length (Mb): 6.26051  
median protein count: 5505  
median GC%: 58.6

**Publications**

1. Adaptation of the pathogen, *Pseudomonas syringae*, during experimental evolution on a native vs. alternative host plant. Meaden S, et al. Mol Ecol 2017 Apr
2. Genome Sequences of Two *Pseudomonas syringae* pv. tomato Race 1 Strains, Isolated from Tomato Fields in California. Thapa SP, et al. Genome Announc 2016 Mar 10
3. Whole-Genome Sequencing of 10 *Pseudomonas syringae* Strains Representing Different Host Range Spectra. Bartoli C, et al. Genome Announc 2015 Apr 30

[More...](#)

**Representative** (genome information for reference and representative genomes)

Reference genome: **Pseudomonas syringae pv. tomato str. DC3000**

**Tools**

BLAST Genome

**Related information**

Assembly

BioProject

Gene

Components

Protein

PubMed

Taxonomy

**Search details**

txid223283[Organism:noexp]

Search See more...

**Recent activity**

[Turn Off](#) [Clear](#)

[Pseudomonas syringae group genomosp. 3](#) Genome

[txid223283\[Organism:noexp\]](#) (1) Genome

[pseudomonas syringae pv. tomato str. dc3000\[orgn\]](#) (1) Taxonomy

[pseudomonas syringae pv. tomato dc3000\[orgn\]](#) (1) Genome

[Pseudomonas syringae pv. tomato DC3000](#) (1) PubMed

Type - go to the tabular format of the annotation, look for H  
find it

Download the gene as fasta format (as a protein, save the  
sequence).

Protein Details for *Pseudomonas syringae* pv. tomato str. DC3000

Sequences: [chr pDC3000A](#) [pDC3000B](#)

Search by locus, locus tag or protein name

HrpL

Selected by HrpL

Name	Accession	Start	Stop	Strand	GeneID	Locus	Locus tag	Protein product	Length	COG(s)	Protein name
chr	<a href="#">NC_004578.1</a>	1542838	1543392	+	<a href="#">1183040</a>	hrpL	PSPTO_1404	<a href="#">NP_791231.1</a>	184	-	RNA polymerase sigma factor HrpL

You are here: NCBI > Genomes & Maps > Genome

**GETTING STARTED**

- NCBI Education
- NCBI Help Manual
- NCBI Handbook
- Training & Tutorials
- Submit Data

**RESOURCES**

- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Variation

**POPULAR**

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

**FEATURED**

- Genetic Testing Registry
- GenBank
- Reference Sequences
- Gene Expression Omnibus
- Genome Data Viewer
- Human Genome
- Mouse Genome
- Influenza Virus
- Primer-BLAST
- Sequence Read Archive

**NCBI INFORMATION**

- About NCBI
- Research at NCBI
- NCBI News & Blog
- NCBI FTP Site
- NCBI on Facebook
- NCBI on Twitter
- NCBI on YouTube
- Privacy Policy

Type - go to the search bar, look for R  
find it

Download the gene as fasta format (as a protein, save the sequence).

(1) From protein tab,

Type NP\_791231.1, NC\_004578.1

What do you get?

(2) From [www.pubmed.com](http://www.pubmed.com) (main page),

Type NP\_791231.1, NC\_004578.1

What do you get?

What are these numbers called?

# BLAST

## Basic Local Alignment Search Tool -

*J. Mol. Biol.* 1990 Oct 5; **215**(3):403-10 —the primary reference for the BLAST algorithm.

---

### Basic Local Alignment Search Tool

Stephen F. Altschul<sup>1</sup>, Warren Gish<sup>1</sup>, Webb Miller<sup>2</sup>  
Eugene W. Myers<sup>3</sup> and David J. Lipman<sup>1</sup>

<sup>1</sup>*National Center for Biotechnology Information  
National Library of Medicine, National Institutes of Health  
Bethesda, MD 20894, U.S.A.*

<sup>2</sup>*Department of Computer Science  
The Pennsylvania State University, University Park, PA 16802, U.S.A.*

<sup>3</sup>*Department of Computer Science  
University of Arizona, Tucson, AZ 85721, U.S.A.*

(Received 26 February 1990; accepted 15 May 1990)

A new approach to rapid sequence comparison, basic local alignment search tool (BLAST), directly approximates alignments that optimize a measure of local similarity, the maximal segment pair (MSP) score. Recent mathematical results on the stochastic properties of MSP scores allow an analysis of the performance of this method as well as the statistical significance of alignments it generates. The basic algorithm is simple and robust; it can be implemented in a number of ways and applied in a variety of contexts including straight-forward DNA and protein sequence database searches, motif searches, gene identification searches, and in the analysis of multiple regions of similarity in long DNA sequences. In addition to its flexibility and tractability to mathematical analysis, BLAST is an order of magnitude faster than existing sequence comparison tools of comparable sensitivity.

Most important and first  
major computational  
biology tool ever  
created...

Cited over 50,000 times

### 1. Introduction

The discovery of sequence homology to a known protein or family of proteins often provides the first clues about the function of a newly sequenced gene. As the DNA and amino acid sequence databases continue to grow in size they become increasingly useful in the analysis of newly sequenced genes and proteins because of the greater chance of finding such homologies. There are a number of software tools for

(Coulson *et al.*, 1987).

Rapid heuristic algorithms that attempt to approximate the above methods have been developed (Waterman, 1984), allowing large databases to be searched on commonly available computers. In many heuristic methods the measure of similarity is not explicitly defined as a minimal cost set of mutations, but instead is implicit in the algorithm itself. For example, the FASTP program (Lipman & Pearson, 1985; Pearson & Lipman, 1988) first finds locally similar regions between two sequences

# BLAST

## Basic Local Alignment Search Tool

Query sequence: R P P Q G L F

Database sequence: D P P E G V V

└─ Exact match is scanned.

Score: -2 7 7 2 6 1 -1

└─ HSP

Optimal accumulated score =  $7+7+2+6+1 = 23$

Query sequence: PQGEFG

Word 1: PQG

Word 2: QGE

Word 3: GEF

Word 4: EFG

# BLAST

Basic Local Alignment Search Tool - types of sequence analysis

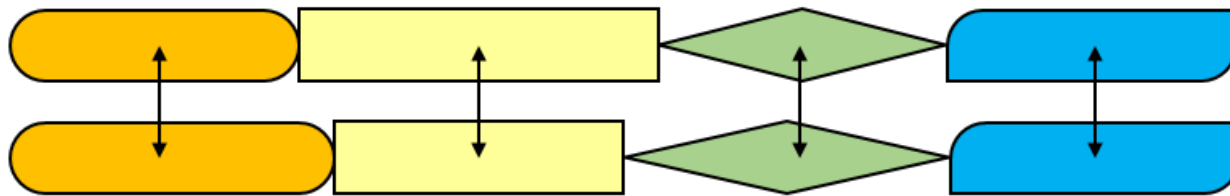
1. Nucleotide Blast (MegaBlast, BlastN, nuc-nuc, db-query)
2. BlastX (translated nucleotide query to protein database)
3. tBlastN (protein query to translated nucleotide database)
4. tBlastX (translated nucleotide query and database)
5. Blastp (protein query and database)

# Alternatives to BLAST

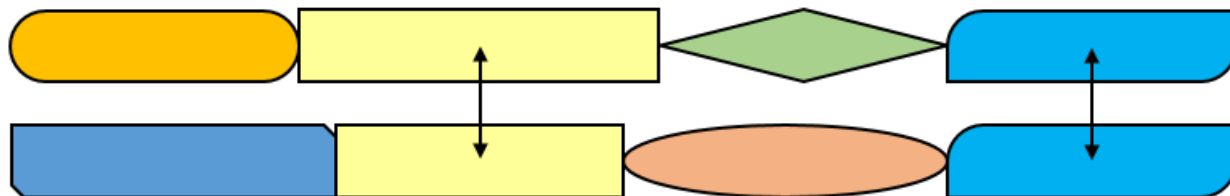
-Pre-blast-

Smith-Waterman algorithm (Local alignment)

Needleman-Wunsch algorithm (Global alignment)



Global Alignment



Local Alignment



# Alternatives to BLAST

-Pre-blast-

Smith-Waterman algorithm (Local alignment)

Needleman-Wunsch algorithm (Global alignment)

Basic differences between SW and NW

	Smith-Waterman algorithm	Needleman-Wunsch algorithm
Initialization	First row and first column are set to 0	First row and first column are subject to gap penalty
Scoring	Negative score is set to 0	Score can be negative
Traceback	Begin with the highest score, end when 0 is encountered	Begin with the cell at the lower right of the matrix, end at top left cell

# Alternatives to BLAST

-Post-blast-

1. BLAT (Blast-like alignment tool)
2. PatternHunter
3. LAST (Local alignment search tool)
4. KLAST
5. Sword (awesome, both fast SW and NW)
6. USEARCH
7. MMseq2
8. DIAMOND
9. Bowtie2, BWA
10. Hmmer (based on HMM)

# BLAST

19

The screenshot shows the NCBI BLAST website. At the top is the BLAST logo and navigation links for NIH, U.S. National Library of Medicine, and NCBI. A search bar is in the top right. The main heading is "Basic Local Alignment Search Tool". Below it, a description states: "BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance." A "Learn more" link is provided. To the right, a "NEWS" box announces a new version (1.4.0) of the BLAST RNA-seq mapping tool, Magic-BLAST, available as of Tuesday, 21 Aug 2018 16:00:00 EST, with a link to "More BLAST news...".

**Web BLAST**

Three main options are shown in boxes:

- Nucleotide BLAST**: nucleotide ► nucleotide (with a DNA double helix icon)
- blastx**: translated nucleotide ► protein (with a right-pointing arrow icon)
- tblastn**: protein ► translated nucleotide (with a left-pointing arrow icon)
- Protein BLAST**: protein ► protein (with a protein ribbon icon)

**BLAST Genomes**

A search bar with the placeholder text "Enter organism common name, scientific name, or tax id" and a "Search" button. Below the bar are links for "Human", "Mouse", "Rat", and "Microbes".

**Standalone and API BLAST**

Three options are listed:

- Download BLAST**: Get BLAST databases and executables (with a download icon)
- Use BLAST API**: Call BLAST from your application (with a server icon)
- Use BLAST in the cloud**: Start an instance at a cloud provider (with a cloud icon)

**Specialized searches**

Go to website - <https://www.ncbi.nlm.nih.gov/BLAST/>

# BLAST

20

BLAST: Basic Local Alignment Search Tool


Home Recent Results Saved Strategies Help

## 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**  
A new version (1.4.0) of the BLAST RNA-seq mapping tool, Magic-BLAST, is now available  
Tue, 21 Aug 2018 16:00:00 EST [More BLAST news...](#)

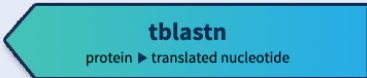
## Web BLAST



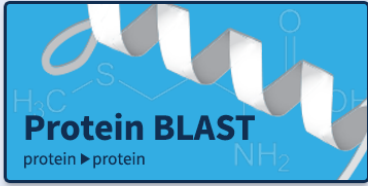
**Nucleotide BLAST**  
nucleotide ► nucleotide



**blastx**  
translated nucleotide ► protein



**tblastn**  
protein ► translated nucleotide




**Protein BLAST**  
protein ► protein

## BLAST Genomes


Enter organism common name, scientific name, or tax id

[Human](#) [Mouse](#) [Rat](#) [Microbes](#)


## Standalone and API BLAST



**Download BLAST**  
Get BLAST databases and executables



**Use BLAST API**  
Call BLAST from your application



**Use BLAST in the cloud**  
Start an instance at a cloud provider

## Specialized searches

Complete a Blastp on your HrpL and RecA sequences

# BLAST

21

https://blast.ncbi.nlm.nih.gov/Blast.cgi

NIH U.S. National Library of Medicine NCBI National Center for Biotechnology Information Sign in to NCBI

**BLAST** » **blastp suite** » RID-USSJZ6G6015 Home Recent Results Saved Strategies Help

**BLAST Results**

[Edit and Resubmit](#) [Save Search Strategies](#) [Formatting options](#) [Download](#) [YouTube](#) [How to read this page](#) [Blast report description](#)

**Job title:** NP\_793795.1 recA protein [[Pseudomonas syringae]...

**RID:** USSJZ6G6015 (Expires on 09-28 12:53 pm)

**Query ID:** lcl|Query\_165272

**Description:** NP\_793795.1 recA protein [[Pseudomonas syringae] pv. tomato str. DC3000]

**Molecule type:** amino acid

**Query Length:** 354

**Database Name:** nr

**Description:** All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects

**Program:** BLASTP 2.8.0+ [Citation](#)

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Related Structures](#) [Multiple alignment](#) [MSA viewer](#)

**New** Analyze your query with [SmartBLAST](#)

**Graphic Summary**

[Show Conserved Domains](#)

Putative conserved domains have been detected, click on the image below for detailed results.

Query seq. Specific hits Superfamilies

recA

Rad51 superfamily

Distribution of the top 100 Blast Hits on 100 subject sequences

Mouse over to see the title, click to show alignments

Color key for alignment scores

<40 40-50 50-80 80-200 >=200

Query

1 70 140 210 280 350

Questions/comments

# BLAST

22

**Descriptions**

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

[Alignments](#) [Download](#) [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> MULTISPECIES: DNA recombination/repair protein RecA [Pseudomonas syringae group]	718	718	100%	0.0	100%	<a href="#">WP_004397594.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas syringae pv. maculicola]	718	718	100%	0.0	100%	<a href="#">KPX75621.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas syringae pv. antirrhini]	718	718	100%	0.0	100%	<a href="#">KPW46983.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas syringae]	717	717	100%	0.0	99%	<a href="#">WP_099264768.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas syringae]	717	717	100%	0.0	99%	<a href="#">WP_060413888.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas syringae pv. helianthi]	717	717	99%	0.0	100%	<a href="#">KPX44783.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas syringae]	716	716	100%	0.0	99%	<a href="#">WP_074842392.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas syringae pv. tagetis]	716	716	99%	0.0	100%	<a href="#">KPY82230.1</a>
<input type="checkbox"/> MULTISPECIES: recombinase RecA [Pseudomonas]	716	716	99%	0.0	100%	<a href="#">WP_024644417.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas syringae]	715	715	100%	0.0	99%	<a href="#">WP_103724441.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas ficusrectae]	714	714	99%	0.0	99%	<a href="#">KPX38027.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas coronafaciens pv. atropurpurea]	714	714	100%	0.0	99%	<a href="#">KPW37033.1</a>
<input type="checkbox"/> DNA recombination/repair protein RecA [Pseudomonas coronafaciens]	714	714	100%	0.0	99%	<a href="#">WP_054989931.1</a>
<input type="checkbox"/> MULTISPECIES: recombinase RecA [Pseudomonas syringae group]	714	714	99%	0.0	99%	<a href="#">WP_007250179.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas coronafaciens pv. oryzae]	714	714	99%	0.0	99%	<a href="#">KPY07991.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas amygdali]	714	714	99%	0.0	99%	<a href="#">KPW35283.1</a>
<input type="checkbox"/> MULTISPECIES: DNA recombination/repair protein RecA [Pseudomonas]	713	713	99%	0.0	99%	<a href="#">WP_004657925.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas coronafaciens]	713	713	99%	0.0	99%	<a href="#">WP_055003380.1</a>
<input type="checkbox"/> MULTISPECIES: recombinase RecA [Pseudomonas syringae group]	713	713	99%	0.0	99%	<a href="#">WP_005891902.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas amygdali pv. morsprunorum]	712	712	99%	0.0	99%	<a href="#">PPS29073.1</a>
<input type="checkbox"/> recombinase RecA [Pseudomonas congelans]	712	712	99%	0.0	99%	<a href="#">WP_096102138.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas savastanoi pv. glycinea]	711	711	99%	0.0	99%	<a href="#">KPX44983.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas syringae pv. corvii]	711	711	99%	0.0	99%	<a href="#">KPW91094.1</a>
<input type="checkbox"/> DNA recombination/repair protein RecA [Pseudomonas savastanoi]	711	711	99%	0.0	99%	<a href="#">WP_004666013.1</a>
<input type="checkbox"/> Protein RecA [Pseudomonas syringae pv. atrofaciens]	711	711	99%	0.0	99%	<a href="#">KPW07061.1</a>
<input type="checkbox"/> MULTISPECIES: DNA recombination/repair protein RecA [Pseudomonas]	710	710	99%	0.0		

[Questions/comment](#)

# BLAST

23

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Related Structures](#) [Multiple alignment](#) [MSA viewer](#)

**New** Analyze your query with [SmartBLAST](#)

## Graphic Summary

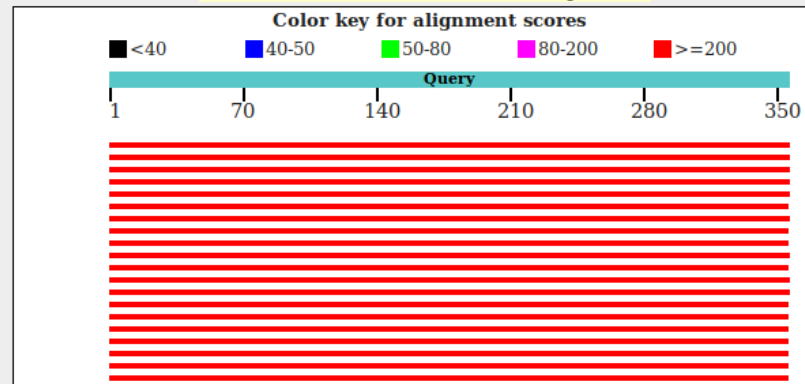
☐ Show Conserved Domains

Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of the top 100 Blast Hits on 100 subject sequences

Mouse over to see the title, click to show alignments



Click the Multiple alignment link, will turn on COBALT

# BLAST

24

https://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi

NIH U.S. National Library of Medicine NCBI National Center for Biotechnology Information Sign in to NCBI

**COBALT** Constraint-based Multiple Alignment Tool Home Recent Results Help

Phylogenetic Tree Edit and Resubmit Back to Blast Results Download

**Multiple Alignment Results - NP\_793795.1 recA protein [[Pseudomonas syringae]... - Cobalt RID UST0HDBZ211 (100 seqs)**

Graphical Overview

processing locations  
Failed to receive reply after 3 tries

Descriptions ☒ Select All [Re-align](#) Alignment parameters

Accession	Description	Links
<input checked="" type="checkbox"/> IcdQuery_165272	NP_793795.1 recA protein [[Pseudomonas syringae] pv. tomato str. DC3000]	
<input checked="" type="checkbox"/> <a href="#">WP_004397594.1</a>	MULTISPECIES: DNA recombination/repair protein RecA [Pseudomonas syringae group]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">KPX75621.1</a>	Protein RecA [Pseudomonas syringae pv. maculicola]	
<input checked="" type="checkbox"/> <a href="#">KPW46983.1</a>	Protein RecA [Pseudomonas syringae pv. antirrhini]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">WP_099264768.1</a>	recombinase RecA [Pseudomonas syringae]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">WP_060413888.1</a>	recombinase RecA [Pseudomonas syringae]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">KPX44783.1</a>	Protein RecA [Pseudomonas syringae pv. helianthi]	
<input checked="" type="checkbox"/> <a href="#">WP_074842392.1</a>	recombinase RecA [Pseudomonas syringae]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">KPY82230.1</a>	Protein RecA [Pseudomonas syringae pv. tagetis]	
<input checked="" type="checkbox"/> <a href="#">WP_024644417.1</a>	MULTISPECIES: recombinase RecA [Pseudomonas]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">WP_103724441.1</a>	recombinase RecA [Pseudomonas syringae]	<a href="#">Related Information</a>
<input checked="" type="checkbox"/> <a href="#">KPX38027.1</a>	Protein RecA [Pseudomonas ficuserectae]	<a href="#">Related Information</a>

Click the Multiple alignment link, will turn on COBALT



# BLAST

25

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Related Structures](#) [Multiple alignment](#) [MSA viewer](#)

**New** Analyze your query with [SmartBLAST](#)

## Graphic Summary

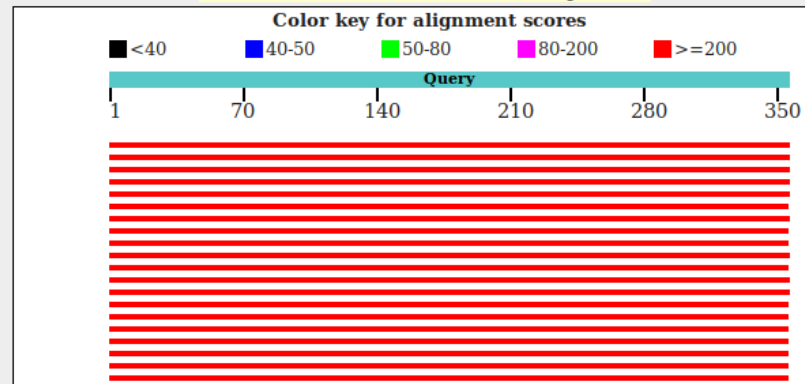
☐ Show Conserved Domains

Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of the top 100 Blast Hits on 100 subject sequences

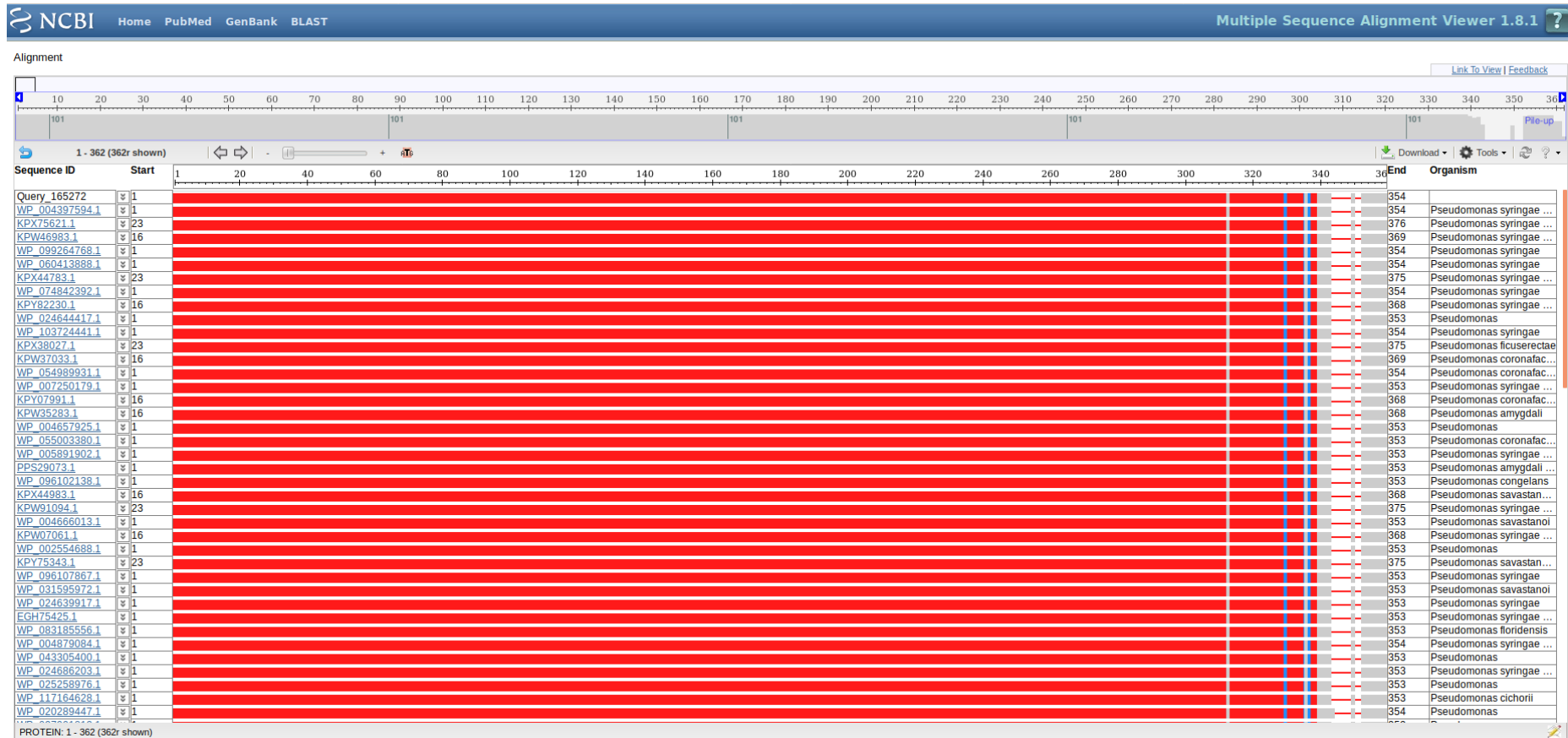
Mouse over to see the title, click to show alignments



Click the MSA viewer

# BLAST

26



# BLAST

27

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Related Structures](#) [Multiple alignment](#) [MSA viewer](#)

**New** Analyze your query with [SmartBLAST](#)

## Graphic Summary

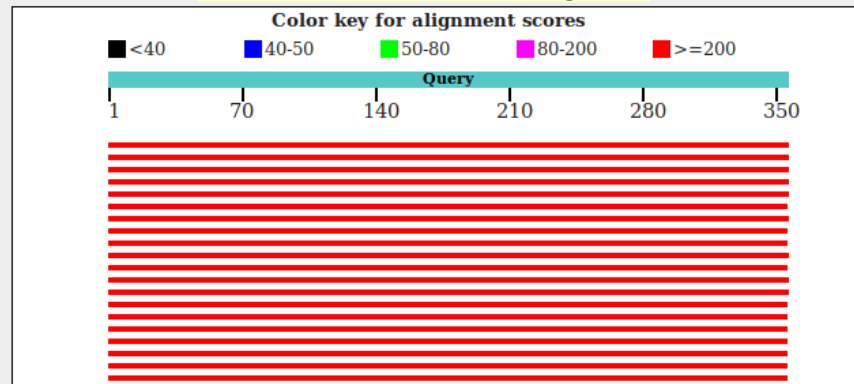
☐ Show Conserved Domains

Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of the top 100 Blast Hits on 100 subject sequences

Mouse over to see the title, click to show alignments



Click the Distance tree results

# BLAST

28

BLAST®

[Home](#) [Recent Results](#) [Saved Strategies](#) [Help](#)

Unexpired BLAST Jobs

Blast Tree View

This tree was produced using BLAST pairwise alignments. [more...](#)

[Reset Tree](#)

BLAST RID [USSJZ6G6015](#)

Query ID [Ic|Query\\_165272](#)

Database [nr](#)

Tree method

Fast Minimum Evolution

Max Seq Difference

0.85

Distance

Grishin (protein)

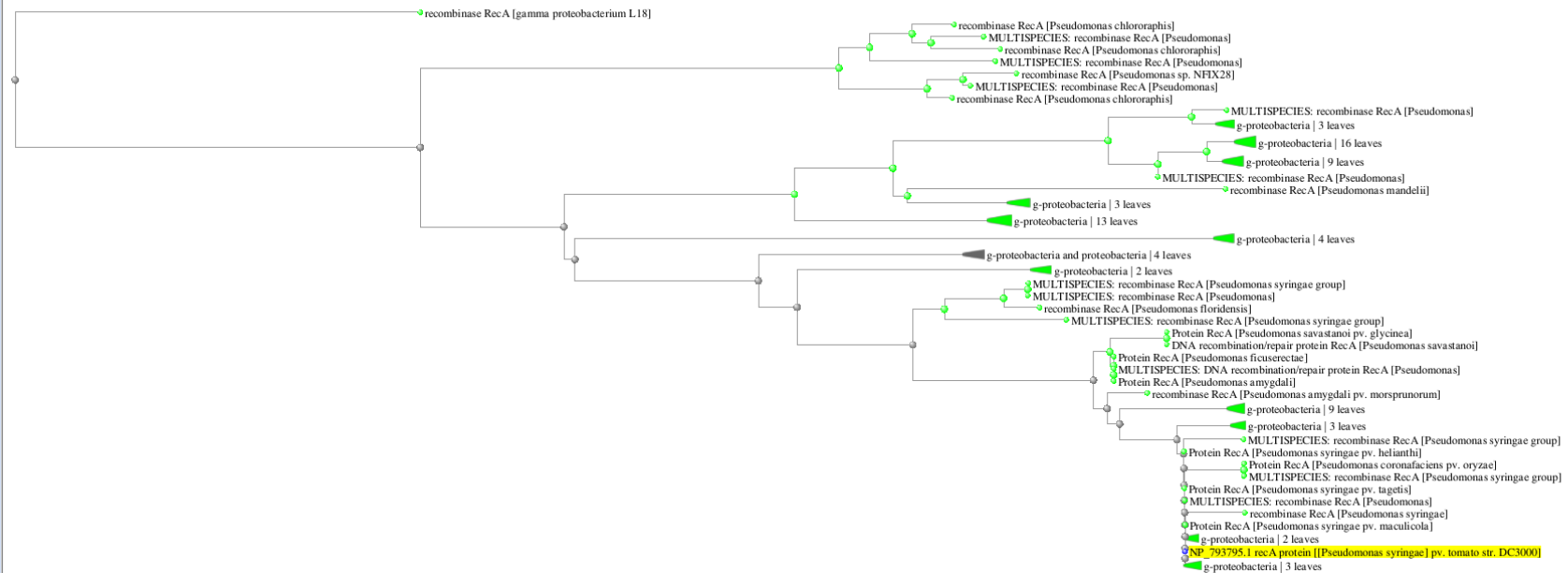
Sequence Label

Sequence Title (if avail)

Find:  all

Mouse over an internal node for a subtree or alignment. Click on tree label to select sequence to download

[Hide legend](#)



Success

Nodes 201(0 selected) View port at (0.0) of 1670x617 0.004