

МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ АСТАНА
Высшая школа естественных наук
ОСНОВЫ БИОИНФОРМАТИКИ
(Практикум)

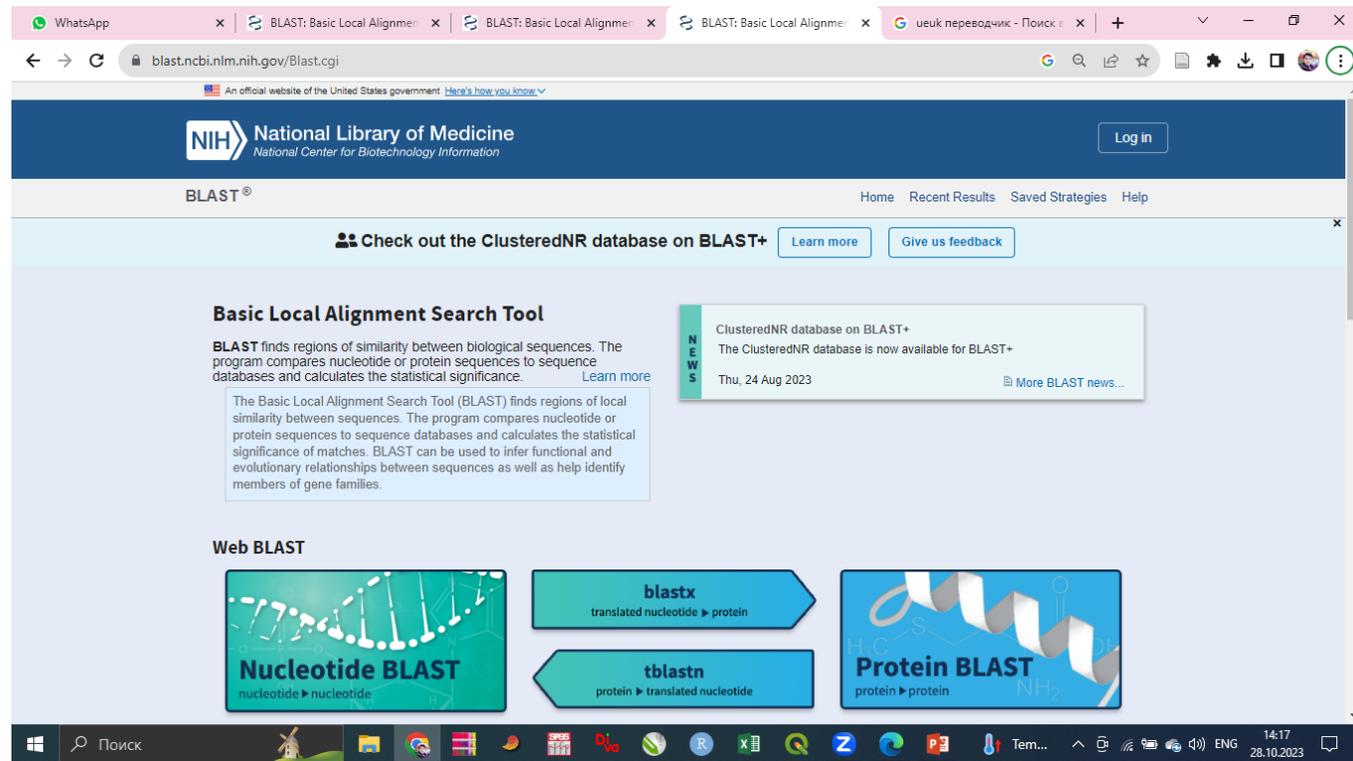
BLAST

ОРАЗОВ АЙДЫН ЕРҒАЛИҰЛЫ
PhD, и.о. ассоциированного профессора Высшей школы
естественных наук международного университета Астана,
Заведующий научно-исследовательской лаборатории
изучения окружающей среды «NatureLaB»

АСТАНА, 2025

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance.

BLAST находит области сходства между биологическими последовательностями. Программа сравнивает нуклеотидные или белковые последовательности с базами данных последовательностей и вычисляет статистическую значимость.



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

Оценка качества находок: E-value

- Для каждой находки вычисляется т.н. **E-value**
- E-value - это математическое ожидание числа находок с таким же или большим весом выравнивания в банке случайных последовательностей того же размера, что и тот банк, в котором велся поиск
- Иными словами, маленькое значение E-value можно интерпретировать как маленькую вероятность того, что находка ошибочная (то есть $E\text{-value} = 0,001 \approx$ вероятность ошибки одна тысячная)

NCBI BLAST

1 Вставить последовательность(и)
2 Выбрать базу данных (Nucleotide collection = GenBank)
3 Выбрать подпрограмму
4

Название организма
Суммарный вес выравниваний
Покрытие запроса
Процент совпадений
Длина находки

Sequences producing significant alignments

select all 92 sequences selected

Описание

Download Select columns Show 100

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

Description	Common Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Cyanea sp. HLC-30135 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Cyanea sp. HL...	1187	1187	100%	0.0	100.00%	658	MG421562.1
<input checked="" type="checkbox"/> Cyanea sp. HLC-30220 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Cyanea sp. HL...	1187	1187	100%	0.0	100.00%	658	MG421848.1
<input checked="" type="checkbox"/> Cyanea sp. 11BIOAK-1621 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	Cyanea sp. 11B...	1187	1187	100%	0.0	100.00%	658	MF121108.1
<input checked="" type="checkbox"/> Cyanea tzetlinii voucher WS3077 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Cyanea tzetlinii	1187	1187	100%	0.0	100.00%	712	KM281994.1
<input checked="" type="checkbox"/> Cyanea tzetlinii voucher WS3036 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Cyanea tzetlinii	1187	1187	100%	0.0	100.00%	721	KM281990.1
<input checked="" type="checkbox"/> Cyanea sp. POG-2014 voucher WS3097 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Cyanea tzetlinii	1186	2372	99%	0.0	100.00%	657	KM983286.1
<input checked="" type="checkbox"/> Cyanea tzetlinii voucher WS3080 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Cyanea tzetlinii	1183	1183	100%	0.0	99.85%	728	KM281997.1

Prunus avium isolate A photosystem Q(B) protein (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His gene, partial sequence; chloroplast

GenBank: OL739702.1

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

LOCUS OL739702 365 bp DNA linear PLN 15-NOV-2022

DEFINITION Prunus avium isolate A photosystem Q(B) protein (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His gene, partial sequence; chloroplast.

ACCESSION OL739702

VERSION OL739702.1

KEYWORDS .

SOURCE chloroplast Prunus avium (sweet cherry)

ORGANISM [Prunus avium](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae; Pentapetalae; rosids; fabids; Rosales; Rosaceae; Amygdaloideae; Amygdaleae; Prunus.

REFERENCE 1 (bases 1 to 365)

AUTHORS Wu,B.-H., Yang,H.-J., Shi,S., Li,F.-F. and Cui,D.-F.

TITLE Prunus sunhangii is a new synonym of P. itosakura var. ascendens

JOURNAL Taiwania 67 (2), 280-284 (2022)

Change region shown

Customize view

Analyze this sequence[Run BLAST](#)[Pick Primers](#)[Highlight Sequence Features](#)[Find in this Sequence](#)**Related information**[Protein](#)[Taxonomy](#)[PopSet](#)**Recent activity**[Turn Off](#) [Clear](#)

1

>OL739702.1 Prunus avium isolate A photosystem Q(B) protein (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His gene, partial sequence; chloroplast

```
AATGCTCATAATTTCCCTCTAGATCTAGCTGCTGTTGAAGTTCCATCTATAAATGGATAAGACTTTGGTC
TTAGTATATACGAGTTCTTGAAAGTAAAGGAGCAATAATAAATTTCTTGTTATATCAAGAGGTTTGGTAT
TGCTCCTTTACTATACTATTTAGTATTTTTTTTTTATTACTACTTAACTTTACTATATTGTTCTATATTTT
TTTTTAGTTAATTTACTACAAAATTTAATTAATAAATATTAATAAAGTTTCAGTTTATTTTATGTTGTATTTTA
TCTTACAAGTAATGATAAATAGTGTAATATTTGTAATAGTACAAGGGGCGGATGTAGCCAAGTGGATCA
AGGCAGTGGATTGTG
```

2

```
AATGCTCATAATTTCCCTCTAGATCTAGCTGCTGTTGAAGTTCCATCTATAAATGGATAAGACTTTGGTCTTAGTATATACG---
AGTTCTTGAA--AGTAAAGGAGCAATAATAAATTTCTTGTTATATCAAGAGGTTTGGTATTGCTCCTTTACT-----ATACTATTTA-----
-----GTATTTTTTTTATTAC---TACTTAACTTTACT---
ATATTGTTCTATATTTTTTTTTTAGTTAATTTACTACAAAATTTAATT-----AAAATAT-
TAAAAGTTTCAGTTTATTTTATGTTGTATTTT-ATCTTACAAGT-AATGATAAATAGTGTAATATTT-----
GTAATAGTACAAGGGGCGGATGTAGCCAAGTGGATCAAGGCAGTGGATTGTG
```

100 sequences selected ?

[Download](#) ▼ [GenBank](#) [Graphics](#)**Prunus avium cultivar Summit chloroplast, complete genome**Sequence ID: [NC_044701.1](#) Length: 157886 Number of Matches: 1[See 1 more title\(s\)](#) ▼ [See all Identical Proteins\(IPG\)](#)**Range 1: 98 to 462** [GenBank](#) [Graphics](#)▼ [Next Match](#) ▲ [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
675 bits(365)	0.0	365/365(100%)	0/365(0%)	Plus/Minus
Query 1	AATGCTCATAATTTCCCTCTAGATCTAGCTGCTGTTGAAGTTCATCTATAAATGGATAA	60		
Sbjct 462	AATGCTCATAATTTCCCTCTAGATCTAGCTGCTGTTGAAGTTCATCTATAAATGGATAA	403		
Query 61	GACTTTGGTCTTAGTATATACGAGTTCCTGAAAGTAAAGGAGCAATAAATAATTTCTTGT	120		
Sbjct 402	GACTTTGGTCTTAGTATATACGAGTTCCTGAAAGTAAAGGAGCAATAAATAATTTCTTGT	343		
Query 121	TATATCAAGAGGTTTGGTATTGCTCCTTActatactatttagtatttttttatttact	180		
Sbjct 342	TATATCAAGAGGTTTGGTATTGCTCCTTACTATACTATTTAGTATTTTTTTTATTACT	283		
Query 181	acttaactttactatattgttctatatttttttagttaatttactacaaaatttaatt	240		
Sbjct 282	ACTTAACTTTACTATATTGTTCTATATTTTTTTTAGTAAATTTACTACAAAATTTAATT	223		
Query 241	aaaatattaaaagtttcagtttattttaTGTGTATTTTATCTTACAAGTAATGATAAAT	300		
Sbjct 222	AAAATATTAAGTTTCAGTTTATTTTATGTTGATTTTATCTTACAAGTAATGATAAAT	163		
Query 301	AGTGTAATATTTGTAATAGTACAAGGGGCGGATGTAGCCAAGTGGATCAAGGCAGTGGA	360		
Sbjct 162	AGTGTAATATTTGTAATAGTACAAGGGGCGGATGTAGCCAAGTGGATCAAGGCAGTGGA	103		
Query 361	TTGTG 365			
Sbjct 102	TTGTG 98			

ССЫЛКИ:

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

<https://lifemap-ncbi.univ-lyon1.fr/>

<https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BB%D1%8B%D0%BD%D1%8C>

<https://ru.wikipedia.org/wiki/%D0%9E%D0%B4%D1%83%D0%B2%D0%B0%D0%BD%D1%87%D0%B8%D0%BA>

<https://xn--80abvyzg.xn--p1ai/%D0%B1%D0%B8%D0%BE%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0/#11>