

生物情報関連

当日使用スライド(抜粋)

当日使用したスライド資料を次頁以降に示す。
今回の実習で用いたツールは以下の通りです。

UniProt (Universal Protein Resource)

<https://www.uniprot.org/>

タンパク質配列および機能に関する包括的データベースである。手動注釈情報や自動注釈情報を統合しており、タンパク質の配列、構造、機能、翻訳後修飾、相互作用など、多角的な情報を提供している。学術研究やバイオインフォマティクス解析における標準的リソースとして広く利用される。

BLAST (Basic Local Alignment Search Tool)

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

配列データベース中から、クエリ配列に類似する配列を高速に検索するためのプログラム群である。局所アライメント手法を用いて、部分的にでも高い一致を示す領域を特定する。配列の同定、機能予測、系統発生解析など、幅広い生物学的解析に用いられる。

ClustalW

<https://www.genome.jp/tools-bin/clustalw>

複数の生物配列（タンパク質またはヌクレオチド配列）を同時にアライメント（整列）するためのツールである。系統樹作成の前処理や保存したアライメントの可視化にも活用される。

バイオフィーマティクス基礎

これまでのまとめ

DNAの構成要素【ヌクレオチド】（4種）

A T G C

タンパク質の構成要素【アミノ酸】（20種）

A F K P T C G L Q V

D H M R W E I N S Y

遺伝子配列の例

ヒト α -アミラーゼ（だ液アミラーゼ）

ATGAAGCTCTTTTGGTTGCTTTTCACCATTGGGTTCTGCTGGGCTCAGTATTCCTCAAATACACAACAAGGACGAACATCTATTGTTTCATCTGTTT
GAATGGCGATGGGTTGATATTGCTCTTGAATGTGAGCGATATTTAGCTCCCAAGGGATTTGGAGGGGGTTCAGGTCTCTCCACCAAATGAAAATGTT
GCCATTCACAACCCTTTTCAGACCTTGGTGGGAAAGATACCAACCAGTTAGCTATAAATTATGCACAAGATCTGGAAATGAAGATGAATTTAGAAAC
ATGGTGACTAGATGCAACAATGTTGGGGTTCGTATTTATGTGGATGCTGTAATTAATCATATGTGTGGTAATGCTGTGAGTGCAGGAACAAGCAGT
ACCTGTGGAAGTTACTTCAACCCTGGAAGTAGGGACTTTCCAGCAGTCCCATATTCTGGATGGGATTTTAAATGATGGTAAATGTAAAACTGGAAGT
GGAGATATCGAGAACTATAATGATGCTACTCAGGTCAGAGATTGTCGTCTGTCTGGTCTTCTCGATCTTGCACTGGGGAAGGATTATGTGCGTTCT
AAGATTGCCGAATATATGAACCATCTCATTGACATTGGTGTTCAGGGTTCAGAATTGATGCTTCCAAGCACATGTGGCCTGGAGACATAAAGGCA
ATTTTGGACAAACTGCATAATCTAAACAGTAACTGGTTCCCGGAAGGTAGTAAACCTTTTCATTTACCAGGAGGTAATTGATCTGGGTGGTGAGCCA
ATTAAAAGCAGTGACTACTTTGGTAATGGCCGGGTGACAGAATTCAAGTATGGTGCAAACTCGGCACAGTTATTCGCAAGTGGAATGGAGAGAAG
ATGTCTTACTTAAAGAACTGGGGAGAAGGTTGGGGTTTTCATGCCTTCTGACAGAGCGCTTGTCTTTGTGGATAACCATGACAATCAACGAGGACAT
GGCGCTGGAGGAGCCTCTATACTTACCTTCTGGGATGCTAGGCTGTACAAAATGGCAGTTGGATTTATGCTTGCTCATCCTTATGGATTTACACGA
GTAATGTCAAGCTACCGTTGGCCAAGATATTTTGAAAATGGAAAAGATGTTAATGATTGGGTTGGGCCACCAAATGATAATGGAGTAACTAAAGAA
GTTACTATTAATCCAGACACTACTTGTGGCAATGACTGGGTCTGTGAACATCGATGGCGCCAAATAAGGAACATGGTTAATTTCCGCAATGTAGTG
GATGGCCAGCCTTTTACAAACTGGTATGATAATGGGAGCAACCAAGTGGCTTTTGGGAGAGGAAACAGAGGATTCATTGTTTTCAACAATGATGAC
TGGACATTTTCTTTAACTTTGCAAACTGGTCTTCCTGCTGGCACATACTGTGATGTCATTTCTGGAGATAAAATTAATGGCAACTGCACAGGCATT
AAAATCTACGTTTCTGATGATGGCAAAGCTCATTTTTCTATTAGTAACTCTGCTGAAGATCCATTTATTGCAATTCATGCTGAATCTAAATTGTAA

アミノ酸配列の例

ヒト α -アミラーゼ（だ液アミラーゼ）

MKLFWLLFTIGFCWAQYSSNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPP
NENVAIHNPFRPWWERYQPVSYKLCTRSGNEDEFNMVTRCNNVGVRIYVDAVINHMCGN
AVSAGTSSTCGSYFNPGSRDFPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLSGL
LDLALGKDYVRISKIAEYMNHLIDIGVAGFRIDASKHMWPGDIKAILDKLHNLNSNWFPEG
SKPFIYQEVIDLGGPEIKSSDYFGNGRVTEFKYGAKLGTVIRKWNGEKMSYLKNWGEKGW
FMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFM LAHPYGFTRVMSSYRWP
RYFENGKDVNDWVGPPNDNGVTKEVTINPDTTTCGNDWVCEHRWRQIRNMVNFRNVVDGQP
FTNWYDNGSNQVAFGRGNRGFIVFNDDWTFSLTLQTGLPAGTYCDVISGDKINGNCTGI
KIYVSDDGKAHFSISNSAEDPFIAIHAESKL

配列が省略表記である利点

配列の比較が簡単になる

MKLFWLLFT**I**GFCWAQY
MKLFWLLFT**F**GFCWAQY

計算処理に最適

軽量なデータ構造 アミノ酸配列1万個 → テキストなら約10KBほど

高速な比較&解析 ソフトウェアで短時間での解析

アルゴリズム開発がしやすい

開発されている
様々なアルゴリズム

BLAST (Basic Local Alignment Search Tool)

類似タンパク質を大規模データベースから高速に検索するローカルアラインメントツール

ClustalW

複数配列を整列させ、進化的関係や保存領域を調べる多重アラインメントツール

HMMER

隠れマルコフモデル (HMM) を用いて、特定の配列モチーフや機能ドメインを検出

AlphaFold

アミノ酸配列からタンパク質の三次元構造を予測する機械学習ベースの構造予測ツール

BLAST

類似タンパク質を大規模データベースから**高速に検索する**ローカルアラインメントツール

BLAST

0) 検索するアミノ酸配列を手に入れよう！

0-1) 検索エンジンで“UniProt”と入力して、検索する。



BLAST

0) 検索するアミノ酸配列を手に入れよう！

0-2) 以下の画面のようなサイトが表示されたら、クリックする。



UniProt

<https://www.uniprot.org>



UniProt

UniProt - Exploring protein sequence and functional information. Join speakers from **UniProt** as they explore this data resource of protein sequence and ...

Proteins (UniProtKB)

Entry, Entry Name, Protein Names, Gene Names, Organism ...



BLAST

0) 検索するアミノ酸配列を手に入れよう！

0-4) サイト内の検索欄に“AMY1A”と入力して、検索する。

*AMY1Aが今回対象にするアミラーゼのタンパク質名です。

Find your protein

UniProtKB ▾

AMY1A|

✕ Advanced | List

Search

Examples: Insulin, APP, Human, P05067, organism_id:9606

BLAST

- 0) 検索するアミノ酸配列を手に入れよう！
- 0-5) 検索結果で “P0DUB6” をクリックする。

UniProtKB 28 results <small>or search "AMY1A" as a Gene Name or Protein Name</small>						
<div> Tools ▾ Download (28) Add View: Cards <input type="radio"/> Table <input checked="" type="radio"/> Customize columns Share ▾</div>						
<input type="checkbox"/> Entry ▲	Entry Name ▲	Protein Names ▲	Gene Names ▲	Organism ▲	Length ▲	
<input type="checkbox"/> P0DUB6	AMY1A_HUMAN	Alpha-amylase 1A [...]	AMY1A, AMY1	Homo sapiens (Human)	511 AA	
<input type="checkbox"/> P00687	AMY1_MOUSE	Alpha-amylase 1 [...]	Amy1, Amy-1-a, Amy1a	Mus musculus (Mouse)	511 AA	

BLAST

0) 検索するアミノ酸配列を手に入れよう！

0-6) 左側下部にある “Sequence” をクリックする。

Function

Names & Taxonomy

Subcellular Location

Disease & Variants

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence

Similar Proteins

P0DUB6 · AMY1A_HUMAN

Proteinⁱ

Alpha-amylase 1A

Amino acids

511 ([go to sequence](#))

Geneⁱ

AMY1A

Protein existenceⁱ

Evidence at protein level

Statusⁱ

UniProtKB reviewed (Swiss-Prot)

Annotation scoreⁱ

(5/5)

Organismⁱ

Homo sapiens (Human)

Entry

Variant viewer160

Feature viewer

Genomic coordinates

Publications

External links

History

Tools ▾

Download

Add

Add a publication

Entry feedback

Functionⁱ

Calcium-binding enzyme that initiates starch digestion in the oral cavity (PubMed:[12527308](#)).

Catalyzes the hydrolysis of internal (1->4)-alpha-D-glucosidic bonds, yielding a mixture of maltose, isomaltose, small amounts of glucose as well as small linear and branched oligosaccharides called dextrins (PubMed:[12527308](#)). 1 Publication

BLAST

0) 検索するアミノ酸配列を手に入れよう！

0-7) “Copy sequence” をクリックすると、アミノ酸配列がPC内に記憶される。

Sequenceⁱ

Sequence statusⁱ | Complete

Sequence processingⁱ | The displayed sequence is further processed into a mature form.

See also sequence in UniParc or sequence clusters in UniRef

Tools ▾ Download Add Highlight ▾ Copy sequence

Length 511

Mass (Da) 57,768

Last updated 2020-12-02 v1

MD5 Checksumⁱ 2E9FF582F8D1ED489F0C6D5662C4898D

102030405060708090

MKLFWLLFTI GFCWAQYSSN TQQGRTSIVH LFEWRWVDIA LECERYLAPK GFGGVQVSPP NENVAIHNPF RPWWERYQPV SYKLCTRSGN

100110120130140150160170180

EDEFNRNMVTR CNNVGVRIYV DAVINHMCGN AVSAGTSSTC GSYFNPGSRD FPAVPYSGWD FNDGKCKTGS GDIENYNDAT QVRDCRLSGL

190200210220230240250260270

LDLALGKDYV RSKIAEYMNH LIDIGVAGFR IDASKHMWPG DIKAILDKLH NLNSNWFPEG SKPFIYQEV I DLGGEPIKSS DYFGNGRVTE

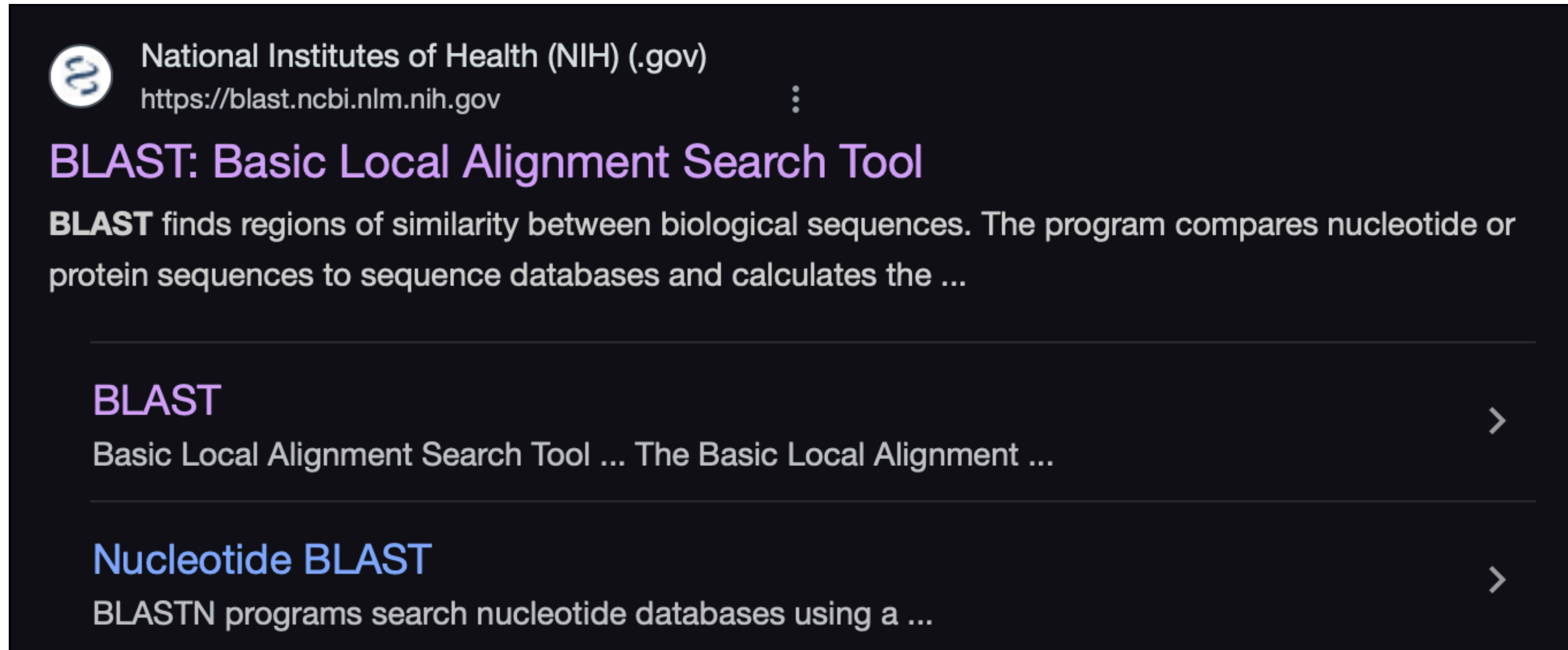
BLAST

1) 検索エンジンで“BLAST”と入力して、検索する。



BLAST

2) 以下の画面のようなサイトが表示されたら、クリックする。



BLAST

3) 画面の右側にある“Protein BLAST”をクリックする。

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

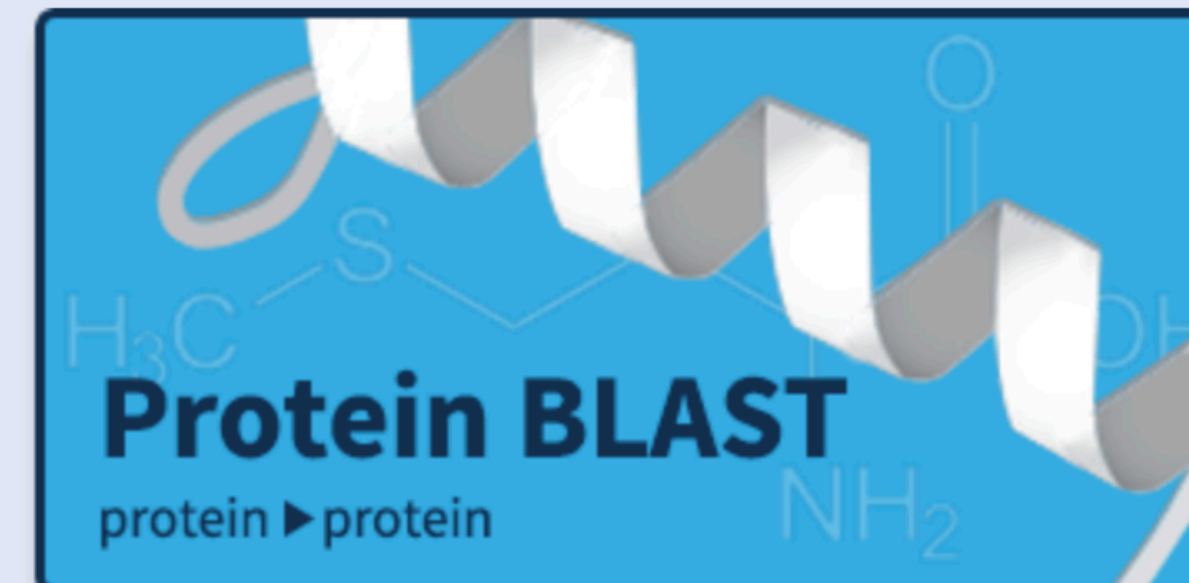
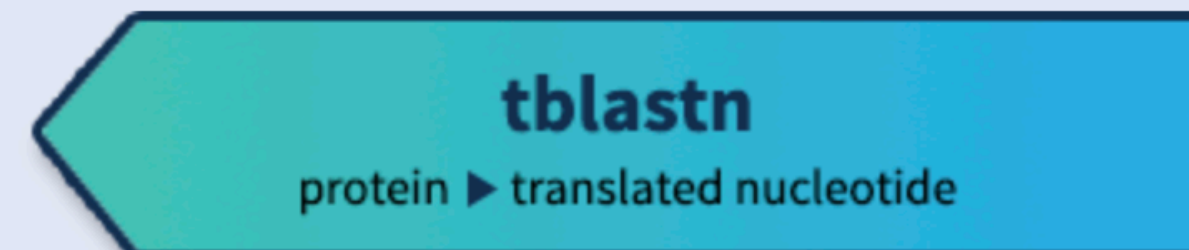
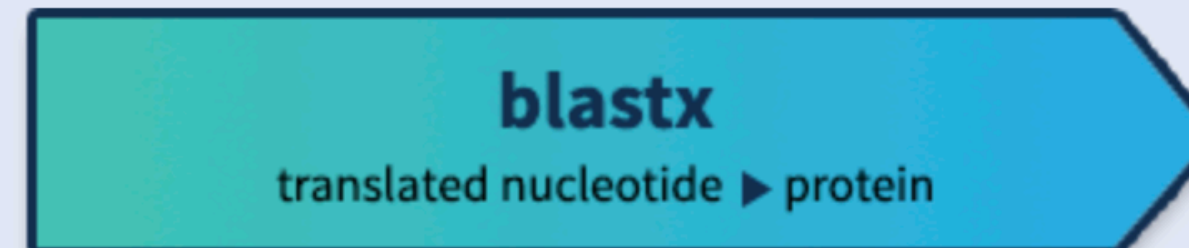
Mon, 17 Mar 2025

Improvements include upgrading to GCP Artifact Registry and better handling of job completion status in kubernetes version 1.30+.

ElasticBLAST 1.4.0 is now available!

[More BLAST news...](#)

Web BLAST



BLAST

- 4) 画面の上部にある “Enter accession number(s), gi(s), or FASTA sequence(s)” の欄内に、コピーした配列をペーストする。

BLAST - programs search protein databases using a protein query.

Enter Query Sequence

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

Query subrange [?](#)

From

To

Or, upload file [?](#)

Job Title

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

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (nr etc.): ☐ Experimental databases

[?](#)

Organism Optional ☐ exclude [Add organism](#)

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

Exclude Optional ☐ Models (XM/XP) ☐ Non-redundant RefSeq proteins (WP) ☐ Uncultured/environmental sample sequences

Program Selection

Algorithm

☐ Quick BLASTP (Accelerated protein-protein BLAST)

☒ blastp (protein-protein BLAST)

☐ PSI-BLAST (Position-Specific Iterated BLAST)

☐ PHI-BLAST (Pattern Hit Initiated BLAST)

☐ DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

Choose a BLAST algorithm [?](#)

[BLAST](#) Search **database nr** using **Blastp (protein-protein BLAST)**

☐ Show results in a new window

BLAST

- 4) 画面の上部にある “Enter accession number(s), gi(s), or FASTA sequence(s)” の欄内に、コピーした配列をペーストする。
- 5) 画面の中央部にある “Organism” の欄内に、“Aspergillus oryzae” と入力すると、候補がたくさん出てくるので、“Aspergillus oryzae RIB40” をクリックする。
- 6) 画面下部の “BLAST” をクリックする。

BLASTP programs search protein databases using a protein query

Enter Query Sequence

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

FVDNHDNQRGHGAGGASILTFWDARLYKMAVGFMALHPYGFTRVMSSYRW
PRYFENGKDVNDWVGPPNDNGVTKEVTINPDTCGNDWVCEHRWRQIRN
MVNFRNVVDGQPFTNWYDNGSNQVAFGRGNRGFVFNDDWTFSLTLQTG
LPAGTYCDVISGDKINGNCTGIKIYVSDDGKAHFSISNSAEDPFIAHAESKL

Query subrange [?](#)

From

To

Or, upload file [?](#)

Job Title

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

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (nr etc.): ☐ Experimental databases

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

Organism [Optional](#)

Exclude [Optional](#)

Aspergillus oryzae ☐ exclude [Add organism](#)

Aspergillus oryzae (taxid:5062) [a will be shown](#) [?](#)

Aspergillus oryzae var. brunneus (taxid:332754)

Aspergillus oryzae RIB40 (taxid:510516)

Aspergillus oryzae 3.042 (taxid:1160506)

Aspergillus oryzae 100-8 (taxid:1197718)

Aspergillus oryzae var. effusus (taxid:1119045)

Aspergillus oryzae var. oryzae (Ahlb.) Cohn, 1884 (taxid:90341)

Aspergillus oryzae var. oryzae (taxid:90341)

☐ Uncultured/environmental sample sequences

Program Selection

Algorithm

Choose a BLAST algorithm [?](#)

[BLAST](#)

Search **database nr** using **Blastp (protein-protein BLAST)**

☐ Show results in a new window

BLAST

Your search is limited to records that include: Aspergillus oryzae RIB40 (taxid:510516)

Job Title

Protein Sequence

RID

[1K0ECWPW013](#) Search expires on 05-07 11:28 am [Download All](#) ▼

Program

BLASTP ? [Citation](#) ▼

Database

nr [See details](#) ▼

Query ID

lcl|Query_7283138

Description

unnamed protein product

Molecule type

amino acid

Query Length

511

Other reports

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

Filter Results

Organism

only top 20 will appear

Type common name, binomial, taxid or group name

+ Use up and down arrows to choose an item from the autocomplete.

Percent Identity

E value

Query Coverage

to

to

to

Filter

Reset

Compare these results against the new Clustered nr database ?

BLAST

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download ▼

Select columns ▼

Show

100 ▼

?

select all

3 sequences selected

GenPept

Graphics

Distance tree of results

Multiple alignment

MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	RecName: Full=Alpha-amylase A type-3; AltName: Full=1,4-alpha-D-glucan glucanohydrolase; AltName: Full=Taka...	Aspergillus oryza...	58.2	58.2	35%	3e-09	28.04%	499	P0C1B4.1
<input checked="" type="checkbox"/>	uncharacterized protein AO090023000944 [Aspergillus oryzae RIB40]	Aspergillus oryza...	57.8	57.8	35%	4e-09	28.04%	499	XP_001821436.1
<input checked="" type="checkbox"/>	Chain A, Alpha-amylase A type-1/2 [Aspergillus oryzae RIB40]	Aspergillus oryza...	57.0	57.0	35%	5e-09	28.04%	478	3VX0_A

複数のタンパク質比較

ヒトのアミラーゼと
より似ているアミラーゼはどれ？

ヒト *Homo sapiens*

MKLFWLLFTIGFCWAQYSSNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVIAIHNPFRPWWERYQPVSYKLCTRSGNEDEFNMVTRCINN VGVRIYVDAVINHMCN
AVSAGTSSTCGSYFNPGRDFPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLSGLLDLALGKDYVRSKIAEYMNHLIDIGVAGFRIDASKHMWPGDIKAILDKLHNLNSNWFPEG
SKPFIYQEVIDLGGPEIKSSDYFGNGRVTEFKYGAKLGTVIRKWNGEKMSYLNWGEWGFMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFM LAHPYGFTRVMSSYRWP
RYFENGKDVNDWVGPPNDNGVTKEVTINPDTCGNDWVCEHRWRQIRNMVNFRNVVDGQPFTNWDNGSNQVAFGRGNRGFIVFNDDWTFSLTLQTGLPAGTYCDVISGDKINGNCTGI
KIYVSDDGKAHFSISNSAEDPFIAIHAESKL

ショウジョウバエ *Drosophila melanogaster*

MFLAKSIVCLALLAVANAQFDTNYASGRSGMVHLFEWKWDDIAAECENFLGPNGYAGVQVSPVNENAVKDRRPWWERYQPISYKLETRSGNEEQFASMVKRCNAVGVRTYVDVVFNHMAA
NGGTGTGGSTASPSSKSYPGVPYSSLDNFPTCAIRNYNDANEVRNCELVGLRDLNQGNSYVQDKVVEFLDHLIDLGVAGFRVDAAKHMWPADLAVIYGR LKNLNTDHGFASGSKAYIVQ
EVIDMGGEAISKSEYTGLGAITEFRHSDSIGKVFRGKNQLQYLTNWGTAWGFAASDRSLVFVDNHDNQRGHGAGGADVLT YKVPKQYKMASAFMLAHPFGTPRVMSSFSFTDTDQGPPTT
DGHNIASPIFNSDNSCSGGWVCEHRWRQIYNMVAFRNTVGSDEIQNWWDNQSNQISFSRGSRGFVAFNNDNYDLNSSLQTGLPAGTYCDVISGSKSGSSCTGKT VTVGSDGRASINIGSS
EDDGVLAIHVNAKL

線虫 *Caenorhabditis elegans*

MLKHLFLFLFIKSSIAYNFYWDKTQTLQNRQTMVHLFEWKWADVAKECENFLQYYGYGAVQVSPPM EHLKAFPNNNYPWWVRYQPVSYKLDSRSGNEQEFQDMVNRCNKVGVRIIVDIV
MNHMVGIGQKSGNGVGSSGSSSFDGTHGVQSFPGPYSLGDFNPNPKCDGDIQGSYQNSAEHVKDCRLVGLLDLNQASATVRAKIVAYLNKLVD MG VAGFRHDASKHMWPQDILNINLNDV
KDLRSDIYGSNQRPFVHEVIDRGGEAVKCGDYFGNGRYTNFNFGAAVSAAAKQQSDWKYLANLGP GYGYGNNEHDVLFIDNHDNQRDSSPYVVTYKDGQKYNLAVGFMLAWPYGYPR
VMSSFAFSYSDQSPPNSGASNDYATTSPTFNGDQTCNSNSGWVCEHRWPAIRQMAKFRSAVQGTAAADIVTDTRRIAFARDGKGFFALNQQDGAWTKIFATNLPAGDYCDHFSGGLDNGK
CVGTKITVRDDHTSYLSVPSNSIVAISLDSKLTPTTPPAPPSGYSTTVILLKKDTNPGQNI FVRGGTSQAHNGKCSTGPYQQSSDPCAIPMYHATTVPFVFAEYLTWSQQDNYLDFEGAE
INQGTHDGAAAFGTPLAYSTNDKTAIEYQPYNKYGPYWI AVL RMDCSKAEQGWFEVKGYETPDIGWEGDIKQGSCSGSVGGSAPFSSINH VAKCGAVNVFTWGSASCVIDSA

ClustalW

複数配列を整列させ、進化的関係や保存領域を調べる多重アラインメントツール

ClustalW

複数配列を整列させ、進化的関係や保存領域を調べる多重アラインメントツール

FASTA形式のデータが必要

>Homo_sapiens
MKLFWLLFTIGFCWAQYSSNTQQGRTSIVHLFEWRWVDIALECYRLAPKGFGGVQVSPPNENVIAIHNPFRRPWWERYQPVSYKLCTRSGNEDEFNMVTRCINNVGVR IYVDAVINHMC GN
AVSAGTSSTCGSYFNPGRDFPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLSGLLDLALGKDYVRSKIAEYMNHLIDIGVAGFRIDASKHMWPGD IKA ILDKLHNLNSNWFPEG
SKPFIYQEVIDLGGEPIKSSDYFGNGRVTEFKYGAKLGTVIRKWNGEKMSYLKNWGEGWGFMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFM LAHPYGFTRVMSSYRWP
RYFENGKDVNDWVGPPNDNGVTKEVTINPD TTCGNDWVCEHRWRQIRNMVNFERNVVDGQPFTNWDNGSNQVAFGRGNRGFIVFNDDWTFSLTLQTGLPAGTYCDVISGDKINGNCTGI
KIYVSDDGKAHFSISNSAEDPFIAIHAESKL

>Drosophila_melanogaster
MFLAKSIVCLALLAVANAQFDTNYASGRSGMVHLFEWKWDDIAAECENFLGPNGYAGVQVSPVNENAVKDRRPWWERYQPISYKLETRSGNEEQFASMVKRCNAVGVRTYVDVVFNHMAA
NGGT YGTGGSTASPSSKSYPGVPYSSLD FNPTCAIRNYNDANEVRNCELVGLRDLNQGNSYVQDKVVEFLDHLIDLGVAGFRVDAAKHMWPADLAVIYGRLKNLNTDHGFASGSKAYIVQ
EVIDMGGEAISKSEYTGLGAITEFRHSDSIGKVFRGKNQLQYL TNWGTAWGFAASDRSLVFVDNHDNQRGHGAGGADVLT YKVPKQYKMASAFMLAHPFGTPRVMSSFSFTDTDQGPPTT
DGHNIASPIFNSDNSCSGGWVCEHRWRQIYNMVAFRNTVGSDEIQNWWDNNGSNQISFSRGSRGFVAFNNDNYDLNSSLQTGLPAGTYCDVISGSKSGSSCTGKTVTVGSDGRASINIGSS
EDDGVLAIHVNAKL

>Caenorhabditis_elegans
MLKHLFLFLFIKSSIAYNFYWYDKTQTLQNRQTMVHLFEWKWADVAKECENFLQYYGYGAVQVSPPM EHLKAFPNNNYPWWVRYQPVSYKLDSRSGNEQEFQDMVNRCNKVGVRIIVDIV
MNHMVGIGQKSGNGVGSSGSSSFDGTHGVQSFPGPYPYSLGDFNPNPKCDGDIQGS DYQNSAEHVKDCRLVGLLDLNQASATVRAKIVAYLNKL VDMGVAGFRHDASKHMWPQDILN I LNDV
KDLRSDIYGSNQRPF AVHEVIDRGGEAVKCGDYFGNGRYTNFNFGAAVSAAAKQQSDWKYLANLGP GYGYGNNEDHDV LNFIDNHDNQRDSSPYVV TYKDGQKYNLAVGFMLAWPYGYPR
VMSSFAFSYSDQSPPN SGASNDYATTSPTFNGDQTCNSNSGWVCEHRWPAIRQMAKFRSAVQGTAAADIVTDTRRIAFARDGKGFFALNQQDGAWTKIFATNLPAGDYCDHFSGGLDNGK
CVGTKITVRDDHTSYLSVPSNSIVAISLDSKLT PYTPPAPPSGYSTTVILLKKDTNPGQNI FVRGGTSQAHNGKCSTGPYQQSSDPCAIPMYHATTVPFVFAEYLTWSQQDNYLDFEGAE
INQGTHDGAAAFGTPLAYSTNDKTAIEYQPYNKYGPYWI AVLRMDCSKAEQGWFEVKGYETPDIGWEGDIKQGS CSGSVGGSAPFSSINHVAKCGAVNVFTWGSASCVIDSA

ClustalW

0) 検索するアミノ酸配列を手に入れよう！

0-1) URL欄に直接 “<https://x.gd/IJV9d>” と入力してEnterを押す。

0-2) FASTA形式のファイルが表示されるので、全ての行を選択してコピーする。

ClustalW

1) 検索エンジンで“ClustalW”と入力して、検索する。



ClustalW

2) 以下の画面のようなサイトが表示されたら、クリックする。



GenomeNet

<https://www.genome.jp> > tools-bin



Multiple Sequence Alignment - CLUSTALW

Enter your sequences (with labels) below (copy & paste): PROTEIN DNA Support Formats: FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF

ClustalW

- 3) 画面内の “Enter your sequences” の欄に、コピーした配列を貼り付ける。
- 4) “Execute Multiple Alignment”をクリックする。

General Setting Parameters:
Output Format: CLUSTAL
Pairwise Alignment: ☒ **FAST/APPROXIMATE** ☐ **SLOW/ACCURATE**
Enter your sequences (with labels) below (copy & paste): ☒ **PROTEIN** ☐ **DNA**
Support Formats: FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF

Or give the file name containing your query
ファイルを選択 ファイル未選択

Execute Multiple Alignment Reset

ClustalW

5) 結果が出力されるので、確認する。

```
CLUSTAL 2.1 multiple sequence alignment

Homo_sapiens      -----MKLFWLLFTIGFCWAQYSSNTQQ
Mus_musculus      -----MKFFLLLSLIGFCWAQYDPHTQY
Danio_rerio       -----MKLLILAALVGLSLAQFDPNTKS
Xenopus_laevis    -----MKVLLLLVAIGLCSAQYNPNTQS
Drosophila_melanogaster -----MFLAKSIVCLALLAVANAQFDTNYAS
Caenorhabditis_elegans -----MLKHLFLFLFIKSSIAYNFYWDKTQTLQ
Aspergillus_oryzae MMVAWWSLFLYGLQVAAPALAATPADWRSQSIYFLLTDRFARTDGSTTAT

Homo_sapiens      GRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPNE-NVAIHNPFRP
Mus_musculus      GRTAIVHLFEWRWVDIAKECERYLAPNGFAGVQVSPNE-NIVVHSPSRP
Danio_rerio       GRTAIVHLFEWRWADIAAECERYLGPNFGGGVQISPPSE-SIVVTNPWHP
Xenopus_laevis    GRTSIVHLFEWRWDDIAAECERYLGPNFGGGVQISPPNE-NIIVTNPYRP
Drosophila_melanogaster GRSGMVHLFEWKWDDIAAECENFLGPNGYAGVQVSPVNE-NAVKDR--RP
Caenorhabditis_elegans NRQTMVHLFEWKWADVAKECENFLQYYGYGAVQVSPMEHLKAFPNNNYP
Aspergillus_oryzae CNTADQKYCGGTWQGIIDKLD-YIQGMGFTAIWITPVTAQLPQTTAYGDA


      .           :           * .:   :   :   :   * :   .:   : : *
      .           :           * .:   :   :   :   * :   .:   : : *
```

ClustalW

6) 画面上部の “Select tree menu” をクリックし、“FastTree”を選択して、“Exec”をクリックする。

CLUSTALW Result

[\[clustalw.aln\]](#)[\[clustalw.dnd\]](#)[\[readme\]](#)

Select tree menu 


Exec

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: Homo_sapiens 511 aa

CLUSTALW Result

[\[clustalw.aln\]](#)[\[clustalw.dnd\]](#)[\[readme\]](#)

FastTree 

Exec

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: Homo_sapiens 511 aa

ClustalW

