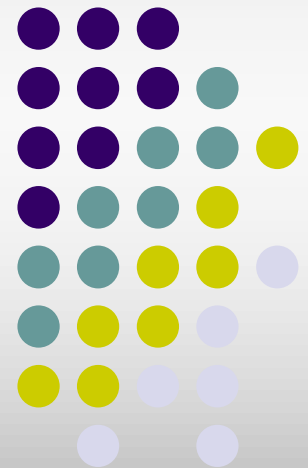

Fasta
Blast



FASTA



- Το **FASTA** (FAST-All – δηλώνει τη γρήγορη σύγκριση πρωτεϊνών ή γρήγορη σύγκριση νουκλεοτιδίων) είναι μηχανή αναζήτησης της αντίστοιχης μίας δοθείσας από το χρήστη σειράς μέσα σε μία βάση δεδομένων σειρών.
- Συγκρίνει την εκάστοτε σειρά (query string) με μία σειρά κάθε φορά, επομένως αναζητώντας να αντιστοιχήσει την query string μέσα σε μία βάση δεδομένων σειρών θα πραγματοποιήσει πολλές αναζητήσεις.
- Το FASTA πραγματοποιεί μία αναζήτηση για την εύρεση τοπικών αντιστοιχίσεων (local alignments) με υψηλή βαθμολογία ανάμεσα στις δύο υπό σύγκριση σειρές, χρησιμοποιώντας για το σκοπό αυτό έναν πίνακα αντικατάστασης. Το FASTA βασίζεται στο ότι σε περίπτωση μίας καλής τοπικής αντιστοίχισης υπάρχει πιθανότητα να υπάρχουν κάποιες υποσειρές οι οποίες να παρουσιάζουν απόλυτη ταύτιση.
- Πριν επιχειρήσει μία χρονοβόρα βελτιστοποιημένη αναζήτηση, το FASTA αναζητά αναγνωρίσεις λέξεων (word hits) ανάμεσα στις δύο σειρές, οι οποίες είναι ενδεικτικές για πιθανές αντιστοιχίσεις.

Ορολογία του FASTA



- **ktup (k tuples)** : Μία ακέραια τιμή που ορίζει το μήκος της λέξης η οποία χρησιμοποιείται για την εύρεση υποσειρών που να ταυτίζονται. Αυξάνοντας το μέγεθος της παραμέτρου ktup μειώνεται η ευαισθησία, επομένως και ο αριθμός των αποτελεσμάτων που δίνει το FASTA.
 - **DNA** : ktup = 4 ή 6
 - **Πρωτεΐνες** : ktup = 1 ή 2(δηλαδή το FASTA αναζητά αντιστοιχήσεις ανάμεσα στις δύο σειρές οι οποίες θα έχουν τουλάχιστον μήκος ενός ή δύο αμινοξέων αντίστοιχα). Το FASTA δεν εξετάζει κάθε λέξη που βρίσκει, αλλά αναζητά τμήματα των σειρών τα οποία περιέχουν πολλές λέξεις, οι οποίες έχουν μικρή απόσταση μεταξύ τους.
- **hot spots** : Οι υποσειρές μήκους ktup, που έχουν βρει αντιστοιχία ανάμεσα στις δύο συγκρινόμενες σειρές. Σε ένα dot plot διαδοχικά hot spots εντοπίζονται στη διαγώνιο.
- **diagonal run** : Μία σειρά από κοντινά hot spots στην ίδια διαγώνιο. Τα κενά ανάμεσα στα hot spots επιτρέπονται.

Fasta

<http://www.ebi.ac.uk/fasta/>



Ορίζεται ο τίτλος της αναζήτησης

1. **interactive:** τα αποτελέσματα θα εμφανιστούν στον browser.
2. **email:** τα αποτελέσματα θα παραδοθούν στο χρήστη μέσω email.

Εισαγωγή της αλληλουχίας σε FASTA format. Δηλ. ">όνομα αλληλουχίας αλληλουχία".

Π.χ.:

```
>Mus musculus protein sequence  
MEAIKKKMQMLKLDKENVLD.....
```

<http://www.ebi.ac.uk/2can/tutorials/protein/sequence8.txt>

YOUR EMAIL	SEARCH TITLE	RESULTS	PROGRAM	DATABASES
<input type="text"/>	Sequence	interactive	fasta3	Protein

GAP PENALTIES	SCORES & ALIGNMENTS	KTUP/HISTOGRAM	DNA STRAND	MATRIX
OPEN: -10 RESIDUE: -2	SCORES: 50 ALIGN: 50	KTUP: 2 HIST: no	none	BLOSUM50

EXPECTATION UPPER VALUE	EXPECTATION LOWER VALUE	SEQUENCE RANGE	DATABASE RANGE	MOLECULE TYPE
10.0	default	START-END	START-END	Protein

Enter or Paste a PROTEIN Sequence in any format:

Upload a file: Αναζήτηση... Run Fasta3 Reset

Fasta

- **PROGRAM** : Το πρόγραμμα που θα χρησιμοποιηθεί από το FASTA. Οι επιλογές είναι:
 - **fasta** : αναζητά παρόμοιες με τη δοθείσα αλληλουχίες μέσα σε μία βιβλιοθήκη αλληλουχιών πρωτεϊνών ή DNA.
 - **fastax/y** : συγκρίνει μία αλληλουχία DNA (μεταφράζοντάς την) με τις αλληλουχίες που υπάρχουν σε μία βάση δεδομένων πρωτεϊνικών αλληλουχιών.
 - **SSEARCH** (local)
 - **GGSEARCH** (global)
 - **GLSEARCH** (global query, local database)

European Bioinformatics Institute [GB] | <https://www.ebi.ac.uk/Tools/seq/fasta/>

Other Protein Databases

STEP 2 - Enter your input sequence

Enter or paste a **PROTEIN** sequence in any supported format

or Upload a file [Εμπεύλη αρχείο](#) [Δεν εμπέυληκ... αρχείο](#) [Use a example sequence](#) [Clear sequence](#) [See more example inputs](#)

STEP 3 - Set your parameters

PROGRAM

FASTA (selected)

FASTX

FASTY

SSEARCH

GGSEARCH

GLSEARCH

☐ Be notified by email (Tick this box if you want to be notified by email when the results are available)

Submit

If you use this service, please consider citing the following publication: [Programmatic access to bioinformatics tools from EMBL-EBI update: 2017](#)

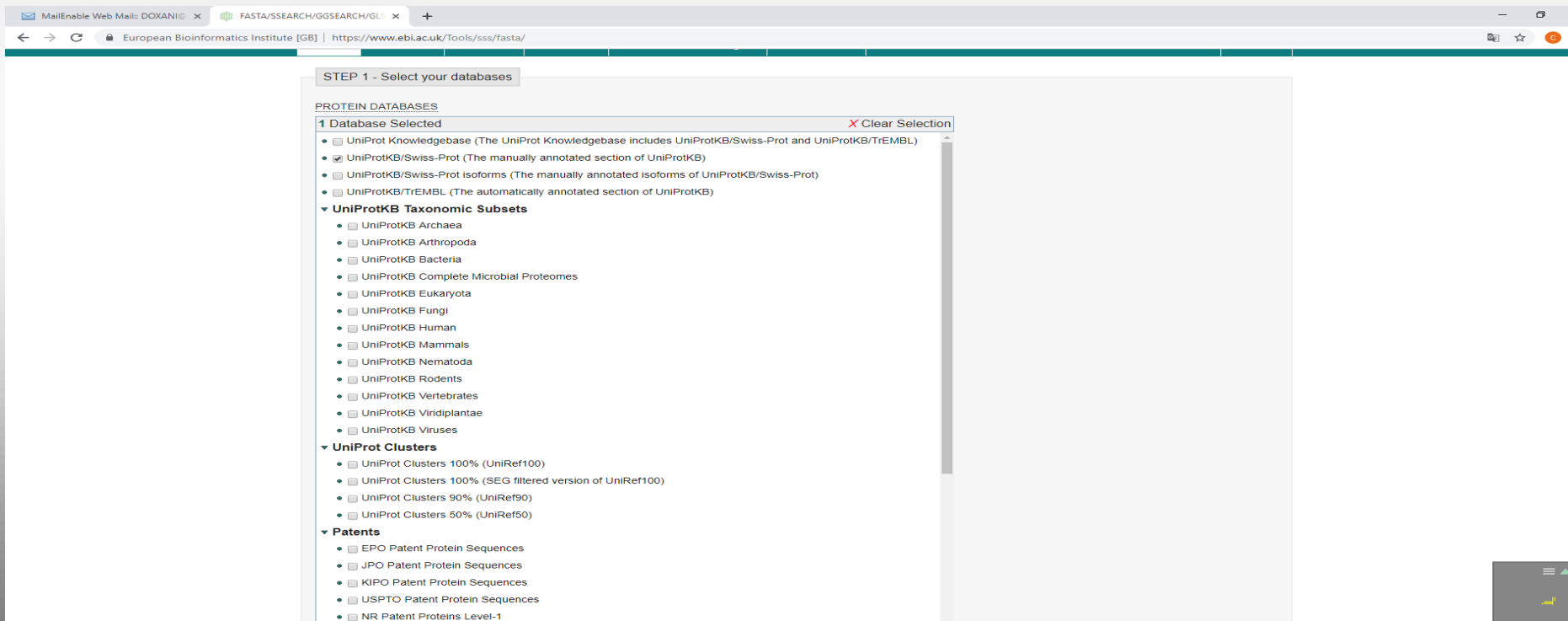
Please read the provided Help & Documentation and FAQs before seeking help from our support staff. If you have any feedback or encountered any issues please let us know via EMBL-EBI Support. If you plan to use these services during a course please contact us. Read our Privacy Notice if you are concerned with your privacy and how we handle personal information.



Fasta



- **DATABASES** : Οι ΒΔ όπου θα αναζητηθεί η αντίστοιχη της δοθείσας αλληλουχίας.
 - DNA FASTA: EMBL, FUNGI, INVERTABRATES, κ.ά.,
 - protein FASTA: UniProt, UniRef, UniParc, swissprot, pdb, κ.ά.



Fasta



- **GAP PENALTIES** : Χρησιμοποιείται η affine μέθοδος.
 - **OPEN PENALTY**: ποινή που θα επιβληθεί στο πρώτο αμινοξύ του κάθε κενού
 - **RESIDUE PENALTY**: ποινή που θα επιβληθεί σε κάθε επιπλέον αμινοξύ (residue) σε κάθε κενό
- **SCORES** : Αριθμός των αντιστοιχήσεων, δηλαδή των αποτελεσμάτων, που θα παρουσιάσει το FASTA από αυτά που βρήκε στη βάση.
- **ALIGNMENTS** : Αριθμός των αντιστοιχήσεων των σειρών – αποτελεσμάτων με την προς εξέταση σειρά που θα παρουσιαστούν.
- **KTUP** : Μήκος της λέξης που θα χρησιμοποιηθεί στην αναζήτηση.
 - protein FASTA : ktup = 2
 - DNA Fasta : ktup = 6

Όσο μεγαλύτερη τιμή δίνεται στο ktup, τόσο μειώνεται η ευαισθησία της αναζήτησης, αλλά ταυτόχρονα μειώνεται και ο χρόνος αυτής.



Fasta

- **DNA STRAND** : Ποίο από τα δύο DNA strands θα χρησιμοποιηθεί στην αναζήτηση μίας αλληλουχίας DNA μέσα σε μία βάση αλληλουχιών DNA. Στο protein FASTA επιλέγεται το none.
- **MATRIX** : Πίνακας αντικατάστασης που θα χρησιμοποιηθεί.
 - protein Fasta: BLOSUM50, BLOSUM62, BLOSUM80, PAM120, PAM250, MDM10, MDM20, MDM40.
Default: BLOSUM50(50% ομοιότητα), καλή επιλογή για την ανίχνευση 'αδύναμων' πρωτεϊνικών ομοιοτήτων. Οι πίνακες PAM χρησιμοποιούνται σε περιπτώσεις αλληλουχιών αμινοξέων.

Μεγαλύτερη τιμή στον πίνακα PAM → αντιστοίχιση σειρών με μεγαλύτερες εξελικτικές αποστάσεις

Μεγαλύτερη τιμή στον πίνακα BLOSUM → αντιστοίχιση σειρών με μεγαλύτερο ποσοστό ομοιότητας.



Fasta

- **EXPECTATION UPPER VALUE E()** :
 - $E() < 0.01$: Οι σειρές σχεδόν πάντα είναι ομόλογες
 - $1 < E() < 10$: Οι σειρές συχνά σχετίζονται.
 - Default: FASTA (protein) $E()=10.0$,
 $E() = \{1.0, 1e-600, 1e-300, 1e-100, 1e-50, 1e-10, 1e-5, 0.001, 0.1, 1.0, 2.0, 5.0, 10.0 \text{ (default), } 20.0, 50\}$
- **EXPECTATION LOWER VALUE E()** : Όσο μικρότερη τιμή δοθεί στο $E()$, τόσο καλύτερα θα είναι τα αποτελέσματα, καθώς οι σειρές που θα προκύψουν θα είναι πιο κοντινές συγγενικά με την υπό εξέταση σειρά. Π.χ. αν $E()=1e-6$, τότε σειρές με $E() < 1e-6$ θα αποκλειστούν. Επομένως η αναζήτηση θα επικεντρωθεί σε σειρές με πιο μακρινή σχέση.
Default protein FASTA: $E()=0$ (δηλαδή δίνει τις βέλτιστες αντιστοιχήσεις)
 $E() = \{0(\text{default}), 1e-600, 1e-300, 1e-100, 1e-50, 1e-10, 1e-5, 0.001, 0.1, 1.0, 2.0, 5.0, 10.0, 20.0, 50\}$.
Εάν ο χρήστης επιλέξει να ορίσει ο ίδιος την παράμετρο αυτή, τότε θα αποκλείσει τις καλύτερες αντιστοιχήσεις και θα επιτραπεί να παρατεθούν πιο μακρινές σχέσεις.



Fasta

- **SEQUENCE RANGE** : Ποια περιοχή μέσα στην υπό εξέταση αλληλουχία πρέπει να χρησιμοποιηθεί στην αναζήτηση.
Π.χ. Για αλληλουχία αποτελούμενη από 380 αμινοξέα, μπορεί να επιθυμεί η αναζήτηση στη βάση δεδομένων να γίνει μόνο χρησιμοποιώντας την περιοχή που βρίσκεται ανάμεσα στις θέσεις 50 έως 200. Τότε **SEQUENCE RANGE** = 50-200.
- **DATABASE RANGE** : Εύρος των αλληλουχιών που θα αναζητηθούν στη βάση.
Π.χ. Για αλληλουχίες που δεν αποτελούνται από περισσότερα από 300 αμινοξέα, τότε **DATABASE RANGE** = 1-300.
Εάν γράψει 1000-3000 δηλώνει μόνο σειρές με τουλάχιστον 1000 και ταυτοχρόνως όχι περισσότερα από 3000 αμινοξέα.
- **MOLECULE TYPE** : Τύπος του molecule που θα χρησιμοποιηθεί στην αναζήτηση.
 - protein FASTA: 'Protein' και δε χρειάζεται να αλλαχθεί για το fasta3
 - nucleotide FASTA: 'Nucleotide'



← → ↺

European Bioinformatics Institute [GB] | https://www.ebi.ac.uk/Tools/sss/fasta/

☆ ⓘ

STEP 3 - Set your parameters

PROGRAM

FASTA

MATRIX

BLOSUM50

GAP OPEN

-10

GAP EXTEND

-2

KTUP

2

EXPECTATION UPPER VALUE

10

EXPECTATION

0 (default)

DNA STRAND

N/A

HISTOGRAM

no

FILTER

none

STATISTICAL ESTIMATES

Regress

SCORES

50

ALIGNMENTS

50

SEQUENCE RANGE

START-END

DATABASE RANGE

START-END

MULTI HSPs

no

SCORE FORMAT

Default

ANNOTATION FEATURES

no

STEP 4 - Submit your job

☐ Be notified by email (Tick this box if you want to be notified by email when the results are available)

Submit

If you use this service, please consider citing the following publication: [Programmatic access to bioinformatics tools from EMBL-EBI update: 2017](#)

Please read the provided Help & Documentation and FAQs before seeking help from our support staff. If you have any feedback or encountered any issues please let us know via [EMBL-EBI Support](#). If you plan to use these services during a course please [contact us](#). Read our [Privacy Notice](#) if you are concerned with your privacy and how we handle personal information.

EMBL-EBI

Services

By topic

By name (A-Z)

Help & Support

Research

Publications

Research groups

Postdocs & PhDs

Training

Train at EBI

Train outside EBI

Train online

Contact organisers

Industry

Members Area

Workshops

SME Forum

Contact Industry

About EMBL-EBI

Contact us

Events

Jobs

News



FASTA Μορφή

Η αλληλουχία σε μορφή FASTA format αποτελείται από:

- Μία γραμμή που ξεκινάει με το σημείο ">", ακολουθούμενο από τον κωδικό ταυτοποίησης της αλληλουχίας.
- Προαιρετικά ακολουθείται από την περιγραφή με μορφή κειμένου της αλληλουχίας.
- Μία ή περισσότερες γραμμές με την αλληλουχία.
- Ένα αρχείο σε μορφή FASTA μπορεί να αποτελείται από περισσότερες από μία αλληλουχίες.
- Η μορφή FASTA μερικές φορές αναφέρεται και ως μορφή "Pearson" (το όνομα του δημιουργού του προγράμματος FASTA και της αντίστοιχης μορφής).

```
>sp|P02700|OPSD_SHEEP_Rhodopsin OS=Ovis aries OX=9940 GN=RHO PE=1
MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLVLGFPINFILTY
VTVQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLTSLHGYFVFGPTGCNLEGFFATLG
GEIALNSLVVLAIERVYVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPPLVGWSRYIP
QGMQCSGALYFTLKPEINNESFVIYMFVVHFSIPLIVIFFCYGQLVFTVKEAAAQQQES
ATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKSSSV
YNFVIYIMMNKQFRNCMLTTLCCKGNPLGDDEASTTVSKTETSQVAPA
```



IUPAC Amino acid Code

IUPAC amino acid code	Three letter code	Amino acid
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Blast

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



- Το **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) είναι ένα πακέτο που αναπτύχθηκε στα πρότυπα λειτουργίας του FASTA.
- Παρέχει προγράμματα για την εύρεση τοπικών αντιστοιχήσεων με υψηλή βαθμολογία ανάμεσα σε μία υπό εξέταση αλληλουχία (query sequence) και σε μία βάση δεδομένων που περιλαμβάνει αλληλουχίες είτε πρωτεϊνικές είτε DNA.
- Το BLAST βασίζεται στο ότι αλληλουχίες ανάμεσα στις οποίες υπάρχει πραγματική αντιστοιχία έχουν μεγάλη πιθανότητα να περιλαμβάνουν μέσα τους μικρά τμήματα αντιστοιχήσεων, ή ταυτίσεις πολλή υψηλής βαθμολογίας.



←

→

↺

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®

HomeRecent ResultsSaved StrategiesHelp

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


BLAST+ 2.9.0 is here!

The latest version has enhanced support for the new database format.

Tue, 02 Apr 2019 17:00:00 EST

[More BLAST news...](#)

Web BLAST



Nucleotide BLAST

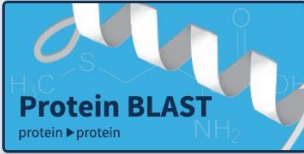
nucleotide → nucleotide

blastx

translated nucleotide → protein

tblastn

protein → translated nucleotide



Protein BLAST

protein → protein

BLAST Genomes

Enter organism common name, scientific name, or tax id

Search


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

10:05 AM




MailEnable Web Mail: DOXANI@ x


FASTA/SSearch/GGSearch/GL x


BLAST: Basic Local Alignment Se x

+


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


**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


SmartBLAST

Find proteins highly similar to your query


Primer-BLAST

Design primers specific to your PCR template


Global Align

Compare two sequences across their entire span (Needleman-Wunsch)


CD-search

Find conserved domains in your sequence


GEO

Find matches to gene expression profiles


IgBLAST

Search immunoglobulins and T cell receptor sequences


VecScreen

Search sequences for vector contamination

CDART

Find sequences with similar conserved domain architecture

Targeted Loci

Search markers for phylogenetic analysis




Multiple Alignment

Align sequences using domain and protein constraints

BioAssay

Search protein or nucleotide targets in PubChem BioAssay

MOLE-BLAST

Establish taxonomy for uncultured or environmental sequences

BLAST is a registered trademark of the National Library of Medicine

NCBI
National Center for Biotechnology Information, U.S. National Library of Medicine
8600 Rockville Pike, Bethesda MD, 20894 USA
Policies and Guidelines | Contact



[Support center](#) [Mailing list](#) [You](#)



Ορολογία του Blast

- **segment pair (ζεύγος τμημάτων)** : Για δύο σειρές S1 και S2 το segment pair είναι ένα ζεύγος από, αυθαίρετου αλλά, ίσου μήκους υποσειρές των S1 και S2, οι οποίες έχουν αντιστοιχισθεί χωρίς κενά.
- **locally maximal segment (τοπικά μέγιστο τμήμα)** : Ένα τμήμα για το οποίο η βαθμολογία που προκύπτει από την αντιστοίχισή του (χωρίς κενά) δεν μπορεί να βελτιωθεί είτε επιμηκύνοντας είτε μικραίνοντάς το.
- **Maximum Segment Pair (MSP) (μέγιστο ζεύγος τμημάτων)** : Στις σειρές S1 και S2 το MSP είναι το segment pair το οποίο έχει την υψηλότερη βαθμολογία απ' όλα τα υπόλοιπα.
- **High-scoring Segment Pair (HSP) (ζεύγος τμημάτων υψηλής βαθμολογίας)** : Για τις σειρές S1 και S2 το HSP είναι το segment pair του οποίου η βαθμολογία αντιστοίχησης είναι ίση ή ξεπερνά ένα ορισμένο όριο (threshold). Επομένως, ένα σύνολο HSP θα προσδιορίζεται από δύο αλληλουχίες (ένα τμήμα της υπό εξέταση σειράς και ένα τμήμα από μία σειρά της βάσης), ένα σύστημα βαθμολόγησης και ένα όριο απόρριψης (σε περίπτωση που το όριο είναι αρκετά υψηλό, τότε το σύνολο θα είναι άδειο).



Blast

Η πρώτη έκδοση του BLAST δεν επέτρεπε την ύπαρξη κενών στις αντιστοιχήσεις των αλληλουχιών. Ακολούθησαν όμως νέες εκδόσεις του BLAST σύμφωνα με τις οποίες επιτρέπονται τα κενά:

- **Gapped BLAST** : επιτρέπει την τοπική αντιστοίχιση με κενά. Επιτρέπει τη συγχώνευση λέξεων με την εισαγωγή κενών ανάμεσά τους. Κάθε νέα αντιστοίχιση θα αποτελείται από δύο λέξεις και τα μεταξύ τους κενά.
- **PSI-BLAST** : Η έκδοση αυτή ξεκινά με το απλό BLAST. Στη συνέχεια παίρνει τα αποτελέσματα και τα χρησιμοποιεί σε πίνακα αντικατάστασης δυναμικού προγραμματισμού. Επαναλαμβάνει την ίδια διαδικασία έως ότου να μην προκύπτουν νέα αποτελέσματα. Το PSI-BLAST είναι καλό για αλληλουχίες που συνδέονται με μακρινή συγγένεια, αλλά συχνά δίνει και λανθασμένες αντιστοιχήσεις.



Κοινές επιλογές WU-Blast και NCBI-Blast

- **PROGRAM** : Το πρόγραμμα που θα χρησιμοποιηθεί από το BLAST. Οι επιλογές είναι:
 - **Protein blast:** συγκρίνει μία πρωτεϊνική αλληλουχία με τις αλληλουχίες σε μία πρωτεϊνική βάση δεδομένων.
 - **Nucleotide blast:** συγκρίνει μία αλληλουχία DNA με τις αλληλουχίες που υπάρχουν σε μία βάση δεδομένων αλληλουχιών νουκλεοτιδίων.
 - **blastx** : συγκρίνει μία αλληλουχία DNA (μεταφράζοντάς την) με τις αλληλουχίες που υπάρχουν σε μία βάση δεδομένων πρωτεϊνικών αλληλουχιών.
 - **tblastn** : συγκρίνει μία πρωτεϊνική αλληλουχία με τις αλληλουχίες σε μία βάση δεδομένων αλληλουχιών νουκλεοτιδίων, τις οποίες έχει μεταφράσει.
 - **tblastx** : συγκρίνει τις πιθανές μεταφράσεις μίας αλληλουχίας νουκλεοτιδίων (DNA) με τις αλληλουχίες σε μία βάση δεδομένων αλληλουχιών νουκλεοτιδίων, τις οποίες επίσης μεταφράσει (αργό).

NCBI-Blast

← → ↺ 🏠 <https://blast.ncbi.nlm.nih.gov/Blast.cgi> 80% ... ☆

⚙️ Συχνές επισκέψεις 📖 Ξεκινώντας 🌐 Suggested Sites 🌐 Web Slice Gallery 🌐 YPOMEDI

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

BLAST Home Recent Results Saved Strategies Help

COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: <https://www.coronavirus.gov> .
Get the latest research from NIH: <https://www.nih.gov/coronavirus> .

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](#)

Search Betacoronavirus Database
We have created a new BLAST database focused on the SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2) Sequences. For further detail please visit [NCBI GenBank](#).
Mon, 03 Feb 2020 10:00:00 EST [More BLAST news...](#)

Web BLAST

Nucleotide BLAST
nucleotide ▶ nucleotide

blastx
translated nucleotide ▶ protein

tblastn
protein ▶ translated nucleotide

Protein BLAST
protein ▶ protein

VODAFONE_WIFI_108
Πρόσβαση στο Internet



NCBI-Blast

- **DATABASES** : Όμοια με το Fasta.
- **MATRIX** : Όμοια με το Fasta.
- **EXP.THR** : (EXpected THReshold) Το όριο που χρησιμοποιείται για τον καθορισμό των HSPs. Όσο πιο μικρό ορίζεται τόσο καλύτερη είναι η αντιστοίχιση που προκύπτει, καθώς μικρότερο όριο δίνει τις αντιστοιχίες με τις μικρότερες πιθανότητες.
- **FILTER** : Μπορεί να ενεργοποιήσει ένα φίλτρο το οποίο να αποκλείει περιοχές της υπό εξέταση αλληλουχίας οι οποίες μπορεί να παρουσιάζουν στατιστικό ενδιαφέρον κατά την αντιστοίχισή τους, αλλά δεν έχουν ενδιαφέρον από βιολογικής απόψεως. Η επιλογή που γίνεται από το BLAST είναι να μη χρησιμοποιήσει κανένα φίλτρο.

NCBI-Blast



← → ↻ 🏠 https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastx&PAGE_TYPE=BlastSearch&BLAS 110% ... 🛡️ ⭐

🔧 Συχνές επισκέψεις 📖 Εκκινώντας 🌐 Suggested Sites 🖼️ Web Slice Gallery 📄 YPOMEDI

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

BLAST » tblastx Home Recent Results Saved Strategies Help

COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: <https://www.coronavirus.gov>.
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.

Translated BLAST: tblastx

blastn blastp blastx tblastn **tblastx**

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Query subrange [From](#) [To](#)

Or, upload file [Αναζήτηση...](#) Δεν επιλέχθηκε αρχείο.

Genetic code Standard (1)

Job Title

Enter a descriptive title for your BLAST search

☐ Align two or more sequences

Choose Search Set

Database Nucleotide collection (nr/nt)

BLAST results will be displayed in a new format by default
You can always switch back to the Traditional Results page.

VODAFONE_WIFI_108
Πρόσβαση στο Internet



Browser address bar: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_

Navigation links: Συχνές επισκέψεις, Εκκινώντας, Suggested Sites, Web Slice Gallery, YPOMEDI

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) **Query subrange**

From
To

Or, upload file [Αναζήτηση...](#) Δεν επιλέχθηκε αρχείο.

Job Title

Enter a descriptive title for your BLAST search

☐ **Align two or more sequences**

Choose Search Set

Database
Non-redundant protein sequences (nr)
Non-redundant protein sequences (nr)
Reference proteins (refseq_protein)
Model Organisms (landmark)
UniProtKB/Swiss-Prot (swissprot)
Patented protein sequences (pataa)
Protein Data Bank proteins (pdb)
Metagenomic proteins (env_nr)
Transcriptome Shotgun Assembly proteins (tsa_nr)

Program Selection
☐ PHI-BLAST (Pattern Hit Initiated BLAST)
☐ DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)
Choose a BLAST algorithm

BLAST results will be displayed in a new format by default
You can always switch back to the Traditional Results page.

VODAFONE_WIFL_108
Πρόσβαση στο Internet

NCBI-Blast



- **ALIGN VIEWS** : Καθορίζει με ποιο τρόπο θα παρουσιαστούν οι αντιστοιχήσεις των αλληλουχιών. Στις δυνατές επιλογές περιλαμβάνονται:
 - **pairwise alignment** (επιλογή του BLAST)
 - **M/S with identities**
 - **M/S without identities**
 - **flat Query – anchored with identities**
 - **flat Query – anchored without identities**



NCBI-Blast

- **OPENGAP** : Ορίζει την ποινή που θα επιβληθεί στο αμινοξύ που 'ανοίγει' το κάθε κενό. Εάν αυξηθεί η ποινή αυτή τότε θα μειωθεί ο αριθμός των κενών στις αντιστοιχήσεις.
- **EXTENDGAP** : Ορίζει την επιπλέον ποινή που θα προστίθεται στο καθορισμένη ποινή κενού για κάθε επιπλέον αμινοξύ στο κενό. Αυτός ο τρόπος χρησιμοποιείται για κενά μεγάλου μήκους. Για να αποφευχθούν τα μεγάλα κενά, αρκεί να αυξηθεί η τιμή / ποινή του extendgap.
- **GAPALIGN** : Οι επιλογές εδώ είναι 'true' και 'false' και ορίζουν εάν το πρόγραμμα θα κάνει αντιστοιχήσεις μέσα σε περιοχές που περιλαμβάνουν κενά.
 1. επιλογή = 'true' —> οι αντιστοιχήσεις συμπεριλαμβάνουν κενά.
 2. επιλογή = 'false' —> μόνο μεμονωμένα HSPs όπου οι δύο αντιστοιχιζόμενες αλληλουχίες ταυτίζονται, και επομένως ανάμεσα στα αποτελέσματα δε θα περιλαμβάνονται αντιστοιχήσεις με κενά.Η επιλογή του BLAST είναι 'true'.



← → ↻ 🏠 https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_ 110% ... 🛡️ ☆

⚙️ Συχνές επισκέψεις 📄 Εκκινώντας 🌐 Suggested Sites 📄 Web Slice Gallery 🌐 YPOMEDI

☐ Show results in a new window

Algorithm parameters

[Restore default search parameters](#)

General Parameters

Max target sequences | 100
Select the maximum number of aligned sequences to display ⓘ

Short queries | ☒ Automatically adjust parameters for short input sequences ⓘ

Expect threshold | 10 ⓘ

Word size | 6 ⓘ

Max matches in a query range | 0 ⓘ

Scoring Parameters

Matrix | BLOSUM62 ⓘ

Gap Costs | Existence: 11 Extension: 1 ⓘ

Compositional adjustments | Conditional compositional score matrix adjustment ⓘ

Filters and Masking

Filter | ☐ Low complexity regions ⓘ

Mask | ☐ Mask for lookup table only ⓘ
☐ Mask lower case letters ⓘ

VODAFONE_WiFi_108
Πρόσβαση στο Internet



Algorithm parameters

General Parameters

Max target sequences: 100 (maximum number of aligned sequences to display)

Short queries: 50 (locally adjust parameters for short input sequences)

Expect threshold: 100

Word size: 250

Max matches in a query range: 1000

Scoring Parameters

Matrix: 20000

Max matches in a query range: 0

Scoring Parameters

Matrix: BLOSUM62

Gap Costs: PAM30 (extension: 1)

Compositional adjustments: PAM70 (positional score matrix adjustment)

Filters and Masking

Filter: BLOSUM45 (regions)

Mask: BLOSUM50 (table only)

BLOSUM90 (mask lower case letters)

Matrix: BLOSUM62

Gap Costs: Existence: 11 Extension: 1

Compositional adjustments: Existence: 11 Extension: 2 (matrix adjustment)

Existence: 10 Extension: 2

Existence: 9 Extension: 2

Existence: 8 Extension: 2

Existence: 7 Extension: 2

Existence: 6 Extension: 2

Existence: 13 Extension: 1

Existence: 12 Extension: 1

BLAST (protein-protein BLAST)

Existence: 11 Extension: 1

Existence: 10 Extension: 1

Existence: 9 Extension: 1

BLAST is a registered trademark of the National Library of Medicine

NCBI

National Center for Biotechnology Information, U.S. National Library of Medicine

8600 Rockville Pike, Bethesda MD, 20894 USA

[Policies and Guidelines](#) | [Contact](#)



MailEnable Web Mail: DOXANI@ x

FASTA/SSSEARCH/GGSEARCH/GL x

Protein BLAST: search protein da x

+

https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome

Choose a BLAST algorithm

BLAST

Search database Non-redundant protein sequences (nr) using Blastp (protein-protein BLAST)

☐ Show results in a new window

Algorithm parameters

Restore default search parameters

General Parameters

Max target sequences

100

Select the maximum number of aligned sequences to display

Short queries

☒ Automatically adjust parameters for short input sequences

Expect threshold

10

Word size

6

Max matches in a query range

0

Scoring Parameters

Matrix

BLOSUM62

Gap Costs

Existence: 11 Extension: 1

Compositional adjustments

Conditional compositional score matrix adjustment

Filters and Masking

Filter

☐ Low complexity regions

Mask

☐ Mask for lookup table only

☐ Mask lower case letters

BLAST

Search database Non-redundant protein sequences (nr) using Blastp (protein-protein BLAST)

☐ Show results in a new window

BLAST is a registered trademark of the National Library of Medicine

Support center

Mailing list

YouTube

NCBI

National Center for Biotechnology Information, U.S. National Library of Medicine

8600 Rockville Pike, Bethesda MD, 20894 USA

Policies and Guidelines | Contact

NATIONAL LIBRARY OF MEDICINE

NIH

USA.gov

10:09 AM

5/4/2019

Στατιστικά μεγέθη σύγκρισης αλληλουχιών



- ***p-value***
 - σχετίζει το αποτέλεσμα μιας αντιστοιχίας με την πιθανότητα να είναι τυχαίο →όσο πιο πολύ προσεγγίζει το μηδέν, τόσο μεγαλύτερη αξιοπιστία υπάρχει ότι το αποτέλεσμα είναι πραγματικό

Στατιστικά μεγέθη σύγκρισης αλληλουχιών



- ***E-value***
- περιγράφει τον αριθμό επιτυχιών (ομοιοτήτων) που αναμένεται να είναι τυχαία στην αναζήτηση μιας βάσης δεδομένων συγκεκριμένου μεγέθους
- όταν το E-value πάρει την τιμή 1 για ένα ταίριασμα, αυτό μπορεί να ερμηνευτεί ότι στην τρέχουσα έρευνα, αναμένεται μόνο από τύχη να βρεθεί μια ομοιότητα με ίδιο αποτέλεσμα.
- Μια τιμή 0 δηλώνει ότι κανένα δεν αναμένεται να είναι τυχαίο, δηλ. είναι απίθανο η αντιστοιχία να είναι από τυχαία ομοιότητα