

## لیست ابزار های مورد استفاده

-1- Pairwise alignment: برای مقایسه دو توالی ها استفاده می شود.

برای استفاده از این ابزار مروگر خود را بر روی آدرس [ebi.ac.uk/tools/psa](http://ebi.ac.uk/tools/psa) قرار دهید. در این صفحه الگوریتم های مختلف برای هم راستایی توالی ها وجود دارد که هر کدام را میتوان بر روی توالی های نوکلئوتیدی یا آمینواسیدی اجرا کرد.

The screenshot shows the EMBL-EBI website with the "Pairwise Sequence Alignment" tool selected. The main content area is titled "Pairwise Sequence Alignment". It includes sections for "Global Alignment", "Local Alignment", and "Genomic Alignment". Each section provides a brief description and links to specific tools like Needle, Stretcher, Water, Matcher, LALIGN, and GenomicAligner. The "Global Alignment" section is currently active. At the bottom, there is a "Feedback" and "Share" button.

**Pairwise Sequence Alignment**

Tools > Pairwise Sequence Alignment

**Pairwise Sequence Alignment** is used to identify regions of similarity that may indicate functional, structural and/or evolutionary relationships between two biological sequences (protein or nucleic acid).

By contrast, **Multiple Sequence Alignment (MSA)** is the alignment of three or more biological sequences of similar length. From the output of MSA applications, homology can be inferred and the evolutionary relationship between the sequences studied.

**Global Alignment**

Global alignment tools create an end-to-end alignment of the sequences to be aligned. There are separate forms for protein or nucleotide sequences.

Needle (EMBOSS)  
EMBOSS Needle creates an optimal global alignment of two sequences using the Needleman-Wunsch algorithm.

Protein Nucleotide

Stretcher (EMBOSS)  
EMBOSS Stretcher uses a modification of the Needleman-Wunsch algorithm that allows larger sequences to be globally aligned.

Protein Nucleotide

**Local Alignment**

Local alignment tools find one, or more, alignments describing the most similar region(s) within the sequences to be aligned. There are separate forms for protein or nucleotide sequences.

Water (EMBOSS)  
EMBOSS Water uses the Smith-Waterman algorithm (modified for speed enhancements) to calculate the local alignment of two sequences.

Protein Nucleotide

Matcher (EMBOSS)  
EMBOSS Matcher identifies local similarities between two sequences using a rigorous algorithm based on the LALIGN application.

Protein Nucleotide

LALIGN  
LALIGN finds internal duplications by calculating non-intersecting local alignments of protein or DNA sequences.

Protein Nucleotide

**Genomic Alignment**

Genomic alignment tools concentrate on DNA (or to DNA) alignments while accounting for characteristics present in genomic data.

GenomicAligner

با توجه به پژوهشی خود الگوریتم مربوطه را انتخاب کنید. در این مثال Global Alignment برای دو توالی پروتئینی نشان داده شده است.

**EMBOSS Stretcher**

Tools > Pairwise Sequence Alignment > EMBOSS Stretcher

**Pairwise Sequence Alignment (PROTEIN)**

EMBOSS Stretcher calculates an optimal global alignment of two sequences using a modification of the classic dynamic programming algorithm which uses linear space.

This is the form for protein sequences. Please go to the nucleotide form if you wish to align DNA or RNA sequences.

STEP 1 - Enter your protein sequences

Enter or paste your first **protein** sequence in any supported format:

Or, upload a file: Choose file no file selected Use a example sequence | Clear sequence | See more example inputs

AND

Enter or paste your second **protein** sequence in any supported format:

Or, upload a file: Choose file no file selected

STEP 2 - Set your pairwise alignment options

MATRIX	GAP OPEN	GAP EXTEND	OUTPUT FORMAT
BLOSUM62	12	2	pair

STEP 3 - 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 plan to use these services during a course please contact us.

توالی های FASTA خود را در پنجره های مرتبطه کپی کنید. با کلیک کردن روی دکمه more options میتوانید تنظیمات همراستایی را از حالت پیش فرض تغییر دهید. ماتریس جهش مورد استفاده و گپ پنالتی از جمله این موارد هستند. در آخر دکمه submit را زده و منتظر شوید تا همراستایی انجام شود. در پایان با چنین صفحه ای مواجه میشوید که خلاصه تنظیمات و نتایج همراستایی را در بر دارد.

The screenshot shows the results of a pairwise sequence alignment between two sequences: HBA\_HUMAN and HBA\_MOUSE. The alignment output is as follows:

```

#####
# Program: stretcher
# Rundate: Fri 24 Aug 2018 16:25:34
# Commandline: stretcher
# Input:
# -stdout
# -sequence emboss_stretcher-I20180824-162531-0081-81745245-p2m.asequence
# -bsequence emboss_stretcher-I20180824-162531-0081-81745245-p2m.bsequence
# -datafile EBLOSUM62
# -gapopen 12
# -gapextend 2
# -afformat3 pair
# -sprotein1
# -sprotein2
# Align format: pair
# Report file: stdout
#####
# =====
# Aligned sequences: 2
# 1: HBA_HUMAN
# 2: HBA_MOUSE
# Matrix: EBLOSUM62
# Gap penalty: 12
# Extend penalty: 2
#
# Length: 142
# Identity: 122/142 (85.9%)
# Similarity: 131/142 (92.3%)
# Gaps: 0/142 (0.0%)
# Score: 648
#
# =====
HBA_HUMAN    1 MVLSPADKTVKAANGKVGAHAGEYGAELERMFLSFPTTKTYFPBFDSL   50
|||||.|.:|:|:|:|||:|.|||..|||||:|||||:|||||:|||:|:|:|:|:|:|:|:|:|:|:|
HBA_MOUSE     1 MVLSEDKSNSKAANGKIGGIGEYGAELERMFASTPTTKTFPHFDVS   50
|||||.|.:|:|:|:|||:|.|||..|||||:|||||:|||:|:|:|:|:|:|:|:|:|:|:|:|
HBA_HUMAN     51 HGSQAVQVGHGKXKVALDTNAVAVVDMPNALSALSDLHAKHLRVPDVNFK 100
|||||.|.:|:|:|:|||:|.|||..|||||:|||||:|||:|:|:|:|:|:|:|:|:|:|:|:|:|
HBA_MOUSE     51 HGSQAVQVGHGKXKVALASAGHLDDLGPGLSALSDLHAKHLRVPDVNFK 100
|||||.|.:|:|:|:|||:|.|||..|||||:|||||:|||:|:|:|:|:|:|:|:|:|:|:|:|
HBA_HUMAN     101 LLSHCLLVTFLASHIPADFTPAVHASLDKFLASVSTVTSKYR    142
|||||.|.:|:|:|:|||:|.|||..|||||:|||||:|||:|:|:|:|:|:|:|:|:|:|:|:|:|
HBA_MOUSE    101 LLSHCLLVTFLASHIPADFTPAVHASLDKFLASVSTVTSKYR    142
|||||.|.:|:|:|:|||:|.|||..|||||:|||||:|||:|:|:|:|:|:|:|:|:|:|:|:|
#=====
#=====

```

۲- Multiple alignment: با استفاده از این ابزار میتوانید چندین توالی را با هم مقایسه کنید. امکان استفاده از خروجی این نوع همراستایی برای کشیدن درخت فیلوژنتیک نیز وجود دارد. برای شروع مرورگر خود را بر روی آدرس [ebi.ac.uk/tools/msa](http://ebi.ac.uk/tools/msa) قرار دهید. در این صفحه الگوریتم های مختلف برای همراستایی چندگانه را مشاهده میکنید. در این مثال از الگوریتم Clustal Omega برای مقایسه چند توالی پروتئین استفاده میکنیم.

The screenshot shows the Multiple Sequence Alignment page on the EMBL-EBI website. It lists several tools for multiple sequence alignment:

- Clustal Omega**: New MSA tool that uses seeded guide trees and HMM profile-profile techniques to generate alignments. Suitable for medium-large alignments. [Launch Clustal Omega](#)
- Kalign**: Very fast MSA tool that concentrates on local regions. Suitable for large alignments. [Launch Kalign](#)
- MAFFT**: MSA tool that uses Fast Fourier Transforms. Suitable for medium-large alignments. [Launch MAFFT](#)
- MUSCLE**: Accurate MSA tool, especially good with proteins. Suitable for medium alignments. [Launch MUSCLE](#)
- MView**: Transform a Sequence Similarity Search result into a Multiple Sequence Alignment or reformat a Multiple Sequence Alignment using the MView program. [Launch MView](#)
- T-Coffee**: Consistency-based MSA tool that attempts to mitigate the pitfalls of progressive alignment methods. Suitable for small alignments. [Launch T-Coffee](#)
- WebPRANK**: The EBI has a new phylogeny-aware multiple sequence alignment program which makes use of evolutionary information to help place insertions and deletions. Try it out at [WebPRANK](#).

Footnote: The tools described on this page are provided using [Programmatic access to bioinformatics tools from EMBL-EBI update: 2017](#). If you have any feedback or encounter any issues please let us know via [EMBL-EBI support](#).

در اولین پنجره Enter or paste a set of FASTA مورد نظر خود را در پنجره مربوطه کپی کنید. دقت داشته باشید که بین پایان یک توالی و آغاز توالی بعد باید یک خط فاصله باشد.

The screenshot shows the Clustal Omega web interface. At the top, there's a navigation bar with links for EMBL-EBI, Services, Research, Training, Industry, About us, and a search bar. On the right, it says "EMBL-EBI Hinxton". Below the header, the title "Clustal Omega" is displayed, followed by a sub-header "Multiple Sequence Alignment". A note below states: "Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between **three or more sequences**. For the alignment of two sequences please instead use our pairwise sequence alignment tools." A note at the bottom of this section says: "Important note: This tool can align up to 4000 sequences or a maximum file size of 4 MB." The main area is divided into three steps: "STEP 1 - Enter your input sequences", "STEP 2 - Set your parameters", and "STEP 3 - Submit your job". In Step 1, there's a text area labeled "Enter or paste a set of PROTEIN sequences in any supported format" and a file upload field. Step 2 contains a table for setting parameters like "DEALIGN INPUT SEQUENCES", "MBED-LIKE CLUSTERING GUIDE", "MAX GUIDE TREE ITERATIONS", and "NUMBER OF COMBINED ITERATIONS". Step 3 has a "Submit" button and a note about contacting support if using services during a course.

پس از پایان کار صفحه‌ی زیر را مشاهده خواهید کرد. با کلیک کردن روی تب‌های مختلف بالای صفحه میتوانید اطلاعات بیشتری راجع به این همراستایی پیدا کنید یا آن را به عنوان ورودی به برنامه‌های دیگر بدهید.

The screenshot shows the Clustal Omega web interface after a job has been submitted. The title "Clustal Omega" is at the top, followed by the sub-header "Multiple Sequence Alignment". Below this, it says "Results for job clustalo-I20180824-163858-0479-29210239-p2m". There are tabs for "Alignments", "Result Summary", "Phylogenetic Tree", and "Submission Details", with "Alignments" being the active tab. Below the tabs are buttons for "Download Alignment File", "Show Colors", "View result with Jalview", "Send to Simple Phylogeny", and "Send to MView". The main content area shows the "CLUSTAL O(1.2.4) multiple sequence alignment" results. It displays two sets of aligned sequences for each of the four entries (P013169, P69905, P01942, P13786). The sequences are color-coded to represent different amino acids. At the bottom, a note says: "PLEASE NOTE: Showing colors on large alignments is slow."



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  
programme

About EMBL-EBI  
Contact us  
Events  
Jobs  
News  
People & groups

- با کلیک بر روی تب Result summary با صفحه‌ی زیر مواجه خواهید شد. هر کدام از لینک‌های موجود در صفحه اطلاعاتی در مورد همراستایی انجام شده به شما میدهد.

The screenshot shows the Clustal Omega results page. At the top, there are tabs for Alignments, Result Summary, Phylogenetic Tree, and Submission Details. Below these, there are sections for Input Sequences, Tool Output, and other generated files. On the right side, there are links to Jalview, Simple Phylogeny, and MView. At the bottom, there is a footer with links to EMBL-EBI services, Research, Training, Industry, and About us, along with copyright information and an Intranet link.

```
# Percent Identity Matrix - created by Clustal2.1
# 
1: tr|O13169|O13169_CYPCA 100.00 50.00 50.70 47.18
2: sp|P69905|HBA_HUMAN      50.00 100.00 85.92 56.34
3: sp|P01942|HBA_MOUSE       50.70 85.92 100.00 56.34
4: sp|P13786|HBAZ_CAPHI     47.18 56.34 56.34 100.00
```

- برای رسم درخت فیلوجنتیک باید بر روی تب Send to simple phylogeny کلیک کنید. این برنامه از نتیجه همراستایی برای رسم درخت استفاده میکند. در صفحه حاصل میتوانید تنظیمات کشیدن درخت را تغییر دهید. از مهم ترین این تنظیمات میتوان به الگوریتم رسم درخت اشاره کرد. با کلیک کردن روی دکمه‌ی submit رسم درخت آغاز میشود.

The screenshot shows the Simple Phylogeny tool interface. It has tabs for Input form, Web services, Help & Documentation, and Bioinformatics Tools FAQ. Below these are sections for STEP 1 (multiple sequence alignment), STEP 2 (phylogeny options), and STEP 3 (submit job). There are also links for course contact and support staff FAQ at the bottom.

## Phylogram

Branch length: ● Cladogram ○ Real



## Phylogenetic Tree

[View Phylogenetic Tree File](#)

```

triO13169jO13169_CYPCA 0.29401
splP69905jHBA_HUMAN 0.07218
splP01942jHBA_MOUSE 0.06866
splP13786jHBAZ_CAPHI 0.23415

```

- خروجی این ابزار به شکل روپرتو است. نام هر شاخه اولین خط در توالی FASTA ای است که به نرم افزار Dadaه شده است. شما میتوانید طول شاخه ها تغییر دهید تا نسبت های واقعی آنها را مشاهده کنید و درخت حاصل را ذخیره کنید.

-۳- BLAST: این ابزار برای جستجوی یک توالی در بین تمامی entry های یک پایگاه داده استفاده میشود.

برای استفاده از این الگوریتم جستجو، آدرس [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov) را در مرورگر وارد کنید و از قسمت Popular Resources بر روی لینک کلیک کنید. در این صفحه باید نوع BLAST را انتخاب کنید. برای جستجوی یک توالی نوکلئوتیدی در پایگاه داده نوکلئوتید از Nucleotide BLAST و برای جستجوی یک توالی پروتئینی در پایگاه داده پروتئین از Protein BLAST استفاده میکنیم. دو نوع دیگر برای جستجو پروتئین در پایگاه داده نوکلئوتید و بالعکس استفاده میشوند. نوع جستجوی مورد نیاز خود را انتخاب کنید.

- در این مثال Protein BLAST نمایش داده شده است. در صفحه‌ی حاصل میتوانید جستجوی خود را انجام دهید و تنظیمات را تغییر دهید.
- با استفاده از دکمه‌ی Choose File فایل FASTA را آپلود کرده یا توالی را در پنجره‌ی بالای صفحه کپی کنید.
  - در قسمت Query subrange میتوانید تنظیم کنید که تنها بخش خاصی از پروتئین مورد نظر شما در پایگاه داده جستجو شود.
  - در قسمت Organism میتوانید موجوداتی را که تمایل دارید که در نتایج جستجو دخیل شده باشد را نشوند را تعیین کنید.
  - در قسمت Max target sequences تعداد hit های نمایش داده شده در خروجی را تغییر میدهد.
  - نتایجی که بالاتر از حد تعیین شده دارند را از خروجی حذف میکند.
  - Word size اندازه کلمات جستجو را تغییر میدهد؛ عدد بزرگتر جستجوی بهتر و طبعاً زمان بیشتر را به همراه دارد.
  - در قسمت Matrix میتوانید ماتریس جهش مورد استفاده را تغییر دهید.
  - امتیاز های منفی شروع و ادامه‌ی گپ را مشخص میکند.
  - پس از اعمال تمامی تنظیمات خود روی دکمه‌ی BLAST کلیک کنید و صبر کنید تا جستجو تمام شود.

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

**BLAST®** » blastp suite Home Recent Results Saved Strategies Help

**Standard Protein BLAST**

blastn blastp blastx tbblastn tbblastx Enter Query Sequence Enter accession number(s), gi(s), or FASTA sequence(s)  Query subrange From  To

Or, upload file Choose File no file selected Job Title Enter a descriptive title for your BLAST search  Align two or more sequences

Choose Search Set Database Non-redundant protein sequences (nr)   Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown.  Exclude  Models (XM/XP)  Non-redundant RefSeq proteins (WP)  Uncultured/environmental sample sequences Entrez Query Enter an Entrez query to limit search  YouTube Create custom database

Program Selection Algorithm Quick BLASTP (Accelerated protein-protein BLAST) **New**  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 Non-redundant protein sequences (nr) using Blastp (protein-protein BLAST)  Show results in a new window

Algorithm parameters General Parameters Restore default search parameters

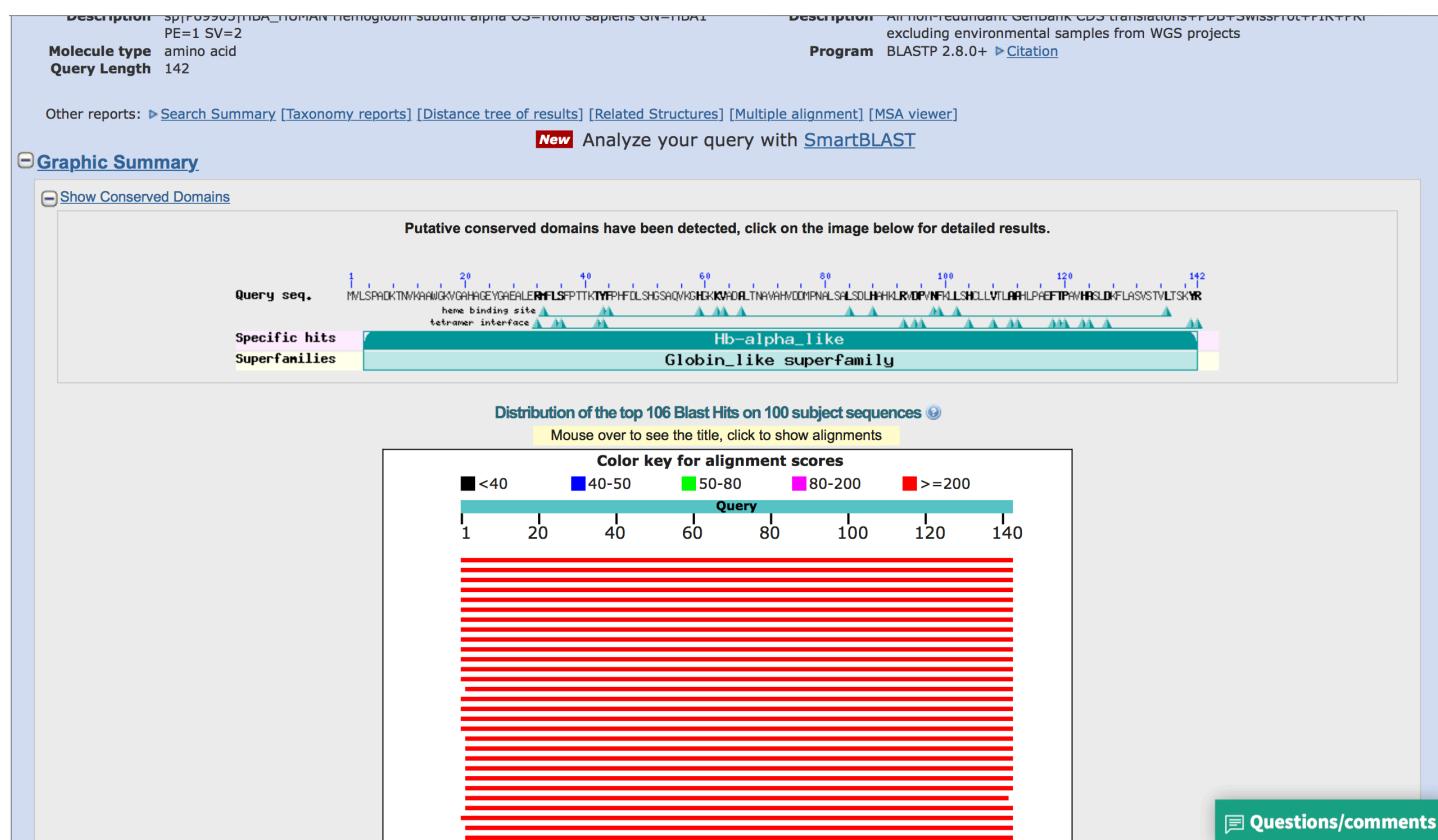
Max target sequences	100 <input type="button" value="▼"/> Select the maximum number of aligned sequences to display <input type="text"/>
Short queries	<input checked="" type="checkbox"/> Automatically adjust parameters for short input sequences <input type="checkbox"/>
Expect threshold	10 <input type="button" value="▼"/>
Word size	6 <input type="button" value="▼"/>
Max matches in a query range	0 <input type="button" value="▼"/>

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

پس از پایان جستجو با صفحه‌ی زیر مواجه خواهید شد. هر کدام از خطوط رنگی مشاهده شده نمایانگر یک توالی هستند که در جستجوی انجام شده به عنوان hit شناخته شده اند. با استفاده از راهنمای رنگی بالای صفحه میتوانید اطلاعاتی نسبت به امتیاز همراستایی برای کل یا بخش‌هایی از hit کسب کنید.



با پایین رفتن در صفحه‌ی نتایج به لیست پروتئین‌های hit خواهید رسید روبروی هر کدام از این توالی‌ها درصد identity و e-value را مشاهده میکنید. با کلیک کردن بر روی لینک accession میتوانید وارد آن entry مشخص در پایگاه داده‌ی ncbi شوید.

Sequences producing significant alignments:										
Select:	All	None	Selected:							
	Alignments				Download	GenPept	Graphics	Distance tree of results	Multiple alignment	Questions/comments
	Description	Max score	Total score	Query cover	E value	Ident				
<input type="checkbox"/>	hypothetical protein [Klebsiella pneumoniae]	288	288	100%	4e-98	100%	WP_108998590.1			
<input type="checkbox"/>	hemoglobin alpha 2 [synthetic construct]	287	287	100%	4e-98	100%	AAX29522.1			
<input type="checkbox"/>	hypothetical protein [Escherichia coli]	287	287	100%	6e-98	100%	WP_108997664.1			
<input type="checkbox"/>	MULTISPECIES: hypothetical protein [Enterobacteriaceae]	286	286	100%	6e-98	100%	WP_108961785.1			
<input type="checkbox"/>	Chain B, A Cis-Proline In Alpha-Hemoglobin Stabilizing Protein Directs The Structural Reorganization Of Alpha-Hemoglobin	286	286	100%	8e-98	100%	3IA3_B			
<input type="checkbox"/>	mutant hemoglobin alpha 2 globin chain [Homo sapiens]	286	286	100%	1e-97	99%	AKZ66543.1			
<input type="checkbox"/>	MULTISPECIES: hypothetical protein [Enterobacteriaceae]	286	286	100%	1e-97	100%	WP_109027992.1			
<input type="checkbox"/>	hemoglobin alpha-2 [Homo sapiens]	285	285	100%	2e-97	99%	AAN04486.1			
<input type="checkbox"/>	alpha-2-globin [Homo sapiens]	285	285	100%	3e-97	99%	AAF72612.1			
<input type="checkbox"/>	TPA: globin C1 [Homo sapiens]	286	286	100%	3e-97	100%	SAI82135.1			
<input type="checkbox"/>	hemoglobin alpha 1-2 hybrid [Homo sapiens]	284	284	100%	5e-97	99%	ABF56145.1			
<input type="checkbox"/>	Chain A, Hemoglobin Thionville: An Alpha-Chain Variant With A Substitution Of A Glutamate For Valine At Na-1 And Having An Acetylated Methionine N <sup>H2</sup> T <sub>e</sub>	284	284	100%	5e-97	99%	1BAB_A			
<input type="checkbox"/>	hemoglobin alpha-1 globin chain [Homo sapiens]	284	284	100%	7e-97	99%	AAK37554.1			
<input type="checkbox"/>	Chain A, Structure Of Haemoglobin In The Deoxy Quaternary State With Ligand Bound At The Alpha Haems	284	284	99%	7e-97	100%	1COH_A			
<input type="checkbox"/>	alpha 2 globin variant [Homo sapiens]	284	284	100%	8e-97	99%	BAD97112.1			
<input type="checkbox"/>	PREDICTED: hemoglobin subunit alpha [Rhinopithecus roxellana]	283	283	100%	9e-97	99%	XP_010380159.1			
<input type="checkbox"/>	hemoglobin subunit alpha [Pongo abelii]	283	283	100%	1e-96	99%	XP_024089067.1			
<input type="checkbox"/>	hemoglobin alpha 1 [Homo sapiens]	283	283	100%	1e-96	99%	AFI57164.1			
<input type="checkbox"/>	Chain A, Hemoglobin (Alpha V1m) Mutant	283	283	99%	2e-96	99%	1BZZ_A			
<input type="checkbox"/>	Chain A, Carbonmonoxy Structure Of Hemoglobin Evans Alphav62mbetawt	283	283	99%	2e-96	99%	4MQC_A			
<input type="checkbox"/>	Chain A, R-State Human Carbonmonoxyhemoglobin Alpha-A53s	283	283	99%	2e-96	99%	1AJ9_A			
<input type="checkbox"/>	Chain A, Structure Of The Human Hemoglobin Mutant Hb Providence (a-gly-c:v1m; B:d:v1m,k82d; Ferrous, Carbonmonoxy Bound)	282	282	99%	2e-96					
<input type="checkbox"/>	RecName: Full=Hemoglobin subunit alpha; AltName: Full=Alnha-nlhin; AltName: Full=Hemoglobin alpha chain	282	282	99%	3e-96					