

DELTA-BLAST Hands-On

BLAST (<http://blast.ncbi.nlm.nih.gov/>), **Basic Local Alignment Search** tool, is a suite of programs for finding similarities between biological sequences. **DELTA-BLAST** is a new, sensitive algorithm, intended for detection of distant protein homologs. Like "BLAST", "DELTA" is an acronym and it stands for **Domain Enhanced Lookup Time Accelerated**.

Outline:

- Compare DELTA-BLAST searches with blastp.
- Search for homologs for a human RAS oncogene family.
- Find which of the human RAS oncogene family are conserved throughout the Eukaryotes.
- Use the longest protein isoforms of these genes to search for homologs in Bacteria.
- Search against a few bacterial genera in the RefSeq protein database.

Step 1: Homologene search

<http://www.ncbi.nlm.nih.gov/homologene/>

Search for clusters with the following requirements:

- conserved in Eukaryotes.
- proteins in the cluster have identifiable conserved domains.
- human gene should be included in the cluster.

The search term for such requirements is:

"txid2759"[Ancestor:noexp] AND "homologene cdd"[Filter] AND human[organism]

Discover Homologs

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search HomoloGene for "txid2759"[Ancestor:noexp] AND "homologene cdd"[Filter] AND human[organism] Go Clear Save Search

Limits Preview/Index History Clipboard Details

History has expired.

Display Summary Show 20 Send to

All: 2843 Fungi: 0 Mammals: 0

Items 1 - 20 of 2843

1: HomoloGene:111024. Gene conserved in Eukaryota

HIST1H4J	histone cluster 1, H4j	<i>Homo sapiens</i>
HIST2H4B	histone cluster 2, H4b	<i>Homo sapiens</i>
HIST4H4	histone cluster 4, H4	<i>Homo sapiens</i>
HIST2H4A	histone cluster 2, H4a	<i>Homo sapiens</i>
HIST1H4E	histone cluster 1, H4e	<i>Homo sapiens</i>
HIST1H4F	histone cluster 1, H4f	<i>Homo sapiens</i>
HIST1H4D	histone cluster 1, H4d	<i>Homo sapiens</i>

Download

Step 2: The Gene database search

From the Homologene search result page, link to the Gene database.

NCBI HomoloGene Discover Homologs

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search HomoloGene for "txid2759"[Ancestor:noexp] AND "homologene cdd"[Fil] Go Clear Save Search

Limits Preview/Index History Clipboard Details

History has expired.

Display Summary Show 20 Send to

All: 28

Items 1

Conserved Domain Links

1: Homo sapiens Gene Links

HIST1 GEO Profile Links histone cluster 1, H4j Homo sapiens

HIST2 Related HomoloGene Links histone cluster 2, H4b Homo sapiens

HIST2 Nucleotide Links histone cluster 4, H4 Homo sapiens

HIST2 EST Links histone cluster 2, H4a Homo sapiens

HIST2 GSS Links histone cluster 1, H4e Homo sapiens

HIST2 OMIA Links

Select the human records.

Gene Gene Search

Limits Advanced

Display Settings: Summary, 20 per page, Sorted by Relevance

Send to: Filter your results:

Results: 1 to 20 of 45156

1: MTHFR - methylenetetrahydrofolate reductase (NAD(P)H) [Homo sapiens]

methylenetetrahydrofolate reductase (NAD(P)H)

Official Symbol: MTHFR

Other Aliases: RP11-56N19.4

Other Designations: 5,10-methylenetetrahydrofolate reductase (NADPH); methylenetetrahydrofolate reductase

Location: 1p36.3

Annotation: Chromosome 1, NC_000001.10 (11845787..11866160, complement)

MIM: 607093

ID: 4524

Order cDNA clone

2: ABCB1 - ATP-binding cassette, sub-family B (MDR/TAP), member 1 [Homo sapiens]

ATP-binding cassette, sub-family B (MDR/TAP), member 1

Official Symbol: ABCB1

Filter your results:

All (45156)

Current Only (45025)

Genes Genomes (45015)

SNP GeneView (18242)

In Variation Viewer (1180)

Manage File

Top Organisms [Tree]

Arabidopsis thaliana (3652)

Danio rerio (3030)

Oryza sativa Japonica Group (2990)

Homo sapiens (2885)

Mus musculus (2041)

All other taxa (29758)

More...

Note the search term in the Gene search box, once the page reloads:

Designation #10 comes from the Homologene link. It will be different each time.

The "Homo sapiens"[porgn: __txid9606] term has been added from selecting the human organism from the taxonomy sorter. This is also reflected in the search details on the right side of the screen.

Gene Gene Search

Limits Advanced

Display Settings: Summary, 20 per page, Sorted by Relevance

Send to: Filter your results:

Results: 1 to 20 of 2885

1: MTHFR - methylenetetrahydrofolate reductase (NAD(P)H) [Homo sapiens]

methylenetetrahydrofolate reductase (NAD(P)H)

Official Symbol: MTHFR

Other Aliases: RP11-56N19.4

Other Designations: 5,10-methylenetetrahydrofolate reductase (NADPH); methylenetetrahydrofolate reductase

Location: 1p36.3

Annotation: Chromosome 1, NC_000001.10 (11845787..11866160, complement)

MIM: 607093

ID: 4524

Order cDNA clone

2: ABCB1 - ATP-binding cassette, sub-family B (MDR/TAP), member 1 [Homo sapiens]

ATP-binding cassette, sub-family B (MDR/TAP), member 1

Official Symbol: ABCB1

Other Aliases: ABC20, CD243, CLCS, GP170, MDR1, P-GP, PGY1

Other Designations: P-glycoprotein 1; colchicin sensitivity; doxorubicin resistance; multidrug resistance protein 1

Location: 7q21.12

Annotation: Chromosome 7, NC_000007.13 (87133179..87342639, complement)

MIM: 171050

ID: 5243

Filter your results:

All (2885)

Current Only (2885)

Genes Genomes (2885)

SNP GeneView (2877)

In Variation Viewer (1180)

Manage Filters

Find related data

Database: Select

Find items

Search details

(#10) AND "Homo sapiens" [porgn]

Add "ras oncogene"[Gene/Protein Name] to the search term in the query box and run the search. The result of the search is 17 human Gene records for genes that have "ras oncogene" in their name. According to our selection, they all have identified conserved domains and they are conserved throughout Eukaryotic kingdom.

In find related data, select protein database.

Link to the Reference Sequence proteins for these genes.

Step 3: Protein records and protein alignment

After selecting find items. Some of the Genes have more than one transcript. These splice variants result in different protein isoforms.

You can select the longest isoforms of each gene by simply looking at the reported protein length on the result summary page. (You can send your selection to clipboard and save the items in My NCBI as a collection).

Check those below:

NP_004153.2 (215 aa), NP_005361.2 (207 aa), NP_112243.1 (201 aa), NP_116235.2 (216 aa), NP_057661.3 (208 aa)

Select the link to the COBALT tool to see their multiple alignments.

NP_116235 1 M--TY AYL FKYIIIGDTGVGKSCLLLQFTDKRFQPVHDLTIGVEFGARMVNIIDCKQIKLOIWDTAGOESFRSIT 72

NP_005361 1 MAKTY DYL FKLLIGDSCVGRKTCVLFREFSEDAFNSTFISTIGIDFKIRTIELDGKRIH

NP_116235 1 M--TY AYL FKYIIIGDTGVGKSCLLLQFTDKRFQPVHDLTIGVEFGARMVNIIDCKQIKLOIWDTAGOESFRSIT 72

NP_004153 1 MASRC[7]NTG[5]FKLVLLGSAVGVKSSLVLRVFKCQFHEFQESTIGAAFLTQTVCCLDDTTV

NP_057661 1 MSAGC DFG[5]FKLVFLCEQSVGRKTSLITRFMYDSFDNTYQATIGIDFLSKTHYLEDRIV

NP_116235 73 RSYRGAAGALLVYDITRRETFNHLTSLWLEDAHQHSSSNMVMILIGNKSDLESREDVVKREGEA

NP_005361 75 TAYYRGAMGIMLVYDITNEKSFNDIRNWIENIEEHASADVEKMLIGNKCDVNDKRVSKERGER

NP_112243 75 SSYRGAHGIIVVYDVTQESYANVVKQLQIEDRYASENVNKLVLGNKSDLTTRKVVNDNTAKR

NP_004153 87 PMYYRCAQAALVVYDITNEESFAPAKNVVKELQRQASPNIVIALSGNKADLANKRAVDFQEAQS

NP_057661 80 PSYIRDSTVAVVVYDITNLNSFQTSKWIDVVRTERCSDVIIMLVGNKTLADKRQITIERGEC

NP_116235 153 ACNVEEAFINTAKEIYRRIQ---QGLFDVHNNEANGIRIGPQQSIST[15]SNSGCC--- 216

NP_005361 155 NHTHNAEELADPTIATVDFVLEGRKQDGGNCGHETEDDQKR- SFFRCV11 207

ras-related protein Rab-2B isoform 1 [Homo sapiens]
>gi|296214433|ref|XP_002753820.1
PREDICTED: ras-related protein Rab-2B isoform 1 [Callithrix jacchus]
>gi|332223469|ref|XP_003260895.1
PREDICTED: ras-related protein Rab-2B-like isoform 1 [Nomascus leucogenys]
>gi|332223471|ref|XP_003260896.1
PREDICTED: ras-related protein Rab-2B-like isoform 2 [Nomascus leucogenys]
>gi|332841820|ref|XP_003314293.1

Result of their alignment, Select NP_116235 protein.

Step 4: Conserved domain and PSSM

Go to the protein record of Rab-2B and select the link to Conserved Domains to see its corresponding conserved domain.

ras-related protein Rab-2B isoform 1 [Homo sapiens]

NCBI Reference Sequence: NP_116235.2

[FASTA](#) [Graphics](#) [Item in clipboard](#)

[Go to](#)

LOCUS NP_116235 216 aa linear PRI 30-JUN-2012

DEFINITION ras-related protein Rab-2B isoform 1 [Homo sapiens].

ACCESSION NP_116235

VERSION NP_116235.2 GI:21361884

DBSOURCE REFSEQ: accession [NM_032846.3](#)

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM [Homo sapiens](#)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE 1 (residues 1 to 216)

AUTHORS Ostrowski,M., Carmo,N.B., Krumeich,S., Fanget,I., Raposo,G., Savina,A., Moita,C.F., Schauer,K., Hume,A.N., Freitas,R.P., Goud,B., Benaroch,F., Hacohen,N., Fukuda,M., Desnos,C., Seabra,M.C., Darchen,F., Amigorena,S., Moita,L.F. and Thery,C.

Change region shown

Customize view

Analyze this sequence

Run BLAST

Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

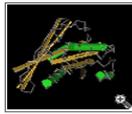
Protein 3D Structure

Crystal Structure
Rab2b
PDB: 2A5J
Source: Homo
Method: X-Ray
Resolution: 1.

The domain is Rab2 and it belongs to the P-loop_NTPase superfamily, link to the domain. You can see active GTP/Mg2+ binding site on the aligned sequences if you scroll down the record. You can also view the PSSM for the CD (link indicated).

A PSSM, or Position-Specific Scoring Matrix, is a type of scoring matrix used in protein BLAST searches in which amino acid substitution scores are given separately for each position in a protein multiple sequence alignment.

cd01866: Rab2, with user query added



Rab GTPase family 2 (Rab2)

Rab2 is localized on cis-Golgi membranes and interacts with Golgi matrix proteins. Rab2 is also implicated in the maturation of vesicular tubular clusters (VTCs), which are microtubule-associated intermediates in transport between the ER and Golgi apparatus. In plants, it regulates vesicle trafficking between the ER and the Golgi bodies and is important to pollen tube growth. GTPase activating proteins (GAPs) interact with GTP-bound Rab and accelerate the hydrolysis of GTP to GDP. Guanine nucleotide exchange factors (GEFs) interact with GDP-bound Rabs to promote the formation of the GTP-bound state. Rabs are further regulated by guanine nucleotide dissociation inhibitors (GDIs), which facilitate Rab recycling by masking C-terminal lipid binding and promoting cytosolic localization. Most Rab GTPases contain a lipid modification site at the C-terminus, with sequence motifs CC, CXC, or CCX. Lipid binding is essential for membrane attachment, a key feature of most Rab proteins. Due to the presence of truncated sequences in this CD, the lipid modification site is not available for annotation.

Links section containing Source: cd00154, Taxonomy: Eukaryota, PubMed: 19 links, Book: 4 links, Protein: Representatives, Specific Protein, Related Protein, Related Structure, Architectures, Superfamily: c109099, BioSystems: 239 links.

Conserved Features/Sites section showing Feature 1: GTP/Mg2+ binding site [chemical binding site]. Evidence includes Comment: The active conformation of Rab is stabilized by interactions between the gamma phosphate of GTP and two critically conserved residues, Thr in switch I and Gly in switch II. Citation: PMID 10196122. Structure: 2A5J: Human Rab2b binds GDP and Mg2+, defined using 3.5 A contacts. Includes buttons for Download Cn3D and Scroll to Sequence Alignment Display.

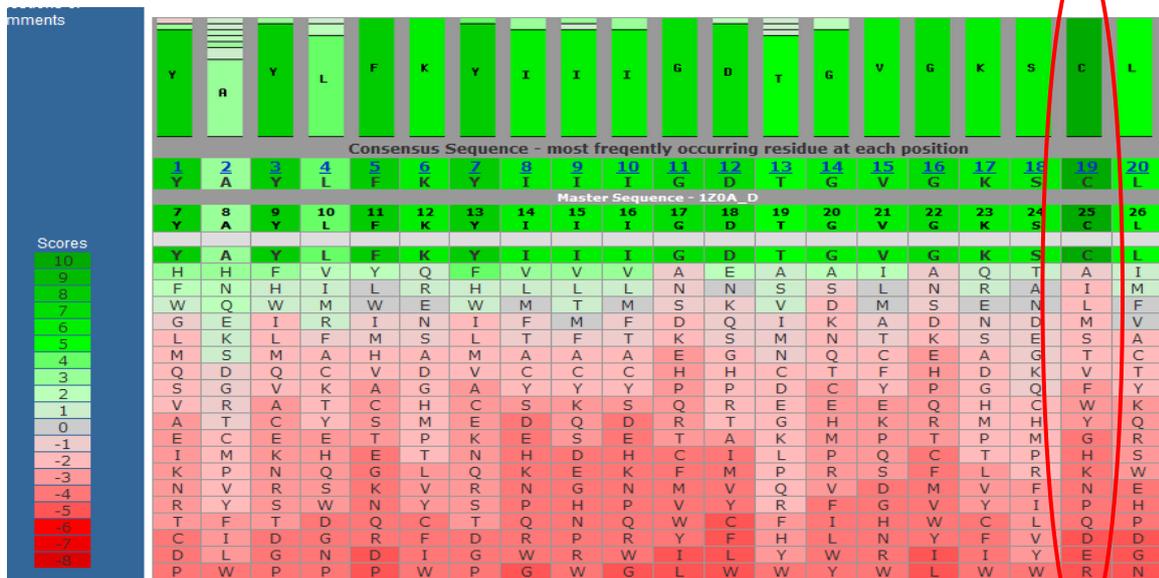
Statistics section showing PSSM-Id: 206658, View PSSM: cd01866, Aligned: 21 rows, Threshold Bit Score: 321.204.

cd01866 is part of a hierarchy of related CD models. Use the graphical representation to navigate this hierarchy. cd01866 is a member of the superfamily c109099.

Feature 1 alignment showing sequence 120A_D and query 2A5J_A with conserved residues highlighted by red circles. Conserved residues include Y, L, F, K, Y, I, I, I, G, D, T, G, V, G, K, S, C, L.

Feature 1 alignment showing sequence 120A_D and query 2A5J_A with conserved residues highlighted by red circles. Conserved residues include L, V, Y, R, E, G, E, A, F, A, R, E, H, G, L, I, F, M, E, T, S, A, K, T, A, C, N, V, E, E, A, F.

From the alignment, you can see the three conserved domains.



Cysteine at position 19 has score 10.

Step 5: BLAST searches

Select the link to the BLAST page from the protein record.

ras-related protein Rab-2B isoform 1 [Homo sapiens]

NCBI Reference Sequence: NP_116235.2

[FASTA](#) [Graphics](#) [Item in clipboard](#)

[Go to:](#)

LOCUS NP_116235 216 aa linear PRI 30-JUN-2012
 DEFINITION ras-related protein Rab-2B isoform 1 [Homo sapiens].
 ACCESSION NP_116235
 VERSION NP_116235.2 GI:21361884
 DBSOURCE REFSEQ: accession [NM_032846.3](#)
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM [Homo sapiens](#)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
 Catarrhini; Hominidae; Homo.
 REFERENCE 1 (residues 1 to 216)
 AUTHORS Ostrowski,M., Carmo,N.B., Krumeich,S., Fanger,I., Raposo,G.,
 Savina,A., Moita,C.F., Schauer,K., Hume,A.N., Freitas,R.P.,
 Goud,B., Benaroch,P., Hacohen,N., Fukuda,M., Desnos,C.,
 Seabra,M.C., Darchen,F., Amigorena,S., Moita,L.F. and Thery,C.

Change region shown

Customize view

Analyze this sequence

Run BLAST

Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

Protein 3D Structure



Crystal Structure
 Rab2b
 PDB: 2A5J
 Source: Homo sapiens
 Method: X-Ray
 Resolution: 1.8 Å

The link leads to the blastp page with the query accession added.

Change the database to search to the Reference proteins.

[NCBI](#) / [BLAST](#) / [blastp suite](#)

Standard Protein

[blastn](#) **[blastp](#)** [blastx](#) [tblastn](#) [tblastx](#)

BLASTP programs search protein databases

Enter Query Sequence

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

From

To

Or, upload file Aucun fichier choisi

Job Title

Enter a descriptive title for your BLAST search

[Align two or more sequences](#)

Choose Search Set

Database [+](#)

Organism [+](#)

[Reference proteins \(refseq_protein\)](#) will be suggested [Exclude](#) [+](#)

[UniProtKB/Swiss-Prot \(swissprot\)](#) will be suggested [Exclude](#) [+](#)

[Patented protein sequences \(pat\)](#) will be suggested [Exclude](#) [+](#)

[Protein Data Bank proteins \(pdb\)](#) will be suggested [Exclude](#) [+](#)

[Metagenomic proteins \(env_nr\)](#) will be suggested [Exclude](#) [+](#)

[Environmental sample sequences](#)

NP_116235.2 is the accession number of the query for the Rab2 isoform.

Add all the information, as you can see in the picture below.

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

From

To

Or, upload file Aucun fichier choisi [?](#)

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

Align two or more sequences [?](#)

Choose Search Set

Database [?](#)

Organism Optional

Escherichia (taxid:561)	<input type="checkbox"/> Exclude	+
Pseudomonas (taxid:286)	<input type="checkbox"/> Exclude	
Rhizobium (taxid:379)	<input type="checkbox"/> Exclude	
Burkholderia (taxid:32008)	<input type="checkbox"/> Exclude	
Listeria (taxid:1637)	<input type="checkbox"/> Exclude	
Streptococcus (taxid:1301)	<input type="checkbox"/> Exclude	

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

Increase the Max target sequences to 500.

Algorithm parameters **Note: Parameter values that differ from th**

General Parameters

Max target sequences [?](#)
Select the maximum number of aligned sequences to display

Short queries Automatically adjust parameters for short input sequences [?](#)

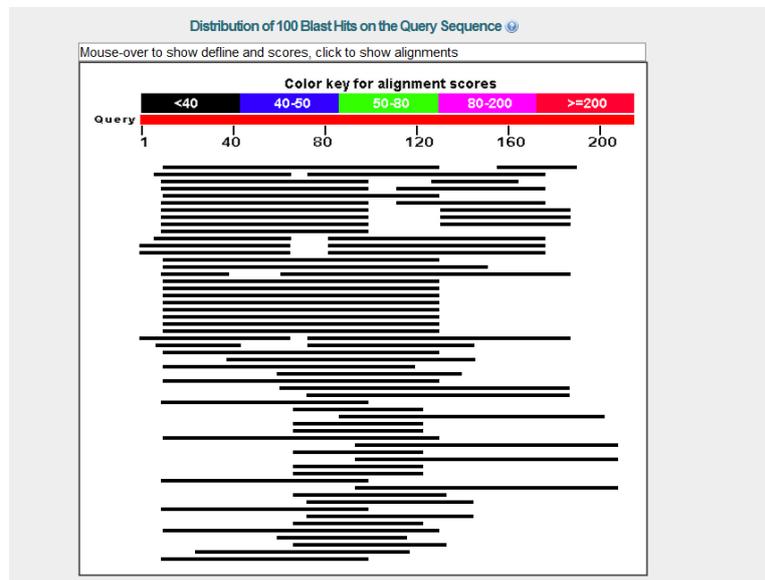
Expect threshold [?](#)

Word size [?](#)

Max matches in a query range [?](#)

BLASTp Output for RAB2:

Alignment result.



```

>ref|ZP_05638623.1| twitching motility protein [Pseudomonas syringae pv. tabaci str.
ATCC 11528]
Length=389

Score = 34.7 bits (78), Expect = 3e-04, Method: Compositional matrix adjust.
Identities = 25/91 (27%), Positives = 39/91 (43%), Gaps = 0/91 (0%)

Query 10  IIGDTGVGKSCSLLQFTDKRFQPVHDLTIGVEFGARMVNIDGKQIKLQIWDIAGQESFR 69
          + +G IG GKS D D R + I +E ++ K I Q SFR
Sbjct 127  LFVGTGSGKSTSAALIDYRNRNASGHIITIEDPVEFIHRHKSIVNQREVGVDIRSF 186

Query 70  SIIRSYRGAAGALLVVDITRRETFNHLTSW 100
          + ++ R A +L+ +I RET H ++
Sbjct 187  AALKNTLRQAPDVILIGEIRDRETMEHALAF 217
    
```

The blastp alignment covers only one of the active site.

DELTA-BLAST for RAB2:

Change the job title, replace blastp by delta.

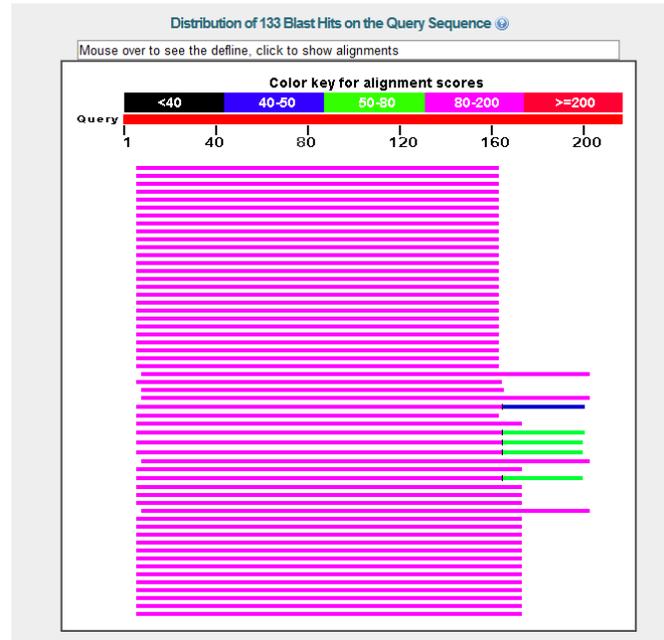
Select Delta-Blast for algorithm.

The screenshot shows the NCBI BLAST search interface. The 'Job Title' field contains 'RAB2B, delta, NP_116235.2'. The 'Algorithm' section has 'DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)' selected. Other visible fields include 'Enter accession number(s), gi(s), or FASTA sequence(s)' with 'ref|NP_116235.2', 'Query subrange' with 'From' and 'To' boxes, and 'Choose Search Set' with 'Reference proteins (refseq_protein)' selected in the database dropdown. The 'Organism' list includes Escherichia (taxid:561), Pseudomonas (taxid:286), Rhizobium (taxid:379), Burkholderia (taxid:32008), Listeria (taxid:1637), and Streptococcus (taxid:1301). The 'Program Selection' section shows 'blastp (protein-protein BLAST)', 'PSI-BLAST (Position-Specific Iterated BLAST)', 'PHI-BLAST (Pattern Hit Initiated BLAST)', and 'DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)'.

Increase the Max target sequences to 5000.

The screenshot shows the 'General Parameters' section of the NCBI BLAST search interface. The 'Max target sequences' dropdown menu is set to '5000'. Below it, the text reads 'Select the maximum number of aligned sequences to display'. Other parameters include 'Short queries' with 'Automatically adjust parameters for short input sequences' checked, 'Expect threshold' set to '10', and 'Word size' set to '3'.

The result:



GENE ID: 915398 trmE | tRNA modification GTPase TrmE
 [Escherichia coli O157:H7 str. Sakai] (10 or fewer PubMed links)

Score = 92.0 bits (227), Expect = 7e-22, Method: Composition-based stats.
 Identities = 37/166 (22%), Positives = 60/166 (36%), Gaps = 19/166 (11%)

```

Query 7   FKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMVNDGKQIKLQIWDTAGQ 65
          K +I G   GKS LL   +   V D   +   ++IDG + L I DTAG
Sbjct 217 MKVVIAGRNNAGKSSLLNALGREAIVVDIAGTTRDVLREHIIHDG--MPLHIIDTAGL 274

Query 66  ESFRS-----ITRSY--YRGAAGALLVVDITRRET FNHLTSWLEDARQHSSNMVIMLI 117
          I R++   A L + D T + +   W +   + + I ++
Sbjct 275 REASDEVERIGIERAWQEI EQADRVLFMVDGTTTDAVDPAEIW-PEFIARLPKLPITVV 333

Query 118 GNKSDLESRRNVKREEGEAFAREHGLIFMETSAKTACNVEEAFINT 163
          NK+D+   E + +G   SAKT V+ +
Sbjct 334 RNKADI-----TGETLGMSEVNGHALINSAKTGEGVDVLRNHL 372
    
```

```

Feature 1
120A_D 7 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMITIDGKQIKLQIWDTAGQESFRSITRSYYRGAAGA 86
query 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMVNDGKQIKLQIWDTAGQESFRSITRSYYRGAAGA 82
2A5J_A 18 GSYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMVNDGKQIKLQIWDTAGQESFRSITRSYYRGAAGA 97
gi 13397937 4 YEYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARLVNLDNKQIKLQIWDTAGQESFRSITRSYYRGAAGA 83
gi 28556900 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMITIDGKQIKLQIWDTAGQESFRSITRSYYRGAAGA 82
gi 30023652 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMITIDNKPIKQIWDTAGQESFRSITRSYYRGAAGA 82
gi 45382561 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMITIDGKQIKLQIWDTAGQESFRSITRSYYRGAAGA 82
gi 27881912 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMITIDGKQIKLQIWDTAGQESFRSITRSYYRGAAGA 82
gi 21645117 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMITIDGKQIKLQIWDTAGQEAFRSITRSYYRGAAGA 82
gi 49257196 3 YAYLFKYIIIGDTGVGKSCLLQLQFTDKRFQPVHDLTIGVEFGARMINIDGKPIKQIWDTAGQESFRSITRSYYRGAAGA 82
    
```

```

Feature 1
120A_D 87 LLVYDITRRDTFNHLTTSWLEDARQHSNSNMVIMLIGNKSDL [3].REVKKEEGEAFAREHGLIFMETSAKTACNVEEAF 164
query 83 LLVYDITRRET FNHLTSWLEDARQHSSNMVIMLIGNKSDL [3].RDVKREEGEAFAREHGLIFMETSAKTACNVEEAF 160
2A5J_A 98 LLVYDITRRET FNHLTSWLEDARQHSSNMVIMLIGNKSDL [3].RDVKREEGEAFAREHGLIFMETSAKTACNVEEAF 175
gi 13397937 84 LLVYDITRRET FNHLNRWLEVRQNSNPHMAIILVGNKSDL [2].REVSAAEQAQFARQNGLIFLETSAKTAKNVEEAF 160
gi 28556900 83 LLVYDITRRDTFNHLTTSWLEDARQHSNSNMVIMLIGNKSDL [3].REVKKEEGEAFAREHGLIFMETSAKTACNVEEAF 160
gi 30023652 83 LLVYDITRRET FNHLASWLEDARQHANAMTIIIGNKSDL [3].RAVSTEEGEQFAKEHGLIFMEASAKTAQNVEEAF 160
gi 45382561 83 LLVYDITRRDTFNHLTTSWLEDARQHSNSNMVIMLIGNKSDL [3].REVKKEEGEAFAREHGLIFMETSAKTASNVEEAF 160
gi 27881912 83 LLVYDITRRDTFNHLTTSWLEDARQHSNSNMVIMLIGNKSDL [3].REVKKEEGEAFAREHGLIFMETSAKTASNVEEAF 160
gi 21645117 83 LLVYDITRRET FNHLTTSWLEDARQHSNSNMVIMLIGNKSDL [3].REVKKEEGEAFAREHGLVFMETSAKTACNVEEAF 160
gi 49257196 83 LLVYDITRRET FNHLTTSWLEDARQHSNSNMVIMLIGNKSDL [3].RDVSRREEGEAFAREHGLIFMETSAKTASNVEEAF 160
    
```

For DELTA-BLAST, all three active sites are in the alignment.