

## Entrez Sequences Quick Start

Peter Cooper,<sup>1</sup> Melissa Landrum,<sup>1</sup> Ilene Mizrahi,<sup>1</sup> and Jane Weisemann<sup>1</sup>

Created: June 29, 2010; Updated: July 5, 2024.

This is a quick start guide for the Entrez Protein, and Nucleotide databases. The instructions here should allow you to quickly begin searching and using the features of the Entrez sequence databases.

- How do I use a simple query, such as a word or a phrase?
- How can I make my search more specific with Boolean operators (AND, OR, NOT)?
- How do I restrict my search to specific subsets of records such as those from a specific organism, molecule type or source database?
- How can I change the format, number, or sorting order of records displayed?
- How do I download sequence records to a file on my computer?
- How can I change the information that is shown such as optional biological features or sequence?
- How do I analyze the sequence data directly or find additional related data?
- How can I search for a sub-sequence, or pattern in a protein or nucleotide sequence?
- How can I locate and highlight a biological feature in a protein or nucleotide sequence?

### How do I use a simple query, such as a word or a phrase?

You can use a protein name, gene name, or gene symbol directly. Searching with a submitter or author name in the following format will produce the best results.

Smith JR (last name followed by initials, no punctuation)

Database identifiers such as accession numbers or gi numbers will directly retrieve the full sequence record.

CAA79696  
NP\_778203  
263191547  
BC043443  
NM\_002020

To find a match to an exact phrase, enclose it in quotation marks.

"contactin associated protein"  
"duchenne muscular dystrophy"

## How can I make my search more specific with Boolean operators (AND, OR, NOT)?

Use the Boolean operator AND to find records that contain every one of your search terms, the intersection of search results.

contactin AND neurofascin                      Protein                      Nucleotide

Use the Boolean operator OR to find records that include one of several search terms, the union of search results.

contactin OR neurofascin                      Protein                      Nucleotide

Use the Boolean operator NOT to exclude records matching a search term

contactin NOT neurofascin                      Protein                      Nucleotide

## How do I restrict my search to specific subsets of records such as those from a specific organism, molecule type or source database?

You can use the *Facets* on the left-hand side of the page to show only certain kinds of records. Follow these links to jump to the Facet of interest: organism, molecule type, source database.

### Facets

Use the facets on the left-hand side of any of the Protein, Nucleotide results pages to restrict the types of records shown.

### Organism

To get records from a specific organism or group of organisms click the appropriate Species filter. You can use the Customize option to add a filter for a particular organism or group of organisms. You use the common or scientific name of a species, strain, or higher taxon as a Filter term. Examples: human, *Mus musculus*, *Drosophila similis*, green plants, bacteria.

Protein
Protein

[Create alert](#) [Advanced](#)

**Species**

- Animals (75,552,140)
- Plants (36,109,409)
- Fungi (63,462,811)
- Protists (16,181,837)
- Bacteria (1,028,848,48)
- Archaea (9,334,488)
- Viruses (58,773,122)
- Customize ...

**Source databases**

- PDB (845,410)
- RefSeq (330,154,754)
- UniProtKB / Swiss-Prot (571,788)
- Customize ...

**Genetic compartme...**

- Chloroplast (4,971,350)
- Mitochondrion (8,174,2)
- Plasmid (7,493,668)
- Plastid (5,617,155)

**Sequence length**

Custom range...

**Molecular weight**

Custom range...

**Release date**

Custom range...

**Revision date**

Custom range...

[Clear all](#)

[Show additional filters](#)

Summary ▾ 20 per page ▾ Sort by Default order ▾
Send to: ▾

See the [results of this search \(770605816 items\)](#) in our new [Identical Protein Groups](#) database.

**Items: 1 to 20 of 1307128206**

<< First < Prev
Page 
of 65356411 Next > Last >>

- [ChbG/HpnK family deacetylase \[Vibrio breoganii\]](#)
- 1. 235 aa protein  
 Accession: WP\_349254351.1 GI: 2741237557  
[PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [GpE family phage tail protein \[Pasteurella atlantica\]](#)
- 2. 36 aa protein  
 GI: 2741237556  
[FASTA](#) [Graphics](#)
- [hypothetical protein ABMQ32\\_27345 \(plasmid\) \[Klebsiella pneumoniae\]](#)
- 3. 210 aa protein  
 Accession: XBP90586.1 GI: 2741237540  
[Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [VirB3 family type IV secretion system protein \(plasmid\) \[Klebsiella pneumoniae\]](#)
- 4. 917 aa protein  
 Accession: XBP90585.1 GI: 2741237539  
[Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [TrbC/VirB2 family protein \(plasmid\) \[Klebsiella pneumoniae\]](#)
- 5. 96 aa protein  
 Accession: XBP90584.1 GI: 2741237538

# Facets

Protein   [Create alert](#) [Advanced](#)

Species [Summary](#) [20 per page](#) [Sort by Default order](#) [Send to:](#)

Animals (75,552,140)  
Plants (36,109,409)  
Fungi (63,462,811)  
Protists (16,181,837)  
Bacteria (1,028,848,48)  
Archaea (9,334,488)  
Viruses (58,773,122)

Customize ...

Source databases  
PDB (845,410)  
RefSeq (330,1...)  
UniProtKB / Sw...  
Prot (571,788...)  
Customize ...

Genetic compartme...  
Chloroplast (4...)  
Mitochondrion  
Plasmid (7,493...)  
Plastid (5,617...)

Sequence length  
Custom range...

Molecular weight  
Custom range...

Release date  
Custom range...

Revision date  
Custom range...

917 aa protein  
Accession: XBP90585.1 GI: 2741237539

See the [results of this search \(770605816 items\)](#) in our new [Identical Protein Groups](#) database.

Items: 1 to 20 of 1307128206

Page 1 of 65356411 [Next >](#) [Last >>](#)

Species

- Animals
- Plants
- Fungi
- Protists
- Bacteria
- Archaea
- Viruses

Add

- 
- Mus musculus
- Mus musculus domesticus
- Mus musculus praetextus
- Mus musculus castaneus
- Mus musculus hortulanus
- Mus musculus musculus
- Mus musculus spretus
- [See all results](#)

[\[Mus musculus \(Mus musculus\)\]](#)

557

[Graphics](#)

[\[Klebsiella atlantica\]](#)

556

[Graphics](#)

[\[Klebsiella \(plasmid\)\] \[Klebsiella\]](#)

0

[Graphics](#)

[\[Klebsiella \(plasmid\)\] \[Klebsiella\]](#)

Protein Protein  [Create alert](#) [Advanced](#)

Species Summary 20 per page Sort by Default order Send to:

Animals (75,552,140)  
Plants (36,109,409)  
Fungi (63,462,811)  
Protists (16,181,837)  
Bacteria (1,028,848,48)  
Archaea (9,334,488)  
Viruses (58,773,122)

See the [results of this search \(770605816 items\)](#) in our new [Identical Protein Groups](#) database.

Items: 1 to 20 of 1307128206

Customize ... Page 1 of 65356411 [Next >](#) [Last >>](#)

Source databases  
PDB (845,410)  
RefSeq (330,1...)  
UniProtKB / S...  
Prot (571,788...)  
Customize ...

Genetic compartme...  
Chloroplast (4...)  
Mitochondrion...  
Plasmid (7,493...)  
Plastid (5,617...)

Sequence length  
Custom range

Molecular weight

**Species**

- Animals
- Plants
- Fungi
- Protists
- Bacteria
- Archaea
- Viruses
- Mus musculus

Add

**Show**

[o breoganii](#)  
657  
[Graphics](#)  
[rella atlantica](#)  
656  
[Graphics](#)  
[\(plasmid\) \[Klebsiella](#)

Taxonomy

Protein   [Create alert](#) [Advanced](#)

Species

Animals (75,552,140)  
Plants (36,109,409)  
Fungi (63,462,811)  
Protists (16,181,837)  
Bacteria (1,028,848,48)  
Archaea (9,334,488)  
Viruses (58,773,122)  
**Mus musculus**  
Customize ...

Source databases  
PDB (845,410)  
RefSeq (330,154,754)  
UniProtKB / Swiss-Prot (571,788)  
Customize ...

Genetic compartme...  
Chloroplast (4,971,350)  
Mitochondrion (8,174,122)  
Plasmid (7,493,668)  
Plastid (5,617,155)

Sequence length  
Custom range...

Molecular weight  
Custom range...

Release date  
Custom range...

Revision date  
Custom range...

[Clear all](#)

See the [results of this search \(770605816 items\)](#) in our new [Identical Protein Groups](#) database.

**Items: 1 to 20 of 1307128206**

[<< First](#)
[< Prev](#)
Page  of 65356411
[Next >](#)
[Last >>](#)

- [ChbG/HpnK family deacetylase \[Vibrio breoganii\]](#)  
1. 235 aa protein  
Accession: WP\_349254351.1 GI: 2741237557  
[PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [GpE family phage tail protein \[Pasteurella atlantica\]](#)  
2. 36 aa protein  
Accession: WP\_349254350.1 GI: 2741237556  
[PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [hypothetical protein ABMQ32\\_27345 \(plasmid\) \[Klebsiella pneumoniae\]](#)  
3. 210 aa protein  
Accession: XBP90586.1 GI: 2741237540  
[Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [VirB3 family type IV secretion system protein \(plasmid\) \[Klebsiella pneumoniae\]](#)  
4. 917 aa protein  
Accession: XBP90585.1 GI: 2741237539  
[Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)



Species [clear](#) Summary ▾ 20 per page ▾ Sort by Default order ▾ [Send to: ▾](#)

Animals (376,274)  
Plants (0)  
Fungi (0)  
Protists (0)  
Bacteria (0)  
Archaea (0)  
Viruses (0)  
✓ **Mus musculus** (376,276)  
[Customize ...](#)

Source databases  
PDB (26,410)  
RefSeq (97,925)  
UniProtKB / Swiss-Prot (17,231)  
[Customize ...](#)

Genetic compartments  
Mitochondrion (7,228)  
Plasmid (24)

Sequence length  
Custom range...

Molecular weight  
Custom range...

Release date  
Custom range...

Revision date  
Custom range...

[Clear all](#)  
[Show additional filters](#)

See the [results of this search](#) (770605816 items) in our new [Identical Protein Groups](#) database.

**Items: 1 to 20 of 376276**

[1st](#) [Prev](#) Page **1** of 18814 [Next](#) [Last >>](#)

**Filters activated: Mus musculus. [Clear all](#)**

[corticotropin-releasing factor receptor 1 isoform 3 \[Mus musculus\]](#)

1. 210 aa protein  
Accession: NP\_001417832.1 GI: 2740888219  
[Nucleotide](#) [PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[ceruloplasmin isoform a precursor \[Mus musculus\]](#)

2. 1086 aa protein  
Accession: NP\_001396576.1 GI: 2274796033  
[BioProject](#) [Nucleotide](#) [PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[ceruloplasmin isoform g precursor \[Mus musculus\]](#)

3. 1057 aa protein  
Accession: NP\_001396579.1 GI: 2274795974  
[BioProject](#) [Nucleotide](#) [PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[ceruloplasmin isoform a precursor \[Mus musculus\]](#)

4. 1086 aa protein  
Accession: NP\_001396575.1 GI: 2274795947  
[BioProject](#) [Nucleotide](#) [PubMed](#) [Taxonomy](#)  
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

You can also use the linked numbers in the Top Organisms list in the right-hand column of search results to filter select records from specific organisms from your results.

**Results by taxon** [▲](#)

**Top Organisms** [\[Tree\]](#)

[Escherichia coli \(76339685\)](#)  
[Klebsiella pneumoniae \(53203659\)](#)  
[Salmonella enterica \(50229600\)](#)  
[Streptococcus pneumoniae \(46548930\)](#)  
[Severe acute respiratory syndrome-related coronavirus \(45251751\)](#)  
[All other taxa \(1035554581\)](#)  
[More...](#)

## Molecule type

In the Nucleotide database you can use the Molecule types facet to limit results to a particular molecule type.

**Molecule types**

genomic

DNA/RNA (275,830,428)

mRNA (230,506,414)

rRNA (873,387)

Customize ...

**Source database**

The Source databases facet allows you to limit to results from a particular database.



**Source databases**  
INSDC (GenBank) (506,574,099)  
RefSeq (113,128,092)  
Customize ...

[Camellia sinensis var. sinensis Seimei DNA, Chr1](#)  
2. **243,588,495 bp linear DNA**  
Accession: BTIW01000001.1 GI: 2741835856  
[BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)

**Sequence Type**  
Nucleotide (51)  
EST (78,010,5)  
GSS (40,839,1)

**Genetic compartment**  
Chloroplast (1,  
Mitochondrion  
Plasmid (346,4  
Plastid (1,877,2

**Sequence length**  
Custom range...

DDBJ  
 EMBL  
 GenBank  
 INSDC (GenBank)  
 RefSeq  
 PDB  
 TPA

**Show**

**Nucleotide**

[Seimei DNA, Chr2](#)  
1828566  
[Taxonomy](#)

[Seimei DNA, Chr3](#)  
1821740  
[Taxonomy](#)

**Source databases**  
PDB (845,410)  
RefSeq (330,149,765)  
UniProtKB / Swiss-Prot (571,788)  
Customize ...

[hypothetical protein, partial \[Streptomyces sp. H10-C2\]](#)

**Genetic compartment**  
Chloroplast (4,  
Mitochondrion  
Plasmid (7,496  
Plastid (5,621,0

**Sequence length**  
Custom range...

**Molecular weight**  
Custom range...

PDB  
 RefSeq  
 UniProtKB / Swiss-Prot  
 DDBJ  
 EMBL  
 GenBank  
 PIR

**Show**

**Protein**

173 aa protein  
Accession: WP\_349307973.1 GI: 2741905073  
[Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

72  
[es sp. H10-C2\]](#)  
1905071  
[TA](#) [Graphics](#)  
[es sp. H10-C2\]](#)

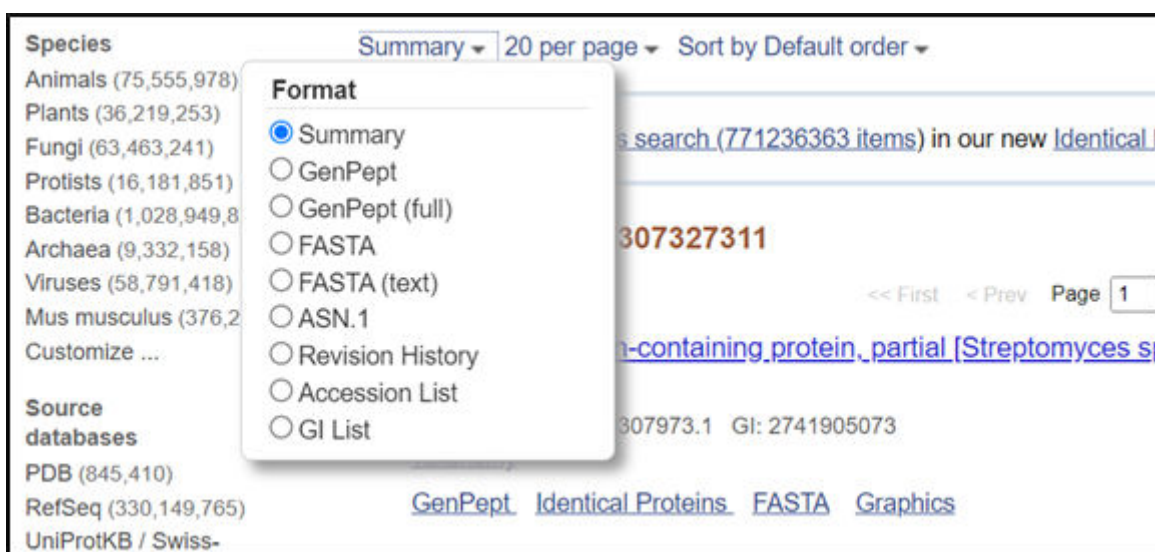
Accession: WP\_349307970.1 GI: 2741905070  
[PubMed](#) [Taxonomy](#)

The source databases for NCBI nucleotide and protein sequences are listed below.

- **Protein:** SwissProt and PIR components of UniProt; Protein Research Foundation (PRF); protein chain sequences from the Protein Data Bank (PDB); and translations of coding regions on sequences in Entrez Nucleotide (RefSeq, International Sequence Database Collaboration – DDBJ / ENA-EMBL / GenBank, Third Party Annotation (TPA).
- **Nucleotide:** International Sequence Database Collaboration (DDBJ / ENA-EMBL / GenBank); NCBI Reference Sequences (RefSeq); Nucleotide sequences from PDB; Third Party Annotation (TPA).

## How do I change the format, number, or sorting order of records displayed?

The menus at the upper left of the results page headed by *Summary*, *20 per page* and *Sort by Default Order* allow you to change the format displayed, the number of records and the sorting order respectively. Click any of these and select the desired format, items per page, or sorting order from the listed radio buttons. The new settings will apply automatically.



## How can I download sequence records to a file on my computer?

Click the *Send to* menu that appears at the upper right of document summaries or record views and select the file radio button. Then choose the desired format from the pull-down list. Click the *Create File* button to save the records.

The screenshot shows the Entrez Sequences interface. On the left, there is a list of species: Animals (75,555,978), Plants (36,219,253), Fungi (63,463,241), Protists (16,181,851), Bacteria (1,028,949,8), Archaea (9,332,158), Viruses (58,791,418), Mus musculus (376,2), and a 'Customize ...' link. Below this is a 'Source databases' section with links for PDB (845,410), RefSeq (330,149,765), and UniProtKB / Swiss-. In the center, a 'Format' dialog box is open, showing radio buttons for 'Summary' (selected), 'GenPept', 'GenPept (full)', 'FASTA', 'FASTA (text)', 'ASN.1', 'Revision History', 'Accession List', and 'GI List'. The background shows a search result for '307327311' with a link to 'n-containing protein, partial [Streptomyces sp]'. At the bottom, there are links for 'GenPept', 'Identical Proteins', 'FASTA', and 'Graphics'.

## How do I change the information that is shown such as optional biological features or sequence?

Open the *Customize View* dialog that appears in the right-hand column of a record display. You can change the kinds of biological features shown and toggle the sequence on or off using the radio buttons and check boxes. Click the *Update View* button to activate the changes.



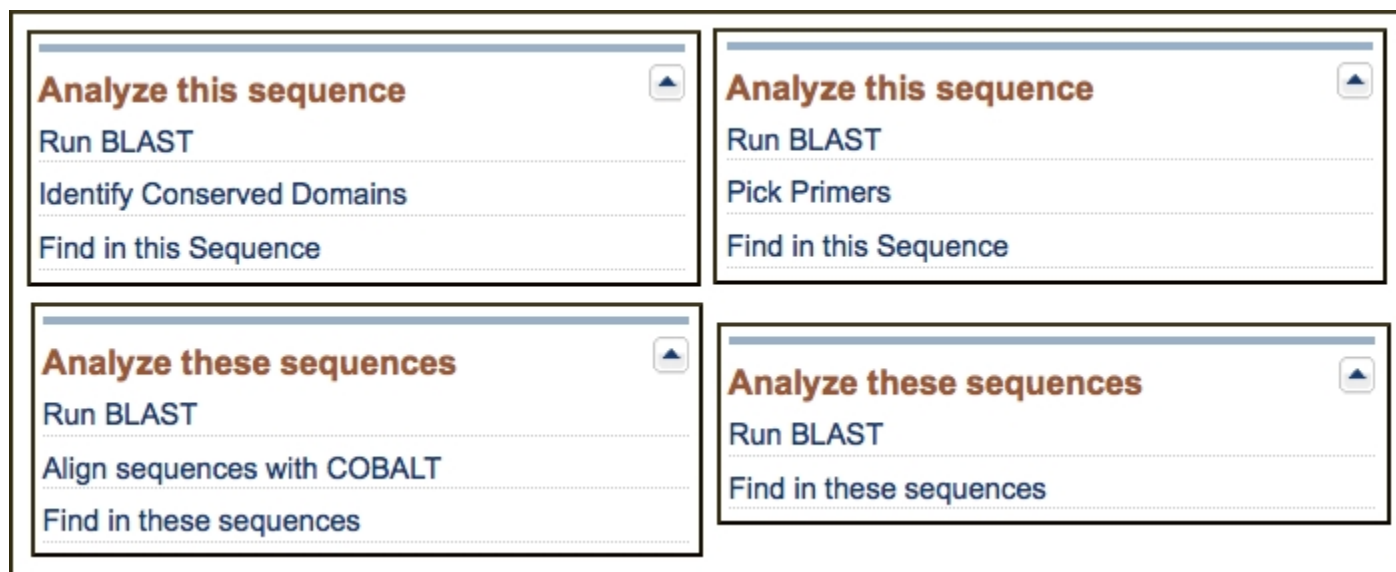
## How can I display a portion of the sequence?

Open the *Change region shown* dialog that appears in the right-hand column of a record display. You can change the kinds of biological features shown and toggle the sequence on or off using the radio buttons and check boxes. Click the *Update View* button to activate the changes.

The 'Change region shown' dialog box is shown. It has a title bar with a close button. Below the title bar are two radio buttons: 'Whole sequence' (selected) and 'Selected region'. Below the radio buttons are two input fields: 'from: begin' and 'to: end'. At the bottom right is an 'Update View' button.

## How do I analyze the sequence data directly or find additional related data?

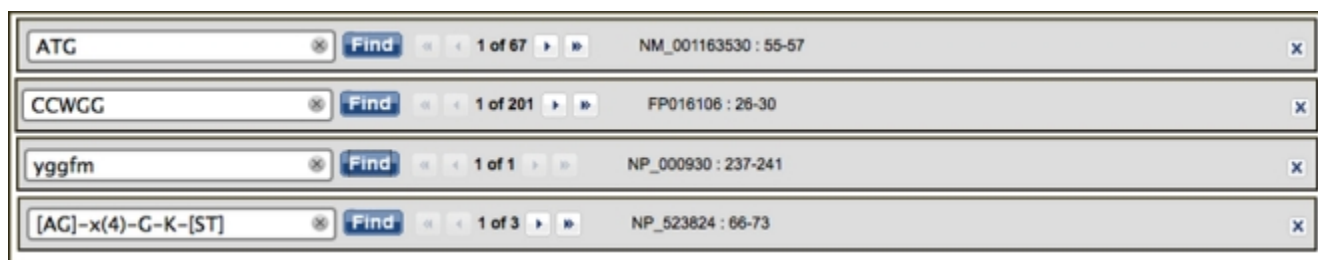
There are direct links to analysis tools including BLAST, Primer-BLAST (Nucleotide), and Conserved Domain Database Search (Protein) in the right-hand column of displayed records.



There are also links to related data in the right-hand column that may provide additional information and pre-computed analyses for the displayed records.

## How can I search for a sub-sequence, or pattern in a protein or nucleotide sequence?

You can access the *Find-in-sequence* feature in the Analysis tools in the right-hand column of single and multiple-record displays. This tool can find sub-sequences or patterns in displayed nucleotide or protein sequences. Clicking the **Find-in-this-Sequence** or **Find-in-these sequences** link opens a search box bar at the bottom of the page.



Find-in-sequence works with single and multiple sequence displays with any format that shows the sequence (GenBank, GenPept, FASTA). The tool can find sub-sequences and patterns typed in the box and works with standard (IUPAC) nucleotide and protein single letter and ambiguity codes as well as [Prosite patterns](#) that match motifs and domain signatures in protein sequences. Valid single letter codes are given below.

| Nucleotide Codes |                     |   |            |
|------------------|---------------------|---|------------|
| A                | adenosine           | Y | T or C     |
| C                | cytidine            | M | A or C     |
| G                | guanine             | W | A or T     |
| T                | thymidine           | R | G or A     |
| N                | A, G, C, or T       | B | G, T, or C |
| U                | uridine (matches T) | D | G, A, or T |
| K                | G or T              | H | A, C, or T |

Table continued from previous page.

|   |        |   |            |
|---|--------|---|------------|
| S | G or C | V | G, C, or A |
|---|--------|---|------------|

| Amino Acid Codes |                      |   |                     |
|------------------|----------------------|---|---------------------|
| A                | alanine              | N | asparagine          |
| B                | aspartate/asparagine | P | proline             |
| C                | cysteine             | Q | glutamine           |
| D                | aspartate            | R | arginine            |
| E                | glutamate            | S | serine              |
| F                | phenylalanine        | T | threonine           |
| G                | glycine              | V | valine              |
| H                | histidine            | W | tryptophan          |
| I                | isoleucine           | Y | tyrosine            |
| K                | lysine               | Z | glutamate/glutamine |
| L                | leucine              | X | any                 |
| M                | methionine           |   |                     |

Find matches by clicking the find button. The first 500 matches are highlighted for each displayed sequence. The first or current match is highlighted in white text on a dark background in the sequence, and its position is shown in the search bar. The other matches are highlighted with a light blue background. The tool ignores spaces and line breaks in the formatted sequence. Clicking the arrow keys jumps to the next or previous match.

## How can I locate and highlight a biological feature in a protein or nucleotide sequence?

You can highlight a feature by clicking on linked feature in the FEATURES table of a displayed nucleotide or protein sequence. A portion of a FEATURES table is shown below for a nucleotide sequence ([NG\\_008957](#)).



| FEATURES               | Location/Qualifiers  |
|------------------------|--|
| <a href="#">source</a> | 1..97661<br>/organism="Homo sapiens"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:9606"<br>/chromosome="X"<br>/map="Xp11.3"   |
| <a href="#">gene</a>   | 5139..95657<br>/gene="MAOA"<br>/gene_synonym="BRNRS; MAO-A"<br>/note="monoamine oxidase A"<br>/db_xref="GeneID:4128"<br>/db_xref="HGNC:HGNC:6833"<br>/db_xref="MIM:309850"   |
| <a href="#">mRNA</a>   | join(5139..5254,32353..32447,42131..42268,60712..60816,<br>61545..61636,77013..77154,80081..80230,80534..80693,<br>81539..81635,85067..85120,89521..89578,90790..90887,<br>92634..92745,92949..93011,93207..95657)<br>/gene="MAOA"<br>/gene_synonym="BRNRS; MAO-A"<br>/product="monoamine oxidase A, transcript variant 1"<br>/transcript_id="NM_000240.4"<br>/db_xref="GeneID:4128"<br>/db_xref="HGNC:HGNC:6833"<br>/db_xref="MIM:309850" |
| <a href="#">exon</a>   | 5139..5254<br>/gene="MAOA"<br>/gene_synonym="BRNRS; MAO-A"<br>/inference="alignment:Splign:2.1.0"<br>/number=1   |

Clicking the feature activates the feature search bar that appears at the bottom of the display and highlights the corresponding residues in the display as shown below for an exon feature in the RefSeq gene record for the MAOA gene (NG\_008957).

```

4921 cgagtgtcag tacaagggtc cgccccgctc tcagtgccca gctccccccg ggtatcagct
4981 gaaacatcag ctccgccct gggcgtccc ggagtatcag caaaagggtt cgccccgcc
5041 acagtcccc gctcccccc ggtatcaaaa gaaggatcgg ctccgcccc gggctcccc
5101 ggggagtiga tagaagggtc cttcccacc ttgccgtcc cactcctgt gcctacgacc
5161 caggagcgtg tcagccaaag catggagaat caagagaagg cgagtatcgc ggccacatg
5221 ttcgacgtag tcgtgatcgg aggtggcatt tcaggtcagt gtggaccgta gcggtggcct
5281 gggggaccct ggccagttag ggtagggga acctacagta gctcttggg tgtttggggg
5341 tctctcatc atccagact atactatag catgcttag cccatatac tccatatac

```

5139..5254  
/gene="MAOA"  
/gene\_synonym="BRNRS; MAO-A"  
/inference="alignment:Splign:2.1.0"  
/number=1

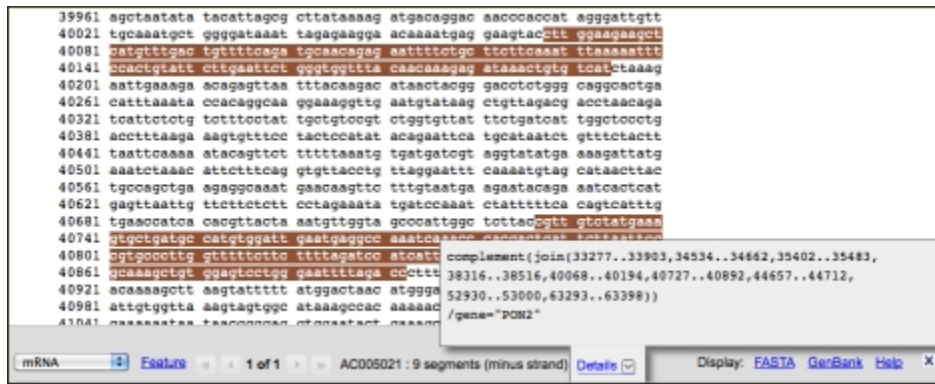
exon  1 of 15 NG\_008957 : 1 segment

Display: [FASTA](#) [GenBank](#) [Help](#)

The “Details” box that shows the annotation from the FEATURES table for the highlighted location can be collapsed if desired by clicking the link. Clicking the “Details” link again re-opens the box.

Discontiguous features that have multiple segments such as mRNA alignments on genomic DNA can also be highlighted. In all cases the number of segments is shown at the right of the sequence accession. Opposite strand features are indicated with the notation “minus strand” to the right of the number of segments of the bar. The image below shows mRNA minus strand feature for the PON2 gene from an annotated BAC clone sequence (AC005021).





## Navigating Using the Feature Highlight Bar

If there is more than one feature of the same type, as in the first example shown above, the navigational arrows on the bar allow jumping to the next, previous, first, and last instances of that feature. The Feature pull-down list at the right-hand side of the bar allows selecting other available feature types. The highlight moves to the next available instance of the selected feature type. The “Feature” link returns the display to the corresponding position in the FEATURES table of the record.

## Displaying Highlighted Regions as Separate Sequences

The FASTA and GenBank links on the right-hand side of the bar present the highlighted sub-sequence in these formats in the Nucleotide or Protein Entrez system and provide a simple means to display and download the corresponding sequence or to forward it to the analysis available analysis tools: BLAST, Primer-BLAST, Find in this Sequence, and Identify Conserved Domains (protein only).