

Entrez Sequences Quick Start

Peter Cooper,¹ Melissa Landrum,¹ Ilene Mizrahi,¹ and Jane Weisemann¹

Created: June 29, 2010; Updated: April 22, 2016.

This is a quick start guide for the Entrez Protein, Nucleotide, Expressed Sequence Tag (EST), and Genome Survey Sequence (GSS) databases. The instructions here should allow you to quickly begin searching and using the features of the Entrez sequence databases.

- Which of the three databases (nucleotide, EST, or GSS) should I search?
- 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?

Which of the three databases containing nucleic acid sequence (Nucleotide, EST, or GSS) should I search?

The Nucleotide, Genome Survey Sequence (GSS), and Expressed Sequence Tag (EST) database all contain nucleic acid sequences. The data in GSS and EST are from two large bulk sequence divisions of GenBank. GSS and EST data are typically uncharacterized, short genomic (GSS) or cDNA (EST) sequences.

Searching any of the three databases will provide links to results in the other. Unless you know that you are trying to find a specific set of EST or GSS sequences, searching the Nucleotide database with general text queries will produce the most relevant results. You can always follow links to results in EST and GSS from the Nucleotide database results.

The screenshot shows the Entrez Sequences search results for the query 'Kinase'. The search was performed in the 'Nucleotide' database. The results page includes a sidebar with filters for Species (Animals, Plants, Fungi, Protists, Bacteria, Archaea, Viruses, Customize) and Molecule types (genomic DNA/RNA). The main content area shows 'Items: 1 to 20 of 4840314' and a summary of 5641676 nucleotide sequences. The first result is 'Arabidopsis thaliana chromosome 1 sequence', which is 30,427,671 bp linear DNA with accession CP002684.1 and GI: 332189094. Links for GenBank, FASTA, and Graphics are provided for this result.

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, GSS, or EST webpages to restrict the types of records shown.

The screenshot shows the NCBI Protein search interface. On the left, there is a 'Facets / Filters' sidebar with the following categories:

- Species**: Animals (18,104,022), Plants (6,528,053), Fungi (8,339,996), Protists (3,215,785), Bacteria (230,345,242), Archaea (2,522,481), Viruses (3,453,766), Customize ...
- Source databases**: PDB (321,441), RefSeq (63,304,148), UniProtKB / Swiss-Prot (551,013), Customize ...
- Genetic compartments**: Chloroplast (696,930), Mitochondrion (3,179,863), Plasmid (840,449), Plastid (807,670)
- Sequence length**: Custom range...
- Molecular weight**: Custom range...
- Release date**: Custom range...
- Revision date**: Custom range...
- [Clear all](#)
- [Show additional filters](#)

The main search results area shows:

- Search type: Protein, Filter: all[filter]
- Buttons: Create alert, Advanced
- Summary: 20 per page, Sort by Default order, Send to: v
- Items: 1 to 20 of 283622511
- Navigation: << First, < Prev, Page 1 of 14181126, Next >, Last >>
- Results list:
 - [PREDICTED: semaphorin-3D-like, partial \[Sinocyclocheilus grahami\]](#)
1. 230 aa protein
Accession: XP_016120447.1 GI: 1020615144
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
 - [PREDICTED: disheveled-associated activator of morphogenesis 2-like, partial \[Sinocyclocheilus grahami\]](#)
2. 238 aa protein
Accession: XP_016120446.1 GI: 1020615139
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
 - [PREDICTED: ch repeat-containing protein 30-like, partial \[Sinocyclocheilus grahami\]](#)
Accession: XP_016120445.1 GI: 1020615134
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
 - [PREDICTED: methylosome subunit pICln-like, partial \[Sinocyclocheilus grahami\]](#)
4. 129 aa protein
Accession: XP_016120443.1 GI: 1020615128
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
 - [PREDICTED: neuroligin-3-like, partial \[Sinocyclocheilus grahami\]](#)
5. 225 aa protein
Accession: XP_016120442.1 GI: 1020615123
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
 - [PREDICTED: serine/threonine-protein kinase 38-like, partial \[Sinocyclocheilus grahami\]](#)
6. 187 aa protein
Accession: XP_016120441.1 GI: 1020615118
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

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 or 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.

Species Summary ▾ 20 per page ▾ Sort by Default order ▾

Animals (18,104,022)
Plants (6,528,053)
Fungi (8,339,996)
Protists (3,215,785)
Bacteria (230,345,242)
Archaea (2,522,481)
Viruses (3,453,766)
Customize ...

Source datab
PDB (321,441)
RefSeq (63,304)
UniProtKB / Sw
Prot (551,013)
Customize ...

Genetic
compartment
Chloroplast (69)
Mitochondrion
Plasmid (840,4)
Plastid (807,67)

Sequence len
Custom range...

Molecular wei
Custom range...

Release date
Custom range...

Revision date
Custom range...

Items: 1 to 20 of 283622511

<< First < Prev Page 1

[PREDICTED: semaphorin-3D-like, partial \[Sinocyclocheilus grahami\]](#)

1. 230 aa protein

Species ×

Animals

Plants

Fungi

Protists

Bacteria

Archaea

Viruses

Add

Mus musculus

Mus musculus

Mus musculus domesticus

Mus musculus praetextus

Mus musculus castaneus

Mus musculus hortulanus

Mus musculus musculus

Mus musculus spretus

See all results

1020615144
[STA](#) [Graphics](#)

[sociated activator of morphogenesis 2-like, p](#)

1020615139
[STA](#) [Graphics](#)

[peat-containing protein 30-like, partial \[Sinoc](#)

1020615134
[STA](#) [Graphics](#)

[subunit pICln-like, partial \[Sinocyclocheilus g](#)

1020615128
[FASTA](#) [Graphics](#)

[3-like, partial \[Sinocyclocheilus grahami\]](#)

GI: 1020615123
[FASTA](#) [Graphics](#)

Species Summary ▾ 20 per page ▾ Sort by Default order ▾

Animals (18,104,022)
Plants (6,528,053)
Fungi (8,339,996)
Protists (3,215,785)
Bacteria (230,345,242)
Archaea (2,522,481)
Viruses (3,453,766)
Customize ...

Items: 1 to 20 of 283622511

<< First < Prev Page 1

[PREDICTED: semaphorin-3D-like, partial \[Sinocyclocheilus grahami\]](#)

1. 230 aa protein

Species ×

Animals

Plants

Fungi

Protists

Bacteria

Archaea

Viruses

Mus musculus

Add

Sequence length

Custom range

Molecular weight

Custom range

Show

1020615144
[STA](#) [Graphics](#)

[associated activator of morphogenesis 2-like, partial \[Sinocyclocheilus grahami\]](#)

1020615139
[STA](#) [Graphics](#)

[repeat-containing protein 30-like, partial \[Sinocyclocheilus grahami\]](#)

1020615134
[STA](#) [Graphics](#)

[subunit pICln-like, partial \[Sinocyclocheilus grahami\]](#)

1020615128
[STA](#) [Graphics](#)

Species Summary ▾ 20 per page ▾ Sort by Default order ▾

Animals (18,104,022)
Plants (6,528,053)
Fungi (8,339,996)
Protists (3,215,785)
Bacteria (230,345,242)
Archaea (2,522,481)
Viruses (3,453,766)
Mus musculus
Customize ...

Items: 1 to 20 of 283622511

<< First < Prev Page 1

[PREDICTED: semaphorin-3D-like, partial \[Sinocyclocheilus grahami\]](#)

230 aa protein
Accession: XP_016120447.1 GI: 1020615144
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[PREDICTED: disheveled-associated activator of morphogenesis 2-like, partial \[Sinocyclocheilus grahami\]](#)

2. 238 aa protein
Accession: XP_016120446.1 GI: 1020615139
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[PREDICTED: leucine-rich repeat-containing protein 30-like, partial \[Sinocyclocheilus grahami\]](#)

3. 263 aa protein
Accession: XP_016120445.1 GI: 1020615134
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[PREDICTED: methylosome subunit pICln-like, partial \[Sinocyclocheilus grahami\]](#)

4. 129 aa protein
Accession: XP_016120443.1 GI: 1020615128
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[PREDICTED: neuroligin-3-like, partial \[Sinocyclocheilus grahami\]](#)

5. 225 aa protein
Accession: XP_016120442.1 GI: 1020615123
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

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

Animals (312,670)
Plants (0)
Fungi (0)
Protists (0)
Bacteria (0)
Archaea (0)
Viruses (0)
✓ **Mus musculus** (312,671)
[Customize ...](#)

Source databases
PDB (11,170)
RefSeq (78,319)
UniProtKB / Swiss-Prot (16,781)
[Customize ...](#)

Genetic compartments
Mitochondrion (2,863)
Plasmid (24)

Sequence length
[Custom range...](#)

Molecular weight
[Custom range...](#)

Release date
[Custom range...](#)

Revision date
[Custom range...](#)

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

Items: 1 to 20 of 312671 << First < Prev Page 1 of 15634 Next > Last >>

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

[lipase_member O4 precursor \[Mus musculus\]](#)

1. **398 aa protein**
Accession: NP_001310315.1 GI: 1020284562
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[integrin alpha-5 isoform 1 preproprotein \[Mus musculus\]](#)

2. **1053 aa protein**
Accession: NP_034707.5 GI: 1020159030
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Gm8978 protein precursor \[Mus musculus\]](#)

3. **399 aa protein**
Accession: NP_001310180.1 GI: 1019286538
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[lipase_member O1 precursor \[Mus musculus\]](#)

4. **399 aa protein**
Accession: NP_001309994.1 GI: 1018976630
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[RecName: Full=DENN domain-containing protein 1B; AltName: Full=Connecdenn 2](#)

5. **766 aa protein**
Accession: Q3U1T9.3 GI: 1018807853
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[RecName: Full=Selenoprotein N; Short=SelN; Flags: Precursor](#)

6. **557 aa protein**
Accession: D3Z2R5.2 GI: 1018740189
[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.

▼ **Top Organisms** [\[Tree\]](#)

Homo sapiens (8301515)

Mus musculus (4852147)

Zea mays (2018857)

Sus scrofa (1621083)

Bos taurus (1559485)

All other taxa (47479836)

[More...](#)

Molecule type

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

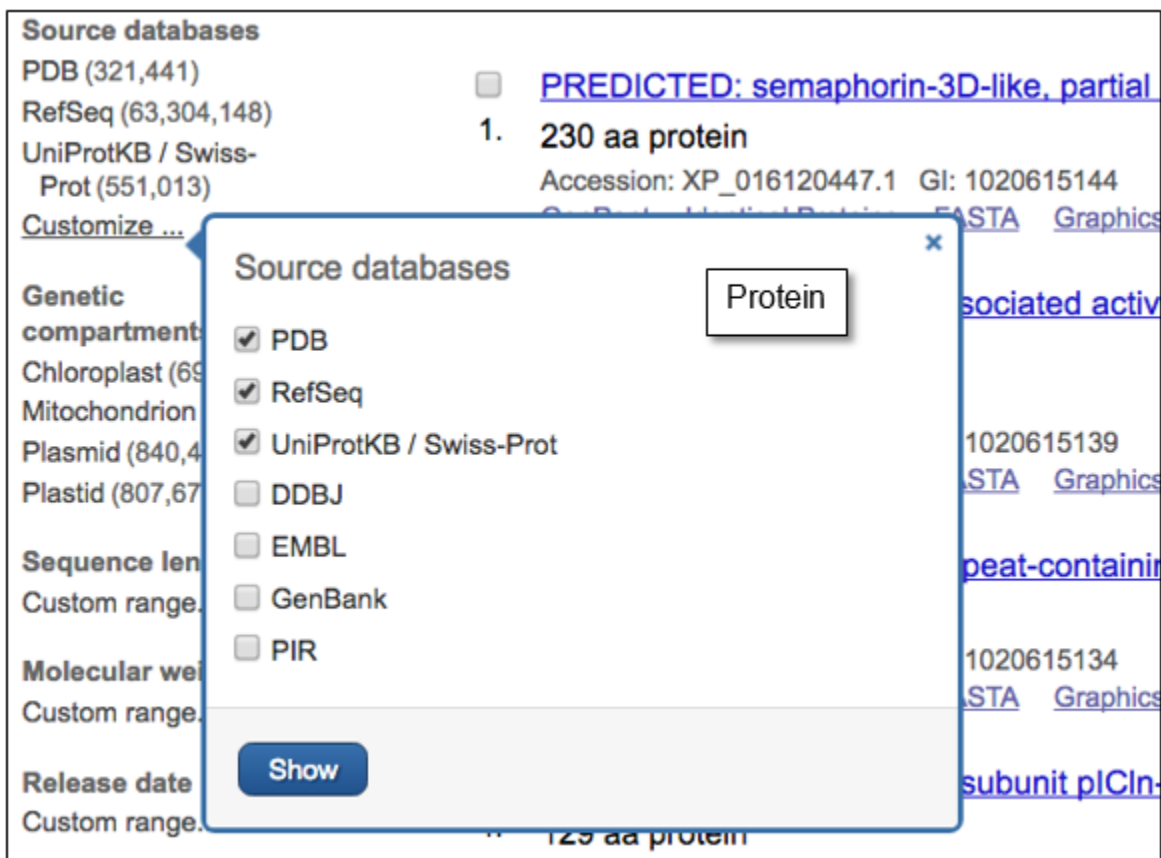
Molecule types

- genomic
 - DNA/RNA (107,209,960)
 - mRNA (38,091,526)
 - rRNA (176,537)
- Customize ...

Source database

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

The screenshot displays the Entrez Sequences interface. At the top right, the accession number KV408990.1 and GI: 1020925611 are shown, along with links for GenBank, FASTA, and Graphics. The main facet area on the left lists various filters: Source databases (INSDC (GenBank) (168,713,572), RefSeq (32,892,401), Customize ...), Genetic compartment (Chloroplast (86), Mitochondrion, Plasmid (150,1), Plastid (946,74)), Sequence length (Custom range.), Release date (Custom range.), and Revision date (Custom range.). A modal dialog box titled 'Source databases' is open, showing a list of source databases with checkboxes: DDBJ, EMBL, GenBank, INSDC (GenBank) (checked), PDB, PIR, RefSeq (checked), and UniProtKB / Swiss-Prot. A 'Show' button is at the bottom of the dialog. A search bar with the text 'Nucleotide' is also visible within the dialog. The background shows a list of search results with accession numbers 925610, 925609, and 925608, and a link for 'Scleropages formosus breed green'.



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

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

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.

The screenshot shows the Entrez Sequences search results page. At the top, there are controls for 'Summary', '20 per page', and 'Sort by Default order'. A 'Format' dropdown menu is open, listing options: Summary (selected), GenPept, GenPept (full), FASTA, FASTA (text), ASN.1, Revision History, Accession List, and GI List. The main content area displays a search result for 'transporter 9 isoform b [Homo sapiens]' with accession number NP_001310848.1 and GI: 1021643673. Below it, another result is partially visible: '[m]C domain-containing protein 8 isoform 3 precursor [Homo sapiens]' with accession number NP_001310848.1 and GI: 1021643671. Navigation links like '<< First', '< Prev', and 'Page 1' are visible.

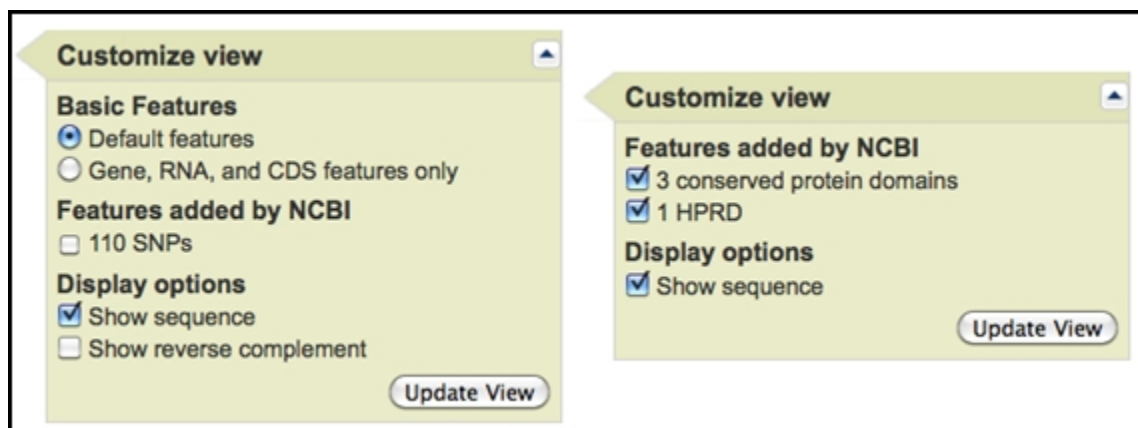
How can I download sequence records to a file on my computer?

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 'Send to' dialog box. At the top, there is a 'Send to:' label with a dropdown arrow. Below this, the 'Choose Destination' section has three radio buttons: 'File' (selected), 'Clipboard', and 'Collections'. Below the destination options, it says 'Download 207 items.' and 'Format'. The 'Format' dropdown menu is set to 'GenBank'. At the bottom, there is a 'Create File' button.

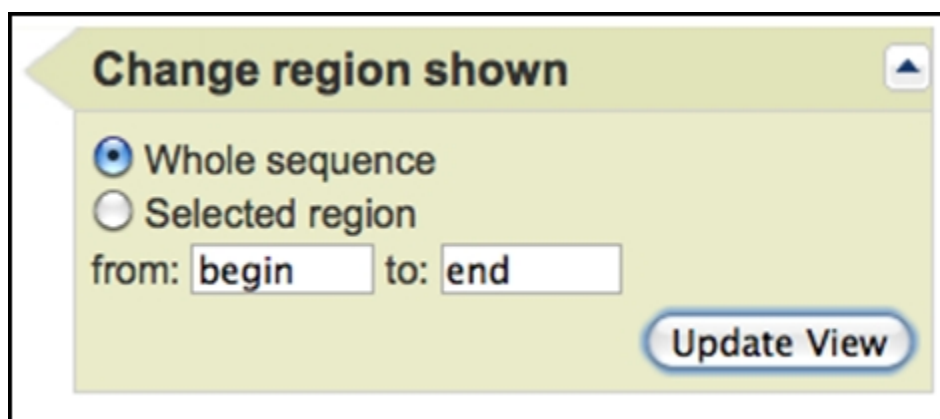
How do I change the information that is shown such as optional biological features or sequence?

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.



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, GSS, and EST), and Conserved Domain Database Search (Protein) in the right-hand column of displayed records.

The image shows a grid of four analysis tool panels. The top-left panel is titled 'Analyze this sequence' and contains links for 'Run BLAST', 'Identify Conserved Domains', and 'Find in this Sequence'. The top-right panel is also titled 'Analyze this sequence' and contains links for 'Run BLAST', 'Pick Primers', and 'Find in this Sequence'. The bottom-left panel is titled 'Analyze these sequences' and contains links for 'Run BLAST', 'Align sequences with COBALT', and 'Find in these sequences'. The bottom-right panel is also titled 'Analyze these sequences' and contains links for 'Run BLAST' and 'Find in these sequences'. Each panel has a small upward-pointing arrow icon in the top right corner.

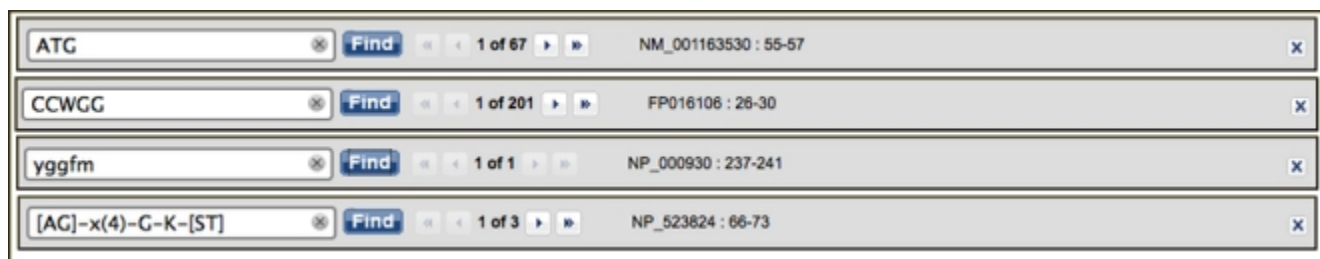
There are also links to related data in the right-hand column that may provide additional information and pre-computed analyses for the displayed records.

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 Discovery column of single and multiple-record displays. This tool can find sub-sequences or patterns in displayed nucleotide or protein sequences.

This image is identical to the one above, showing a grid of four analysis tool panels. The top-left panel is titled 'Analyze this sequence' and contains links for 'Run BLAST', 'Identify Conserved Domains', and 'Find in this Sequence'. The top-right panel is also titled 'Analyze this sequence' and contains links for 'Run BLAST', 'Pick Primers', and 'Find in this Sequence'. The bottom-left panel is titled 'Analyze these sequences' and contains links for 'Run BLAST', 'Align sequences with COBALT', and 'Find in these sequences'. The bottom-right panel is also titled 'Analyze these sequences' and contains links for 'Run BLAST' and 'Find in these sequences'. Each panel has a small upward-pointing arrow icon in the top right corner.

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 [Prosites 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
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
source	1..97660 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="X" /map="Xp11.3"
STS	3835..4158 /standard_name="PMC130047P4" /db_xref="UniSTS:270611"
gene	5001..95660 /gene="MAOA" /note="monoamine oxidase A" /db_xref="GeneID:4128" /db_xref="HGNC:6833" /db_xref="MIM:309850"
mRNA	join(5001..5254,32353..32447,42130..42267,60711..60815, 61544..61635,77012..77153,80080..80229,80533..80692, 81538..81634,85066..85119,89520..89577,90789..90886, 92633..92744,92948..93010,93206..95660) /gene="MAOA" /product="monoamine oxidase A" /transcript_id="NM_000240.2" /db_xref="GI:33469954" /db_xref="GeneID:4128" /db_xref="HGNC:6833" /db_xref="MIM:309850"

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).

The screenshot displays the NCBI Entrez Sequences interface for the MAOA gene (NG_008957). A feature search bar is active, showing a list of features with their coordinates and qualifiers. The 'exon' feature at coordinates 32353..32447 is highlighted with a red arrow. Below the search bar, the corresponding nucleotide sequence is shown, with the residues from 32353 to 32447 highlighted in orange. The sequence is as follows:

```

32221 gaa...tga
32281 tgtgtagacc tggtaaac tggatttta agcatttga tggtaacttg ctctttttg
32341 tttttctttt aqtaactatc gttgccaacg tottgaciga atatggogit arlgttttg
32401 ttttgaagc tgggacagc gttgagggaa gaacataac tatnagggta agtgattta
32461 atacttaccat gtaatgtaat atctactca aactacaagt ggcaccactg gaaatcacag
32521 ggtatagaga ggtcttaaga gttcatgcaq ttaagccgtt tgtctccat agcctaaaq
32581 taotccocat aaataagagt ood
32641 ctgaatttca gtggagtga ttc
32701 attccatac tagagacgta aat
32761 tgtactggaa gtgactatac acc
32821 ottagattoc ctatttatoc aat
  
```

The interface also shows a navigation bar at the bottom with options for 'exon', 'CDS', 'gene', 'mRNA', and 'STS'. The 'exon' feature is selected, and the display shows 'FASTA' format. The search bar at the bottom right contains the text 'Feature' and '2 of 15'.

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).

The screenshot displays a genomic DNA sequence with several lines of text. A red highlight bar is positioned over a segment of the sequence, indicating a feature. The feature bar includes a pull-down menu with 'mRNA' selected. A 'Details' box is open, showing the feature's coordinates and description: 'complement(join(33277..33903,34534..34662,35402..35483,38316..38516,40068..40194,40727..40892,44657..44712,52930..53000,63293..63398)) /gene='PON2''. The sequence coordinates are 39961 to 41041. The feature is labeled as 'mRNA' and is on the 'minus strand'.

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).