

SelenoDB: insert genes

didac.santesmasses@crg.cat

2016

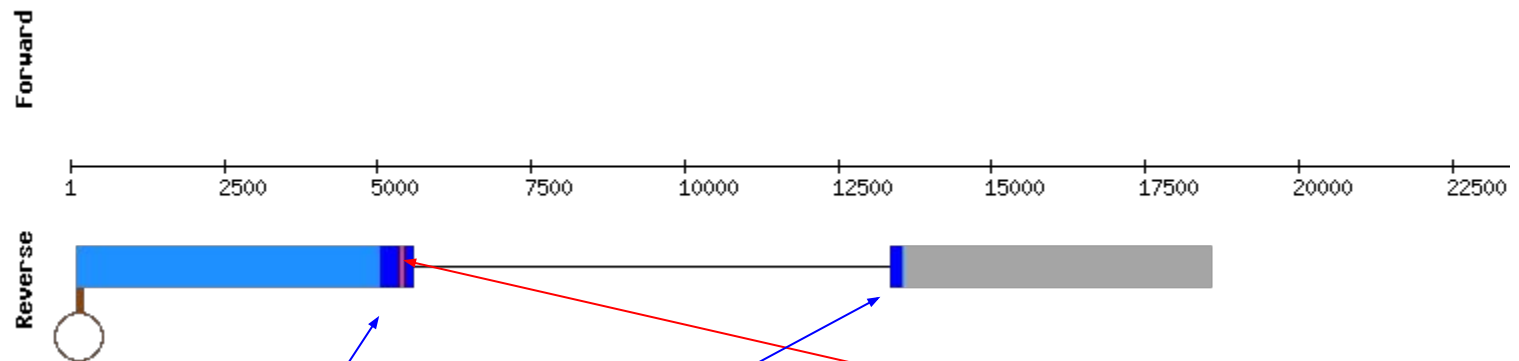
GFF (general feature format)

- Include all your genes predictions in gff format
 - <http://www.ensembl.org/info/website/upload/gff.html>
- Each gene in a separate file (SPS2.gff, SeIP.gff, ...)
- Selenoproteins
 - cds (coding sequence, Sec position)
 - SECIS (if available)
- Machinery genes and Cys homologues
 - cds
- Example:

```
24 program cds 25166504 25167076 . - . profile "DI"; label "selenocysteine"; sec_position "25166902";
24 program cds 25174843 25175064 . - . profile "DI"; label "selenocysteine";
24 program secis 25161608 25161682 . - . SeB.1 Infernal_score=31.09 Covels_score=20.81
```

Feature (cds or secis) Start, end (relative to genome position, subseq(x)) Selenocysteine position (relative to genome position, subseq(x))

GFF (general feature format)



```
24 program cds 25166504 25167076 . - . profile "DI"; label "selenocysteine"; sec_position "25166902";
24 program cds 25174843 25175064 . - . profile "DI"; label "selenocysteine";
24 program secis 25161608 25161682 . - . SeB.1 Infernal_score=31.09 Covels_score=20.81
```

SECIS and Sec position

<http://seblastian.crg.es/>

Selenoprotein prediction server

Welcome to the SECISearch3/Sebastian server. Mouse over the different fields to display information about them in this box.

SECIS prediction
SECISearch3

search also complementary strand
 filter improbable structures
 generate SECIS images (dpi: 150)
 predict SECIS type

SECISearch3 method:

Infernal
score threshold: 10

Covels

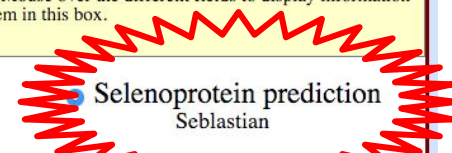
Original SECISearch

Upload your sequence file:
 No file chosen

or paste it here:

```
>SEQ00000033_1.0 # Sequence # Selenophosphate synthetase 2 (SPS2) # Homo sapiens # complete # Reverse
TGCTATGGTGGTTTGGCTGCACCTATCAACCTGTCACTAGGTTTTAAGCCCCGCATGCCT
TAGCTGTTTGTCCCTAATGCTCTCCCTCCCTTTGCCCCCAACCCCAATGGGCCCTGGTGT
CCTATGTTCCCCCTCCCTGTGTCATGTGTTCTCATTTGTTCAACTCCCACTATGAGTGA
GAACATGCAAGTGTGTTGTTTCTGTTCTGTTAGTTGCTGAGAAATGATGGCTTCCA
GCTTCATCCATGTCCTGCAAGGACATGAACATCTTTTATGGCGACATGGTCTCA
CTTCTATTTATTTATTTTGGAGACAGACTCGGCTGTGACCCAGGCTGGAGTGCAG
TGGCATAGTCTTGGCTCAGTCCAGCTCCACCTCCCGGGTTACGCCGTTTCCCTGCCTC
AGCCTCCCAAGTAGCTGGGACTACAGGTGCCTGCCACCACGCCCGGCTAATTTTTTGTAT
```

[About](#) [Contact us](#)



fastasubseq output
sequence



SECIS and Sec position

<http://seblastian.crg.es/>

Output page

[Need help with output?](#)

Total number of SECIS elements predicted: 1

Number of known selenoproteins predicted: 1

You can download here these output files:

- [Summary of predictions](#)
- [FASTA amino acid sequence of selenoproteins](#)
- [FASTA coding sequence of selenoproteins](#)
- [GFF file \(Genomic Feature Format\)](#)
- [FASTA sequence of SECIS elements](#)

or scroll down to inspect each result.

Selenoprotein id: 1 Category: known selenoprotein

Protein prediction

Predicted by: exonerate Blastx evalue: 0.0

Query protein: gil15011844|reflNP_036380.2| selenide, water dikinase 2 [Homo sapiens] >gil172044671|lspQ99611.3|SPS2_HUMAN
RecName: Full=Selenide, water dikinase 2; AltName: Full=Selenium donor protein 2; AltName: Full=Selenophosphate synthase 2
>gil14717790|gblAAC50958.2| selenophosphate synthetase 2 [Homo sapiens] >gil34783113|gblAAH02381.3| Selenophosphate synthetase
2 [Homo sapiens] >gil34785063|gblAAH16643.1| Selenophosphate synthetase 2 [Homo sapiens]

Positions on query: 1-448 Query length: 448

Target name: SEQ00000033_1.0

Positions on target: 3754-5097 Strand: -

```
Query  MAEASATGACGEAMAAAEAGSSGPAGLTLGRSFSNYRPFEPQALGLSPSWRLTGFSGMKGUGCKVPQBALL
Target MAEASATGACGEAMAAAEAGSSGPAGLTLGRSFSNYRPFEPQALGLSPSWRLTGFSGMKGUGCKVPQBALL
agggtgaggtgggagggggtgcggtacgcattatcctgccgtgcacatccagtgaaagttagccggcc
tcaccccgcggactcccaagccgcctctgggtcaagctacactgtcgggtcgtcgtaggggatacaact
ggacgggcccaggaggaccgcgctgcgcgcgcgcggggcccggggcccgggcccgggcccgggcccggg
Query  KLLAGLTRPDVRPPLGRGLVGGQEEASQEAGLPAGAGPSPFFPALGIGMDSCVIPLRHGGLSLVQTTDF
Target KLLAGLTRPDVRPPLGRGLVGGQEEASQEAGLPAGAGPSPFFPALGIGMDSCVIPLRHGGLSLVQTTDF
accggcaccggcccgcggcggggtcgggcccgggcccacatcgcgagagttagaccagggctcgcaagtt
attcgtcgcattgctgggttggaaaccaacgtccgcgcctcctctgtacgttctgaggtcttaccatt
acggaggggcccggcgcgctcgaggccaccggaagcccctacgccggcccggcggcggcggcggcggcggc
```

```
Query  YPLVEDPYMMGRIACANVLSLDYAMGITECDNMLMLLSVSQSMSEEREKVTPLMVKGRDAAEEGGTAV
Target YPLVEDPYMMGRIACANVLSLDYAMGITECDNMLMLLSVSQSMSEEREKVTPLMVKGRDAAEEGGTAV
tctgggctaagcagtgagcagctgagaagtgaatcacagacaaggcgagaccagagtcggggggagg
acttaacattggtcgcattgatactgtcagaattttgtgagtgaaagaatcctttagtaccaaggcct
ccgaatccgggcatccgggtccccgggtgtccgggaccgccgctgtggacagagacgcactgtgtgaaggag
```

http://selenodb.crg.cat/

SelenoDB

Release **2.0**

Home Statistics Download Documentation Help

Chart Search Reports for Keyword Advanced

Welcome

Please, query the SelenoDB 2.0 database for:

- Selenoproteins
- Cysteine-containing homologs or selenium machinery genes

Please, access SelenoDB 1.0 [here](#)

Add genes [here](#) (permission required)

Chart Search Species for Family Advanced

Home Cite License Disclaimer Faq Contact us

[Home](#)

[Statistics](#)

[Download](#)

[Documentation](#)

[Help](#)

Author

email address *

Species

-- select a species -- *

Gene name

-- select a gene family -- Forward Reverse *

Promoter

Promoter start - Promoter end

Exons

Exon start - Exon end + Add *

Protein

Sec / U Protein start - Protein start *

Secis

Secis start - Secis end + Add

Residue

Sec / U Residue start - Residue end + Add

CHR - Sequence start on the c - Sequence end on the ch *

Choose File No file chosen

[gff format example](#) *

Fetch Sequence

Sequence

Submit

[Home](#)

[Cite](#)

[License](#)

[Disclaimer](#)

[Faq](#)

[Contact us](#)

Author	<input type="text" value="email address"/>
Species	<input type="text" value="Test test"/>
Gene name	<input type="text" value="SPS"/> <input checked="" type="radio"/> Forward <input type="radio"/> Reverse
Promoter	<input type="text" value="168885"/> - <input type="text" value="173884"/>
Exons	<input type="text" value="173885"/> - <input type="text" value="175228"/> <input type="button" value="+ Add"/> <input type="button" value="- Remove"/>
Protein	<input type="text" value="N/A"/> <input type="text" value="173885"/> - <input type="text" value="175228"/> <input type="button" value="Sec / Cys homologue label"/>
Secis	<input type="text" value="175790"/> - <input type="text" value="175866"/> <input type="button" value="+ Add"/> <input type="button" value="- Remove"/>
Residue	<input type="text" value="Sec / U"/> <input type="text" value="Residue start"/> - <input type="text" value="Residue end"/> <input type="button" value="+ Add"/>
Sequence	<input type="text" value="KN993176.1"/> - <input type="text" value="173885"/> - <input type="text" value="175228"/> <input type="button" value="Choose File"/> <input type="text" value="SPS.1.selenocysteine.gff"/> gff format example <input type="button" value="Fetch Sequence"/> <pre>GGTAAATGATAGAACAGAGAAAGACCAGGGTATGGGATCACAAG GAGGGACTTCCTCTTGAAGATAATAGTITTCTTGGCAGAGAATA GCCATTCCAAAGGGGAATGGAGAGCAGCccctgttagcccaggctggatg cagtggccattcaCAGGTGCCATCGTAGCAGCTATGGCCTTGAACCT TGGACtccaatgatcctcctgcctcagcctccaactagctgaggctacaggctcAAC CACCATGTGTGGCTCTTGTCTTGTGGCAAGAGCTGATGCACATG GGCTTTCACATTTGCCCCAGATATGAAGCTGAAGCTTCTGTCACT</pre> <p>Translated: MAEVSATGACCEAAMAAMAAEGSSDPAGLTLGRNFSNYRPFEPQALGLSPSWRLTGFSGM KGUGCKVPQETLLKLLAGLTRPNVRPPLGRGLVGGQEEASQETGAGPSPTFPALGIGMDS CVIPLRHGGLSLVQTTDFFYPLVEDPYMMGRIACANVLSDLYAMGITECDNMLLVSQS SMSEERERKVTPLMVKGFDAAEEGCTAVTGGQTVVNPWIIIGGVATVVCQPNFIMPDS AVVGDVVLTKPLGTQVAVNAHQWLDNPERWNVKVMVISREVELAYQEAMFNMATLNRT AAGLMHTFNAAATDITGFGILGHSQNLAQQORNEVSVFIHNLPIIAKTAASIKASGRFG LLQCTSAETSGGLLICLPREQAARFCSEIKSSKYGEGHQAWIVGIVEKGNRFARIIDKPR VTEVLRRCATAAAFAPDNSNASSEPCS@</p> <input type="button" value="Submit"/>