

MUTATION
Surveyor®
DNA Variant Analysis Software

Selecting and Saving
Preferred Transcript

SOFTGENETICS®

Software PowerTools for Genetic Analysis

www.softgenetics.com / info@softgenetics.com

Select "Gene" from the dropdown menu.

Enter the name of gene and click the search button

The screenshot shows the NCBI website interface. At the top left, the NCBI logo and navigation links for 'Resources' and 'How To' are visible. A search bar is located at the top right, with a 'Search' button highlighted by a red box. Below the search bar, a dropdown menu is open, showing 'Gene' selected and 'CXCR4' entered in the search field. The dropdown menu also lists 'Recent' and 'All' categories with various database options. A red box highlights the 'Gene' option in the dropdown. On the right side of the page, there are sections for 'Popular Resources' (including PubMed, Bookshelf, etc.) and 'NCBI Announcements' (including 'Nucleic Acids Research Database 2015 Issue illustrates NCBI databases, updates and future plans'). At the bottom of the page, a navigation bar shows a series of numbered links (1-8) with the number 5 highlighted.

Gene
[Save search](#) [Advanced](#)

[Show additional filters](#)

[Clear all](#)

Gene sources
 Genomic

Categories
 Alternatively spliced
 Annotated genes
 Non-coding
 Protein-coding
 Pseudogene

Sequence content
 CCDS
 Ensembl
 RefSeq
 RefSeqGene

Status
 Current only

Chromosome locations
 Select...

[Clear all](#)

[Show additional filters](#)

Display Settings: Tabular, 20 per page, Sort

Did you mean CXCR4 as a gene symbol?
 Search Gene for [CXCR4](#) as a symbol.

Results: 1 to 20 of 769

i Filters activated: Current only. [Clear all](#) to

Name/Gene ID	Description	Location	Aliases	MIM
<input checked="" type="checkbox"/> CXCR4 ID: 7852	chemokine (C-X-C motif) receptor 4 [<i>Homo sapiens</i> (human)]	Chromosome 2, NC_000002.12 (136114349..136118155, complement)	CD184, D2S201E, FB22, HM89, HSY3RR, LAP-3, LAP3, LCR1, LESTR, NPY3R, NPYR, NPYRL, NPYY3R, WHIM	162643
<input type="checkbox"/> Cxcr4 ID: 12767	chemokine (C-X-C motif) receptor 4 [<i>Mus musculus</i> (house mouse)]	Chromosome 1, NC_000067.6 (128588199..128592313, complement)	CD184, Cmkar4, LESTR, PB-CKR, PBSF/SDF-1, Sdf1r, b2b220Clo	
<input type="checkbox"/> Cxcr4 ID: 60628	chemokine (C-X-C motif) receptor 4 [<i>Rattus norvegicus</i> (Norway rat)]	Chromosome 13, NC_005112.4 (45314952..45318856, complement)		
<input type="checkbox"/> CXCR4 ID: 395324	chemokine (C-X-C motif) receptor 4 [<i>Gallus gallus</i> (chicken)]	Chromosome 7, NC_006094.3 (30263671..30266354, complement)		
<input type="checkbox"/> CXCR4 ID: 707329	chemokine (C-X-C motif) receptor 4 [<i>Macaca mulatta</i> (Rhesus monkey)]	Chromosome 13, NC_007870.1 (116017490..116020734)		
<input type="checkbox"/> CXCR4 ID: 281736	chemokine (C-X-C motif) receptor 4 [<i>Bos taurus</i> (cattle)]	Chromosome 2, AC_000159.1 (61582125..61585838)	BOS_1835	
<input type="checkbox"/> CXCR4 ID: 483900	chemokine (C-X-C motif) receptor 4 [<i>Canis lupus familiaris</i> (dog)]	Chromosome 19, NC_006601.3 (38874650..38877740, complement)		

Locate organism in the description column and click gene name to hyperlink to gene report.

Filters

▼ To
 Hor
 Mu:
 Rat
 Ma
 Hur
 All
 More.

Find
 Datab

Fin

Search

CXCR

Sea

Rece

Q C

Q C

C [H

Select transcript from “Genomic regions, transcripts, and products” section.

NC accession files contain multiple transcripts/isoforms.

Genomic regions, transcripts, and products

Go to reference sequence details

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

Genomic Sequence: **NC_000002.12 chromosome 2 reference GRCh38 Primary Assembly**

- NC_000002.12 chromosome 2 reference GRCh38 Primary Assembly
- NG_011587.1 RefSeqGene
- NC_018913.2 chromosome 2 alternate CHM1_1.1
- AC_000134.1 chromosome 2 alternate HuRef
- NC_000002.11 Chromosome 2 Reference GRCh37.p13 Primary Assembly

NC_000002.12: 136M

136,118,500 136,116,500 136,116 K 136,115,500 136,115 K 136,114,500 136,114 K

Genes, NCBI Homo sapiens Annotation Release 106

NH_003467.2 CXCR4 NH_001008540.1 NP_003458.1 NP_001008540.1

CCDS Features, Release 17 (NCBI Annotation Release 106 compared to Ensembl Release 76)

CCDS46420.1 CCDS33295.1

Genes, Ensembl release 77

ENS600000121966

dbSNP 142 (Homo sapiens Annotation Release 106) all data

ClinVar Short Variations based on dbSNP 142 (Homo sapiens Annotation Release 106)

dbVar ClinVar Large Variations

Cited Variants, dbSNP 142 (Homo sapiens Annotation Release 106)

RNA-seq exon coverage, aggregate (filtered, log2 scaled), NCBI Homo sapiens Annotation Release 106

RNA-seq intron-spanning reads, aggregate (filtered, log2 scaled), NCBI Homo sapiens Annotation Release 106

RNA-seq intron features, aggregate (filtered), NCBI Homo sapiens Annotation Release 106

After transcript is selected, click the GenBank option to the right of the page.

Genomic regions, transcripts, and products

Go to [reference sequence details](#)

Genomic Sequence: NC_000002.11 Chromosome 2 Reference GRCh37.p13 Primary Assembly

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)



1. Navigate to "Send to"
2. Select File
3. Select GenBank format
4. Select Create file
5. Save file to computer

The image shows a screenshot of the NCBI Nucleotide search interface. The search results for "Homo sapiens chromosome 2, GRCh37.p13 Primary Assembly" are displayed. A "Send to" menu is open, showing options for "File", "Clipboard", "Collections", and "Analysis Tool". The "File" option is selected, and the "Format" dropdown is set to "GenBank". The "Create File" button is highlighted with a red box. In the background, a Windows File Explorer window is open, showing the "Downloads" folder with a file named "CXCR4_NC.gb" (18 KB, GB File) created on 1/21/2015 at 3:31 PM.

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search

Advanced Help

Display Settings: GenBank

Send to: Choose destination shown

Choose Destination

- File
- Clipboard
- Collections
- Analysis Tool

Download 1 items.

Format

GenBank

Create File

FASTA Graphics

Go to:

LOCUS NC_000002 3807 bp DNA linear CON 13-AUG-2013

DEFINITION Homo sapiens chromosome 2, GRCh37.p13 Primary Assembly.

Basic Features

- Default features
- Gene, RNA, and CDS features only

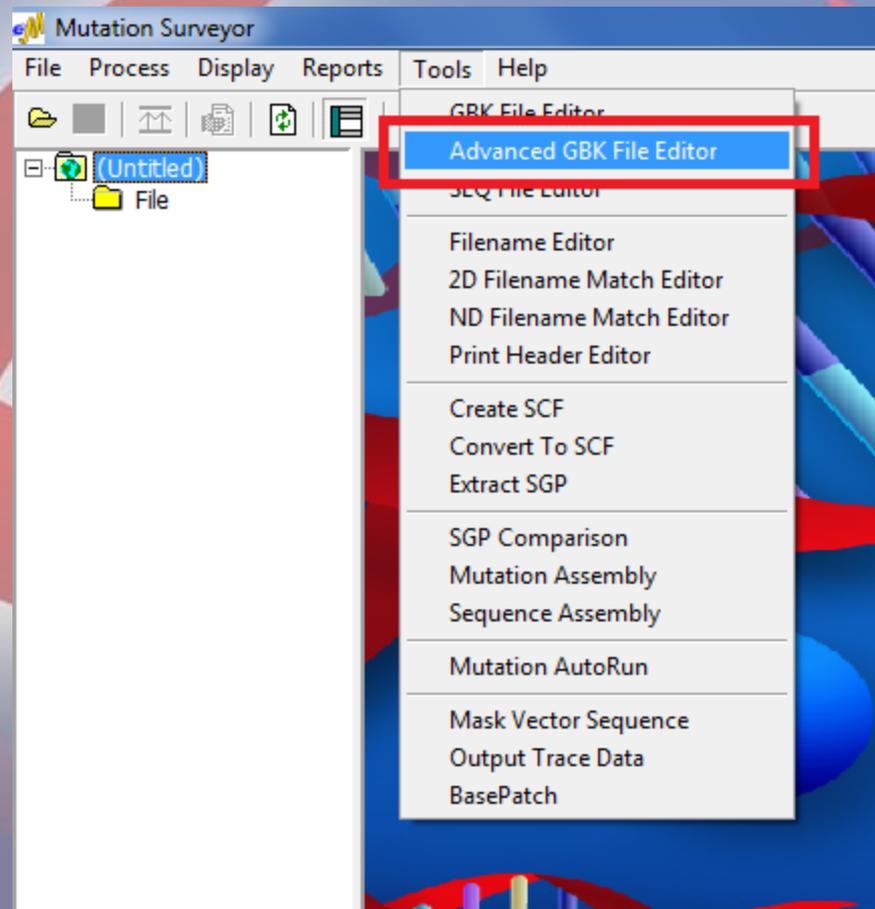
Display options

- Show sequence
- Show reverse complement

Update View

Launch Mutation Surveyor

Navigate to Tools > Advanced GBK File Editor



Open GenBank file

Isoform information will populate in left hand column

Click NP number to select isoform- a check mark will indicate that it is selected

Click the "Sequence" tab to review transcript

The screenshot shows the 'Advanced GBK File Editor' interface. On the left, a tree view displays the file structure for 'CXCR4_NC.gb', including 'Gene', 'CXCR4', 'CDS', 'mRNA', and 'Variations'. The 'NP_001008540.1' entry is selected and highlighted with a blue box and a checkmark. The main window is divided into 'Basic Information' and 'Sequence' tabs. The 'Sequence' tab is active, showing the following sequence data:

Base Count: A: 832 C: 974 G: 1005 T: 996 Others: 0
Sequence Start: 1 Amino Acids Reading Frame: 1 Amino acid Start 1 mRNA Start Index 1

Sequence:

```
2161 TAAATTGCTT TAAAAATTTT TTTAACTGG GTTAAATGCTT GCTGAATTGG AAGTGAATGT
2221 CCATTCCTTT GCCTCTTTTG CAGATATACA CTTCAGATAA CTACACCGAG GAAATGGGCT
2281 CAGGGGACTA TGACTCCATG AAGGAACCCCT GTTCCCGTGA AGAAAATGCT AATTTCAATA
2341 AAATCTTCTT GCCCACCATC TACTCCATCA TCTTCTTAAC TGCCATTGTG GGCAATGGAT
2401 TGGTCATCCT GGTGATGGGT TACCAGAAGA AACTGAGAAG CATGACGGAC AAGTACAGGC
2461 TGCACCTGTC AGTGGCCGAC CTCCTCTTTG TCATCAGCTT TCCTTCTGG GCAGTTGATG
2521 CCGTGGCAAA CTGGTACTTT GGAACCTCC TATGCAAGGC AGTCCATGTC ATCTACACAG
2581 TCAAACCTCTA CAGCAGTGTG CTCATCCTGG CCTTCATCAG TCTGGACCGC TACCTGGCCA
2641 TCGTCCACGC CACCAACAGT CAGAGGCCAA GGAAGCTGTT GCCTGAAAAG GTGGTCTATG
2701 TTGGCGTCTG GATCCCTGCC CTCCTGTGTA CTATTCCCGA CTTCATCTTT GCCAACGTCA
2761 GTGAGGCAGA TGACAGATAT ATCTGTGACC GCTTCTACCC CAATGACTTG TGGGTGGTTG
2821 TGTTCAGTTC TCAGCACATC ATGGTTGGCC TTATCCTGCC TGGTATTGTC ATCCTGTCTT
2881 GCTATTGCAT TATCATCTCC AAGCTGTGAC ACTCCAAGGG CCACCAGAAG CGCAAGGCC
2941 TCAAGACCAC AGTCATCCTC ATCCTGGCTT TCTTCGCTG TTGGCTGCCT TACTACATTG
3001 GGATCAGCAT CGACTCCTTC ATCCTCCTGG AAATCATCAA GCAAGGGTGT GAGTTTGAGA
3061 ACACTGTGCA CAAGTGGATT TCCATCACCG AGGCCCTAGC TTCTTCCAC TGTTGTCTGA
3121 ACCCCATCCT CTATGCTTTT CTTGGAGCCA AATTTAAAAA CTCTGCCAG CACGCATCA
3181 CCTCTGTGAG CAGAGGGTCC AGCCTCAAAG TCCTCTCCAA AGGAAAAGCGA GGTGGACATT
3241 CATCTGTTTC CACTGAGTCT GAGTCTTCAA GTTTTCACTC CAGCTAACAC AGATGTAAAA
3301 GACTTTTTTT TATACGATAA ATAACTTTTT TTTAAGTTAC ACATTTTTTC GATATAAAAG
3361 ACTGACCAAT ATTGTACAGT TTTTATTGCT TGTGGATTG TTGCTTTGTG TTTCTTTAGT
3421 TTTTGTGAAG TTTAATTGAC TTATTTATAT AAATTTTTTT TGTTCATAT TGATGTGTGT
3481 CTAGGCAGGA CCTGTGGCCA AGTTCCTAGT TGCTGTATGT CTGCTGGTAG GACTGTAGAA
3541 AAGGGAAGT AACTTCCAG AGCGTGTAGT GAATCAGTA AAGCTAGAAA TGATCCCCAG
3601 CTGTTTATGC ATAGATAATC TCTCCATTCC CGTGGAAACG TTTTCTGTT CTTAAGACGT
3661 GATTTTGCTG TAGAAGATGG CACTTATAAC CAAAGCCCAA AGTGGTATAG AAATGCTGGT
3721 TTTTCAGTTT TCAGGAGTGG GTTGATTTCA GCACCTACAG TGTACAGTCT TGTATTAAGT
3781 TGTTAATAAA AGTACATGTT AAACCTTA
```

Legend mRNA: █ CDS: █ SNP: AG CG GT AC AT CT Indel

Select File-Save As

Select the option to save “Selected Gene Region, Selected mRNA Region”

Rename file and select the directory to save the file

The image shows a screenshot of the 'Advanced GBK File Editor' software. The main window displays a DNA sequence with various regions highlighted in green (mRNA) and red (CDS). The 'File' menu is open, and the 'Save As...' option is highlighted with a red box. A 'Save As' dialog box is overlaid on the main window, also with a red box around the 'Selected Gene Region, Selected mRNA Region' radio button. The dialog box shows the following options:

- Mode:
 - Selected Gene Region
 - Selected Gene Region, Selected mRNA Region
 - Seq Files for Selected LDS
- File name: CXCR4 Region:1..3807 Length:3807

The main window also shows a 'Basic Information' tab with the following data:

Base Count: A: 832 C: 974 G: 1005 T: 996 Others: 0
Sequence Start: 1 Amino Acids Reading Frame: 1 Amino acid Start 1 mRNA Start Index 1

Sequence:
2161 TAAATTGCTT TAAAAATTTT TTTTAACTGG GTTAATGCTT GCTGAATTGG AAGTGAATGT
2221 CCATTCCTTT GCCTCTTTTG CAGATATACA CTTCAGATAA CTACACCGAG GAAATGGGCT
2281 CAGGGGACTA TGACTCCATG AAGGAACCCCT GTTCCCGTGA AGAAAAATGCT AATTTCAATA
2341 AAATCTTCTT GCCCACCATC TACTCCATCA TCTTCTTAA TGGCATTGTG GGCAATGGAT
2401 TGGTCATCCT GGTCATGGGT TACCAGAAGA AACTGAGAAG CATGACGGAC AAGTACAGGC
2461 TGCACCTGTC AGTGG
2521 CCGTGGCAAA CTGGT
2581 TCAACCTCTA CAGCA
2641 TCGTCCACGC CACCA
2701 TTGGCGTCTG GATCC
2761 GTGAGGCAGA TGACA
2821 TGTTCAGTT TCAGC
2881 GCTATTGCAT TATCA
2941 TCAAGACCAC AGTCA
3001 GGATCAGCAT CGACT
3061 ACACTGTGCA CAAGT
3121 ACCCCATCCT CTATG
3181 CCTCTGTGAG CAGAG
3241 CATCTGTTT CACTG
3301 GACTTTTTTT TATA
3361 ACTGACCAAT ATTGT
3421 TTTTGTGAAG TTTAA
3481 CTAGGCAGGA CCTGT
3541 AAGGGAACTG AACAT
3601 CTGTTTATGC ATAGA
3661 GATTTTGTCT TAGAA
3721 TTTTCAGTTT TCAGG
3781 TGTTAATAAA AGTAC

Legend mRNA: [green box] CDS: [red box]

Two files will save:

filename.gbk = Full GenBank file for selected isoform

*filename*_mRNA.gbk = mRNA file for selected isoform, introns removed

