

SNP Toolbox

SNP Toolbox User Manual

- [1 About SNP Toolbox](#)
- [2 Installation](#)
- [3 Basic Functions](#)
 - [3.1 SNP Terminology](#)
 - [3.2 SNP Window Components](#)
 - [3.2.1 SNP Navigator](#)
 - [3.2.2 Sequence View](#)
 - [3.2.3 SNP Report](#)
 - [3.3 Main Menu Overview](#)
 - [3.4 Global Toolbar](#)
 - [3.5 Working with Sessions](#)
 - [3.5.1 Creating New Session](#)
 - [3.5.2 Load Session File](#)
 - [3.5.3 Load Recently Used Session](#)
 - [3.5.4 Saving Session](#)
- [4 Manipulating](#)
 - [4.1 SNP Navigator Manipulating](#)
 - [4.1.1 All Variations Filter](#)
 - [4.1.2 Effect Filter](#)
 - [4.1.3 Gene Location Filter](#)
 - [4.1.4 Variation Characteristic Filter](#)
 - [4.2 Sequence View Manipulating](#)
 - [4.2.1 Going To Position](#)
 - [4.2.2 Toggling View](#)
 - [4.2.3 Capturing Screenshot](#)
 - [4.2.4 Zooming Sequence](#)
 - [4.2.5 Creating new Ruler](#)
 - [4.2.6 Showing and Hiding Translations](#)
 - [4.2.7 Selecting Sequence](#)
 - [4.2.8 Copying Sequence](#)
 - [4.3 SNP Report Manipulating](#)

1. **About SNP Toolbox**

A fast and effective tool for analysis of genome variations in human chromosomes. It works on Windows, Mac OS X or Linux and requires only a few clicks to install.

2. **Installation**

3. **Basic Functions**

1. **SNP Terminology**

Database

Database contains sequences, their genes and matrix of damage effect.

Session

A session is created by user. It contains a database of annotated sequences, variations and computed matrix of damage effect.

Sequence View

Sequence View aimed to visualize sequences with their genes.

SNP Navigator

SNP Navigator shows the list of variations.

SNP Report

SNP Report shows the variation detailed information and mapping this variation in gene(s).

SNP Navigator

PublicID	Position	Ref	Obs	Chr#
chr22v2	42524947	C	T	chr22
chr22v1	19951271	G	A	chr22
chr19v2	45411941	T	C	chr19
chr19v1	45412079	C	T	chr19
chr16v4	31104509	C	G	chr16
chr16v3	48258198	C	T	chr16
chr16v2	89986117	C	G	chr16
chr16v1	53820527	T	A	chr16
chr15v3	28365618	A	G	chr15
chr15v2	28344238	A	G	chr15
chr15v1	78894339	G	A	chr15
chr12v1	21331549	T	A	chr12
chr11v3	113270828	G	A	chr11

Total: 38

Sequence View (sequence with genes)

chr16 [dna] 20m 40m 60m 90 354 753

Gene (347) uc002efh

Gene (517) uc002efg uc002efj.1 uc002efi.1

Gene (764) uc002efc.1 uc002efg uc002efj.1 uc002efi.1

variation (4)

48 081 724 48.2m 48258198 48.3m 48350k 48 434 672

48 258 181 48258190 48258.2k 48 258 214

S A T T P E Y T G N
L C H L L A R V W Q C
T C T G C C A C T T A C T G G C C G A G T A C A C T G G C A A T G
48 258 181 48258190 48258.2k 48 258 214
A G A C G G T G A A T G A C C G G C T C A T G T G A C C G T T A C
R G S V P G L V S A I
D A V * Q G S Y V P L A
C W F S A P T C O C H

SNP chr16v3 SNP Report

General information

Chromosome: chr16
Position: 48258198
Variation: C->T
Overlapped Genes Overview:

2. SNP Window Components

When you load a session the main window SNP opens, which consists of the SNP Navigator, Sequence View and SNP Report.

1. SNP Navigator

SNP Navigator shows the list of all variations. For each variation the following information is available:

PublicID – variation ID in the program.

Position – variation position in the chromosome.

Ref – reference nucleotide in the chromosome.

Obs – observed nucleotide in the chromosome.

Chr# – a name of a chromosome.

Filter All variations Filter Settings

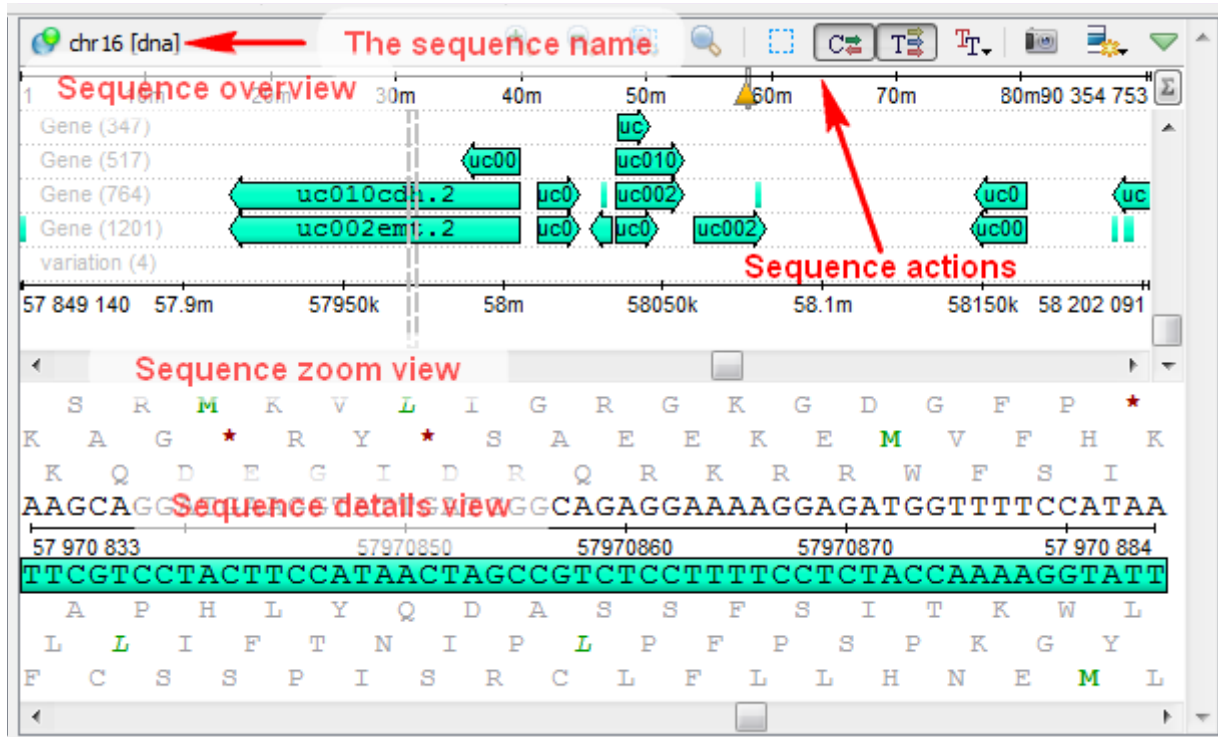
PublicID	Position	Ref	Obs	Chr#
chr22v2	42524947	C	T	chr22
chr22v1	19951271	G	A	chr22
chr19v2	45411941	T	C	chr19
chr19v1	45412079	C	T	chr19
chr16v4	31104509	C	G	chr16
chr16v3	48258198	C	T	chr16
chr16v2	89986117	C	G	chr16
chr16v1	53820527	T	A	chr16
chr15v3	28365618	A	G	chr15
chr15v2	28344238	A	G	chr15
chr15v1	78894339	G	A	chr15
chr12v1	21331549	T	A	chr12
chr11v3	113270828	G	A	chr11
chr11v2	116663707	G	A	chr11

The SNP Navigator information can be sorted by each of its fields (PublicID, Position, Ref, Obs, Chr#). To do this, click on the corresponding field.

2. Sequence View

1. Sequence View Components

The *Sequence View* aimed to visualize and edit sequences along with their genes. After the view is opened you can see a set of buttons. In the picture below these buttons are pointed by the "Sequence actions" arrow. For sequence the small toolbar with actions for the sequence and the following areas are available:



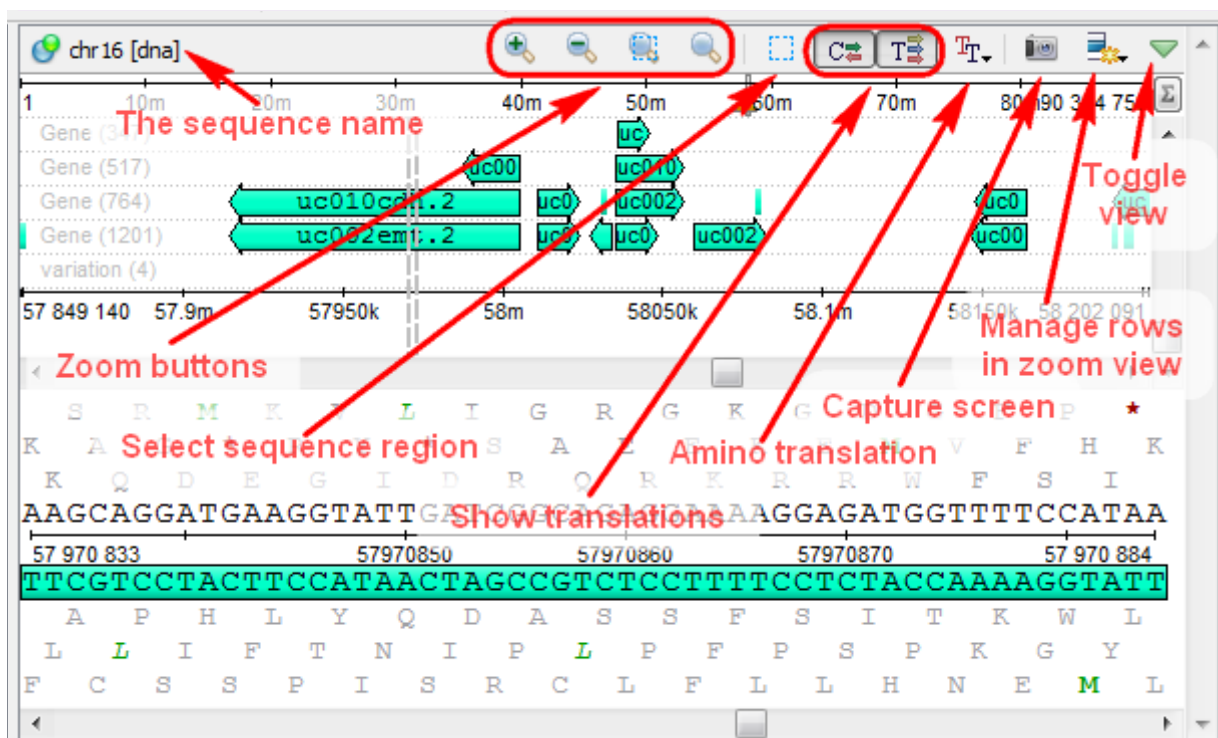
Sequence overview - shows the sequence in whole and provides handy navigation in the *Sequence zoom view* and the *Sequence details view*.

Sequence zoom view - provide flexible tools for navigation in large annotated sequence regions.

Sequence details view - a supplementary component of the *Sequence overview*. It is used to show sequence content without zooming.

2. Sequence Toolbar

A brief description of the sequence toolbar buttons is shown on the picture below:



3. Sequence Overview

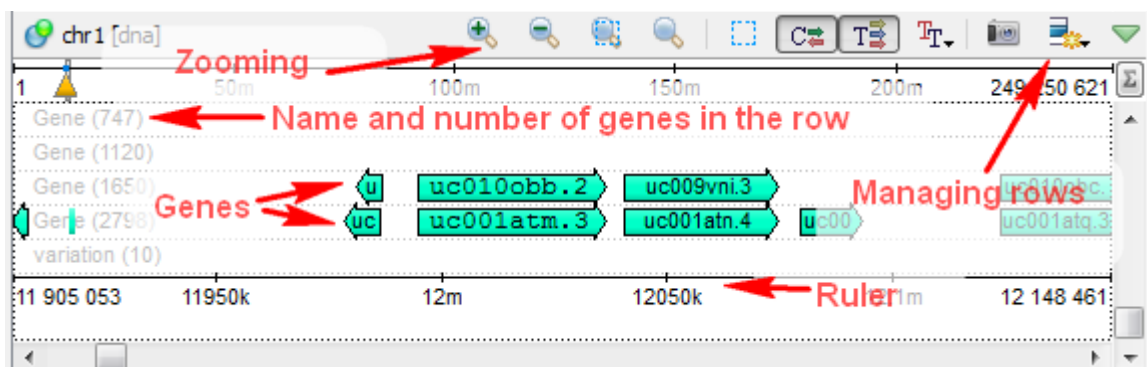
The *Sequence overview* is an area of the *Sequence View* below the sequence toolbar. It shows the sequence in whole and provides handy navigation in the *Sequence zoom view* and the *Sequence details view*.



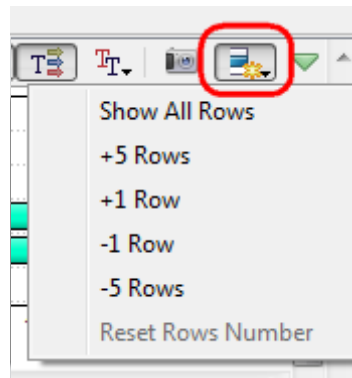
When the sigma button is pressed, density of genes in the sequence is shown.

4. Sequence Zoom View

The *Sequence zoom view* is designed to provide flexible tools for navigation in sequence regions when there are many genes. The most Sequence zoom view space is used to visualize genes for the sequence. The genes are organized in rows by their names. For every row the name and the total number of genes in the row are shown with a light grey text at the left part of the area.



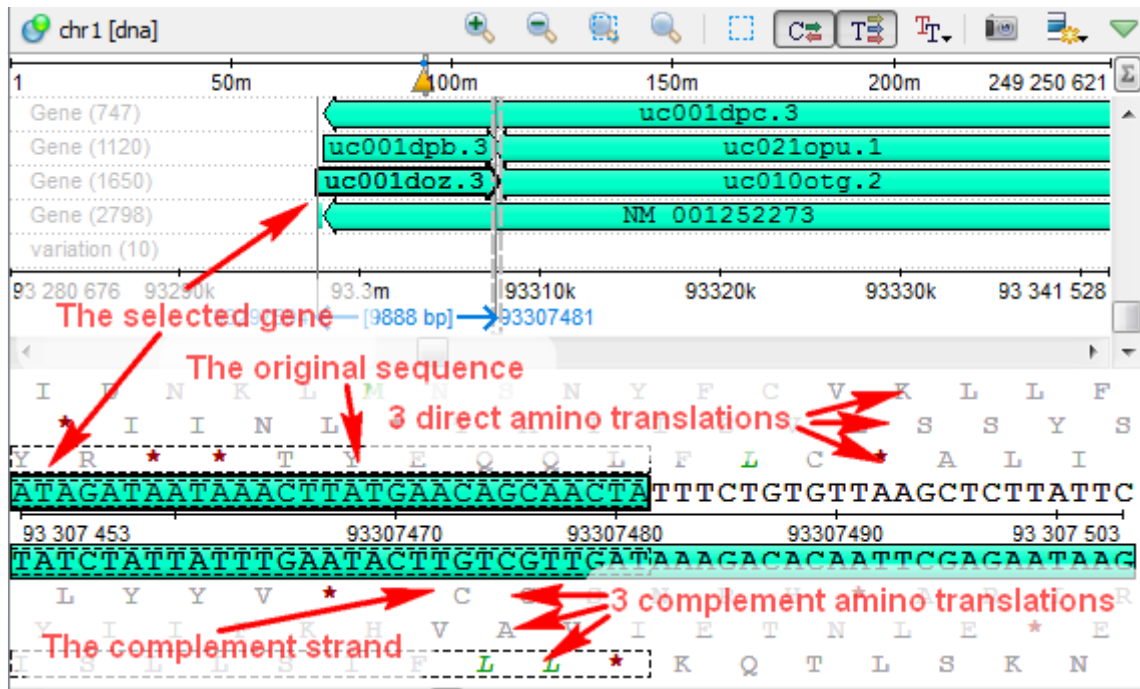
Below the gene rows there is a ruler to show coordinates in the sequence. The *Zoom View* contains not more than 20 rows by default. The rest rows are available by scrolling. To change this behavior use the *Manage Rows in Zoom View* menu button on a sequence toolbar:



When the *Show All Rows* item is checked all available genes are always shown. You can also add rows by selecting the *+5 Rows* and *+1 Row* items and remove rows by selecting the *-5 Rows* and *-1 Row* items. To restore the default number of rows select the *Reset Rows Number* item.

5. Sequence Details View

The *Sequence details view* is a supplementary component of the *Sequence overview*. It is used to show sequence content without zooming. Every time you double click the sequence in the *Sequence overview* area or select a gene, the corresponding sequence position is made visible in the *Sequence details view*. The *Sequence details view* automatically shows complement strand and amino translation frames.



3. SNP Report

The *SNP Report* shows detailed information about variation. In order to open the *SNP Report* double-click on any variation in the *SNP Navigator*. On the picture below you can see a map of variations in the sequence:



SNP chr16v2

General information

Chromosome: chr16
Position: 89986117
Variation: C-->G
Overlapped Genes Overview:

Legend: ■ - exon area ■ - intron area ■ - CDS area

Gene uc002fpe.4

Name: uc002fpe.4	Exons: 89984287..89987385
Accession: Q01726	Description: Homo sapiens melanocortin 1 receptor (alpha melanocyte stimulating hormone receptor) (MC1R), mRNA.
Region: 89984287..89987385	Location: CDS. Exon: 89984287..89987385
Strand: +	Variation: C --> G Tolerance Score: 0 (DAMAGING) Position in protein: 151 Codon: CGC => GGC Translation: R => G
CDS: 89985667,89986620	

Gene uc002fpf.2

Name: uc002fpf.2	Exons: 89985259..89986616,89986998..89987145,89998979..89999087,89999876..89999986,90001137..90002505
Accession: Q13509	Description: Homo sapiens tubulin, beta 3 class III (TUBB3), transcript variant 1, mRNA.
Region: 89985259..90002505	Location: CDS. Exon: 89985259..89986616
Strand: +	Variation: C --> G Tolerance Score: 0 (DAMAGING) Position in protein: 151 Codon: CGC => GGC Translation: R => G
CDS: 89985667,90002212	

SNP Report includes the following information:

General information

The general information includes the information about sequence:

Chromosome – name of a chromosome.

Position – position of a variation.

Variation - the reference and the observed nucleotides.

Overlapped Genes Overview – shows genes overview with exon, intron, CDS areas.

Gene parameter will be displayed for each gene where there is the required variation.

Gene

The *Gene* includes:

Name – the name of a gene.

Accession - gene accession.

Region – the length of gene.

CDS – the length of CDS (coding region of a gene).

Exons - exon regions.

Description – description of a gene.

Location - variation location in the gene. You may change splice site length in the general properties.

Variation Effect – shows the following information:

Tolerance Score: the value of a damage effect (if it is in a database).

Position in protein: position in a protein.

Codon: reference codon in a chromosome and observed codon.

Translation: reference translation and observed translation.

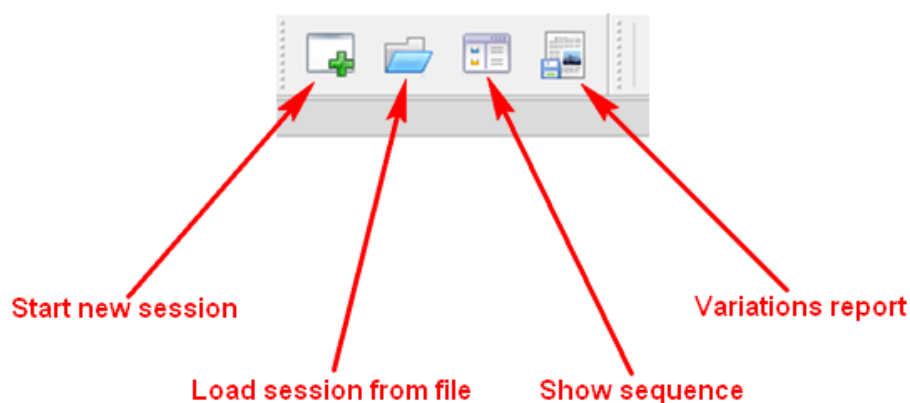
There are not *Variation Effect* components if database hasn't information about damage effect or if a gene is non-protein coding.

3. Main Menu Overview

Пункт меню	Описание
File	A set of session level operations. Example: create, load, close a session.
Settings	Preferences
Window	A list of active windows and basic manipulations with the windows. Example: close active view, close all windows, tile windows, cascade windows, next window, previous window.
Help	Information about program, user manual, quick start guide.

4. Global Toolbar

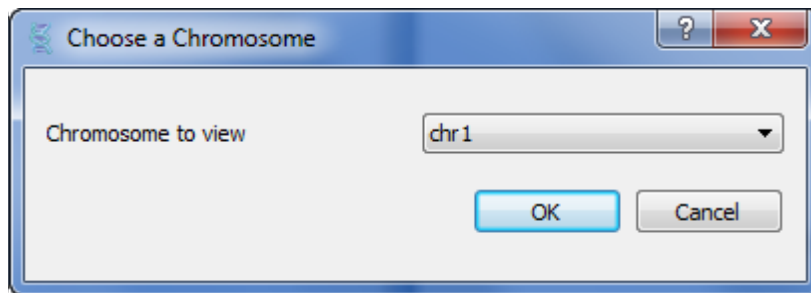
The following options are available from the global toolbar:



Start new session - create a new session.

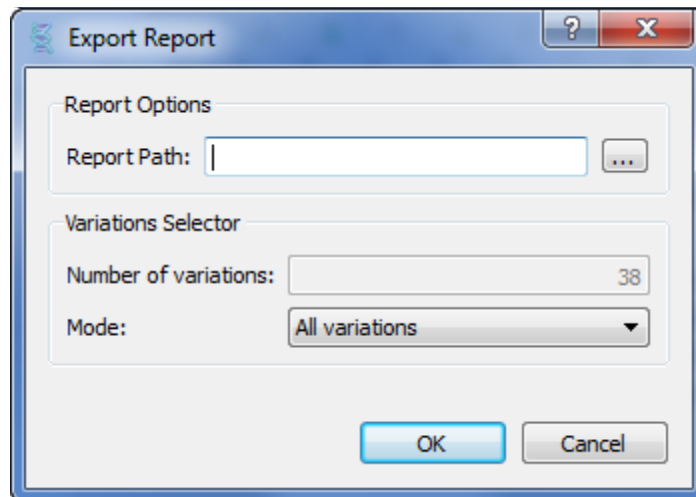
Load session from file - upload created session from a file.

Show sequence - shows the required sequences.



Choose a chromosome and click "OK" button.

Variations report - create and export the required report.



The following parameters are available:

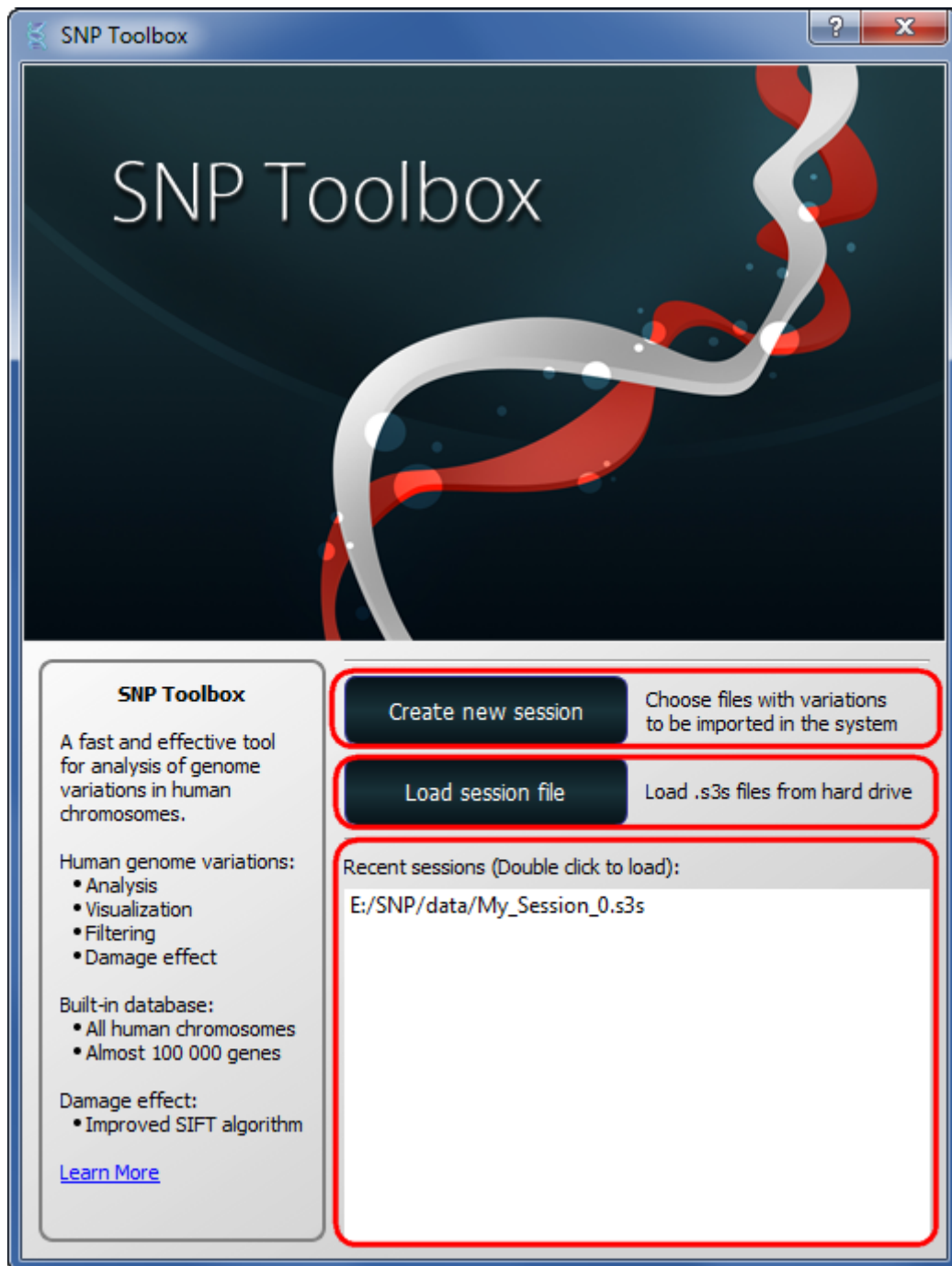
Report Path - path for saving a report.

Number of variations - the number of variations which will be included into report. This option is available only for *First variations* mode.

Mode - parameter of number of variations which will be included into report: *All variations* or number of *First variations* or *Selected variations*.

5. Working with Sessions

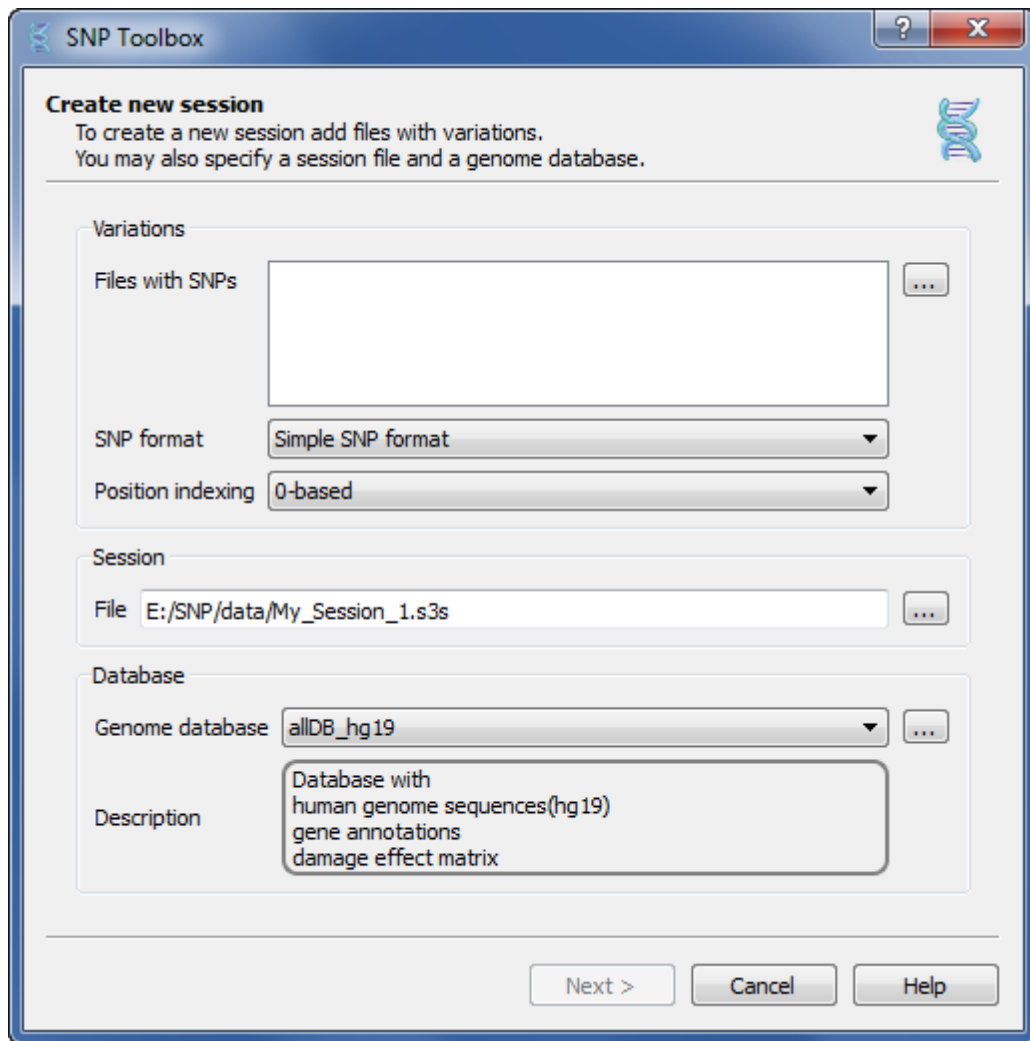
When you run the program the following dialog appears:



You can create a new session, load a session from a file, or select and load recent session.

1. Creating New Session

To create a new session, click on the *Create new session* button. The following window will appear:



The following parameters are available:

Files with SNP - files with variations.

SNP format - variations format.

There are two variations formats:

- Simple SNP format
- VCF4 (<http://www.1000genomes.org/node/101>)

Position indexing - start position indexing of variations in files (0-based or 1-based).

File - file for saving a session.

Genome database - file with a database.

Description - description of the database.

2. Load Session File

To load a session from a file click on the *Load session file* button, choose a file with session and click *Open* button. If a session is already open, firstly click the *Start New Session* button on the global toolbar and after that *Load session from file* button.

3. Load Recently Used Session

To load recently used session double click on the required session from the list of recently used sessions. If a session is already open, firstly click the *Start New Session* button on the global toolbar and after that choose a recently used session.

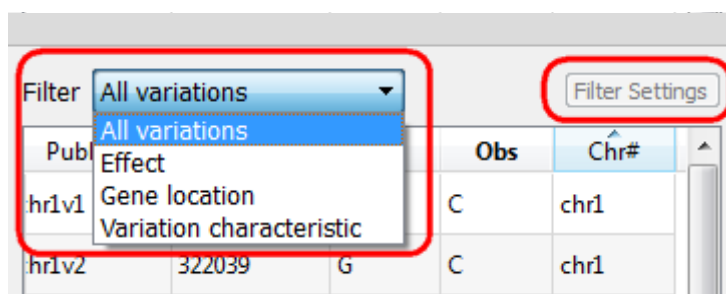
4. Saving Session

All of new sessions saving automatically and appear in the *Recently sessions* list as a *.s3s file.

4. Manipulating

1. SNP Navigator Manipulating

The data displayed in SNP Report can be filtered in various ways. To do this choose the required filter in the *SNP Navigator*.



The parameters of filters can be configured with the option *Filter Settings*. If you change the settings, the filter is applied it to all loaded variations, but not to the previous result of filtration.

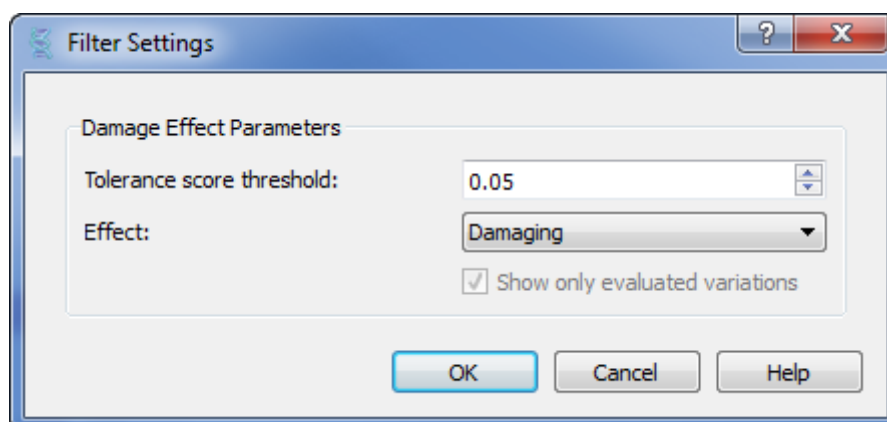
1. All Variations Filter

All variations filter shows all available variations. Because the *All variations* filter shows all variations, the *Filter settings* option is not active for this filter.

2. Effect Filter

Effect filter shows variations with selected value of the damage effect. To estimate the damage effect of variations using the SIFT algorithm [Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1, 31 (13) :3812-4.], which estimate the damage effect of using a multiple alignment of protein where is found the SNP (this variation) among the available homologues.

The following settings parameters are available:



Tolerance Score Threshold – score threshold of the damage effect. It can be from 0 to 1. A variation is considered to be damaging, if at least one gene (in which there is this variation) the value of the damage effect for this gene is less than the specified value of *Tolerance Score Threshold*.

Effect - effect mode. It can be *Damaging* mode which shows the damaged proteins or *Tolerated* which shows not damaged proteins.

Show only evaluated variations – shows only evaluated variations, i. e. if database has information about the damage effect.

3. Gene Location Filter

Gene location filter shows variations which there are in the required region. Using this filter you can filter variations by *Gene Related Location* or by *Gene Name*:

The screenshot shows a 'Filter Settings' dialog box. It is divided into two main sections: 'Gene Related Location' and 'Gene Name'. The 'Gene Related Location' section is currently selected and contains several dropdown menus and input fields. The 'Gene Name' section contains three empty text input fields. At the bottom of the dialog are three buttons: 'OK', 'Cancel', and 'Help'.

Gene Related Location - gene location of variation.

General Location - general gene location of variation:

All – variations in a chromosome.

In Gene – variations in a gene.

Out of Gene – variations out of a gene.

Gene Promoter - variation in a promoter region of a gene.

Gene Coding Role - coding type of gene. This parameter available only for *In Gene* of *General Location* parameter.

All – all genes.

Protein Coding – genes are coding a protein.

Non-Protein Coding – genes are not coding of protein.

Location In Gene - region of a gene with variation. This parameter available only for *In Gene* of *General Location* parameter.

Exon – variations in an exon.

Intron – variations in an intron.

Splice site - variations in a splice site.

Whole Gene – variations in a whole gene.

Location In CDS - CDS location of variation. This parameter available only for *In Gene* of *General Location* parameter.

In CDS – variations in a CDS.

Out of CDS – variations out of a CDS.

Out of CDS. 3'-end - variations out of CDS (for 3' end).

Out of CDS. 5'-end - variations out of CDS (for 5' end).

Promoter length - length of a promoter region

Splice site interval - splice site interval from the start(end) of exons.

Gene Name - name of a gene:

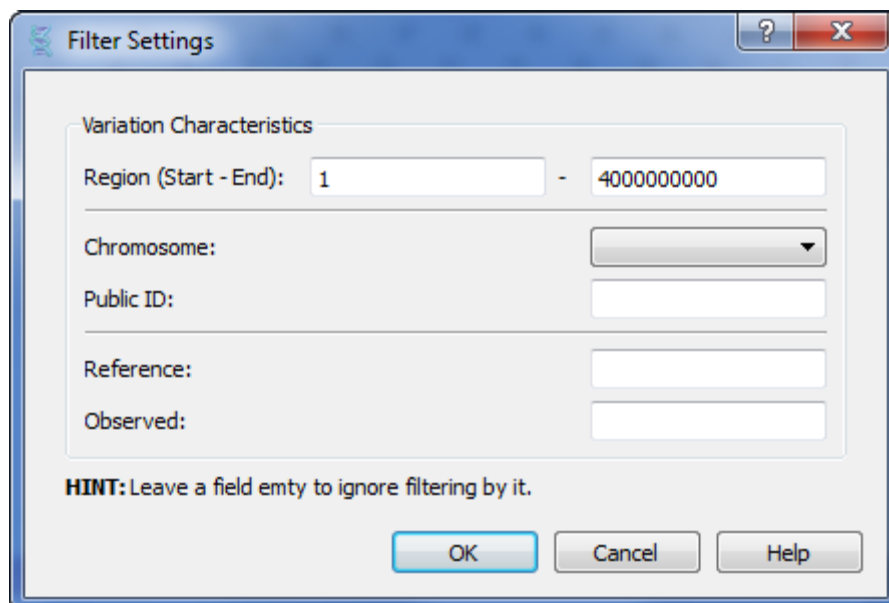
Protein Accession – name of a protein.

Gene Description – description of a gene.

Gene Name - name of a gene.

4. **Variation Characteristic Filter**

Variation characteristic filter shows variations with the required parameters only.



The following parameters are available:

Region – shows variations from the required region.

Chromosome – shows variations from the required chromosome.

Public ID – shows variations with the required ID.

Reference – shows variations with a required reference nucleotide.

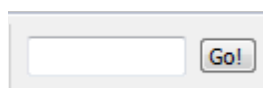
Observed – shows variations with a required observed nucleotide.

You can input one or more filter parameters.

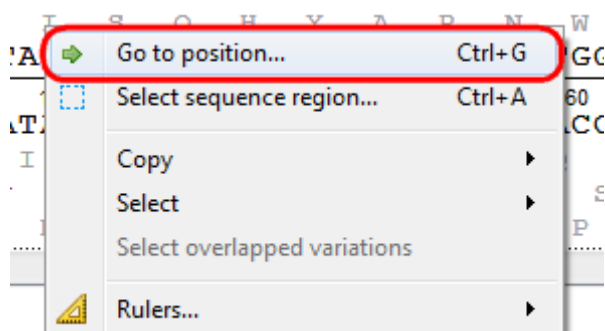
2. Sequence View Manipulating

1. Going To Position

To go to a position, use the global actions toolbar:



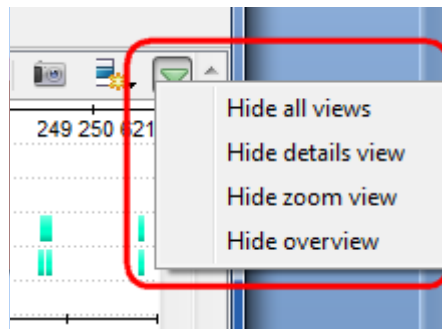
Or use the *Go to position* context menu item.



Also you can use the shortcut *Ctrl-G*.

2. Toggling View

It is possible to switch the *Sequence overview*, *Sequence zoom view* and the *Sequence details view* visibility using the rightmost button in the toolbar:



The sequence can be removed from the view using the same menu.

3. Capturing Screenshot

Use a sequence toolbar *Capture screen* button to save a screenshot of the sequence:



Available file formats are *.jpg, *.png and *.tiff.

4. Zooming Sequence

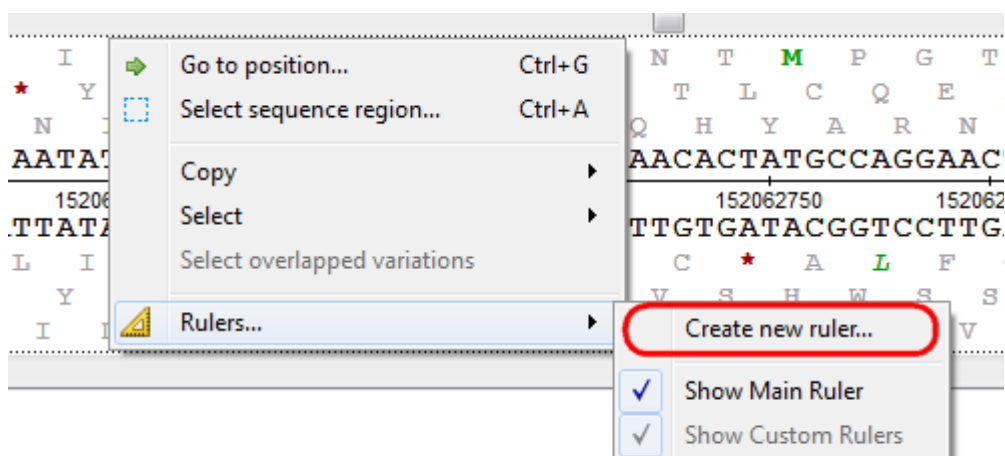
To zoom a sequence in the *Sequence zoom view* you can use one of the zoom button on the sequence toolbar:



There are standard *Zoom In* and *Zoom Out* buttons. Additionally you can zoom to a selected region using the *Zoom to Selection* button. To restore the default view of the *Sequence zoom view* (when the sequence is not zoomed) use the *Zoom to Whole Sequence* button.

5. Creating new Ruler

You can create any number of additional rulers by clicking the *Ruler-->Create new ruler* context menu item:



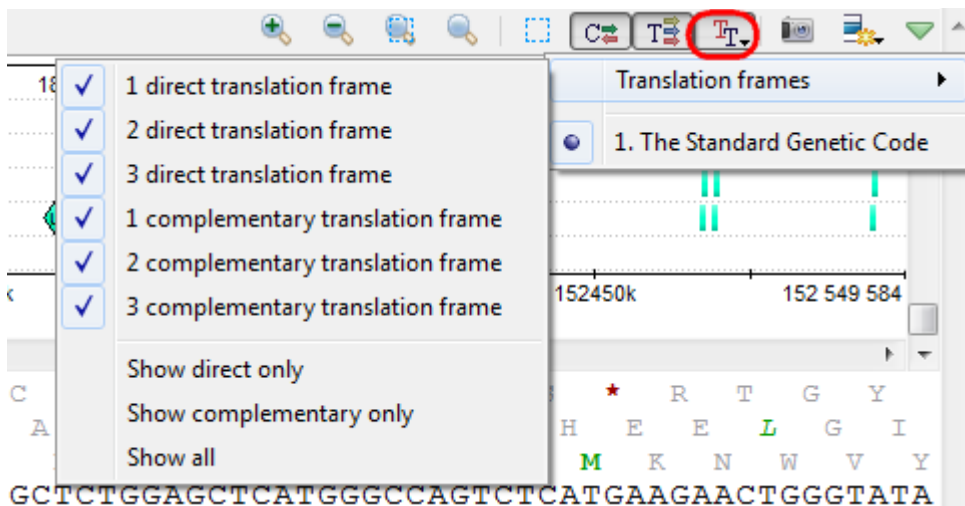
The new ruler will be shown right above the default one.

6. Showing and Hiding Translations

You can turn on / off the direct and complement amino translations visualization in the *Sequence details view* using the Show complement strand and the Show amino translations toolbar buttons.

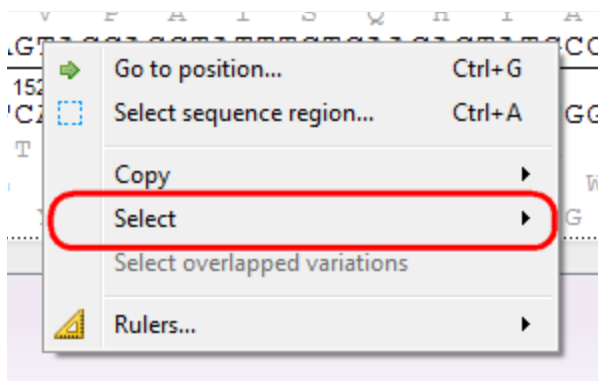


Also to show or hide translation frames you can use the *Amino translation* button:

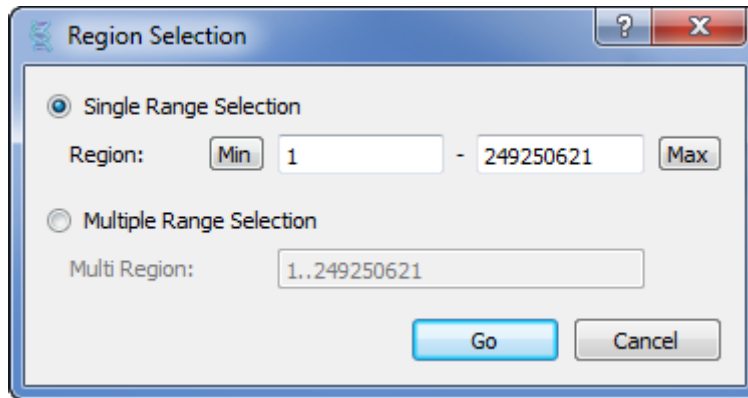


7. Selecting Sequence

You can use different items from the *Select* submenu of the context menu to select a sequence.



Selecting the *Sequence region* context menu item opens the *Select range dialog*:

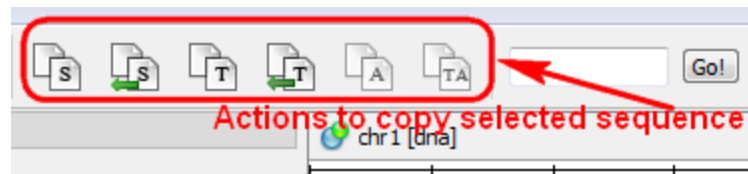


Here you can specify the *Single range selection* or *Multiple range selection*. You can open the same dialog using the *Select sequence region* button on a sequence toolbar or using the *Ctrl-A* key sequence.

8. Copying Sequence

The selected sequence region or amino translations can be copied to clipboard:

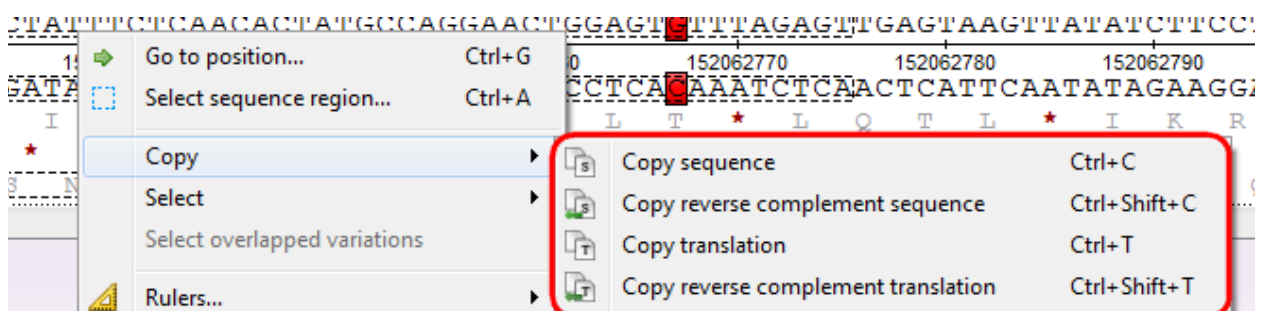
1. By pressing the corresponding buttons in the global toolbar.



2. Using the following shortcuts:

- *Ctrl-C* - copies direct sequence strand
- *Ctrl-T* - copies direct amino translation
- *Ctrl-Shift-C* - copies reverse-complement sequence
- *Ctrl-Shift-T* - copies reverse-complement amino translation

3. Using the *Copy* submenu of the context menu:



3. SNP Report Manipulating

To open the SNP Report double-click on any variation in the SNP Navigator. Also you can *Collapse/Uncollapse all* fields of the *SNP Report* and *Print report*. Also you can see the *User Manual*. For this click on the correspondingly buttons:

