User Guide to



DNASTAR, Inc. 2024

Table of Contents

Welcome to GenVision Pro	6
GenVision Pro Tutorials	7
Try It! – Analyze an E. coli assembly	8
Part A: Start a new session and rename a sequence using data fields	9
Part B: Work with feature tracks	12
Part C: Investigate sequence coverage	14
Part D: Work with region tracks	16
GenVision Pro Interface	20
Welcome screen	22
Views	24
Analysis view	25
Change the Analysis view color scheme	28
Wrap the Analysis view display	32
Genome view	33
Overview	34
Reports view	36
Assembly Report	37
Per-Template Report	40
Enrichment Report	44
Structural Variation view	47
Choosing and understanding Structural Variation view columns	49
Filtering in the Structural Variation view	53
Variants view	54
Choosing and understanding Variants view columns	57
Filtering in the Variants view	64
Text view	67
Table view	69
Pairwise view	71
Console view	73
Panels	75
Style panel	76
Overview section	78
Sequence section	79
Analysis section	80
Variants section	82
Tracks panel	83
Layout section	85
Options section	86
Features panel	87

Places panel	88
Favorites section	90
Recent section	92
Received section	93
Explorer panel	94
Jobs panel	
Details panel	
When no session, or a blank session, is ope	n
When a non-empty session is open	
When one sequence is selected	101
When a portion of a sequence is selected	
When multiple sequences are selected	
When one track is selected	
When multiple tracks are selected	
When one feature is selected	
When multiple features are selected	
Button tools	111
Navigation tools	113
Create or Open a Session	116
Create a new session	117
Open an existing session	118
Change the default application for opening files	119
Create a new assembly using SeqMan NGen	
Launch a session from within ArrayStar	121
Make a session active	
Add Data	
Add assemblies	
Add sequences	
•	
Add tracks	
Search for Sequences Online	
•	
-	
•	
	rez)
·	
	153

Save and reopen saved search results	154
Specify the Experiment or Chromosome to Display	155
Work with Tracks	156
Feature tracks	158
Reference track	160
Coverage tracks	161
Ruler tracks	
Sashimi tracks	165
Alignment tracks	168
Variants tracks	171
Region tracks	172
Customize the Look and Layout	173
Customize the appearance and layout of views	174
Customize the layout of other window components	177
Save a custom layout and apply it to other projects	179
Use Basic Editing Commands	180
Search Within the Project	181
Make a Selection	183
Rename Assemblies and Tracks	185
Create names automatically using specified data fields	186
Rename a single sequence manually	189
Copy	190
Export	191
Export an image of a view	192
Edit an exported image in PowerPoint	194
Export data to a file	195
Save, Close and Exit	197
Appendix	199
Supported file types	200
Set preferences	201
Set Internet preferences	202
Set temporary file preferences	203
IUPAC codes	204
Installed Lasergene file locations	206
Turn off usage logging	208
Troubleshoot the error "A different version is running"	209

Clean cached files on startup	210
Research references	211

Welcome to GenVision Pro

GenVision Pro supports visualization and analysis of multiple samples from a single genome. In RNA-seq, ChIP-seq and copy number variation (CNV) studies, where coverage and peak plots are compared and used to interpret gene expression and regulation, the ability to display and compare genome-wide data sets for multiple samples makes GenVision Pro an essential tool

With GenVision Pro, you can visually navigate between single-segment chromosomes and/or contigs of a genome, then export publication-quality images.

The following video shows how easy it is to use GenVision Pro:

These support pages contain written help and tutorials, as well as video tutorials.

- For help **INSTALLING** Lasergene, see our separate <u>Installation Guide</u>.
- For written TUTORIALS with free data, see GenVision Pro Tutorials.
- <u>Click here</u> for a list of all topics that include **VIDEOS**.
- To **PRINT** the current page of the User Guide, click the printer icon in the top right corner ().
- To download a **PDF** of the entire User Guide, scroll to the bottom of the table of contents on the left, and press **Download as PDF**.

GenVision Pro Tutorials

GenVision Pro tutorials are listed below. Any data needed can be downloaded for free using the link provided in the tutorial.

- Analyze an E. coli assembly Links to the first of four parts.
- Analyze Copy Number Variation (CNV) results This tutorial is located in the SeqMan NGen User
 Guide. The workflow begins with assembly in SeqMan NGen and uses ArrayStar and GenVision Pro
 for downstream analysis.

Try It! - Analyze an E. coli assembly

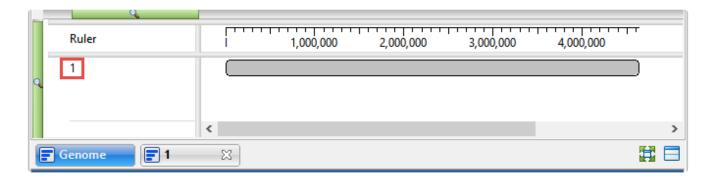
The following tutorial uses free data from the DNASTAR website to demonstrate a typical workflow in GenVision Pro. The tutorial is in four parts, which should be done in order.

Begin with Part A: Start a new session and rename a sequence using data fields.

Part A: Start a new session and rename a sequence using data fields

In this first part of the GenVision Pro tutorial, you will start by downloading and extracting the tutorial data from the DNASTAR website. Next, you will begin a new session and add an *E. coli* assembly produced by DNASTAR's <u>SeqMan NGen</u>. You will then create a custom label for the sequence based on the accession number and sequence description.

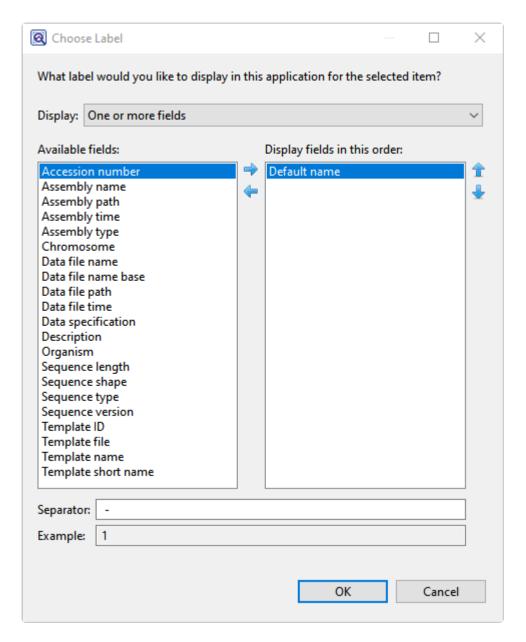
- 1. Download the tutorial data extract the folder to any suitable location (e.g., your desktop).
- 2. Launch GenVision Pro, which will open at the Welcome screen.
- 3. Drag the newly-extracted project *MG1655_genome.assembly* and drop it on the GenVision Prowindow.
- 4. Click on the Genome tab.



Note that the name of the active sequence is '1.' In the next steps, you will be changing this name to something more useful.

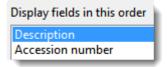
5. Right-click on sequence '1' and choose **Rename with Fields**. This opens the Choose Label dialog.

Note that the **Display** is set to **One or more fields**. This setting lets you automate the naming process for any number of selected sequences using fields such as sequence length, organism, or description.



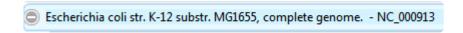
The left pane shows available fields, and the right pane shows fields that will be displayed as part of the sequence name. Initially, **Default name** is the only field listed in the right pane.

- 6. Select and organize the fields to display as part of the sequence name:
 - a. Click on the left arrow () to remove **Default name** from the right-hand pane.
 - b. Use Ctrl/Cmd+click to select Accession number and Description from the list on the left. Click the right arrow () to move them to the right-hand pane.
 - c. Click on **Description** on the right to select it. Then use the up or down arrows on the right (**) to move **Description** to the top of the list.



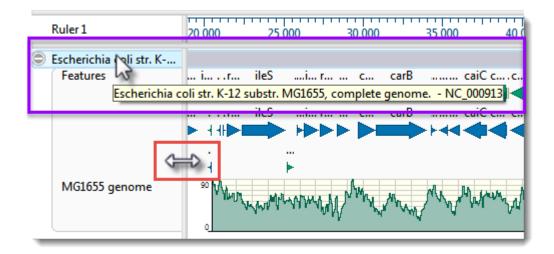
d. Click **OK** to save changes and exit from the dialog.

The sequence name in both the Overview and Analysis view has changed from '1' to:



However, you may only be able to see the first part of the name.

7. To see a tooltip displaying the full name, hover your cursor over the visible portion of the name (see area boxed in purple, below).



8. To view the entire name, drag the pane divider (boxed in red, above) to the right. Then return the pane divider to its original position for ease of viewing the sequence section in the next part of the tutorial.



Note: Even if the full name is not visible when viewed in the GenVision Pro window, it will be shown in the Details panel, the Navigation bar's chromosome selector (the dropdown menu) and in printouts of the views.

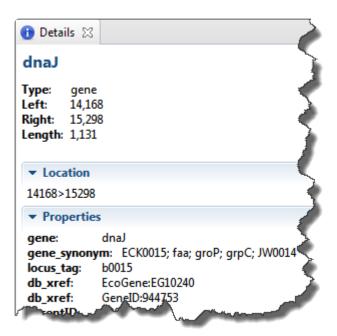
9. To edit the name field manually, right-click on the name and select **Rename**. This opens the same dialog as before, but Display is now set to Custom label. The current name appears in the Label field. In that field, change "Escherichia" to "E." and remove the comma and the words "complete genome." Click OK.

Proceed to Part B: Work with feature tracks.

Part B: Work with feature tracks

Features are displayed as colored arrows in the Overview and the Analysis view. In this part of the tutorial, you will learn to toggle between displaying/hiding features, and to view features that are not currently visible.

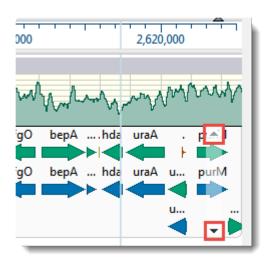
1. To see an information balloon for any feature, hover over the feature. Try clicking on the feature. This displays the same information in the Details panel on the right, where it can be copied and pasted into a text document.





Note:*If you can't see the Details panel, click *View > Reset Panel Layout.

- 2. Right-click on the word Features in either view and select Show Track Options. This command opens both the Tracks panel and the Options section associated with the selected track type; in this case, Features.
- 3. Note that there are two "Feature" check boxes in the Tracks panel; one under **Overview details** and one under Sequence details. Experiment with unchecking and rechecking these two Features boxes to toggle feature display in the views. Check both boxes before moving to the next step.
- 4. The GenVision Views display several rows of features, by default. If you wish to see additional rows of features, use the up/down triangles near the right end of the features display. These arrows are connected by a semi-transparent blue bar.



An alternative is to hover the mouse over the bottom of the track, then grab the resize bar and drag it downward.



Proceed to Part C: Investigate sequence coverage.

Part C: Investigate sequence coverage

GenVision Pro's navigation tools consist of a "sequence selector," a text box, and a **Go** button. In this part of the tutorial, you will use the navigation tools to go to a particular region of low coverage that begins around 4,314,000.

1. In the <u>navigation tools</u> area near the top of the window, hover your cursor over the sequence selector (drop-down menu) and the navigation field text box, in turn. Note that each provides tool tips.

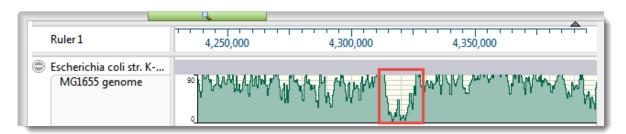


2. Enter **4314000** into the text box. (Alternatively, depending on your locale and computer settings, you could enter **4,314,000** or **4.314.000**.)

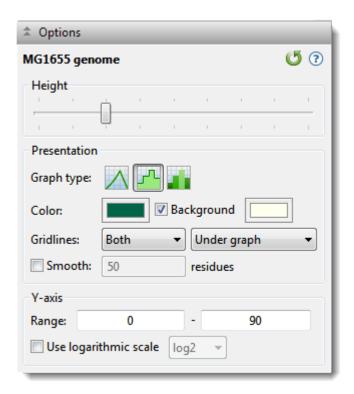


Press Go or Return/Enter.

3. In the Analysis View, slide the horizontal zoom slider until the Coverage track (entitled "MG1655 genome") roughly matches the image below. Note the obvious area of low coverage (boxed in red, below).



4. To optimize the Coverage track options, right-click on the name of the track and choose **Show Track Options**.



- a. Next to **Graph type**, click the bar graph icon (). Note that low-coverage areas like the area found above now appear in dark green, making them easier to recognize.
- b. Experiment with other available options. For example, try changing the **Height** slider and/or the **Y-axis Range**, select a different **Graph type** or **Color**, and experiment with the **Gridlines** dropdown menus. Check the **Use logarithmic scale** button and observe how the Coverage track changes.
- c. When you are finished, revert to the default values using the **Restore default settings** tool ().

Proceed to Part D: Work with region tracks.

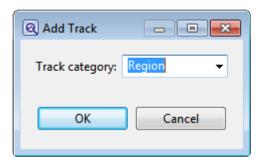
Part D: Work with region tracks

In this part of the tutorial, you will load two BED files associated with the E. coli assembly.

- MG1655_m56_v3.prophage-regions.bed. contains ten regions or intervals corresponding to the locations of prophages within the genome of this particular strain of E. coli.
- MG1655_m56_v3.Sigma32_regulon.bed contains regions corresponding to genes that are known to be induced by heat shock (i.e., high temperatures).

Each BED file will appear as a separate Region track in GenVision Pro, just like the Features and Coverage tracks discussed in earlier parts of this tutorial.

- 1. BED files must be added one at a time. Select File > Add Track. Navigate to and double-click on MG1655 m56 v3.prophage-regions.bed.
- 2. Leave the **Track category** set to **Region** and press **OK**.



3. Repeat the above steps to add MG1655_m56_v3.Sigma32_regulon.bed.

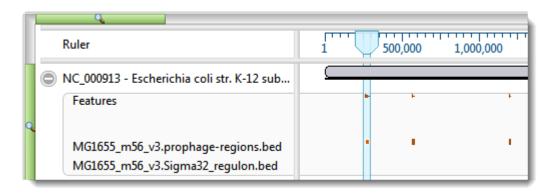


Note: Tracks can be renamed by right-clicking on their names in the Tracks panel (not in the views) and choosing Rename or Rename with Fields. This is the same basic procedure used to rename sequences in Part A of this tutorial. You can rename the tracks if you wish, but it is not necessary for this tutorial.

Now that the tracks are loaded into the session, the next step is to find regulon members that reside within prophages. One way to do this is to use the <u>Overview</u> to manually navigate to each of the ten prophage regions in the BED file and visually check whether any regulon regions overlap prophage regions.

- 4. In the Overview, move the horizontal zoom slider to the left, if it is not already there. Even zoomed out the whole way, you should be able to see 10 orange regions in the MG1655 m56 v3.prophageregions.bed track.
- 5. In the Analysis view, move the horizontal zoom slider until the viewport (the semi-transparent vertical

bar in the Overview) is narrower than 1 cm in width. Then drag the Viewport so that it is over the left-most orange region.



In the Analysis view, adjust the horizontal <u>zoom</u> slider so that the entire CP4-6 prophage region is visible. Note that the regulon track below it does not contain any data at this point, indicating that there are no overlapping regulons in this region.



6. Repeat the previous step for each of the other nine prophage regions. You should see overlapping regulons at two of them.

One is the fifth prophage from the left ("Qin"):



The other is the ninth from the left ("CP4-57"):

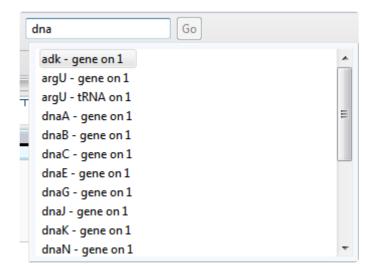


7. Under CP4-57, click on the orange arrow to see the regulon member's region name and coordinates in the <u>Details panel</u>.



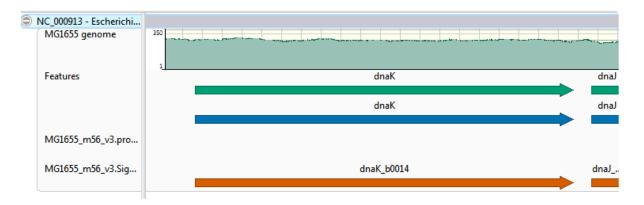
Imagine you are interested in the gene dnaK, and want to see if is a member of the regulon.

8. In the navigation area at the top of the window, start typing **dnaK**. As you type, a list of matches appears.



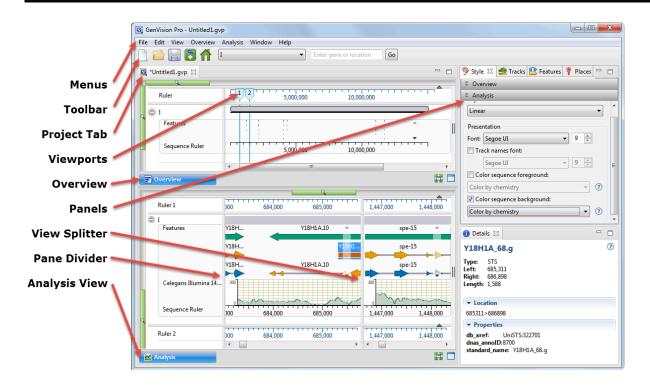
You can select **dnaK** from the list by double-clicking on it. Or simply finish typing the name manually and press **Go** or **Enter/Return**.

In the Analysis view, you can see that dnaK is indeed a member of the regulon.



Congratulations on completing the tutorial!

GenVision Pro Interface



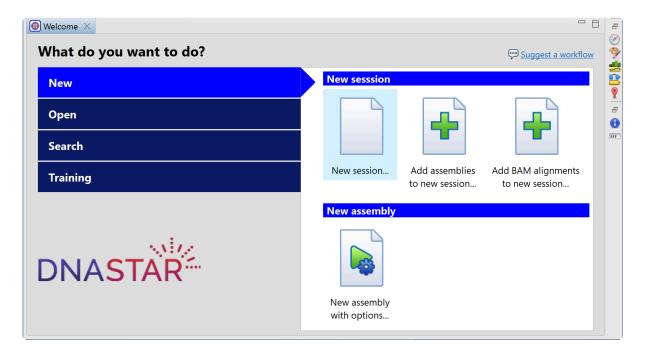
The following table describes components of the GenVision Pro user interface, along with links to more indepth information about each item:

Item	Description		
Menus	The GenVision Pro menus are: GenVision Pro (available only on Macintosh), File , Edit , View , Overview , Analysis , Window and Help . Some menu commands can also be initiated by keyboard shortcuts, button tools, navigation tools, or right-click options.		
Button tools	The GenVision Pro button tools are above the Overview on the left.		
Navigation tools	The GenVision Pro navigation tools are above the Overview, to the right of the button tools.		
<u>Sessions</u>	In GenVision Pro, each session is displayed in a separate tab with a .gvp extension.		
Overview	Displays any enabled tracks, including a ruler and features.		
Genome view	Lets you choose the chromosome (or other item) to analyze in the other views.		
Analysis view	Shows a reference sequence and enabled <u>tracks</u> .		
<u>Panels</u>	The <u>Style</u> , <u>Tracks</u> , <u>Features</u> , <u>Places</u> and <u>Details</u> panels provide details about the views on the left,		

	or affect what is displayed there.
View splitter	The view splitter is used to divide the Analysis view or Overview vertically into multiple panes.
Header/ Footer	A header and footer are collapsible sections above and below the selected reference sequence. The Overview provides a header, and the Analysis view provides a header and footer.
Sequence Ruler	Sequence rulers pertain only to the sequences just above them.

Welcome screen

The Welcome "project" opens when you launch GenVision Pro and is a central location for opening or creating sessions or accessing help. As you open sessions, the Welcome "project" will remain as the leftmost tab unless you manually move or close it. We recommend leaving the Welcome tab in place, since it has time-saving shortcuts to useful commands.



Tab on left	Option on right	Shortcut to this menu command:		
	New session	File > New GenVision Pro Session. See Create a new session.		
	File > New GenVision Pro Session followed by File > Add Assemblies. See <u>Create</u> a new session and <u>Add assemblies</u> .			
assembly screen. Choose a workflow and proceed through the SeqMan NGen		File > Add BAM Alignments. See Add BAM files.		
		File > New Assembly with Options . This launches SeqMan NGen at the Workflow screen. Choose a workflow and proceed through the SeqMan NGen wizard to create a new assembly that can be opened in GenVision Pro. See <u>Create a new assembly using SeqMan NGen</u> .		

Open	Open previously- saved session	File > Open. See Open an existing session.	
	Recent documents	File > Recent Documents.	
	Search	Search > Search. See <u>Search for sequence matches</u> .	
Search	Search for text in NCBI Nucleotide	Search > Search for Text In > NCBI Nucleotide. See Search for sequence matches.	
	Search for text in NCBI Protein	Search > Search for Text In > NCBI Protein. See Search for sequence matches.	
Training	User guide	Help > User Guide. Opens this User Guide.	
Training	Tutorials	Help > Tutorials. Opens this User Guide at the topic GenVision Pro Tutorials.	

Views

GenVision Pro views are the large windows on the left of the interface. The following views may be available depending on circumstances. For instance, you may need to select something or perform a search of the BLAST or Entrez databases before a particular view is available.

- Genome view
- · Analysis view
- Overview
- Text view
- Table view
- Console view

You can use the <u>Style panel</u> and other methods to <u>customize the look and layout</u> of these views.

Analysis view

The Analysis view shows tracks for a single sequence. The view is empty unless the session contains only one sequence; or until a single sequence has been chosen in the <u>Overview</u> or <u>Explorer panel</u>, or by using the <u>navigation tools</u> or <u>Places panel</u> bookmarks. At that point, the Analysis view appears automatically and is populated with data.

To access the Analysis view at any time, click on the **Analysis** tab, the dark blue tab in the image below, or use **View > Analysis**.



By default, the Analysis view displays a header and footer, along with a sequence and its corresponding data tracks. Tracks in the Analysis view can have different appearances depending on the zoom level.

The following table shows tasks associated with the Analysis view:

Task	How To
Display the search tools	To display the <u>search</u> tools in the upper left of the Analysis view, use the Find tool ().
	To sort read files (the files in the left margin of the view) according to the selected criteria, use the Sort reads tool ().
Sort constituent sequences	 By Name – Sorts read files in alphanumeric order by name. This type of sorting is normally only used for Sanger reads, which are often user-named. This is a shortcut to the Sequence > Sort Reads > By Name command.

• By Position – Sorts read files by their position along the consensus. This is a shortcut to the **Sequence > Sort Reads > By Position** command. • By Variant – Brings all the reads containing variant bases to the top of the alignment stack. This is a shortcut to the Sequence > Sort Reads > By Variant command. When you choose this option, SegMan Ultra automatically loads and calculates the SNPs. Note that if you are working with a large .sqd project, this calculation can take a considerable amount of time. Use the **Compact text** tool (to adjust the zoom level to maximize the number of bases while Adjust the keeping the text legible. Use the **Restore default zoom** tool () to adjust the zoom level to the zoom level (wider) default setting, making the text more comfortable to read. Use the Export data tool () and choose between Export Aligned Sequences and Export Consensus. This tool acts as a shortcut to the File > Export Data > Aligned Sequences and **Export** data File > Export Data > Consensus commands and opens the same popup dialogs. See Export data to a file for information on using the dialog. Use the **Export image** tool (). This tool acts as a shortcut to the **File > Export Image > Export** image **Analysis** command. See Export an image of a view for more information. Change the color Use the Style panel's Sequence section and/or Analysis section. scheme. font, or layout Click on the small dark gray arrows near the upper right corners of each section. Show/hide the header & footer TTGGGTTTA Click Analysis > Show Guideline. The presence of a check mark directs GenVision Pro to display the guideline. A vertical guideline is displayed by default, and is a navigational aid to help in guiding your eye between sequences and tracks. Show a a a g c c t a a g c c t a a vertical cTel52S.1 guideline II (1>15,279,421) 15,279,421 bp 2L52.2

Hide a vertical guideline

Click **Analysis > Show Guideline**. The absence of a checkmark means that only an abbreviated version of the guideline will remain visible in the header and footer.

Change the Analysis view color scheme

The foreground and background color schemes for the <u>Analysis view</u> are specified in the <u>Analysis section</u> of the <u>Style panel</u>. To open the section, click on the expand bar entitled **Analysis**, or choose **View > Style > Analysis**. Each available color scheme is described below:

Color Scheme	Description	Nucleotide Legend	Amino Acid Legend
Color by Chemistry	The default color scheme used by multiple DNASTAR applications. Amino acids are colored according to their side chain chemistry. While Color by Chemistry can be used for nucleotide sequences too, we recommend instead using a Solid background color for nucleotides and then selecting a Comparison like Color only nonconserved residues, Color only differences from reference, etc.	A C G T/U	Aromatic - Phenylalanine (F), Tryptophan (W), Tyrosine (Y) Acidic (negatively charged) - Aspartate (D), Glutamate (E) Basic (positively charged) - Arginine (R), Histidine (H), Lysine (K) Nonpolar (aliphatic) - Alanine (A), Glycine (G), Isoleucine (I), Leucine (L), Methionine (M), Proline (P), Valine (V) Polar (neutral) - Cysteine (C), Asparagine (N), Glutamine (Q), Serine (S), Threonine (T)
Physiochemical: Shapely	Matches the RasMol amino acid and RasMol nucleotide color schemes, which are, in turn, based on Robert Fletterick's "Shapely models."	A C G T	Alanine (A) Arginine (R), Lysine (K) Asparagine (N), Glutamine (Q) Aspartate (D), Glutamate (E) Cysteine (C), Methionine (M) Glycine (G) Histidine (H) Isoleucine (I), Leucine (L), Valine (V) Phenylalanine (F), Tyrosine (Y) Proline (P) Serine (S), Threonine (T) Tryptophan (W)

Physiochemical: Zappo	Colors residues according to their physico-chemical properties, and is also used in JalView (Waterhouse et al., 2009).	A C G T U	Aliphatic/hydrophobic - Alanine (A), Isoleucine (I), Leucine (L), Methionine (M), Valine (V) Aromatic - Phenylalanine (F), Tryptophan (W), Tyrosine (Y) Conformationally special - Glycine (G), Proline (P) Cysteine (C) Hydrophilic - Asparagine (N), Glutamine (Q), Serine (S), Threonine (T) Negatively Charged - Aspartate (D), Glutamate (E) Positively Charged - Arginine (R), Histidine (H), Lysine (K)
Physiochemical: UGene	<u>Unipro UGENE</u> coloring scheme.	C G T/U gap	Positive (KRH) Aromatic (FWY) Large aliphatic hydrophobic (ILM) Small hydrophobic (ST) GP EX

Spectral Colors: Taylor	Taken from Taylor (1997) and also used in JalView (Waterhouse et al., 2009).	C G T/U	Alanine (A) Arginine (R) Asparagine (N) Aspartate (D) Cysteine (C) Glutamate (E) Glutamine (Q) Glycine (G) Histidine (H) Isoleucine (I), Leucine (L) Lysine (K) Methionine (M) Phenylalanine (F) Proline (P) Serine (S) Threonine (T) Tryptophan (W) Tyrosine (Y) Valine (V)
Clustal X	Taken from Thompson et al. (1994). This color scheme is only available for protein alignments.	N/A	Hydrophobic (ACFHILMVWY) > 60% conservatism Negatively charged (DE) > 50% conservatism Postively charged (KR) > 60% conservatism Polar (STQN) > 50% conservatism Cysteine > 85% conservatism Glycine 85% conservatism Proline 85% conservatism Aromatic (FYW) 50% conservatism

Grayscale – Percentage Identity	This option uses only black/white/gray and is suitable for publications requiring grayscale images.	0-40% 41-60% 61-80% 81-100%	0-40% 41-60% 61-80% 81-100%
Percentage Identity	This option uses only shades of slate blue.	0-40% 41-60% 61-80% 81-100%	0-40% 41-60% 61-80% 81-100%

Wrap the Analysis view display

To wrap the display in the Analysis view:

In the Style panel's <u>Analysis section</u>, select **Automatic wrapping** from the **Layout** drop-down menu. The <u>Analysis view</u> will immediately wrap horizontally.

If you had been using a **Linear** layout and had already <u>split the sequence portion of the view vertically</u>, choosing **Automatic wrapping** will change vertical splits to horizontal ones. Also, the scroll bars will move from the bottom of the view to the right.

When the Analysis view is split, the <u>Overview</u> provides a separate "<u>viewport indicator</u> for each section. Splits are ignored when <u>exporting</u> an image, although the layout is preserved.

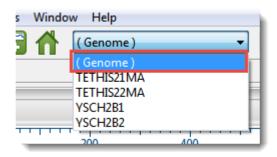
Genome view

The <u>Genome view</u> can display multiple experiments, chromosomes, contigs or sequences simultaneously.



To access the Genome view, do any of the following:

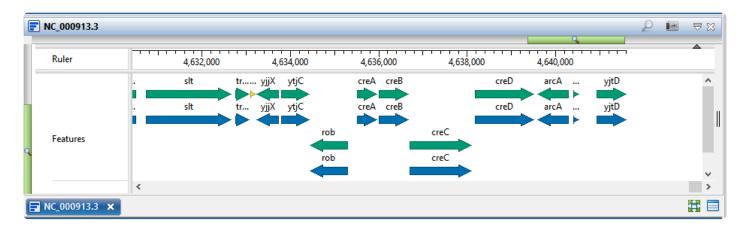
- Press the Home (button tool.
- Use the View > Genome > Show command.
- · Click on the Genome tab.
- Use the <u>navigation tools</u> to choose (**Genome**) from the drop-down menu.



Double-click on an item in this view to display it in the Analysis view and Overview.

Overview

The Overview shows an overview of the chromosome currently selected in the <u>Genome view</u> and its features. To access the view, click on the **Overview** tab (the dark blue part of the image, below) or use **View > Overview > (Name) > Show**.

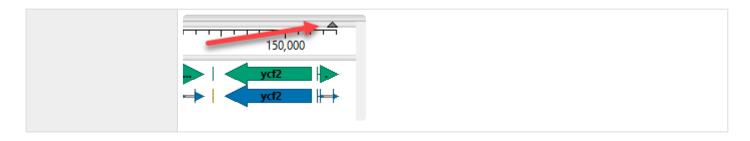


This view was released in Lasergene 17.3 and updated in 17.5, but still has limited features. We are developing a variety of very useful tracks for this view that will be available in Lasergene 18.0 and beyond. In the meantime, the Overview primarily serves as a way to view a selected chromosome at a different zoom level than is being used in the <u>Analysis view</u>.

The appearance of a ruler and/or features are controlled through "Overview"-related checkboxes in the <u>Tracks panel</u>. See <u>Work with Tracks</u> for more information.

The following table shows tasks associated with the Overview:

Task	How To	
Display the search tools	To display the <u>search</u> tools in the upper left of the Analysis view, use the Find tool (
Export image	Use the Export image tool (). This tool acts as a shortcut to the File > Export Image > Analysis command. See <u>Export an image of a view</u> for more information.	
Change the color scheme, font, or layout	Use the Style panel's Overview section.	
Show/hide the header & footer	Click on the small dark gray arrows near the upper right corners of each section.	



Reports view

The Reports view consists of two separate reports, each displaying data related to a single assembly or other data source.

To display the Reports view, select a row in the **Experiments** section of the <u>Explorer panel</u> and click the **Show Reports** () tool or use the **View > Reports** command. Then use the drop-down menu near the top left of the table to choose which version of the report to display.

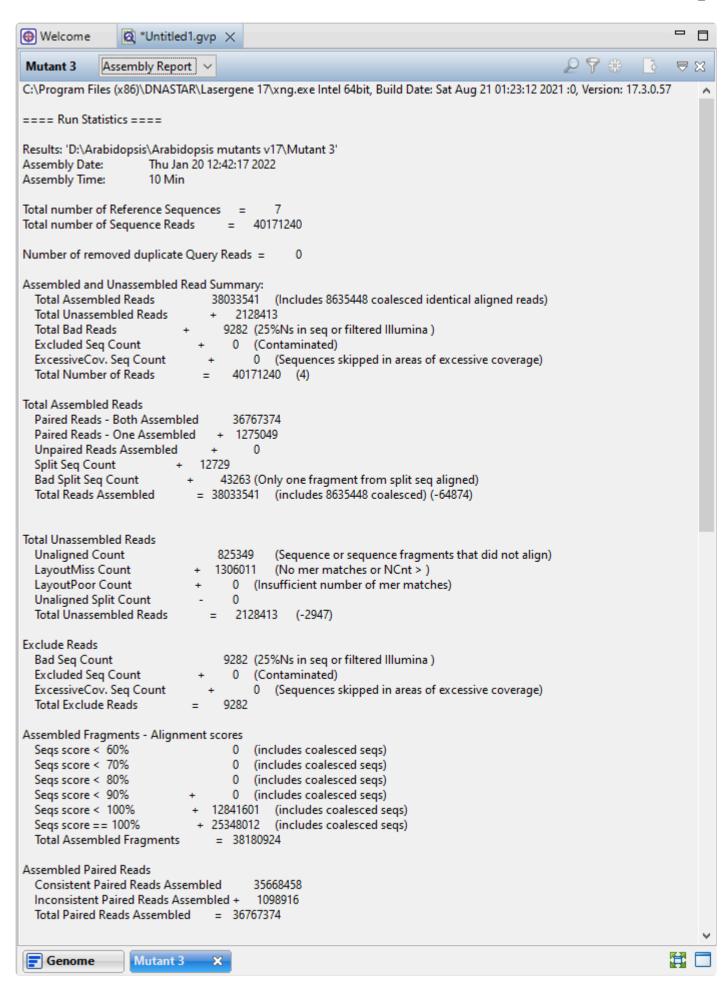
For information on viewing and understanding each of the two displays, see the topics:

- Assembly Report
- Per-Template Report

Assembly Report

The Assembly Report is one of several reports in the <u>Reports view</u>. This report summarizes the assembly time, the parameters used in constructing the contig, and each stage in the progress of an assembly, as well as the median pair distance and standard deviation for each set of a given sized pair library in an assembly.

To display this report, select a row in the **Experiments** section of the <u>Explorer panel</u> and click the **Show Reports** () tool or use the **View > Reports** command. Then use the drop-down menu near the top left of the table to choose **Assembly Report**.



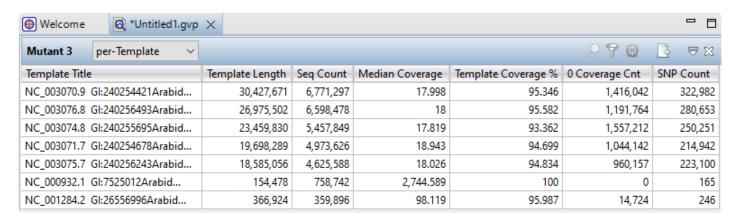
The table below describes some parameters of interest in this report.

Parameter	Description
Total Assembled Reads	Contains information about reads that made it into the assembly. "Coalesced identical aligned reads" indicates the number of reads that were treated as duplicates.
Total Unassembled Reads	Lists reasons why certain reads were not included.
Found SNP Cnt (incl. indel lengths)	(Variant assemblies only) Shows the number of SNPs that were found.
Assembly Parameters	(Variant assemblies only) Shows the SNP-calling parameters specified in the SeqMan NGen wizard for this assembly.
SCRIPT	Below the word "SCRIPT" is the entire assembly script used in SeqMan NGen when creating this assembly.

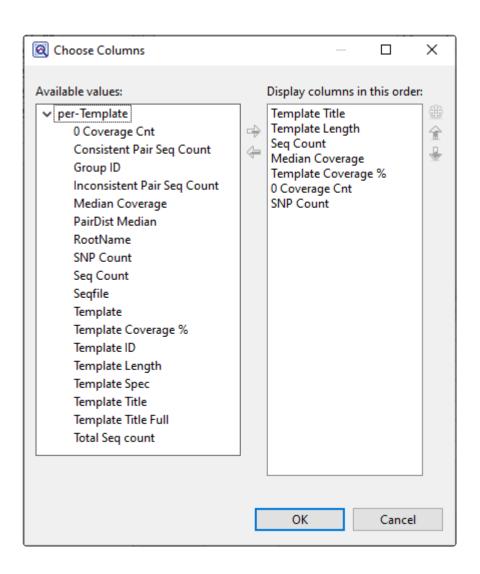
Per-Template Report

The Per-Template Report is one of the reports available in the Reports views and is available for some types of .assembly projects. This report displays the contents of a file written by SeqMan NGen's XNG assembler called perTemplateResults.txt. The report's table can be customized to display a wide variety of data columns in any desired order.

To display this report, select a row in the **Experiments** section of the <u>Explorer panel</u> and click the **Show Reports** () tool or use the **View > Reports** command. Then use the drop-down menu near the top left of the table to choose **per-Template**.

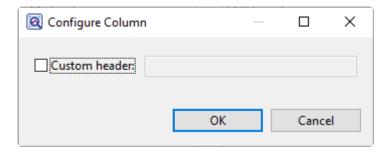


To select which columns to display in this table, or to rename or reorder the columns, click the **Choose or rearrange columns** () tool. The Choose Columns dialog appears.



Available columns are on the left, while currently-applied columns are on the right.

- <u>To add a column to the display</u>, select its name on the left and press the right arrow key to move it to the right.
- <u>To remove a column from the display</u>, select its name on the right and press the left arrow key to move it to the left.
- <u>To change the order of displayed columns</u>, select the column name you wish to move on the right, then use the up/down arrows to place it in the desired order.
- <u>To create a custom header for a displayed column</u>, select its name on the right, then choose the **Configure column** tool (). In the popup dialog, type in the desired name and press **OK**.



Descriptions of each column appear below in alphabetical order.

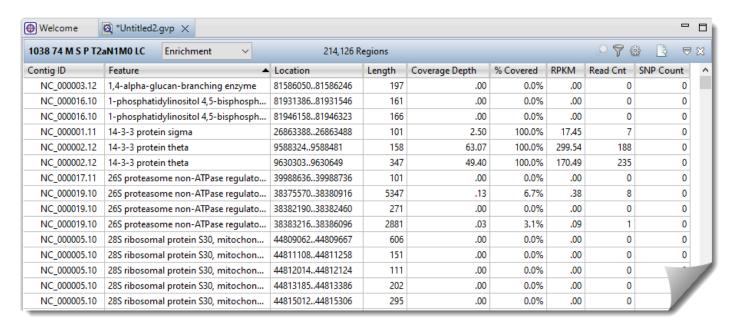
Column Name	Description			
0 Coverage Cnt	The number of bases in the reference sequence with zero coverage in the alignment.			
Consistent Pair Seq Count	For paired end sequencing (i.e. Illumina), the number of aligned read pairs that are correctly oriented and within the specified minimum and maximum distances.			
Group ID	Experiment Name assigned by the user when setting up the assembly in SeqMan NGen.			
Inconsistent Pair Seq Count	For paired end sequencing (i.e. Illumina), the number of aligned read pairs that are incorrectly oriented and/or outside the specified minimum and/or maximum distances.			
Median Coverage	The median depth of coverage across the individual contig.			
PairDist Median	The median distance between paired reads in the data set.			
RootName	The file name and complete file path of the template file without the .bam extension.			
SNP Count	The number of potential SNPs that were identified.			
Seq Count	The number of sequences in the template file.			
Seqfile	The file name and complete file path of the template file, including the .bam extension.			
Template	The file name and complete file path of the template file-of-files, including the .template.fof extension.			
Template Coverage %	The percentage of bases in the template covered by aligned sequence reads.			
Template ID	Internal identifier of the template used by DNASTAR for trouble-shooting.			
Template Length	The length of the template in base pairs.			

Template Spec	File pathway to the reference sequence.			
Template Title	The complete field of the reference sequence.			
Template Title Full	(DNASTAR internal use only)			
Total Seq Count	The total number of sequences in the layout for that template. Includes assembled and unassembled reads. Both halves of a split read are counted whether they aligned or not.			

Enrichment Report

The Enrichment Report displays exome and gene panel data and is one of several reports available in the Reports views. This report is only available for exome or gene panel assemblies that utilized BED files during the assembly process.

To display this report, select a row in the **Experiments** section of the **Explorer panel** and click the **Show Reports** () tool or use the **View > Reports** command. Then use the drop-down menu near the top left of the table to choose **Enrichment**. The report may take several minutes to populate.



The header displays the number of targeted regions (214,126 regions in the example above). The table can be customized to display a variety of data columns in any desired order (see second table, below, for a complete list).

The following table shows ways to customize the table or export table data.

Task	How to				
Jump to a specific region in the Analysis view	Double-click on any row in the table. To see the regions displayed as data tracks in the Analysis view, see Region tracks .				
Filter to locate regions	Press the Filter () tool in the top right of the view. The filter area opens at the top left of the view.				

Filter to find regions with insufficient coverage

with insufficient coverage

Show region if average coverage is below:

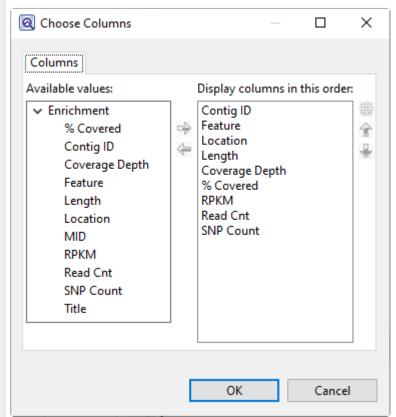
• Enter an upper threshold for average coverage (see **Coverage Depth** in the table below for definition). Only regions that do not meet the threshold will be shown.

or % covered below:

• Enter an upper threshold for % covered (see % **Covered** in the table below for definition). Only regions that do not meet the threshold will be shown.

Press **Enter** or click anywhere on the report to update it with your customized filter parameters.

Press the **Change options in this view** tool () in the top right of the view. To select which columns to dispersion this table, or to rename or reorder the columns, click the **Choose or rearrange columns** () tool. The Choose Columns dialog appears.

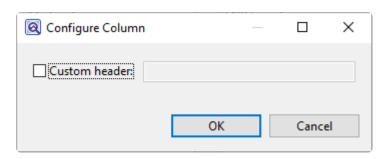


Add, rearrange, rename, or remove table columns

Available columns are on the left, while currently-applied columns are on the right.

- <u>To add a column to the display</u>, select its name on the left and press the right arrow key to move to the right.
- <u>To remove a column from the display</u>, select its name on the right and press the left arrow key to move it to the left.

- <u>To change the order of displayed columns</u>, select the column name you wish to move on the right then use the up/down arrows to place it in the desired order.
- To create a custom header for a displayed column, select its name on the right, then choose the **Configure column** tool (). In the popup dialog, type in the desired name and press **OK**.



Export data from the table

Press the **Export data** tool () in the top right of the view. This opens a save dialog in which you can save tabular data in comma-separated (.csv) or tab-separated (.tab) format. See Export data to a file for details.

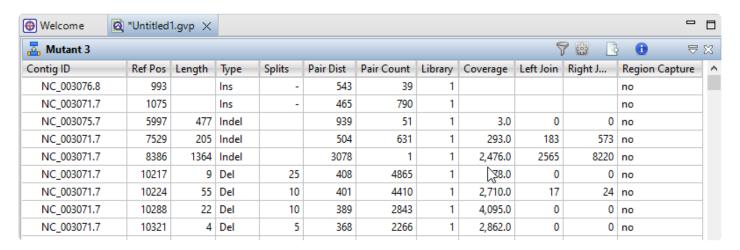
Descriptions of available columns appear in alphabetical order below:

Column Name	Description			
% Covered	The percent of the targeted region covered by aligned sequences.			
Contig ID	The name of the contig in which the exome was found.			
Coverage Depth	The average depth of coverage across the targeted region.			
Feature	The feature corresponding to the targeted region.			
Length	Total length of the exome in bases.			
Location	The range of sequence associated with the targeted region, including gaps.			
MID	Displays exons separately for each MID sample.			
RPKM	The RPKM value averaged over the set of replicate samples. Only available if replicates and replicate sets were specified in SeqMan NGen.			
Read Cnt	The number of reads in the exome.			
SNP Count	The number of variants in the exome.			
Title	The <title> field of the exon.</td></tr></tbody></table></title>			

Structural Variation view

In addition to variants and small insertions and deletions, genetic variation can also involve large scale rearrangements. These rearrangements may include large insertions and deletions, inversions, and translocations — collectively known as *structural variations* (*SV's*). During a templated data assembly with DNASTAR's SeqMan NGen, potential insertions and deletions are coded as such in the assembly output. When the assembly project is opened in GenVision Pro, the encoded SV information can be viewed in tabular format in the Structural Variation view. This view is only available for reference-guided assemblies in *.assembly* format.

To access the Structural Variation view, choose a contig from the <u>Explorer panel</u>. Then either choose **View** > **Structural Variation** > **(Contig Name)** or click on the **Show Table of Structural Variants** tool ().



For definitions of each data column, as well as instructions for showing, hiding, renaming or rearranging data columns, see Choosing and understanding Structural Variation view columns.

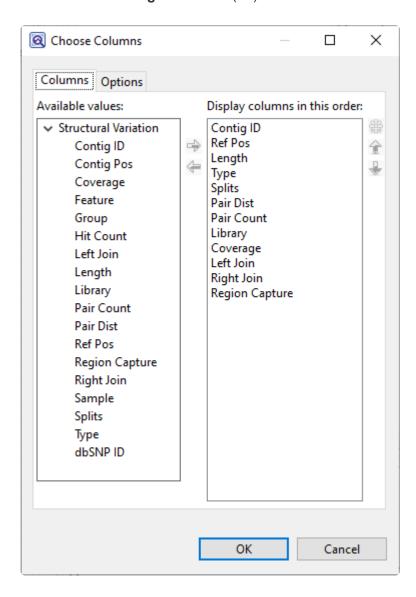
The table below describes tasks related to the Structural Variation view:

Task	How To
Open the Analysis view at the structural variation's position	Double-click on any row in the table.
Add, rearrange, rename, or remove table columns	Press the Change options in this view tool () in the top right of the view. For instructions and a description of each column type, see <u>Choosing and understanding Structural Variation view columns</u> .
Choose which items should appear in the view	Press the Filter () tool in the top right of the view. See <u>Filtering in the Structural</u> <u>Variation view</u> for detailed information.
Export data from the	Press the Export data tool () in the top right of the view. This opens a save dialog

table in which you can save the tabular data in comma-separated (.csv) or tab-separated (.tab) format. See Export data to a file for details.				
Toggle between showing/hiding the	Use the Show quick details tool (1) in the top right of the view.			
table header	0 confirmed 1 rejected 1,000,167 putative 0 mixed 49,401 filtered			

Choosing and understanding Structural Variation view columns

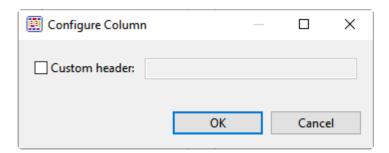
To select which columns to display in the <u>Structural Variation view</u>, or to rename or reorder them, click the **Choose or rearrange columns** () tool. The Choose Columns dialog appears.



Available columns are on the left, while currently-applied columns are on the right.

- <u>To add a column to the display</u>, select its name on the left and press the right arrow key to move it to the right.
- <u>To remove a column from the display</u>, select its name on the right and press the left arrow key to move it to the left.
- <u>To change the order of displayed columns</u>, select the column name you wish to move on the right, then use the up/down arrows to place it in the desired order.

• To create a custom header for a displayed column, select its name on the right, then choose the **Configure column** tool (). In the popup dialog, type in the desired name and press **OK**.



Descriptions of each column appear below in alphabetical order.

Column Name	Description				
Contig ID	The name of the contig in which the SV was found.				
Contig Pos	The position where the SV was detected in gapped coordinates. For insertions, a single coordinate is given. For deletions, the first and last coordinates of the deleted region are given.				
Coverage	The average depth of coverage across a putative deleted or indel region. The Coverage value is the mean depth of coverage over all columns, between the two edges of the deletion, in the original (unsplit) assembly. A "good" deletion will normally have much lower coverage than the flanking regions of the assembly. In cases where the deleted region is composed of non-repetitive sequence, the Coverage value will typically be zero or near zero. In cases where the deleted region is composed of repetitive sequence (e.g. an insertion sequence [IS] element), this value will typically be some fraction of the average coverage of the entire assembly. The exact value will depend on how many instances of the repeat are in the reference and the genome being sequenced. Note that clicking a row in the Contig Structural Variation report will take you to the Analysis view column just prior to the left edge of the deletion. However, Coverage does not measure the depth of coverage of that column, as the column is excluded from the range over which the mean is taken.				
	Coverage FAQ:				
	Q: In the SV Report for a gap-closure workflow, why does the Coverage column display low values (e.g., 0-7) for the majority of reference positions?				
	A: Coverage refers to the read depth in the original templated assembly and is calculated by summing the total number of aligned bases in that region and then dividing by the length of the region. Coverage is therefore expected to be low for deletions and indels. It is important to note that the Structural Variation Report is static and does not change after being created in SeqMan NGen.				

Q: Why doesn't the Analysis view for an SV with a Coverage of 4 have four sequences of coverage? A: That's because the deleted region is an insertion sequence (IS) repeat that is present elsewhere in the reference and in the data set. Since this is repeated sequence, SegMan NGen's assembler sees the match and places only a portion of the repeated reads in that spot. We can tell the region is deleted by the edges that form on both ends, the number of split reads defining the deletion endpoints, the relatively low depth in the deleted area and the lack of left and right joining pairs (pairs with one end outside the deletion and the other end in the putative deleted region). In cases where the deleted region is composed of unique sequence, the coverage will be zero or very near zero. The dbSNP rs ID, if available, for positions with known variants. Double-clicking on the entry opens dbSNP ID the corresponding page at dbSNP. The feature(s) affected by the SV. For deletions affecting multiple features, the first and last feature are displayed separated by an ellipsis (e.g. polB...ilvH). To see a complete list of affected genes for **Feature** a given SV, select the SV and use the Features > Show Feature Table command. Note that the Contig Structural Variation report does not explicitly show features of type "DNA_SPLIT," as this feature type merely denotes the presence a structural variation. Specifies whether the variant was called from the NGS or the Sanger data. The NGS group is named using the Read technology selected in the SeqMan NGen wizard. If you are investigating a Group putative variant that was identified in the initial run of the NGS data, you can find that position here and open it in the Analysis view to check whether there is a Sanger-called variant at the same position. **Hit Count** Number of reads used to calculate the Structural Variation. (DNASTAR internal use only) Left Join Length Total length of the structural variation in bases. A number corresponding to the library with information in that row. This column is present only for Library assemblies with multiple mate pair libraries of different insert sizes. General information on each library can be seen on a per contig basis using the Contig > Contig Info command. Pair The number of mate pairs spanning the SV. For assemblies with multiple mate pair libraries of Count different insert sizes, a separate row for each library with spanning pairs is shown for each SV. The median distance between mate pair reads spanning the SV. For deletions, this distance is approximately the average distance between all pairs in the assembly plus the size of the deleted segment. For insertions smaller than the insert size of the mate pair library, this distance is **Pair Dist** approximately the average distance between all pairs in the assembly minus the size of the inserted segment. For insertions larger than the insert size of the mate pair library, no distance is reported because no pairs will span the insertion. For assemblies with multiple mate pair libraries of different insert sizes, a separate row for each library with spanning pairs is shown for each SV. Ref Pos the position where the SV was detected in ungapped reference coordinates.

Region Capture	Whether or not the identified region overlaps a segment in the region capture file (i.e., the BED file).			
Right Join	(DNASTAR internal use only)			
Sample	The sample name.			
Splits	The number of "split reads" defining the breakpoints of a deletion.			
Туре	The type of SV: Ins (insertion), Del (deletion), or Indel (substitution).			

Filtering in the Structural Variation view

From the <u>Structural Variation view</u>, you can open a filter dialog to specify which SVs are shown in the table. Start by pressing the **Filter** tool (**7**) in the top right of the view. The filter area opens at the top left of the view.



- To the right of **Min splits**, enter the minimum number of "split reads" needed to determine the breakpoints of a deletion. Leave this section blank if you do not want to set a threshold.
- To the right of **Min coverage**, enter the minimum depth of coverage needed to determine an indel. Leave this section blank if you do not want to set a threshold.
- Check or uncheck the boxes next to Ins, Del and Indel to show/hide insertions, deletions, or both types (indels).

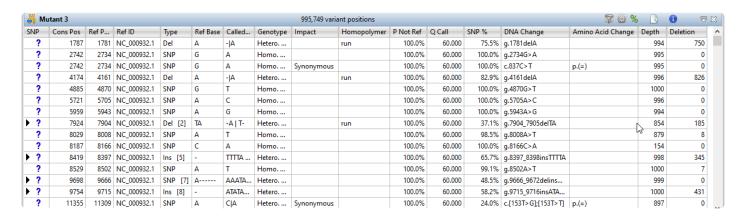
Click anywhere on the Structural Variation view to update it with your customized filter parameters.

Variants view

The Variants view provides a table with a row for each variant position. This view is only available for reference-guided assemblies in .assembly format.

To access the Variant view, choose a contig from the <u>Explorer panel</u>. Then either choose **View > Variants >** (Contig Name) or **Variants > Show Variant Table** or right-click on the selected contig and choose **Show Variant Table**.

Each row in the table summarizes information for all of the variant bases in an aligned column. If you are viewing a SeqMan NGen assembly, only variants meeting the SNP filter stringency (**High**, **Medium** or **Low**) that you specified in the SeqMan NGen wizard are displayed.



For definitions of each data column, as well as instructions for showing, hiding, renaming or rearranging data columns, see Choosing and understanding Variants view columns.

The table below describes tasks related to the Variants view:

Task	How To
Open the Analysis view at the variant position	Double-click on any row in the table. This is especially useful in situations such as an amplicon assembled against a much longer reference. In this case, the <u>Analysis view</u> may appear virtually empty when viewed on its own. However, double-clicking on a row in this table will open the Alignment view at a populated position.
Show, hide, rename or rearrange data columns	For instructions, see Choosing and understanding Variants view columns.
Select a	Use the commands in the Variants menu or right-click menu options with the same names to

category of variants

select a large group of variants at once. Regardless of the current selection, you can select Variants > All or Variants > None. If you have already made a selection in the table, you can select everything that is not currently selected by using **Variants > Invert Selection**.

Note that as of Lasergene 17.5, the option **Select All of Type** is not yet functional in GenVision Pro.

The leftmost column of the table (represented by a small triangle in the header) denotes the status of a variant. Initially, all variants are marked as putative.

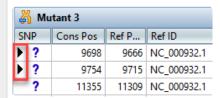
Change the status of selected

variants

- ? Putative (the initial status for all variants in the table)
- ✓ Confirmed
- × Rejected

To change the status of variants, select one or more rows, then right-click on the leftmost column and choose Confirm (Reject, Question) Selected Variants. Alternatively, you can make the selection, then use the following menu commands: Variants > Confirm (Reject, Question) Selected Variants.

If there were multiple reference bases at the location of a variant, the leftmost column will feature a small triangle in that row. Click on the triangle to show the multiple reference bases.



Hide/show multiple bases

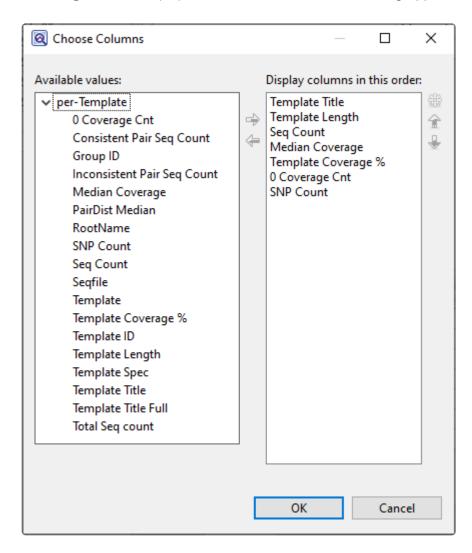
In .assembly projects, variants in adjacent columns are coalesced into a single insertion or deletion if they are of the same type, and if at least 80% of the reads with the called variant in one column have a variant in the adjacent column. Each coalesced multiple-base insertion or deletion references can be opened to reveal individual variants by clicking the corresponding triangle to the left of the SNP column. After clicking a triangle, information on each position of the insertion/deletion is displayed in a separate row.

F	?	8419	8397	NC_000932.1	Ins [5]	-	ППА	Hetero. Ref.
	?	8419	8397		Ins	-	TJ-	Hetero. Ref.
	?	8420	8397		Ins	-	TJ-	Hetero. Ref.
	?	8421	8397		Ins	-	TJ-	Hetero. Ref.
	?	8422	8397		Ins	-	TJ-	Hetero. Ref.
	?	8423	8397		Ins	-	A -	Hetero. Ref.
	?	8529	8502	NC_000932.1	SNP	Α	T	Homo. Variant
Þ	?	9698	9666	NC_000932.1	SNP [7]	A	AAATATC	Hetero. Ref.

Choose which items should appear in the view	Press the Filter () tool in the top right of the view. See <u>Filtering in the Variants view</u> for detailed information.
Add, rearrange, rename, or remove table columns	Press the Change options in this view tool () in the top right of the view. For instructions and a description of each column type, see <u>Filtering in the Variants view</u> .
Show counts as a percent	Press the Show counts as percent tool (%) in the top right of the view. This affects columns named Cnt .
Export data from the table	Press the Export data tool (in the top right of the view. This opens a save dialog in which you can save the tabular data in comma-separated (.csv) or tab-separated (.tab) format. See <u>Export data to a file</u> for details.
Toggle between showing/ hiding the table header	Use the Show quick details tool (i) in the top right of the view. 0 confirmed 1 rejected 1,000,167 putative 0 mixed 49,401 filtered

Choosing and understanding Variants view columns

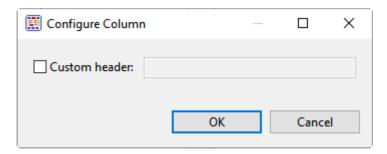
To select which columns to display in the <u>Variants view</u>, or to rename or reorder them, click the **Choose or rearrange columns** (tool. The Choose Columns dialog appears.



Available columns are on the left, while currently-applied columns are on the right.

- <u>To add a column to the display</u>, select its name on the left and press the right arrow key to move it to the right.
- <u>To remove a column from the display</u>, select its name on the right and press the left arrow key to move it to the left.
- <u>To change the order of displayed columns</u>, select the column name you wish to move on the right, then use the up/down arrows to place it in the desired order.

• To create a custom header for a displayed column, select its name on the right, then choose the **Configure column** tool (). In the popup dialog, type in the desired name and press **OK**.



Descriptions of each column appear below in alphabetical order.

Column Name	Description
Amino Acid Change	 This column is only available if Show Codon Bases & Distance to feature is unchecked. The column shows the change(s) in the amino acid sequence, using the nomenclature established by the Human Genome Variation Society and The Sequence Ontology Project. This includes: Conservative in-frame insertions. Example: p.K2_M3insQSK denotes that the sequence GlnSerLys (QSK) was inserted between amino acids Lysine-2 (K) and Methionine-3 (M). Disruptive in-frame insertions. Example: p.C28delinsWV denotes a 3 bp insertion in the codon for Cysteine-28, generating codons for Tryptophan (W) and Valine (V). Conservative in-frame deletions. Example: p.(C28_M30del) a deletion of three amino acids, from Cysteine-28 to Methionine-30. Disruptive in-frame deletions. Example: p.(C28_M30delinsL) denotes a 9 bp deletion including 2 bp from the codon for Cysteine-28 and 1 bp from the codon for Methionine-30 resulting in replacement of C28 to M30 with leucine (L).
Called Base	The dominant variant in the aligned column. In the case of a heterozygote call, both bases at the position are shown, separated by a vertical bar. For multi-base insertions, the inserted string is shown. For multi-base deletions, the deleted bases are represented with dashes (-).
Coding Feature Distance	Shows whether variants are within or near a named feature, and the distance from that feature. For .assembly files and certain .sqd files (e.g., from de novo or special templated workflows), the following color scheme may be used: • Gray + feature name + – Variant is within the named feature. • Pink + arrow + feature name – Distance from the variant to the closest upstream

	coding feature.
	 Orange + arrow + feature name – Distance from the variant to the closest downstream coding feature.
	Coding Feature Distance CT <- 146 CT <- 104 CT <- 103 CT <- 94 CT <- 90 CT <- 66 139 -> CT CT <within> CT <within> CT <within></within></within></within>
Codon	When a translated feature is present on the reference sequence at the position of a variant, a codon change is displayed. The codon and amino acid translation is shown for the reference sequence and compared to the codon and amino acid translation for the selected variant. The position number of the amino acid change is also displayed.
Cons Pos	The position on a gapped contig corresponding to this chromosome (for SeqMan Ultra and SeqMan NGen assemblies only).
COSMIC	The Catalogue of Somatic Mutations in Cancer (<u>COSMIC</u>) ID for positions with known variants. Double-clicking on the entry opens the corresponding page at COSMIC. For human assemblies only.
Cross Compare	(DNASTAR internal use only)
dbSNP ID	The dbSNP rs ID, if available, for positions with known variants. Double-clicking on the entry opens the corresponding page at <u>dbSNP</u> .
Deletion	The number of deleted bases in the Indel.
Depth	The number of reads overlapping the aligned column. Since this calculation disregards bases below the quality threshold, the Alignment View may show a greater number of sequences than the Depth shown in the Variants Summary Report. The default quality threshold for assembly in SeqMan NGen is 5.
DNA Change	Change(s) in the DNA sequence affecting either CDS features or splice sites are indicated using the nomenclature established by the Human Genome Variation Society (HGVS). A "c." prefix, followed by coordinates taken from the ORF, denotes a change in a CDS feature. For example: • Substitutions. Example: c.76A>C denotes that at nucleotide 76 an A is changed to a C.

- Insertions within coding regions. Example: c.76_77insT denotes that a T is inserted between nucleotides 76 and 77.
- Deletions within coding regions. Example: c.76_78delACT denotes an ACT deletion from nucleotides 76 to 78.

A "g." prefix followed by genomic coordinates denotes a change in the intronic region of a splice site.

Note: When a multibase variant affects both the intron and exon portions of a splice site, it is represented under two separate entries: one with g. coordinates and the other with c. coordinates.

Feature Name

If a variant is located within an annotated feature in the reference sequence, the feature type and name are displayed. A single nucleotide change may sometimes be reported as affecting multiple overlapping features. These can include different overlapping genes on the same or opposite strands, as well as alternatively spliced messages from the same gene. In this case, SeqMan NGen produces multiple VCF Variant table entries at the same position, one for each reported feature. A bracketed number follows the **Feature Name** to indicate which isoform from the Feature view table was used (e.g., TP53 [2]).

Note: If a non-gene feature ("mRNA", "CDS", etc.) exists in the template file, but has no corresponding "gene" feature, SeqMan NGen adds the "gene" feature automatically during assembly. The locations of any automatically added "gene" annotations are indicated by asterisks (*) in this column.

For variants within a gene feature, the feature type is shown in the following order of precedence: CDS, mRNA, Gene. If Show Codon Bases Distance to feature is selected, this column also contains a feature designation if the variant is within 150 bases of the nearest exon. Therefore, it is possible for a variant that is in a gene to also be listed as a CDS, mRNA, etc. When Show Codon Bases is checked, the Feature Type column will also show the distance to the nearby exon and an arrow indicating the direction of the feature. Feature types for different variant locations are shown below:

Feature Type

- gene Within a gene feature, but not included in an mRNA or CDS feature for that gene. (Variants within the intron portion of a splice site are indicated as CDS features.)
- CDS Within an exon or splice site.
- mRNA In the 5' and 3' untranslated portions of an mRNA.

Genotype

When the "Diploid" SNP detection method is used in a SeqMan NGen assembly, there are four possibilities: 1) homozygous variant (both alleles have the same base and it is different

from the reference), 2) reference (both alleles have the same base and it is the same as the reference), 3) heterozygous reference (two different alleles are called, one with the same base as the reference, the other with a variant base), and 4) heterozygous not reference (two different alleles, neither of which match the reference base). It is quite rare for the reference case to occur in the table. This only happens in cases where there is sufficient evidence of the possibility of a variant to pass the filtering threshold, but where the evidence is still quite weak. These cases are usually eliminated by even modest filtering. When the Haploid SNP detection method is used, only variant and reference are possible.

Note: In this column, if one or more of the adjacent variants is called as a heterozygote, the coalesced variant is also called a heterozygote. Therefore, for a coalesced variant to be called homozygous, all positions must be called homozygous.

The <u>Genomic Evolutionary Rate Profiling</u> (GERP) score representing the calculated evolutionary constraint at that position. GERP data is automatically delivered when you use DNASTAR's human template package prior to performing a templated assembly in SeqMan NGen. To limit the size of the data file required, only positions with scores of 1.0 or greater are displayed.

GERP

GERP is a tool that provides a score for each position in the human genome that estimates whether that position is under purifying selection or not (Davydov et al. 2010). GERP uses alignments between the human genome and 33 other mammalian genomes to quantitate the position-specific constraint in terms of rejected substitutions, defined as the difference between the neutral rate of substitution and the observed rate, estimated by maximum likelihood. Substitutions in sites under selection are assumed to be more deleterious than those not under selection. Scores range from negative values to ~6. Positions with scores below or near zero are not under selection. Conversely, the more positive the score, the more constrained the position. GERP information can be useful in evaluating the impact of non-synonymous variants in coding regions and the impact of changes in or near promoter elements, among others.

Homopolymer

Impact

Indicates whether the variant occurs within a homopolymeric run, which is defined as two or more identical bases in a row. When using Pacific Biosciences (PacBio) or Ion Torrent data, SeqMan Pro may not list all homopolymeric indels. When possible, insertions or deletions are placed at the 5' end (top strand) of the run during alignment.

The impact of the variant or indel on the genome, displayed as one of the following values:

- Synonymous No amino acid changes.
- Non Synonymous Amino acid substitution only.
- Nonsense Amino acid to translational stop.
- Frameshift An indel within a coding region and which is not a multiple of 3,

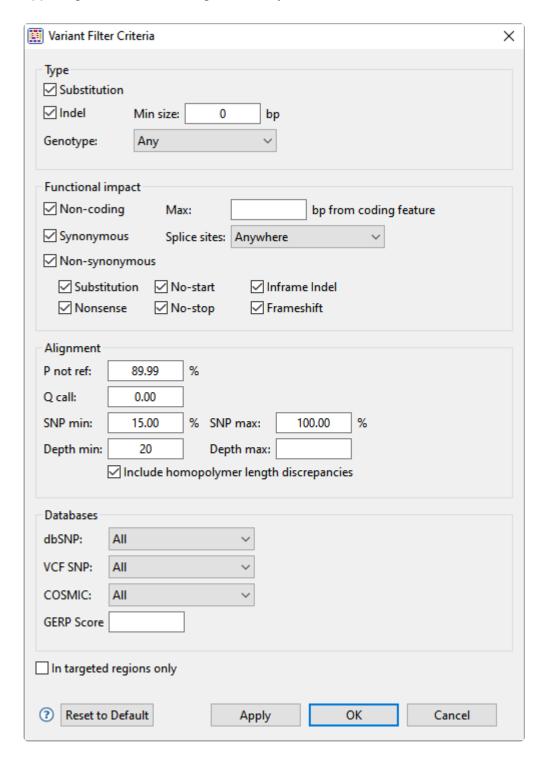
Copyright DNASTAR, Inc. 2022

	thereby changing the reading frame.
	No Start – A change that disrupts the start codon.
	 No Stop – A change that converts a stop codon to an amino acid, and thereby extends the reading frame.
	• Inframe Insertion – An insertion within a coding region whose length is divisible by 3. The type is followed by the word Conservative if the insertion occurs between two codons, and Disruptive if it occurs with a codon.
	 Inframe Deletion – A deletion within a coding region whose length is divisible by 3. The type is followed by the word Conservative if the insertion occurs between two codons, and Disruptive if it occurs within two codons.
	If sorting by the Impact column, the column is ordered by severity. For example, a Frameshift is more severe than a Nonsense change.
PDB ID	Worldwide Protein Data Bank (PDB) ID number.
P Not Ref	The probability that this position does not match the reference. For combined SNPs and indels, P not ref will be the minimum of the P not refs in the used columns.
Q Call	The Phred-like quality score of the called genotype. It is a measure of the confidence that the SNP is present in the sample on a 0-60 log10 scale. For combined SNPs and indels, Q call will be the minimum of all available columns at that reference position.
Ref Base	The reference sequence base in this position. For multi-base deletions, the reference sequence of the string is shown, beginning with the base at the Ref Pos coordinate. If there is no reference sequence present, the Ref Base column displays the most frequently occurring non-ambiguous base at this position. If no such base exists, the consensus base at this position is shown.
Ref ID	Reference sequence or chromosome.
Ref Pos	Reference position that does not include gaps. Coordinates matching entries in the VCF Variant table are shown. For deletions, Ref Pos is the genomic coordinate of the first deleted base. For insertions, Ref Pos is the genomic coordinate of the base preceding the insertion.
Region Capture	Indicates whether the variant occurs within a region specified in the .bed or manifest file used. Values are Yes and No .
Residue Count (A Cnt, etc.)	The number of bases of this type called in the aligned column. A dash (-) represents the reference base.
Sample	The sample name.
SNP	A manual evaluation score for the SNP, with a question mark being the default. See Change

	the status of a selected variant in the table below for a legend and instructions.
SNP %	The percentage of the sequence at this position in the assembly which varied from the reference.
Splice	Variant is in or near an exon splice site. Splice site variations are changes to the 5' ("donor") or 3' ("acceptor") consensus splice site sequences. The DNA sequence for the donor is 5'-AGGTRAGT-3' and for the acceptor is 5'-YYYYYYYYCAGGT-3'. Note that the AG dinucleotide on the 5' end of the donor and the GT dinucleotide on the 3' end of the acceptor are within the exon. Therefore, changes at these positions can also cause changes in the amino acid sequence of the resulting proteins. Changes in the intron portion of the splice site are marked as "Splice" in the column while those in the exon portion of the site are labeled "Splice in CDS." Only the position where the change occurs is considered, not the identity of the base.
Trace%	(DNASTAR internal use only)
Transcript ID	Transcript ID number from ENSEMBL.
Туре	Specifies the variation type as SNP, Del (deletion) or Ins (insertion). For .assembly files and certain SQD files (e.g., from de novo or special templated workflows), pink typeface may be used to indicate non-synonymous variants.
User ID	Positions corresponding to a custom VCF Variant Table are labeled with the ID from that set.

Filtering in the Variants view

From the <u>Variants view</u>, you can open a filter dialog to specify which variants are shown in the table. To open the Variant Filter Criteria dialog, use the **Variants > Filter** command, press the **Filter** tool (**V**) in the upper-right of the view, or right-click anywhere in the view and choose **Filter**.



In the Type section:

Choose to display **Substitutions** and/or **Indels** by checking the boxes next to their name(s). If you want to see only variants greater than a certain length threshold, type that threshold into the **Min size** box.

Use the **Genotype** drop-down menu to limit the display to **Annotated SNPs only** or **Novel SNPs only**. To display both types, keep the default setting of **Any**.

In the Functional impact section:

Choose whether to display indels having the following functional impacts:

- **Non-coding** The indel appears in a non-coding region of the sequence, and will therefore have no functional impact.
- **Synonymous** The indel does not cause an amino acid change, and will therefore have no functional impact.
- Non Synonymous Amino acid substitution only.
 - **Substitution** Any change from one nucleotide to another, regardless of effect or lack thereof.
 - Nonsense Amino acid to translational stop.
 - No Start A change that disrupts the start codon.
 - No Stop A change that converts a stop codon to an amino acid, and thereby extends the reading frame.
 - **Inframe Indel** An insertion or deletion within a coding region whose length is divisible by 3. The type is followed by the word Conservative if the indel occurs between two codons, and Disruptive if it occurs with a codon.
 - **Frameshift** An indel within a coding region and which is not a multiple of 3, thereby changing the reading frame.

In the Alignment section:

- P not ref The probability that the called base at this position is not the reference base. For coalesced variants, this value is equal to the minimum value of all "child" values. The minimum allowed value is 30%.
- Q call The Phred-like quality score of the called genotype. It is a measure of the probability that the called genotype is correct.
- **SNP min / SNP max** The minimum and maximum percentage of reads that should contain a SNP at a given position.

- Depth min / Depth max The minimum and maximum depth of reads needed to include the SNP.
- Include homopolymer length discrepancies If you are using Illumina or Sanger data, we recommend leaving this box checked. If using Ion Torrent data, we recommend unchecking the box. If you uncheck the box, GenVision Pro will remove all homopolymeric run length variants from the table.

In the Databases section:

- Use the dbSNP, VCF SNP and COSMIC drop-down menus to filter in/out subsets of variants from these databases.
- To filter out variants with a **GERP score** below a specified threshold number, enter that number in the text box.
- To filter out variants from locations other than the targeted regions (i.e., from a .bed or manifest file used in the assembly), check the box next to **In targeted regions only**.

If you wish to return to default settings, click **Reset to Default**.

To save your settings, click **Apply** or **OK**, or simply press **Enter**.

Text view

To access the Text view, <u>perform a BLAST search</u>. When the search is complete, the Text and <u>Table</u> views are displayed in a popup window that is separate from the main GenVision Pro window. To reopen the views at a later time, click on the **Status** link in the <u>Jobs panel</u>.

An example Text view is shown below:

```
BLASTN 2.10.0+
Reference: Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb
Miller (2000), "A greedy algorithm for aligning DNA sequences", J
Comput Biol 2000; 7(1-2):203-14.
Database: Nucleotide collection (nt)
         71,357,103 sequences; 466,239,233,881 total letters
Query= 1(4636670>4638166)
Length=1497
RID: E6W6KWF8013
                                                              Score
                                                                          E
Sequences producing significant alignments:
                                                             (Bits)
                                                                       Value
CP053604.1 Escherichia coli strain NEB10-beta chromosome, complet... 2765
                                                                        0.0
CP050289.1 Escherichia coli strain pK19EC149 chromosome
                                                                        0.0
                                                              2765
CP055669.1 Escherichia coli strain RHB30-C19 chromosome, complete...
                                                              2765
                                                                        0.0
CP059137.1 Escherichia coli strain OP50 chromosome
                                                              2765
                                                                        0.0
CP059043.1 Escherichia coli strain SY3626C1 chromosome, complete ... 2765
                                                                        0.0
CP058948.1 Escherichia coli strain SY3626C5 chromosome, complete ... 2765
                                                                        0.0
>CP053604.1 Escherichia coli strain NEB10-beta chromosome, complete genome
Length=4667764
Score = 2765 bits (1497), Expect = 0.0
 Identities = 1497/1497 (100%), Gaps = 0/1497 (0%)
 Strand=Plus/Plus
              ATGGGATATAGCCTGAGGGGCCTGTAATGCGTATCGGCATGCGGTTGTTGCTGGGCTATT
Query 1
              Sbjct 4662782 ATGGGATATAGCCTGAGGGGCCTGTAATGCGTATCGGCATGCGGTTGTTGCTGGGCTATT
                                                                     4662841
              TTTTACTGGTGGCGGTGGCAGCCTGGTTCGTACTGGCCATTTTTGTCAAAGAAGTTAAAC
Query 61
              Sbjct 4662842 TTTTACTGGTGGCGGTGGCAGCCTGGTTCGTACTGGCCATTTTTGTCAAAGAAGTTAAAC
                                                                     4662901
              CGGGCGTGCGAAGAGCAACGGAGGGGACGTTGATCGACACCGCAACGTTGCTGGCGGAGC
Query 121
```

In order, the view shows the following:

- The version of BLAST used in the search (e.g., "BLASTP 2.9.0+").
- References for the BLAST search engine.
- A description of the BLAST database used in the search, including the number of sequences.
- Information about the Query sequence, including its length and the Request ID (RID) number.
- A table of sequences producing significant alignments.
- Pairwise alignments of each significant match with summary information on the match.

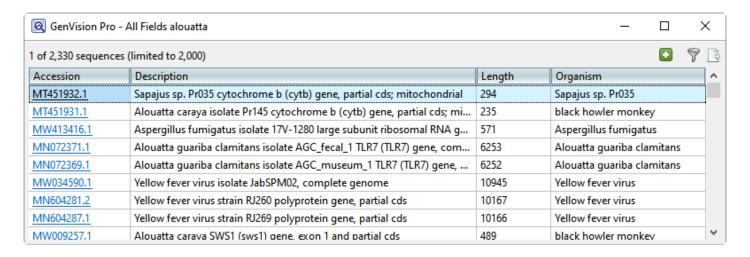
Tasks pertaining to the Text view:

Task	How to
Make selections in the view	See Make a Selection.
Copy all text	Right-click anywhere in the view and choose Copy .
Select all text	Right-click anywhere in the view and choose Select All .

Table view

To access the Table view, <u>perform a BLAST search</u>. When the search is complete, the <u>Text</u> and Table views are displayed together in a single popup window that is separate from the main GenVision Pro window. To reopen the views at a later time, click on the **Status** link in the <u>Jobs panel</u>.

If no significant matches were identified, the Table view will be empty. An example of a populated Table view is shown below:



Columns in the view's table are described below. Only a subset of these columns will be available, depending on whether you did a sequence or text search.

Column	Description
Accession	The NCBI "Accession" number for the match. Press the link to open the corresponding page on the NCBI website.
Description	The NCBI "Definition" field text.
Score	The similarity score for the match. (See the <u>NCBI definition for raw score</u>). In general, a higher score denotes a better match.
E-Value	The chance that the match is expected to occur by chance. (See the <u>NCBI definition for E value</u>). In general, a lower score denotes a better match.
%Coverage	The percentage of the query sequence that aligns to a match in the BLAST database. Example: Query has 596 bases, of which 565 align to a match in the NCBI database. **Coverage = 565/596 = 95%.
%Identity	The percentage of bases that match when the query and result sequence are aligned. (NCBI definition).
	Example: Query has 565 bases that align to the match and 3 bases that don't match. %Identity

	= 562/565 = 99.5%.
Мар	Graphic showing areas of agreement between the query and the match.
Length	(Entrez search only) The length of the sequence in bp.
Organism	(Entrez search only) The name of the organism.

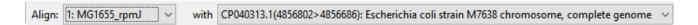
Tasks pertaining to the Table view:

Task	How To
If a BLAST or Entrez search was performed	
Launch the NCBI page for matching database entry	Click the hyperlink in the table's Accession column.
Copy all information from a search result row	Right-click on the row and choose Copy .
Save the sequence match in .gbk format	Right-click on the row and choose Download sequences . In the ensuing Save As dialog, choose a location in which to save the sequence. By default, the sequence is named using its Accession number.
Add sequences or matches	Use the Add sequences or matches tool (). Sequences are added back into the host application, not the job's document.
Filter results from a sequence or text search	Use the Filter results tool ($\overline{\mathbb{Y}}$). Learn more about <u>sequence search</u> or <u>text search</u> filtering.
Export data from the view	Use the Export data tool (), described in the topic <u>Export data to a file</u> .
	Only if a BLAST search was performed
To save results	Click the Save as tool (). A save dialog opens and allows you to save the search results as a <i>.search</i> format file. Use the command Search > Open Search File to open the search results again in the Table, Pairwise and Text views.
To display a pairwise alignment using a particular search result	If the Table view is open, click on the desired row in the Table view to see the corresponding pairwise alignment in the Pairwise view. As an alternative, open the Pairwise view manually and choose the desired target from the with drop-down menu.

Pairwise view

After the <u>Jobs panel</u> shows the **Status** of a <u>BLAST search</u> or <u>Entrez search</u> is **Done**, pressing the "Done" hyperlink opens the search results simultaneously in the Text, <u>Table</u>, and <u>Pairwise</u> views. All three views are contained in a large pop-up window that is separate from the main GenVision Pro window. The Pairwise view consists of three areas: toolbar, summary, and sequences.

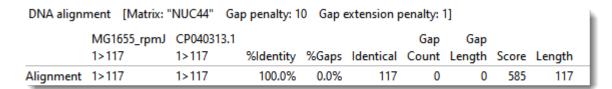
Toolbar area:



This area consists of two drop-down menus and one tool. The **Align** drop-down menu is used to specify the target sequence, which should be the longest sequence of the two. The **with** drop-down menu is used to specify the query sequence. Specifying a different sequence in either menu will initiate a new pairwise alignment.

The Export image* tool (), located on the right of the header, lets you export the summary and sequences areas.

Summary area:



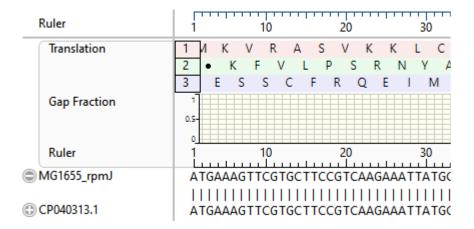
The summary area describes the alignment engine and settings that were used to create the pairwise alignment.

Value	Description
% Identity	Number of matching residues divided by the sum of all residues and gap characters. The quotient is multiplied by 100 to give the similarity as a percent.
% Gaps	The gap length divided by the consensus length. The quotient is multiplied by 100 to give the similarity as a percent.
Identical	Number of residues that are identical in both sequences.
Gap Count	Number of gaps (of any length) in either sequence of an alignment.

Gap Length	Number of gap characters in an alignment.
Score	The cumulative score for an alignment derived by adding the matrix-derived match scores and subtracting gap penalties.
Length	Number of residues in the consensus, including gaps that match in both sequences.

Sequences area:

The sequences area consists of nucleotide or amino acid codes for each sequence on the right, and their labels on the left. The order is always "target sequence" first, and "query sequence" second. A variety of optional <u>data tracks</u> can be displayed in this section.



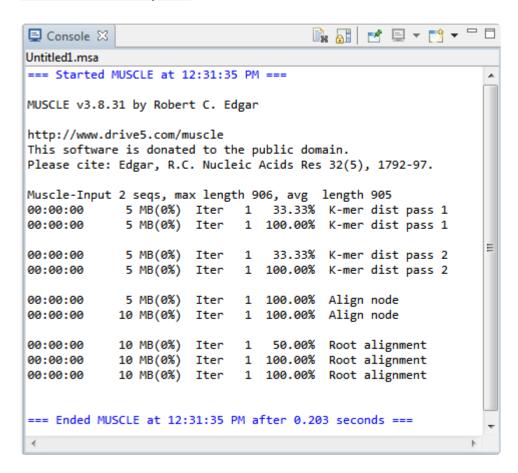
Console view

The Console view displays information, messages and warnings about the operation in progress, or an operation that was recently performed.

To display the console for a particular project:

Click the tab with the project name, then open its console using the **View > Console** command. Each view will have a different console.

To use the Console panel:



Each console has a maximum number of characters. If this maximum is exceeded, the oldest characters disappear in order to make room for new characters.

Right-click on the Console tab to access commands allowing you to move, hide, rearrange, or restore panels and views. (See <u>Customize the layout of other window components</u>).

The following tools are available in the upper right of the Console view:

Tool	Name	Description
	Clear Console	Clears all text from the active Console.
	Scroll Lock	Stops auto-scrolling caused by alignment output. When alignment output is long, this may fail due to a limit on the amount of text in the console.
	Word Wrap	Controls whether or not lines wider than the console panel should be wrapped.
ď	Pin Console	Keeps the current Console active even if another project is selected.
₽ ▼	Display Selected Console	Consoles can be displayed for one or more open projects. If more than one alignment is in progress, clicking the arrow to the right of this icon opens a drop-down menu from which the desired Console can be brought to the front.
<u>□</u>	Open Console	Opens a new Console. If more than one alignment is in progress at the same time, you can open a new Console and pin it to a particular run using Pin Console (above).
	Minimize	Minimizes the Console.
	Maximize	Maximizes the Console.

Panels

Settings panels allow you to optimize the look, contents and layout of the GenVision Pro view. To learn what can be done within each panel and their "areas" (expandable sections within each panel), click the links below:

- Style
- Tracks
- Features
- Places
- Details

Style panel

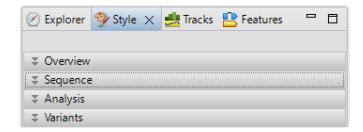
The Style panel lets you customize view formatting and layout.

To open this panel, do any of the following:

- Click on the **Style** tab (Style).
- Use the View > Style > Style command.
- Right-click within a view and choose Show Style Panel.

To change settings in the Style panel:

The Style panel has four sections, each pertaining to different views.



Click the links below to learn about each of the sections in the Style panel:

- The Overview section contains options affecting the Overview.
- The <u>Sequence section</u> contains options affecting the <u>Analysis view</u>.
- The <u>Analysis section</u> contains options affecting the wrapping and color schemes used in the <u>Analysis view</u>.
- The <u>Variants section</u> contains options affecting the <u>Variants view</u>.

To save styles and apply them to future sessions:

Once you have set up styles, you can save them and then apply them to future sessions, either by default or only when you specify that they be applied. You can choose whether to save and apply styles for just one particular view or for all views.

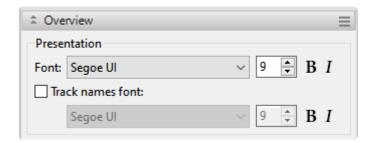
Note that some of the instructions below refer to the "hamburger" menu. To access commands on these menus, click on the expand bar of interest in the Style panel. Then click on the tool consisting of three horizontal lines.



Task	How To
Save all current styles to a file that can be applied to future sessions	Use View > Document Style > Save Style.
Save the styles from one particular view to a file that can be applied to that view in future sessions	Click the "hamburger" menu on the right of a Style panel expand bar and choose Save (Expand bar name) Style .
Save all current styles to a file that will be applied as the default for future sessions	Use View > Document Style > Save Style as Default.
Save the current styles from one particular view to a file that will be applied as the default for that view in future sessions	Click the "hamburger" menu on the right of a Style panel expand bar and choose Save (Expand bar name) Style as Default .
Apply styles that were saved earlier to the current sessions	Use View > Document Style > Apply Style
Apply styles that were saved for this view previously and apply them to the same view in the current sessions	Click the "hamburger" menu on the right of a Style panel expand bar and choose Apply (Expand bar name) Style .
Restore all DNASTAR default styles to the current sessions	Use View > Document Style > Restore DNASTAR Default Style.
Restore the DNASTAR default style for one particular view to that view in the current sessions	Click the "hamburger" menu on the right of a Style panel expand bar and choose Restore DNASTAR Default (Expand bar name) Style.

Overview section

The Overview section contains options affecting the <u>Overview</u>. To access this section, click on the <u>Style</u> <u>panel</u> expand bar entitled **Overview**, or choose **View > Style > Overview**.

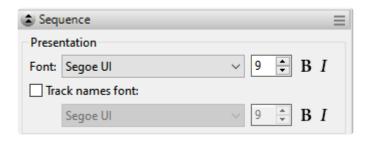


Customization for font style and size applies to all text in the Overview. Individual rows cannot have different styles or sizes.

Task	How to
Change font options for the ruler numbers in the Overview	Use the Font drop-down menu to choose a font for the ruler numbers in the Overview. Recently selected fonts (if any) appear at the top of the menu, and other available fonts appear below the row of dashes. Enter the preferred font size by typing a number between 4-72 in the box at right, or by using the corresponding up and down arrows (). To make the font Bold and/or <i>Italic</i> , press the B and/or <i>I</i> buttons, respectively.
Select a different font for sequence and track labels in the	First, check the box next to Track names font . If desired, use the corresponding drop-down menu to change the font from the current value. Enter the preferred font size by typing a number between 4-72 in the box at right, or by using the corresponding up and down arrows (Note: When Track names font is unchecked and tracks are open, Font and Font size
Overview.	(described above) affect everything in the view. When Track names font is checked and tracks are open, Font and Font size affect only the rulers, while Track names font and Track names font size affect sequence labels and track labels.

Sequence section

The Sequence section contains options affecting the <u>Analysis view</u>. To access this section, click on the <u>Style panel</u> expand bar entitled **Sequence**, or choose **View > Style > Sequence**.

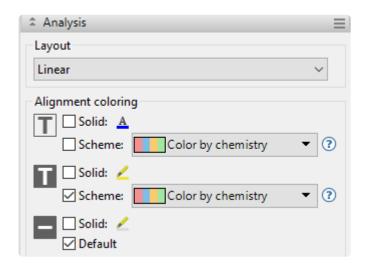


Customization for font style and size applies to all text in the Analysis view. Individual rows cannot have different styles or sizes.

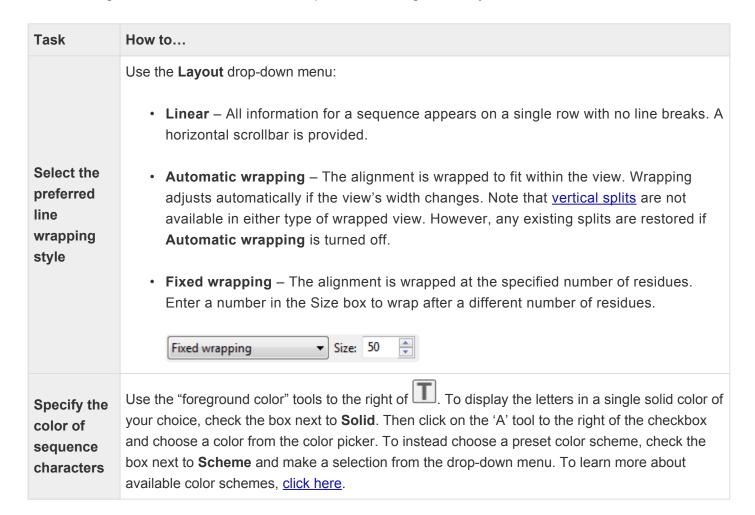
Task	How to
Specify font options for the ruler numbers in the Analysis view	Use the Font drop-down menu. Recently selected fonts (if any) appear at the top of the menu, and other available fonts appear below the row of dashes. Enter the preferred font size by typing a number between 4-72 in the box at right, or by using the corresponding up and down arrows (). To make the font Bold and/or <i>Italic</i> , press the B and/or <i>I</i> buttons, respectively.
Select font options for other items in the Analysis view	First, check the box next to Track names font , then use the corresponding drop-down menu to change the font from the current value. Enter the preferred font size by typing a number between 4-72 in the box at right, or by using the corresponding up and down arrows (). To make the font Bold and/or <i>Italic</i> , press the B and/or <i>I</i> buttons, respectively.

Analysis section

The Analysis section contains options affecting the wrapping style for rows in the <u>Analysis view</u>. To access it, click on the <u>Style panel</u> expand bar entitled **Analysis**, or choose **View > Style > Analysis**.



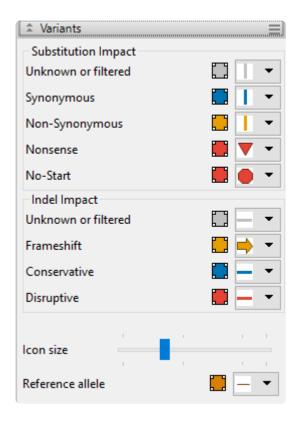
The following table shows tasks that can be performed using the Analysis section.



Specify the background color (boxes around each letter)	Use the "background color" tools to the right of . To use a single solid color of your choice, check the box next to Solid . Then click on the 'highlighter pen' tool to the right of the checkbox and choose a color from the color picker. To instead choose a preset color scheme, check the box next to Scheme and make a selection from the drop-down menu. To learn more about available color schemes, <u>click here</u> .
Specify the background color for gaps	Use the "background color for gaps" tools to the right of . To use a single solid color of your choice, check the box next to Solid . Then click on the 'highlighter pen' tool to the right of the checkbox and choose a color from the color picker. To choose the default (light gray), check the box next to Default . To learn more about available color schemes, <u>click here</u> .

Variants section

The Variants section contains options affecting how variants are displayed in the <u>Analysis view</u> when <u>Variants tracks</u> are applied. To access it, click on the expand bar entitled **Variants** in the <u>Style panel</u>, or choose **View > Style > Variants**.



Most of this section consists of a list of variant types, with a color chooser box and a drop-down menu to the right of each. To change the color for a type of variant, click on its color chooser box. To change the shape used to represent a type of variant, make a selection from the drop-down menu.

The same decisions can be made regarding how the **Reference allele** is represented.

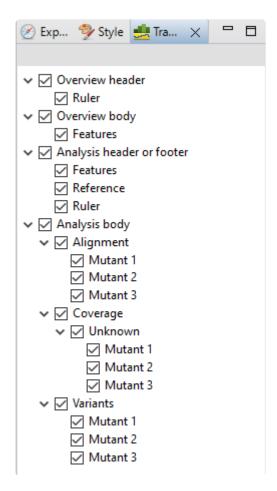
Use the Icon size slider if you want to increase/decrease the size of the icons in the Analysis view.

Tracks panel

The Tracks panel lets you manage <u>tracks</u> displayed in the views.

To open this panel, do either of the following:

- Click on the Tracks tab (Tracks).
- Use the View > Tracks > Tracks command.



The uppermost part of the Tracks panel lets you specify which tracks are visible, and configure their options and appearance. Put check marks next to available tracks that you wish to display, and remove check marks next to tracks that you wish to hide. For information about each track type, including user-editable options for the track, see Work with Tracks.



Note: Tracks are added to the views in the same order in which their boxes are checked in the Tracks panel. This means that the position of a track may change when its box is unchecked and then checked again.

To move up or down the list of tracks, either use the mouse wheel or use the vertical scrollbar on the right.

To see the detail tracks that have been checkmarked for display, expand one or more rows in the <u>Overview</u> and/or <u>Analysis view</u> by clicking the relevant plus sign icon (①).

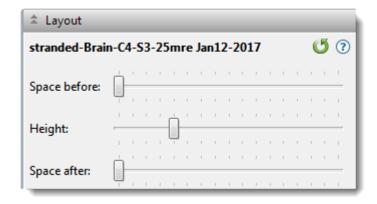
Below the list of tracks are two expandable sections:

- Layout lets you change the height of a track and create blank space above and/or below the track.
- Options lets you optimize how tracks appear in the Overview and Analysis view.

Layout section

The Layout section lets you change the height of a track and the space above and/or below a track. Changes made in this section are applied to all currently selected tracks.

Access this section by clicking on the <u>Tracks panel</u> expand bar entitled **Layout**, or by choosing **View > Tracks > Layout**.



- To control the amount of blank space above and/or below each of the selected track(s), use the Space before/after sliders.
- To control the heights of the selected tracks, use the **Height** slider. This option is enabled only when all selected tracks support user-scalable height.
- To return the sliders in this dialog to their default settings, click the **Restore default settings** tool ().

Options section

The Options section is used to edit settings for the track type currently selected in the <u>Tracks panel</u>, and varies in appearance depending upon the selection. Access the Options section by clicking on the Tracks panel expand bar entitled **Options** or by choosing **View > Tracks > Options**.

See the following topics for discussions of each track type and its corresponding Options section.

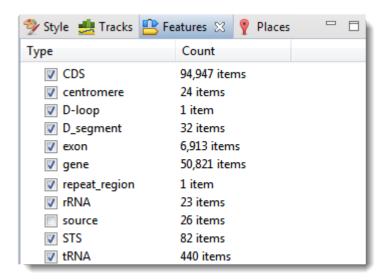
- Feature tracks
- Reference track
- Ruler tracks
- · Sashimi tracks
- Alignment tracks
- Coverage tracks
- Variants tracks

Features panel

The Features panel lets you specify which feature types should be visible in the <u>Overview</u> and the <u>Analysis</u> <u>view</u> when the feature and region <u>tracks</u> are visible.

To open this panel, do any of the following:

- Click on the Features tab (Features).
- Use the View > Features command.
- Right-click within a view and choose Show Features Panel.



The **Type** column shows the feature type, while the **Count** column shows the number of each feature type in the session. **Count** includes all sequences, whether selected or unselected. By default, all feature types except **Source** are checked.

To display or hide a particular feature type:

Check or uncheck its box in the **Type** column.

To display or hide all features:

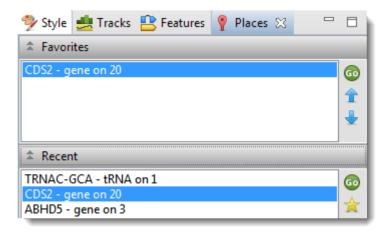
Right-click anywhere in the Features panel and select Show All or Hide All.

Places panel

The Places panel and the GenVision Pro <u>navigation tools</u> work together so you can quickly navigate to a chromosome or sequence range that you have visited before.

To open this panel, do either of the following:

- Click on the Places tab (Places).
- Use the View > Places > Places command.



This panel consists of two or three expandable sections:

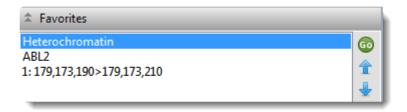
- Recent lists all locations that you visited after performing a search using GenVision Pro's navigation tools.
- <u>Received</u> lists data exported to GenVision Pro from ArrayStar (e.g. from ArrayStar's Gene or SNP Tables) using the File > Send Selection To > GenVision Pro command. If the GenVision Pro session does not contain items exported from ArrayStar, the Received section is disabled.
- <u>Favorites</u> is similar to the bookmarks folder for a web browser. It contains a list of previously-visited locations that you have decided you may wish to visit again.

The following video shows the Places panel in action in another DNASTAR application, MegAlign Pro:

Favorites section

The Favorites section is similar to the bookmarks folder for a web browser. It contains a list of previously-visited locations that you have decided you may wish to visit again. Favorites are listed in the same order in which they were added, by default, though you can change the order manually (see table below). The Favorites list is saved whenever you save a session.

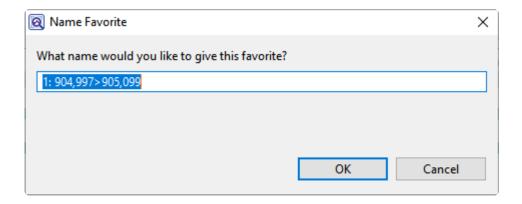
Access this section by clicking on the Places panel expand bar entitled **Favorites** or by choosing **View > Places > Favorites**.



To nominate an entry to the list:

Do any of the following:

- Use the <u>Recent</u> or <u>Received</u> sections of the <u>Places panel</u>. Click the "Recent" and "Received" links to learn more.
- Choose Analysis > Add Place to Favorites or right-click within the Analysis view and choose Add Place to Favorites. The following popup appears.



The popup shows the selected coordinates. If you would like to give this location a name, you can type it over the coordinates. It's also fine to keep the coordinates themselves as the name. Press **OK**.

The following table shows tasks that can be done from within the Favorites section:

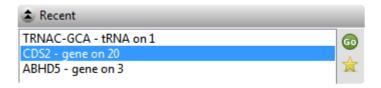
Task

To return to a location in this section	Double-click on the item or make a selection from the list and press the Go to this gene or location button ().
To move a bookmarked location further up the Favorites list	Make a selection from the list and press the Move this place up arrow ().
To move a bookmarked location further down the Favorites list	Make a selection from the list and press the Move this place down arrow ().
To rename a selected item	Click on an item in the Favorites section to select it. Then right-click on the same item and choose Rename Favorite . Type a name into the box and click OK .
To remove a selected item from the Favorites list	Click on an item in the Favorites section to select it. Then right-click on the same item and choose Remove Favorite from This List .
To remove all items from the list	Right-click in the Favorites section and select Clear Favorites .

Recent section

When you perform a search for a gene, feature or numerical sequence location using the <u>navigation tools</u> or the **Edit > Go to Place** command, each found location is automatically stored in the Recent section of the <u>Places panel</u>.

Access this section by clicking on the Places panel expand bar entitled **Recent** or by choosing **View > Places > Recent**.



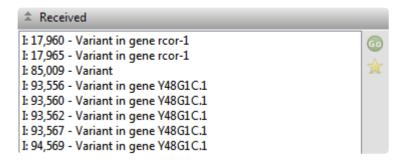
Recent locations are listed in the same order in which they were added, and the list is saved whenever you save a session.

- To return to a location on the list, make a selection from the list and press the **Go to this gene or** location button (©).
- To save a location on the list to the <u>Favorites section</u> (similar to bookmarking a page on the Internet), make a selection from the list and press the **Add this place to Favorites** button (). Alternatively, make a selection from the list, then right-click on the item and choose **Add This Place to Favorites**.
- To remove all items from the Recent list, right-click in the Recent section and select Clear Recent.

Received section

After opening a SeqMan NGen assembly in DNASTAR's ArrayStar application, you can send up to 100 data rows from its Gene Table or SNP Table directly to GenVision Pro using the **File > Send Selection To > GenVision Pro** command. When you do this, each row exported from the ArrayStar table is automatically stored as a "bookmark" in the Received section of the <u>Places panel</u>.

Access this section by clicking on the Places panel expand bar entitled **Received** or by choosing **View > Places > Received**.

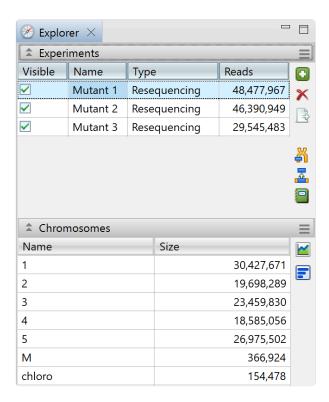


Received locations are listed in the same order in which they were added, and the list is saved whenever you save a session.

- To return to a location on the list, make a selection from the list and press the **Go to this gene or** location button ().
- To save a location on the list to the <u>Favorites section</u> (similar to bookmarking a page on the Internet, make a selection from the list and press the **Add this place to Favorites** button (). Alternatively, make a selection from the list, then right-click on the item and choose **Add This Place to Favorites**.
- To remove all items from the Received list, right-click in the Recent section and select Clear Received.

Explorer panel

To open the Explorer panel, click on the **Explorer** tab (or use the *View > Explorer * commands.



The Experiments section shows a list of the assemblies or other data sources loaded into the current GenVision Pro session and allows you to add additional files or delete existing ones. Some of the data added here as part of an .assembly file is parsed as <u>tracks</u> that you can display in the <u>Analysis view</u> and/or <u>Overview</u>.

Column	Description
Visible	Check the box to display the experiment in the views or uncheck the box to hide it.
Name	The name of the assembly or other data source.
Туре	The sample type. If you open an older assembly or session, or if GenVision Pro cannot determine the type, this will appear as Misc .
Reads	The number of reads in the assembly or chromosome.

The following table shows how to use the tools in the Experiments section.

Task	How to
Add a new .assembly to the session	Press the Add experiments tool (). To add another data type,

	see Add Data.
Delete one or more data sources from the session	Select the table rows you wish to remove and press the Remove this data source from session tool (X).
Export the selected data source	Select a single row in the table and press the Export Selected Assembly tool ().
Display the selected data source in the Variants view	Select a single row in the table and click the Show Variants tool (
Display the selected data sources in the Structural Variation view	Select one or more rows in the table and click the Show Structural Variation tool ().
Display the selected data sources in the Reports view	Select one or more rows in the table and click the Show Reports tool ().

The Chromosomes section shows a table of chromosomes in the currently-selected assembly. This section allows you to easily maneuver between chromosomes in a multiple sample project.

Column	Description
Name	The name or number of the chromosome.
Size	The chromosome length in base pairs.

The following table shows how to use the tools in the Chromosomes section.

Task	How to
Display the <u>Overview</u> and <u>Analysis view</u> for one selected chromosome	Double-click on the table row.
Display only the <u>Analysis view</u> for a selected chromosome	Select a row in the table and click on the Show analysis view of selected chromosome tool (). This command can be used in tandem with the command in the row below to show the Analysis view for one chromosome and the Overview for a different chromosome.
Display only the Overview for a selected chromosome	Select a row in the table and click on the Show overview of selected chromosome tool (). This command can be used in tandem with the command in the row above to show the Analysis view for one chromosome and the Overview for a different chromosome.

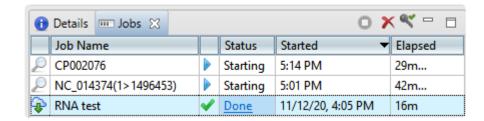
Jobs panel

After initiating an NCBI <u>BLAST</u> or <u>Entrez</u> search, you can monitor its progress and download its eventual output from the Jobs panel. This panel is not visible in a project until you have started running at least one search.

To open the Jobs panel:

Do any of the following:

- Click on the Jobs tab.
- Press Ctrl+Alt+J (Win) or Option+Cmd+J (Mac).
- Use View > Jobs or Search > Show Jobs.



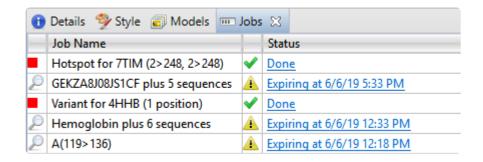
The Jobs panel consists of a header and a table with the following columns:

Column	Description	
Job Name	<u>Text searches</u> are all named with the field and term of the first search criterion followed by "plus # terms," where # represents the number of additional criteria. <u>Sequence searches</u> use the name specified in the <u>Job screen</u> of the Search wizard. The same Job Name can be used for more than one prediction.	
Status	A word and icon indicating the status of the search. • Started – Run or Run Job were pressed in the Search wizard • Queued – ▶ – Submitted and placed into the queue • Running – ▶ – Running • Done – ✓ – Job complete; results available • Failed – ※ – Job failed • Stopped – ※ – Job stopped by user	
Started	The date and time the job began.	
Elapsed	The total elapsed time between starting and completing the job.	

Tasks that pertain to the Jobs panel:

Task	How To
To open the results of one or more jobs	Click on the Done (Done) or Expiring links for that job. See below this table for more information on job expiration. You can also open a saved .search file using Search > Open Search File .
To log into Cloud Assemblies (allows you to see all jobs associated with your username)	Use the Not logged into Assemblies on the Cloud tool (\$\struct\$). If you are already logged in, the tool changes to have a green checkmark (\$\struct\$).
To stop one or more jobs that are in progress	Select one or more rows with a Status of Started , Queued or Running . Then press the Stop tool () in the header above the table. Once stopped, a job cannot be restarted from the Jobs panel. You must instead <u>initiate a new search</u> .
To delete a finished job	Select one or more rows with a Status of Done , Failed or Stopped . Then press the Delete tool (\nearrow) in the header above the table.
Top open/reopen the progress dialog for a job	Select one or more rows with a Status of Started , Queued or Running . Then right-click and choose Open Progress dialog .

Sequence (BLAST) and text (Entrez) search jobs are intended to be ephemeral and expire after 7 days. The Jobs panel notifies you of upcoming expiration dates.



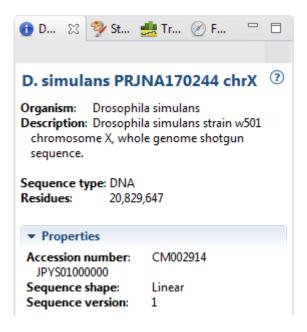
All jobs are recorded in the location specified in the <u>Temporary File preferences</u> dialog. If you are a Macintosh user and elect to keep the default save location, the operating system may remove jobs after a reboot. To avoid this, we recommend choosing a different location in which to save temporary files.

Details panel

The Details panel shows details of the application, selection or session. Details are available for all <u>tracks</u> in the <u>Overview</u> and <u>Analysis view</u>, as well as the <u>Tracks panel</u>.

To open this panel, do any of the following:

- Click on the **Details** tab (10).
- Use the View > Details command.
- Right-click within a view and choose Show Details Panel.



The appearance of this panel varies according to the situation:

- When no session, or a blank session, is open
- When a non-empty session is open
- When one sequence is selected
- · When a portion of a sequence is selected
- · When multiple sequences are selected
- · When one track is selected
- · When multiple tracks are selected
- · When one feature is selected
- · When multiple features are selected

Text in the Details panel can be copied using <u>basic editing commands</u> and pasted into a text editor, etc.

When no session, or a blank session, is open

When no <u>session</u>, or only a blank session, is open, the <u>Details panel</u> displays links to help you get started in GenVision Pro.

	Getting Started in GenVision Pro	Click on the links to:
Heading	Create a new project Open a file	* Create or open a session
	Add assemblies	* Add assemblies to the session
	Watch a training video	* Watch a <u>training video</u>

When a non-empty session is open

If a <u>session</u> is open and <u>data has been added</u>, the <u>Details panel</u> contains the following information:

Heading	CEU gene panel assembly.gvp Sequence type: DNA	The session file name, in blue.
Sequences	▼ Sequences Sequences: 26 Sequence type: DNA Min. sequence length: 16,569 Max. sequence length: 249,250,621	* Sequences – Number of sequences in the assembly. * Sequence type – Type of sequences in the assembly (DNA or RNA). * Min. and Max. sequence lengths – Lengths of the shortest and longest sequences.
Actions	▼ Actions Add assemblies Add tracks	* Add assemblies – Shortcut to menu command File > Add Assemblies. * Add tracks – Shortcut to menu command File > Add Track.

When one sequence is selected

If a single sequence has been selected in any view, the <u>Details panel</u> may contain a subset of the following information:

The label used in the Analysis view and Overview, in blue. * Organism – The organism containing the chromosome, if ? available. IV Organism: Caenorhabditis elegans * **Description** – A description of the Description: Caenorhabditis elegans Bristol N2 chromosome IV, complete sequence. Heading selected chromosome, if available. Sequence type: DNA * Sequence type – Type of Residues: 17,493,829 sequences in the assembly (DNA or RNA). * Nucleotides/Residues – Number of nucleotides or residues in the selected chromosome. Note that the **Properties** for a sequence show information about a single assembly (not all assemblies) loaded into the session. The information here pertains to the first Properties loaded assembly that covered the NC 003282 Accession number: chromosome of interest. Assembly name: Celegans Illumina 14-0-indexed Assembly path: C:\Users\Public\Documents\DNASTAR * Accession number – The NCBI \Sample Data\Demo SegMan NGen\Celegans_Illumina_14accession number of the selected 0-indexed.assembly **Properties** Assembly time: 7/7/16 10:50 AM sequence. Whole Genome Assembly type: Sequence shape: Unknown Sequence version: * Assembly name – The name of the Template name: NC_003282 assembly where the sequence or Template short name: IV chromosome is located. * **Assembly path** – The file path to that assembly. * **Assembly time** – The time the

SeqMan NGen assembly finished. * **Assembly type** – The type of assembly performed in SegMan NGen, if available (e.g., Exome/Gene Panel). * Sequence shape – Linear or circular, if specified in the reference sequence. * Sequence version – The NCBI version number (the part following the decimal after the accession number) for the sequences in the chromosome or contig. * **Template name** – SeqMan NGen's name for the assembly template, usually the same as its NCBI accession number. Example: In the image shown at left, the Template name is the accession number of the human 'X' chromosome. * **Template short name** – May match the label displayed in the Overview or the default name originating from the Genome Template Package used to create the assembly. * **Data specification** – The name of ▼ Data the sequence file, including extension. specification: CEU_Gene_Panel_12_2_release-22.template-seq Data file time: 2/5/15 3:53 PM Data * **Data file time** – The time the sequence file was created or modified. path: C:\Users\yildizs\Desktop\CEU_Gene_Panel_12_2_relea se.assembly\CEU_Gene_Panel_12_2_release-22.template-seq * **Data file path** – The path to the sequence file in the assembly. ▼ Actions * Rename – Shortcut to **Edit** > **Actions** Rename... Rename. Rename with Fields...

	* Rename with Fields – Shortcut to Edit > Rename with Fields.
--	---

When a portion of a sequence is selected

When a portion of a sequence has been <u>selected</u> in the <u>Analysis view</u> (e.g., by double-clicking on a feature, or by dragging across the sequence with the mouse), the <u>Details panel</u> contains the following information:

Heading	Selected Regions Sequence:1 Residues: 17,230,913 Sequence selection: 8>17230891, 17230910>17230938	* Sequence – The name of the sequence from which the region was selected. * Residues – The total number of residues in the selection. * Sequence selection – The inclusive coordinates for the beginning and end of each selection.
Actions	▼ Actions Copy Copy as FASTA	* Copy – Shortcut to the menu command Edit > Copy. * Copy As FASTA – Shortcut to the menu command Edit > Copy As FASTA. See Copy for more information about these commands.

When multiple sequences are selected

If multiple sequences have been <u>selected</u> in the <u>Genome view</u>, the <u>Details panel</u> contains the following information:

Heading	2 sequences Sequence type: DNA Min. sequence length: 198,022,430 Max. sequence length: 243,199,373	* Number of sequences currently selected, in blue. * Sequence type – DNA or RNA (if known). * Min./Max. sequence length – Lengths of the shortest and longest sequences in the current selection.
Actions	▼ Actions Rename with Fields	* Rename with Fields – Shortcut to Edit > Rename with Fields command.

When one track is selected

If a single <u>track</u> has been <u>selected</u> in the <u>Overview</u> or <u>Analysis view</u>, the <u>Details panel</u> contains the following information.

Heading	CEU Gene Panel 12 2 release Summary: This track displays a value at each position in the sequence.	 The heading contains: Type of track or assembly name, in blue. Summary – Basic information about what is displayed in the selected track.
Data	■ Data Assembly name: CEU Gene Panel 12 2 release Assembly type: Unknown Assembly time: 2/5/15 3:56 PM Assembly path: C:\Users\Public\Documents\DNASTAR\Sample Data\Demo SeqMan NGen\CEU_Gene_Panel_12_2_release.assembly	 Some tracks may have a Data section that includes the source of data. Assembly name – The name of the assembly where the sequence or chromosome is located. Assembly type – The type of assembly performed in SeqMan NGen, if available (e.g., Exome/Gene Panel). Assembly time – Date and time the SeqMan NGen assembly finished. Assembly path – The file path to that assembly.
Actions	▼ Actions Rename Rename with Fields Show track layout Show track options	Only a subset of links may be available, depending on the selected track. • Rename – Shortcut to Edit > Rename. • Rename with Fields – Shortcut to Edit > Rename with Fields. • Show track layout – Shortcut to View > Tracks > Layout. • Show track options – Shortcut to View > Tracks > Options.

When multiple tracks are selected

If two or more <u>tracks</u> have been <u>selected</u> in the <u>Overview</u> or <u>Analysis view</u>, the <u>Details panel</u> contains the following information.

Heading	2 tracks Summary: Multiple tracks are selected. They can be moved or hidden together.	 The heading contains: The number of tracks currently selected (in blue). Summary – Description of options available for the multiple track selection.
Actions	▼ Actions Rename Rename with Fields	 Rename – Shortcut to Edit > Rename. Rename with Fields – Shortcut to <u>Edit > Rename with Field</u>.

When one feature is selected

If a single feature has been selected in a track in the <u>Overview</u> or <u>Analysis view</u> (e.g., by clicking one time on the feature), the <u>Details panel</u> contains the following information. Note that the tooltip displayed when you hover over a feature is a subset of the Details panel information for that feature.

Heading	galK Type: CDS Left: 788,054 Right: 789,202 Length: 1,149	 The heading contains: The feature name, in blue. Type – Feature type. Left and Right – The inclusive coordinates of the beginning and end of the feature range. Length – Length of the selected feature in nucleotides or residues.
Location	▼ Location 788054<789202	The location of the feature, as a set of one or more segments.
Properties	product: galactokinase gene: galK gene_synonym: ECK0746; galA; JW0740 function: enzyme; Degradation of small molecules: Carbon compounds locus_tag: b0757 protein_id: NP_415278.1 db_xref: GI:16128725 db_xref: ASAP:ABE-0002568 db_xref: UniProtKB/Swiss-Prot:P0A6T3 db_xref: ECOCYC:EG10363 db_xref: GeneID:945358 codon_start: 1 EC_number: 2.7.1.6 GO_component: GO:0005737 - cytoplasm GO_function: GO:0042280 - cell surface antigen activity, host-interacting GO_process: GO:0016052 - carbohydrate catabolic process; GO:0009242 - colanic acid biosynthetic process; GO:0009269 - response to desiccation parentID: b0757 transl_table: 11	This section shows properties of the selected feature. The properties included vary based on the type of feature and the annotations present in the original data file.

Translation	▼ Translation MSLKEKTQSLFANAFGYPATHTIQAPGRVNLIGEHTD YNDGFVLPCAIDYQTVISCAPRDDRKVRVMAADYE NQLDEFSLDAPIVAHENYQWANYVRGVVKHLQLR NNSFGGVDMVISGNVPQGAGLSSSASLEVAVGTVL QQLYHLPLDGAQIALNGQEAENQFVGCNCGIMDQ LISALGKKDHALLIDCRSLGTKAVSMPKGVAVVIINS NFKRTLVGSEYNTRREQCETGARFFQQPALRDVTIE EFNAVAHELDPIVAKRVRHILTENARTVEAASALEQ GDLKRMGELMAESHASMRDDFEITVPQIDTLVEIVK AVIGDKGGVRMTGGGFGGCIVALIPEELVPAVQQAV AEQYEAKTGIKETFYVCKPSQGAGQC	This section is displayed only for some CDS features. When present, it displays the NCBI /translate feature qualifier.
Notes	▼ Notes note: ELKS/RAB6-interacting/CAST family member 2; Derived by automated computational analysis using gene prediction method: BestRefSeq, Gnomon.	 This section is not available for all features. note – The NCBI feature qualifier of the same name.
Actions	▼ Actions Select Subsequence You can also double-click on this feature.	Select Subsequence – * Shortcut to the menu command topic=Menu Commands *Edit > Select Subsequence. Double-clicking on the feature fulfills the same function. See Make a Selection for more information.

For additional information about working with features, see <u>Features panel</u>.

When multiple features are selected

If two or more features have been selected in a track in the <u>Overview</u> or <u>Analysis view</u>, the <u>Details panel</u> contains the following information:

Heading	3 features Type: Region, source, Protein Left: 1 Right: 374 Min. Length: 368 Max. Length: 374	 Number of features currently selected, in blue. Type – List of all feature types represented in the selection. Left and Right – The inclusive coordinates of the beginning and end of the feature range. Min. and Max. Length – Lengths of the shortest and longest features in the selection.
Properties	▼ Properties gene: PRR23B db_xref: HGNC:33764, GeneID:389151,	Shows a subset of the qualifiers that all selected features have in common. This section does not show qualifiers that appear elsewhere in in the Details panel when a single feature is selected (e.g., /translation and /note)

Button tools

At the top of the GenVision Pro window are eight button-style tools, as well as a set of <u>navigation tools</u>. In the bottom right corner of each view, GenVision Pro offers two additional button-style tools for selecting the screen mode.

You can toggle between showing/hiding the toolbar using the command **View > Show/Hide Toolbar**.

The following table describes the ten button-style tools.

Name	Tool	Keyboard/Mouse Option	Description
Create a new session		Shortcut to the File > New GenVision Pro Session command. See <u>Create</u> <u>a new session</u> for details.	
New assembly with options		Shortcut to the File > New Assembly with Options command. See <u>Create</u> a new assembly using <u>SeqMan NGen</u> for details.	
Open a file		Shortcut to the File > Open command. See <u>Open an existing session</u> for details.	
Save session		Along with Ctrl/Cmd+S , this is a shortcut to the File > Save Session command. See <u>Save</u> , <u>Close and Exit</u> for details.	
Add assemblies to session	=	Shortcut to the File > Add Assemblies command. See <u>Add assemblies</u> for details.	
Add sequences to session	ACC +	Shortcut to the File > Add Sequences command. See <u>Add sequences</u> for details.	
Download and add a genome to this session	\	Shortcut to the File > Download and Add Genome command. See <u>Add genomes from the NCBI</u> for details.	
Add tracks to session	+	Shortcut to the File > Add Track command. See <u>Add tracks</u> for details.	
Show whole genome		Same as opening the <u>Genome view</u> or choosing (Genome) from the <u>navigation tools</u> drop-down menu. All available sequences (i.e., the whole genome) will be displayed in the Overview. This is default state of the	

	Overview when a selection has not yet been made.
Full screen	Shortcut to the View > Full Screen command. Undo by clicking the Esc key or the Exit Full Screen icon () at the bottom of the full-screen view.
Change layout	Lets you choose how many views are displayed and how they are laid out. See <u>Customize the appearance and layout of views</u> for details.

Navigation tools

GenVision Pro's navigation tools are located just to the right of the button tools and consist of a "sequence selector" (the drop-down menu), a text box, a **Go** button, and two green arrows. By default, **(Genome)** is shown in the sequence selector, causing the entire genome to be displayed in the <u>Genome view</u>. The <u>Analysis view</u> and <u>Overview</u> remain empty until a different selection is made.

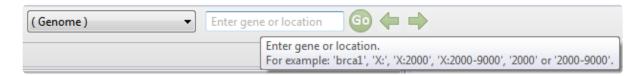
To move the cursor to a particular position or range using the search bar at the top of the GenVision Prowindow:

Use the search bar in the header of the GenVision Pro window, just to the left of a round "Go" button.

See the table below for syntax examples, and the contexts in which they may be used. If you are currently displaying all chromosomes (see above), you may limit the search to a particular one by including it in the query syntax. GenVision Pro will automatically update the drop-down menu to correspond to the chromosome specified by the query.

Context	Туре	To go to
	chr1:	chromosome 1
All contexts	3:	chromosome 3
	chr1:42	position 42 on chromosome 1
	chr2:300-433	range 300-433 on chromosome 2
When viewing a chromocome	451	position 451 of current chromosome
When viewing a chromosome	23-4343	range 23-4343 of the current chromosome
When viewing a genome	3	chromosome or contig 3

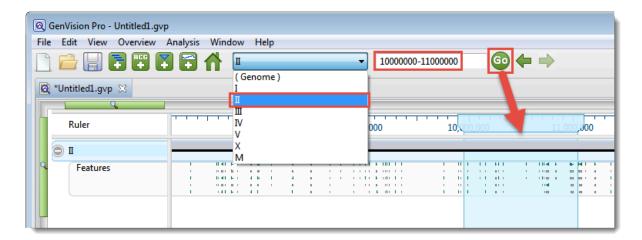
If you forget how to enter a query, hover over the sequence selector or text box to reveal a list of sample queries.



A variety of separators (e.g., spaces, < or > symbols) are accepted between the endpoints in a range. For example, all of the following are accepted: 300-433, 300 433, 300.433, 300>43 and 433<300. You can also paste in a range obtained from a variety of external sources, including SeqNinja and the <u>National Center for Biotechnology Information</u> (NCBI).

Example: The following image shows an example where both the sequence selector and the text box were

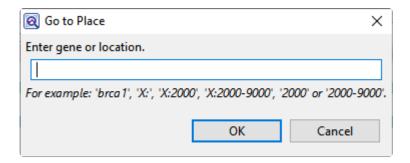
used to perform a search.



After pressing the **Go** button, the location will automatically be saved to the <u>Recent section</u> of the <u>Places</u> <u>panel</u>. This section is similar to the "recent items" cache in other computing contexts, and allows you to return to the same location in the future.

To move the cursor to a particular position or range using a menu command:

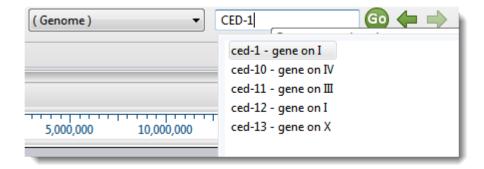
Choose **Edit > Go to Place** to open up a popup dialog with a similar search bar.



In the popup dialog, type a position or range into the text box (refer to the section above for details) and press **OK**.

To move the cursor to a particular feature:

Type the feature name into the text box. As you begin typing, possible matches will appear in a list below. To pick one of them, use the up/down arrow keys or the mouse. Press the return key or **Go to gene or**location specified button (Go) to go to the selected item. You can also go to an item by double-clicking on it in the list.



After pressing the **Go** button, the location will automatically be saved to the <u>Recent section</u> of the <u>Places</u> <u>panel</u>. This section is similar to the "recent items" cache in other computing contexts, and allows you to return to the same location in the future.

To go back to the last recent place:

Click the left arrow (). The arrow is only enabled if another location was visited prior to the current location. Note that more recently-visited places are higher up on the list.

To go forward to the next recent place:

Click on the right arrow (). The arrow is only enabled if you returned to this location after first visiting another location.

Create or Open a Session

Because GenVision Pro does not store data when you close it, the information shown in the interface is called a "session," rather than a "project." Once you open GenVision Pro, you can open one or many sessions. Each session is displayed in a separate tab and is initially assigned a unique name (Untitled[#].gvp).

For more information, see the following topics:

- Create a new session
- Open an existing session
 - Change the default application for opening files
- Launch a session from within ArrayStar
- Make a session active

Create a new session

To create a new, untitled GenVision Pro session:

Do any of the following:

- From the <u>Welcome screen</u>, choose **New session** to open a new blank session; or **Add assemblies to new session** to open a session and add sequences at the same time.
- From the Details panel, click the link Create a new session.
- From the menu, choose File > New GenVision Pro Session.
- Press Ctrl/Cmd+N.

Once you create a session and <u>add sequences</u> or <u>add assemblies</u>, the information is initially shown in the <u>Genome view</u>. Double-click on a chromosome in that view to open its <u>Overview</u> and <u>Analysis view</u>. The appearances of the views are determined by options in the <u>Style panel</u>.

To open an additional GenVision Pro window:

Do either of the following:

- Click on the **Welcome** tab and choose **New session**.
- Use the Window > New Window command.

Open an existing session

To open an existing GenVision Pro (.gvp) session :

Do any of the following:

- From the Welcome screen, click the **Open** tab, then click on **Open session**.
- From the Details panel, click the link **Open a session**.
- Use the File > Open command.
- Press Ctrl+O (Win) or Cmd+O (Mac).

Once a project is open in GenVision Pro, you can open additional sessions using the same command.

Notes:

- If you created a session in GenVision Pro 17.2.1 or earlier and attempt to open it with version 18.0 or later, read tracks will not be available in the <u>Tracks panel</u>, and a warning may appear. In this case, we recommend closing the session and <u>creating a new session</u> instead.
- If you moved or renamed any of the assembly projects, the track file(s) or index file (if any) prior to reopening a session, you may be prompted to supply the new name(s) or location(s) of those files.
- If you always want to open .gvp files using GenVision Pro, see Change the default application for opening files.

Change the default application for opening files

Windows 7 and Windows 8:

- 1. Open the Control Panel.
- 2. Go to Programs > Default Programs > Associate a file type or protocol with a specific program.
- 3. Select extension you wish to change in the **Extensions** list and then click the **Change program** button.
- 4. Click the **Browse** button in the Open With dialog box.
- Navigate to and select the application you want the files to open in. (Lasergene applications are located by default in C:\Program Files\DNASTAR\Lasergene 'x' or C:\Program Files (x86)\DNASTAR\Lasergene 'x', where 'x' represents the version number).
- 6. Click Open.
- 7. Click **OK** to close the Open With dialog box, and then click **Close**.

Macintosh:

- 1. Select a file of the type you wish to change in Finder.
- 2. Go to File > Get Info.
- 3. Click the dropdown arrow in the **Open With** field, and navigate to the application you want the files to open in. (Lasergene applications are located by default in **Applications:DNASTAR:Lasergene 'x'**, where 'x' represents the version number).
- 4. Click the **Change All** button to make the application you selected the default for all files of this type.

Create a new assembly using SeqMan NGen

To create a new assembly from within the GenVision Pro application, do either of the following:

- From the Welcome screen, click the New tab on the left and then choose New assembly with options on the right.
- From within a GenVision Pro session, click on the **New assembly with options** tool in the top left (
- Choose the menu command File > New Assembly with Options.

All of the above cause SeqMan NGen to launch at the Welcome screen. Choose the desired workflow from the Workflow screen and proceed through the wizard to set up the assembly. For more information, consult the <u>SeqMan NGen User Guide</u>.

The assembly status will appear in GenVision Pro's <u>Jobs panel</u>. Depending on the data type and workflow, the results may launch automatically in SeqMan Ultra once the assembly finishes. If they don't, you can open them in SeqMan Ultra by clicking the blue link in the **Status** column. To learn more about using SeqMan Ultra, consult the <u>SeqMan Ultra User Guide</u>.

Launch a session from within ArrayStar

To view and analyze a selected entry from ArrayStar:

- 1. Create or open a project in DNASTAR's ArrayStar application.
- Open any table and locate a row with a position of interest. Select the row and choose File > Send
 Selection To > GenVision Pro or right-click on the row and choose Send Selection To GenVision
 Pro. GenVision Pro will launch and load the corresponding contig(s) or chromosome(s). If your project
 contains multiple experiments, all experiments will be loaded into GenVision Pro.
- 3. (optional) To move to a new location on the sequence in GenVision Pro, repeat Step 2 using a different table row.

The following video shows this workflow in action:

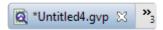
This next video shows an RNA-Seq workflow that begins with assembly in SeqMan NGen, proceeds to downstream analysis in ArrayStar, and then uses ArrayStar's **Send Selection to GenVision Pro** command to open the results in GenVision Pro. Once in GenVision Pro, the Sashimi track is applied. (GenVision Pro section begins at 2:41).

Make a session active

To make a session active:

Do any of the following:

- Click+ on any session tab to make that session active.
- Click on the Window menu and choose the name of the session from the list.
- Press **Ctrl+**; (Win) or **Cmd+**; (Mac) to open a "session chooser." You may also open the session chooser by pressing the >> icon that appears to the right of the session tabs. This icon is displayed when multiple sessions are open and the GenVision Pro window is too small to display each one in an individual tab.



The session chooser appears as a yellow popup with a line for typing in text at the top, and a list of open sessions below. Use your mouse to click on the session you want to make active. If you do not have a mouse, make a selection using the keyboard arrow keys, then press **Enter**.



To instead search for a session using a filter, type text into the text field and press **Enter**. The text field supports the "?" and "*" wildcards.

Add Data

Once you have created a GenVision Pro session, the next step is to add data.

To add this type of data	see this topic	Where data type appears in GenVision Pro	How to view the data
.assembly files created in SeqMan NGen	Add assemblies	Experiments section of the Explorer panel	
.seq files .sbd and .sbp files created in SeqBuilder Pro .msa files created in MegAlign Pro Reference chromosomes in .fasta, .fas, .gb, or .gbk format	Add sequences	Chromosomes section of the Explorer panel	Double-click on a row to make the selected experiment active in the Analysis view and Overview.
Genomes from NCBI	Add genomes from NCBI	Chromosomes section of the Explorer panel	
BED files (.bed) Wiggle files (.wig) GenVision box files (.txt)	Add tracks	Tracks panel	Check the box next to the track name to view this data in the <u>Analysis view</u> .

Add assemblies

Choose File > Add Assemblies

To add one or more SegMan NGen .assembly packages to the session:

Do any of the following:

- Drag & drop one or more assemblies from your computer's file explorer onto an open GenVision Pro session.



For all but the first option on the list above, you will be prompted to select an assembly (or folder of assemblies) from the file browser. After making a selection, press **OK**. Note that double-clicking is not supported in this situation.

Each assembly you add will be listed in the Experiments section of the <u>Explorer panel</u>. Double-click on the assembly to view it in the <u>Overview</u> and <u>Analysis view</u>.

Errors when adding assemblies:

The following table shows warning messages that you might encounter when adding an assembly to a session.

Warning	Explanation
Please pick an assembly or a directory containing assemblies.	The item selected was not an assembly project or a folder containing an assembly project.
Unable to add duplicate data source.	The selected assembly has already been added to the current session.
Most of the sequences being added from the assembly " <assembly name="">" do not match those in the document. Add data anyway?</assembly>	Most contigs or chromosomes in the selected assembly do not match those of assemblies already in the session. Click OK to add the assembly anyway, or Cancel to close the dialog without adding the assembly.
Feature index not found for assemblies. Index features now?	If this prompt appears, we recommend choosing Yes . Choosing Yes simply causes a "feature index" file to be quickly created and added to the original assembly file package. A detailed explanation of "feature index" appears below this table.

"Feature index" description:

During assembly, SeqMan NGen 14.0 (and later) automatically includes a features index file as part of the assembly package. GenVision Pro uses the file for rapid access to features. The presence of a features index allows you to navigate to a gene by name and to view features in the Features panel.

An assembly built in an older version of SeqMan NGen will initially lack this feature index. When you add such an assembly to a GenVision Pro session, this message asks whether you wish to have a feature index created automatically. If you click **Yes**, a file will be created and stored within the original assembly data. If you click **No**, no file will be created, and the assembly will be added to the session without indexing. This will cause loss of the functionality described in the paragraph above. If you decline to create the feature indices, you may be prompted again when the assembly project is used a different session.

If you add a set of assemblies built using equivalent reference templates, and respond **Yes** to this prompt, a feature index will be created for just one member of the set. In this case, the genome templates for other members of the set are not loaded into GenVision Pro. Note that if you later add one of the non-indexed assemblies to a different GenVision Pro session, you will again be prompted to index, as the new GenVision Pro session will no longer have access to the indexed assembly. Also, if you add assemblies that have additional chromosomes from those added earlier, one of the "expanded" assemblies will also need to be indexed.

Example: Say you start a GenVision Pro session by adding some human female assemblies from SeqMan NGen v. 13.0. You respond **Yes** to the indexing prompt, and a feature index is created for one of the assemblies. This index will lack feature indexing for the Y chromosome, since all samples so far are female. If you then add human male assemblies that have the additional chromosome Y, one of the male assemblies will need to be indexed.

Note: In order for GenVision Pro to add the index file, you must have "write" permission for the assembly package. If you do not, someone with Administrator status may need to do this step for you. Alternatively, you can copy to the assembly to a location where you have write permission.

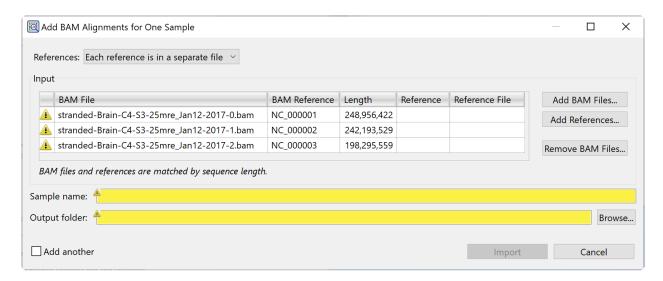
Add BAM files

If you create a templated assembly in SeqMan NGen, the .assembly package output contains the reference sequence(s) as well as feature information. By contrast, BAM files (the output for assemblies with Geneious, CLC Workbench and others) do not have either type of information. This means that when you add BAM files to a GenVision Pro session, you will always need to upload one or more reference sequences. The reference can be a genome template package downloaded from DNASTAR, or another reference of your choice. If you wish to upload features and your reference sequence files do not contain those, you will also need to upload a .gff file containing features.

In the instructions below, note that Step 3 requires you to make a choice based on the type of reference files you will be using. Each choice causes a different version of the dialog to appear, and each version is described separately in that step.

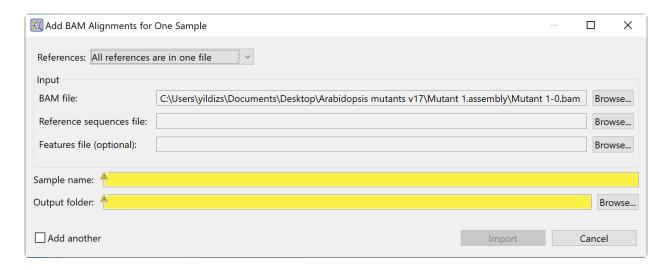
To add BAM files to the GenVision Pro session:

- 1. Do any of the following:
 - Drag & drop one or more .bam files from your computer's file explorer onto an open GenVision Pro session.
 - From the New tab on the Welcome screen, choose Add BAM alignments to new session.
 - Press the Add BAM alignments to session tool ().
 - Choose File > Add BAM Alignments.
- 2. A file chooser is displayed. Navigate to and select the desired .bam alignment(s), then press Open.
- 3. Use the **References** drop-down menu to choose either **Each reference is in a separate file** or **All** references are in one file.
 - If your references are a set of .gbk or .fasta (or .fas) files, often one file per chromosome, choose **Each reference is in a separate file**. The dialog appears as in the image below.



Start by importing references in the upper part of the dialog. Note that yellow "warning" triangles are shown to the left of any sequence that does not yet have an associated reference sequence. At first, all sequences will have the warning triangles. Start by importing references until all the warnings have disappeared. GenVision Pro automatically assigns references to their corresponding samples. If GenVision Pro detects that the references are from the wrong organism or otherwise incorrect, it will not assign them to any samples.

- To add one or more references sequences, press Add References. Once you have added references, this button changes to Replace References.
- To add additional .bam files, press Add BAM Files.
- To remove a file that is already in the Input table, select the sample row and press
 Remove BAM Files.
- If your reference consists of one .fasta (or .fas) file with a separate .gff file containing features, choose All references are in one file. The dialog appears as in the image below.



- To add a reference sequence, use the Browse button to the right of Reference sequences file.
- To add an optional feature file in .gff format, use the Browse button to the right of Features file (optional).
- To switch to a different BAM file, use the Browse button to the right of BAM file.
- 4. Name the sample by typing text into the Sample name field. This field is highlighted in yellow until you input a name.
- 5. When you import BAM files, GenVision Pro reprocesses them using the same XNG assembly engine used by SegMan NGen. In essence, GenVision Pro does a "mini" assembly. Therefore, you need to specify a folder for this assembly output. Choose the Output folder using the Browse button on the right. The **Output folder** field is highlighted in yellow until you choose a location.
- 6. Press Import to import the samples listed in the Input section above. If you wish to add another sample file in .bam format, check the box next to Add another before pressing Import. The additional sample must use the same reference sequence(s) that you already imported in the earlier steps above.

After pressing Import, the BAM file import will show up as an "in progress" job in the Jobs panel. Small jobs will take only a few minutes, while samples using the human genome as the reference can take about an hour to process. While the job is in progress, you will receive a warning if you attempt to close GenVision Pro. However, you can work on additional GenVision Pro sessions while your job is processing. You can also track the progress of the job in the Console view.



Note: As of Lasergene 17.5, GenVision Pro cannot currently call variants in BAM projects.

Add sequences

You can add a sequence to the session if it is in any of the following formats: sequence file (.seq), SeqBuilder Pro format (.sbd), MegAlign Pro format (.msa), Genbank (.gbk) or FASTA (.fas and _.fasta).

To add sequences:

Do any of the following:

- Drag & drop one or more sequence files from your computer's file explorer onto an open GenVision Pro session.
- Use the Add sequences to project button tool (
- Choose File > Add Sequences.

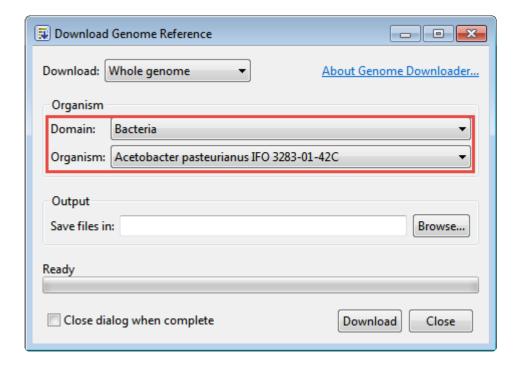
Each sequence you add will be listed in the Chromosomes section of the <u>Explorer panel</u>. Double-click on the sequence to view it in the <u>Overview</u> and <u>Analysis view</u>.

Note that you can add additional sequences at any time. However, once sequences are added, they cannot be deleted from the session. Note that it usually takes longer to add sequences to GenVision Pro than it does to add assemblies.

Add genomes from NCBI

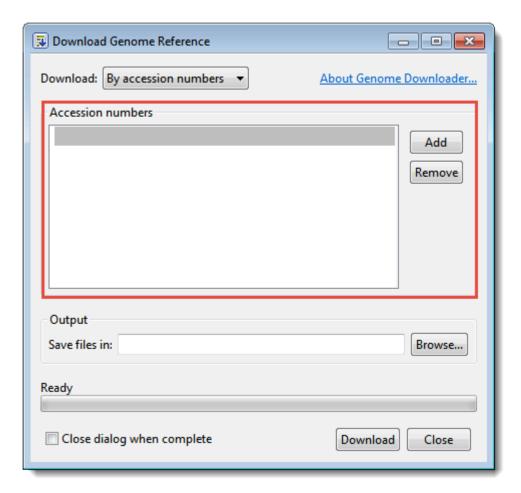
To import a genome from the <u>National Center for Biotechnology Information</u> (NCBI) database via organism name or accession number:

- Choose File > Download and Add Genome or press the Download and add a genome to the session <u>button to</u> (
- 2. In the Download Genome Reference dialog, use the **Download** dropdown menu to choose between downloading a **Whole genome** or **By accession numbers**.
 - If you select **Whole genome**, GenVision Pro will retrieve the most recent build of the selected genome. Use the next two drop-down menus to select the **Organism type** and **Organism**.

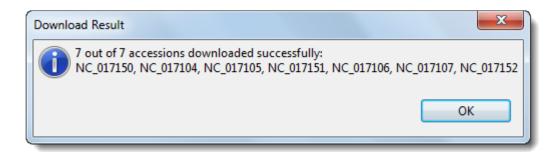


GenVision Pro will download all the reference sequences from the NCBI Entrez Genome Project database for the selected genome. These downloads may include auxiliary genomes such as mitochondria and chloroplasts. They may also include some contigs which have not yet been placed by the genome finishing process.

If you select By accession numbers, the Organism section disappears and is replaced by an
Accession numbers entry area. Type an accession number or paste it from your clipboard, then
press Add to add a number to the list. Continue adding numbers, as desired. Multiple accession
numbers should be separated using a space, comma, semi-colon or line break. To remove an
accession number from the list, select it and click Remove.



- 3. Click **Browse** to select a name and location in which to save the downloaded genome files.
- 4. If you do not need to download additional genomes, you can check **Close dialog when complete**. Otherwise, leave the box unchecked to keep the dialog open after initiating the current download.
- 5. Press **Download**. Once the download is complete, a message like the one below will appear.



6. Click **OK** to close the dialog and add the accessions to the GenVision Pro session. If you checked **Close dialog when complete**, the Download genome reference dialog will also close. Otherwise, it will remain open so that you can download additional NCBI genomes.

Downloaded genomes will be listed in the Chromosomes section of the <u>Explorer panel</u>. To learn how to view the genome in the <u>Overview</u> or <u>Analysis view</u>, see <u>Work with Tracks</u>.



Tip: Some large genomes may take extra time to download. In the meantime, you can work on a different GenVision Pro session by clicking on its tab.

Add tracks

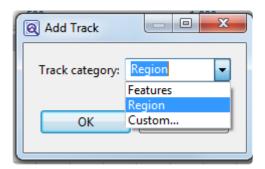
A data track in BED file (.bed), GenVision box file (.txt) or Wiggle (.wig) format can be imported into a GenVision session.

To add tracks:

- 1. Do any of the following:
 - Drag & drop one or more track files (see above) on an open GenVision Pro session
 - Use the Add tracks to project <u>button tool</u>



- Choose File > Add Track.
- 2. In the file chooser, navigate to and select the desired file. There are three allowable file types: BED (.bed), GenVision file (.txt; formatted as a histogram or arrow panel) and Wiggle file (.wig). For instance, you could use a targeted region BED file from a SeqMan NGen assembly, or a box file from DNASTAR's original GenVision application. Press **Open**.
- 3. In the ensuing Add Track dialog, use the drop-down menu to select the Track category for the file you chose in the previous step. Choices will include a subset of Feature, Region, Coverage or Custom. You can type a custom value directly into the field. Optionally, you can choose the Custom menu item to clear the field.



GenVision Pro will use the **Track category** to place the track in the hierarchy in the <u>Tracks panel</u> and determine default track options.

4. Press **OK** to add the track to the current session.

To learn how to view track data in the Overview or Analysis view, see Work with Tracks.

The following table shows error messages that may occur when attempting to import track data:

Error Message	Explanation
Some data could not be mapped to document members.	This message appears if some of the sequences identified in the track data could not be matched automatically to sequences in the session.
This file is already loaded in this document.	The selected file has already been added to the current session.
Could not open file.	The file does not contain recognizable track data.

Search for Sequences Online

This application lets you search <u>National Center for Biotechnology Information</u> (NCBI) databases for matches to text or sequence that you specify.

- Query part or all of a nucleotide or protein sequence against one of NCBI's sequence databases using the <u>Basic Alignment Search Tool (BLAST)</u>.
- Query text or accession numbers against entries NCBI's Entrez database.

Once a search has been initiated, you can use the <u>Jobs panel</u> to monitor the search, stop an in-progress search, or open search results in the <u>Text</u> and <u>Table</u> views.

Tools available in the Search wizards:

Some or all of the following optional tools may be available in the bottom left corner of the Search wizards:

- 1. To display the help for this topic, press the question mark icon (②).
- 2. To enter an NCBI Access ID, press the **Set Access ID** (tool. This causes the Access ID dialog to pop up. See <u>Access ID dialog</u> for information.
- 3. To submit an idea for a new feature, press the **Suggest a feature** () tool.
- 4. To view information about your service plan, press the **Show service plan details** () tool. For sequence or text search services, the ensuing dialog describes the state of your service plan and any warning or error text messages. The license control has three states: **Active**, **Expiring**, and **Expired**.

Search for sequence matches (BLAST)

The BLAST Search wizard is used to <u>search</u> a query sequence against one of NCBI's databases. The sequence can be part or all of an active sequence, another sequence file on your computer, or typed/pasted-in text. The wizard consists of four consecutive screens, each described in a separate topic:

- Query screen
- Search screen
- · Options screen
- Job screen

To launch the BLAST Search wizard, use one of the following commands or tools. Some commands let you launch the wizard with the workflow pre-selected and the Search list already pre-populated.

Command, shortcut or tool	Description	
Search > Search		
Use the "Change search strategy" tool () from the Text, Table or Pairwise views	Launches the wizard at the Query screen.	
From the Welcome tab, choose Search and then Search for sequences .		
Search > Sequence Search with Selection	Launches the wizard at the Query screen, whose table is pre-	
Press Ctrl/Cmd+B	populated with the currently-selected portion of sequence.	
Search > Sequence Search with All Sequences	Launches the wizard at the Query screen, whose table is pre- populated with all sequences in the document.	

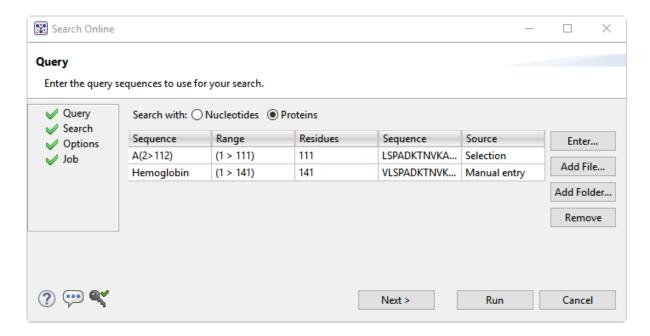
Search wizard: Query

The Query screen is the first of four screens in the <u>BLAST Search wizard</u>. The Query screen is where you add the query sequence, i.e., the sequence for which you wish to locate matches in the NCBI BLAST database.

At the top of the screen, next to "Search with," select the **Nucleotides** or **Proteins** button.

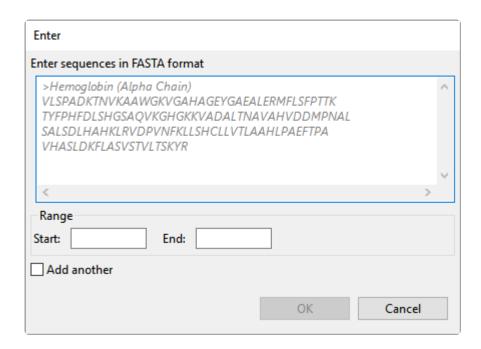
To add one or more query sequence files or a folder of sequences:

Press the **Add** or **Add Folder** buttons. In the file explorer, navigate to and select the desired file(s)/folder(s), and then click **Open**. If you add a multi-file sequence or multiple sequences, each sequence will appear on a separate row in the table.



To paste or type in the query sequence:

Press **Enter**.



In the Enter dialog:

- 1. Type or paste in the desired sequence following the "dimmed" example in the dialog. A header line is optional.
- 2. (optional) If you wish to use only a range of this sequence, specify the **Start** and **End** positions.
- (optional) If you want to add another query sequence after this one, check the Add another box. After clicking OK to the first sequence, a new Enter dialog will open to add the next query. Otherwise, leave this box unchecked.
- 4. Press **OK**. If you add additional sequences, each sequence will appear on a separate row in the table.

To remove a file that has already been added:

Select the unwanted file from the list and click Remove.

To learn about the optional tools in the bottom left corner of the screen, see the bottom of this topic.

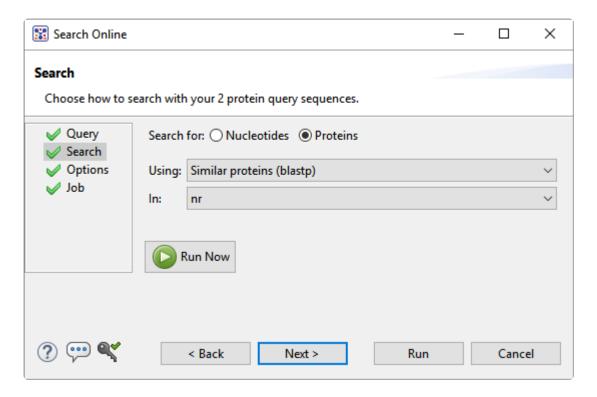
Once you have set up the query sequence(s) as desired, do either of the following:

- To choose the BLAST program to use and database to search, press Next to continue to the <u>Search</u> screen.
- To run the search directly, press Run.

• To close the Search wizard without starting a search, press **Cancel**.

Search wizard: Search

The Search screen is the second of four screens in the <u>BLAST Search wizard</u>. The Search screen is used to choose the BLAST program and database.



- Specify whether you wish to search for Nucleotides or Proteins.
- Select the desired BLAST program from the **Using** drop-down menu. For information on the different BLAST programs, see the <u>BLAST Program Selection Guide PDF</u>. Available options depend on the combination of the selections you made in **Search with** (<u>Query screen</u>) and **Search for** (this screen).
 - If you selected **Proteins** on both screens, the **Using** menu contains the unchangeable default of **Similar proteins** (blastp).
 - If you selected Nucleotides on both screens, Choices are Highly similar sequences (megablast), Similar nucleotides (blastn), and Similar translated nucleotides (tblastx).
 - If you selected Nucleotides in the Query screen and Proteins in the Search screen, the Using menu contains the unchangeable default of Similar proteins (blastx).
 - If you selected **Proteins** in the Query screen and **Nucleotides** in the Search screen, the **Using** menu contains the unchangeable default of **Similar translated nucleotides (tblastn)**.
- Select the desired BLAST database from the **In** drop-down menu. Available choices depend on whether you selected **Nucleotides** or **Proteins** in this screen.

To learn about the optional tools in the bottom left corner of the screen, see the bottom of this topic.

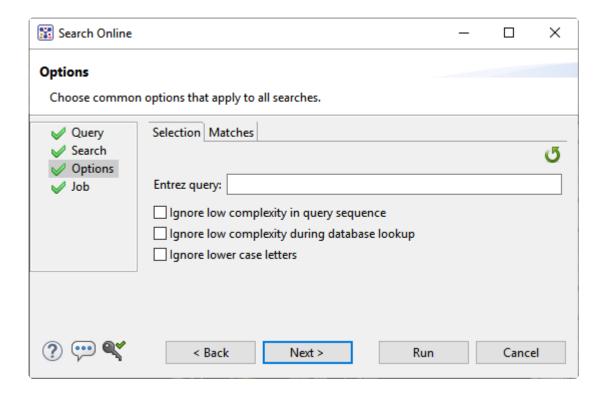
Once you have made the desired selections, do any of the following:

- To go back to the Query screen, press Back.
- To add additional parameters to the search, press **Next** to continue to the <u>Options screen</u>.
- To start the search immediately, press the **Run Now** button () or **Run** button. The <u>Jobs panel</u> in the application will open and contain information on the status of the search. During the search, a progress dialog is displayed where you can cancel the search using the **Stop** button, hide the dialog using the **Hide** button, or set up another query using the **Settings...** button. Use the <u>Jobs panel</u> to monitor an <u>online search</u>, stop an in-progress search, or open search results in the <u>Text</u> and <u>Table</u> views.
- To close the Search wizard without starting a search, press Cancel.

Search wizard: Options

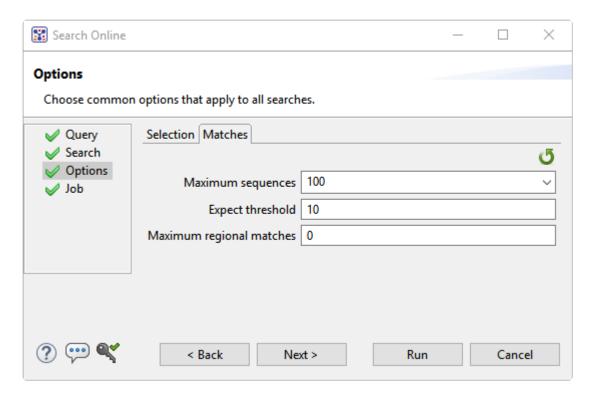
The Options screen is the third of four screens in the <u>BLAST Search wizard</u>, and has two tabs: **Selection** and **Matches**. These tabs allow you to customize miscellaneous parameters relating to the BLAST search.

Click the Selection tab to make it active.



In the **Selection** tab:

- To filter the database you're searching using a text term, enter a text string in the **Entrez** query field.
- To filter low-complexity regions in the query, check the box next to Ignore low complexity in query sequence. The blastn search program will use DUST for filtering; all other programs will use <u>SEG</u>.
- To mask the query while producing seeds used to scan database (but not for extensions), check the box by Ignore low complexity during database lookup.
- To mask any letters that were lower-case in the FASTA input, check the box next to Ignore
 lower case letters
- · Click the Matches tab to make it active.



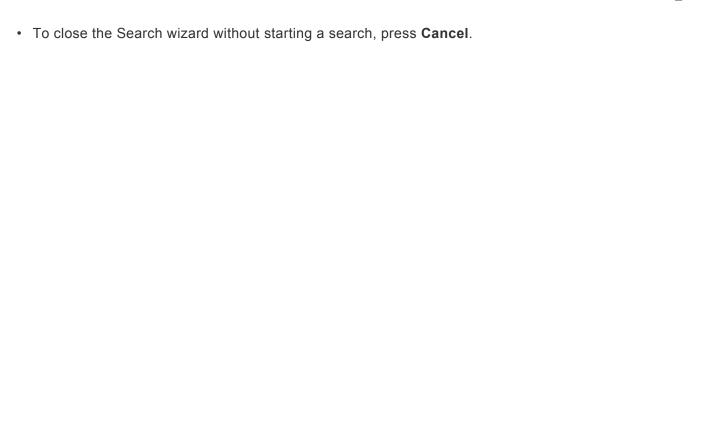
In the **Matches** tab:

- Enter a number next to Maximum sequences, the maximum number of aligned sequences to display. The actual number of alignments may be greater than this number. The default is 100.
- Enter a number beside Expect threshold, the expected number of chance matches in a random model. The default is 10.
- Enter a number next to Maximum regional matches to limit the number of matches to a query range. The default is 0.

To learn about the optional tools in the bottom left corner of the screen, see the bottom of this topic.

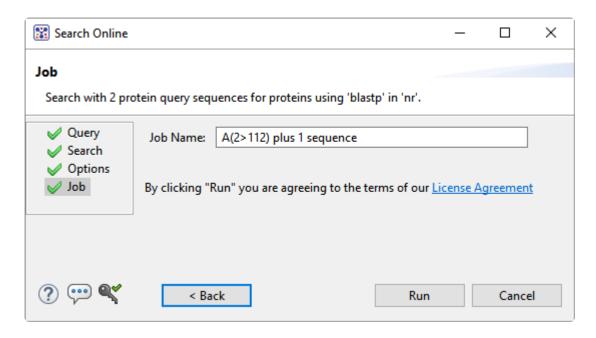
Once you have made the desired selections, do any of the following:

- To continue filtering, press Next to continue to the <u>Job screen</u>.
- To start the search immediately, press the **Run Now** button () or **Run** button. The <u>Jobs panel</u> in the application will open and contain information on the status of the search. During the search, a progress dialog is displayed where you can cancel the search using the **Stop** button, hide the dialog using the **Hide** button, or set up another query using the **Settings...** button. Use the <u>Jobs panel</u> to monitor an <u>online search</u>, stop an in-progress search, or open search results in the <u>Text</u>, <u>Table</u> and <u>Pairwise</u> views.



Search wizard: Job

The Job screen is the fourth of four screens in the <u>BLAST Search wizard</u>, This screen allows you to overwrite the default search name with a custom name and initiate the BLAST search.



The text box contains an automatically-generated name for the search job. If you wish, you can overwrite this default name with a custom name.

To learn about the optional tools in the bottom left corner of the screen, see the bottom of this topic.

After entering a custom job name (or keeping the default name), do either of the following:

- To start the search, press the Run button. The <u>Jobs panel</u> in the application will open and contain information on the status of the search. During the search, a progress dialog is displayed where you can cancel the search using the **Stop** button, hide the dialog using the **Hide** button, or set up another query using the **Settings...** button. Use the <u>Jobs panel</u> to monitor an <u>online search</u>, stop an inprogress search, or open search results in the <u>Text</u>, <u>Table</u> and <u>Pairwise</u> views.
- To close the Search wizard without starting a search, press Cancel.

Search for text or accession number matches (Entrez)

The Entrez Search wizard is used to search a text query or accession number against NCBI's Entrez databases for nucleotides or proteins.

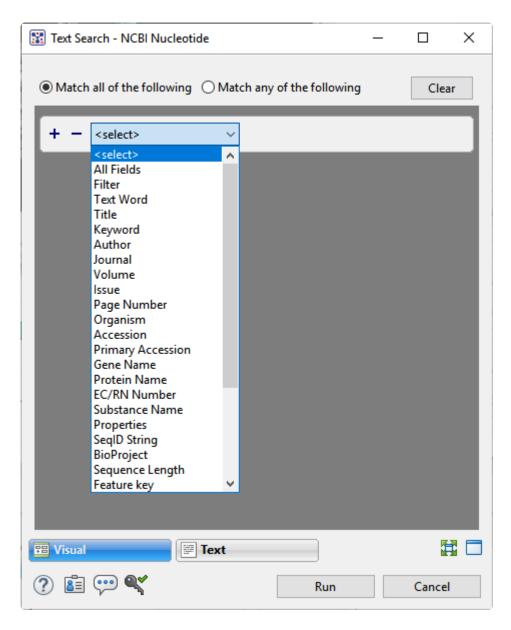
To launch the Entrez search dialog, choose Search > Text Search in > NCBI Nucleotide or Search > Text Search in > NCBI Protein. Alternatively, open the Welcome tab, choose Search and then click either Search for text in... link.

The Visual view is displayed by default and is described below.



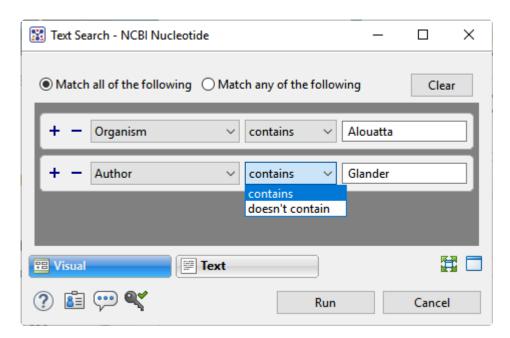
Note: The Visual view is only for non-nested queries using the same type of operator (AND, OR). To perform a more complicated search, press the **Text** tab near the bottom left corner and create the search in the Text view.

1. Use the <select> drop-down menu to choose a criterion. The field headings vary depending upon the online database that you are using. Choosing All fields will cause the application to search every word of every listing for the guery term in the database.



Once you make a selection, a second drop-down menu and an empty text box appear to the right.

- In the second drop-down menu on the row, choose **contains** or **doesn't contain** (see image below).
- In the text box, type the text that <u>should</u> be found in any results (in the case of **contains**) or <u>should not</u> be found in any results (in the case of **doesn't contain**).
- 2. To add another criterion, press the plus icon (+). To delete a criterion, press the minus icon (-).

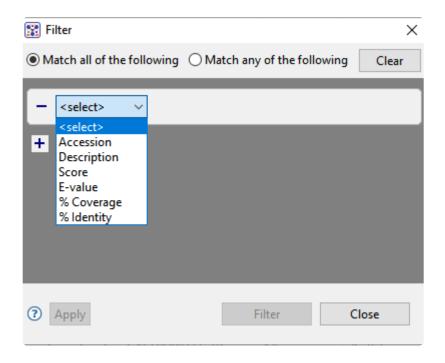


- 3. Once you have set up all desired criteria, go to the top of the dialog and choose whether to **Match all** of the following or **Match any of the following** criteria.
- 4. To learn about the optional tools in the bottom left corner of the screen, see the bottom of this topic.
- 5. Click **Run** to send the query to the database. To instead exit the dialog without searching, press **Cancel**.

During the search, a progress dialog is displayed. Once the search is completed, any matches are displayed in the <u>Jobs panel</u>.

Filter sequence search results

When you press the **Filter results** tool (\mathbb{Y}) in the header above the <u>Table</u> or <u>Pairwise</u> views, there are two options: **Filter** and **Clear Filter**. Clicking on the **Filter** option opens the following dialog:



• Use the **<select>** drop-down menu to choose a filter type, then enter the desired options.

Filter type	Description	Options
Accession	The Accession number.	
Description	The NCBI "Definition" field text.	
Score	The similarity score for the match. (See the <u>NCBI definition for raw score</u>). In general, a higher score denotes a better match.	
E-value	The chance that the match is expected to occur coincidentally. (See the NCBI definition for E value). In general, a lower score denotes a better match.	
% Coverage	The percentage of the query sequence that aligns to a match in the BLAST database.	
	Example: Query has 596 bases, of which 565 align to a match in the NCBI database. %Coverage = 565/596 = 95%.	
% Identity	The percentage of bases that match when the query and result sequence are aligned. (NCBI definition).	
	Example: Query has 565 bases that align to the match and 3 bases that don't	

match. **%Identity** = 562/565 = 99.5%.

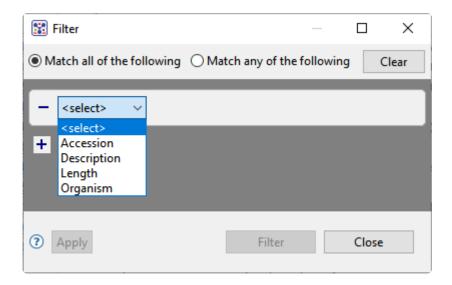
- To add another filter, press the plus icon (+). To delete a filter row, press the minus icon (-).
- Once you have set up all desired queries, go to the top of the dialog and choose whether to **Match all** of the following or **Match any of the following** queries.

Once you are finished setting up filters, do any of the following:

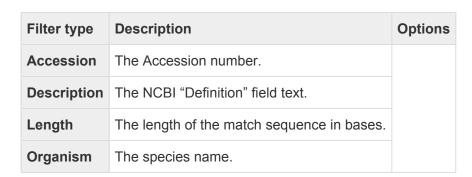
- To apply the filters, press Apply or Filter.
- To clear all filters and start over, press Clear.
- To exit the dialog without filtering, press **Close**.

Filter text search results

When you press the **Filter results** tool ($\overline{\mathbb{Y}}$) in the header above the <u>Table</u> or <u>Pairwise</u> views, there are two options: **Filter** and **Clear Filter**. Clicking on the **Filter** option opens the following dialog:



Use the <select> drop-down menu to choose a filter type, then enter the desired options.



- To add another filter, press the plus icon (+). To delete a filter row, press the minus icon (-).
- Once you have set up all desired queries, go to the top of the dialog and choose whether to Match all
 of the following or Match any of the following queries.

Once you are finished setting up filters, do any of the following:

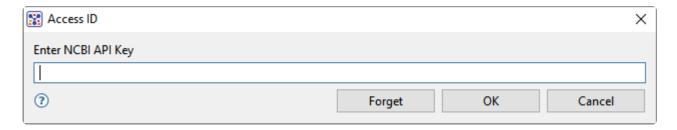
- · To apply the filters, press Apply or Filter.
- To clear all filters and start over, press Clear.
- To exit the dialog without filtering, press Close.

Access ID dialog

As of December 2018, NCBI requires an API key for faster access to NCBI services, or to run more searches concurrently. To get a key, start by <u>setting up an NCBI account</u>. After registering your account, click on your user name and request your key using **API Key Management**. Once you have your key, you can enter it in Lasergene applications to enable NCBI's premium search access.

To add your key to a Lasergene application:

- Begin an <u>Entrez database search</u> using Search > Text Search in > NCBI Nucleotide or Search >
 Text Search in > NCBI Protein.
- 2. Press the **Set Access ID** () tool in the bottom left corner of the dialog.



- If an NCBI API key is not already present in the text box, enter your key and press **OK**.
- If an NCBI API key is present but you no longer use it, press **Forget** to remove it from the application's memory. Then enter the desired key and press **OK**.
- To exit without saving your changes, press Cancel.

Save and reopen saved search results

To save the <u>Table</u>, <u>Pairwise</u> and <u>Text</u> views for the current search, use the **Save as** tool (in the top right of the Table view. The results are saved in .search format.

To reopen saved results, choose **Search > Open Search File** and navigate to the .search file.

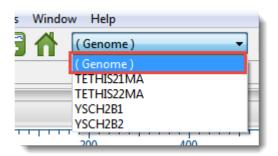
Specify the Experiment or Chromosome to Display

The <u>Genome view</u> can display multiple experiments, chromosomes, contigs or sequences simultaneously. However, only one of these at a time may be displayed in the <u>Analysis view</u> and <u>Overview</u>.

To display Genome view rows for all uploaded session data:

Do any of the following:

- Press the **Home** (1) button tool.
- Use the View > Genome > Show command.
- Click on the Genome tab.
- Use the <u>navigation tools</u> to choose (**Genome**) from the drop-down menu.



To display an individual chromosome, contig or sequence in both the Overview and Analysis view:

Do any of the following:

- · Double-click on any item in the Genome view.
- Double-click on any item in the Experiments or Chromosomes sections of the Explorer panel.
- Use the <u>navigation tools</u> to make a selection other than **(Genome)** from the drop-down menu.

Work with Tracks

Tracks are rows of information pertaining to a header, footer, or individual sequence row in certain GenVision Pro <u>views</u>. Tracks may be displayed or hidden by check or unchecking boxes in the <u>Tracks panel</u>.

The following tracks can be applied, removed or (in some cases) edited:

- Feature
- Reference
- Ruler
- Sashimi
- Alignment tracks
- Coverage
- Variant tracks

Some tracks may change in appearance depending upon the zoom level.

Where does the data used to build a track come from? When you open an .assembly file (or folder of these files) in GenVision Pro, the information contained in the file is parsed out for display as data tracks. For example, as of Lasergene 17.3, GenVision Pro can create the following tracks for each experiment or assembly: a reference sequence, reference ruler, reference features, read tracks, coverage plots, and Sashimi plots (if present). If the assembly project involved BED or WIG files, these are also available as tracks. Track data can also be imported manually or can be calculated by the application itself (e.g., Ruler and Coverage tracks).

The following table shows tasks that affect the display of tracks within the views:

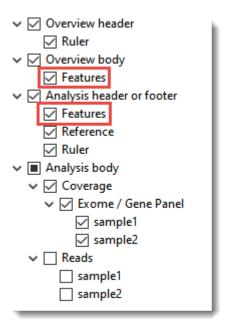
Task	How to		
Display a track in the views	To display a specific track in all views, check the box next to the name of the track in the <u>Tracks panel</u> .		
Hide one or more tracks in the views	 To hide a specific track from all views, uncheck the box next to the name of the track in the Tracks panel. To hide a specific track in one view only, right-click on the track within that view and choose Hide Track. Alternatively, select a track and use Edit > Delete or press the Delete key. 		
Change the	In the Analysis view or Overview, hover your mouse over the bottom middle of a track to		

vertical height of a track	enable a slider. Grab the slider and pull to reveal more or less of that track.
Add white space around a track	To add white space around a track in the <u>Analysis view</u> or <u>Overview</u> , select the track name in the <u>Tracks panel</u> , then use the sliders in the <u>Layout section</u> .
Access the options for a particular type of track	 Pright-click on a track in a view and choose Show Track Options. Click View > Tracks > Options or open the Tracks panel and click the Options bar (Options), then specify the track of interest by clicking on its name in the Tracks panel.
Automatically organize tracks in the Analysis view	Use Analysis > Order Tracks Ascending (or Descending) By > Category (or Name). These commands are especially useful if you are viewing a large multi-sample project with multiple tracks applied to the view. Sort by Name to group tracks by experiment name. Sort by Category to group tracks by track type.
Change the order of a selected track in a view	Open a sequence row to view its track. Then drag a track label (e.g. "Features") to the desired position among the other tracks and drop it there. Note that if you turn off some tracks in the Tracks panel and add them again later, they will appear in the order in which their boxes were checked.

Feature tracks

One or more Feature tracks can be added to the <u>Overview</u> and/or to the <u>Analysis view</u> header or footer. Feature tracks may come from the sequences in an assembly, or can be <u>imported separately</u> using **File > Add Track**.

For detailed information about applying tracks, see <u>Work with Tracks</u>. These are the Track panel items to check if you want to display one or both Feature tracks:



Features are displayed with the same default color scheme in both GenVision Pro and MegAlign Pro.

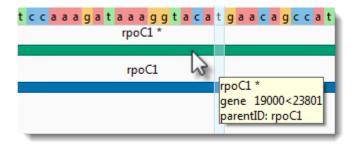
Feature Type	Color
gene	green
CDS	blue
RNA	orange
repeat	yellow
transcription	red-purple
all others	orange-red

To resize the track vertically, use the slider that appears if you hover over the bottom middle of the track data. You can also resize the track and change the amount of white space above/below the track using sliders in the <u>Layout section</u>.

To view a tool tip with details about a region or feature, hover over it with the mouse. Or select one or more features or regions by clicking on them. You can then view information about them in the <u>Details panel</u>.

Inferred features:

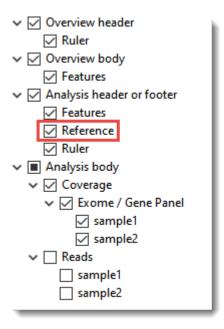
An asterisk next to a feature name indicates that the feature was inferred by SeqMan NGen. For example, if you assembled data in SeqMan NGen using a template that contained a CDS feature—without a corresponding gene feature—then "gene" would be inferred at that position, and would be displayed in the Feature track. However, the inferred gene would be marked with an asterisk.



Reference track

If you are viewing a templated assembly, you can choose to display a Reference track in the <u>Analysis view</u> header or footer.

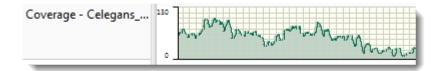
For detailed information about applying tracks, see <u>Work with Tracks</u>. This is the Track panel item to check if you want to display a Reference track:



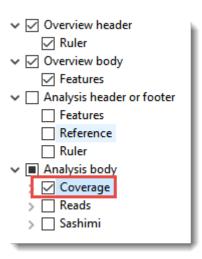
The Reference track does not have editable options. However, you can change the amount of white space above/below the track using sliders in the <u>Layout section</u>.

Coverage tracks

If data are available, you can choose to display a numeric track in the body of the <u>Analysis view</u>. As of Lasergene 17.3, the main numeric track type is the Coverage track.



For detailed information about applying tracks, see <u>Work with Tracks</u>. This is the Track panel item to check if you want to display a Coverage track:



Sources of data for numeric tracks:

- Numeric data of any type imported via a Wiggle file (.wig) using File > Add Track (see Add tracks).
- Coverage tracks imported with a SeqMan NGen .assembly package (see <u>Add assemblies</u>). This track
 displays a plot of values along one or more sequences in the session. Values are obtained
 automatically from assemblies added to the session. When the Coverage track is visible, it consists of
 a graph showing the depth of sequence coverage for each nucleotide position along the Analysis
 view.

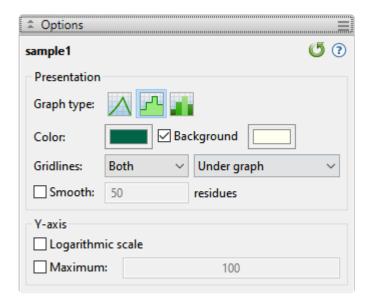
To change Coverage track options:

Coverage tracks can be customized using tools that are replicated in two locations. The first location is the Options section. To access it, select the **Coverage** track in the Tracks panel and then click the **Options** tab to expand the Options section.

Once you have checked a Coverage track in the <u>Tracks panel</u>, it is displayed in the <u>Analysis view</u>. The second location where the tools are replicated can be accessed by selecting the Coverage track in that view. A toolbar will pop up to the top-left of the track.



You can use your mouse to grab the popup on its left end and drag it to any desired location in the Analysis view. Options provided in the popup are a subset of all options available in the Options section.



- Select from three different Graph types (Line, Skyline, Bar) by clicking on the corresponding icon.
- By default, numeric graphs are displayed in green with a pale yellow background. To choose another
 graph color, click on the box to the right of Color to make a selection from a color chart. To choose
 another background color, click the color box to the right of the Background checkbox. To turn off the
 background color, uncheck the checkbox.
- Use the first **Gridlines** drop-down menu to select which grid-lines to display: **Both** (horizontal & vertical), **Horizontal** (only) or **None**. Use the second drop-down menu to choose whether grid-lines should be visible above (**Over graph**) or below (**Under graph**) other parts of the graph.
- Check the Smooth box if you wish to average the values over a specified number of residues.
 Smoothing is computed by averaging values in a sliding window of the specified number of residues.
 Enter the number of residues in the adjacent text box (maximum=5000). Inputting a high value will smooth the graph curves, while a low value will highlight local differences.
- Check the Logarithmic scale box if you want to use a logarithmic, rather than a linear scale. The log of 0 is not defined. Therefore, log values for the track are calculated as follows: For y > 0, y' = log(y); otherwise, it is displayed as a 0.
- If the default display range is not be suitable for a numeric track, you may see an empty graph or a

graph with the peaks cut off. To correct this, check the **Maximum** box and enter an integer into the associated text box. You may need to try several values before finding the ideal one.

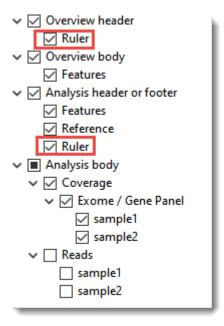
Click if you wish to return to the default values.

To resize the track vertically, use the slider that appears if you hover over the bottom middle of the track data. You can also resize the track and change the amount of white space above/below the track using sliders in the <u>Layout section</u>.

Ruler tracks

One or more Ruler tracks can be added to the <u>Overview</u> header and/or to the <u>Analysis view</u> header or footer.

For detailed information about applying tracks, see <u>Work with Tracks</u>. These are the Track panel items to check if you want to display one or both Ruler tracks:



The Ruler tracks do not have editable options. However, you can add white space above and/or below the ruler using sliders in the <u>Layout section</u>.

Sashimi tracks

Analysis of mRNA isoform expression is a key aspect of understanding gene expression. For many genes, the change in isoform ratios for a given gene produce a more significant physiological effect than the overall expression level of that gene. Isoforms are typically visualized as Sashimi plots, a way of quantitatively visualizing splice junctions for mRNA sequences aligned to an annotated genomic reference. Sashimi plots allow you to screen differentially spliced exons in genomic regions of interest. If data are available, you can choose to display a Sashimi track in the <u>Analysis view</u>.

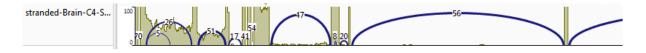
For detailed information about applying tracks, see <u>Work with Tracks</u>. This is the Track panel item to check if you wish to display a Sashimi plot:

✓ ✓ Overview header
✓ Ruler
✓ ✓ Overview body
✓ Features
Analysis header or footer
Features
Reference
Ruler
✓ ■ Analysis body
> Coverage
> Reads
Sashimi

When the Sashimi track is visible in the Analysis view, it consists of two components:

- A bar graph showing the abundance of each exon
- Arcs representing the number of reads split across the junction, also referred to as "junction depth."

Multiple Sashimi plots from multiple samples can be displayed simultaneously.



Scroll to the bottom of this page to see two videos showing Sashimi tracks in action.

Sashimi track data can come from the sequences in an assembly, or can be <u>imported separately</u> using **File**> **Add Track**. When data are available for a given sample, GenVision Pro automatically replaces the

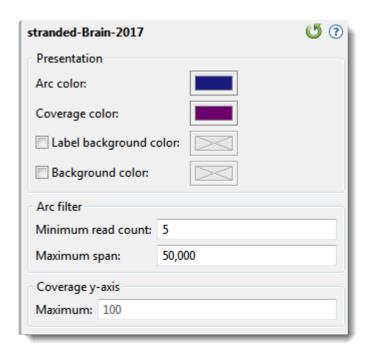
Coverage track with a Sashimi plot, which contains the same Coverage histogram, but also allows for "arcs" to be displayed. The number at the top of each arc represents the number of split reads; the thickness of each arc is proportional to this number. Note that split reads do not necessarily correspond to introns; other types of structural variation can also cause SegMan NGen to split reads. Annotations are not considered

when creating the Sashimi track.

For a brief video showing RNA-Seq data analysis using the Sashimi track, see the second video in <u>Launch a session from within ArrayStar</u>.

To change Sashimi track options:

To learn how to access the options section for this track, see Options section.



- By default, Sashimi graphs have navy blue arcs and purple coverage histograms. To choose different colors, click on the **Arc color** and/or **Coverage color** box and make a selection from a color chart.
- By default, arc numbers ("labels") and the Sashimi plot, as a whole, do not have background colors. To add colors, check **Label background color** and/or **Background color**, then click on the corresponding color box and make a selection from a color chart. To turn the colors off again, uncheck either or both check-boxes.
- **Minimum read count** lets you set a threshold for the minimum number of split reads needed to display an arc. If this value is set to zero, no filtering will occur.
- **Maximum span** allows you to set a threshold for the maximum allowable arc length, in bases. To turn off this filter, delete any numbers from the text box, leaving it blank.
- Under Coverage y-axis, set a **Maximum** limit for the height of the y-axis on the coverage graph. This setting is equivalent to the Y-axis range maximum in the Coverage tracks.

Click of if you wish to return to the default values.



Note: The height of track labels is specified using the presentation Font settings in View > Style > Analysis. To change the height of the graph and the white space surrounding it, use the sliders in the Layout section.

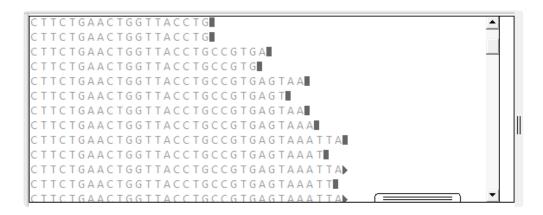
This video shows how to apply and interpret Sashimi plots in GenVision Pro:

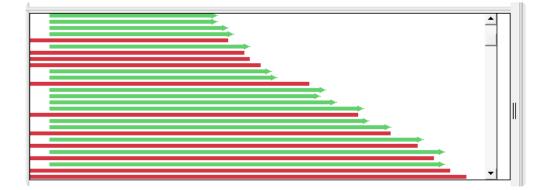
This video shows an RNA-Seq workflow that begins with assembly in SeqMan NGen, proceeds to downstream analysis in ArrayStar, and then uses ArrayStar's Send Selection to GenVision Pro command to open the results in GenVision Pro. Once in GenVision Pro, the Sashimi track is applied. (GenVision Pro section begins at 2:41).

Alignment tracks

If you are viewing an assembly created in SeqMan NGen, you can add a track to the <u>Analysis view</u> showing each of the aligned reads in the selected experiment using the method described in <u>Work with Tracks</u>. Briefly, you can display an Alignment track by opening the <u>Tracks panel</u> and checking one or more of the boxes to the left of **Analysis body > Alignment > (sample names)**.

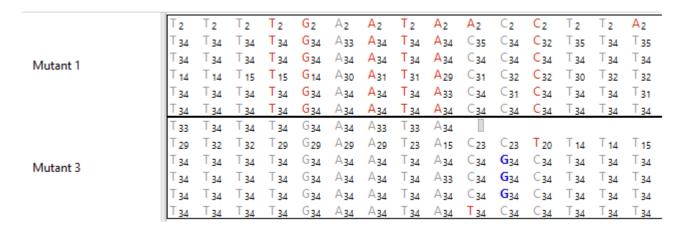
Depending on the zoom level in the view, the Alignment track will show base letters (above) or green and red arrows representing forward and reverse reads, respectively (below).





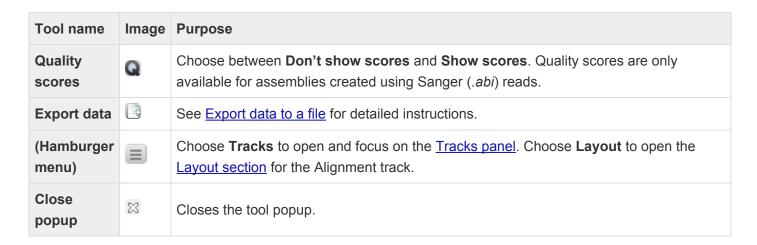
To resize the track vertically, use the slider that appears if you hover over the bottom middle of the track data. You can also resize the track and change the amount of white space above/below the track using sliders in the <u>Layout section</u>.

You can also display quality score information on the Alignment tracks as shown in the image below.



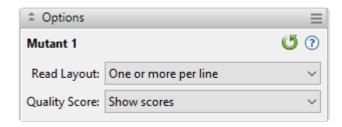
Unlike most other track types, Alignment tracks have their own tools. To see the Alignment track tools, first apply an Alignment track to the Analysis view, then select the Alignment track in that view. The toolbar will pop up at the top-left of the track. You can use your mouse to grab the popup on its left end and drag it to any desired location in the Analysis view.





To change Alignment track options:

Select the **Alignment** track in the Tracks panel and then click the **Options** tab to expand the Options section.



• Use the lowermost **Read Layout** menu to choose **One sequence per line** or **One or more per line**. If you choose **One sequence per line**, only one sequence will appear in each line, preceded or

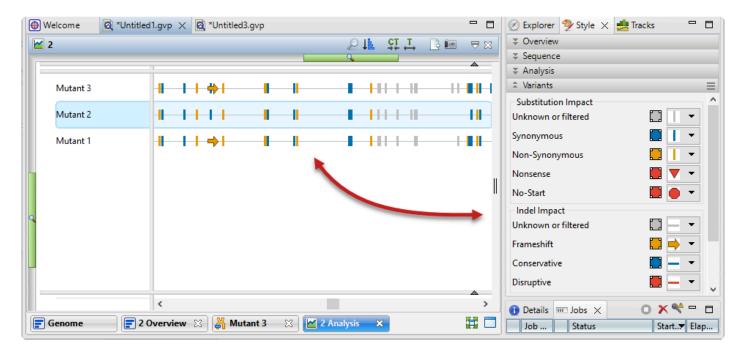
followed by blank space before and after the sequence. If you choose **One or more per line**, GenVision Pro saves vertical space by putting any sequences that do not interfere/overlap with another sequence on the same row as another sequence. These sequences are placed in areas that would otherwise have consisted only of blank space.

- Use the lowermost Quality Score menu to choose Don't show scores, Show scores, or Average scores.
- Click if you wish to return to the default values.

Variants tracks

If you are viewing an assembly created in SeqMan NGen, you can add a Variants track to the <u>Analysis view</u> using the method described in <u>Work with Tracks</u>. Briefly, you can display this track by opening the <u>Tracks</u> <u>panel</u> and checking one or more of the boxes to the left of **Analysis body > Variants > (sample names)**.

You may need to use the green horizontal zoom slider to see the variants clearly. The styles used to represent each type of variant are controlled by choices made in the <u>Variants section</u> of the <u>Style panel</u>.

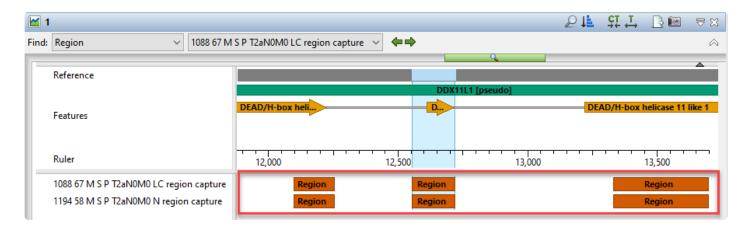


To resize the track vertically, use the slider that appears if you hover over the bottom middle of the track data. You can also resize the track and change the amount of white space above/below the track using sliders in the <u>Layout section</u>.

Region tracks

If you created an exome or gene panel assembly in SeqMan NGen using a BED file with targeted regions, GenVision Pro lets you display a Region track in the <u>Overview</u> and/or <u>Analysis view</u>. To display the Regions track, check the associated checkboxes in the <u>Tracks panel</u>. There may be multiple feature and/or region tracks available under the "parent" track.

Region tracks show each region as an orange bar. You may have to adjust the horizontal zoom level to make the bars visible.



To find and navigate between regions, do either of the following:

- Use the **Find** tool () in the Overview or Analysis view, choose **Region** from the Find drop-down menu, then use the green arrows to find the next/previous region.
- Open the <u>Enrichment Report</u> and click on any row to navigate to that region in the Overview and Analysis view.

To view a tool tip with details about a region or feature, hover over it with the mouse. Or select one or more features or regions by clicking on them. You can then view information about them in the <u>Details panel</u>.

Region tracks do not have customizable options. However, you can also add white space above and/or below features tracks using sliders in the <u>Layout section</u>.

Customize the Look and Layout

The following topics describe the many ways you can customize the look and layout of GenVision Pro.

- <u>Use the Style panel to change the look of a view</u> (font styles and sizes, colors, number of decimal places, etc.)
- Customize the appearance and layout of views
- Customize the layout of other window components
- Apply information tracks to a view

Once you have customized the look and layout as desired, you can:

- Save Styles for individual views and apply them to other sessions
- Save the entire layout and apply it to other sessions

Customize the appearance and layout of views

The following table lists ways in which you can customize the appearance and layout of views.

Task	How to	
Show a view	Use View > (View Name) > (chromosome or contig number) > Show. The third item in the command at left may or may not appear. Example: to see the Analysis view for Chromosome 12, choose View > Alignment > 12 > Show.	
Resize a track in a view	In the Analysis view or Overview, hover your mouse over the bottom middle of a track to enable a slider. Grab the slider and pull to reveal more or less of that track.	
Add white space around a track in a view	To add white space around a track in the <u>Analysis view</u> or <u>Overview</u> , select the track name in the <u>Tracks panel</u> , then use the sliders in the <u>Layout section</u> .	
Change the style (fonts, colors, calculation metrics and more)	Use the <u>Style panel</u> and its subsections. Customization for font style and size applies to all text in a view. Individual rows cannot have different styles.	
Move one or more contig rows as a group	In the <u>Overview</u> , you can rearrange contigs one at a time or as contiguous groups using drag and drop. First, select the desired rows using Ctrl/Cmd+click or Shift+click . Next, drag the selected row(s) to the desired location and release the mouse button.	
Detach a view	Only certain views can be detached. To do this, right-click on the view tab and choose Detach (View Name). Alternatively, use View > (View Name) > (chromosome or contig number) > Detach. Detaching a view turns it into a free-floating window.	
Reattach all detached views to their original locations	Use View > Reattach All Views.	
Show the active view in full-screen mode	Use the View > Full Screen command or the the Full screen tool at the bottom right corner of the view.	

Exit from full-	Click the Esc key, use the Exit Full Screen tool (at the bottom of the full-screen view,		
screen mode	or press Ctrl/Cmd+Shift+F.		
Configure the number and location of visible views.	Use the View > Document Layout command or the Change layout tool (image varies) in the bottom right corner of the view to access possible layouts. The following options are available: • One - One - To show only the current view. • Two Horizontal (Two Horizontal) - To show the current view on top and one additional view below it. • Two Vertical (Two Vertical) - To show the current view on the left, and one additional view to the right. • Three Horizontal (Three Horizontal) - To show the current view on top and two additional views below it. • Three - Left (Three - Left) - To show the current view on the left and two additional views stacked one on top of the other on the right. • Three - Right (Three - Right) - To show the current view on the right and two additional views stacked one on top of the other on the left. • Three - Top (Three - Top) - To show the current view on the top and two additional views side-by-side on the bottom. • Three - Bottom (Three - Bottom) - To show the current view on the bottom and two additional views side-by-side on the top.		
Zoom in or out on a view	Both the Overview and Analysis view have independent vertical and horizontal zoom sliders. Grab any green slider with your mouse and drag left/right or up/down to zoom in or out in the view.		
Split a view vertically	Drag the vertical splitter tool (), found at the right edge of the view, in a leftward direction. Splits are ignored when exporting an image, although the layout is preserved. To remove a split, drag the divider as far as you can to the right of the view and drop it there.		
To view two or	By default, only one session at a time is displayed, even if several sessions are open. To		

more sessions simultaneously on a split screen	view two or more sessions simultaneously on a split screen, drag a session tab in a downward, right or left direction. When you see a black arrow, release the mouse button. Sessions can be re-docked in their original positions by reversing this process.
Return to the default window display	Choose View > Reset Panel Layout.

Customize the layout of other window components

Controls for this application's view, panels and expand bars are described in the following table. For information that pertains only to views, also see <u>Customize the appearance and layout of views</u>.

Task	"How To"		
Show or open a hidden item	Use the associated View menu command, click on a closed expand bar or click the Restore icon (Restore).		
Show/Hide the side panels	Double-click on the project name to toggle between showing and hiding the side panels on the right of the window.		
Display panel sub-sections	Click on a closed expand bar (medium gray bar) or use the associated View menu command. If you can't see all the information in an expanded section, close other expanded sections and/ or increase the size of the whole panel as described elsewhere in this table.		
Hide panel sub-sections	Click on an open expand bar. Alternatively, click the "hamburger menu" () on the right of an expand bar and choose Collapse All to close all sections or Collapse Others to close all but this section.		
Maximize an item	use the Maximize icon (.		
Change the relative sizes of panes	Drag the divider bar between them to the left, right, up or down.		
Hide or close a visible item	Use the Close icon () or Minimize icon (), or click on an open expand bar. To close one or more tabs, right-click on a project, view or panel tab and choose Close , Close Others , Close Tabs to Right , or Close All . Some of these options are only available for certain tabs.		
Move a view, panel or project to another location	Right-click on the view tab and choose Move (View Name) to (Location) . In some cases, you can use your mouse to grab a project, view or panel tab, then drag and drop it in the desired location.		
Scroll vertically	Use the vertical scrollbar or the mouse wheel.		
Scroll horizontally	Use the horizontal scrollbar, or hold down the Alt key (Win) or Option key (Mac) and drag the mouse.		

Return to the default side panel layout	Choose View > Reset Side Panel Layout . This command restores the default layout of panels, including opening any default panels that have been closed; and restoring panels that have been minimized or maximized. It will not reattach a detached view. In that case, use View > Reattach All Views instead.
Return to the default layout for the entire window	Choose View > Document Layout > Restore DNASTAR Default Layout.

Save a custom layout and apply it to other projects

You can change the layout of GenVision Pro's views and panels as described in <u>Customize the look and layout</u>.

Once you have arrived at a layout you like, you can save it and then apply it to future projects, either by default or only when you specify that it be applied.

Task	How To
Save the current layout to a file that can be applied to future projects	Use View > Document Layout > Save Layout.
Save the current layout to a file that will be applied as the default layout for future projects	Use View > Document Layout > Save Layout as Default.
Apply a layout that was saved earlier to the current project	Use View > Document Layout > Apply Layout.
Restore the DNASTAR default layout to the current project	Use View > Document Layout > Restore DNASTAR Default Layout.

Use Basic Editing Commands

The following basic editing commands are available in GenVision Pro:

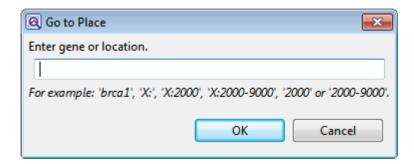
Command	Keyboard Shortcut	Right- Click Option	Description
Edit > Cut	Ctrl/Cmd+X	Cut	This command is not enabled in GenVision Pro.
Edit > Copy	Ctrl/Cmd+C		To <u>copy</u> text or data to the clipboard. A warning message will appear if you attempt to copy an extremely large amount of data to the clipboard.
Edit > Copy As FASTA	Ctrl/ Cmd+Shift+C	Copy as FASTA	To <u>copy</u> text or data to the clipboard in FASTA file format. This command is only available if a subsequence is selected.
Edit > Paste	Ctrl/Cmd+V		To paste text from the clipboard into a text field. This command is disabled in all but a few circumstances.
Edit > Delete	Delete key		To remove selected text from a text box, or to remove one or more selected tracks from a <u>view</u> . Using Edit > Delete or pressing the Delete key has the same effect as unchecking the corresponding box(es) in the <u>Tracks panel</u> .
Edit > Select All	Ctrl/Cmd+A		If used in the Overview, selects all chromosomes. If used in the Analysis view, selects all of the active chromosomes in both views.

Search Within the Project

To navigate to a specific position:

Do any of the following:

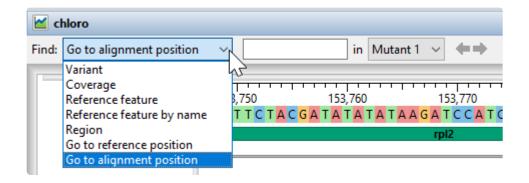
- Open the <u>Places panel</u> and double-click on a previously-stored location.
- Use Edit > Go to Place or press Ctrl/Cmd+J. In some cases, the right-click option Go to Place may
 also be available. In the Go to Place dialog, type the desired feature name, numerical position, or
 numerical range. Then click OK.



To navigate to a location or feature using the Search bar:

The search tools are available from the <u>Overview</u> and <u>Analysis view</u>. By default, the search tools are hidden.

1. Display the tools by clicking the **Find** tool ($\stackrel{\frown}{\sim}$) at the top right of either view or use **Edit > Find**. This causes the search tools to be displayed on the top left of that view.



Use the **Find** drop-down menu to choose the item to search for. Available search target options
depend on the project type (.sqd vs. .assembly), assembly type (de novo vs. templated), and the view
(Analysis view vs. Overview). The table below shows possible options in alphabetical order.



Coverage	Search for the edge of a region that is below the minimum threshold coverage.
Go to alignment position	Go to a specific numerical position on the aligned sequence rows.
Go to reference position	Go to a specific numerical position on the reference sequence.
Reference feature	Searches for and selects the next or previous feature on the reference sequence. To see these features in the view, be sure to check the Reference > Features box(es) in the <u>Tracks panel</u> .
Reference feature by name	Search for and select a feature by entering its name and by optionally limiting the search to particular types, such as gene , CDS , etc.
Region	(This functionality is not yet implemented as of Lasergene 17.5.)
Variant	Locates bases that have been identified by SeqMan NGen's SNP-calling detector as variants. Variant is available for reference-guided assembly projects only.

- 3. Different search options may cause additional drop-down menus or text boxes to appear, allowing you to type in a text string of query bases, choose a feature category, etc. In some cases, a box is provided to **Allow ambiguity**. Check the box if you want to include IUPAC ambiguity codes in your search.
 - If you type a query that is "too wide" for the search box, you can hover over the search box to view the entire query. Note that long query sequences may take considerable time to locate.
- 4. Use the green right/left arrows () to jump to the next/previous find. After reaching either end of an alignment, SeqMan Ultra wraps around and continues searching from the other end. Each time a feature is located, it is selected in all views until you press one of the green arrows again.

Make a Selection

Within the <u>Analysis view</u>, you can select an entire chromosome, a portion of one, or a specific portion of chromosome corresponding to a feature. The active selection is highlighted in light blue. The selection can then be copied to the clipboard in standard or FASTA format.

To select	Do this		
One or more rows	In either the Overview or Analysis view, click on a row label to make it the active selection, and click again to deselect it. To select multiple rows, use Ctrl/Cmd+click or Shift+click .		
The entire active chromosome	Use Edit > Select All.		
A portion of the active chromosome	Drag the mouse across residues in the sequence track while holding down the left mouse button.		
Only the portion of a single sequence associated with a section of track (e.g. a Gap Fraction peak, etc.)	Drag the mouse across the open track while holding down the left mouse button.		
One or more tracks	Click on a track to select it. To add additional tracks to the selection, use Ctrl/Cmd+click or Shift+click .		
One or more features and corresponding sequence	Click on a feature to select it. To add additional features to the selection, use Ctrl/Cmd+click .		
The portion of a single sequence associated with a feature	To expand a selection consisting of one or more features to a selection containing just the sequence corresponding to the feature(s), do any of the following: • Click on the feature to select it, then use Edit > Select Subsequence. • Click on the feature to select it, then right-click and choose Select Subsequence.		
	 Double-click on the feature under the sequence. In the <u>Details panel</u>, click on the Action Select Subsequence. 		
Additional continuous items, when the first endpoint has been selected	Use Ctrl/Cmd+Shift and click on the other endpoint.		
	Hold down the Ctrl/Cmd key while clicking on the items to toggle on or off. For subsequence selection, hold down the Ctrl/Cmd key while dragging the mouse.		
Toggle selection of items on or off	For subsequence selection, hold down the Ctrl/Cmd key while dragging the		
	For subsequence selection, hold down the Ctrl/Cmd key while dragging the		

expanded to encompass additional bases

new endpoint. The selection will expand to use the chosen base as the new endpoint. If there are multiple selected ranges, the one closed to the chosen base will be expanded.

Rename Assemblies and Tracks

A selected assembly or track can be renamed manually or by using specified data fields. See the following topics for details:

- Create names automatically using specified data fields
- Rename a single sequence manually

Create names automatically using specified data fields

The Edit > Rename with Fields command lets you create labels for one or more selected sequences automatically, using specified fields such as sequence length, organism, or description. The custom labels are then displayed for those sequences within the GenVision Pro application.

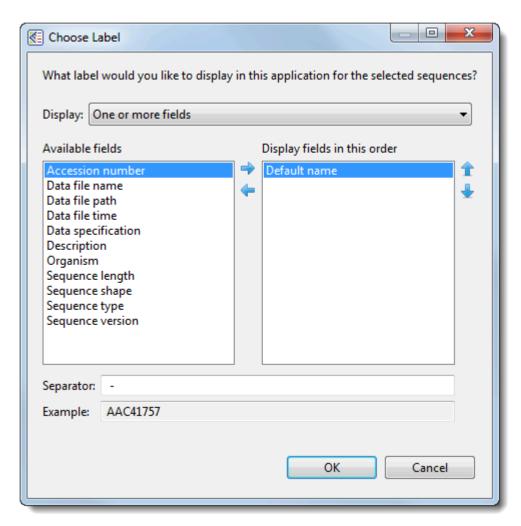
To rename sequences using specified data fields:

- 1. Select one or more sequence names from any GenVision Pro view.
- 2. Right-click on the selection and choose Rename with Fields, or instead choose Edit > Rename with Fields from the main menu. Making no selection prior to using the main menu command is equivalent to selecting all sequences.

The Choose Label dialog opens with the **Display** drop-down menu set to **One or more fields**.



Note: If you selected a single sequence and wish to rename it manually, select Custom label from the drop-down. Then follow the instructions in Rename a single sequence manually.



The left pane shows available fields, and the right pane shows fields that will be displayed as part of the sequence name. Initially, **Default name** is the only field listed in the right pane. An example showing the appearance of the name using the specified field(s) appears in the **Example** box at the bottom of the dialog.

- 3. Select and organize the fields you wish to display as part of the sequence name:
 - To add or remove items, move them between panes using the left and right arrows (< →).
 Use the up and down arrows on the right (→) to specify the order in which displayed fields should appear.
 - By default, a hyphen (-) is included between each of the displayed fields. If desired, type in a
 different Separator. Or remove all text in the Separator box to include nothing between the
 fields.
 - The **Example** text box shows what the label will look like with the current selections. This text can be manually edited, if desired.
- 4. When you are satisfied with the choices made in this dialog, click **OK** to save changes and exit from

the dialog. Or click **Cancel** to leave the dialog without saving any changes.

Rename a single sequence manually

The Edit > Rename command lets you manually enter a label (name) to display for those sequences within the GenVision Pro application.

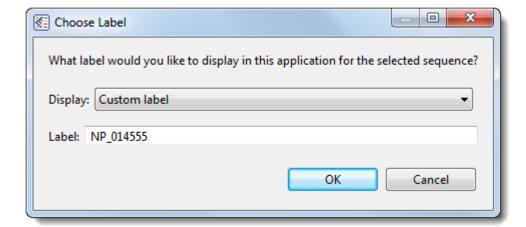
To manually rename a single selected sequence:

1. Select a single sequence name from any GenVision Pro view, then right-click on it and choose **Rename**. Alternatively, select the sequence name and choose **Edit > Rename**.

The Choose Label dialog opens with the **Display** drop-down menu set to **Custom label**.



Note: To instead name the sequence using specified data fields, select One or more fields from the drop-down. Then follow the instructions in Creating names automatically using specified data fields.



- 2. Type the desired name into the **Label** text box.
- 3. Click **OK** to save changes and exit from the dialog, or **Cancel** to leave the dialog without saving any changes.

Copy

In many cases, once you have <u>selected</u> text, data, sequences or subsequences, you can copy the information to the clipboard. The copy functions described below are accessible using menu commands, context menu commands (i.e., "right-click options"), and/or keyboard shortcuts.

- Edit > Copy or Ctrl/Cmd+C Copies the selected text, just as it appears in the view.
 - If a portion of sequence or one or more features are selected, this command copies the selected subsequence(s) as one long string, including the gaps. The subsequence is copied as one long string, without line breaks. If the selection extends across multiple members of the alignment, names are included as prefixes.
 - If a selection is made in the Style panel and in dialogs, use Ctrl/Cmd+C.
- Edit > Copy As FASTA Copies the selected text or data in FASTA format.

Export

For detailed steps on exporting data and images from GenVision Pro, see the following topics:

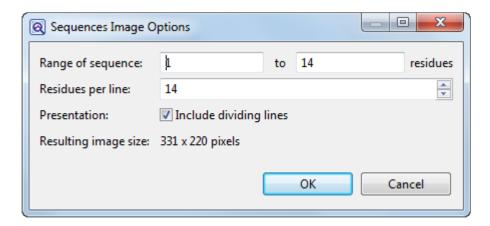
- Export an image of a view (and edit it in PowerPoint)
- Export data to a file a file}

Export an image of a view

GenVision Pro lets you export an image of the <u>Overview</u> or <u>Analysis view</u> for use in a text editor or similar program.

- 1. Do either of the following:
 - Press the Export image tool () in the upper right corner of the view you want to export.
 - Use File > Export Image > Overview or File > Export Image > Analysis.
- 2. Specify the file name and choose the export format: Adobe PDF (.pdf), Microsoft PowerPoint-optimized (.pptx), PNG image (.png) or JPG image (.jpg, .jpeg).
 - If you choose Adobe PDF format and a particular font is not supported there, a similar font will be substituted for the original one automatically.
 - If you choose the format that is optimized for use with Microsoft PowerPoint, the GenVision Pro
 image is converted into an editable set of objects that can be grouped, ungrouped and moved
 within the target application. To constrain an object (e.g., a GenVision Pro track) within
 Microsoft PowerPoint to move only vertically, rather than laterally, hold down the Shift key.

If exporting the Analysis View, the following dialog appears. If exporting the Overview, an abbreviated version appears.



- 3. Choose export options.
 - Use the Range of sequence boxes to specify the range of residues that should appear in the image. The default is the current viewport range.
 - ("Analysis" only) Type in the number of **Residues per line**. By default, GenVision Pro places the entire range on a single line.

• ("Analysis" only) Check the **Include dividing lines** box if you wish to include separators between the header, sequences, and footer.

The **Resulting image size** updates automatically based on the options you choose.

4. Click **OK** to export the image using the specified options or Cancel to close the dialog without exporting.



Note: If you saved the image in PowerPoint format, you can edit the image after opening it in PowerPoint. See Edit an exported image in PowerPoint.

Edit an exported image in PowerPoint

The following video shows how to edit <u>exported</u> images of your alignment or genomic tracks within Microsoft PowerPoint:

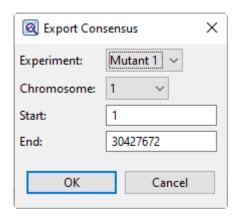
Export data to a file

Data can be exported using the **Export data** tool () located in the top right of some GenVision Pro views and panels. The table below shows export options available from each view, as well as menu commands that can be used to export aligned sequences or the consensus sequence.

To export	Click tool in	Available options	Notes
All of the aligned sequences	Analysis view	Export Aligned Sequences	See complete instructions below this table. Instead of using the tool, you can also use the File > Export Data > Export Aligned Sequences menu command or right-click within the Analysis view and choose Export Aligned Sequences.
The consensus sequence	Analysis view	Export Consensus	See complete instructions below this table. Instead of using the tool, you can also use the the File > Export Data > Export Consensus menu command or right-click within the Analysis view and choose Export Consensus.
Contents of the Variants view table	Variants view	NA	Save data in either comma- or tab-separated value (.csv or .tab) formats.
Matching regions of displayed matches, but not the query sequence itself	Table view	Export Data > Matches without Gaps	Data is saved as an ungapped FASTA (.fasta) file.
Displayed matches aligned against the query	Table view	Export Data > Aligned Matches	Data is saved as a gapped FASTA (.fasta) file.
Contents of the Table view	Table view	Export Data > Table	Data is saved in either comma- or tab-separated value (.csv or .tab) formats.
Contents of the Text view	Table view	Export Data > Text	Data is saved in text (.txt) format.
A file of selected sequence matches	Table view	Download Sequences	Data is saved in GenBank (.gbk) format.
A selected data source in BAM format	Explorer panel	NA	Data is saved in BAM (.bam) format.

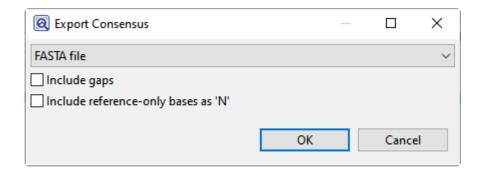
Exporting a consensus sequence or the aligned sequences:

When you export a consensus sequence, the following popup opens:



- Use the Experiment and Chromosome dropdown menus to choose the data source and chromosome name/number.
- Use the **Start** and **End** areas to optionally define a subrange to export.
- Press OK.

A new popup appears:



- Use the drop-down menu to choose from FASTA file, GenBank file or Lasergene DNA file formats.
- If you wish to preserve gaps in the consensus, add a checkmark next to **Include gaps**.
- If you wish to include bases that were only found in the reference sequence and not in any read sequences, add a checkmark next to **Include reference-only bases as 'N'**.
- · Press OK.
- A Save As dialog appears. Choose a name and filetype and press Save or OK.

Save, Close and Exit

To save one or more sessions:

To save one or more sessions or to overwrite a previously saved version of the session in GenVision Pro (.gvp) format

- To save all open sessions, use File > Save All Sessions or press Ctrl/Cmd+Shift+S.
- To save only the active session, do any of the following:
 - Use the Save Session button tool (). If you want to preserve the earlier version of a session, or store a copy of the session with a different name or location, use File > Save Session As instead.
 - Choose File > Save Session
 - Press Ctrl/Cmd+S.

In all cases, GenVision Pro saves to the desktop by default.



Note: When you add Wiggle files using File > Add Track, they are indexed for fast retrieval. When the session is saved, these indices are written to a folder called <Session Name> .gvp.index. When you reopen the session, GenVision Pro uses these indices for faster performance. If you move or copy the .gvp file to another location, you can move the .gvp.index folder also; this avoids the need to recompute indices upon reopening the session. Alternatively, you can move just the .gvp file, and the indices will be automatically recomputed.

To close the selected session:

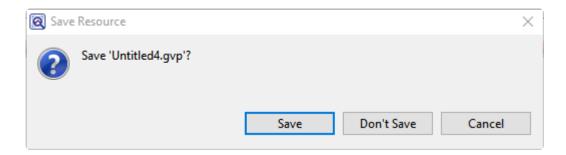
Use File > Close Editor or press Ctrl/Cmd+W. If multiple sessions are open, first select the tab corresponding to the session you wish to close.

To quit the application:

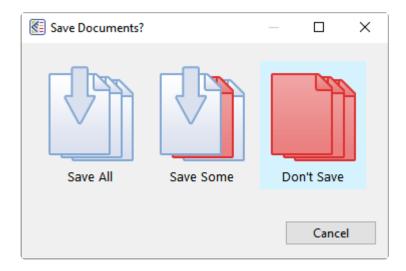
- Windows: Use File > Exit or Ctrl+Q.
- Macintosh: Use GenVision Pro > Quit GenVision Pro or Cmd+Q.

Closing or quitting when there are unsaved sessions:

If you select **File > Close**, or if you quit GenVision Pro when just one unsaved session is open, GenVision Pro will open the Save Resource dialog. Click **Save** to save changes, **Don't Save** to close/quit without saving changes, or **Cancel** to return to GenVision Pro without closing/quitting.



If you select **File > Close**, or if you quit GenVision Pro when multiple unsaved sessions are open, GenVision Pro will open the Save Documents dialog.



To save all changed projects, choose **Save All**. To choose which projects to save, choose **Save Some**. To exit from GenVision Pro without saving any sessions, choose **Don't Save**. Click **Cancel** to return to GenVision Pro without saving the sessions.

Appendix

The Appendix contains the following topics:

- View a list of supported file types
- <u>Set preferences</u> for the <u>Internet</u> or for <u>specifying temporary file locations</u>.
- View <u>IUPAC codes</u>
- See the installed Lasergene file locations
- Turn off usage logging
- Troubleshoot the error
- · Clean cached files on startup
- See the list of <u>research references</u> related to this application

Supported file types

For a list of file formats supported by this DNASTAR product, please see the <u>File Formats</u> page of our website.

Set preferences

To open the Set Preferences dialog on Macintosh, choose **GenVision Pro > Preferences** (macOS 12 and lower) or **GenVision Pro > Settings** (macOS 13 and higher). On Windows, choose **Edit > Preferences**.

You can set preferences in GenVision Pro for Internet, and temporary files.

Set Internet preferences

The Internet Preferences dialog lets you choose whether you wish to be notified about Lasergene updates.

To access the dialog on Windows, use **Edit > Preferences**. On Macintosh, use **GenVision Pro > Preferences** (macOS 12 and lower) or **GenVision Pro > Settings** (macOS 13 and higher). If the Internet section is not already active, click on the word "Internet" on the left.



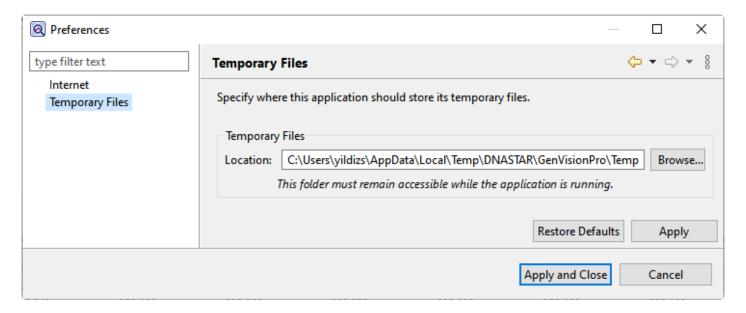
Check the box if you would like GenVision Pro to display a notification when a newer version of Lasergene is available. Uncheck the box if you do not wish to receive this notification.

Click **Apply** and **OK** to save your changes or **Cancel** to ignore any changes. Click **Restore Defaults** to return to the default settings for all four drop-down menus.

Set temporary file preferences

GenVision Pro creates and uses temporary files while a session is open. The need for temporary file space grows with the number and length of the sequences and track data in open sessions. If you expect the temporary files will be too large to fit in the default location, you may specify a different location for these files using the Temporary Files Preferences dialog.

To access the dialog on Windows, use **Edit > Preferences**. On Macintosh, use **GenVision Pro > Preferences** (macOS 12 and lower) or **GenVision Pro > Settings** (macOS 13 and higher). If the Temporary Files section is not already active, click on the phrase "Temporary Files" on the left.



To change the temporary file location, close all GenVision Pro sessions and then return to this dialog. Select the **Browse** button, navigate to the new location and click **OK**. As noted in the dialog, the location you choose must remain accessible while GenVision Pro is running. The new location will be used for all sessions that are subsequently created or opened.

Click **Apply** and **OK** to save your changes or **Cancel** to ignore any changes. Click **Restore Defaults** to return to the default (temporary) storage location.

IUPAC codes

IUPAC codes for amino acids:

1-Letter Code	3-Letter Code	Amino Acid
Α	Ala	Alanine
В	Asx	Aspartic Acid or Asparagine
С	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
Н	His	Histidine
1	lle	Isoleucine
J	Xle	Leucine or Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
0	Pyl	Pyrrolysine
Р	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
Т	Thr	Threonine
U	Sec	Selenocysteine
V	Val	Valine
W	Trp	Tryptophan
Х	Xaa	Unspecified or unknown
Υ	Tyr	Tyrosine

Z	Glx	Glutamic Acid or Glutamine

IUPAC codes for nucleotides:

Symbol	Description	Bas	es Re	prese	ented
А	adenosine/adenine	А			
С	cytidine/cytosine		С		
G	guanosine/guanine			G	
Т	thymidine/thymine				Т
U	uridine/uracil				U
W	weak	Α			Т
S	strong		С	G	
M	amino	Α	С		
K	keto			G	Т
R	purine	Α		G	
Υ	pyrimidine		С		Т
В	not A		С	G	Т
D	not C	Α		G	Т
Н	not G	Α	С		Т
V	not T	Α	С	G	
N or -	any base (not a gap)	Α	С	G	Т



The non-standard nucleotide codes "I" (Inosine) or "X" (xanthine) are supported in DNASTAR's PrimerSelect and EditSeq applications only.

Installed Lasergene file locations

The following file names use 'x' to represent the version number.

File Category	Application	Path	
	ArrayStar	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x\ArrayStar	
Application ²	SeqNinja (command line)	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x\SeqNinjaCL	
Application	All others	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x	
		Macintosh: /Applications/ DNASTAR/Lasergene x	
Data Manager ² (<i>DMx</i> , <i>DMx</i> .exe)	SeqBuilder Pro, Protean 3D, GeneQuest, MegAlign,	Windows: C:\Program Files (x86) \DNASTAR\ Lasergene x	
	MegAlign Pro, PrimerSelect	Macintosh: /Applications/ DNASTAR/Lasergene x	
		Windows: C:\Program Data\DNASTAR\ DataManager	
Data Manager State File	SeqBuilder Pro, Protean 3D, GeneQuest, MegAlign, MegAlign Pro, PrimerSelect	Macintosh: ~/Library/ Application Support/ DNASTAR/DataManager,	
		~/Library/Preferences/ DNASTAR/DataManager	
		Windows: C:\Program Files (x86)\DNASTAR\License Manager	
License Manager ²	All	Macintosh: /Applications/ DNASTAR/License Manager	

Server License File (<i>Iservrc</i>), Server License Manager ² , <i>Server Executables</i> (<i>_Iserv, Iservnt.exe</i>)	All	Windows: C:\Program Files (x86)\DNASTAR-LicenseServer\Server Macintosh: ~/Library/ DNASTAR-LicenseServer
Standalone & Trial Licenses (*.license), License Server Client License (*.lshost), Key Server Client License (*.keyhost)	All	Windows: C:\Program Data\DNASTAR\Licenses Macintosh: ~/Library/ Application Support/ DNASTAR/Licenses
	Protean 3D, Navigator, SeqNinja (DNA*), GenVision Pro	Windows: C:\Users\ <user>\DNASTAR Macintosh: ~/Library/ DNASTAR</user>
Preferences	ArrayStar	Windows: C:\Users\ <user>\AppData\ Roaming\DNASTAR\ ArrayStar</user>
	All others	Windows: C:\Users\ <user>\AppData\ Local\DNASTAR\ Macintosh: ~/Library/ Preferences</user>



AppData is a hidden folder in Windows. To unhide the folder, go to **Organize > Folder** and **Search Options > View > Show Hidden files and folders**.

Turn off usage logging

By default, <u>usage logging</u> is enabled in Lasergene version 11 and later. To opt out of usage logging on Windows, launch the DNASTAR Navigator and go to **View > Preferences > Lasergene**. On Macintosh, go to **DNASTAR Navigator > Preferences > Lasergene** (macOS 12 and below) or **DNASTAR Navigator > Settings** (macOS 13 and higher). Uncheck the box and click **Apply** and then **OK**.

Troubleshoot the error "A different version is running"

If you receive the error "A different version of Lasergene is running" when you try to launch a Lasergene application, but another version of Lasergene actually isn't running, there may be a lingering .state file from another version that needs to be deleted.

The solution is to locate and delete files named STARDM*.state, ...where '*' represents the version number.

These files may exist in the following directories:

Windows 7 and Windows 8:

- C:\Users\username\AppData\Local\DNASTAR\DataManager
- C:\ProgramData\DNASTAR\DataManager

Macintosh:

- Hard Drive:Library:Preferences:DNASTAR:DataManager
- · Hard Drive:Users:username:Library:Preferences:DNASTAR:DataManager

If you find any of the *.state* files listed above, delete them. Once the files are deleted, the error should no longer appear.

Clean cached files on startup

If you are a Windows user, you may execute a "-clean" in the unusual event that cached data has become corrupted:

- 1. Close GenVision Pro if it is currently open.
- 2. Launch the Command Line (C:\Windows\system32\cmd.exe).
- 3. Type the path to the GenVision Pro application, followed by a space and the -clean command.
 - C:\Program Files (x86)\DNASTAR\DNASTAR\Lasergene 15\GenVision Pro.exe -clean

Research references

Taylor WR (1997). "Residual colours: a proposal for aminochromography." Protein Engineering vol.10 no.7 pp.743 – 746. (See the <u>PDF file.</u>) ["Taylor" color scheme]

Waterhouse AM, Procter JB, Martin DMA, Clamp M and Barton GJ (2009). "JalView version 2: a multiple sequence alignment and analysis workbench." Bioinformatics 25:1189-1191. (See on <u>PMC</u>.) [Discusses methodology later used to create the "Taylor" and "Zappo" color schemes]