



Mutation *Surveyor*TM

DNA Variant Analysis Software

Version 2.5

Mutation Surveyor/Explorer[®] Updates of Version 2.5

We are pleased to inform you of the new functionalities and reporting capabilities that are included in the latest version of Mutation Explorer and Mutation Surveyor.

Each addition to the program is the result of needs requested and developed in collaboration with users such as yourself.

Our goal has been and remains to quickly respond to user's growing list of requirements, in order to assure continued improvement in the speed and accuracy of your discovery or diagnostic activity.

As always, we solicit and appreciate your comments on how we can continue to improve Mutation Surveyor and Explorer. Please do not hesitate to contact us.

Please take a few minutes to review the following short descriptions of these new enhancements, some we are sure will assist you in your research or diagnostic activities.

BioGene Ltd

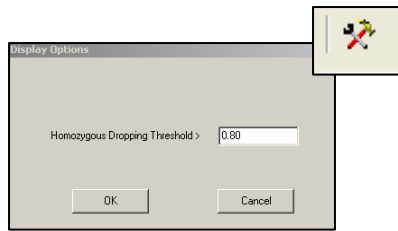
6 The Business Centre, Harvard Way
Kimbolton, Cambridgeshire
PE28 0NJ
0845 1300 950
www.BioGene.com
info@BioGene.com

Version 2.5 Updates:

- Improved Het-Indel alignment and sensitivity algorithms
- New tool for multiple primer data alignment
- Additional printing & reporting capabilities
- New & Enhanced "Whole Gene" mutational analysis tools
- Sequence text & consensus reporting
- Basecall scoring
- Basecall editing
- Automated downloading of "GenBank" sequence text
- Automated Mobility Shift correction
- Enhanced false positive Filter
- mRNA-cDNA nomenclature converter
- New genotype reporting
- Somatic & Germline comparison reporting
- High sensitivity detection of Mosaic peaks
- New report customization tool
- New FASTA file trace converter
- Easy addition of non-reported mutations to GBK files.

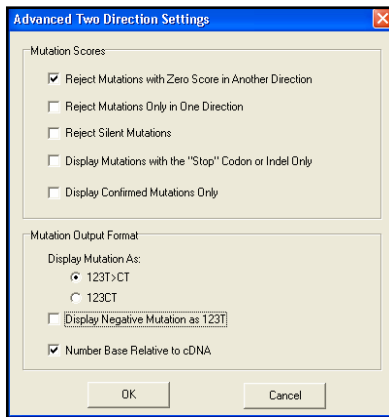
Paired 2 direction ORT

The **Copy Samplename Enabled** option allows the user to copy / paste specific file names by right clicking on the Sample File.



The **Display Options** icon activates a **Display Options** window. In this window users may alter the **Homozygous Dropping Threshold**. This value determine the level at which a mutation is identified as either homozygous or heterozygous.

Advanced 2 direction ORT

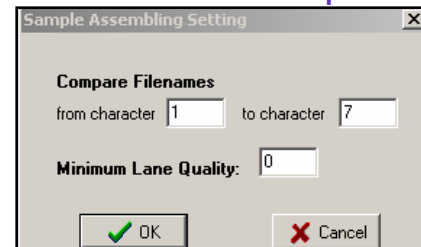


Due to the nature of electrophoresis, an artifact peak may show as real data. However, artifact peaks rarely occur at the same position in both forward and reverse. This rule is used to reject a mutation peak shown only in one direction. To eliminate mutations appearing in only one direction, **click** on the **Advanced 2D** icon. A dialogue box will activate with options for both **Mutation Scores** (reject with zero in other direction, reject mutations in only one direction, reject silent mutations, display mutations with a stop codon or indel only, display confirmed mutations only) and **Mutation Output Format** (display as, display negative mutation as, display mutation type, number base relative to cDNA). Check those additional filters that you wish to use. This will change the ORT by eliminating a number of mutations; those spaces will contain an "n.a" instead of the mutation code information.

Whole Gene Output Table

No.	Sample File	Gene	Exon	Ref	Start	End	Size	Qual	Mut#1	Mutation1	Mutation2		
1	H_AD-0003n>PCR2322_021n_F.ref	FLT3	3	96499	97067	568	17	1	n.a		97016C>A461		
2	H_AD-0003n>PCR2322_021n_R.ref	FLT3	3	96521	97068	548	12	1	n.a		97016C>A465		
3	H_AD-0003n>PCR2322_022n_F.ref	FLT3	3	96214	96700	487	19	0	n.a		n.a		
4	H_AD-0003n>PCR2322_022n_R.ref	FLT3	3	96293	96723	431	12	0	n.a		n.a		
5	H_AD-0003n>PCR2322_023n_F.ref	FLT3	3	94626	95146	521	19	1	86042G>C410		n.a		
6	H_AD-0003n>PCR2322_023n_R.ref	FLT3	3	94646	95170	524	10	1	86042G>C410		n.a		
7	H_AD-0016n>PCR2322_021n_F.ref	FLT3	3	96499	97044	546	12	1	n.a		97016C>A459		
8	H_AD-0016n>PCR2322_021n_R.ref	FLT3	3	96524	97068	545	9	1	n.a		97016C>A490		
9	H_AD-0016n>PCR2322_022n_F.ref	FLT3	3	96490	96624	135	6	0	n.a		n.a		
10	H_AD-0016n>PCR2322_022n_R.ref	FLT3	3	96230	96723	494	11	0	n.a		n.a		
11	H_AD-0016n>PCR2322_023n_F.ref	FLT3	3	94626	95146	521	19	1	86042G>C463		n.a		
12	H_AD-0016n>PCR2322_023n_R.ref	FLT3	3	94646	95170	523	10	1	86042G>C465		n.a		
13	H_AD-0020n>PCR2322_021n_F.ref	FLT3	3	96499	97044	546	13	1	n.a		97016C>A448		
14	H_AD-0020n>PCR2322_021n_R.ref	FLT3	3	96524	97068	545	10	1	n.a		97016C>A490		
15	H_AD-0020n>PCR2322_022n_F.ref	FLT3	3	96214	96700	487	15	0	n.a		n.a		
16	H_AD-0020n>PCR2322_022n_R.ref	FLT3	3	96290	96723	474	10	0	n.a		n.a		
17	H_AD-0020n>PCR2322_023n_F.ref	FLT3	3	94626	95140	515	12	1	86042G>C452		n.a		
18	H_AD-0020n>PCR2322_023n_R.ref	FLT3	3	94646	95170	523	12	1	86042G>C484		n.a		
18									8996	113	0.67	100.0%	100.0%

Tools within the Whole Gene Output Table



Along the top toolbar in this table there are 6 icons. The first 2, **Copy** and **Save**, are standard function. The third, **Display Options**, allows users to reject mutations present in only one trace, reducing false positives. The fourth icon, **Whole Gene Sample Assembling**, opens a **Sample Assembling Setting** window. This tool enables Mutation Explorer/Surveyor to perform its assembling function based on the location of a user-specified filename substring. All distinct substrings from all sample files in that character interval will be grouped. The output table displays a separate row for each distinct substring, including the number of mutations in all the files that contain that distinct substring and the mutation itself. Low quality lanes can be excluded by setting a **Minimum Lane Quality**. The range can be set from 0-100, any lane not meeting the quality setting will be excluded from the assembly.

No.	Sample	Mut#1	Mutation1	Mutation2
1	-0003	11	8325delA	8332delA
2	-000h	1	8452G>AG A>A/A\$1	
3	-0016	2	8452G>AG A>A/A\$1;8452G>A A>A/A\$3 DX	
4	-0020	2	8452G>A A>A/A\$14E (8452G>AG A>A/A\$1	
5	-0024	2	8284G>AG R>R/R\$1;8452G>AG A>A/A\$4	
6	-0025	1	8452G>AG A>A/A\$3	
7	-0033	2	8452G>A A>A/A\$105 (8452G>AG A>A/A\$2	
8	-0077	2	8452G>A A>A/A\$69 DX (8452G>AG A>A/A\$2	
9	-00rh	5	8326-8327insT	8452G>A D>N\$42.0I

EXAMPLE:

Continuing with the grouping used above, there were 13 distinct substrings in all the sample files starting from character 5 and going to character 9. The number of mutations for each substring is in the next column, and the mutations are shown in the remaining columns.

In the **Sample Assembling Table**, click **“Display in Matrix”** to view a Matrix Display of the mutations.

No.	Mutations	YCP009	YCP010	YCP011	Total# (% of Mutals)
1	26deA	YES	NO	NO	1 (33.33%)
2	229C>CT_p.R28RC	NO	YES	NO	1 (33.33%)
3	245G>AG_p.G33EG	YES	YES	NO	2 (66.67%)
4	372A>G_p.L75LL	YES	YES	NO	2 (66.67%)
5	372A>G_p.L75L	NO	NO	YES	1 (33.33%)
6	490A>G	YES	YES	NO	2 (66.67%)
7	490A>G	NO	NO	YES	1 (33.33%)
		4	4	2	

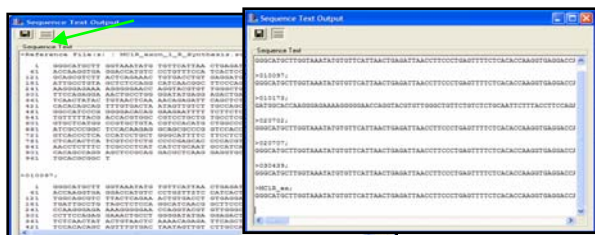
To view the alleles for each mutation, click the **“Display Alleles”** box. Both of these tables can be saved as text files (*.txt) and imported into Microsoft Excel for printing.

No.	Mutations	YCP009	YCP010	YCP011	Total# (% of Mutals)
1	26deA	YES	NO	NO	1 (33.33%)
2	229C>CT_p.R28RC	CC	CT	CC	1 (33.33%)
3	245G>AG_p.G33EG	AG	AG	GG	2 (66.67%)
4	372A>G_p.L75LL	AG	AG	AA	2 (66.67%)
5	372A>G_p.L75L	AA	AA	GG	1 (33.33%)
6	490A>G	AG	AG	AA	2 (66.67%)
7	490A>G	AA	AA	GG	1 (33.33%)
		4	4	2	

The fifth icon is **Consensus Sequence Text Output**. This tool allows users to search for and group sequence strings within a certain range; this range is set using the **Compare Sample** boxes to input the starting and ending base pairs. Users must enter the starting character and ending character for the location of the substring they want to use for grouping. **Compare Filenames** and **Minimum Lane Quality** work as discussed above for **Whole Gene Sample Assembling**.

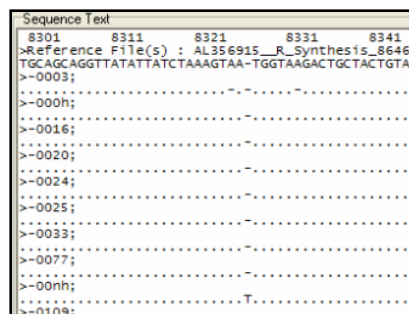
In the output table, the reference is shown at the top, and the sequence text for the base pair region selected is displayed below for each substring.

To view the sequence without numbering or spaces click the icon depicted below by the green arrow.



The sixth icon in this toolbar is **Sequence Text Comparison**. This tool requires the exact same parameters as the **Sequence Text** function discussed above; the settings are set and act the same. The difference lies in the output files. **Sequence Text Comparison** displays the reference sequence as a long row of characters, and then each group of distinct substrings is listed in a new row below the reference. The sample groups are compared with the reference sequence. A black DOT (.) indicates a match between the reference and sample. If there is a difference between any sample and the reference, the base pair in the sample is listed in that position.

A black dash (–) indicates that there is no base pair in a specific location.



Conventions for Het Indels: Inserted bases are indicated by the following coding:

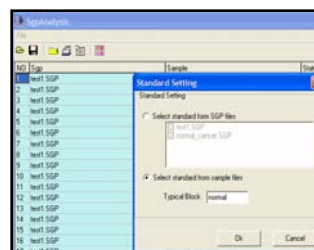
- Z= inserted A
- Q=inserted G
- U=inserted C
- I=inserted T

Text Output Report: Show Contig in Text Format

To view the electropherogram for specific mutations in all the samples, click the desired mutation indicated by the red arrow. The **Contig Trace Figure** window will appear which allows you to scroll through the samples at that mutation position (shown by the vertical grey line).



SGP Comparison Analysis



This tool was designed to provide users with a report that specifically compares normal samples with cancerous samples. The report is completely color-coded to simplify identification of inconsistent mutations between the standard and sample traces.

Prior to using the SGP Analysis tool, it is necessary for the user to create SGP files for comparison. For example one SGP file contains patient samples with normal cells and another SGP file contains cancerous cells from the same patients. The SGP Analysis tool provides a report which directly compares the mutation findings between SGP projects.

To activate the SGP Analysis Tool go to the **Tools** menu and choose **SGP Analysis**. Load the desired SGP projects by clicking the **Load Files** icon indicated by the green arrow. Next, group files by clicking the **Filename Assembly** icon indicated by the black arrow. Last, the user must designate the standard by clicking on the **Choose Standard** icon shown by the blue arrow. You may either choose an entire project as the standard or choose a character or string of characters for the standard.

Below is an example of an SGP Report comparing normal cells with cancer cells from patient samples. After choosing the correct standard as described above, click the **Compare Mutation Output** icon indicated above by the yellow arrow.

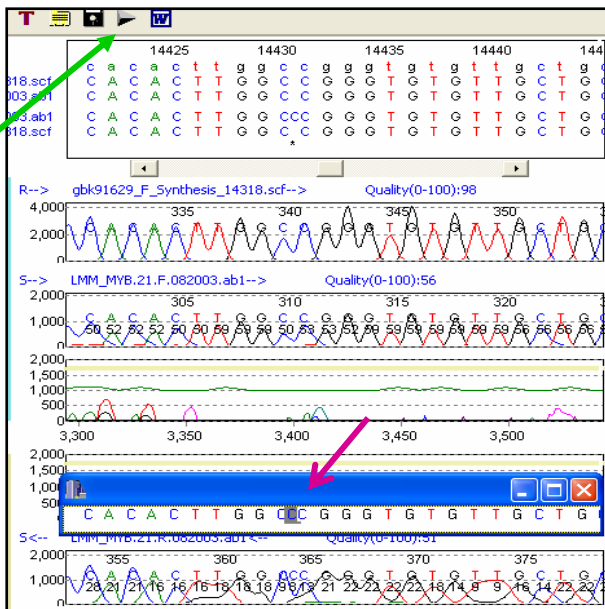
NO	Sgp	Sample	Mutation1	Mutation2	Mutation3	Mutation4	Mutation5	Mutation6	Mutation7	Mutation8	Mutation9
80	normal_cancer	96C_Z161.A435									
81	test1	SGP_507	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
82	test1	SGP_507	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
83	normal_cancer	507	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
84	normal_cancer	507	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
85	test1	SGP_510	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	255596A
86	test1	SGP_510	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	255596B
87	normal_cancer	510	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	255596B
88	normal_cancer	510	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	255596B
89	test1	SGP_512	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
90	test1	SGP_512	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
91	normal_cancer	512	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
92	normal_cancer	512	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
93	normal_cancer	ATC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
94	normal_cancer	ATC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
95	normal_cancer	ATC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
96	test1	SGP_ATC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
97	normal_cancer	ATC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
98	test1	SGP_ATC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
99	normal_cancer	GMC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
100	normal_cancer	GMC	23231.G111								

The light blue and light yellow backgrounds represent the standard (ex: normal samples) and the dark blue and yellow backgrounds represent the samples (ex: cancer samples). The mutations with a purple background are GenBank reported mutations, those with an orange background are inconsistent in the cancer samples and the mutations with green background are inconsistent in the standard (normal).

The user may export all samples with inconsistent mutations by click the icon shown below by the arrow.

Base-call Editing

To further minimize false positive mutation calls, the software has an option that allows the user to edit base calls in the electropherogram. In some cases the original base call is wrong and the sequence needs to be edited. In the example below there is an extra C which needs to be deleted. To edit the base, right click on the position and a window will open up as indicated by the red arrow. After editing the base, click the "Run" arrow indicated by the green arrow. To save the base call change in the trace file, click the save icon. The software will automatically correct the base call for similar patterns in the other sample traces.



Contig Option Elements

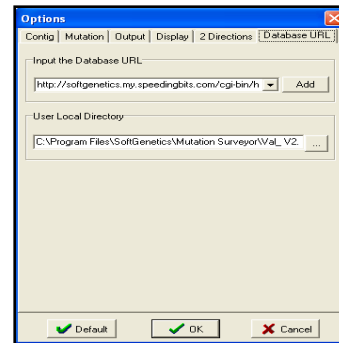
Base patch corrects for mobility shift within the sample traces
GenBank Reference Comparison: controls the files used for reference in a project. If this option is unchecked then the software will not use the GenBank sequence as a reference for comparison. Rather, the software will compare the sample traces with the reference files input in the second panel labeled "Reference Files." If this option is checked, the software will use the GenBank sequence as a reference for direct comparison to the sample traces.

Mutation Option Elements

Mutation Detection determines the level of sensitivity the software uses to call mutations. When high sensitivity is checked, the software can detect mutations in areas of high frequency and noise. If the "Allow Computer to Delete Mutations" box is checked then the software will delete illogical mutations and those with high background noise.

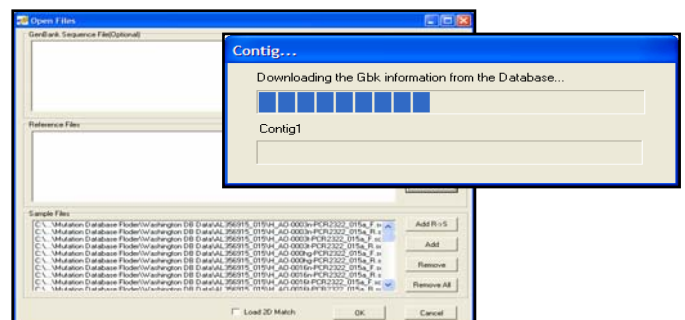
Database URL

The **Database URL** function streamlines the process by which users keep track of and enter human genomic sequences (GenBank files) stored in a reference database.



Input the Database URL allows users to type in, copy and paste in, or select a previously added URL for use as a reference file. There are two steps to using this function – entering URLs to the Options window and using it correctly when entering files. To designate a URL for a specific human genomic sequence, users need to open the **Database URL** window from the **Options** menu and then to type in, copy and paste in, or select a previously added URL for use as a reference file. Click on **OK** to load this information. When this URL database is set up correctly, select the **Open Files** icon and enter Sample Files as usual. Users should then leave the GenBank and Reference Files blank. Mutation Explorer/Surveyor will recognize the absence of a GenBank sequence and reference file and automatically connect to the online database through the stored URL to get the Genomic reference sequence.

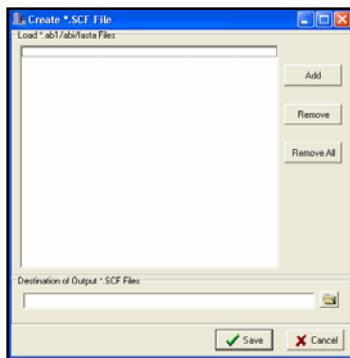
After clicking the Run button, the Genomic sequence will be downloaded from the online database through the selected URL.



User Local Directory: allows users to type in, copy and paste in, or select a previously created Filename Header. Under the Tools menu, there is an option, Print Header Editor, which allows the user to customize the header information to be printed on the mutation report. The saved header file can then be input into the **User Local Directory** window. The user must enter a personalized file or the software will use the default header. This will also save the gbk files in the directory.

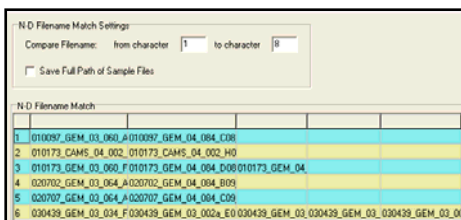
Converting between formats

Convert AB1/ABI/fasta to Scf (to transform trace files into a space saving format) Since AB1 trace files contain raw data and processed sequence trace data they are larger than Scf file which contain only the processed trace data. Converting AB1 files to Scf files saves data space. From the **Tools** menu item, select **Convert ABI/AB1/FASTA to SCF**. This opens a 2-part dialog window. Files to be converted are entered into the **Load ab1/abi/fasta Files** portion of the window. The destination for the new files is entered into the **Destination of Output Scf Files window**.



ND Filename Match Editor

The ND Filename Match Editor is similar in operation and function to the 2D Filename Match Editor, but it is used for projects with multiple primers.



Navigating within the Editor

Open displays a standard-format file entry window where users select several files. Once file names have been entered, click **Open** to begin the editor function. Mutation Explorer/Surveyor will separate the files into columns based upon the character comparison chosen.

Refresh will group files according to the designated number of matching characters. Files will be put into the same row as their match. This allows users to realign files manually if necessary.

Clear will clear all work from the window. This option should be used with caution as unsaved files can be lost.

Save will save the *.txt file created by this editor to a user-specified directory. This file location will be used for later entry of these files into a project.

Gbk File Editor Elements

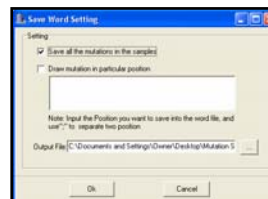
Feature 2 contains details about the gene (start and stop), coding sequence (CDS) and variation. CDS contains the following information: Join__Exons, Exon Start and Stop and Exon Size, Region of Interest Start and Stop and ROI size, gene, codon_start, product, protein_ID, db_xref, transition. Variation contains where the variation occurs and notes. You may manually input variations and save them within the gbk file. When this file is used as a reference, the added variations will be represented by a green tick mark above the mutation position in the mutation electropherogram.



Saving and Printing

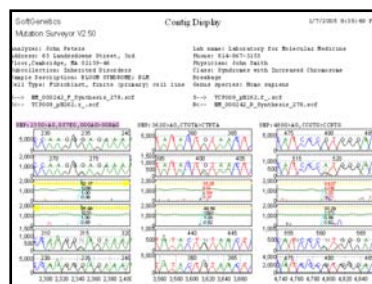
Saves the mutation screen into a Microsoft Word File.

To save the entire screen into a Microsoft Word document, set the screen as you would like it to be saved ie with or without nucleotides, amino acids, peak table, browser, and click the Microsoft Word icon. The following window will appear with options for saving the mutations and for the location of the Output File.



Clinical Report

You may also choose to print a Clinical Report, which is a custom-formatted display via the Print Sample icon located on the main toolbar. A clinical report is formatted such that the page header contains the sample information. The report prints all of the possible mutation points listed in the GenBank file and real mutations. The header information can be found under the **Tools** menu, **Print Header Editor**. The custom-formatted display presents just the area around the selected mutation with a user-set header (see Print Header Editor in Chapter 6 Editors for more on this tool).



Printing Heterozygous Indels

In order to print the electropherogram displaying heterozygous indels, open the Heterozygous Insertions and Deletions screen through the icon. The screen below shows the Heterozygous indel detection screen. To see a print preview click the Print Preview icon indicated by the red arrow, where you can then either Print or go to the Page set-up window.

