Tutorial

Find Actionable Variants With QIAseq DNA Panels

February 25, 2022

Sample to Insight

QIAGEN Aarhus · Silkeborgvej 2 · Prismet · 8000 Aarhus C · Denmark
digitalinsights.qiagen.com · ts-bioinformatics@qiagen.com

Find Actionable Variants With QIAseq DNA Panels

This tutorial uses the capabilities of the CLC Genomics Workbench and Biomedical Genomics Analysis plugin to find actionable variants - even at very low frequencies - in targeted sequencing data generated using a BRCA1 and BRCA2 QIAseq panel kit.

This tutorial covers in just a few steps all the following:

- Import Illumina paired reads in the workbench.
- Find low frequency variants with the Targeted DNA workflow via the **Analyze QIAseq Samples** guide.

Prerequisites For this tutorial, you must be working with CLC Genomics Workbench 22.0 with the Biomedical Genomics Analysis plugin installed. How to install plugins is described here: http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Install.html. For the last section, "Upload your results to QCI Interpret", you will need an active QCI Interpret subscription (not a trial license).

Import the reads

1. First, download the sample data from our website:

   http://resources.qiagenbioinformatics.com/testdata/QIAseq_DNA_tutorial.zip

   and unzip it in the location of your choice on your computer. Note that the folder contains more files than you will need to use, but that can be useful if you are working with older versions of the workbench and plugin.

2. Start the *CLC Genomics Workbench*.

3. Import the reads via the toolbar: **Import** () | **Illumina**

   - Select the two fastq paired files.
   - Under "Import files and options" (figure 1), ensure that:
     - **Paired reads** and **Discard read names** checkboxes are checked.
     - In the "Paired read orientation" section, ensure the **Paired-end (forward-reverse)** option is checked.
     - Set the **Minimum distance** to 1 and the **Maximum distance** to 1000 (default values).
     - Click **Next**.

4. Click **Next** and choose the folder you wish to save the reads to (you can create a new folder dedicated to this tutorial for example) before clicking **Finish**.

   The R1 and R2 files will be merged as a single file of paired reads during import.

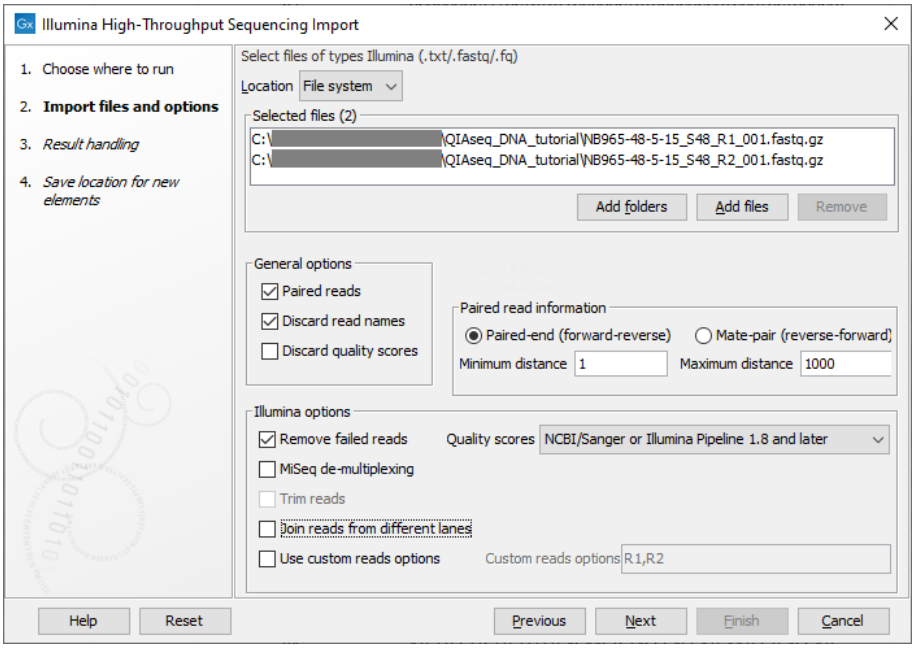


Figure 1: *Import the reads.*

Run the Targeted DNA Workflow

We will now run the BRCA1 and BRCA2 Panel (DHS-102Z) workflow with all the settings set to their default values.

1. Open the **Analyse QIAseq Samples** guide from:

   Toolbox | **Template Workflows** | **Biomedical Workflows** ( ) | **QIAseq Sample Analysis** ( ) | **Analyze QIAseq Samples** ( )

2. Select the "Targeted DNA" tab at the top (figure 2).

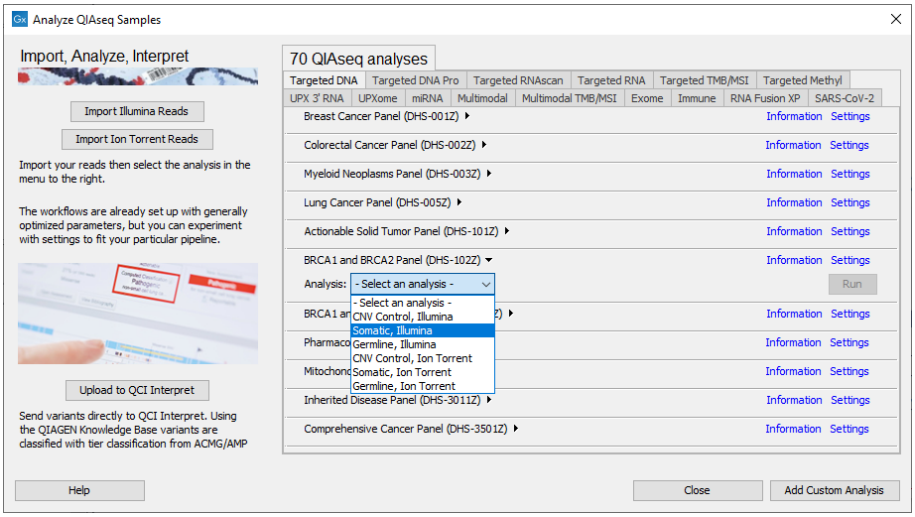


Figure 2: *Select the right configuration of the DHS-102Z workflow.*

3. Click on the arrow to the right of the BRCA1 and BRCA2 Panel (DHS-102Z) workflow, select the option **Somatic, Illumina**, and click **Run**. If you are connected to a *CLC Genomics Server*, you will be asked where to run the workflow. We recommend that you choose the *CLC Genomics Server* when possible.

4. The first dialog will require for you to acquire the relevant reference data set (figure 3). Click on **Download to Workbench**, then on **Next**. If this data is already available on your system you will simply not see this step.

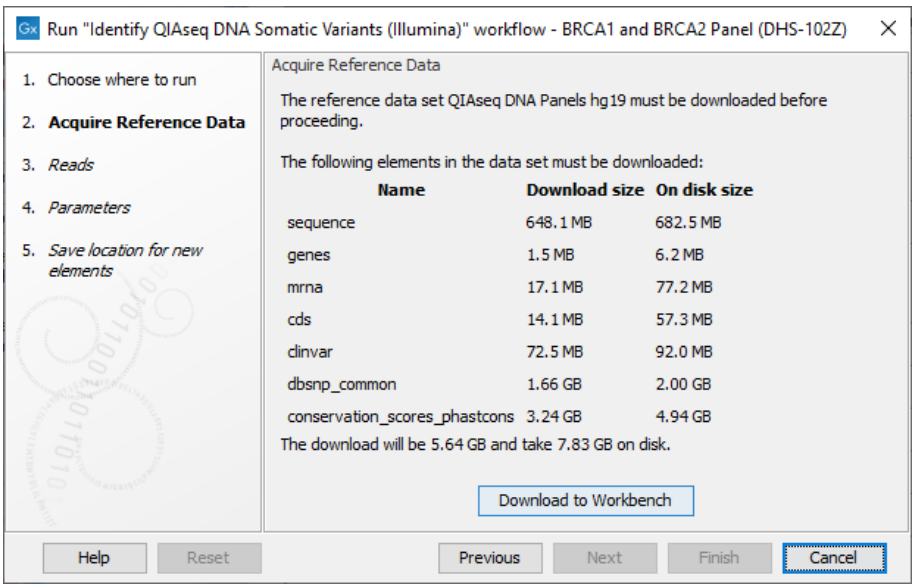


Figure 3: *Acquire reference data set by clicking Download to Workbench.*

5. Select the sequencing reads that should be analyzed (figure 4) and click **Next**.

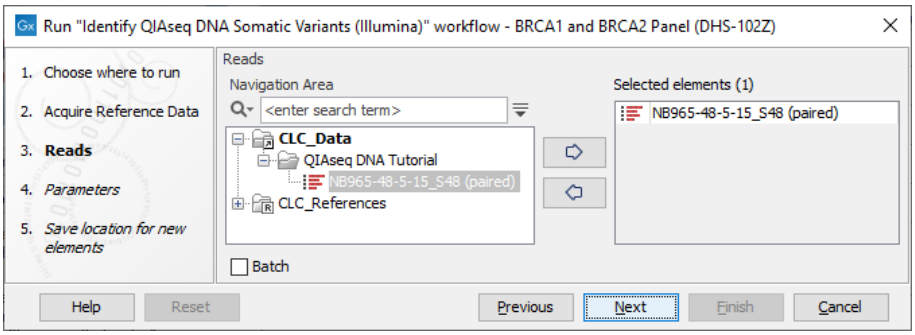


Figure 4: *After clicking Run, select the sequencing reads by double-clicking on the file name or by clicking once on the file name and then on the arrow pointing to the right hand side.*

6. In the next dialog (figure 5), keep the default values for the parameters and click **Next**.

7. In the final wizard step, choose to **Save** the results of the workflow and specify a location in the Navigation Area before clicking **Finish**.

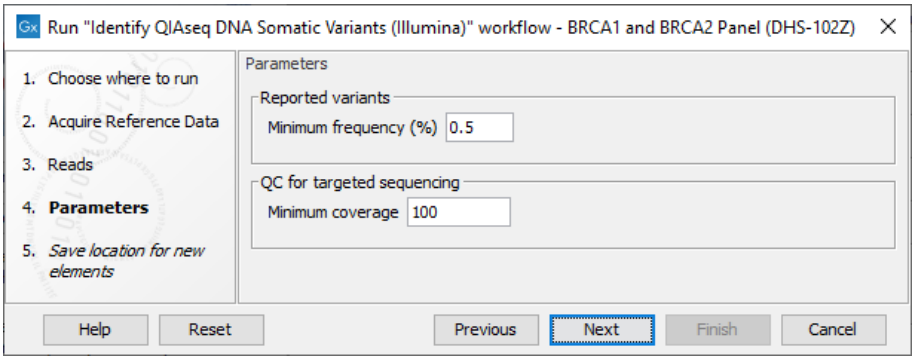


Figure 5: *Use the default values for parameters. These have been optimized for somatic variant detection.*

Output from the BRCA1 and BRCA2 Panel Analysis

The BRCA1 and BRCA2 Panel Analysis workflow produces the following outputs (most of which are available in a subfolder as seen in figure 6):

- a Genome Browser View ( )
- several reports ( ) of which some sections are also included in a Combined Report ( )
- a read mapping of the UMI Reads ( )
- several tracks: variant tracks before and after filtering, structural variant tracks and an amino acid track that allows identification of variants.

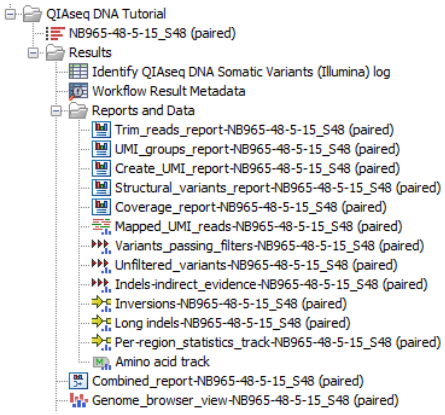


Figure 6: *Output from the BRCA1 and BRCA2 Panel Analysis workflow.*

In general, it is very important to inspect whether the data meets quality requirements such as coverage and the expected number of reads per UMI, but we do not go through these steps in this tutorial. To learn more about Quality Control for the Identify QIAseq DNA Variants see the manual: http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsanalysis/current/index.php?manual=Quality_Control_Identify_QIAseq_DNA_Variants_workflow.html.

Double-click on the Genome Browser view output. The file opens as a split screen with the list of tracks in one view and the filtered variant table below. In the table, filter for variants for which

the allele is not identical to the reference: open the advanced filter by clicking on the arrow in the top right corner of the table. Fill in the filtering fields with "**Reference allele**", "**=**" and type "**no**" (figure 7).

QIAseq panels allow the detection of very low frequency variants. To observe this is in the tutorial data, sort the "**Frequency**" column in the variant table in ascending order by clicking on the header of that column. Variants with a frequency ("Count" divided by "Coverage") even lower than 2 appear at the top of the table.

Clicking on a particular row in the variant table causes the Genome Browser to zoom in to the nucleotide level with the focus on that variant position. This allows you to view the coverage and the context of this variant looking e.g., at the Mapped UMI Reads, Amino acid track, and so on.

Chromosome	Region	Type	Reference	Allele	Reference ...	Length	Zygosity	Count	Coverage	Frequency
13	32911050^32911051	Insertion	-	AAAC	No	4	Heterozygous	2	376	0.53
13	32913419	SNV	G	C	No	1	Heterozygous	3	449	0.67
13	32911040	SNV	C	G	No	1	Heterozygous	3	377	0.80
17	41231352	SNV	T	A	No	1	Heterozygous	3	361	0.83
17	41256177	SNV	T	A	No	1	Heterozygous	4	475	0.84
13	32913454	Deletion	T	-	No	1	Heterozygous	2	218	0.92
13	32913848	SNV	A	T	No	1	Heterozygous	4	413	0.97
13	32911561	SNV	C	G	No	1	Heterozygous	4	363	1.10
13	32920993^32920994	Insertion	-	A	No	1	Heterozygous	6	449	1.34

Figure 7: *Track List.*

Upload your results to QCI Interpret

It is now possible to upload one of the variant tracks to QCI Interpret. For this section, you need an active QCI Interpret subscription.

1. To start the upload, go to:

   Toolbox | **Tools** | **QIAseq Panel Expert Tools** ( ) | **QCI Interpret Integration** ( ) | **Upload to QCI Interpret** ( )

   Alternatively, open the **Analyse QIAseq Samples** guide and click on **Upload to QCI Interpret**.

2. In the first dialog (figure 8), select the variant track containing variant passing filters, from which a VCF will be uploaded to QCI Interpret. You can also select structural variant tracks (e.g. Inversions and Long indels) but for this tutorial data there are no detected inversions or long indels.

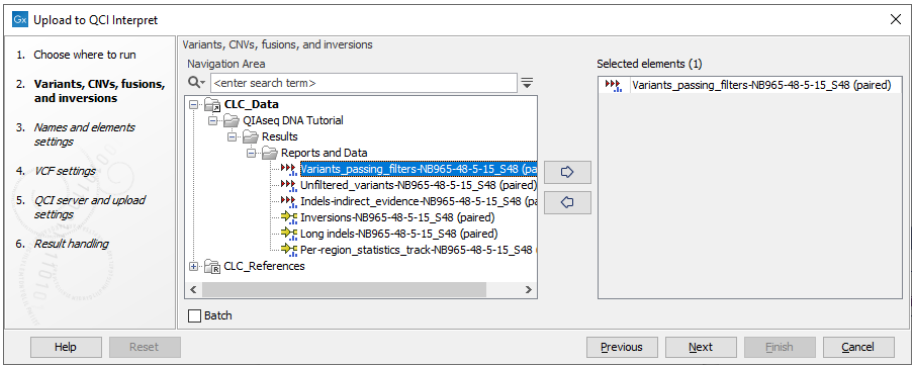


Figure 8: *Select the variants file you would like to to upload to QCI Interpret.*

3. In the next step it is optional to type in a Sample name, SubjectID and/or Project as seen within QCI Interpret. Select the reference sequence to hg19. This sequence can be found in the "Reference Data tab", under the "QIAGEN Active". Find here the QIAseq DNA Panels hg19 and under "sequence", the Homo_sapiens_sequence_hg19 (figure 9).

4. In the next step you can specify VCF export settings. For this upload, choose "Somatic" under "Prefill VCF Settings".

5. In the Server and upload settings dialog (figure 10), select your QCI region. Under QCI Interpret login, select "Browser" and click "Log in". This will open a new browser window where username and password associated with your QCI Interpret account can be typed in. When you are logged in, you will be told to return to the CLC Workbench.

6. Click on **Finish**.

Note: Clicking on **Finish** in the wizard is the start of the steps needed to transfer variants to QCI Interpret. The tool will output a report containing the link to the QCI Interpret sample list page (this page will also open when the tool is complete) (figure 11).

To analyze the uploaded data in QCI Interpret, create a new test selecting the uploaded sample and fill in the QCI Interpret Interview pages.

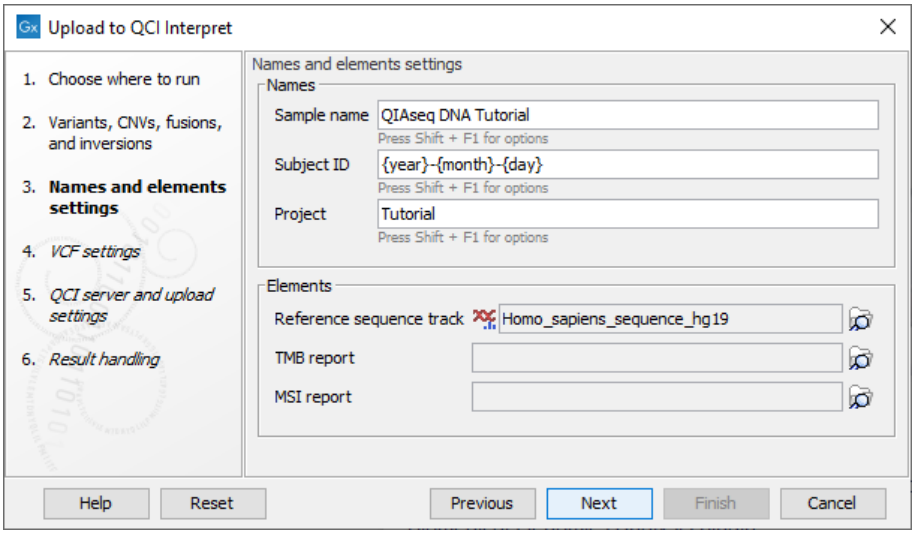


Figure 9: *Provide names for the QCI Interpret sample and select the reference sequence.*

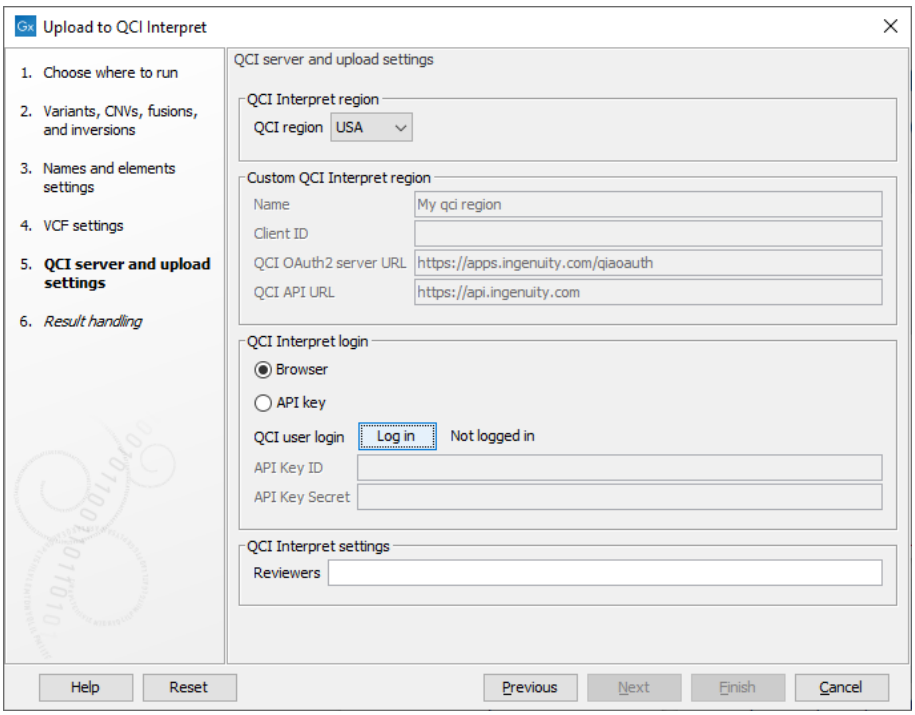
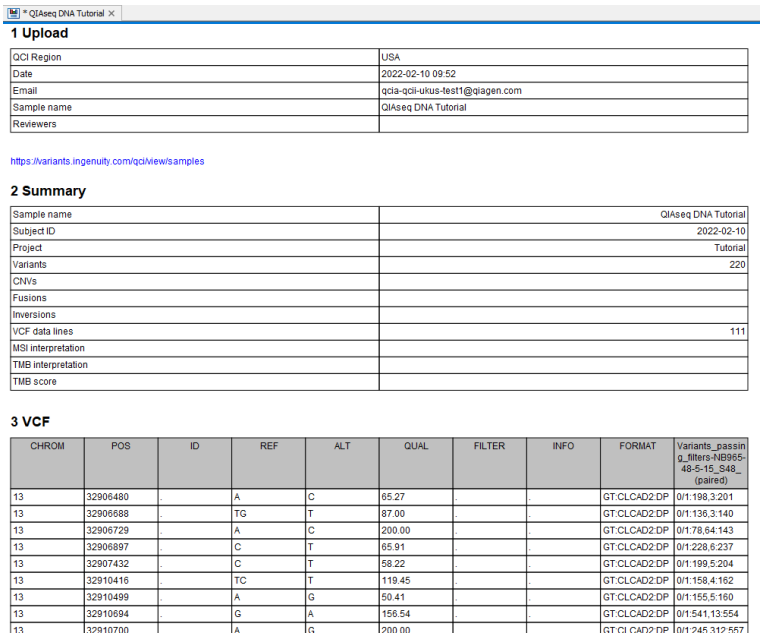


Figure 10: *Configure the connection to the QCI Interpret server dialog.*

Please refer to the QCI Interpret documentation to learn more about the interpretation of the results (figure 12).



* QIAseq DNA Tutorial

1 Upload

QCI Region	USA
Date	2022-02-10 09:52
Email	qcia-qcii-ukus-test1@qiagen.com
Sample name	QIAseq DNA Tutorial
Reviewers	

https://variants.ingenuity.com/qci/view/samples

2 Summary

Sample name	QIAseq DNA Tutorial
Subject ID	2022-02-10
Project	Tutorial
Variants	220
CNVs	
Fusions	
Inversions	
VCF data lines	111
MSI interpretation	
TMB interpretation	
TMB score	

3 VCF

CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	Variants_passing_filters-NB965-48-5-15_S48_ (paired)
13	32906480	.	A	C	65.27	.	.	GT:CLCAD2:DP	0/1:198,3:201
13	32906688	.	TG	T	87.00	.	.	GT:CLCAD2:DP	0/1:136,3:140
13	32906729	.	A	C	200.00	.	.	GT:CLCAD2:DP	0/1:78,64:143
13	32906897	.	C	T	65.91	.	.	GT:CLCAD2:DP	0/1:228,6:237
13	32907432	.	C	T	58.22	.	.	GT:CLCAD2:DP	0/1:199,5:204
13	32910416	.	TC	T	119.45	.	.	GT:CLCAD2:DP	0/1:158,4:162
13	32910499	.	A	G	50.41	.	.	GT:CLCAD2:DP	0/1:155,5:160
13	32910694	.	G	A	156.54	.	.	GT:CLCAD2:DP	0/1:541,13:554
13	32910700	.	A	G	200.00	.	.	GT:CLCAD2:DP	0/1:245,312:557

Figure 11: *Example of a report containing a link to the QCI Interpret sample list.*

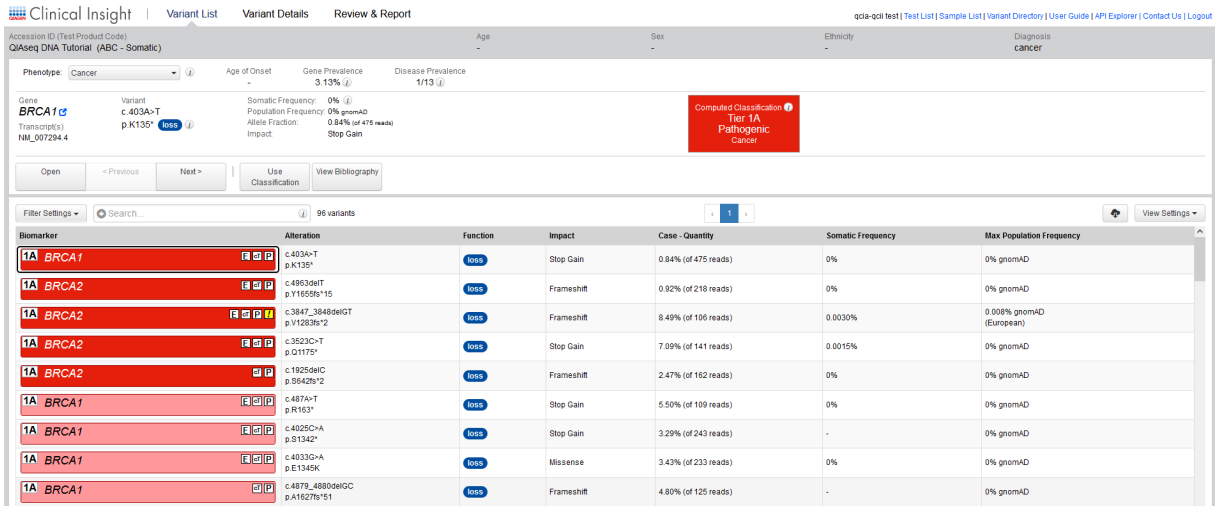


Figure 12: *Results of the interpretation analysis from the QCI Interpret website.*