

Find Actionable Variants With QIAseq DNA Panels

February 25, 2022

Sample to Insight

\_

QIAGEN Aarhus  $\cdot$  Silkeborgvej 2  $\cdot$  Prismet  $\cdot$  8000 Aarhus C  $\cdot$  Denmark digitalinsights.qiagen.com  $\cdot$  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: <a href="http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Install.html">http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Install.html</a>. 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 ((2) | 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.



🐼 Illumina High-Throughput S	equencing Import >
<ol> <li>Choose where to run</li> <li>Import files and options</li> </ol>	Select files of types Illumina (.txt/.fastq/.fq) Location File system v Selected files (2)
3. Result handling	C:\QIAseq_DNA_tutorial\\B965-48-5-15_548_R1_001.fastq.gz C:\QIAseq_DNA_tutorial\\B965-48-5-15_548_R2_001.fastq.gz
4. Save location for new elements	Add <u>fo</u> lders <u>A</u> dd files Remove
	General options         Paired reads         Discard read names         Discard quality scores         Paired -end (forward-reverse)         Mate-pair (reverse-forward)         Minimum distance         Minimum distance         Illumina options         Remove failed reads         Quality scores         NCBI/Sanger or Illumina Pipeline         1.8 and later         MiSeq de-multiplexing         Trim reads         Join reads from different lanes         Use custom reads options         Custom reads options
Help Reset	Previous Next Einish Cancel

Figure 1: Import the reads.

#### **Run the Targeted DNA Worklflow**

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).

mport, Analyze, Interpret	70 QIAseq analyses	
States and the Constant	Targeted DNA Targeted DNA Pro Targeted RNAscan Targeted RNA Targeted TMB/MSI	Targeted Methyl
	UPX 3' RNA UPXome miRNA Multimodal Multimodal TMB/MSI Exome Immune RNA	Fusion XP SARS-CoV-2
Import Illumina Reads	Breast Cancer Panel (DHS-001Z) >	Information Settings
Import Ion Torrent Reads	Colorectal Cancer Panel (DHS-002Z) >	Information Settings
oort your reads then select the analysis in the nu to the right.	Myeloid Neoplasms Panel (DHS-003Z) >	Information Settings
workflows are already set up with generally	Lung Cancer Panel (DHS-005Z) >	Information Settings
nized parameters, but you can experiment settings to fit your particular pipeline.	Actionable Solid Tumor Panel (DHS-101Z) >	Information Settings
	BRCA1 and BRCA2 Panel (DHS-102Z) 🔻	Information Settings
Compared Constant Pathogenic Networks of Constant	Analysis: - Select an analysis - V	Run
(in the year)	- Select an analysis - BRCA1 ar CNV Control, Illumina Somatic, Illumina	Information Settings
A REAL PROPERTY AND A REAL	Pharmaco Germline, Illumina	Information Settings
	CNV Control, Ion Torrent Mitochonc Somatic, Ion Torrent Germine. Ion Torrent	Information Settings
Upload to QCI Interpret	Inherited Disease Panel (DHS-3011Z) >	Information Settings
variants directly to QCI Interpret. Using QIAGEN Knowledge Base variants are ified with tier classification from ACMG/AMP	Comprehensive Cancer Panel (DHS-3501Z) >	Information Settings

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.

Gx Run "Identify QIAseq DNA S	Somatic Variants (Illumina)" workf	low - BRCA1 and	BRCA2 Panel (DHS-102Z)	×		
1. Choose where to run	Acquire Reference Data					
2. Acquire Reference Data	The reference data set QIAseq DNA Panels hg 19 must be downloaded before proceeding.					
3. Reads	The following elements in the data	set must be dowr	nloaded:			
A Downstown	Name	Download size	On disk size			
4. Parameters	sequence	648.1 MB	682.5 MB			
5. Save location for new elements	genes	1.5 MB	6.2 MB			
elements	mrna	17.1 MB	77.2 MB			
C.S.C.	cds	14.1 MB	57.3 MB			
(US)	dinvar	72.5 MB	92.0 MB			
The Carl State State	dbsnp_common	1.66 GB	2.00 GB			
0	conservation_scores_phastcons	3.24 GB	4.94 GB			
TO TO	The download will be 5.64 GB and take 7.83 GB on disk.					
and have a start						
	Download to Workbench					
Help Reset	Previous	Next	Finish Cancel			

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.

Gx Run "Identify QIAseq DI	VA Somatic Variants (Illumina)" workflow - BRCA1 and Reads	d BRCA2 Panel (DHS-102Z) X
1. Choose where to run	Navigation Area	Selected elements (1)
2. Acquire Reference Data	Q.▼ <enter search="" term=""> =</enter>	Image: NB965-48-5-15_S48 (paired)
<ol> <li>Reads</li> <li>Parameters</li> </ol>	CLC_Data C→ QLAseq DNA Tutorial C→ SUB965-48-5-15_S48 (paired) C→ Ch2 References	
5. Save location for new elements	Batch	
Help Reset		Next Einish Cancel

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**.



2.	Choose where to run Acquire Reference Data	Parameters           Reported variants           Minimum frequency (%)	
3.	Reads	QC for targeted sequencing	
4.	Parameters	Minimum coverage 100	
5.	Save location for new elements		

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 (W) 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: <a href="http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsanalysis/current/index.php?manual=Quality\_Control\_Identify\_QIAseq\_DNA\_Variants\_workflow.html">http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsanalysis/current/index.php?manual=Quality\_Control\_Identify\_QIAseq\_DNA\_Variants\_workflow.html</a>.

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.



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 ( $\widehat{kar}$ ) | QCI Interpret Integration ( $\widehat{kar}$ ) | Upload to QCI Interpret ( $\overline{kar}$ )

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

 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.

. Choose where to run	Variants, CNVs, fusions, and inversions Navigation Area		Selected elements (1)	
Variants, CIVs, fusions, and inversions     Names and elements settings     VCF settings     QCI server and upload settings     Result handling	Center search term>         QC       Center search term>         QLC_Data         QLAseq DNA Tutorial         QLAseq DNA Tutorial         Provide Search term>         QLAseq DNA Tutorial         Provide Search term>         Provide Search term> <td>paired) S48 (pa</td> <td></td> <td>ers-H8965-48-5-15_S48 (paired)</td>	paired) S48 (pa		ers-H8965-48-5-15_S48 (paired)
	Batch			

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.



Gx Upload to QCI Interpret			×
1. Choose where to run	Names and eleme	ents settings	
2. Variants, CNVs, fusions, and inversions	Sample name	QIAseq DNA Tutorial Press Shift + F1 for options	
3. Names and elements	Subject ID	{year}-{month}-{day} Press Shift + F1 for options	
settings 4. VCF settings	Project	Tutorial Press Shift + F1 for options	-1
5. QCI server and upload settings	Elements Reference sec	quence track 🎇 Homo_sapiens_sequence_hg19	ò
6. Result handling	TMB report	, in the second se	à
	MSI report		õ
Help Reset		Previous Next Finish Cancel	

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

Gx Upload to QCI Interpret			×
1. Choose where to run	QCI server and upload setti	ings	
2. Variants, CNVs, fusions, and inversions	QCI Interpret region QCI region USA ~	]	
<ol> <li>Names and elements settings</li> </ol>	Custom QCI Interpret reg	-	
4. VCF settings	Name Client ID	My qci region	
5. QCI server and upload settings		https://apps.ingenuity.com/qiaoauth	
6. Result handling	QCI API URL	https://api.ingenuity.com	
	QCI Interpret login     O API key     QCI user login     API Key ID     API Key ID     API Key Secret      QCI Interpret settings     Reviewers	n Not logged in	
Help Reset		Previous Next Einish Cance	

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).



VIAseq DNA Tut QCI Region 2022-02-10 09:52 Date Email Sample nar Reviewers qcia-qcii-ukus-test1@qiagen.com QIAseq DNA Tutorial https 2 Summary Sample name Aseq DNA Tuto Subject II Project Variants CNVs ibject ID Fusions Inversions VCF data lines MSI interpretat TMB interpreta TMB score 3 VCF FILTER CHRON POS REF ALT QUAL INFO Variants g\_filters 48-5-1 (pa GT:CLCAD2:DP 0/1:198,3:201 GT:CLCAD2:DP 0/1:136,3:140 GT:CLCAD2:DP 0/1:78,64:143 32906480 32906688 32906729 32906897 32907432 32910416 65.27 87.00 TG 200.00 65.91 58.22 119.45 GT:CLCAD2:DP 0/1:228,6:237 GT:CLCAD2:DP 0/1:199,5:204 GT:CLCAD2:DP 0/1:158,4:162 50.41 156.54 200.00 32910499 32910694 32910700 GT:CLCAD2:DF 0/1:155,5:16 GT:CLCAD2:DP 0/1:541,13:554 GT:CLCAD2:DP 0/1:245.312:557

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

Clinical Insight   Variant Ust Variant Details Review & Report other Control Up Log						
ccession ID (Test Product Code) IAseq DNA Tutorial (ABC - Somatic)		Age -		Sex -	Ethnicity -	Diagnosis cancer
Phenotype: Cancer - J	Age of Onset Gene Prevalence Disease Pr - 3.13% () 1/13					
Gene         Variant           BRCA1 C         c.403A>T           Transcript(s)         p.K135* (loss) ()           NM_007294.4         ()	Somatic Frequency: 0% () Population Frequency: 0% gnomAD Allele Fraction: 0.84% (of 475 wads) Impact Stop Gain			Computed Classification 7 Tier 1A Pathogenic Cancer		
Open < Previous Next >	Use View Bibliography Classification					
Filter Settings -	() 96 variants			- e - 1 - 5		• View Settings
Biomarker	Alteration	Function	Impact	Case - Quantity	Somatic Frequency	Max Population Frequency
1A BRCA1	E eT P c.403A>T p.K135*	loss	Stop Gain	0.84% (of 475 reads)	0%	0% gnomAD
1A BRCA2	E et P c.4963delT p.Y1655fs*15	loss	Frameshift	0.92% (of 218 reads)	0%	0% gnomAD
1A BRCA2	E et P // c.3847_3848delGT p.V1283fs*2	loss	Frameshift	8.49% (of 106 reads)	0.0030%	0.008% gnomAD (European)
1A BRCA2	C.3523C>T p.Q1175*	loss	Stop Gain	7.09% (of 141 reads)	0.0015%	0% gnomAD
1A BRCA2	c.1925delC p.S642fs*2	loss	Frameshift	2.47% (of 162 reads)	0%	0% gnomAD
1A BRCA1	E.eT.P c.487A>T p.R163*	loss	Stop Gain	5.50% (of 109 reads)	0%	0% gnomAD
1A BRCA1	E of P c.4025C>A p.S1342*	loss	Stop Gain	3.29% (of 243 reads)		0% gnomAD
1A BRCA1	E.er P c.4033G>A p.E1345K	loss	Missense	3.43% (of 233 reads)	0%	0% gnomAD
1A BRCA1	c.4879_4880delGC p.A1627fs*51	loss	Frameshift	4.80% (of 125 reads)		0% gnomAD

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