

QIAGEN GeneRead Panel Analysis Plugin

USER MANUAL

User manual for QIAGEN GeneRead Panel Analysis Plugin 1.12

Windows, Mac OS X and Linux

November 7, 2017

This software is for research purposes only.

QIAGEN Aarhus Silkeborgvej 2 Prismet DK-8000 Aarhus C Denmark



Contents

1	Introduction to QIAGEN GeneRead Panel Analysis Plugin	4
	1.1 Specify GeneRead DNAseq Gene Panel	5
	1.2 How to run the QIAGEN GeneRead Panel Analysis ready-to-use workflow	6
	1.3 Trim primers and their dimers of mapped reads	14
2	Install and uninstall plugins	19
	2.1 Install	19
	2.2 Uninstall	20

Chapter 1

Introduction to QIAGEN GeneRead Panel Analysis Plugin

QIAGEN GeneRead Panel Analysis Plugin is a ready-to-use workflow that can identify and annotate variants in Targeted Amplicon Sequencing data generated with GeneRead DNAseq Gene Panels. The GeneRead DNAseq Gene Panels can either be standard panels focused on a specific set of genes or can be customized to include genes tailored to specific research interests.

QIAGEN GeneRead Panel Analysis Plugin is bundled with target primers and target regions from QIAGEN GeneRead DNAseq Gene Panels 2.0, but it is also possible to use the QIAGEN GeneRead Panel Analysis Plugin if you are working with customized gene panels.

Also included is a tool for trimming primers and their dimers. This tool is described in section 1.3. The QIAGEN GeneRead Panel Analysis ready-to-use workflow and the Trim Primers and their Dimers of Mapped Reads tool are installed in the toolbox as illustrated in figure 1.1.



Figure 1.1: The workflow and tool are found in the toolbox.

The QIAGEN GeneRead Panel Analysis ready-to-use workflow covers all steps from read mapping to annotation of the variants and therefore performs both secondary and tertiary analysis.

The first step in the ready-to-use workflow is mapping of the sequencing reads to the human reference sequence. This is followed by a local realignment step, which is included to improve the variant detection that follows directly after a primer trimming step. After variant detection, the variants are annotated with gene names, exon numbers, amino acid changes, conservation scores, information from relevant variants present in the ClinVar database, and information from common variants present in the common dbSNP, HapMap, and 1000 Genomes database. Furthermore, a detailed target regions mapping report is created that allows inspection of the coverage and mapping specificity in the target regions.

The QIAGEN GeneRead Panel Analysis Plugin assumes that the sequences used as input do not contain adapters as the removal of adapters is often done directly on the sequencing machine. If adapters have not been trimmed off, please do so before proceeding with your analysis by using the Trim Reads tool (from the "Prepare Raw Data" folder) with the "Automatic read-through adapter trimming" option enabled.

1.1 Specify GeneRead DNAseq Gene Panel

Before running the QIAGEN GeneRead Panel Analysis ready-to-use workflow, you must first specify which GeneRead DNAseq Gene Panel has been used for targeted sequencing. In cases where you are using a customized panel, you need to have imported and saved the primers and the targeted regions files in the Navigation Area of the workbench before hand. To learn how to import the primer file, see http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsworkbench/current/index.php?manual=Import_Primer_Pairs.html. To import the target regions file, simply use the Standard Import tool.

To specify the relevant GeneRead DNAseq Gene Panel, go to:

Toolbar | Data Management (骨)

This will open the wizard shown in figure 1.2.

First select the **hg19 Reference Data Set** and click on the button labeled Download. When the download is over, click on the button Apply. Then select **QIAGEN GeneRead Panels hg19** and click on the button labeled Download. When the download is over, click on the button Apply.

Two extra folders are now in the CLC_References/homo_sapiens folder: "target_primers" and "target_regions" (see figure 1.3). Each folder contains elements specific to each commercially available QIAGEN GeneRead Panels kit.

To create a Reference Data Set specific to one panel in particular, click on Create Custom Set. This opens a pop up window where both "Target Primers" and "Target Regions" are represented by a drop down menu (figure 1.4). Select the option "Custom" to open another window where you can select the relevant panel from the CLC_References folder.

It is possible to select multiple target primers or target regions simultaneously if you work with multiple GeneRead DNAseq Gene Panels. Similarly, if you are using customized GeneRead DNAseq Gene Panels you can choose your customized primers and target regions provided that you had saved them before in your Navigation Area. In any case, remember that only the gene panels that are selected under "Data Management" will be available when you run the QIAGEN

ANJ YELYELETTIN.	HATSING POINGTAGTCCCTTTA	AGr.	ACIAGEN'com	Workspi	ace Plugins Data Manage	ement World
Bx Manage Reference Data for Installed Workflows						×
QIAGEN Reference Data Library Reference Data Sets				Manage Re Free space on CLC_Re Free space on tempora	ference Data: Locally ferences location: 43.5 ary folder location: 43.5	6 GB 6 GB
	QIAGEN GeneRea Version: 1, Reference D	id Panels hg19 lata Set	Export	3 Download Dela	Size on disk	
Hg38 Ensembl v81, db5NP v142, ClinVar 20150901				Dominous Dec	(and (app))	
Image: hg38 Ensembl v80, dbSNP v142, ClinVar 20150629	Reference Data included: Reference Data	Version	Download Size	On Disk Size	Applied	
	Target Primers	qiagen_v2.01	1.3 MB	6.9 MB	Yes	
hg19 Ensembl v74, dbSNP v138, ClinVar 20131203	Affects these workflows (1):	qiagen_V2.01	223 ND	1012 NB	162	
QIAGEN GeneRead Panels hg 19 Ensembl v74	Does not affect these workfl	ows (70):				•
Help					Close	

Figure 1.2: Open the Data Management and download QIAGEN GeneRead Panels hg19.



Figure 1.3: The folders "target_primers" and "target_regions" are available in your CLC_References data folder.

GeneRead Panel Analysis ready-to-use workflow.

Once you have selected the target primers and regions, do not forget to edit the name of your Custom Data Set before saving it. The new data set now appears under the Custom Reference Data Set tab of the Data Management window (figure 1.5). Click on Apply before you close the "Manage Reference Data" wizard. You can always go back and make changes if necessary.

1.2 How to run the QIAGEN GeneRead Panel Analysis ready-to-use workflow

The QIAGEN GeneRead Panel Analysis ready-to-use workflow can be found in the toolbox under "Targeted Amplicon Sequencing":

	Create Custom Reference	e Data Set		×
	Name	QIAGEN GeneRead Panels hg19		
	Organism	homo_sapiens		
	Chromosomal Extension	Full human genome		
	Annotation Type	Includes target regions and primers		
	Reference:		Version:	Selected data:
	Target Primers		qiagen_v2.01	15 elements selected
	Target Regions		qiagen_v2.01	15 elements selected
			qiagen_v2_chr_17	
Select feature tracks			qiagen_v2.01_chr_17	
Navigation Area	Sele	cted elements (1)	giagen_v2_chr_21	
NGHS-102X_Human_BRCA1_	→\$ NGHS-102X Human BRCA1 and BRCA ▲ → S NGHS-003X Human Myeloid Leukemia 495 NGHS-102X Human Breat and BRCA1 and BRCA ▲			
a target_regions	agen v2.01 dr 22			
Gigen_v2.01	Cancer -]
NGHS-002X_Human_Colorec	tal_Cancer			
NGHS-003X_Human_Myeloid	Leukemia 🗸 🗐			
·				Save Cancel
<pre>Q+ <enter search="" term=""></enter></pre>	_			
		-		
	OK X Cancel			

Figure 1.4: Select the relevant target primers and target regions and click "OK".

Manage Reference Data fo	or Installed Workflows						×
QIAGEN Reference Data Library	Custom Reference Data Sets				Free Free	Manage Refi space on CLC_Refer space on temporary	erence Data: Locally - ences location: 259.65 GB folder location: 259.65 GB
Reference Data Sets O QIAGEN GeneRead Panels hg19 - human myeloid leukemia Ensembl v74			QIAGEN GeneRea Version: 1, Reference D	nd Panels h Pata Set	ng19 - human myek	pid leukemia	Size on disk
		Create Custom Set	Maariaa	Download Size	On Dials Gine	Apply	
			Target Primers	Custom	N/A	-	No
		-	Target Regions	Custom	N/A	-	No
Help							Close

Figure 1.5: The newly created Custom Reference Data Set can be seen under the Custom Reference Data Set tab. Do not forget to edit the data set'a name and to apply it before starting the workflow.

Toolbox | Ready-to-Use Workflows | Targeted Amplicon Sequencing () | Somatic Cancer (TAS) () | QIAGEN GeneRead Panel Analysis ()

1. Double-click on the QIAGEN GeneRead Panel Analysis ready-to-use workflow (figure 1.1) to run the analysis.

If the QIAGEN GeneRead Analysis Plugin has been installed on a CLC Server you are connected to via your Workbench, then you will be asked where you would like to run the analysis. We recommend that you run the analysis on a CLC Server when possible. Click **Next**.

2. Select the sequencing reads that should be analyzed (figure 1.6).

If you would like to analyze more than one sample you can choose to run the analysis in batch mode. This is done by ticking "Batch" in the lower left side of the wizard and selecting the folder(s) that holds the data you wish to analyze. If you have your sequencing data in separate folders, you should choose to run the analysis in batch mode. You can learn more about batch analysis in the Biomedical Genomics

	Navigation Area	Selected elements (1)
2. Select sequencing reads	Whole_genome_sequencing targeted_amplicon_sequencing example_set demo gene_panel_prostate_cancer results2 iii MS+3, S3, L001, R1, 001 (paired) Q* Batch	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

Figure 1.6: 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.

Workbench user manual (see http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsworkbench/current/index.php?manual=Batch_processing.html).

3. In the next window, specify which of the available 1000 Genomes populations to use in the analysis by clicking on the browse folder symbol ()) on the right-hand side of the wizard. Figure 1.7 shows the default settings where all three available 1000 Genomes populations are selected. This is the default setting if all three populations have been selected under Data Management as described in the Biomedical Genomics Workbench user manual (http://resources.qiagenbioinformatics.com/manuals/ biomedicalgenomicsworkbench/current/index.php?manual=Download_configure_ reference_data.html).

	QIAGEN GeneRead Panel	al Analysis	
 Choose where to run Select sequencing reads 	1000 Genomes population	000GENOMES_phase_1_AMR, 1000GENOMES_phase_1_EU	JR 🝺
3. 1000 Genomes population	Navigation Area	Select Workflow Input Selected elements (3)	
MINAR MINI		Cancel OK Previous Next Finish Canc	el

Figure 1.7: The 1000 Genomes population(s) that are selected and available as the default choice in the wizard are the population(s) that have been selected under Data Management. To remove populations that are not relevant for this analysis, click on the populations that is to be deselected and click on the arrow pointing to the left-hand side.

If you have selected only one population in the Data Management (in this example the European population), this wizard step will be skipped.

- 4. In the Map Reads to Reference wizard step (figure 1.8), you can configure the read mapper by setting the "Cost of insertions and deletions" to either "Affine gap cost" (default) or "Linear gap cost".
 - **Linear gap cost** The cost of a gap is computed directly from the length of the gap and the insertion or deletion cost. This model often favors small, fragmented gaps over long contiguous gaps.
 - Affine gap cost An extra cost associated with opening a gap is introduced such that long contiguous gaps are favored over short gaps.



Figure 1.8: In this wizard step you can set the "Cost of insertions and deletions" to either "Affine gap cost" (default) or "Linear gap cost".

5. In the next wizard window (figure 1.9), you can restrict the calling of InDels and Structural Variants to the targeted regions only. All available gene panels are selected as the default choice. By clicking on the plus symbol (+) on the right-hand side of the wizard it is possible to adjust the number and type of gene panels to use as target regions and thereby restrict the variant calling to only the sequences that have been targeted in your sequencing experiment. Click on the button labeled Next when you are done.

• •	QIAGEN GeneRead Panel Analysis
 Choose where to run Select sequencing reads 1000 Genomes population Map Reads to Reference InDels and Structural Variants 	InDels and Structural Variants Configurable Parameters Restrict calling to target regions NGHS-002X_Human_Colorectal_Cancer Locked Settings
?	Previous Next Finish Cancel

Figure 1.9: In this wizard step you can specify the targeted regions matching your read mapping.

- 6. Specify the target primers for primer trimming in the **Trim Primers and their Dimers of Mapped Reads** window (figure 1.10). If you would like to add more GeneRead DNAseq Gene Panel target primers, this can be done using "Data Management" as described in section 1.1. It is also possible to either enable or disable the parameter "Only keep reads that have hit a primer". Note that it is enabled by default.
- 7. In the **Low frequency Variant Detection** wizard step (figure 1.11), you can specify the parameters for variant detection.

• •	QIAGEN GeneRead	Panel Analysis
1. Choose where to run	Trim Primers and their Dimers of Mappe	ed Reads
2. Select sequencing reads	Primer track	NGHS-002X_Human_Colorectal_Cancer
3. 1000 Genomes population	Only keep reads that have hit a prime	er 🧭
4. Map Reads to Reference	 Locked Settings 	
 InDels and Structural Variants 		
6. Trim Primers and their Dimers of Mapped Reads		
?		Previous Next Finish Cancel

Figure 1.10: Select the primer track from the drop-down list.

	QIAGEN GeneRead Panel Anal	ysis			
Choose where to run	Low Frequency Variant Detection				
. choose where to run	Configurable Parameters				
2. Select sequencing reads	Required significance (%)	1.0			
3. 1000 Genomes population	Ignore positions with coverage above	1,000,000,000			
. Map Reads to Reference	Restrict calling to target regions	Selected 15 elements.			
. InDels and Structural	Ignore broken pairs				
Variants	Ignore non-specific matches	Reads			
. Trim Primers and their Dimers of Mapped Reads	Minimum read length	20			
Low Frequency Variant	Minimum coverage	10			
Detection	Minimum count	2			
	Minimum frequency (%)	5.0			
	Base quality filter				
	Read direction filter				
	Direction frequency (%)	5.0			
	Relative read direction filter				
	Significance (%)	1.0			
Of a	Read position filter				
	Significance (%)	1.0			
- Sanda	Remove pyro-error variants				
0	In homopolymer regions with minimum length	3			
10	With frequency below	0.8			
10 Manuture	Locked Settings				
? 5	Prev	ious Next Finish Cancel			

Figure 1.11: In this wizard step the parameters for variant detection can be adjusted.

Please see the Biomedical Genomics Workbench user manual for a description of the different parameters that can be adjusted in the variant detection step. A description of the "Low Frequency Variant Detection" tool can be found here: http://resources. qiagenbioinformatics.com/manuals/biomedicalgenomicsworkbench/current/index.php?manual=Low_Frequency_Variant_Detection.html.

8. In the QC for Target Sequencing wizard step (figure 1.12), you must specify your target region. As the default choice, all available target region tracks that were selected under Data Management are selected. The number and kind of target regions can be adjusted at this step. Note that as we already described for the "Trim Primers and their Dimers of Mapped Reads" wizard step, the options you get in the list that is accessed via the plus symbol (♣) in this window are the target regions that were selected under "Data Management".

You can also specify:

• Minimum coverage This will be used to provide the length of each target region that

٠		QIAGEN Ger	neRead Panel Analysis
1. 2. 3. 4. 5. 6. 7. 8.	Choose where to run Select sequencing reads 1000 Genomes population Map Reads to Reference InDels and Structural Variants Trim Primers and their Dimers of Mapped Reads Low Frequency Variant Detection QC for Target Sequencing	QC for Target Sequencing Configurable Parameters Track of Target Regions Minimum coverage Ignore non-specific matches Ignore broken pairs Locked Settings	NGHS-002X_Human_Colorectal_Cancer 30 2 2
	? 🥱		Previous Next Finish Cancel

Figure 1.12: Specify your target regions and adjust the parameters if desired.

has at least this coverage.

- Ignore non-specific matches and/or broken pairs When these are applied reads that are non-specifically mapped or belong to broken pairs will be ignored.
- 9. The next two wizard window, Add Information from 1000 Genomes Project and Add Information from Hapmap, you can specify the populations that fit your dataset. Indeed, detected variants are annotated with a range of different data in this ready-to-use workflow, but for databases that provide data from more than one population (i.e., HapMap and the 1000 Genomes Project), the populations relevant to the data set can be specified by the user.

First, the variants are annotated with information from the 1000 Genomes Project (see figure 1.13).

• •	QIAGEN GeneRead Panel Analysis	
1. Choose where to run	Add Information from 1000 Genomes Project	
2. Select sequencing reads	Configurable Parameters Known variants track 1000GENOMES_phase_1_EUR	φ
 3. 1000 Genomes population 4. Map Reads to Reference 	Locked Settings	
5. InDels and Structural Variants		
Trim Primers and their Dimers of Mapped Reads		
7. Low Frequency Variant Detection		
8. QC for Target Sequencing		
9. Add Information from 1000 Genomes Project		
? 9	Previous Next Finish Can	cel

Figure 1.13: Select the relevant population from the list or use all three populations that have already been selected.

From the list that can be accessed by clicking on the plus symbol (\clubsuit) you can choose the population that matches the population your samples are derived from. Please note that the populations available from the drop-down list can be specified with the Data Management (\bigcirc) function found in the top right corner of the Workbench.

Note! Under "Locked settings" you can see that "Automatically join adjacent MNVs and SNVs" has been selected. The reason for this is that many databases do not report a succession of SNVs as one MNV as is the case for the Biomedical Genomics Workbench,

and as a consequence it is not possible to directly compare variants called with Biomedical Genomics Workbench with these databases. In order to support filtering against these databases anyway, the option to *Automatically join adjacent MNVs and SNVs* is enabled. This means that an MNV in the experimental data will get an exact match, if a set of SNVs and MNVs in the database can be combined to provide the same allele. This assumes that SNVs and MNVs in the track of known variants represent the same allele, although there is no evidence for this in the track of known variants.

10. Click on the button labeled Next and do the same to annotate with information from HapMap (figure 1.14).



Figure 1.14: Select the relevant population from the list or use all populations that have already been selected.

11. Finally, in the last wizard step (shown in figure 1.15), pressing the button Preview All Parameters allows you to preview all parameters (see figure 1.16), but to make any changes, you must use the button Previous and Next to reach the relevant wizard window. If no change is necessary, choose to save the results and click **Finish**.

Output from the QIAGEN GeneRead Panel Analysis The QIAGEN GeneRead Panel Analysis tool produces seven different outputs:

- Target region reads track with the locally realigned trimmed reads (=)
- Target region coverage track ()
- Coverage report (<u></u>)
- Amino Acid Changes track (M)
- Annotated variant track (
)
- Genome Browser View (
- Log table (

• • •	QIAGEN GeneRead Panel Analysis
1. Choose where to run 2. Select sequencing reads	Result handling
3. 1000 Genomes population	
 Map Reads to Reference InDels and Structural Variants 	Workflow parameters
 Trim Primers and their Dimers of Mapped Reads 	Preview All Parameters
7. Low Frequency Variant Detection	Result handling Open
8. QC for Target Sequencing	O Save
 Add Information from 1000 Genomes Project Add Information from 	Log handling Open log
HapMap 11. Result handling	
? 4	Previous Next Finish Cancel
_	

Figure 1.15: Check the selected parametes by pressing "Preview All Parameters".

Note! We advise you to not delete any of the produced files individually as some of them are linked to each other. If you would like to delete an experiment, please always delete all of generated files from one experiment at the same time.

When looking at the results of the analysis, a good place to start is the target region coverage report (\Rightarrow) to see whether the coverage is sufficient in the regions of interest (e.g. >30). Please also check that at least 90% of the reads are mapped to the human reference sequence and that the majority of the reads map to the targeted region.

Open the Genome Browser View file (11) to get an overview of the identified variants (see 1.17).

The Genome Browser View includes the annotated variants in context to the human reference sequence, genes, transcripts, coding regions, targeted regions, mapped sequencing reads, relevant variants in the ClinVar database as well as common variants in common dbSNP, HapMap and 1000 Genomes databases. Finally, a track with conservation scores shows the level of nucleotide conservation around each variant.

The conservation scores are based on a multiple alignment with a range of different vertebrates. The conservation in the region around each variant is particularly relevant when you consider the potential importance of the individual variants. A high conservation score could indicate that the variant is located in a region of the genome that is of great importance.

The annotated variant track can also be shown in table view. To open the table, double-click on the name of the variant track in the left side of the Genome Browser View (when opened in the View Area). The annotated variant table includes all variants and the added information/annotations (see 1.18).

In figure 1.19 the annotated variant table and the Genome Browser View are shown in split view. The annotated variant table and the Genome Browser View are connected and when selecting a variant from the table by clicking on a row in the table, the Genome Browser View will automatically put the selected variant into focus. In figure 1.19 the "Zoom to base level"

Dels and Structural Variants		
P-Value threshold 0.	0001	
Maximum number of mismatches 3		
Filter variants		
Minimum number of reads 2		
Restrict calling to target regions 👘	NGHS-002X_Human_Colorectal_Cancer	
ocal Realignment		
Realign unaligned ends	\checkmark	
Multi-pass realignment	2	
Guidance-variant track	is defined by InDels and Structural Variants	
Force realignment to guidance-varia	ts	
Primer track	Ra NGHS-002X_Human_Colorectal_Cancer	
Reference	Prove Homo_sapiens_sequence_hg19	
Minimum primer overlap length	9	
Allow dangling 3' end base	v	
Minimal primer overlap fraction	0.5	
Only keep reads that have hit a prime	r 🗸	
Additional bases to trim	2	
ow Frequency Variant Detection		
Required significance (%)	1.0	
Ignore positions with coverage above	1,000,000,000	
Export to:	Tab delimited text (.txt)	

Figure 1.16: Preview all parameters. At this step it is not possible to introduce any changes, it is only possible to view the settings.

function (1), marked with a red arrow in the lower right corner of the View Area, has been used to zoom in on the variant.

The added information can support identification of candidate variants for further research. For example common genetic variants (present in the HapMap database) or variants known to play a role in drug response or other relevant phenotypes (present in the ClinVar database) can easily be singled out using the table.

Also, identified variants that are unknown in the ClinVar database can be for example prioritized based on amino acid changes. A high conservation level on the position of the variant between many vertebrates or mammals can also be a hint that this region could have an important functional role, with variants with a conservation score of more than 0.9 (PhastCons score) that should be prioritized higher. Filtering of the variants based on their annotations can be facilitated using the table filter in the top right side of the table.

Please note that in case none of the variants are present in ClinVar or dbSNP, the corresponding annotation column headers are missing from the result.

1.3 Trim primers and their dimers of mapped reads

The Trim Primers and their Dimers of Mapped Reads tool is used in the QIAGEN GeneRead Panel Analysis ready-to-use workflow. It is also provided as a separate tool to be used for targeted



Figure 1.17: Genome Browser View to inspect identified variants in the context of the human genome and external databases.

Rows: 33	Table \	/iew: Genon	ne										Filte	.r 🔤 🦷
Chromosome	Region	Туре	Reference	Allele	Reference	Length	Zygosity	Count	Coverage	Frequency	Probability	Forward re	Reverse re	Forw
1	160786670	SNV	Α	G	No	1	Heterozygous	292	560	52.14	1.00	292	292	
1	160786670	SNV	Α	A	Yes	1	Heterozygous	268	560	47.86	1.00	268	268	
3	178922274	SNV	С	Α	No	1	Heterozygous	303	572	52.97	1.00	303	302	
3	178922274	SNV	С	С	Yes	1	Heterozygous	267	572	46.68	1.00	267	266	
3	178941854	Deletion	т	-	No	1	Heterozygous	523	1018	51.38	1.00	523	523	
3	178941854	SNV	т	т	Yes	1	Heterozygous	494	1018	48.53	1.00	494	494	8
4	46329655	SNV	A	т	No	1	Heterozygous	596	1193	49.96	1.00	596	596	
4	46329655	SNV	A	Α	Yes	1	Heterozygous	591	1193	49.54	1.00	591	591	
4	46329723	SNV	т	Α	No	1	Heterozygous	597	1193	50.04	1.00	596	597	
4	46329723	SNV	т	т	Yes	1	Heterozygous	593	1193	49.71	1.00	591	593	
5	152697706	SNV	С	т	No	1	Heterozygous	151	467	32.33	1.00	151	151	
5	152697706	SNV	С	С	Yes	1	Heterozygous	316	467	67.67	1.00	316	316	
в	128750597	SNV	A	G	No	1	Heterozygous	406	791	51.33	1.00	406	406	
В	128750597	SNV	A	Α	Yes	1	Heterozygous	384	791	48.55	1.00	384	384	
9	21968199	SNV	С	G	No	1	Homozygous	145	145	100.00	1.00	145	145	
10	89692732	Deletion	т	-	No	1	Heterozygous	46	442	10.41	1.00	46	46	
10	89692732	SNV	т	т	Yes	1	Heterozygous	395	442	89.37	1.00	395	395	
12	6945914	SNV	С	G	No	1	Heterozygous	134	225	59.56	1.00	134	134	
 III 														•

Figure 1.18: The annotated variant track opened in table view from the Genome Browser View. The table makes it easier to inspect identified variants in detail.

amplicon sequencing experiments with many targets (and as a consequence many primers). To be able to trim off the primers used in your sequencing experiment you must know the primer sequences as you will need to specify which target primer sequence file to use.

The Trim Primers and their Dimers of Mapped Reads not only trims off primers but also makes use of the primer pairs in the trimming process to predict and trim possible primer dimerizations. The prediction is based on the primer pairs, the reference, and user settings that are described later in this section. Removal of primers and their dimers from the mapped reads ensures that no bias is introduced in the variant calling as would be the case if the primers and dimers were considered to be part of the sequencing reads.



Figure 1.19: The annotated variant table and the Genome Browser View shown in split view.

When running the Trim Primers and their Dimers of Mapped Reads tool, primers of the reads are trimmed first. The tool then looks for primer dimerization artifacts and trim them.

Primer trimming The user can specify the fraction of the primer that must overlap with a read's aligned bases in order to record a primer hit.

The Trim Primers and their Dimers of Mapped Reads is strict regarding primer position, i.e., if a primer begins after the read at the 5' end, it will not be considered for trimming. The read's unaligned bases are taken into account. For example, if the primer begins two positions before the read's first aligned base and the read has three unaligned bases, the primer is said to begin after the read. Similarly, a primer that ends before the read at the 3' end is not considered. Again unaligned bases are taken into account.

Primer dimer trimming The primer dimer trimming is done in two steps.

In the first step, all primers are compared against each other for possible primer dimerization. The user may specify the minimum number of bases that needs to bind for primers to dimerize and amplify. After the first step, a list of possible primer dimerizations have been compiled for each primer.

In the second step, the actual trimming is performed. All reads are examined, and if the read was trimmed by a primer, *p*, and the read starts with the sequence predicted by one of *p*'s possible primer dimerizations, it is assumed that the read has a primer-dimer artifact. The tool proceeds to trim the read so the artifact is unaligned. In the case where the read only consists of the primer-dimer artifact sequence, the read will be discarded.

The Trim Primers and their Dimers of Mapped Reads can be found in the toolbox:

Toolbox | Resequencing Analysis (a) | Trim Primers and their Dimers of Mapped Reads (a)

Select read mapping track In the first wizard step (figure 1.20), you are asked to select the read mapping. If you would like to analyze more than one read mapping, you can choose to run the analysis in batch mode by ticking the "Batch" box in the lower left corner of the wizard and then selecting the folder that hold the read mappings you want to analyze.

Trim Primers and their Dim	ers of Mapped Reads	—
1. Choose where to run	Select read mapping track Navigation Area	Selected elements (1)
2. Select read mapping track	the genome_sequencing targeted_amplicon_sequencing demo demo demo gene_panel_prostate_cancer for results2 units2 uni	▲ M54-3_S3_L001_R1_001 (paired) (Re
? (S)	Batch	Next / Einish X Cancel

Figure 1.20: Select files to import.

Specify trim parameters In the next wizard step (see figure 1.21), specify the parameters for the trimming tool.

CR Trim Primers and their Dim	ers from Mapping	×
1. Choose where to run	Specify trim parameters Primer trim parameters	
2. Select read mapping track	Primer track	Ĵ
3. Specify trim parameters	 Only keep reads that have hit a primer 	
00000000000000000000000000000000000000	Primer dimer trim parameters Reference Minimum primer overlap length 9 IV Allow dangling 3' end base	õ
Party of the second sec	Other parameters Additional bases to trim 0	
?	← <u>Previous</u> → <u>N</u> ext ✓ Einish X <u>C</u> ancel	

Figure 1.21: Select your primer location file and choose whether you want to keep or discard reads with no matching primers.

• Primer trim parameters

- Primer track Click on the folder icon on the right-hand side of the wizard to select your primer location file.
- Minimal primer overlap fraction Specifies the fraction of the primer that must overlap with the read's aligned bases in order to record a primer hit. Setting the fraction to 0.0 will disable this requirement.
- Read handling configuration If you tick "Only keep reads that have hit a primer", reads with no matching primers will be discarded.
- Primer dimer trim parameters
 - Reference Click on the folder icon on the right-hand side of the wizard to select your reference location file.
 - Minimum primer overlap length The minimum number of bases that needs to bind for primers to dimerize and amplify.
 - Allow dangling 3' end base If you tick "Allow dangling 3' end base", a mismatch is allowed in the primer dimerization at the 3' end.
- Other parameters
 - Additional bases to trim This number of nucleotides will be trimmed off a read right after the primer. This trimming is not done on reads for which primer dimer artifacts were identified.

Result handling In the last wizard window (figure 1.22), choose to save the result of the primer trimming and click **Finish**.

CR Trim Primers and their Dim	ers of Mapped Reads
1. Choose where to run	Result handling
2. Select read mapping track	
3. Specify trim parameters	Cutput options
4. Result handling	
	Result handling
6	Open
Cler C)	Save
1001	Log handling
and a second sec	Open log
ALCONTRACTOR	
ANDRIA.	
?	← Previous → Next ✓ Einish X Cancel

Figure 1.22: Select output options.

Output of the Trim Primers and their Dimers of Mapped Reads tool The output is a read file from which primers and their dimers have been trimmed off. The name of the output is the one from the original input file with "trimmed reads" appended to it. In the last wizard step it is also possible to save a track with the primer dimers that were used to trim reads. The track contains information on why the primer dimer was predicted and the number of times it was used to partially trim a read or remove a read. A read is removed if the read only consists of the primer dimer.

Chapter 2

Install and uninstall plugins

QIAGEN GeneRead Panel Analysis is installed as a plugin.

Note: In order to install plugins and modules, the Workbench must be run in administrator mode. On Linux and Mac, it means you must be logged in as an administrator. On Windows, you can do this by right-clicking the program shortcut and choosing "Run as Administrator".

Plugins are installed and uninstalled using the plugin manager.

Help in the Menu Bar | Plugins... (😫) or Plugins (😫) in the Toolbar

The plugin manager has two tabs at the top:

- Manage Plugins. This is an overview of plugins that are installed.
- Download Plugins. This is an overview of available plugins on QIAGEN Aarhus server.

2.1 Install

To install a plugin, click the **Download Plugins** tab. This will display an overview of the plugins that are available for download and installation (see figure 2.1).



Figure 2.1: The plugins that are available for download.

Select QIAGEN GeneRead Panel Analysis to display additional information about the plugin on

the right side of the dialog. Click **Download and Install** to add the plugin functionalities to your workbench.

Accepting the license agreement

Part of the installation involves checking and accepting the end user license agreement (EULA) as seen in figure 2.2.

IND USER LICENSE AGREEMENT FOR QIAGEN AARHUS SOFTWARE. Bisulfite Sequencing 1.2.0	-
Recitals	Ε
.1 This End-User License Agreement ("EULA") is a legal agreement between you (either an individual person or a single legal ntity, who will be referred to in this EULA as "You") and QIAGEN Aarhus A/S, CVR-no.: 28 30 50 87 for the software woducts that accompanies this EULA, including any associated media, printed materials and electronic documentation (the Software Product").	
.2 The Software Product also includes any software updates, add-on components, web services and/or supplements that JAGEN Aarhus may provide to You or make available to You after the date You obtain Your initial copy of the Software Product to the extent that such items are not accompanied by a separate license agreement or terms of use. By installing, opying, downloading, accessing or otherwise using the Software Product, You agree to be bound by the terms of this EULA. If fou do not agree to the terms of this EULA, do not install, access or use the Software Product.	
.3 A static license to the software product allows the software product to be used locally on one specific computer system. When using a static license only a single instance of the software product can be running at any given time.	
.4 The software product is for research purposes only.	
SOFTWARE PRODUCT LICENSE	
.1 The Software Product is protected by intellectual property laws and treaties. The Software Product is licensed, not sold.	
2.2 You may install and use one copy of the Software Product on one single computer, device, workstation, terminal, or other ligital electronic or analogue device ("Device").	
.3 You are not licensed to do any of the following:	
.3.1 You may not sell, license or distribute copies of the Software Product on a stand-alone basis or as part of any collection, wordnet or service of any kind	-
I accept these terms	

Figure 2.2: Read the license agreement carefully.

Please read the EULA text carefully before clicking in the box next to the text **I accept these terms** to accept. If requested, fill in your personal information before clicking **Finish**.

If QIAGEN GeneRead Panel Analysis is not shown on the server but you have the installer file on your computer (for example if you have downloaded it from our website), you can install the plugin by clicking the **Install from File** button at the bottom of the dialog and specifying the plugin *.cpa file saved on your computer.

When you close the dialog, you will be asked whether you wish to restart the workbench. The plugin will not be ready for use until you have restarted.

2.2 Uninstall

Plugins are uninstalled using the plugin manager:

Help in the Menu Bar | Plugins... (😫) or Plugins (😫) in the Toolbar

This will open the dialog shown in figure 2.3.

The installed plugins are shown in the **Manage plugins** tab of the plugin manager. To uninstall, select QIAGEN GeneRead Panel Analysis and click **Uninstall**.

If you do not wish to completely uninstall the plugin, but you do not want it to be used next time you start the Workbench, click the **Disable** button.

P M Manage Plugins			
Provider: QIAGEN Aa Support contact: sup Version: 1.0 beta 7 (Peak finding using shape info	hape Tools Plugin [Beta] arhus port-dcbio@qiagen.com Build: 170213-1058-156573) rmation		▲ E
		Uninstall	Disable
Provider: QIAGEN Aa Support contact: sup Version: 1.5 (Build: 1	arhus port-dcbio@qiagen.com 70220-1616-157004)		
Rename files in batch by add	ing a prefix or a number.	Uninstall	Disable
Provider: Support contact: sup Version: 1.2 (Build: 1	ng port-ddbio@qiagen.com 70222-1036-157093)		
Tools to analyse bisulfite-con	verted reads.		-
Help Proxy Settings	Check for Updates	Install from File	Close

Figure 2.3: The plugin manager with plugins installed.

When you close the dialog, you will be asked whether you wish to restart the workbench. The plugin will not be uninstalled until the workbench is restarted.