Tutorial

Fusion Detection Using QIAseq RNAscan Panels

November 21, 2017

Sample to Insight

QIAGEN Aarhus · Silkeborgvej 2 · Prismet · 8000 Aarhus C · Denmark
Telephone: +45 70 22 32 44 · www.qiagenbioinformatics.com ·
AdvancedGenomicsSupport@qiagen.com

Fusion Detection Using QIAseq RNAscan Panels

This tutorial uses the capacities of the Biomedical Genomics Workbench and the QIAseq Targeted Panel Analysis plugin to detect fusion events in sequences generated using a QIAseq Targeted RNAscan Panel. Note that the plugin includes an Analyze QIAseq Panels guide that facilitates the configuration and the running of QIAseq panels secondary analysis workflows, but we choose to show you in this tutorial how to proceed without using the guide.

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

- Import Illumina paired reads in the workbench.
- Generate a trim adaptor list to remove the sequencing primers off the reads.
- Find fusion events with the **Detect QIAseq RNAscan Fusions** ready-to-use workflow.

This tutorial makes use of a subset (reads mapping only to chromosome 10) of the sequencing reads obtained using a QIAseq Targeted RNAscan Panel.

Prerequisites

For this tutorial, you must be working with the Biomedical Genomics Workbench 5.0 or higher. You must also have installed the QIAseq Targeted Panel Analysis plugin version 1.0 or higher. How to install plugins is described here: http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsworkbench/current/index.php?manual=Installing_plugins.html.

Download and import data

Go through the following steps to download and import the data into the Workbench.

1. Download the sample data from our website: http://resources.qiagenbioinformatics.com/testdata/Fusion_Gene_Detection_tutorial.zip and unzip it in the location of your choice on your computer.
2. Start the *CLC Genomics Workbench*.
3. Import the reads via the toolbar: **File** | **Import** () | **Illumina**
    - Select the two fastq paired files.
    - Under "General Options" section, ensure the **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 on the button labeled **Save** in the wizard page that appears, choose the folder you wish to save the reads to (you can create a new folder dedicated to this tutorial for example) and click **Finish**.

Data management configuration

In order to do this tutorial, you need to use the Reference Data Set made specifically for this tutorial:

1. Click on the Data Management icon in the top right corner of the workbench (1) (figure 1).
2. In the QIAGEN Reference Data Library (2), open the Tutorial Reference Data Sets tab (3), select the **Fusion Detection Using QIAseq RNAscan Panels** data set, and click on the button **Download** (4).

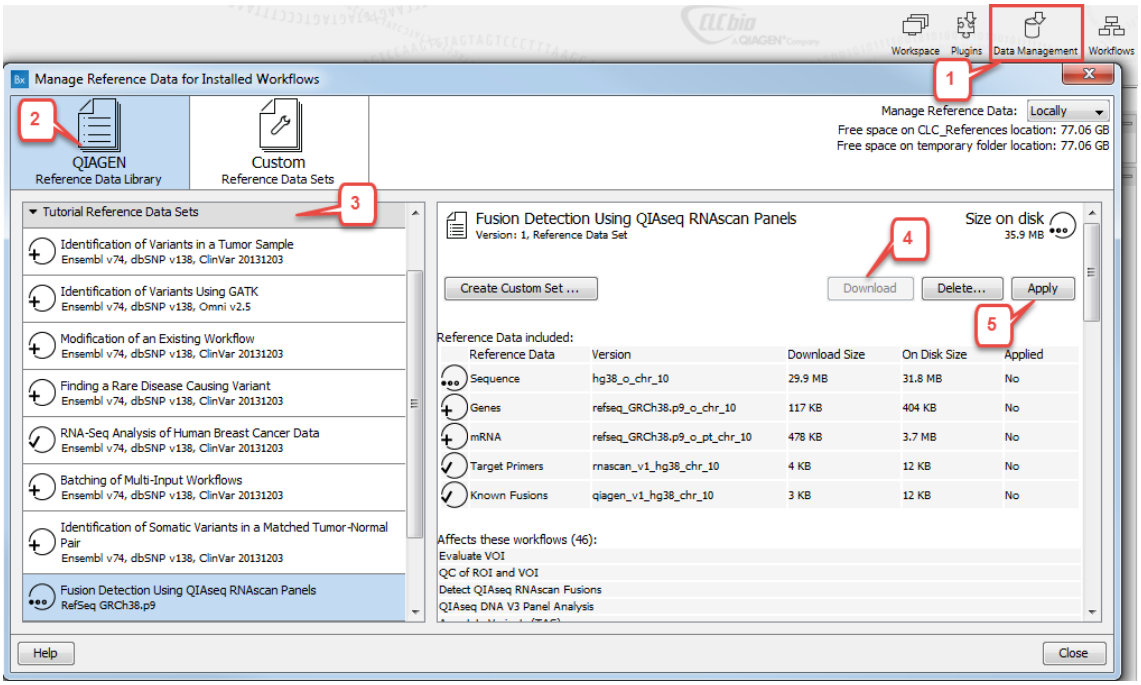


Figure 1: *Downloading the reference data.*

3. When the data is downloaded (characterized by check icons next to each item included in the dataset), click "Apply" (5).

The workflow is now set up to work with the adequate references. In addition, QIAseq Targeted RNAscan Panels specific primer tracks are now available in the CLC_References folder located in the Navigation Area.

Prepare your Adaptor trim list

As remaining adaptor sequences in the reads might lead to bias in downstream data analysis, we recommend to trim them off using a Trim adapter list. Trim adapter lists are most easily generated using the "Set up adapters" feature from the Analyze QIAseq Panels interface, as the Targeted RNAscan template already includes a Poly A, Poly G and the QIAseq Universal sequences.

1. Open Analyze QIAseq Panels here:

   Ready-to-Use Workflows | **QIAseq Panel Analysis** | **Analyze QIAseq Panels**

2. Click on Set up adapters (figure 2).

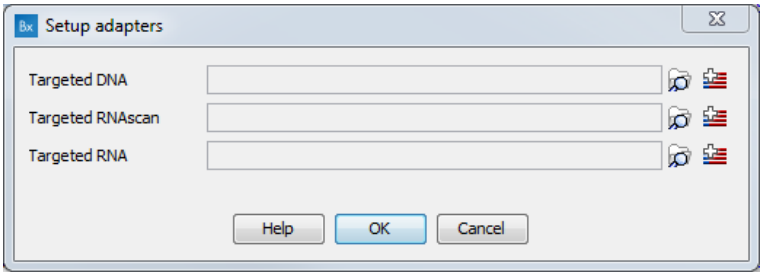


Figure 2: *Application-specific trim adapter lists.*

3. Click on the Trim adapter list button to the right of the Targeted RNAscan field.

4. Then follow the instructions given in the dialog (figure 3), i.e., insert the Nextera transposase adapter sequences (Read 1 and Read 2) by copying them from page 12 of the Illumina Customer Sequence Letter found at this address: http://support.illumina.com/downloads/illumina-customer-sequence-letter.html. [1]

   Read 1 (5'-3'): copy and paste here Nextera Transposase Adapter Read 1 sequence
   Read 2 (5'-3'): copy and paste here Nextera Transposase Adapter Read 2 sequence

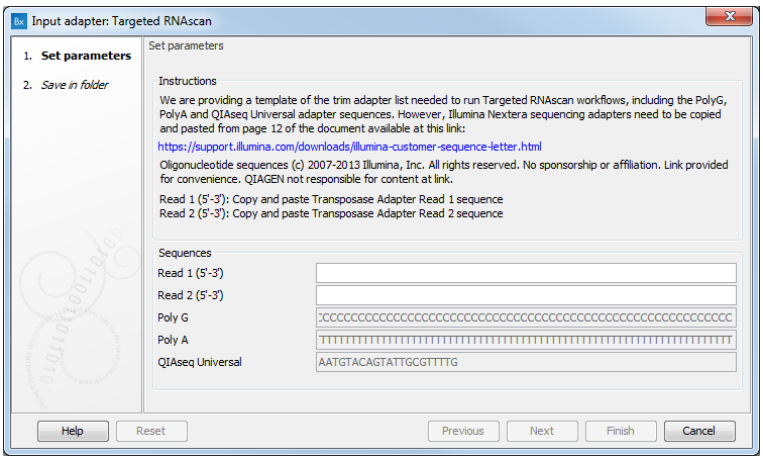


Figure 3: *RNAScan specific trim adapter list template.*

5. Click Next, and select a location in the Navigation Area where you want the trim adapter list to be saved before clicking Finish. Remember to save the list on the server if this is where you intend to run the workflows.

6. In the Set up adapters window, click on the Browse button and select the Trim adapter list you just saved.

Run Detect QIAseq RNAscan Fusions

The Detect QIAseq RNAscan Fusions ready-to-use workflow can be found in the Toolbox here:

[1] Oligonucleotide sequences © 2007-2013 Illumina, Inc. All rights reserved. No sponsorship or affiliation. Link provided for convenience. QIAGEN not responsible for content at link.

Toolbox | Ready-to-Use Workflows | QIAseq Panel Analysis | QIAseq Analysis Workflows | Detect QIAseq RNAscan Fusions

1. Double-click on the Detect QIAseq RNAscan Fusions ready-to-use workflow to run the analysis.

   If the QIAseq Targeted Panel Analysis plugin has been installed on a CLC Server you are connected to via the workbench, 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.

2. Select the sequencing reads that should be analyzed (figure 4).

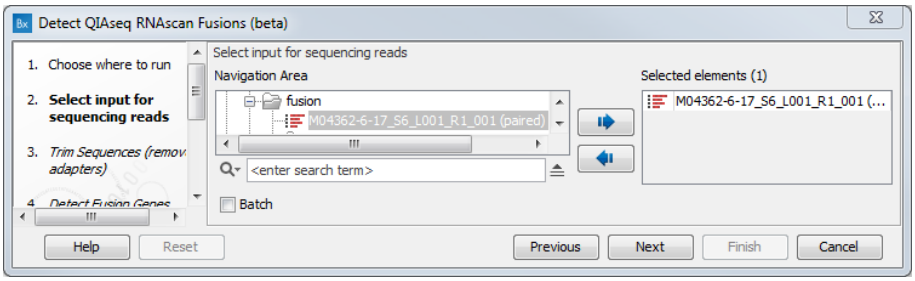


Figure 4: *Select the sequencing reads by double-clicking on the file name or by clicking once on the file name and then on the arrow pointing to the right hand side.*

3. Select the **Trim adapter list** you created to trim the remaining PCR adapters off your reads (figure 5).

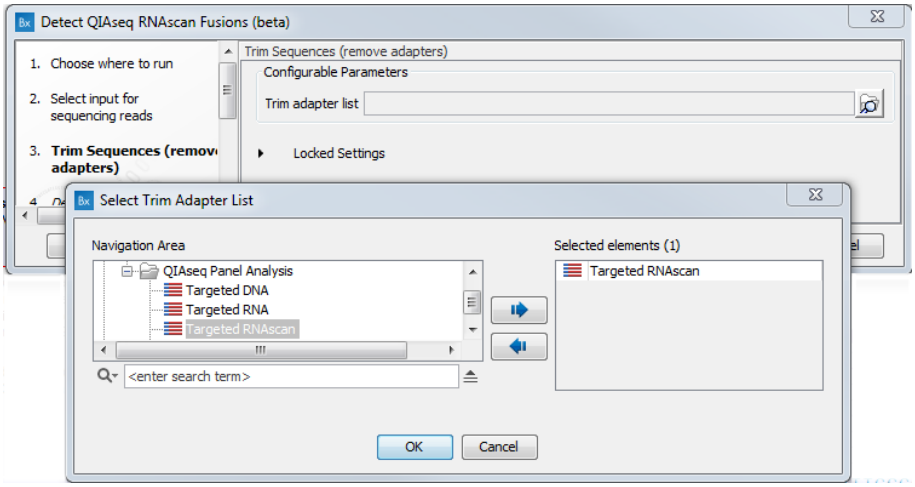


Figure 5: *Select the Trim adapter list you generated ahead of running the workflow.*

4. Leave the parameters from the "Detect Fusion Genes" dialog as they are set by default and click **Next**.

5. The tutorial Reference Data Set only contains the primer track specific to the panel you used to generate the sequenced reads (figure 6), so you do not need to select a specific track form a drop down menu.

6. Finally, in the last wizard step, choose to **Save** the results of the workflow and specify a location in the Navigation Area before clicking **Finish**.

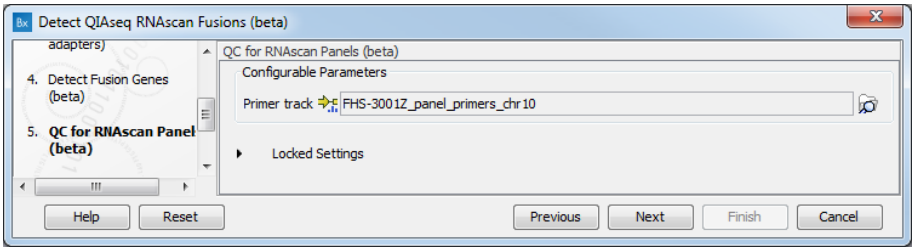


Figure 6: *Select the primer file you imported ahead of running the workflow.*

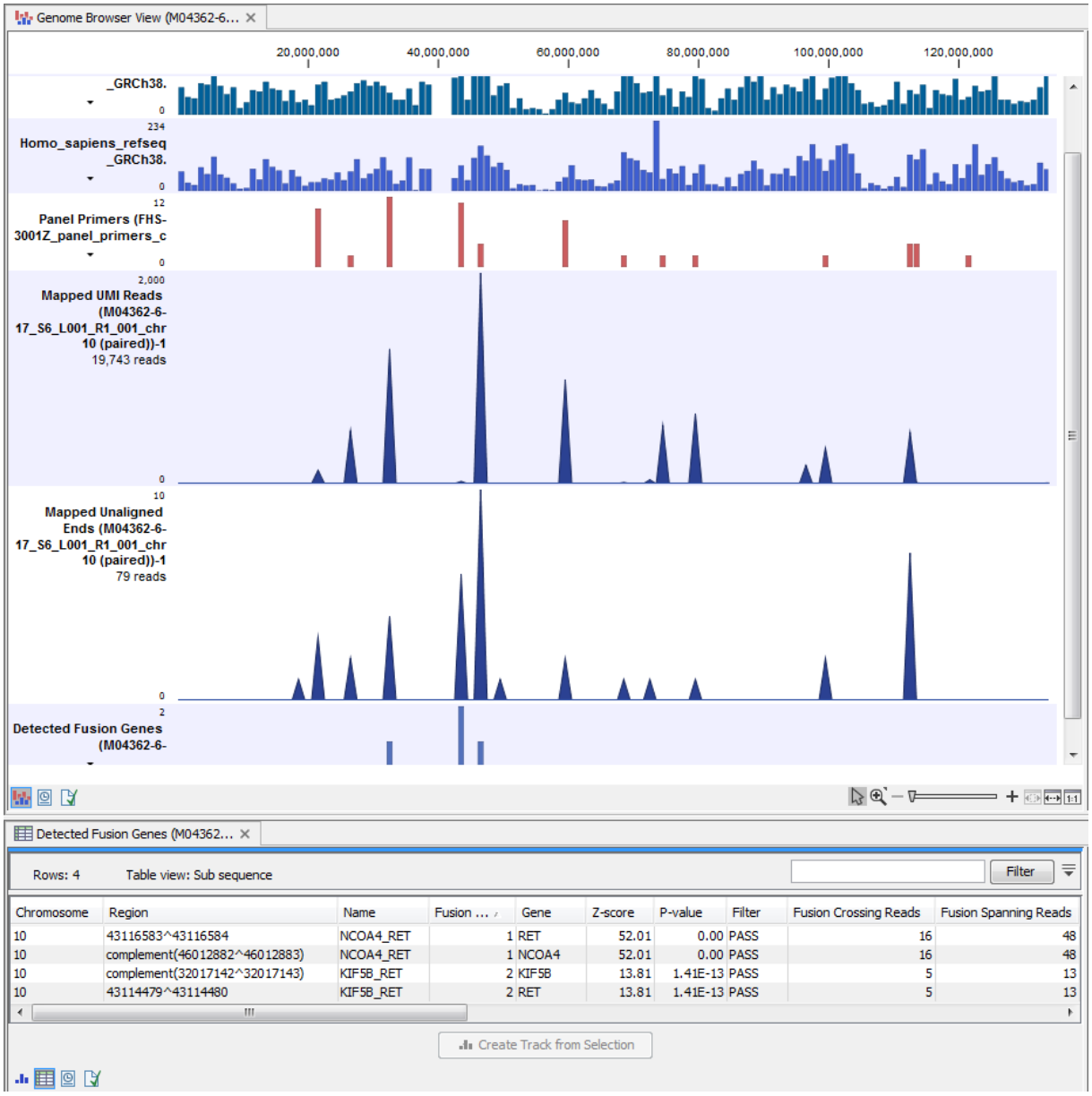
Output from the Detect QIAseq RNAscan Fusions workflow

The Detect QIAseq RNAscan Fusions workflow produces a series of output. Before checking the detected fusions, users are usually recommended to review the QC for RNAscan Panels Report, and in particular, to verify that no DNA contamination was found in the sample. Also the four reference genes included in the RNAscan panels should have been detected and their expression to that of the target genes should be reviewed (Target gene vs reference gene coverage ratio [%]). In the case of this tutorial, the data was limited to chromosome 10 only and the QC for RNAscan Panels Report is therefore incomplete.

The easiest way to review the results is to open the Genome Browser View (figure 7).

Double-click on the Detected Fusion Genes track name (to the left of the Genome Browser View) while pressing the Ctrl button. The fusion track will open as a table in split view, below the Genome Browser View. You can sort and filter the table using one of the metrics. Fusions that get the **Filter** annotation **PASS** have not been caught by any filter and are therefore the most trustworthy fusions. Sorting the table using the **Fusion Rank** column will move the most trustworthy fusions to the top of the table. Finally, clicking on a fusion event in the table will zoom in to its location in the UMI read mapping, allowing you to review the UMI reads supporting the detected fusion. Note that each fusion event is represented by two rows in the table, allowing you to visualize successively both breakpoints of each fusion. You can estimate the evidence for the fusion by studying the Mapped UMI Reads that aligns to one side of a fusion, and the Mapped Unaligned Ends that, after re-mapping, align to the other side of the fusion.

Note that the Mapped UMI track displays the mapping of UMI reads to the transcript sequences. Paired reads still in proper pairs are blue. Forward reads are green, reverse reads are red. Unaligned ends of mapped reads are shown in a lighter shade of the respective read color. Multi-match reads, i.e., reads that could have been mapped to another position within the mapping criteria used, are in yellow.



Chromosome	Region	Name	Fusion ...	Gene	Z-score	P-value	Filter	Fusion Crossing Reads	Fusion Spanning Reads
10	43116583^43116584	NCOA4_RET	1	RET	52.01	0.00	PASS	16	48
10	complement(46012882^46012883)	NCOA4_RET	1	NCOA4	52.01	0.00	PASS	16	48
10	complement(32017142^32017143)	KIF5B_RET	2	KIF5B	13.81	1.41E-13	PASS	5	13
10	43114479^43114480	KIF5B_RET	2	RET	13.81	1.41E-13	PASS	5	13

Figure 7: *A Genome Browser View showing the NCOA4_RET fusion.*