Tutorial

Modification of an Existing Workflow

June 20, 2019

Modification of an Existing Workflow

Ready-to-use workflows are provided in *CLC Genomics Workbench* for different applications and scenarios. However, there may be situations where you would like to extend an analysis or customize it to special needs. In these situations it can be relevant to modify an existing workflow, or to customize a reference data set. In this tutorial we will first modify the **Identify QIAseq DNA Variants** ready-to-use workflow (available when the Biomedical Genomics Analysis plugin is installed) to adapt it to germline applications.

Modifying a workflow

1. To modify a workflow, we are first going to open the workflow layout in the **View Area**. This can be done by right-clicking on the name of the workflow in the Toolbox, and selecting the option "Open Copy of Workflow" (see figure 1).
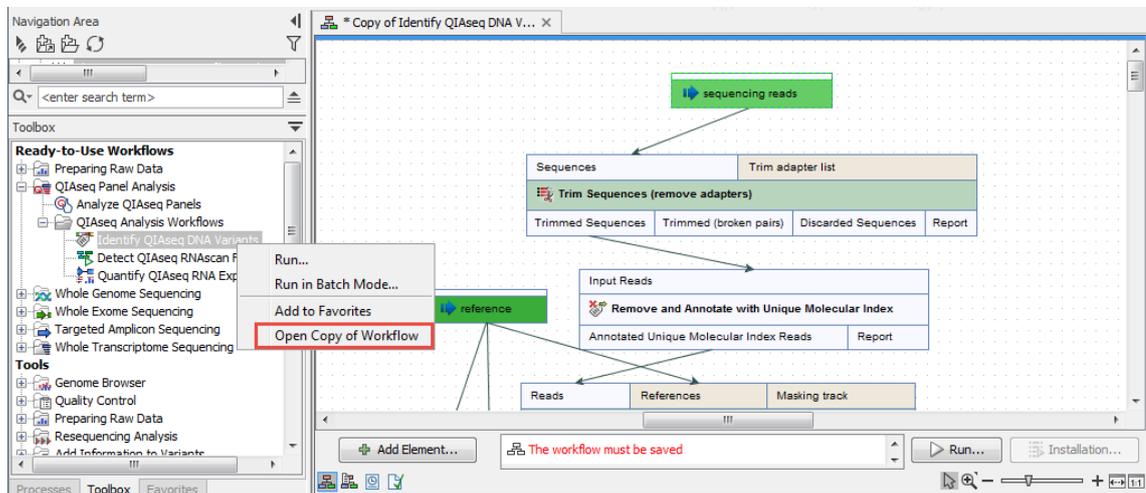

Figure 1: *Open a workflow layout in the open workflow editor.*

2. Use the **Find** field in the right side panel of the workflow view, type in "low" to find the tool **Low Frequency Variant Detection**. Remove the tool by right-clicking on its name (in the green tile area) and selecting the option **Remove** (see figure 2). Note that when a tool is deleted from a workflow, arrows connected to it are also removed. It may be good to remember what were the input and output of a tool before removing it, so similar connections can be reestablished later on with the replacement tool.
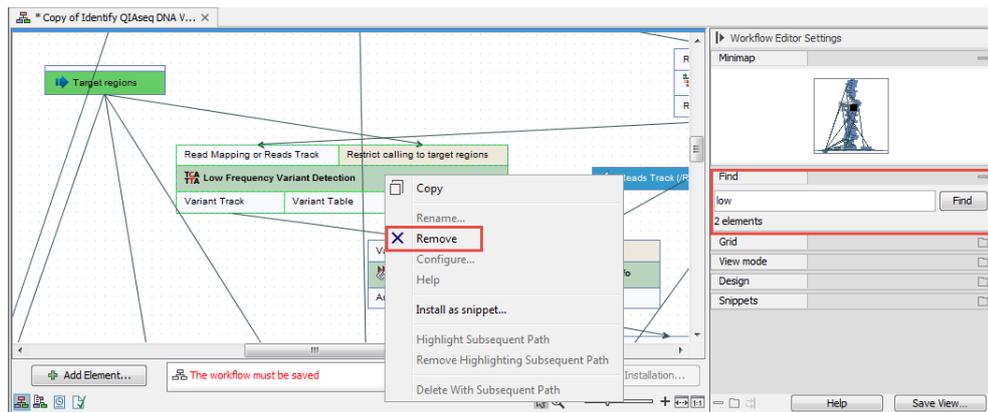
Figure 2: *Remove the tool Low Frequency Variant Detection.*

3. Next we would like to replace the tool with another variant caller, the Fixed Ploidy Variant Detection tool. Right-click in the workflow view area and select in the menu that pops up the option **Add Elements**. In the dialog that opens, you can type "ploidy" in the upper field to easily locate the tool in the Toolbox (see figure 3). Click **OK**.
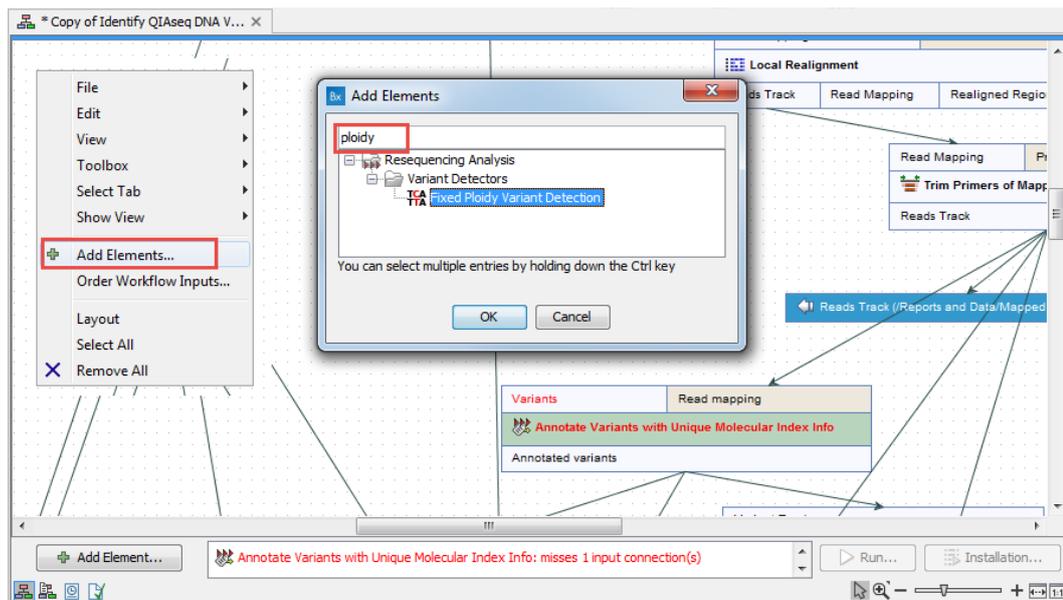

Figure 3: *Insert the tool Fixed Ploidy Variant Detection in the workflow editor.*

4. Connect the various inputs to the new tool exactly as they used to be connected with the Low Frequency Variant Detection (figure 4).

   - Select the green tile input called "Target regions", and while holding down the mouse button on this box, drag the arrow that appears to the tool input tile labeled "Restrict calling to target regions". When you let go of the mouse button, the two tools will be connected with an arrow.

   - Proceed the same way with the "Reads Track" output from the Trim Primers of Mapped Paired End Reads: connect the arrow to the variant caller tool input labeled "Read Mapping or Reads Track".

   - Finally, select the variant caller output "Variant Track" and connect it to the next tool called Annotate Variants with Unique Molecular Index Info.
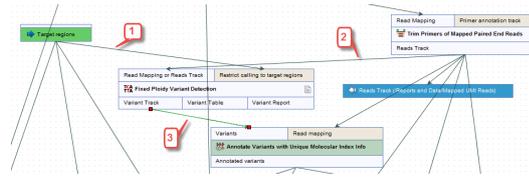
Figure 4: *Connect the relevant inputs and output for the Fixed Ploidy Variant Detection tool.*

5. Save the workflow by using **Save** in the top toolbar. Choose a location in the **Navigation Area** and enter the name *Identify QIAseq DNA Variants for Germline Applications* (figure 5).
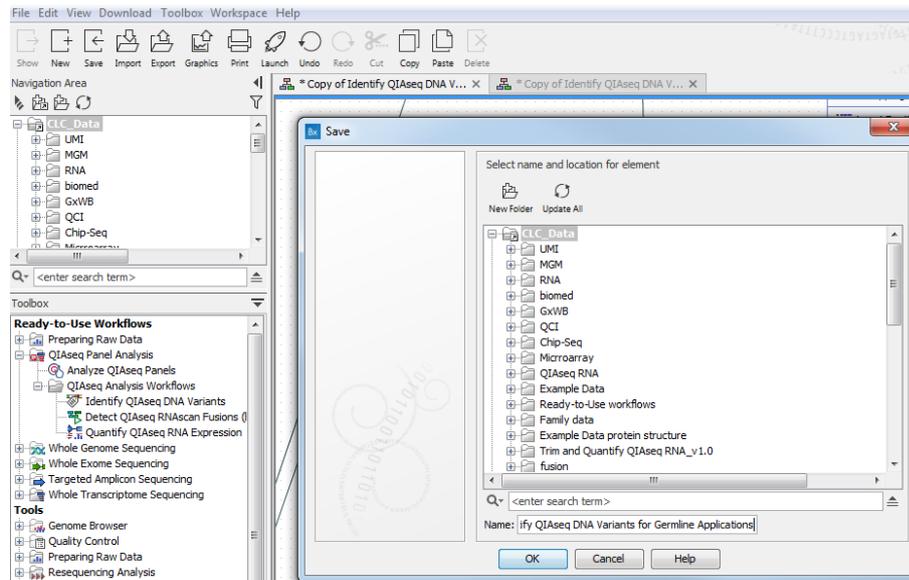

Figure 5: *Rename and save the workflow.*

6. After the workflow has been saved, click on **Installation** in the lower part of the workflow editor to install the workflow in your workbench. Fill in your affiliation details as in figure 6: name, email, institution etc. Under "Workflow description" it is a good idea to write a detailed workflow description. Click on the button labeled **Next**.
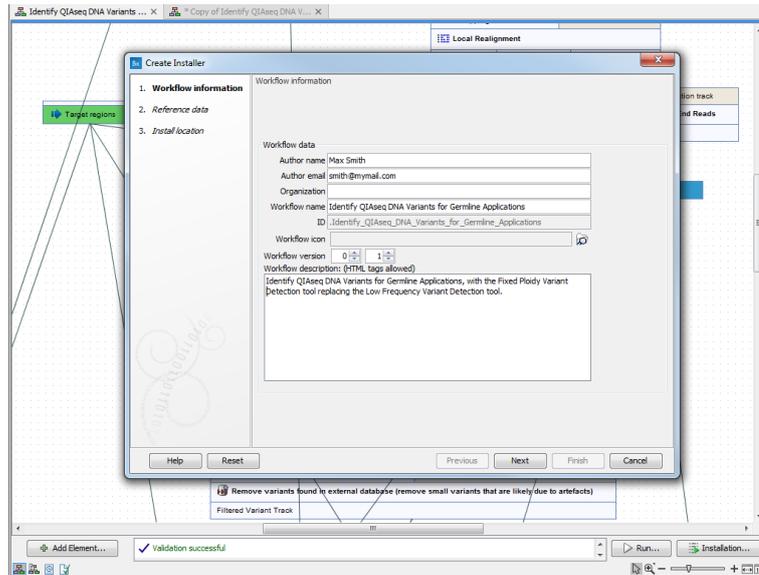
Figure 6: *Fill in your email adress and give the workflow a name.*

7. In the Reference Data window, you can choose between "Ignore", "Reference" or "Bundle" the references. When selecting **Bundle** the reference data are packed into the installer and the workflow will be configured to use exactly these data. This is in contrast to **Ignore** where no specific reference data are packed into the installation file of the workflow. Instead, the workflow will use the reference data that you have downloaded and specified with the Reference Data Manager.

8. Choose to **Install the workflow on your own computer**.

9. Click on the button labeled **Finish**.

You can now find your new workflow in the toolbox in the "Workflows" folder:

   Toolbox | **Tools** | **Workflows**   .