

Create an Automated, Customized NGS Analysis and QC Report with NextGENe® Software

December 2012

Kevin LeVan, Megan Manion, Sean Liu, Ni Shouyong, and CS Jonathan Liu

Introduction

NextGENe software v2.3.1 includes a new automated approach for creating a single customizable report to enhance your NGS analysis pipeline. The Summary Report can be configured to include both analysis and quality control information. And this report can be set up either prior to or following processing of your samples. The new reporting capability streamlines the analysis pipeline while affording desired flexibility. Reporting configurations can be saved and easily applied to subsequent analysis projects. Report configuration is a simple matter of pointing and clicking on the required report elements, no scripting is required.

Procedure

The Summary Report of NextGENe Viewer displays and exports the information of several reports within a single window. It is viewed in the Reports section of the main display by selecting Summary Report from the Reports menu (see Figure 2) or by clicking on the Show Report  button.

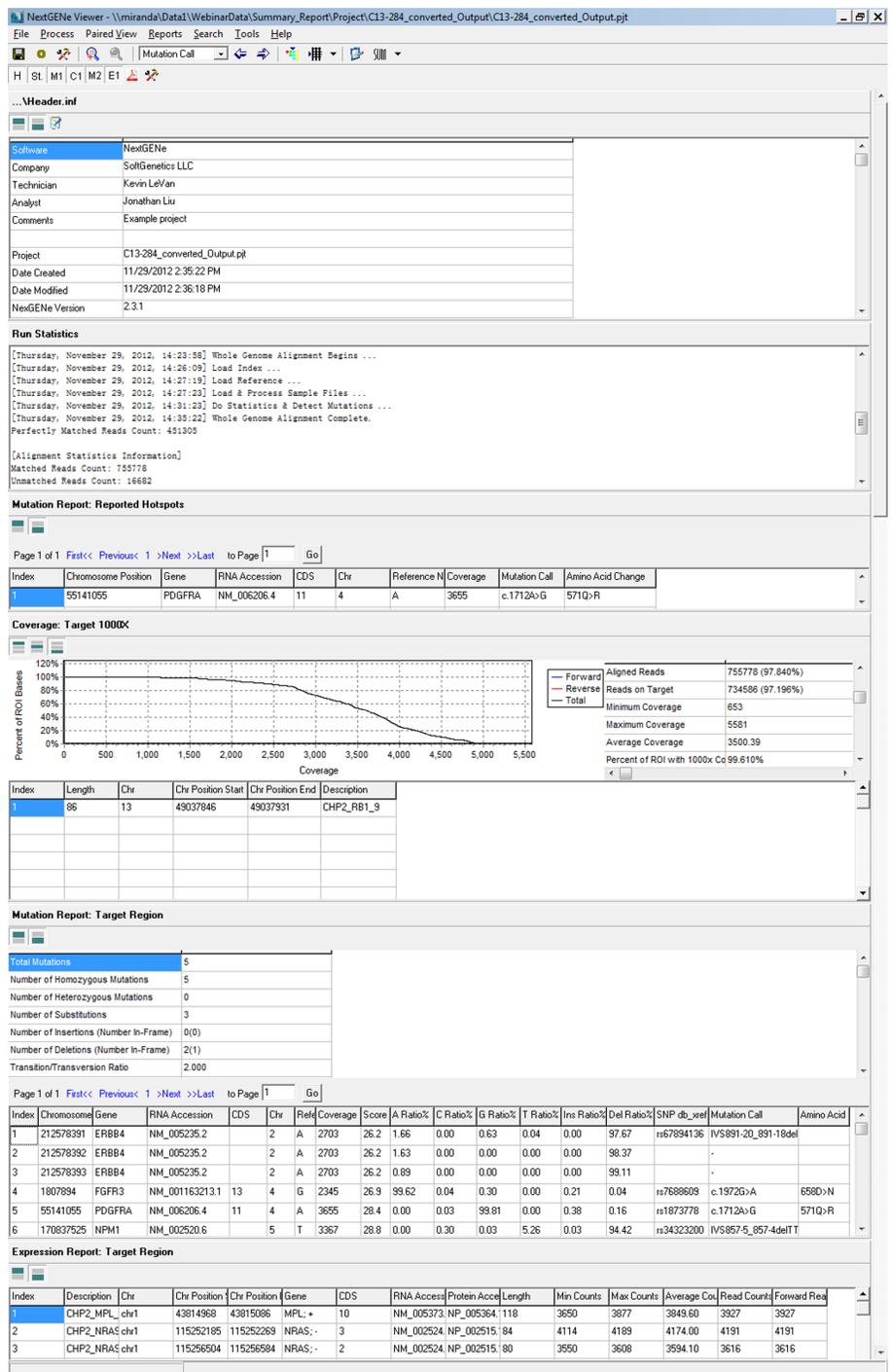


Figure 1: Typical Summary Report showing customized header, run statistics, mutation calls, and coverage statistics.

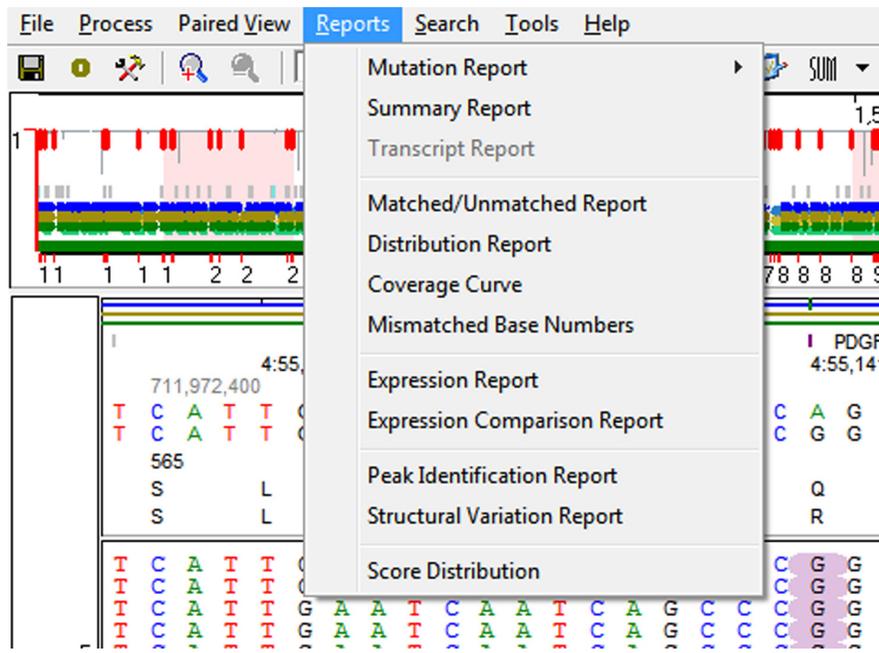


Figure 2: Click on Summary Report from the Reports menu to display this report in the main view of NextGENe Viewer.

There are two methods for configuring the Summary Report, manually through the NextGENe Viewer’s Summary Report settings dialog or automatically while setting up projects.

Manually Generate Summary Report

One method for generating this report is through the Summary Report Settings dialog, which can be opened by clicking on the last button of the report’s toolbar, (see Figure 3). Select the report that you want to add and set the customized configuration file. Configuration files can be generated by clicking on the Edit button.

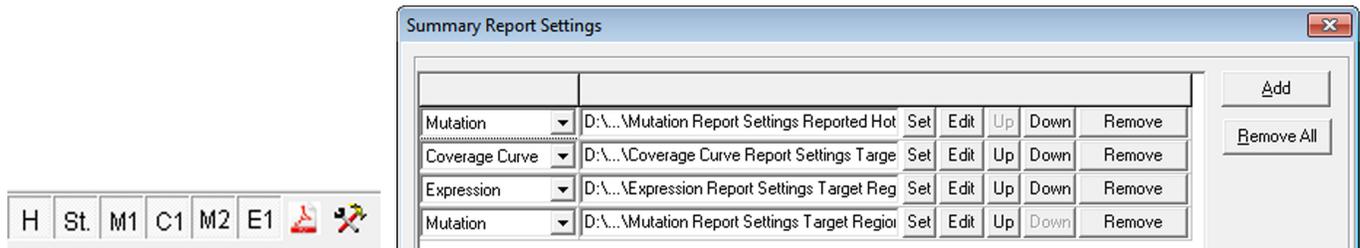


Figure 3: Summary Report toolbar is shown on left. The first several buttons allow you to show/hide Header, Statistics, and other configured sections. Click the next button to save the Summary Report as a PDF. Click the last button to open the Summary Report Settings dialog, shown on right. Easily select the reports of interest by choosing the settings.

Multiple sections of the Summary Report can be configured from the same report settings. For example, Figure 3 shows two Mutation Report sections, lines 1 and 4. The configuration file for each line is unique, one section is set to show only mutation calls within hotspots and the other section is set to show all mutation calls within the target region. Any available report can be displayed as multiple sections of the Summary Report.

Automatically Generate Summary Report

The other method for generating the Summary Report is an automated approach. You can set up NextGENe’s Project Wizard or the NextGENe AutoRun Tool to automatically configure the desired Summary Report.

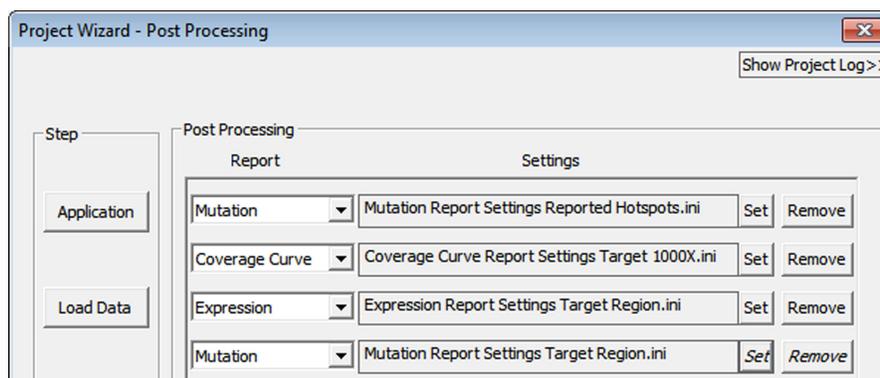


Figure 4: The last step of the Project Wizard is Post-Processing, where you can configure the reports to have automatically generated. These reports are added to the Summary Report.

Within the NextGENe Project Wizard, there is a Post-Processing step (see Figure 4). This post-processing step enables you to automatically create a list of reports from customized settings and all of these reports are automatically added to each section of the Summary Report. Similarly to the Summary Report Settings dialog, the Post-Processing dialog contains a drop-down box of available reports and the path to a configuration file for that report. Within the NextGENe AutoRun's Job File Editor, there is also an option to configure Post-Processing reports.

Results

The Summary Report as shown in Figure 1 contains many sections including a header, run statistics, 2 lists of variants with different filtering and displays, and sections including coverage information.

The Header includes the project name, date and time and software version. It has also been customized to include the names of the technician and analyst and additional comments.

From the Run Statistics section, you can see that the alignment of over $\frac{3}{4}$ million reads was aligned in about 8 minutes and the run concluded after an additional 4 minutes of calculations. Almost $\frac{1}{2}$ million reads aligned perfectly and less than 20,000 reads were unmatched. Many additional statistics are included in the Run Statistics section.

This example contains two separate groupings of variant calls. The first section of variants includes a single reported mutation call within a list of hotspot positions. The second includes all identified variations within the target region. Each of these sections has also been configured to display different columns of information for each variant.

Coverage information is displayed in two separate groupings, the Coverage and Expression sections. The Coverage section notes that over 97.8% of the reads aligned to the human build 37 reference, and that 97.1% of the aligned reads mapped to the target region. More than 99.6% of the target region exceeds the 1000X minimum coverage requirement. The table beneath the coverage plot shows a single 85bp region where the coverage falls below the 1000X target depth. A graphical representation of coverage depth versus percent of target region is also shown in the Coverage section. The Expression section lists coverage statistics for each of the 207 regions as specified in the target BED file, including the minimum and maximum depth of coverage per region and the total number of mapped reads per region.

Discussion

The Summary Report offers the ability to create a single display that contains a customized header, project information, run statistics, quality information and variants of interest, all in a single report that can easily be saved as a PDF. This report can be automatically configured before you even do your runs.

The NextGENe Viewer has many reports in addition to the Mutation Report, Coverage Report, Expression Report and Structural Variation Report, and expansion of the Summary Report to include other reports is already in progress.

Acknowledgement

Datasets used in this application note are available on the Ion Torrent Community website.