# Create an Automated, Customized NGS Analysis and QC Report with NextGENe<sup>®</sup> 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 Sum • button.

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

NextGENe Viewer	- \\miran	da\Data1\Webi	inarData\Si	ummar	y_Repo	ort\Pro	oject\C13	-284_cor	overted_0	utput\C13	-284_cor	nverted_Outp	out.pjt				_	8
<u>File P</u> rocess Paire	d <u>V</u> iew @ I_⊾	Reports Sear	ch <u>I</u> ools	Help	, 田	<b>⊧</b>	🐢 qin	•										
H St M1 C1 M2	E1 🔏	2		<b>v</b>   1	н		JUI 😋											
\Header inf																		
oftware	NevtG	FNe																-
onware ompany	SoftGr	enetics LLC																
echnician	Kevin	LeVan																
nalyst	Jonath	ian Liu																
omments	Examp	le project																
pject	C13-2	34_converted_0	)utput.pjt									_						
ate Created	11/29	/2012 2:35:22 P	M															
ate Modified	11/29	/2012 2:36:18 P	м															
exGENe Version	2.3.1																	-
un Statistics																		
Thursday, November Thursday, November	29, 201	2, 14:23:58]	Whole Gene Load Index	ome Ali	gnmen	t Begi	ins											^
hursday, November	29, 201	2, 14:27:19]	Load Refer	rence .														
nursday, November Thursday, November	29, 201	2, 14:27:23]	Do Statist	tics à	Detec:	t Muta	tions											
hursday, November rfectly Matched R	29, 2013 eads Cour	2, 14:35:22] nt: 451305	Whole Gend	ome Ali	gnmen'	t Comp	olete.											Ξ
lignment Statisti	cs Infor	mation]																
atched Reads Count	: 755778																	
watched Keads Cou																		
utation Report: Re	ported H	otspots																
age 1 of 1 First<< Pi	revious< 1	>Next >>Last	t to Page	1	Go													
dex Chromoso	me Position	n Gene	RNA Acce	ssion	CDS	CI	nr	Referer	nce N Cov	erage M	utation Ca	all Amino A	cid Change					^
55141055		PDGFRA	NM_0062	06.4	11	4		A	3655	5 c.	1712A>G	571Q>R						-
overage: Target 10	000X																	
120%												Esquard	Aligned Re	ads	755778 (97	.840%	)	
80%					~							- Reverse	Reads on T	arget	734586 (97	.196%	)	-
5 60% 40%							4					- Total	Minimum Co	overage	653			
20%													Maximum C	overage	2500.20			
0 50	1,0	00 1,500	2,000	2,500	3,00	00	3,500	4,000	4,500	5,000	5,500		Percent of	ROI with 1000	Ix Co 99.610%			-
	6	[C1 . D		D	Cove	erage							•				,	
Idex Length	13	49037848	6 49	r Positic 037931	n End	CHP2	RB1 9	-										Ē
								-										
								_										
								_										-
lutation Report: Ta	raet Rea	ion																
atal Mutations		6				-												~
umber of Homozygous	Mutations	5				-												
mber of Heterozygou	s Mutation	s 0				-												
umber of Substitutions		3																
umber of Insertions (N	umber In-F	rame) 0(0)				_												
umber of Deletions (Nu	Imber In-Fr Ratio	ame) 2(1)				_												
		2.000		•		1												*
rage 1 of 1 First<< Pr	evious< 1	>Next >>Las	t to Page	1	Lio (	1	- Ic	la c i	*  0.5	* 10 5 1	v I z m		do un vi	daup :	hun er			_
dex Chromosome Ge	ne DD4	HNA Accession	CDS	Chr	Refe	L'overa	ge Score	A Ratio	% C Ratio	G Ratio	T Rati	Ins Ratio	© Del Ratio%	SNP db_xrel	Mutation Call	104.1	Amino Acid	l ^
212578391 ER	BB4	NM_005235.2		2	A .	2703	26.2	1.66	0.00	0.63	0.04	0.00	98.37	1867894136		-18del		-
212578393 FR	BB4	NM_005235.2		2	A	2703	26.2	0.89	0.00	0.00	0.00	0.00	99.11					
1807894 FG	FR3	NM_001163213	3.1 13	4	G	2345	26.9	99.62	0.04	0.30	0.00	0.21	0.04	rs7688609	c.1972G>A		658D>N	
55141055 PD	GFRA	NM_006206.4	11	4	A :	3655	28.4	0.00	0.03	99.81	0.00	0.38	0.16	rs1873778	c.1712A>G		571Q>R	
	M1	NM_002520.6		5	T :	3367	28.8	0.00	0.30	0.03	5.26	0.03	94.42	rs34323200	IVS857-5_857-4	4delTT		-
170837525 NF	Larget B	egion																
170837525 NF																		
170837525 NF																		
170837525 NF xpression Report:	n Chr	Chr Positi	ion Chr Pos	ition   G	ene	0	DS	RNA A	ccess Prot	ein Acce Le	ngth	Min Counts	Max Cour	its Average 0	Cou Read Counts	Forwar	rd Rea	<b>_</b>
170837525 NF xpression Report: dex Descriptio CHP2_MF	n Chr PL_chr1	Chr Positi 43814968	ion Chr Pos 8 438150	ition   G	ene PL;+	C 1	DS D	RNA A NM_00	ccess Prot 5373. NP_	ein Acce Le 005364. 11	ngth 8	Min Counts 3650	Max Cour 3877	its Average 0 3849.60	Cou Read Counts 3927	Forwar 3927	rd Rea	
i 170837525 NF ixpression Report: idex Descriptio CHP2_MF CHP2_NF	n Chr PL_ chr1	Chr Positi 43814968 11525218	ion Chr Pos 8 438150 85 115252	ition G 86 M 269 N	ene PL;+ RAS;-	C 11 3	DS D	RNA A NM_00 NM_00	ccess Prot 5373. NP_ 2524. NP_	ein Acce Le 005364. 11 002515. 84	ngth 8	Min Counts 3650 4114	Max Cour 3877 4189	ts Average 0 3849.60 4174.00	Cou Read Counts 3927 4191	Forwar 3927 4191	rd Rea	1



SoftGenetics LLC 100 Oakwood Ave. Suite 350 State College, PA 16803 USA Phone: 814/237/9340 Fax 814/237/9343 www.softgenetics.com email: info@softgenetics.com





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.

	Summary Report Settin	ngs							×
								Add	
	Mutation 💌	D:\\Mutation Report Settings Reported Hot	Set	Edit	Up	Down	Remove	Bemove	
	Coverage Curve 💌	D:\\Coverage Curve Report Settings Targe	Set	Edit	Up	Down	Remove	<u></u>	<u>~</u>
	Expression 💌	D:\\Expression Report Settings Target Reg	Set	Edit	Up	Down	Remove		
St. M1 C1 M2 E1 👗 🛠	Mutation 💌	D:\\Mutation Report Settings Target Region	Set	Edit	Up	Down	Remove		

**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.



Н



Project Wizard - Po	st Processing		<b>×</b>
			Show Project Log>>
_	- Deat Dragoning		
Step	Post Processing	0	
	Report	Settings	
Application	Mutation	▼ Mutation Report Settings Reported Hotspots.ini	Set Remove
	Coverage Curve	Coverage Curve Report Settings Target 1000X.ini	Set Remove
Load Data	Expression	Expression Report Settings Target Region.ini	Set Remove
	Mutation	Mutation Report Settings Target Region.ini	Set Remove

**Figure 4:** The last step of the Project Wizard is Post-Processing, where you can configure the reports to have autmatically 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 <sup>3</sup>/<sub>4</sub> million reads was aligned in about 8 minutes and the run concluded after an additional 4 minutes of calculations. Almost <sup>1</sup>/<sub>2</sub> 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.



