Analyzing NGS Data with NextGENe® Software Pipeline Tool

January 2012

Megan Manion, Kevin LeVan, John McGuigan, Shouyong Ni, CS Jonathan Liu

Introduction

Next Generation Sequencing Technologies allow for the sequencing of multiple samples in short time frames. The use of barcoding or multiplexing techniques increases the number of samples that can be processed on each machine run. Because of this high output, setting up analysis parameters for the output samples can be time consuming and tedious.

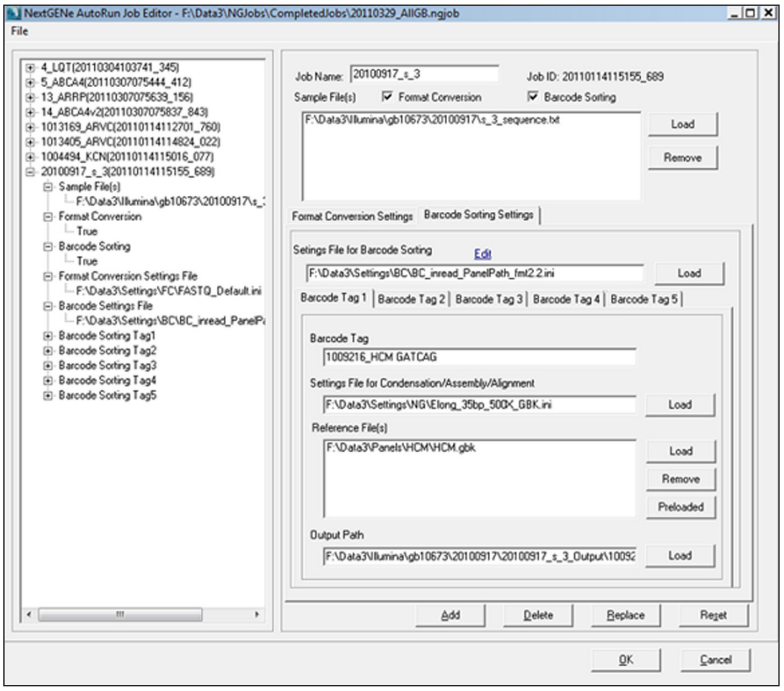
NextGENe Software now includes a specialized Pipeline Tool, NextGENe AutoRun, which provides enhanced, streamlined batch processing. This tool allows users to quickly set up multiple jobs to be analyzed consecutively. Multiple samples can then be analyzed unattended.

Procedure

Create Job File

1. Open NextGENe Auto Run Tool by going to Tools > NextGENe AutoRun.
2. When NextGENe Auto Run opens, go to Tools > Job File Editor.
3. Provide a name for the job in the Job Name field.
4. If sample files are not in fasta format, select Format Conversion. If samples files are in fasta format skip to Step 6.
5. Specify settings for format conversion:
   a. Click Load to load a pre-saved *.ini file with the desired settings OR
   b. Click Edit to manually select desired settings.
6. If sample data includes barcodes for multiple samples, select Barcode Sorting. If sample files do not include barcode, skip to Step 8.
7. Specify settings for barcode sorting :
   a. Click Load to load a pre-saved *.ini file with the desired settings OR
   b. Click Edit to manually select desired settings.
8. Specify Reference File(s)*.
   a. Click Load to select a reference file or files in Fasta or GenBank format OR
   b. Click Preloaded to select an indexed reference genome
9. Specify analysis settings by loading a pre-saved *.ini file*.
10. Specify directory and folder name for output files*.
11. To add another job to the file, click Add.
12. When finished, click Save to Save the Job File.

*When barcode sorting is used, these options can be set for each sample individually so different settings or different reference files can be used for each sample.



Processing Job Files

1. Click on the Settings icon ( ) in NextGENe AutoRun to specify the folder where the job file(s) are saved under Job File Detecting Directory.
2. To begin processing the job file(s), click on the Detect icon ( ) or go to File > Detect.

Discussion

NextGENe's AutoRun Tool allows users to quickly set up job files to specify sample files, reference files and settings for the analysis of multiple projects. NextGENe AutoRun scans the job files to check for the availability of all files needed for processing a job. When files are available, NextGENe AutoRun launches NextGENe to begin processing the job. Jobs are processed consecutively. If some files are not available, NextGENe AutoRun will periodically re-check for the files at the time interval specified in the NextGENe AutoRun settings. Job files can be easily edited to add additional jobs or edit jobs that have not yet been processed. Also, additional jobs can be added to the queue while NextGENe is currently processing another set of jobs. NextGENe will automatically begin the next list of jobs that are queued once the first list is complete.

NextGENe can be used to analyze next generation sequencing data from Illumina GA, MiSeq & HiSeq systems, Roche/ 454 GS FLX, FLX Titanium and Junior, Applied Biosystems' SOLiD System, Ion PGM and PacBio platforms. Multiple application types are available including SNP/Indel Analysis, RNA-Seq, de novo assembly, and ChIP-Seq. Results are displayed in the interactive NextGENe Viewer, providing a detailed visualization not found in other commercial programs like Lasergene's SeqMan Pro, CLC Bio & DNASTAR's NGEN or in open-source tools like TopHat, Bowtie & BWA.

Acknowledgement

We would like to thank GeneDx for their suggestions in the development of the NextGENe AutoRun Tool.

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