BECKMAN COULTER
Life Sciences

Automation of the Agilent SureSelect XT Library Prep Kit with the Human All Exon V6 Panel on the Biomek NGeniuS Next Generation Library Prep System

Introduction

In the past three decades, methods for sequencing nucleic acids have improved significantly in speed and accuracy. Despite the technological advances in sequencing, preparation of libraries for Next-generation sequencing (NGS) remains a tedious process. The library preparation process can take anywhere from 2.5 hours to several days to complete, depending on the type of library created. Great care must be taken to keep accurate records of which samples received which index, while manual pipetting can lead to errors in creation of libraries with incorrect indices. Many library preparation protocols have strict time constraints and limited safe stopping points, leading to long, inconvenient preparation procedures. While chemistry companies are optimizing kits to make the process less dense, more controlled, and easier to execute, the current processes still require care and precision.

Agilent's SureSelect XT DNA Library Prep Kit combined with the Human All Exon V6 probe panel prepares hybrid-capture libraries suitable for sequencing on Illumina platforms. The Human All Exon V6 probe panel consists of biotinylated baits covering approximately 20,000 coding regions (60 Mb) of the human genome, allowing for the simultaneous assessment of the estimated 20,000 to 50,000 variants thought to exist in the human exome. The DNA input for this automated method is 200 ng and the SureSelect XT Human Exon V6 kit can be used to process both genomic DNA (gDNA) and Formalin-Fixed Paraffin-Embedded DNA (FFPE DNA) samples.

The first step in the library preparation protocol is Covaris shearing (target fragment size is 150-200bp), followed by an AMPure cleanup. An optional QC step allows the user to verify the size of the sheared DNA on a fragment analyzer such as the Agilent Bioanalyzer. The ends of the DNA fragments are then repaired and cleaned up again, followed by A-tailing and another AMPure cleanup. Libraries are ligated to sequencing adapters, purified again, PCR amplified, and purified again. Following library preparation, the libraries are subjected to both fragment analysis and a fluorometric assay to determine concentration. 750 ng of library is then concentrated to a final volume of 3.4 µL and hybridized for a minimum of 16 hours with the Human All Exon V6 probe panel. Following hybridization, the library fragments bound to the probe panel are then captured using streptavidin beads, washed to remove library fragments not within the targeted region, amplified, and subjected to a final purification producing libraries with a final size of approximately 250-350bp. The final libraries are compatible with any Illumina sequencing platform.

In this application note we demonstrate that the automated processing on the Biomek NGeniuS Next Generation Library Prep System is equivalent to the manual preparations at input concentration of 200 ng (gDNA) and 200 ng (FFPE DNA). The hands-on time for processing is reduced, handling errors are reduced, and the interactions/touchpoints are reduced. The Biomek NGeniuS system is a sample preparation system that allows users to configure settings for library preparation, select stopping points and upload sample data to a csv file and mark it ready to run by the instrument in a process called batch

creation. The information provided by the user for the batch is then converted into a work aid that gives detailed instructions on how to thaw reagents and where to load them on the system, as well as any safety warning associated with handling of the reagents. Biomek NGeniuS software monitors loading of consumables and reagents in real time, limiting possibilities for user error. Following reagent loading, the Biomek NGeniuS system aliquots the reagents to an appropriate temperature-controlled storage (chilled, warm, or ambient). Finally, the user loads the reaction vessel containing the samples onto the instrument and executes the batch. When the batch is complete, the user can pool the libraries and then dilute and denature the libraries as required for sequencing.

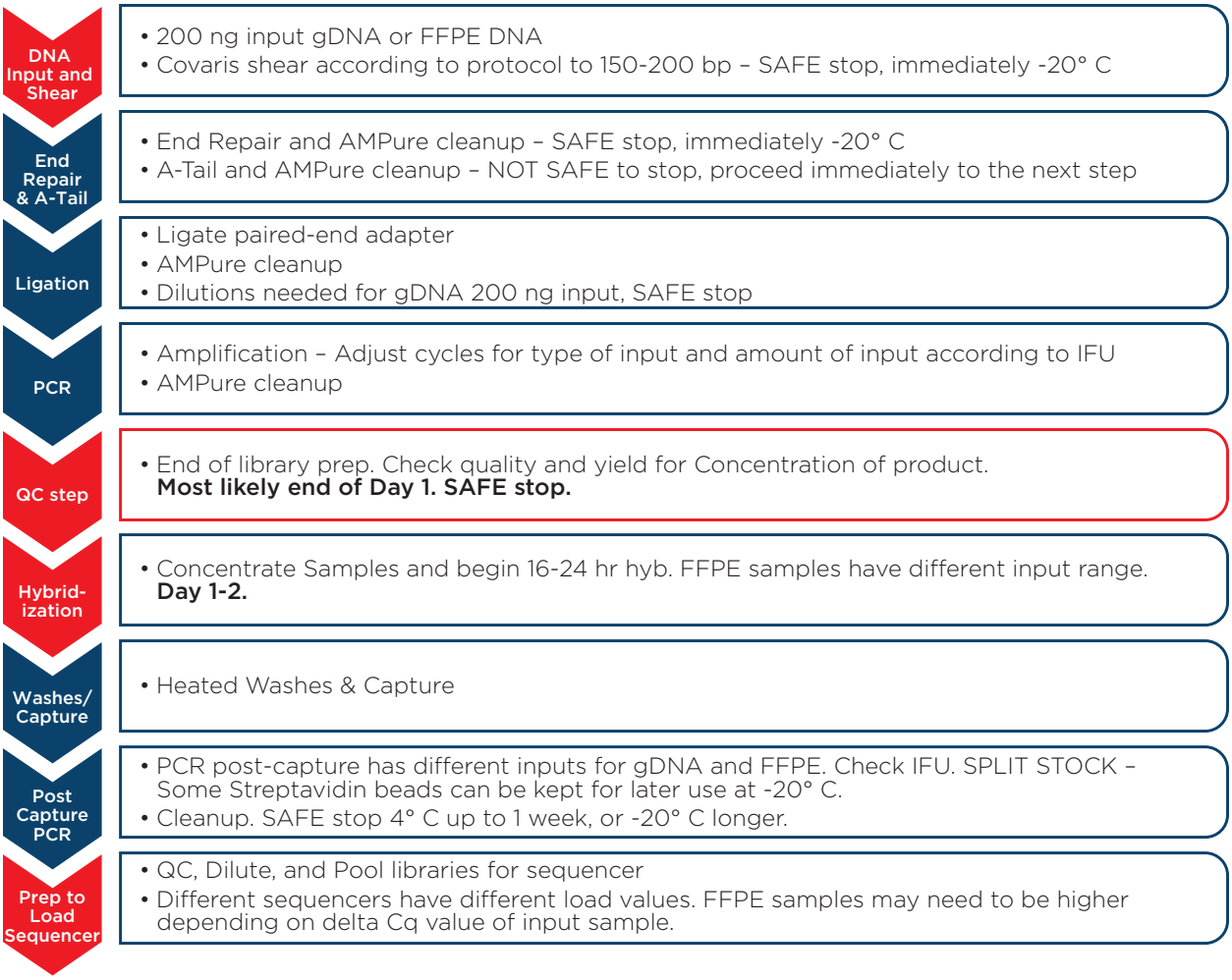

Figure 1. Workflow for Agilent SureSelect XT library prep kit with Human All Exon V6 panel protocol. Red chevrons indicate off-instrument steps. Blue chevrons indicate steps performed by the Biomek NGeniuS system.

Methods

1) Run Setup

When the samples are ready to run, a batch is set up in the Biomek NGeniuS customer portal. The first step is to select the **+create** button to create a batch to be run on the system (Figure 2). Next, the Agilent SureSelect XT Human All Exon V6 App was selected to process samples. The setup is broken up into four sections: Batch info (name of batch and number of samples to be run), App Settings, Sections, Sample Data. The Batch name is a unique run name for the samples being processed. Number of samples is any number between 4 to 16 for this application, as indicated by the light grey numbers below the input box. App Settings contains variables specific to the library kit that may be changed between runs or may be locked by the lab administrator. Batch info and App settings are shown in Figure 3. Table 1 lists the app settings and descriptions of each setting.

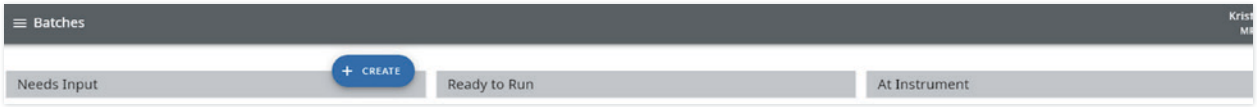

Figure 2. The **+create** button in the above figure is used to begin a new batch setup.

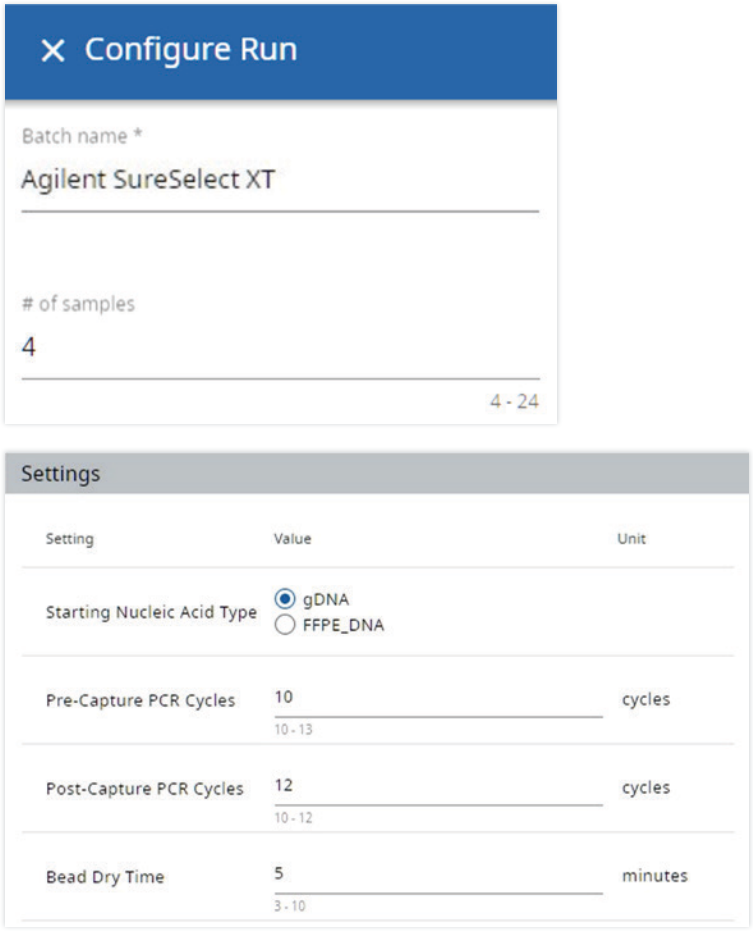

Figure 3. Batch info and app settings for batch run.

Setting	Description
Starting Nucleic Acid Type	The type of nucleic acid used as starting material for the current batch (gDNA or FFPE).
Pre-Capture PCR Cycles	The number of cycles of PCR amplification to perform on library.
Post-Capture PCR Cycles	The number of cycles of PCR amplification of enrichment product.
Bead Dry Time	The default is 5 minutes. This may change based on environmental conditions like humidity.

Table 1. Application settings and description of each setting.

The next section of data to be filled out is **Sections** (Figure 4). Agilent SureSelect XT has eight potential sections and users can select where to start in the process using the **Start at section** dropdown. This method also contains two **Off System** sections, which are completed off the Biomek NGeniuS system before reloading back onto the system with instrument prompts used for this process. All the starting points are determined by safe stops defined in the instructions for use of the library prep kit, which are also suitable for a safe stop on the instrument. The blue slider to the left of the sections allows the user to select a safe stop to end processing of samples. The instrument is designed to run unattended, but users can elect to stop processing and store samples safely before resuming the run at the next shift.

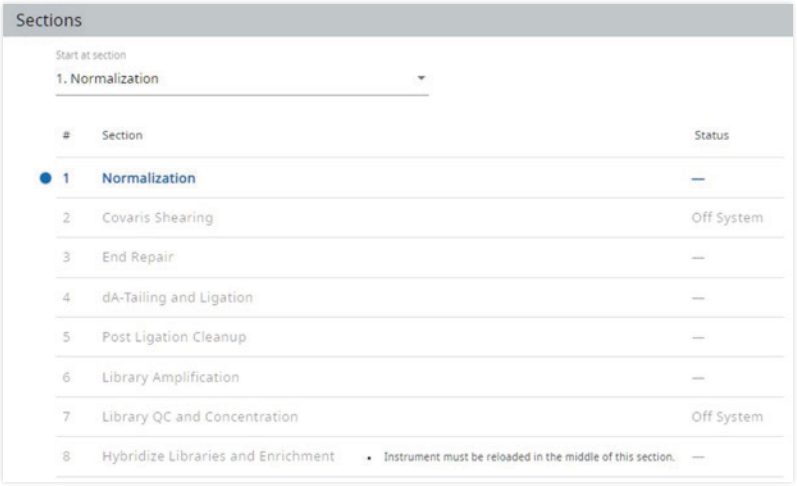

Figure 4. Sections of Agilent SureSelect XT application. Sections the Biomek NGeniuS Library Preparation System is not able to perform on deck are marked "Off System."

The final step in setting up a batch to run is to input the sample data (Figure 5). In the sample data section, users can click the **DOWNLOAD SAMPLE DATA TEMPLATE** and fill in the appropriate information. This is a .csv file that is filled out and uploaded into the sample data by clicking the **Upload** button. Users can utilize tool tips to determine what information goes into each column by hovering over the header of each column. Agilent SureSelect XT has 3 different data pieces that are required for plate-based index processing. The first column is the **Sample_ID** of each sample. The second item, **Index**, is the index to be used for that particular sample. The final column, initialConcentration, is the concentration of DNA that will be placed into each well for dilution and processing for library preparation. Once the data is entered in the template and saved, the user clicks the **Upload** button. If there are any unexpected values in the sample data file, a red box will appear indicating the source of the problem. Users can fix the data file and upload again if needed. The final step is to click the **Ready to run** button in the top right of the screen. The batch can be initiated at any Biomek NGeniuS system within the same tenant.

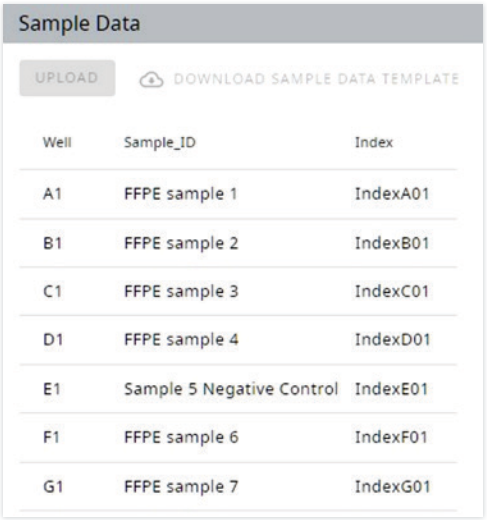

Figure 5. Sample Data information. Shown is the 7 sample Horizon HD798 FFPE sample data.

2) Library Preparation

Several runs were performed on the Biomek NGeniuS system with inputs of 200 ng (Table 7) using reagents, equipment, and consumables detailed in Tables 8, 9, and 10. System requested reagents from the SureSelect XT Exon v6 kit, bulk reagents, and Biomek NGeniuS consumables were loaded onto the system for processing. Variables selected for processing are as seen in Table 3. The system processed dilution of samples to a standard concentration to be used for library concentration based on the total mass of input DNA. The DNA was sheared using a Covaris S220 sonicator using conditions described in Table 2. The sheared DNA samples were placed on the Biomek NGeniuS system along with reagent master mixes to prepare the initial libraries.

Setting	Value
Duty Factor	10%
Peak Incident Power	175
Cycles per Burst	200
Treatment Time	360 Seconds (240 for FFPE samples)
Bath Temp	4-8° C

Table 2. Settings for Covaris Shear for 200ng DNA input.

After all reagents had been aliquoted to proper storage locations, the user was instructed to remove reagents and notified of an estimated time of completion for the library prep, based on selections the user inputs at the start of the protocol. After completion of the runs, automated samples were analyzed on an Agilent 2100 Bioanalyzer system for both the pre-hybridization capture libraries and the post-hybridization capture libraries. The Qubit High Sensitivity assay (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer system were used to determine sample yields. Sequencing was performed on an Illumina NovaSeq using a 2x150bp PE run with multiple lanes of a NovaSeq S4 flowcell. Data was analyzed on Illumina BaseSpace using the DRAGEN Germline App on BaseSpace (v4.0.3) aligning the reads to the human reference genome (hg38) using default settings with each library being randomly down sampled to 30 million reads during alignment. Variant calling and enrichment metric generation was performed using Agilent SureCall software (version 4.2.2.3) with the SureSelect Human All Exon V6 r2 (Design ID S07604514) panel design.

Sample Input	Sample Type	Pre-Capture PCR Cycles	Post-Capture PCR cycles	Bead Dry Time (minutes)	Adapter Dilution?	Input for Hyb (ng)	Post Capture PCR DNA input (µL)
200ng	Genomic DNA – 4 samples	10	10	5	Ten-fold dilution	750 ng	14 µL
200ng	FFPE – 7 samples	10 (if $\Delta\Delta C_q$>4, 13 cycles)	12	5	No dilution	500-750 ng (FFPE)	30 µL
200ng	Genomic DNA – 16 samples	10	10	5	Ten-fold dilution	750 ng	14 µL

Table 3. Method variables and selections for Agilent SureSelect XT library prep kit with Human All Exon V6 panel.

Results and Discussion

After completion of the runs by the Biomek NGeniuS system, libraries were assessed on the Agilent 2100 Bioanalyzer or Agilent 4200 TapeStation (Agilent Technologies, Inc.) to determine fragment size. Library prep yield masses were measured with Qubit fluorometric quantification (Thermo Fisher Scientific). Libraries were sequenced using an Illumina NovaSeq instrument using a 2x150 PE run on multiple lanes of an S4 flowcell (Illumina Inc.). Sequencing results returned 3.77 billion reads.

Four libraries (including one negative control) were prepared from Human NA12878 gDNA samples (Coriell Institute). Pre-capture libraries showed an average size of 565 bp and an average concentration of 77 ng/µL. Post-capture libraries showed an average size of 360 bp and an average concentration of 4.7 ng/µL (Figure 6). An average of 88.5% of reads mapped to the covered regions of the panel and an average of 90.1% of reads mapped to the covered regions of the panel ± 200 bp (Table 4).

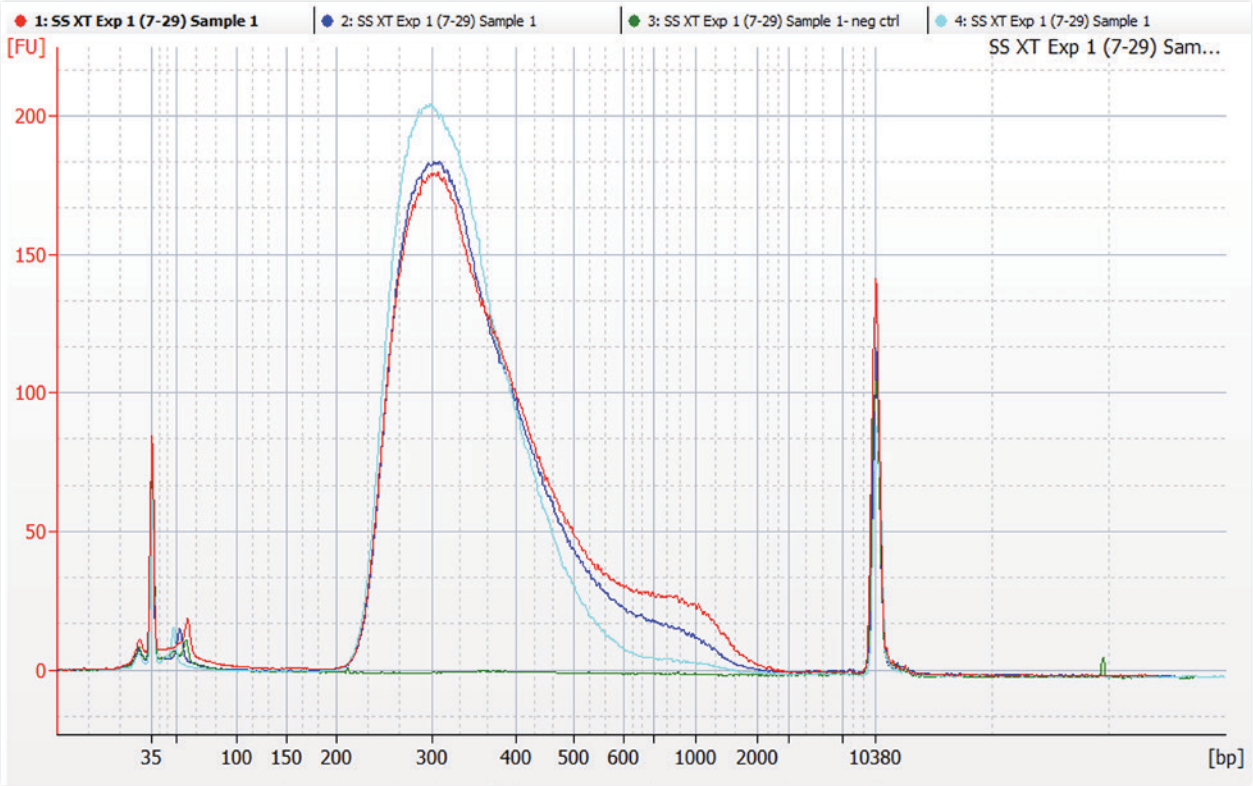

Figure 6. Post-capture NA12878 gDNA (Four samples including one negative control) library traces for Agilent SureSelect XT libraries assayed with Agilent TapeStation 2100 system with Agilent High Sensitivity DNA chip.

Seven libraries prepared from Horizon HD798 Quantitative Multiplex Reference Standard fcDNA (mild) samples (including one negative control). Pre-capture libraries showed an average size of 304 bp and an average concentration of 71.9 ng/µL. Post-capture libraries showed an average size of 337.5 bp and an average concentration of 5.9 ng/µL. (Figure 7). An average of 88.8% of reads mapped to the covered regions of the panel and an average of 90.6% of reads mapped to the covered regions of the panel ± 200 bp (Table 4).

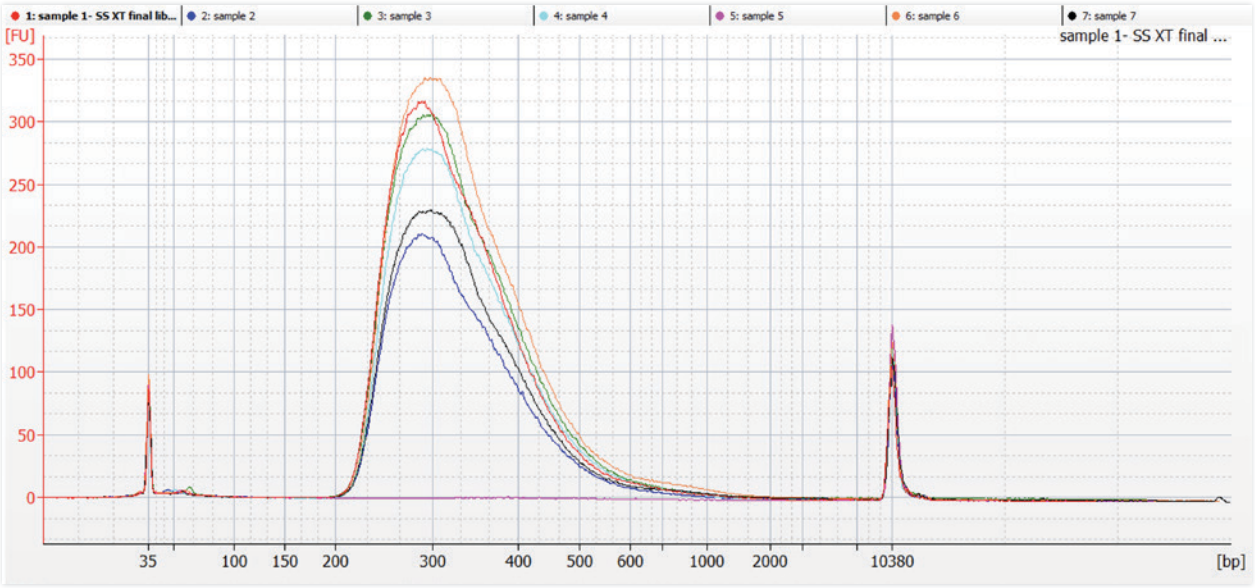

Figure 7. Post-capture HD798 FFPE (Seven samples run including one negative control) library traces for Agilent SureSelectXT libraries assayed with Agilent TapeStation 2100 system with Agilent High Sensitivity DNA chip.

Sixteen libraries (including one negative control) were prepared from Human NA12878 gDNA samples (Coriell Institute). Pre-capture libraries showed an average size of 335 bp and an average concentration of 77.4 ng/µL. Post-capture libraries showed an average size of 324.6 bp and an average concentration of 6.3 ng/µL. (Figure 8). An average of 90.6% of reads mapped to the covered regions of the panel and an average of 92.3% of reads mapped to the covered regions of the panel ± 200 bp (Table 4).

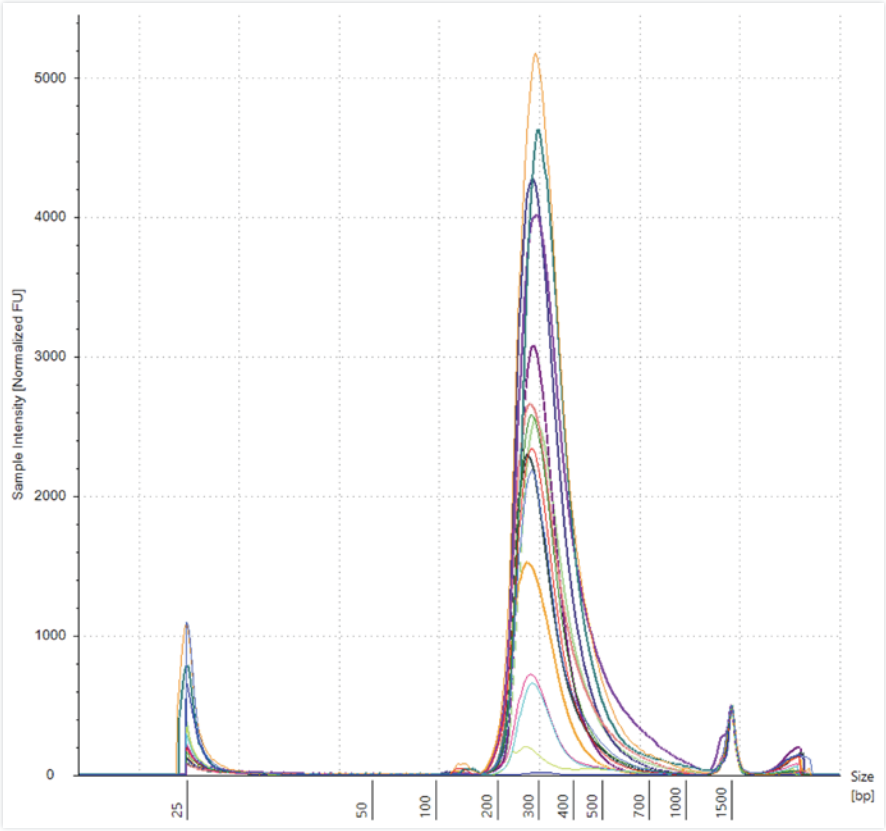

Figure 8. Post-capture NA12878 gDNA (Sixteen sample run including one negative control) library traces for Agilent SureSelectXT libraries assayed with Agilent TapeStation 4200 system with D1000 High Sensitivity ScreenTape.

Library and Sequencing Run	Average Total Reads	Average Read Depth in Target Region	Average % Covered Regions with Zero Coverage	Average % Reads in Covered Regions	Average % Reads +/- 100 bp Covered Region	Average % Reads +/- 200 bp Covered Region
200 ng NA12878 gDNA	194,467,634	365	2.4%	88.5%	89.7%	90%
200 ng HD798 FFPE	158,348,030	306	2.15%	88.8%	90.3%	90.6%
200 ng NA12878 gDNA	149,221,079	289.2	2.35%	90.6%	92%	92.3%

Table 4. Average sequencing depth and region coverage information.

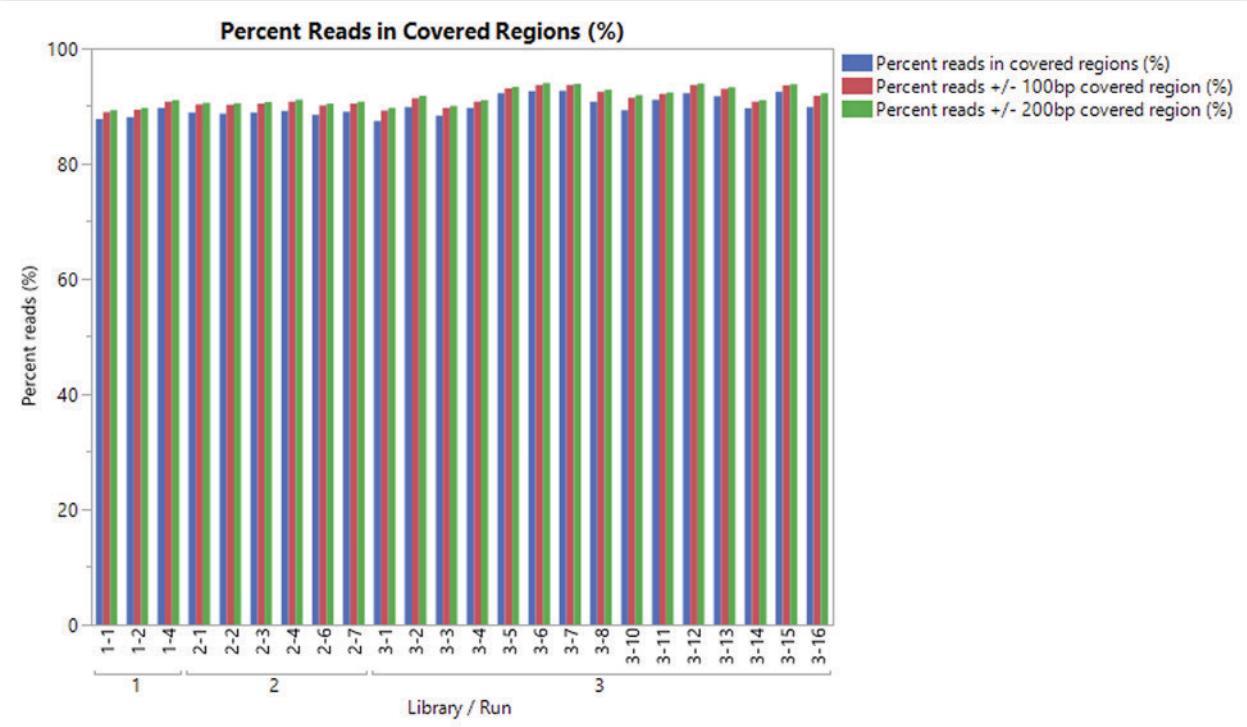

Figure 9. Percent Reads in covered regions with ± 100 bp padding and ± 200 bp padding by library.

For all libraries, 95% of analyzable target regions were covered by at least 20 reads following down sampling (Figure 10). Low numbers of off target reads contribute to this efficiency, an example of which is provided in Figure 11, which shows the read alignment depth around a specific single nucleotide variant in the PIK3CA gene.

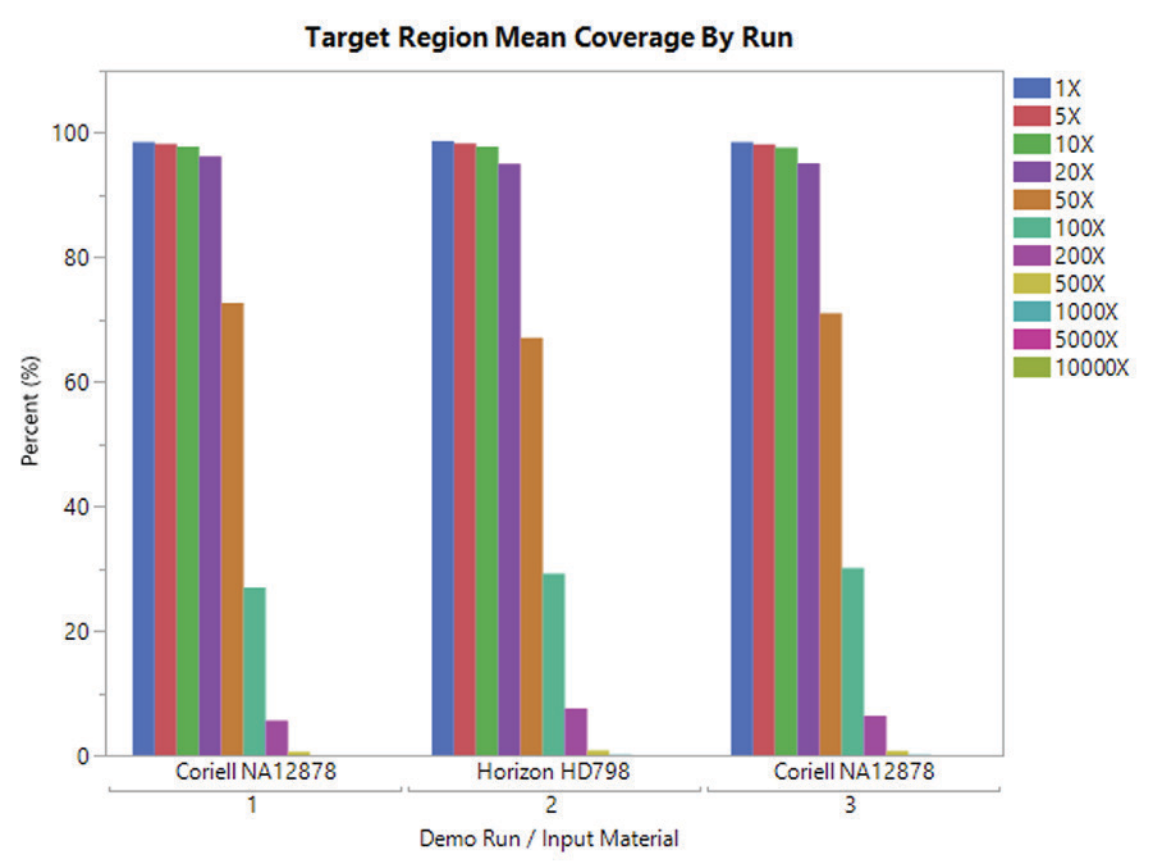

Figure 10. Average Target Region Coverage Depth Per Run.

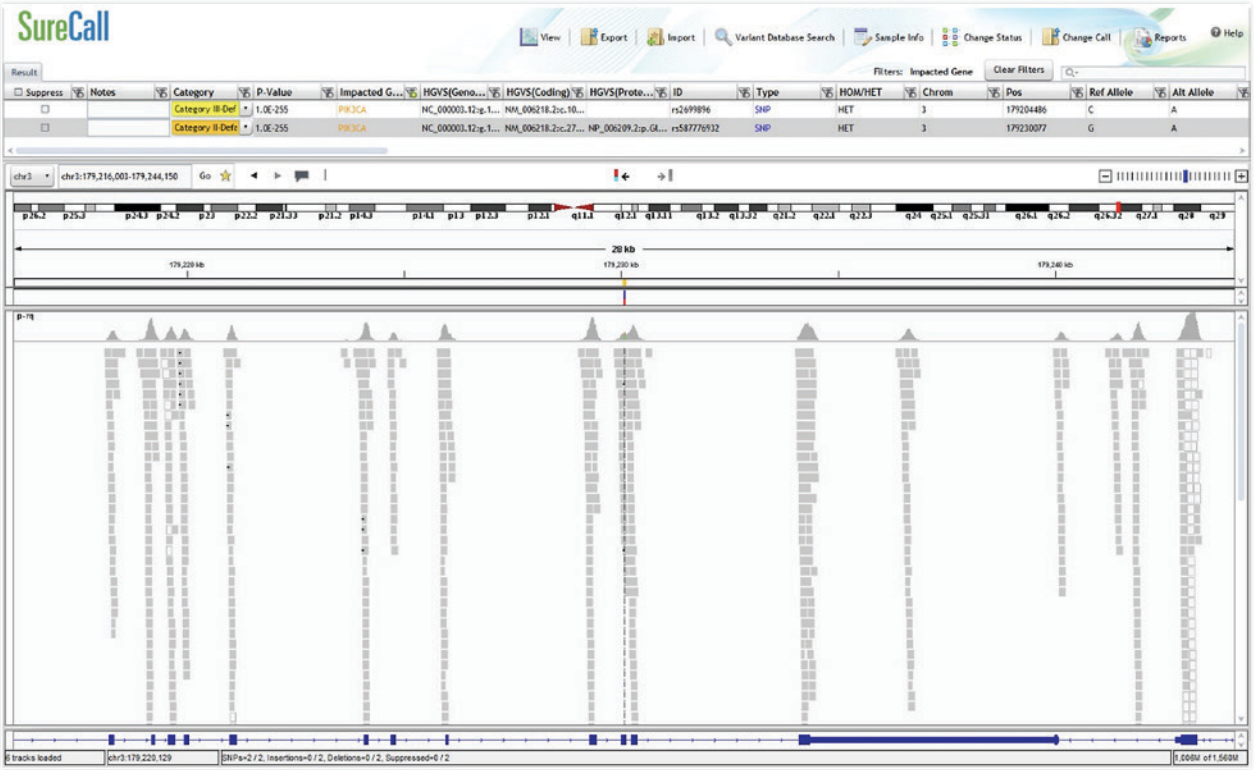

Figure 11. Triage view of a single nucleotide variant in the PIK3CA gene identified in an Agilent SureSelectXT library produced by the Biomek NGeniuS system.

Conclusion

The advent of NGS has helped achieve cost-efficient sequencing of exomes of the 20,000+ human genes. This is a critical step in advancing the field of translational medicine, allowing researchers to identify key genetic differences both between and within patients. One key bottleneck is the laborious and time-intensive nature of DNA library preparation steps prior to sequencing. To address this issue, we automated the Agilent SureSelect XT library prep kit with the Human All Exon V6 panel using a Biomek NGeniuS Next Generation Library Prep System.

The Biomek NGeniuS Next Generation Library Prep System allows users to configure settings for library preparation, select stopping points and upload sample data to a file and mark it ready to run by the instrument. The automated workflow minimizes the chances of introducing operator error within the run, while simultaneously reducing user hands-on time and increasing potential productivity. The on-deck thermal cycler can support long hybridization steps to minimize evaporation and to ensure that the samples and reagents are maintained at the correct temperature.

Libraries generated using the Agilent SureSelect XT library prep kit with the Human All Exon v6 panel on the Biomek NGeniuS system show uniform size distribution on the either the Agilent 2100 Bioanalyzer or Agilent 4200 TapeStation system (Figure 6) and fall within the recommended library size range of the Agilent SureSelect XT library prep kit.

We demonstrated the Biomek NGeniuS Next Generation Library Prep System can successfully produce high-quality whole genome sequencing libraries suitable for sequencing on the Illumina platforms using the Agilent SureSelect XT library prep kit with the Human All Exon V6 panel.

Materials

Sample	Vendor	Part Number
NA 12878 Human gDNA	Coriell Institute	RM8398
Quantitative Multiplex Reference Standard fcDNA (mild)	Horizon Discovery	HD798

Table 5. Sample types and inputs used in preparations of samples Agilent SureSelect XT library prep kit with Human All Exon V6 panel.

Reagents	Manufacturer	Part Number
SureSelect XT reagent kit, MSQ, 16	Agilent Technologies, Inc.	G9612A
SureSelect XT Human All Exon V6, 16 reactions	Agilent Technologies, Inc.	5190-8863
Qubit High Sensitivity Kit	Thermo Fisher Scientific	Q32854
Bioanalyzer High Sensitivity Kit	Agilent Technologies, Inc.	5067-4627
AMPure XP Beads	Beckman Coulter Life Sciences	A63882
PCR grade Water	Invitrogen-Life Technologies	10977-023
Ethanol	AmericanBio	AB00515-0500
Herculase II Fusion DNA Polymerase (includes dNTPs and 5x buffer), 200 reactions	Agilent Technologies, Inc.	600677
BioReagent-grade Mineral Oil	Sigma	M5904-5ML
Dynabeads MyOne Streptavidin T1, 2ml	Thermo Fisher Scientific	65601
1x Low TE Buffer	Thermo Fisher Scientific	12090-015

Table 6. Reagents used in preparation of libraries with Agilent SureSelect XT Human All Exon V6 app template.

Equipment	Manufacturer
Biomek NGenius NGS Sample Prep System	Beckman Coulter Life Sciences
Next Seq Sequencer	Illumina Inc.
Allegra X-14 Centrifuge	Beckman Coulter Life Sciences
Qubit	Thermo Fisher Scientific
Bioanalyzer	Agilent Technologies, Inc.
Vacufuge, SpeedVac	Eppendorf
Covaris S220 Sonicator	Covaris

Table 7. Equipment used in sample preparation and processing Agilent SureSelect XT library prep kit with Human All Exon V6 panel samples.

Consumable	Manufacturer	Part Number
1025 µL Filter Tips	Beckman Coulter Life Sciences	C59585
70 µL Filter Tips	Beckman Coulter Life Sciences	C62712
Biomek NGenius 24 well plates/lids	Beckman Coulter Life Sciences	C62706
Biomek NGenius Bulk reservoirs	Beckman Coulter Life Sciences	C62707
Biomek NGenius seal pads	Beckman Coulter Life Sciences	C70665
Biomek NGenius reagent plugs	Beckman Coulter Life Sciences	
Qubit Tubes	Thermo Fisher Scientific	Q32856
Covaris AFA microtube	Covaris	520045

Table 8. Consumables required for sample processing.

References

1. SureSelect XT Low Input Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library, Version C0, 2018. Agilent Technologies, Inc. G9703-90000.

2022-GBL-EN-100608-v3