Technical Overview

Agilent
Trusted Answers

Comparison of DNA Sample QC for NGS Workflows with the Agilent Fragment Analyzer and Bioanalyzer Systems

Introduction

Next-generation sequencing (NGS) is an essential tool in molecular biology laboratories for the analysis of nucleic acid samples in numerous basic, applied, translational, and clinical research settings. All NGS library preparation protocols have steps throughout the workflow where quality control (QC) checkpoints are needed to ensure downstream success. They help save time and resources by identifying samples that are of poor quality. Sample quality information can also aid in troubleshooting or optimizing library preparation protocols. Recommended QC checkpoints include initial DNA or RNA sample input, after fragmentation, after adapter ligation, and the final library before sequencing. Robust NGS sequencing workflows require the preparation of high-quality libraries for optimal data generation. Checking the quality of the final library for proper size, molarity, and artifacts ensures that it meets sequencing requirements. Agilent provides several instruments to fit the specific nucleic acid sample QC needs for numerous NGS application workflows.

In this technical overview, pre-capture and final libraries were compared on both the Agilent Fragment Analyzer and Bioanalyzer systems with their corresponding DNA assays to demonstrate data equivalency across platforms. The DNA assays selected have similar analytical specifications, allowing for seamless comparison between the instruments (Table 1 and 2). Both systems support standard-sensitivity (SS) and high-sensitivity (HS) kits that cover a wide range of sample concentrations that occur at different QC checkpoints throughout library preparation. Starting DNA material can vary greatly in concentration depending on whether its source is from fresh material or ancient samples. The HS kits are ideal for conservation of low-concentration samples and can be utilized at all recommended QC checkpoints. Alternatively, the SS kit has the advantage of analyzing high-concentration samples while eliminating time-consuming dilution steps. Having the option of both SS and HS kits simplifies the library preparation workflow. The SS and HS kits are compared in this technical overview.

Analytical specifications

Sizing, quantification, and molarity of NGS pre-capture and final libraries were compared between SS and HS Fragment Analyzer and Bioanalyzer kits. The pre-capture libraries were analyzed with the Agilent Fragment Analyzer Small Fragment kit and Bioanalyzer DNA 1000 kit (Table 1). Both kits are suited for the resolution of highly concentrated, small NGS libraries ranging from 50 to 1,000 bp in length. The final libraries were analyzed with the Agilent Fragment Analyzer HS NGS Fragment kit (1-6000 bp) (HS NGS Fragment kit) and the Bioanalyzer High Sensitivity DNA kit (HS DNA kit), which are ideal for traditional short-read NGS libraries ranging from 100 to 6,000 bp in length (Table 2).

Experimental

NGS pre-capture libraries 97, 103, and 111

The NGS pre-capture libraries 97, 103, and 111 were prepared with Agilent SureSelect XT HS kit (p/n 5191-6764). Average sizing, concentration, and molarity were compared on the Agilent 5200 Fragment Analyzer system (p/n M5310AA) with the Small Fragment kit (p/n DNF-476) (n=3) and Agilent 2100 Bioanalyzer system (p/n G2939BA) with the DNA 1000 kit (p/n 5067-1504) (n=3). Sample concentration was determined by the Fragment Analyzer system, the Bioanalyzer system, and the Qubit dsDNA HS assay (Thermo Fisher Scientific, p/n Q32854) and diluted in 1x TE buffer (pH 7.4) to ensure sufficient amount of material for electrophoretic analysis in triplicates for the study. A region or smear analysis from 170 to 1,000 bp was utilized for comparison of size, concentration, and molarity between the systems.

NGS final libraries 20, 21 and 22

NGS final libraries 20, 21, and 22 were prepared from Universal Reference Mouse RNA (Agilent, p/n 740100) using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England BioLabs, p/n E7530S). Following cleanup with SPRI beads to eliminate dimer artifacts, the concentration of each library was determined with the Fragment Analyzer system, the Bioanalyzer system, and a Qubit 4.0 Fluorometer with the dsDNA HS Assay kit (Thermo Fisher Scientific, p/n Q32851). The libraries were analyzed using the 5200 Fragment Analyzer system with the Agilent HS NGS Fragment kit (1-6000 bp) (p/n DNF-474) (n=9) and the 2100 Bioanalyzer system with the Agilent HS DNA assay (p/n 5067-4626) (n=3). A region or smear analysis from 150 to 3,000 bp was utilized for comparison of size, concentration, and molarity between the systems.

Table 1. Comparison of analytical specifications of the Agilent Small Fragment kit and the DNA 1000 kit used to analyze NGS pre-capture libraries.

Analytical Specifications	Fragment Analyzer System Small Fragment kit	Bioanalyzer System DNA 1000 kit
Sizing range	50 to 1,500 bp	25 to 1,000 bp
Sizing accuracy	±5% or better[1,3]	±10%[4]
Sizing precision	2% CV[1,2,3]	5% CV[4]
Quantitative range (DNA fragments)	0.1 to 10 ng/µL[1]	0.5 to 50 ng/µL[4]
Quantitative range (smears)	5 to 100 ng/µL[2]	-
Quantitative accuracy	±25%[1]	±20%[4]
Quantitative precision	15% CV[2]	25 to 500 bp: 15% CV[4] 500 to 1,000 bp: 5% CV[4]

[1] *Results using 400 bp DNA fragment standard in 1x TE buffer.*
[2] *Results using sheared gDNA with smear range from 10 - 1,400 bp in 1x TE buffer.*
[3] *Results using DNA Ladder in 1x TE buffer.*
[4] *Results using the respective DNA ladder as sample.*

Table 2. Comparison of analytical specifications of the Agilent HS NGS Fragment kit (1-6000 bp) and the High Sensitivity DNA kit used to analyze final libraries.

Analytical Specifications	Fragment Analyzer System HS NGS Fragment kit	Bioanalyzer System HS DNA kit
Sizing range	100 to 6,000 bp	50 to 7,000 bp
Sizing accuracy	±5% or better[1]	±10%[2]
Sizing precision	2% CV[1]	5% CV[2]
Quantitative range (DNA fragments)	5 to 500 pg/µL[1]	5 to 500 pg/µL[2]
Quantitative range (smears)	50 to 5,000 pg/µL[1]	-
Quantitative accuracy	±25%[1]	±20%[2]
Quantitative precision	15% CV[1]	5 to 2,000 bp: 15% CV[2] 2,000 to 7,000 bp: 10% CV[2]

[1] *Results using DNA fragments and smears in 1x TE buffer.*
[2] *Results using the respective DNA ladder as sample.*

Results and discussion

Comparison of NGS pre-capture libraries

Sizing of pre-capture libraries

Electrophoretic analysis of the pre-capture libraries is one of the main QC checkpoints in the workflow of the Agilent SureSelect target enrichment system. Monitoring the pre-capture library size distribution ensures that the samples meet the qualifications for the next step in the library preparation process of hybridization with a target specific probe (capture library). Electrophoretic separation of NGS pre-capture libraries 97, 103, and 111 was compared between the Fragment Analyzer system with the Small Fragment kit and the Bioanalyzer system with the DNA 1000 kit. As shown in the overlays, the electrophoretic traces of all three libraries on both platforms exhibited similar distribution (Figure 1).

There are two different approaches to assign a size to a library. The peak size is the highest point or largest size in the smear. A typical library has a bell-shaped curve and the peak size would be directly in the middle of the library smear. Alternately, a library size can be assigned by smear range analysis. A base pair range that includes the entire library is chosen and the average smear size is calculated based on the percentage of fragments at each base pair throughout the library. The smear range provides a very accurate average size and is advantageous for libraries with smearing or tailing to the left or right.

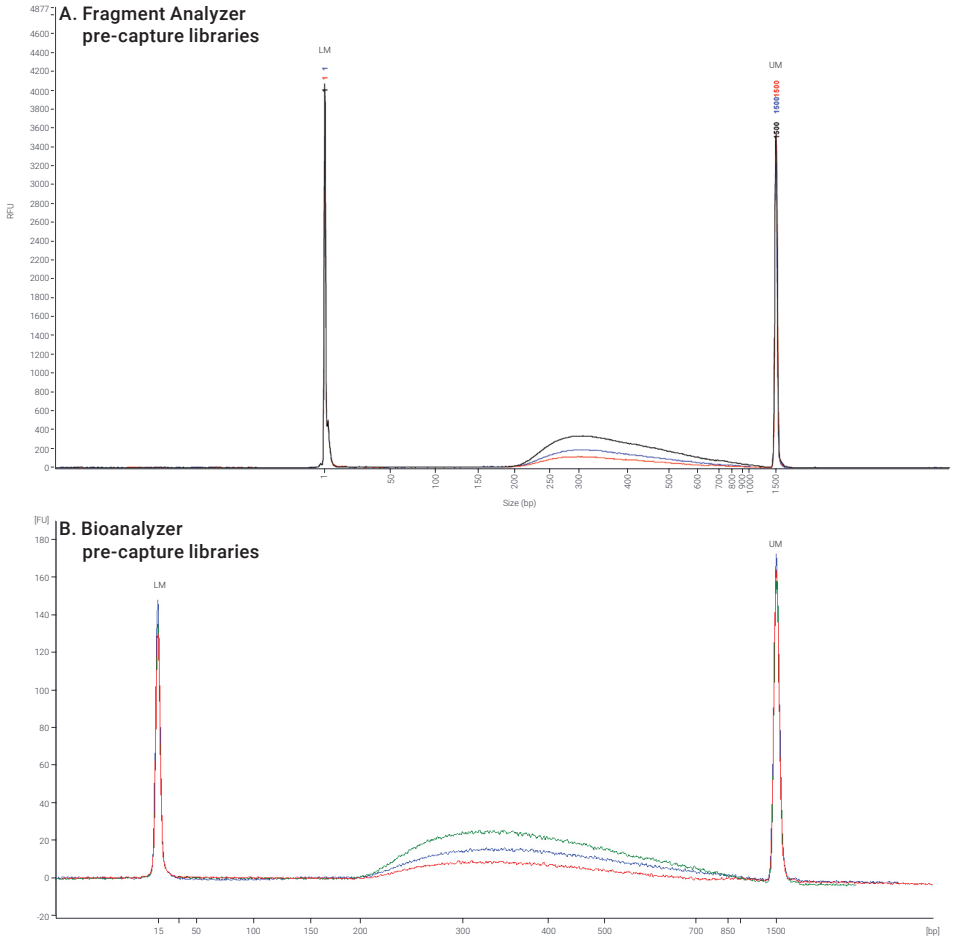

Figure 1. Electropherogram overlays of NGS pre-capture libraries 97, 103, and 111 analyzed on (A) Agilent Fragment Analyzer system with the Small Fragment kit (library 111: black; library 97: blue; library 103: red) and (B) Agilent Bioanalyzer with DNA 1000 kit (library 111: green; library 97: blue; library 103: red).

To determine an average size for each NGS pre-capture library, an identical smear range from 170 to 1,000 bp was defined for both automated electrophoresis platforms using the respective Fragment Analyzer ProSize data analysis software and Bioanalyzer 2100 Expert software features. The example electropherograms for library 111 with defined smear or region analysis of 170 to 1,000 bp on the Fragment Analyzer system (A) and the Bioanalyzer system (B) are shown in Figure 2. The average smear size for the three NGS pre-capture libraries remained consistent, averaging 9.5% difference between the two DNA assays (Figure 3). Individually, both kits demonstrated reproducible sizing with a maximum coefficient of variation (%CV) or precision of 2.0% and 2.1%, respectively. Sizing precision between the two assays was comparable and met the corresponding analytical specifications of each kit.

The SureSelect XT HS2 manual provides qualification guidelines for QC steps throughout the library preparation process[1]. Pre-capture libraries prepared from high-quality gDNA and 2×150 reads should have an average smear size between 330 to 450 bp. Both kits reported an average library size within the recommended range. These pre-capture libraries would be eligible for the next step of hybridization, in the library preparation process.

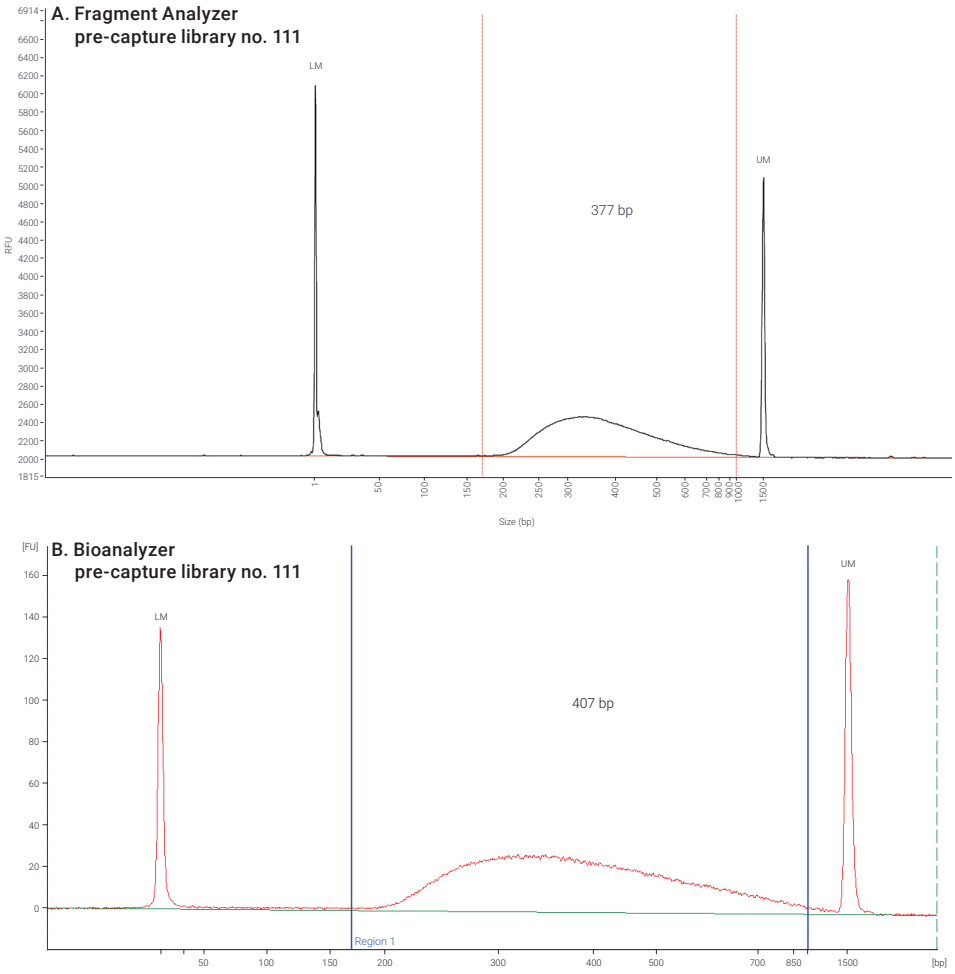

Figure 2. Electropherograms of NGS pre-capture library 111 with defined smear range from 170 to 1,000 bp analyzed on (A) Agilent Fragment Analyzer system with the Small Fragment kit and (B) Agilent Bioanalyzer system with DNA 1000 kit.

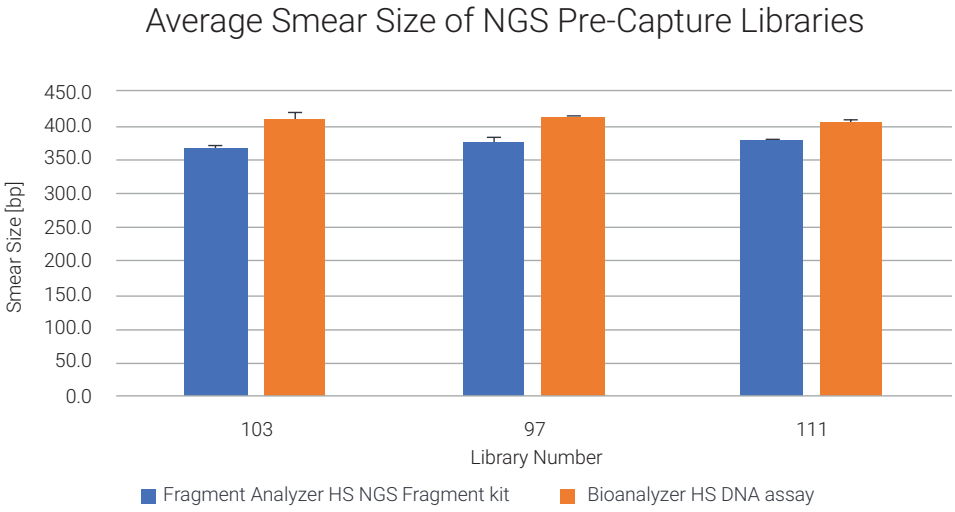

Figure 3. The average smear size of NGS pre-capture libraries 97, 103, and 111 was determined by setting the analytical range from 170 to 1,000 bp using the respective Agilent Fragment Analyzer ProSize data analysis software and Agilent 2100 Expert software smear analysis features.

Quantification of pre-capture libraries

Pre-capture libraries can be analyzed on both the Small Fragment kit and DNA 1000 kit without dilution, saving time in the QC library preparation workflow. For this study, the stock pre-capture libraries 97, 103, and, 111 were diluted to approximately 5, 10, and 15 ng/µL, within the recommended sample input range of both kits, and analyzed on both automated electrophoresis platforms. Comparison of the reported average library concentrations demonstrated high similarity between each kit and the Qubit (Figure 4). Both kits met the respective analytical specifications, with an 18% error compared to Qubit. In addition, the precision or %CV for both platforms did not exceed 15%.

Molarity of pre-capture libraries

The SureSelect XT HS2 library preparation protocol supports deep multiplex sequencing. Unique indexes are added to pre-capture libraries during a PCR amplification step, allowing the user to perform optional pooling with up to 32 samples. Libraries are combined based on their individual molarities and a final pool is prepared at a molarity specific to the particular sequencing application. Thus, successful sequencing of a library pool relies upon accurate molarity measurements. The molarity is automatically reported for both instruments. The molarity values for each pre-capture library were compared between the kits and were found to report similar molarity results (Figure 5) with precision (%CV) below 16% for the Small Fragment kit and below 11% for the DNA 1000 kit.

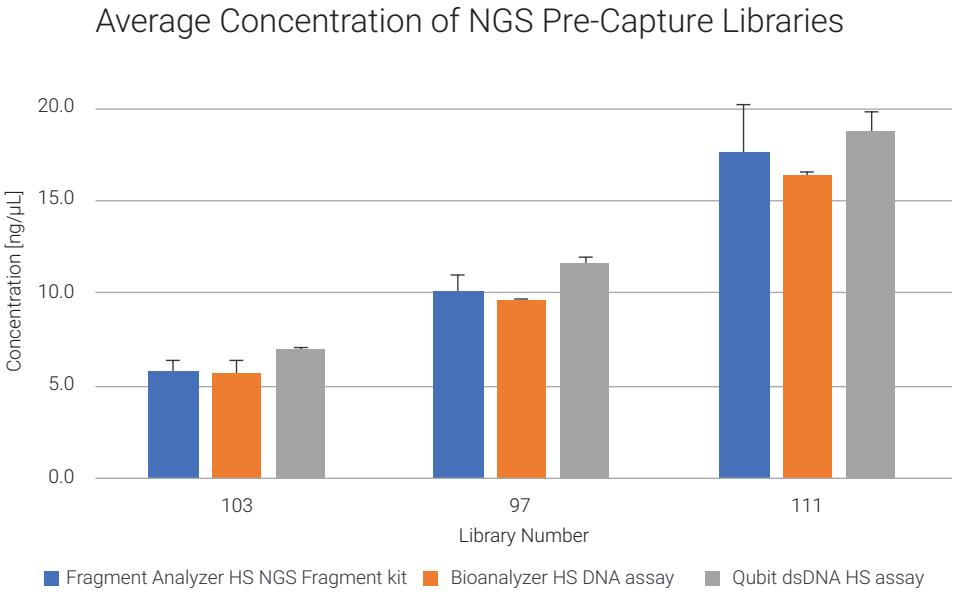

Figure 4. Comparison of concentrations for NGS pre-capture libraries 97, 103, and 111 obtained on the Agilent Fragment Analyzer with the Small Fragment kit, the Agilent Bioanalyzer system with the DNA 1000 kit, and the Qubit 4.0 with the dsDNA HS assay.

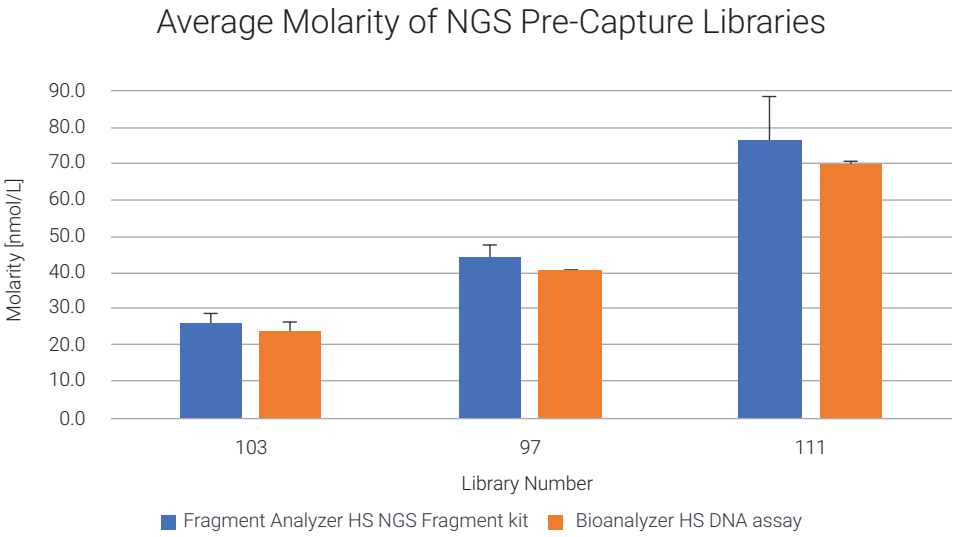

Figure 5. Comparison of the average molarity values for NGS pre-capture libraries 97, 103, and 111 obtained on the Agilent Fragment Analyzer and the Bioanalyzer systems using the Agilent Small Fragment kit and DNA 1000 kit, respectively.

The Small Fragment kit and DNA 1000 kit for the Fragment Analyzer and the Bioanalyzer systems offered comparable analysis results without the need for dilution. Sizing of the pre-capture libraries is an important intermediate QC step in the SureSelect XT HS2 library preparation. Both kits reported a similar average smear size for the libraries, as well as providing consistent concentrations and molarity. Either instrument and its corresponding SS kit will provide reliable QC data for pre-capture libraries.

Comparison of the NGS final libraries

Sizing of final library

Library preparation kits are designed to construct libraries within a specified size range required by the sequencing application. It is important to check the library size and screen for any artifacts to ensure successful sequencing. Adapter dimers are the most common artifact found in final libraries and are the result of self-ligation without a library insert. These dimers form clusters very efficiently and consume valuable space on the sequencing flow cell. Adapter dimers limit sequencing reads produced from the libraries and negatively impact sequencing data quality. Sequencing applications often specify the percent of adapter dimer allowed in a library sample to ensure successful sequencing data. The NEBNext Ultra RNA Library Prep Kit for Illumina manual sets qualification guidelines for QC steps throughout the library preparation process[2]. Noted in the manual is the need to check for the presence of primers and adapter dimers. If a peak at approximately 80 bp (primers) or 128 bp (adapter dimer) is present, an additional AMPure XP Bead cleanup step is required.

NGS final libraries 20, 21, and 22 were separated on the Fragment Analyzer system with the HS NGS Fragment kit and on the Bioanalyzer system with the HS DNA kit. Both kits support similar sizing ranges of 50 to 6,000 bp and 50 to 7,000 bp, respectively. The three libraries displayed similar electropherogram traces (Figure 6). A small amount of adapter dimer was detected in libraries 21 and 22 on the Fragment Analyzer system (A) and the Bioanalyzer system (B). The percentage of adapter dimer averaged less than 0.005% of the total molarity (pmol/L), indicating that both instruments have very high sensitivity, enabling them to detect such a small amount of adapter dimer contamination after cleanup with SPRI beads.

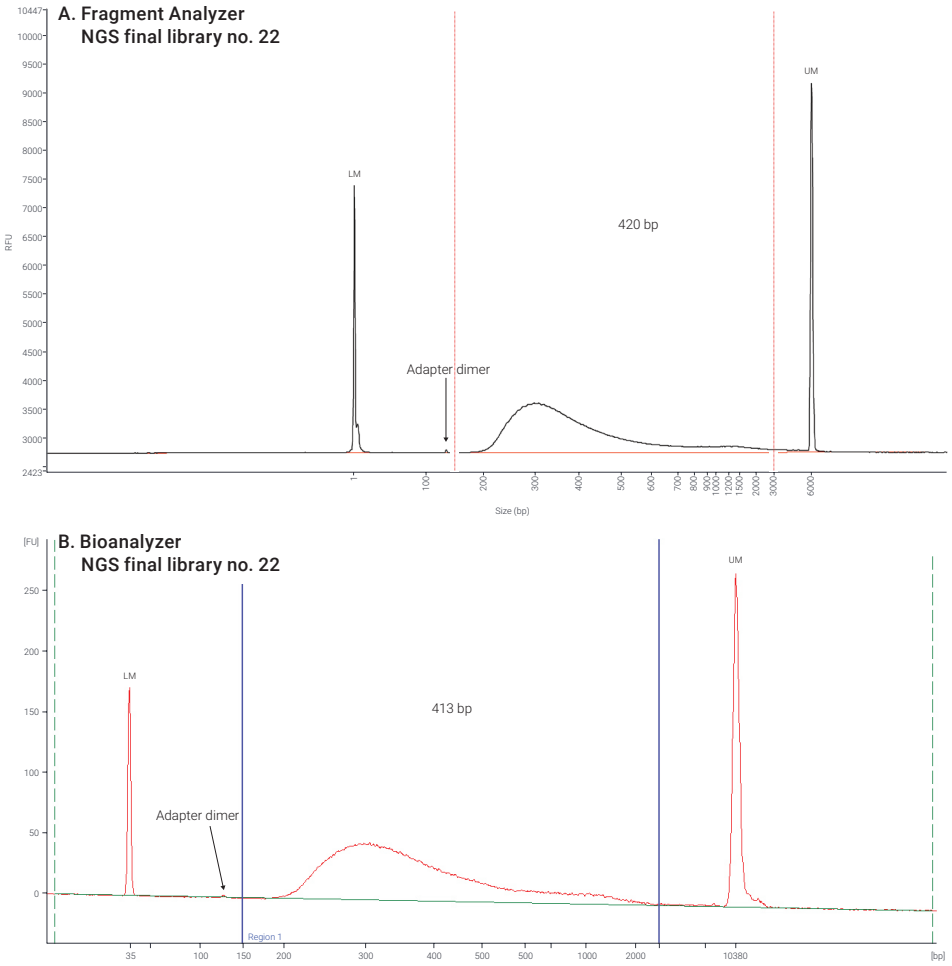

Figure 6. NGS final library 22 separated (A) on the Agilent Fragment Analyzer system with the HS NGS Fragment kit (1-6000 bp) and (B) on the Agilent Bioanalyzer system with the High Sensitivity DNA kit. The two instruments displayed similar traces and detected small amounts of adapter dimers.

The average size of each library was determined by smear or region analysis set at 150 to 3,000 bp. The Fragment Analyzer and Bioanalyzer systems reported similar average sizes between the three libraries, with small differences ranging from only 9 to 22 bp between the systems (Figure 7). Precision was very tight, with a percent CV below 3.5% for both instruments.

Quantification of final library

The concentration of the NGS final libraries 20, 21, and 22 was measured on the Fragment Analyzer system with the HS NGS Fragment kit, on the Bioanalyzer system with the HS DNA kit, and on the Qubit 4.0 with the dsDNA HS assay for comparison. The three instruments reported very similar concentrations for all three libraries (Figure 8). Quantification accuracy when compared to the Qubit was below 6% error for both instruments. In addition, both the Fragment Analyzer and Bioanalyzer systems displayed very tight precision with 6% CV, well within kit specifications.

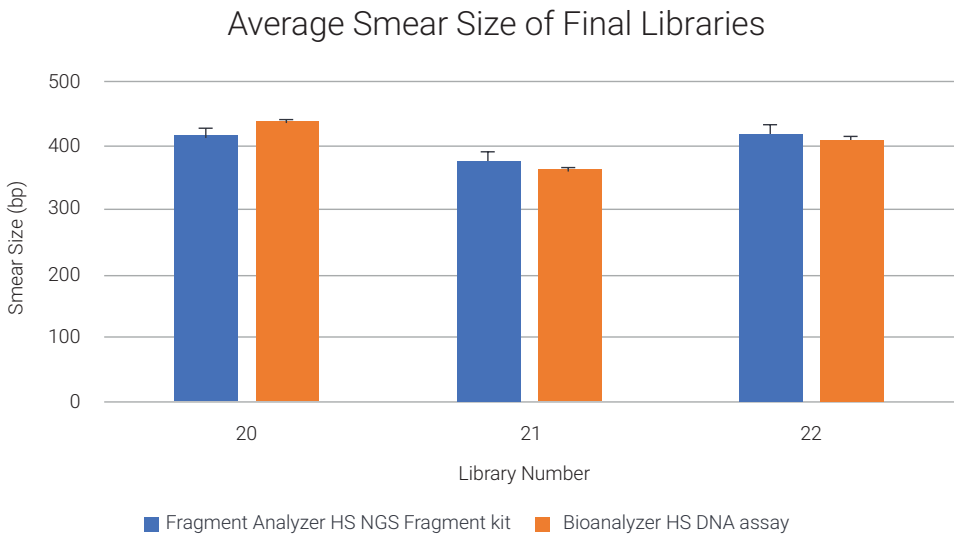

Figure 7. Average smear sizing of final libraries 20, 21, and 22 on the Agilent Fragment Analyzer system with the HS NGS Fragment (1-6000 bp) kit and on the Agilent Bioanalyzer system with the High Sensitivity DNA kit. The instruments reported similar sizing with high precision.

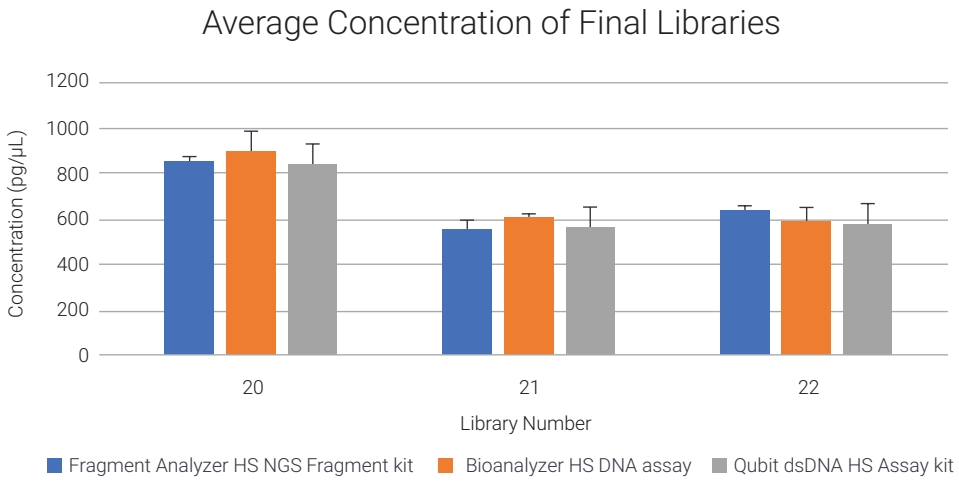

Figure 8. Average concentration of final libraries 20, 21, and 22 on the Agilent Fragment Analyzer system with the HS NGS Fragment (1-6000 bp) kit, on the Agilent Bioanalyzer system with the High Sensitivity DNA kit, and the Qubit with dsDNA HS assay.

Molarity of final library

The amount of library loaded onto a flow cell is an important parameter of the Illumina sequencing system, as it influences the density of the clusters that form. Loading too little library leads to under-clustering in the flow cell and results in lower data output. In contrast, loading too much library allows clusters to form too close together (over-clustering), resulting in poor image resolution and affecting the quality of the sequencing data. Over-clustered flow cells have lower Q30 scores and reduced data output. Molarity is used to measure the amount of DNA loaded, considering both sizing and concentration. Differences in both sizing and concentration can greatly change the molarity of the final library, affecting the input amount for sequencing. The Fragment Analyzer and Bioanalyzer systems reported similar molarities for all three libraries, with the precision (Figure 9) averaging below 6% CV for both instruments. The small difference in molarity indicates consistent measurements in sizing and concentration between the two instruments.

Both the Fragment Analyzer and Bioanalyzer systems provided reliable QC analysis of NEBNext final libraries with similar size, quantification, and molarity for monitoring NGS library preparation and flow cell loading. The high sensitivity capabilities of both instruments allowed for minute detection of adapter dimers in the final library samples.

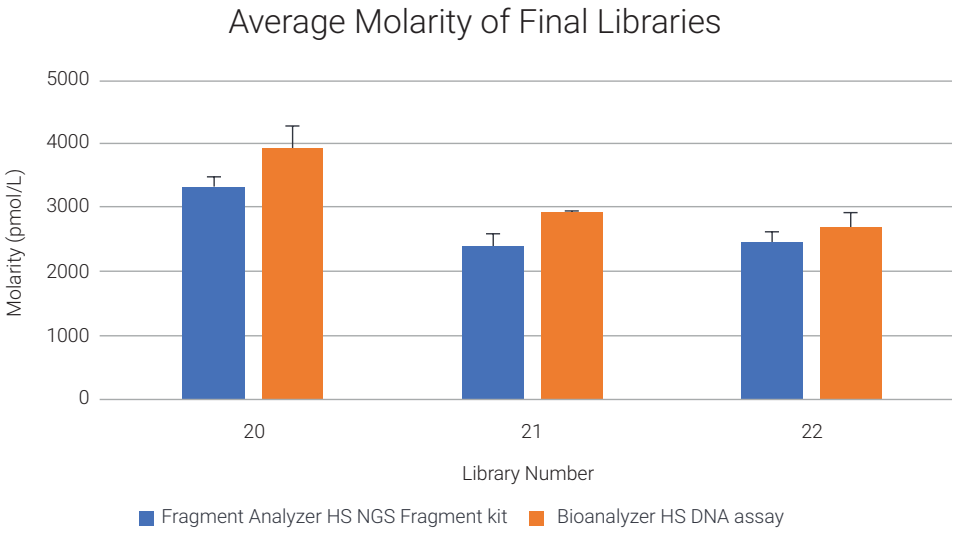

Figure 9. Average molarity of final libraries 20, 21, and 22 on the Agilent Fragment Analyzer system with the HS NGS Fragment (1-6000 bp) kit and on the Agilent Bioanalyzer system with the High Sensitivity DNA kit.

Conclusion

Quality control throughout the NGS library preparation workflow is critical to ensure successful downstream sequencing. The Agilent Fragment Analyzer and Bioanalyzer systems provide reliable QC analysis for monitoring size, quantity, and molarity of the sample at different steps in the NGS workflow for two different library kits. Standard-sensitivity and high-sensitivity kits were compared between the instruments to demonstrate compatibility across the platforms. The Small Fragment kit of the Fragment Analyzer system and the DNA 1000 kit of the Bioanalyzer system reported similar sizing, quantification, and molarity of NGS pre-capture libraries. In addition, the HS NGS Fragment kit of the Fragment Analyzer system and the HS DNA kit of the Bioanalyzer system also reported similar sizing, quantification, and molarity for NGS final libraries. Reliable and comparable QC results can be obtained from either the Fragment Analyzer system or the Bioanalyzer system, providing ease of mind when gathering data between two labs or transferring QC protocols from one system to the other.

References

1. SureSelect XT HS2 DNA System DNA Library Preparation and Target Enrichment for Illumina Paired-End Multiplexed Sequencing, *Agilent Technologies Manual*, publication number G9983-90000, **2020**.
2. NEBNext Ultra RNA Library Prep Kit for Illumina, *New England BioLabs Manual*, publication number E7530S/L, **2020**.

www.agilent.com/genomics/automated-electrophoresis

For Research Use Only. Not for use in diagnostic procedures.
PR7000-7525

This information is subject to change without notice.

Printed in the USA, October 15, 2020
5994-2459EN

Agilent
Trusted Answers