Application Note

Genomics

Agilent
Trusted Answers

Quality Control of NGS Libraries with Daisy Chains

Authors

Carmen Rothmund,
Roche Diagnostics
Deutschland GmbH

Vera Rykalina and Rainer Nitsche,
Agilent Technologies

Abstract

The Agilent automated electrophoresis solutions offer fast and reliable quality control of next-generation sequencing (NGS) libraries. The Bioanalyzer, TapeStation, and Fragment Analyzer systems with dedicated high sensitivity assays allow unambiguous detection of daisy chains and accurate sizing of the target library. In this application note, we provide a recommendation for which assay to employ to reliably visualize daisy chain products in NGS libraries. Furthermore, we emphasize the consistency between the instruments and reproducibility of analysis confirmed by results on a set of two-fold serial dilutions.

Introduction

DNA library preparation for next-generation sequencing (NGS) is a standardized procedure that generally includes fragmentation of input material and ligation of functional adapters to the resulting DNA fragments. To increase the proportion and total amount of adapter-ligated DNA fragments, and to incorporate functional platform-specific sequences, most of the library construction protocols require a certain number of PCR amplification cycles. The extent of library amplification, however, should be optimized and reduced to a minimum to prevent undesired PCR artifacts. Furthermore, excessive amplification after substrate depletion can result in the formation of so-called ‘daisy chains’, or heteroduplex molecules. These heteroduplex molecules are comprised of partially single-stranded DNA fragments, hybridized to each other via their complimentary adapter sequences (Figure 1). Daisy chains migrate more slowly during electrophoretic analysis and cannot be correctly quantified by fluorometric methods. However, accurate qualification and quantification of NGS libraries is crucial to avoid over- or underclustering of the sequencing flow cell and thus guarantees the best sequencing performance.

Experimental

NGS libraries were generated with the KAPA HyperPlus kit (Roche, p/n 07962401001), using 100 ng of genomic DNA (Roche, p/n 11691112001). Library 1 was amplified with five PCR cycles according to the manufacturer’s protocol[1], whereas library 2 was deliberately overamplified with four additional rounds of thermocycling. The amplified samples were subjected to 1X bead-based cleanup and eluted in 50 µL of 10 mM Tris-HCl (pH 8.0). Final libraries were quantified using the Qubit 2.0 Fluorometer with the dsDNA HS assay kit (Thermo Fisher Scientific, p/n Q32854). Based on the Qubit records, a set of two-fold serial dilutions were prepared for each library. Electrophoretic sample preparation and analysis were performed according to the manufacturer’s protocols and recommendations[2,3,4]. The 2100 Bioanalyzer system (p/n G2939BA) with the High Sensitivity DNA kit (p/n 5067-4626), 4200 TapeStation system (p/n G2991AA) with the High Sensitivity D5000 ScreenTape (p/n 5067-5592) and reagents (p/n 5067-5593), and 5200 Fragment Analyzer system (p/n M5310AA) with the HS Small Fragment kit (p/n DNF-477) from Agilent Technologies were used for analysis of NGS libraries.

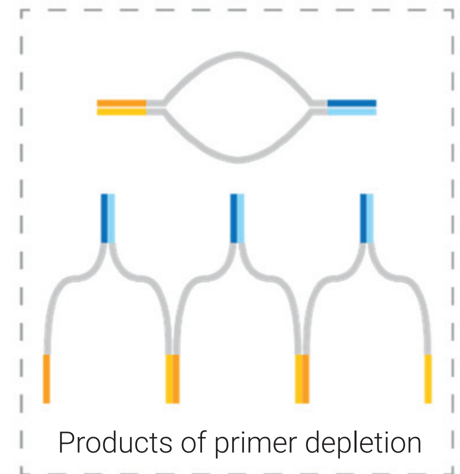

Figure 1. Schematic representation of daisy chains (used by permission of Roche Diagnostics Deutschland GmbH). The daisy chains are comprised of strands of library molecules with non-complementary inserts annealed to each other via their complimentary adapter sequences (highlighted in color).

Results and discussion

Electrophoretic profiles

Two KAPA HyperPlus libraries generated with different levels of amplification were analyzed on the Bioanalyzer (Figure 2A), TapeStation (Figure 2B), and Fragment Analyzer (Figure 2C) systems using the High Sensitivity DNA kit, High Sensitivity D5000 ScreenTape assay, and HS Small Fragment kit, respectively. As shown in the overlays (Figure 2), both libraries demonstrated the desired size distribution with a pronounced secondary product in library 2. The excessive amplification of library 2 resulted in the formation of daisy chains, which were observed as an additional higher molecular weight product in all electropherograms. The daisy chains migrated more slowly through the gel matrix and were easily detected by all Agilent automated electrophoresis instruments using the respective assays. However, additional testing revealed that daisy chains cannot be reliably detected by the High Sensitivity D1000 ScreenTape assay (TapeStation) and the HS NGS Fragment kit (Fragment Analyzer). Thus, we recommend using the Bioanalyzer system with the High Sensitivity DNA kit, the TapeStation systems with the High Sensitivity D5000 ScreenTape assay, and the Fragment Analyzer systems with the HS Small Fragment kit for quality control of NGS libraries that possibly contain daisy chains.

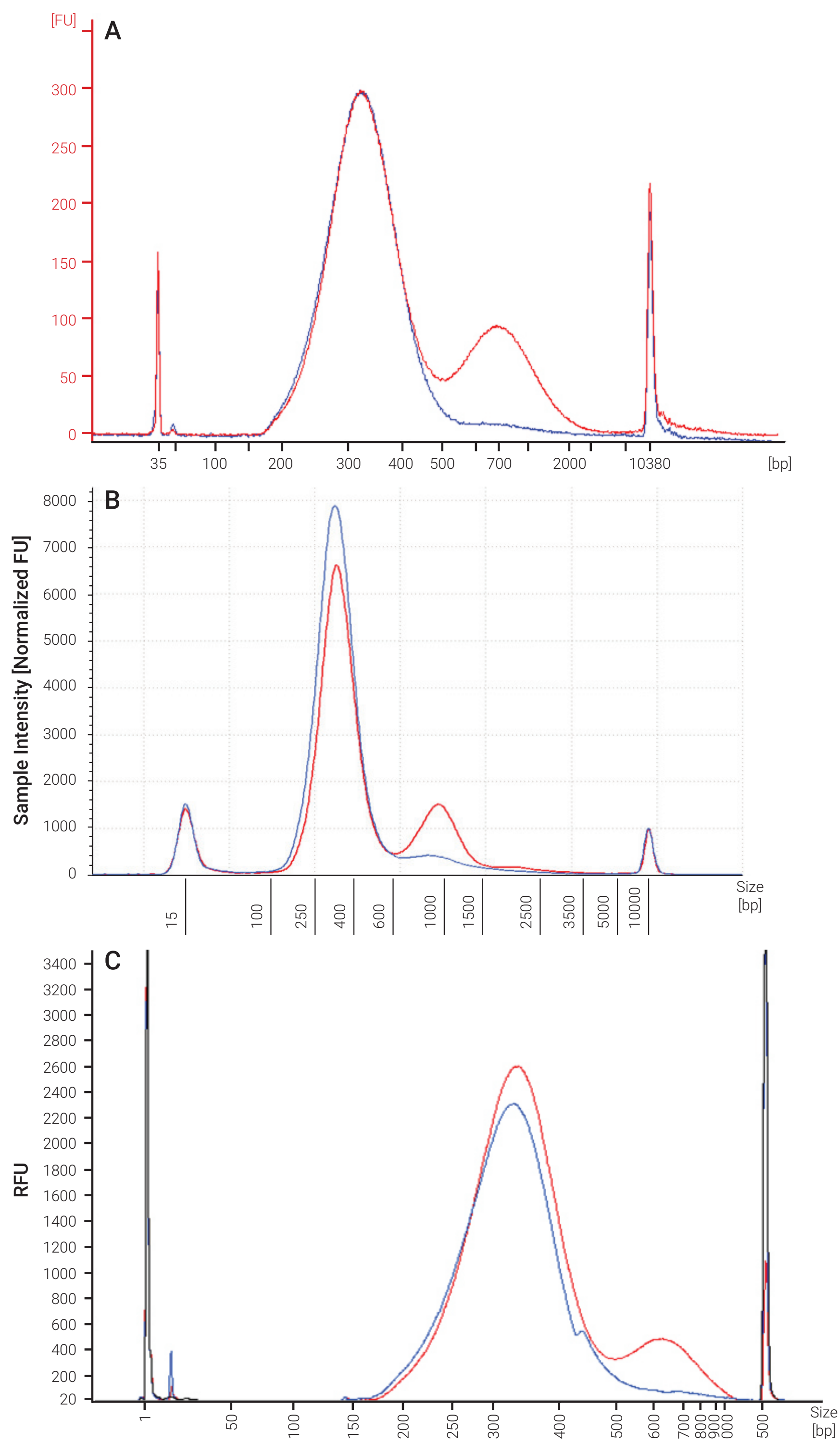

Figure 2. Electropherogram overlays of the Kapa HyperPlus libraries 1 (blue) and 2 (red) analyzed with the Agilent 2100 Bioanalyzer (A), 4200 TapeStation (B), and 5200 Fragment Analyzer (C) systems using the High Sensitivity DNA kit, High Sensitivity D5000 ScreenTape assay, and HS Small Fragment kit respectively.

Table 1. Library size compared between different platforms and sample concentration.

Dilution Series [ng/µL]	Average Size [bp]		
	Agilent 2100 Bioanalyzer System	Agilent 4200 TapeStation System	Agilent 5200 Fragment Analyzer System
	High Sensitivity DNA kit	High Sensitivity D5000 ScreenTape Assay	HS Small Fragment kit
4	332	341	330
2	333	339	332
1	333	341	330
0.5	332	354	331

Sizing

The KAPA HyperPlus Kit from Roche recommends the Agilent automated electrophoresis solutions over conventional slab gels to monitor the size distribution of precapture and final libraries[1]. Table 1 compares the average size for library 2 across the three electrophoresis platforms. Regions can be easily established based on the individual electropherogram traces using the region analysis feature of the corresponding Agilent analysis software. Figure 3 illustrates an example of the region settings within the TapeStation software used for analysis of the library size distribution. The boundaries are set to flank the library smear, excluding artifactual daisy chain products. Quality control of NGS libraries through size distribution analysis with all used assays verified the expected average library size in the range from 300 to 350 bp. The consistency of the results was demonstrated by analysis of a set of two-fold serial dilutions. Figure 4 shows the overlay of library 2 analyzed in four different concentrations on the 2100 Bioanalyzer instrument with the High Sensitivity DNA kit. All dilution points had highly similar average library size across the three automated electrophoresis systems.

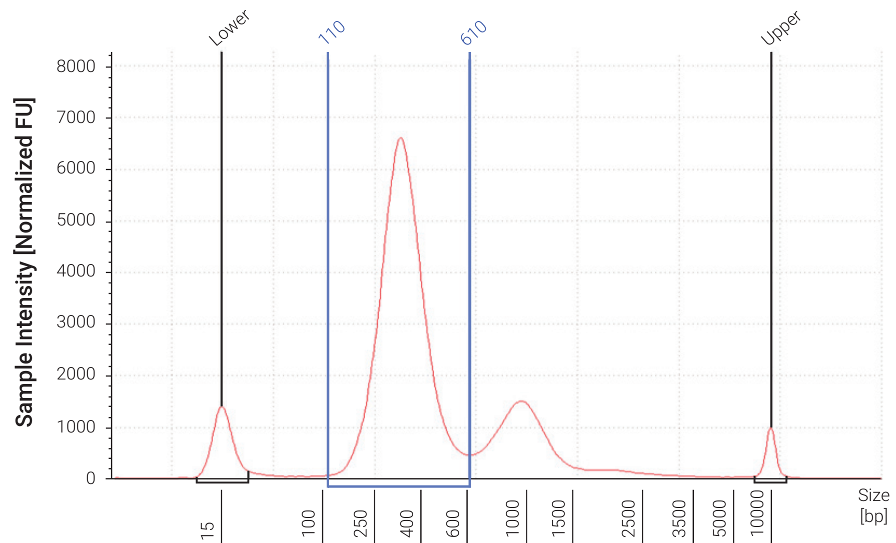

Figure 3. The region analysis feature of the TapeStation software was utilized for analysis of the size distribution of library 2 (4 ng/µL). The region from 110 to 610 bp was established visually based on the shape of the electropherogram trace.

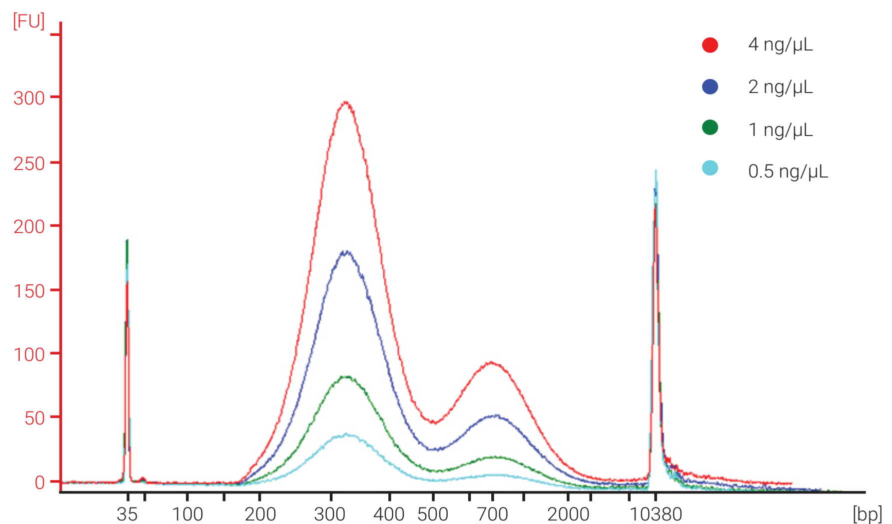

Figure 4. Bioanalyzer electropherogram overlay of the dilution series for library 2. All dilution points exhibit a target smear of the NGS library and a secondary product corresponding to daisy chain conformation. At higher concentrations, daisy chains are more pronounced.

Quantification and sequencing

The DNA High Sensitivity kits and assays for automated electrophoresis systems are designed to provide an accurate quantification of double-stranded DNA molecules, and the results are confirmed to be consistent with fluorometric-based methods[5]. However, daisy chains contain significant portions of intramolecular single-stranded DNA, and using assays optimized for double-stranded DNA will lead to underestimation of the library concentration. This will result in overloading of the sequencing flow cell, and thus a negative impact on the sequencing quality. In contrast, library quantification by qPCR can be an alternative method even when daisy chains are observed in the generated NGS libraries. qPCR denatures library fragments into single-stranded DNA molecules, allowing for an accurate measure of both the target library and secondary daisy chain products.

Normally, reamplification of a library with daisy chains with new PCR reagents including primers for 2–3 PCR cycles can help to resolve the daisy chain conformation[1]. This simple procedure can also be used to distinguish daisy chain products from other PCR artifacts. Nevertheless, the primary goal is to avoid the formation of daisy chains by limiting the number of PCR cycles to a minimum and enabling accurate quantification with fluorometric-based methods featuring Agilent automated electrophoresis instruments.

The sequencing of NGS libraries even with the presence of daisy chains can be done without significant effects, providing that library concentration is correctly measured before loading on the flow cell. The denaturation step prior to flow cell binding removes the artificial daisy chain conformation and provides unimpeded probe hybridization and bridge amplification. The daisy chain molecules, being single stranded after denaturation, exhibit the desired size distribution.

Conclusion

The Agilent automated electrophoresis product portfolio enables effective and robust quality control of NGS libraries with daisy chain molecules for the Agilent Bioanalyzer system with the High Sensitivity DNA kit, the TapeStation systems with the High Sensitivity D5000 ScreenTape assay, and the Fragment Analyzer systems with the HS Small Fragment kit. These High Sensitivity kits and assay provided excellent resolution for simple identification of both the desired NGS library and the unwanted daisy chains generated with the KAPA HyperPlus kit. The Bioanalyzer, TapeStation, and Fragment Analyzer systems offer accurate sizing, which is an essential parameter for the calculation of the library molarity. The average library size showed a high level of correlation across the instruments and the entire dilution series.

References

1. KAPA HyperPlus Kit Technical Data Sheet, *KAPA Roche*, publication number KR1145-v4.17, **2017**.
2. Agilent High Sensitivity DNA Quick Start Guide, *Agilent Technologies*, publication number G2938-90322, **2020**.
3. Agilent High Sensitivity D5000 ScreenTape Quick Guide for TapeStation Systems, *Agilent Technologies*, publication number G2991-90151, **2018**.
4. Agilent HS Small Fragment Kit Quick Guide for the 5200 Fragment Analyzer System, *Agilent Technologies*, publication number M5310-92477, **2018**.
5. Monitoring Library Preparation for Next-Generation Sequencing in Systems Biology Omics Analysis, *Agilent Technologies Application Note*, publication number 5994-0946EN, **2019**.

www.agilent.com/genomics/automated-electrophoresis

For Research Use Only. Not for use in diagnostic procedures.

This information is subject to change without notice.

PR7000-7416

© Agilent Technologies, Inc. 2020

Printed in the USA, August 1, 2020
5994-2233EN