Application Note

Genomics

Agilent

Trusted Answers

FFPE Sample Quality Control for the MGISEQ-2000 Sequencing Platform with the Agilent TapeStation System

Authors

Biao Huang,
MGI Tech Co. Ltd.,

Lisa Houston,
Agilent Technologies, Inc.

Abstract

Formalin-fixed paraffin embedding (FFPE) is one of the more common ways to preserve clinical samples, and FFPE tumor samples serve as valuable study materials for clinical and translational medicine research. To ensure successful and reliable sequencing of FFPE-derived DNA, it is necessary to conduct quality control (QC) of the initial DNA and fragmented samples at key points during next-generation sequencing (NGS) library construction. In this application note, the Agilent 4150 TapeStation system was used to conduct sample QC of the entire library construction process for five FFPE DNA samples used on the MGI Tech MGISEQ-2000RS sequencing platform. High-quality sequencing results were achieved by stepwise QC throughout the library preparation process, including the initial sample, pre- and post-PCR steps, and the final library, helping to ensure high-quality sequencing.

Introduction

Formalin-fixed paraffin embedding (FFPE) is one of the most common methods for preserving clinical samples, as it can effectively maintain the integrity of the tissue for histopathological research and morphological observation. However, the quality of DNA extracted from FFPE samples often varies, due to issues such as formaldehyde cross-linking and nucleic acid degradation. Therefore, quality assessment of FFPE samples is particularly important for high-sensitivity and high-cost applications such as high-throughput sequencing[1]. The Agilent 4150 TapeStation objectively evaluates the degradation of FFPE DNA samples using the DNA Integrity Number (DIN), and automatically calculates the size distribution and concentration of fragments[2]. The library preparation protocol can be adjusted and optimized, in a timely manner, based on quality and size data obtained from the 4150 TapeStation to ensure that the quality of the final library meets the sample loading requirements.

Experiment

Materials

An FFPE reference sample was obtained from GeneWell (p/n GWC0301). FFPE tumor sections from thyroid cancer, cervical cancer, colon cancer, and breast cancer were obtained from anonymous donors. DNA was extracted using the MGISP-NE32 automatic nucleic acid extractor (p/n 950-000020-00) and MGIEasy FFPE Genomic DNA Extraction Prepacked Kit (p/n 940-000113-00) from MGI Tech. Ultrasonic fragmentation of the DNA samples was performed using ME220 Focused-ultrasonicator from Covaris (p/n 500444). The MGIEasy Universal DNA Library Prep Kit was used for library construction. The Agilent SureSelect Human All Exon V8 kit (p/n 5191-6877) and Agilent SureSelect XT HS2 DNA Target Enrichment kit (p/n G9987A) were used as the NGS target enrichment probes and hybridization reagents. The Agilent 4150 TapeStation system was employed as the DNA quality control analysis platform, along with Agilent Genomic DNA ScreenTape assay (p/n 5067-5365) and reagents (p/n 5067-5366), the Agilent D1000 ScreenTape assay (p/n 5067-5582) and reagents (p/n 5067-5583), and Agilent High Sensitivity D1000 ScreenTape assay (p/n 5067-5584) and reagents (p/n 5067-5585)

DNA extraction

Nucleic acid was extracted from all samples using MGISP-NE32 with MGIEasy FFPE Genomic DNA Extraction Prepacked Kit in accordance with the standard protocol[3]. The integrity and concentration of the initial genomic DNA samples was determined on the TapeStation system using the Genomic DNA ScreenTape and corresponding reagents according to the protocol[6].

Library preparation

The SureSelect Human All Exon V8[4] was used in combination with MGIEasy Universal DNA Library Prep Kit for standard library construction[5]. The D1000 ScreenTape and High Sensitivity D1000 ScreenTape were used together with the supporting reagents for QC of pre-PCR and post-PCR libraries, respectively. The concentration and yield were determined using Qubit 4.0 (Thermo Fisher p/n Q33236) and Qubit dsDNA HS Assay kit (p/n Q32854). The target smear sizes were analyzed with the TapeStation, according to the standard experimental protocol[7, 8].

Whole exome sequencing

60 ng of each of the four samples and the FFPE standard were pooled post-PCR for a total of 300 ng. The pooled PCR products were heat denatured and circularized with a splint oligo sequence, forming the single strand circle DNA (ssCir DNA) as the final library. The ssCir DNA was amplified to make a DNA nanoball (DNB) with 300 to 500 copies of one molecule. The DNBs were loaded into the patterned nanoarray, and pair-end 100 base reads were generated and sequenced by combinatorial Probe-Anchor Synthesis (cPAS) on the MGISEQ-2000 platform. The sequencing data was qualified if the yield exceeded 550 M reads, and the sequencing quality (Q30) was greater than 80%.

Results and discussion

Quality of genomic DNA derived from FFPE samples

The FFPE nucleic acid extraction workflow includes deparaffinization, washing, and elution (Figure 1). The deparaffinization step is very important to DNA integrity. Too long of a deparaffinization may lead to nucleic acid degradation, while too short, or other operations that are not conducive to deparaffinization, may result in low extraction efficiency and impact DNA integrity.

The MGISP-NE32 automated FFPE genomic DNA (gDNA) extraction system can generate gDNA that is suitable for NGS. The concentration of DNA fragments was measured by Qubit. The DNA quality (DIN) and size of the main peak were analyzed by the TapeStation with the Genomic DNA ScreenTape and related reagents (Table 1 and Figure 2). Serious degradation was observed in some of the samples, which is indicated by low DIN values.

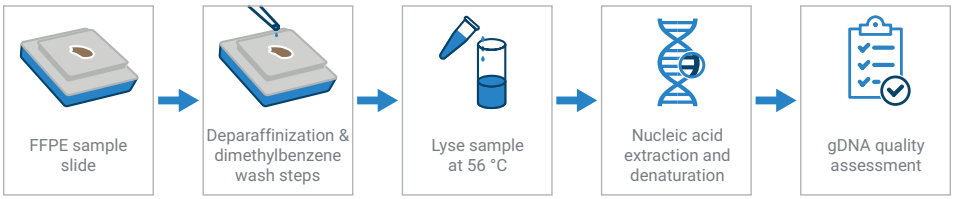

Figure 1. MGIEasy FFPE Genomic DNA Extraction kit workflow.

Table 1. FFPE-extracted DNA was analyzed for quality (DIN) and size (bp) on the Agilent 4150 TapeStation system. Concentration (ng/μL) was determined on the Qubit. The size of the main peak was determined by Agilent TapeStation software, and corresponds to the highest peak as generated from the TapeStation software.

No.	Type of Tissue	DIN	Concentration (ng/μL)	Main Peak (bp)
FFPE-1	Cervical	2.9	28.6	1,919
FFPE-2	Colorectal	4.5	84.8	2,241
FFPE-3	Thyroid	2.6	42.4	731
FFPE-4	Breast	2.2	10.2	574
Control	Standards	7.9	32.4	24,312

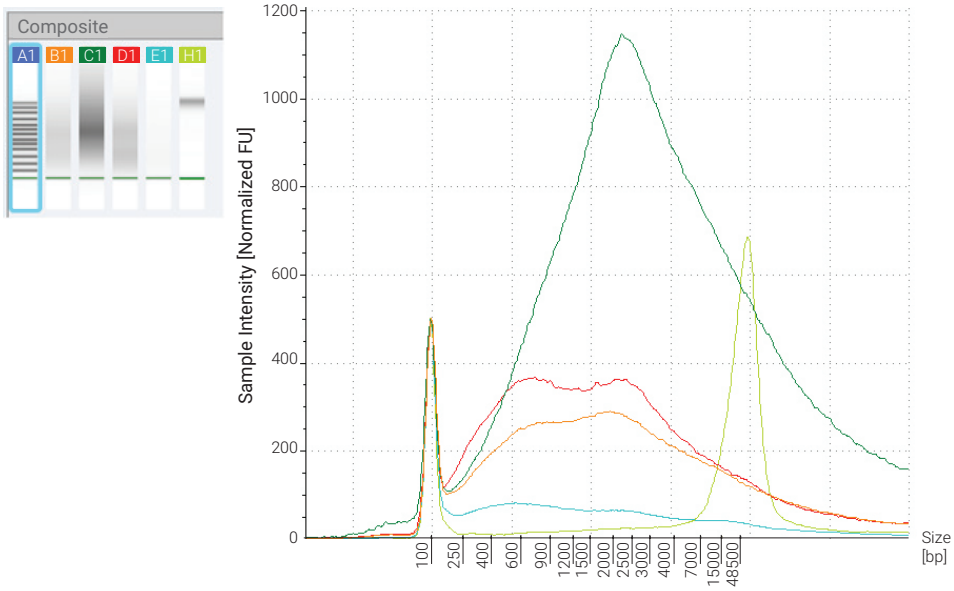

Figure 2. Analysis of purified genomic DNA from FFPE samples using the Agilent 4150 TapeStation system with the Agilent Genomic DNA ScreenTape assay and corresponding reagents. The left panel is the gel image, and the right panel is the electropherogram. A1 is the ladder provided along with the gDNA kit (100 bp-48.5 kb), B1-E1 correspond to FFPE samples 1-4, and H1 is the FFPE DNA standard.

Quality of precapture libraries

The FFPE gDNA samples were used for library preparation, starting with shearing the samples to approximately the same size using Covaris ultrasonication fragmentation. The shearing methods were previously optimized based on the size of the samples as obtained by the TapeStation. Approximately 400 ng of each sample was subjected to ultrasonic fragmentation. Fragmentation treatment time for each sample varied based on size. Genomic DNA samples larger than 13 kb were treated for 155 seconds, while samples smaller than 13 kb were treated for 125 seconds. All other shearing parameters were the same regardless of the sample size, and are as follows: peak incident power 75 W, duty factor 20%, with 1,000 cycles for burst.

Libraries were constructed from each fragmented sample using the MGIEasy Universal DNA Library Prep Kit following standard procedures. To maintain consistency, each library was subject to 12 PCR cycles. The precapture libraries were analyzed for fragment size distribution and concentration using the D1000 ScreenTape and supporting reagents (Table 2). For all samples, the size of the peaks were consistent at 200-300 bp (Figure 3).

Table 2. The amount of ultrasonic fragmented DNA used to construct the pre-PCR library, and the concentration, yield, and fragment sizes of the pre-PCR library. Concentration and yield were determined by Qubit. The size of the pre-capture library main peak was determined by the Agilent 4150 TapeStation, and corresponds to the highest peak.

No.	Pre-Library Concentration (ng/μL)	Pre-Library Yield (ng)	Pre-Library Main Peak (bp)
FFPE-1	24.4	732	281
FFPE-2	65.0	1,950	312
FFPE-3	13.4	402	273
FFPE-4	4.3	130	260
Control	75.2	2,256	324

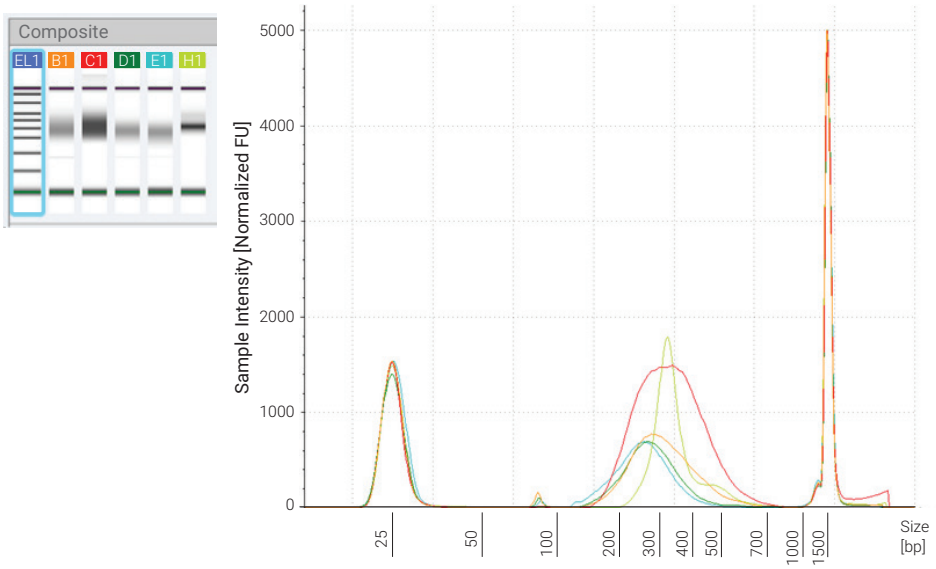

Figure 3. Analysis of the precapture libraries obtained with the Agilent 4150 TapeStation system, Agilent D1000 ScreenTape assay, and related reagents. The left panel is the gel image, and the right panel is the electropherogram. The main peak of the library corresponds to the highest point, and is about 300 bp for all samples, with the fragments mainly concentrated within 200-500 bp.

Quality of final library

The precapture samples were pooled to create the final library for hybridization capture and sequencing. The precapture samples were pooled at a combination of 300 ng: 300 ng: 300 ng: 100 ng (samples 1 to 4 respectively). The FFPE-4 precapture sample had a lower concentration than the others, necessitating a lower amount used for pooling in order to create a total of 1,000 ng for hybridization capture. A separate library was constructed from 1,000 ng of the standard. Twelve cycles of PCR amplification were performed on the pooled samples and standard libraries to produce the final library. Hybridization capture was performed using the SureSelect Human All Exon V8 and the SureSelect XT HS2 DNA kit. The QC of the final library was conducted using the TapeStation System with HS D1000 ScreenTape and corresponding reagents. The final library sizes all ranged from 200 – 500 bp, with total amount larger than 300 ng, which meets the requirement for sequencing on the MGI platform (Table 3 and Figure 4).

Quality of the sequencing results

The final library was sequenced on the MGISEQ-2000RS system using the FCS PE100 Rapid Sequencing Kit. Each sample data were normalized to 7.5 Gb for analysis. All data met the acceptance criteria, with the sequencing yield exceeding the standard of 550 M reads on MGISEQ-2000, and the sequencing quality (Q30) was greater than 90% before and after filtration (Figure 5). The results demonstrate that the quality of sample extraction and the uniformity and complexity of the sequencing library created from target enrichment are optimal.

Table 3. Amount of precapture library used for hybridization capture, post enrichment amplification cycles, and concentration and yield of the final library. DNA quantity and concentration were determined by Qubit. The size (bp) of the final library was determined by the Agilent 4150 TapeStation system, Agilent High Sensitivity D1000 ScreenTape assay, and supporting reagents.

No.	DNA Quantity (ng)	Post-PCR Amplification Cycle	Final Library Concentration (ng/µL)	Final Library Yield (ng)	Final Library Main Peak (bp)
4-in1 Pooled FFPE Sample Control	1000	12	19.3	482	276
Control	1000	12	17.8	445	304

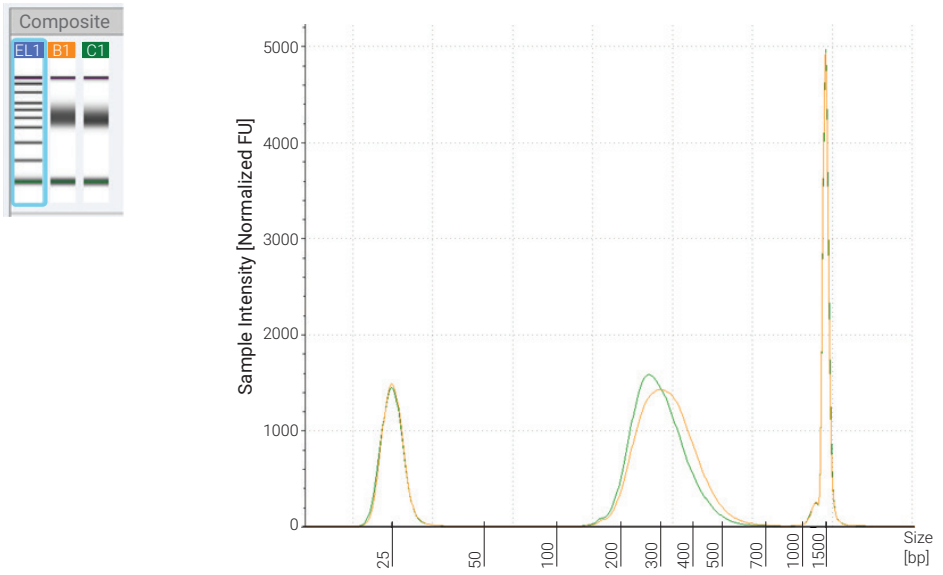

Figure 4. Quality of the final library analyzed using the Agilent 4150 TapeStation system with the Agilent High Sensitivity D1000 ScreenTape and supporting reagents. B1 is the standard, and C1 is the pooled FFPE sample.

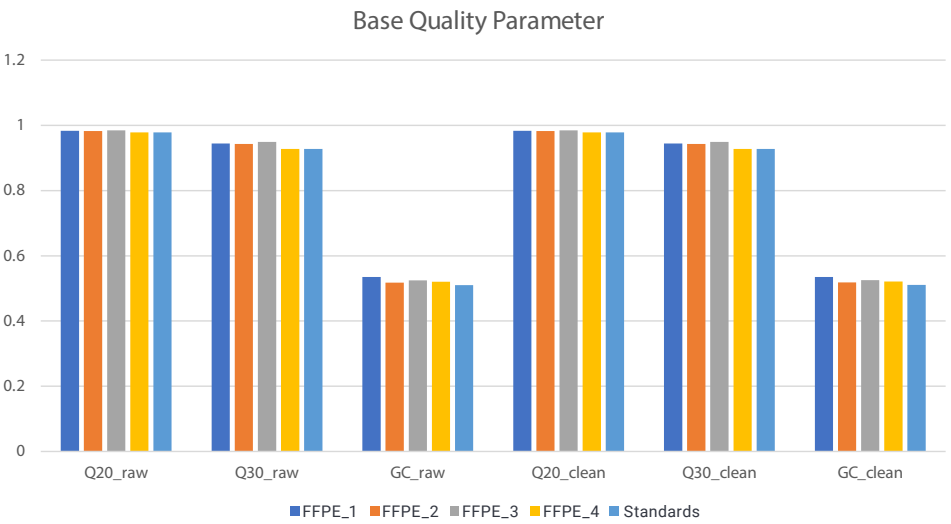

Figure 5. Sequencing quality parameters for four FFPE samples and standards.

Conclusion

In the sequencing sample preparation workflow based on the MGI sequencing platform, the Agilent 4150 TapeStation system can be used for QC at a variety of steps. First, the TapeStation can be used to analyze the size distribution of genomic DNA derived from FFPE samples. These results can be used to determine sample fragmentation parameters. Second, the TapeStation can be used after fragmentation, to verify the size distribution of the sheared samples. Third, in the library construction process, the TapeStation system can evaluate the size distribution of the precapture libraries. Fourth, the final target-enriched libraries can be analyzed on the TapeStation to ensure that the size of the library meets the requirements for the sequencing platform.

In this application note, the final size of the libraries and sequencing results from the FFPE DNA extracted by the MGIEasy FFPE Genomic DNA Extraction prepackaged kit met the gDNA requirements for NGS sequencing. A high-quality whole-exome sequencing library can be obtained by combining the Agilent SureSelect Human All Exon V8 probe and the Agilent SureSelect XT HS2 DNA hybridization capture reagents with the MGIEasy Universal DNA Library Prep Kit, as demonstrated by the high-quality TapeStation sequencing results. Quality control with the TapeStation aids in achieving successful sequencing with the MGISEQ system.

References

1. Tang, W.; David, F. B.; Wilson, M. M.; Barwick, B. G.; Leyland-Jones, B. R.; Bouzyk, M. M. DNA Extraction from Formalin-Fixed, Paraffin-Embedded Tissue. *Cold Spring Harbor Protocols*, **2009**.
2. High Throughput Genomic DNA Assessment by the Agilent 4200 TapeStation System. *Agilent Technologies technical overview*, publication number 5991-6629EN, **2016**.
3. MGIEasy FFPE Genomic DNA Extraction Prepacked Kit (NE32) (V1.0) Doc. #H-940-000071-00-00.
4. High Performance Exome Built on Advanced and Proven Technology. *Agilent Technologies data sheet*, publication number 5994-3154EN, **2021**.
5. MGIEasy Universal DNA Library Prep Set Instruction Manual (version A4) Doc. #: B02-034.
6. Genomic DNA ScreenTape Assay for TapeStation Systems. *Agilent Technologies quick guide*, publication number SD-UF0000092 Rev. B., **2018-2021**.
7. D1000 ScreenTape Assay for TapeStation Systems. *Agilent Technologies quick guide*, publication number SD-UF0000094 Rev. C., **2018-2021**.
8. High Sensitivity D1000 ScreenTape Assay for TapeStation Systems. *Agilent Technologies quick guide*, publication number SD-UF0000091 Rev. C., **2018-2021**.
9. High-throughput (Rapid) Sequencing Set User Manual (V3.0) Doc. DNBSEQ-G400RS.

www.agilent.com/genomics/automated-electrophoresis

PR7000-8670

Published in the USA, June 1, 2022
5994-4787EN

Agilent
Trusted Answers