Application Note

Nucleic Acid Analysis

Agilent

Trusted Answers

The DNA Integrity Number (DIN) Provided by the Agilent Genomic DNA ScreenTape Assay Allows for Streamlining of NGS on FFPE Tissue

Authors

Hyunju Jung
Sumin Ji
Sanghoon Song
Yeji Park
Theragen Etex Co., Ltd., Suwon, Korea (info@totalomics.com)

Ji-Won Yang
Agilent Technologies, Inc. Korea

Eva Schmidt
Agilent Technologies, Inc.
Waldbronn, Germany

Abstract

Sequencing of genomic DNA (gDNA) from formalin-fixed paraffin-embedded (FFPE) archived tissue can be challenging, as the obtained material is often of variable quality. This study demonstrates that the DNA Integrity Number (DIN) obtained by the quality control (QC) of gDNA using the Agilent Genomic DNA ScreenTape assay has allowed for a pronounced saving of sequencing and sample preparation overhead. Out of a total of 751 FFPE samples, a subset of 197 were tested for a correlation of various next-generation sequencing parameters against the DIN. A correlation was identified between DIN and the key parameters of on-target rate and coverage at 10x. A QC threshold of ≥ 3 DIN was therefore set, which consequently excluded 65% (n = 488) of the total sample set and saved a significant amount of time and effort.

Introduction

Formalin-fixed paraffin-embedded (FFPE) tissue is widely used for clinical sample preservation and archiving. These archives provide a valuable resource for genomics studies, and are beginning to be successfully exploited by next-generation sequencing (NGS) applications. DNA extraction from FFPE samples has proven to be challenging, and often results in low amounts of suitable DNA for NGS applications.

Before subjecting genomic DNA (gDNA) extracted from FFPE tissue to the NGS workflow, it is highly recommended to perform a quality control (QC) step to determine the integrity and concentration of the extracted gDNA. The preparation of NGS libraries includes additional QC steps to assess the fragment length distribution of the processed sample (Figure 1).

The Agilent 2200 TapeStation system and the DNA ScreenTape assays can be used for sample QC at several steps of the NGS workflow, such as shearing of gDNA and library amplification. This application note focuses on the QC of gDNA extracted from FFPE tissue with the 2200 TapeStation system and the Genomic DNA ScreenTape assay. For gDNA samples, the Agilent 2200 TapeStation analysis software (revision A01.05) automatically calculates the DNA concentration and provides the DNA Integrity Number (DIN)[1]. This numerical assessment of the gDNA integrity ranges from 1 to 10, and is displayed directly under the gel image as well as in the sample table. A high DIN indicates highly intact gDNA, and a low DIN a strongly degraded gDNA.

Experimental

Materials

The S2 Focused-ultrasonicator was obtained from Covaris Inc. (Woburn, MA, USA). The Agilent 2200 TapeStation system (p/n G2965AA) with the Agilent TapeStation analysis software (revision A01.05), Agilent Genomic DNA ScreenTape (p/n 5067-5365) and Agilent Genomic DNA Reagents (p/n 5067-5366), and Agilent SureSelect XT reagent kit - HSQ (p/n G9611A) were obtained from Agilent Technologies Inc. (Santa Clara, CA, USA). The HiSeq 2500 system was purchased from Illumina, Inc. (San Diego, CA, USA).

Samples

gDNA was extracted from FFPE blocks of cancer tissue, then subjected to the SureSelect XT workflow.

Library preparation and sequencing

The gDNA was fragmented using ultrasonication (10% duty cycle, intensity 5, six cycles of 60-second treatment, at 4 to 7 °C). The SureSelectXT protocol provided by the manufacturer was followed for library preparation. In brief, the extracted gDNA was sheared to produce small fragments, followed by the library preparation using sequencer-specific adaptors and indexes. The samples were then hybridized with biotinylated RNA library baits. Targeted regions were isolated using magnetic streptavidin beads. In a final step, the samples were amplified before sequencing. Sequencing was performed with the HiSeq 2500 system according to the manufacturer's instructions.

QC analysis with the Agilent 2200 TapeStation system

A gDNA QC was performed using the 2200 TapeStation system in combination with the Genomic DNA ScreenTape assay, including the automatic determination of DIN, according to the manufacturer's instructions[2].

Figure 1 provides an overview of the overall SureSelect XT sequencing sample preparation workflow, including the sample QC steps.

Results and Discussion

As an initial QC step, gDNA samples extracted from FFPE cancer tissue were analyzed with the Genomic DNA ScreenTape assay. The 2200 TapeStation analysis software (revision A.01.05) automatically displays a DIN as a measure of gDNA integrity. In total, 751 samples were analyzed, and the DIN ranged from 1 to 7.8.

Following the workflow in Figure 1, gDNA samples were sheared, further processed, and finally subjected to sequencing using the HiSeq 2500 system. An on-target rate above 70% was defined as the sequencing success parameter. In addition, the deduplication rate (target above 80%), and coverage at 10x (target above 90%) were analyzed as additional sequencing quality criteria.

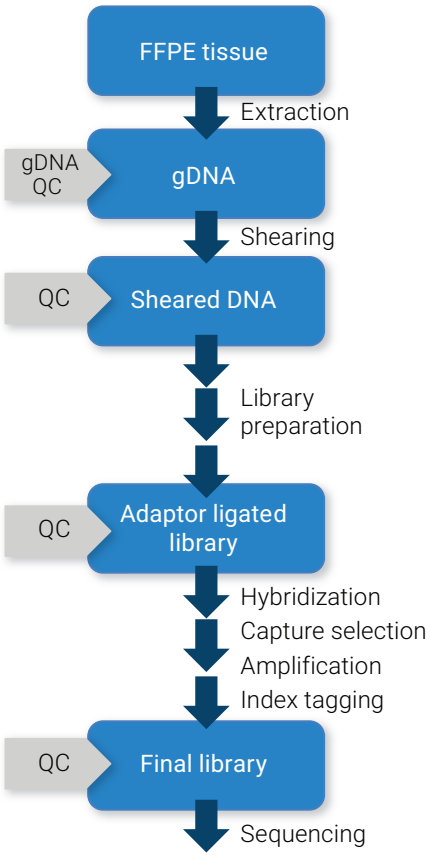

Figure 1. Overview of the overall Agilent SureSelect XT sequencing sample preparation workflow, indicating the recommended sample

As part of the NGS library preparation workflow, input gDNA is fragmented (Figure 1). Therefore, it could be anticipated that the original size distribution of a sample has no significant effect on the quality of the sequencing results. To test this, 197 samples of the large sample pool were randomly selected and further analyzed to determine if there was a correlation between the initial gDNA integrity, calculated as DIN, and the sequencing quality criteria. This subset of samples had a DIN range of 1.5 to 6.4.

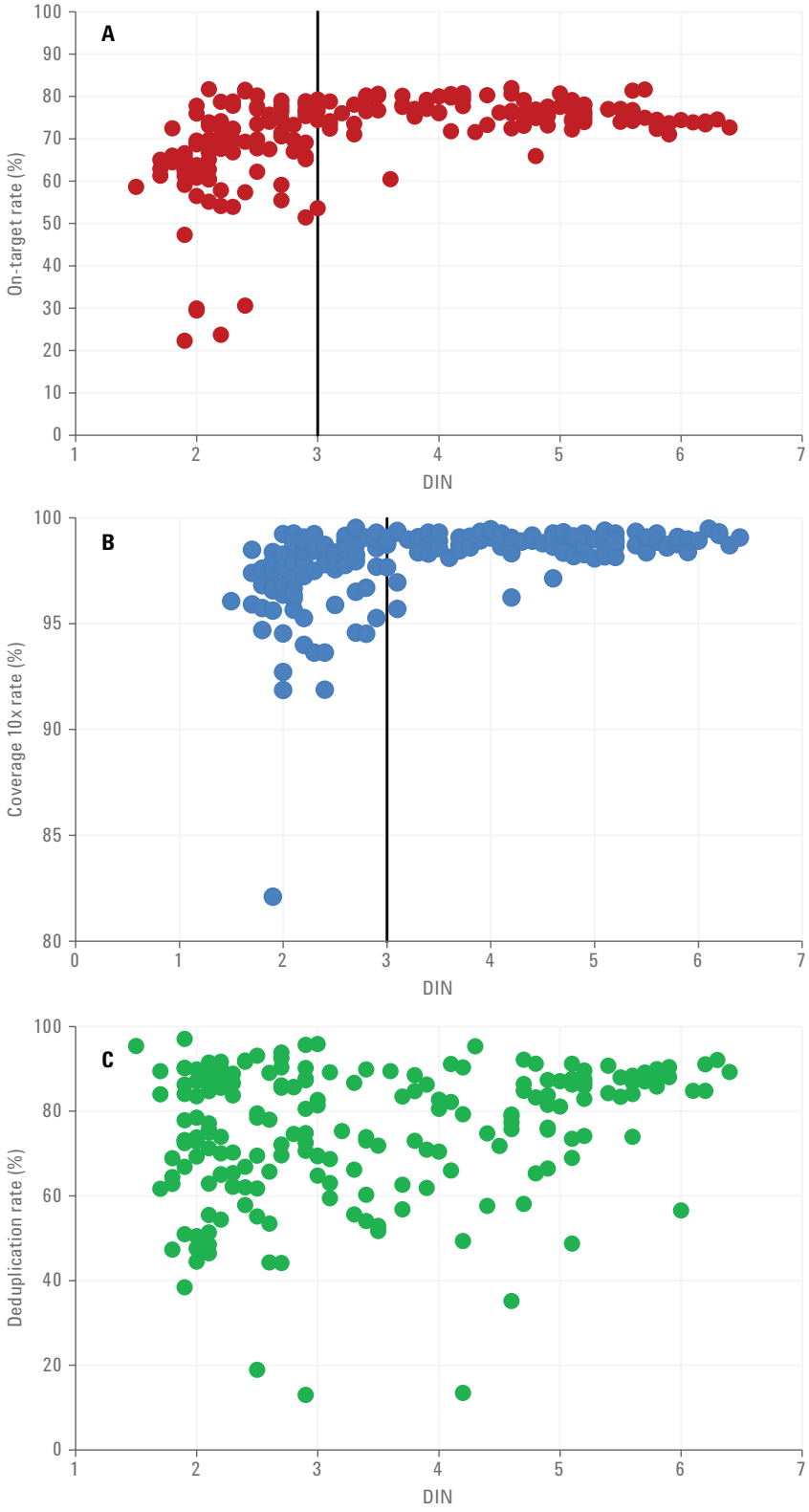

Figure 2. The DIN determined with the Agilent 2200 TapeStation system and the Agilent Genomic DNA ScreenTape assay plotted against on-target rate (A), the coverage rate at 10x (B), and the deduplication rate (C) for 197 samples. There is a greater variation of both on-target rate and coverage rate at 10x when the DIN is below 3. There is no noticeable correlation between the deduplication rate and the DIN.

A total of 63 samples (31%) of the analyzed sample subset did not meet the on-target rate criteria of at least 70%. Among these, the integrity analysis using DIN showed that 61 samples (97%) received a DIN below 3 (Figure 2A). We further analyzed the %coverage rate at 10x for possible correlation to the integrity assessment. The majority of samples did pass the 90% target coverage rate at 10x, but greater variation was observed for samples with the DIN below 3 (Figure 2B). In contrast, Figure 2C shows that there is no clear correlation between the DIN and the deduplication rate.

Based on subjective interpretation of the electropherograms or gel images, it would have been difficult to reject samples, especially when handling results in a high-throughput format. The determined DIN threshold allows objective exclusion of samples from further processing and sequencing, and ensures correct repeatability of the DNA integrity assessment. In addition to the DIN, the amount of gDNA extracted from the FFPE tissue is critical for the success of the downstream sequencing. The 2200 TapeStation system determines the DIN and the total sample concentration in a single step.

Figure 3 shows the respective electropherogram traces and gel images analyzed with the Genomic DNA ScreenTape assay and the 2200 TapeStation system of selected samples from the correlation study sample subset.

Table 1 summarizes detailed results for the gDNA samples shown in Figure 3. Three samples with DIN > 3 showed > 70% on-target rate, and the three samples DIN < 3 showed a diminished on-target rate. Based on the analysis of the initial sample integrity and the correlation to the on-target and coverage 10x rate, it was decided to set a DIN threshold of > 3 for subsequent sequencing. Among the correlation data set (n = 197), 49% did not pass the DIN > 3 criteria. Among the total sample set (n = 751), 65% did not pass the DIN > 3 criteria. Samples with a DIN below 3 were not further processed and sequenced. The described NGS sample preparation workflow, including the several QC steps is relatively work intensive, and the cost per sample is relatively high, therefore, it is very advantageous to be able to qualify samples for acceptance and rejection criteria using DIN.

Table 1. The DIN and the sequencing quality criteria obtained for the six samples shown in Figure 3.

Sample	DIN	On-target rate % > 70 %	10x Coverage rate % > 90 %	Deduplication rate %
1	4.1	79.7	98.6	66.0
2	5.2	78.0	99.0	74.2
3	5.0	80.7	98.1	87.1
4	2.2	23.8	94.0	65.0
5	2.9	51.5	97.7	95.7
6	1.9	47.4	96.6	97.1

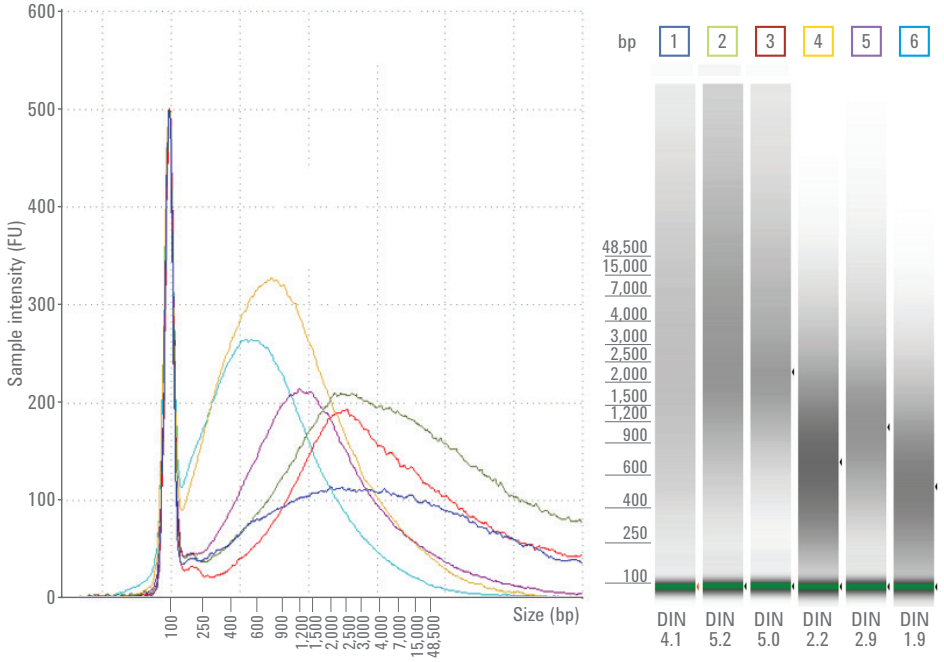

Figure 3. QC for gDNA extracted from FFPE tissue was performed using the Agilent 2200 TapeStation system and the Agilent Genomic DNA ScreenTape assay. The overlay of the electropherograms of different samples and the gel image with the determined DIN are shown. Representative samples of sufficient DIN > 3 (samples 1–3) and insufficient DIN < 3 (samples 4–6) were selected for this figure.

Conclusion

The automatically determined DIN provided by the Agilent 2200 Tape Station system for the Agilent Genomic DNA ScreenTape assay provides an optimal tool to screen gDNA samples extracted from FFPE tissue in a high-throughput format. The DIN correlates with key sequencing quality metrics and, thus, presents integrity thresholds for the processing of FFPE samples. The DIN can be integrated as a selection criteria on whether or not to proceed individual samples for the downstream workflow. Thus, using DIN, specific sample integrity requirements can be established for respective NGS applications and easily communicated to customers or collaborators.

References

1. DNA Integrity Number (DIN) with the Agilent 2200 TapeStation System & Genomic DNA ScreenTape, *Agilent Technologies*, publication number G5991-5258EN, **2014**.
2. Agilent Genomic DNA ScreenTape System Quick Guide, *Agilent Technologies*, publication number G2964-90040 rev.C, **2014**.

www.agilent.com/genomics/tapestation

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

PR7001-1330

This information is subject to change without notice.

Published in the USA, September 1, 2023
5991-5360EN

Agilent
Trusted Answers