Application Note

Genomics

Agilent

Trusted Answers

Quality Control of Cell-free DNA Samples Analyzed with Next-Generation Sequencing

Authors

Christina Geörg and Katrin Pfütze,
Sample Processing Lab,
German Cancer Research Center DKFZ,
Heidelberg, Germany

Ute Ernst, Nicolle Diessl, and
Marie Beckhaus
Genomics and Proteomics Core Facility,
German Cancer Research Center DKFZ,
Heidelberg, Germany

Elisa Viering and Chava Pocernich,
Agilent Technologies, Inc.

Abstract

Next-generation sequencing (NGS) is an essential tool for analysis of biosamples in clinical research. Cell-free DNA (cfDNA) has become an important input material for NGS as a result of the noninvasive collection methods from blood and urine. However, analysis of cfDNA can introduce challenges due to its low concentration and possible contamination from high-molecular weight (HMW) DNA, both necessitating reliable quality control. The Agilent 4200 TapeStation system is a vital quality control (QC) tool in NGS workflows. The TapeStation systems and Cell-free DNA ScreenTape assay provide a %cfDNA quality metric for determining the quality of input cfDNA for downstream processes. The Sample Processing Lab (SPL) collaborated with the Genomics and Proteomics Core Facility (GPCF), both part of the German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ) to track 13 cfDNA samples with the TapeStation system from initial QC with the %cfDNA metric, through the NGS workflow, to the final sequencing results. All 13 samples reported a %cfDNA greater than 83% with successful library preparation and sequencing metrics.

Introduction

Liquid biopsy is a noninvasive way to collect samples from biofluids such as plasma, serum, urine, and cerebrospinal fluid for serial assessments. The cfDNA found in these biofluids is often DNA released from apoptotic cells. The DNA is fragmented between the histone complexes into various sized cfDNA. The different cfDNA fragments are referred to as the mono-nucleosome, di-nucleosome, and tri-nucleosome peaks. The majority of cfDNA is in the mono-nucleosome fragment, around 170 bp in length. cfDNA is often used for screening, following disease progression, and monitoring effectiveness of disease treatment. Sequencing cfDNA can provide information on tumor-derived mutations since a small fraction of the cfDNA is circulating tumor DNA (ctDNA), derived from tumor cells. The genetic information from ctDNA contains biomarkers that can be used to monitor cancer progression.

The SPL provides expertise in the isolation of DNA, RNA, and proteins and quality control analysis of these extractions from different sources of material such as tissue biopsies, FFPE tissue blocks, blood, and cells. The SPL works closely with the GPCF, transferring suitable samples for sequencing. Currently, the SPL is developing and optimizing extraction procedures of circulating cell-free tumor DNA from liquid biopsies, including automated extraction processes.

It is essential to know the quality of incoming samples for the experimental success of library preparation and sequencing. HMW DNA in the cfDNA sample may interfere with library yield and sequencing quality. After extraction of cfDNA, the SPL performed cfDNA sample quality control with the 4200 TapeStation system and the Cell-free DNA ScreenTape. The Cell-free DNA ScreenTape assay was developed specifically for cfDNA analysis. They investigated the %cfDNA quality metric[1], which was introduced with the Cell-free DNA ScreenTape assay as a way to screen the quality of cfDNA samples for NGS library preparation and sequencing.

Experimental

The SPL received blood collected in Fusion2 cfDNA BCT (Hiss Diagnostics p/n 230256) and PAXgene Blood ccfDNA (PreAnalytiX, a Qiagen/BD company, p/n 768115) collection tubes. cfDNA was extracted from 8 to 9 mL of plasma with the QIAmp MinElute ccfDNA Midi kit (Qiagen, p/n 55284) and eluted in 50 µL of water. Quality control analysis was performed with the Agilent 4200 TapeStation system and the Agilent Cell-free DNA ScreenTape (p/n 5067-5630) and Cell-free DNA reagents (p/n 5067-5631). Concentration was determined with both the Qubit and the dsDNA HS assay (p/n Q32851) and the 4200 TapeStation and Cell-free DNA assay. Libraries were prepared with the Agilent SureSelectXT Low Input Target Enrichment system (p/n 5191-4080) with 25 or 50 ng of input cfDNA. Final libraries went through QC analysis with the 4200 TapeStation and the HS D1000 assay (ScreenTapes p/n 5067-5584, reagents p/n 5067-5585). Sequencing was done with the Illumina HiSeq 4000 system and a paired-end 100 bp run. Sequencing data QC thresholds are mostly based on Illumina specifications.

Results and discussion

Quality control of extracted cfDNA

Sizing

Thirteen cfDNA samples were screened for QC on the 4200 TapeStation system with the Cell-free DNA ScreenTape assay, which was developed specifically for cfDNA analysis. The electrophoresis separations with the Cell-free DNA ScreenTape assay all displayed the mono- and di-nucleosome peaks (Figure 1A) with the majority of samples showing the tri-nucleosome peak (Figure 1B). It is common with cfDNA samples to observe sample to sample variation regarding the number of nucleosome fragments present. The TapeStation analysis software automatically reported a size for the mono-nucleosome peak. Sizing of the mono-nucleosome peak across all the samples was similar, ranging from 170 to 195 bp (Figure 2).

Quantification

Total cfDNA concentration was measured on the 4200 TapeStation system with the Cell-free DNA ScreenTape assay and on the Qubit with the dsDNA HS assay. Qubit has two assays available for measuring low-concentration DNA samples: the dsDNA HS assay (p/n Q32851) and the newer 1X dsDNA HS assay (p/n Q33230). Previously, we showed that total cfDNA concentrations measured with the Cell-free DNA ScreenTape assay are highly comparable with the Qubit 1X dsDNA HS assay, ddPCR, and qPCR, while the Qubit dsDNA HS assay consistently reported higher concentrations[2].

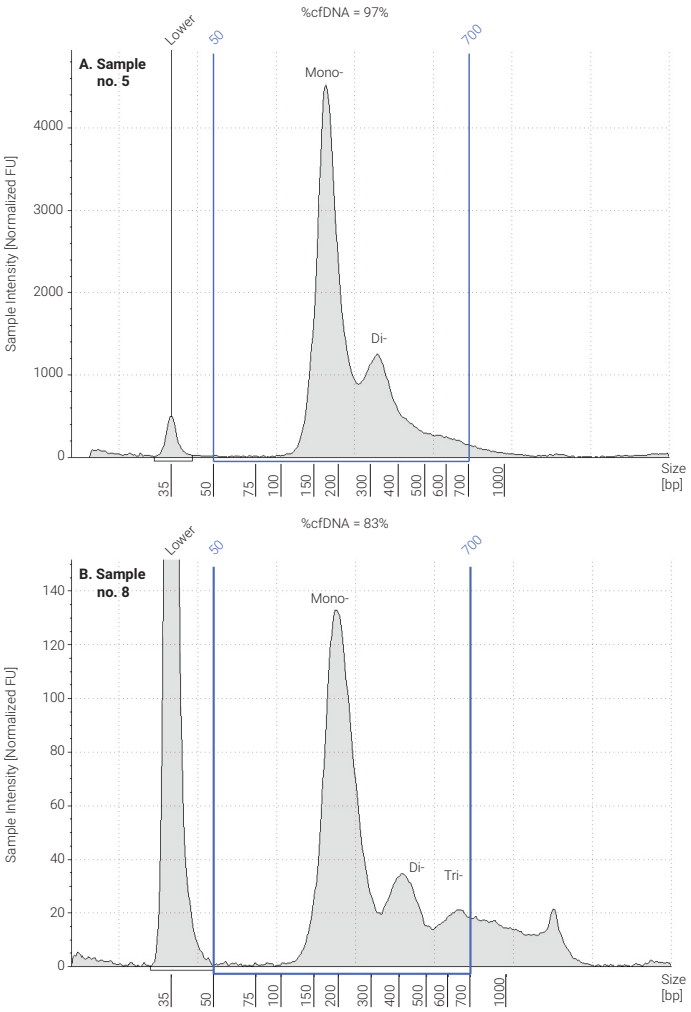

Figure 1. cfDNA samples analyzed with the Agilent 4200 TapeStation system and the Cell-free DNA ScreenTape assay. (A) Sample no. 5 displayed the mono- and di-nucleosome peak. (B) Sample no. 8 displayed the mono-, di-, and tri-nucleosome peaks.

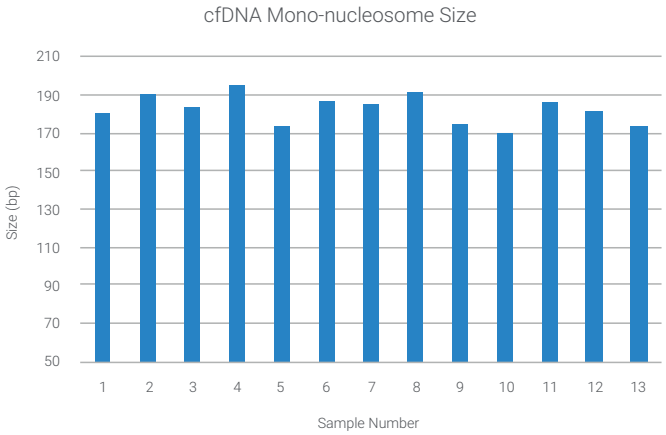

Figure 2. Mono-nucleosome fragment size from cfDNA samples analyzed with the Agilent 4200 TapeStation system and the Cell-free DNA ScreenTape assay.

In this current study, quantification of the total cfDNA samples was consistently higher with the Qubit dsDNA HS assay compared to the Cell-free DNA ScreenTape assay (Figure 3), consistent with the previous studies.

The TapeStation system has the distinct advantage of separation analysis, allowing for sizing and quantification of the different components that make up a sample. Region analysis set by the user provides detailed information about each component of interest in the total sample. Concentration of the cfDNA portion of the sample was reported by region analysis, set at 50 to 700 bp, and compared to the total sample (Figure 4). There is no way for the Qubit to discern the concentration of the different components of a sample since it does not perform any type of separation and can only measure the total concentration of a sample.

Percent cfDNA

The %cfDNA metric was developed for the TapeStation instrument as an easy-to-use screening tool for determining the quality of cfDNA samples for NGS library preparation and sequencing. The %cfDNA metric reports the percent of cfDNA fragments from 50 to 700 bp in the total sample. This range encompasses the mono-, di-, and tri-nucleosome cfDNA fragments. The metric ranges from 0 to 100%. HMW DNA can be extracted with cfDNA and interfere with downstream processes. A high percent cfDNA indicates a high-quality sample with very little HMW DNA. Likewise, a low %cfDNA indicates the presence of HMW DNA. All the samples from the SPL reported a %cfDNA between 83 to 97% cfDNA, indicating minimal HMW DNA present (Figure 5).

Library preparation

All samples continued through library preparation. Due to the low concentration available from cfDNA extraction, samples no. 3, 8, and 9 had a lower input of approximately 25 ng cfDNA for library preparation compared to 50 ng for the majority of samples. A threshold of 10 ng sample input had previously been established for the SureSelectXT Low Input Target Enrichment system[3] by the DKFZ. The number of PCR cycles was adapted to the input amount and fragmentation was excluded due to the small fragment size of cfDNA. DKFZ did not incorporate size selection to eliminate HMW DNA, since the %cfDNA was high for all samples, indicating minimal HMW DNA present.

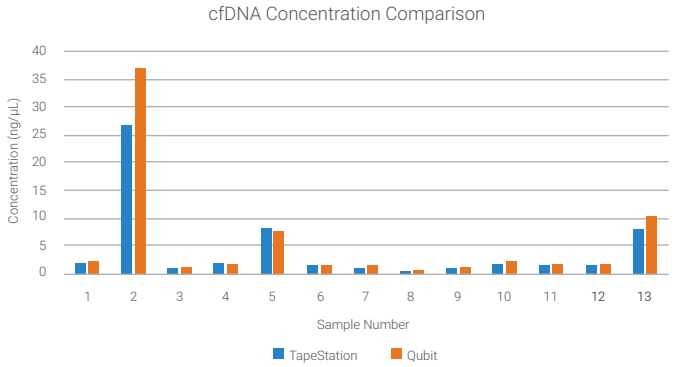

Figure 3. Total cfDNA sample concentration was measured on the Agilent 4200 TapeStation system with the Cell-free DNA Screen Tape assay and on the Qubit with the dsDNA HS assay.

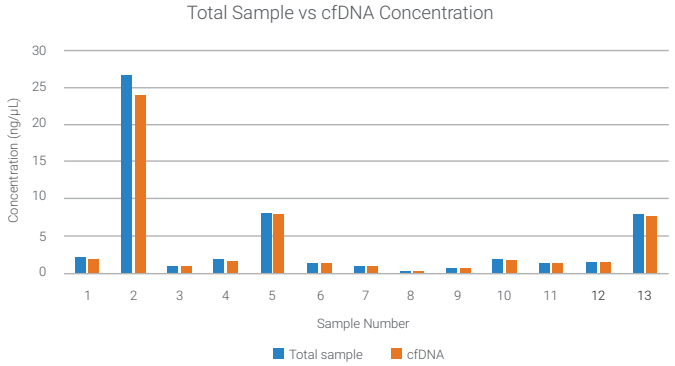

Figure 4. The Agilent 4200 TapeStation system with the Cell-free DNA Screen Tape assay measured both the total sample concentration and the cfDNA of the sample by region analysis set at 50 to 700 bp, allowing for comparison to the total sample concentration.

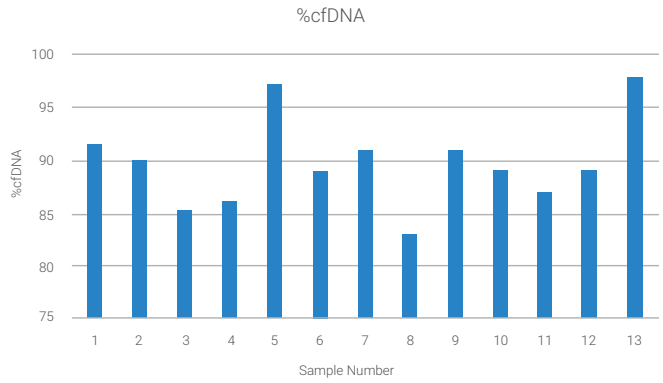

Figure 5. %cfDNA metric for samples analyzed with the Agilent 4200 TapeStation system and the Cell-free DNA Screen Tape assay. All samples reported a %cfDNA above 83%.

QC was performed again after library preparation with the TapeStation and the HS D1000 assay to ensure quality libraries. All libraries displayed a shape similar to the cfDNA trace, with most showing three nucleosome peaks or a smear in the region of the tri-nucleosome peak (Figure 6). For example, cfDNA sample no. 5 (Figure 1A) displayed two peaks and only a small smear in the tri-nucleosome region, which was mimicked in the library trace. Library QC required the smear size to be between 200 to 400 bp, a concentration of 2 ng/µL, and with less than 0.5% adapter-dimer present for patterned flow cell.

The library size was assessed by region analysis, taking into consideration the entire library sample. All samples except no. 8 fell within the library size QC threshold of 200 to 400 bp (Figure 7). Sample no. 8 was slightly outside the range, at 436 bp. It also reported the lowest %cfDNA, at 83%, indicating the presence of a small amount of HMW DNA. Samples with slightly higher %cfDNA at 85% fell within the library size QC threshold. All samples were below the required 0.5% adapter-dimer (data not shown).

The library concentration was compared on the Qubit and TapeStation. All but one sample measured below the 2 ng/µL library concentration threshold on the TapeStation. Sample no. 11 reported exactly at 2 ng/µL. The varied amount of input cfDNA did not affect the final library concentration, as all libraries reported a similar concentration.

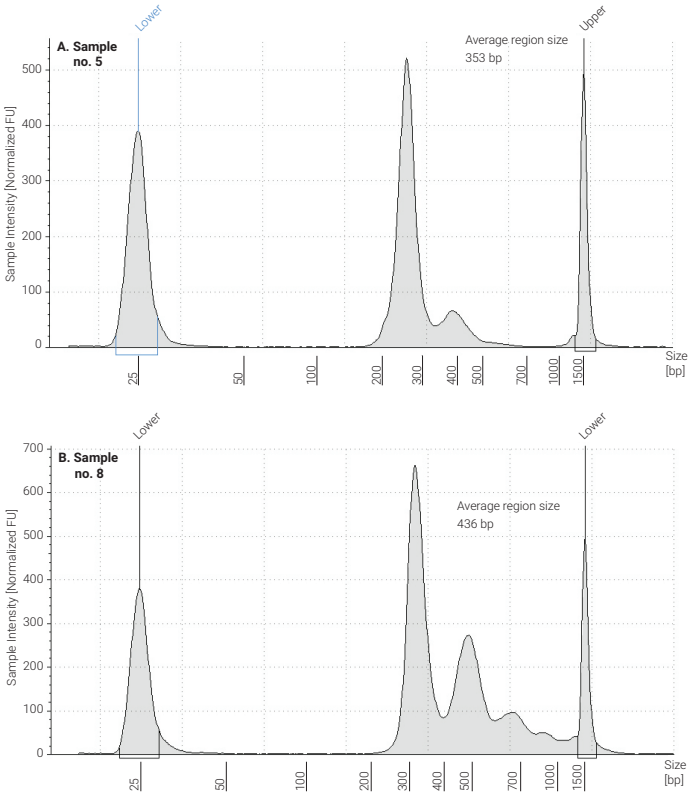

Figure 6. Library samples analyzed with the Agilent 4200 TapeStation system and the HS D1000 assay. (A) sample no. 5 (B) sample no. 8. The three peaks represent the mono-, di-, and tri-nucleosome cfDNA peaks with adapters attached.

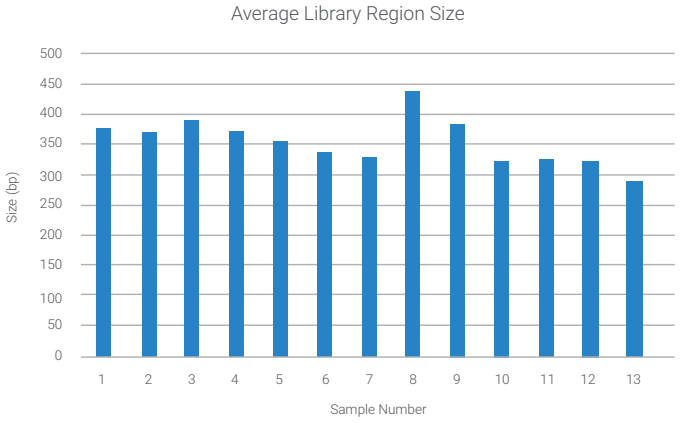

Figure 7. Average region size of the libraries was assessed with the Agilent 4200 TapeStation system and the HS D1000 assay with the region analysis setting.

In addition, all samples reported comparable concentrations between the Qubit and TapeStation, except sample no. 3, which reported a much higher concentration on the Qubit dsDNA HS assay (Figure 8). Excluding sample no. 3, the percent difference from the Qubit ranged from 0.35 to 15%, averaging 6.7% difference, demonstrating comparable results between the two instruments. The percent difference was well within the range of the HS D1000 assay and all samples continued through to sequencing.

Sequencing QC metrics

Table 1 summarizes and compares the sequencing QC metrics of all 13 samples. Sequencing QC thresholds for cfDNA have not yet been defined due to the lack of statistical information, but the threshold for OnTarget rate > 50% has been set for the SureSelect XT HS Human ALL Exon v7 protocol. General thresholds for successful sequencing, including mapping rate > 99%, duplicates </= 20%, and insert size > 140 bp, were followed. Only three samples reported one of the sequencing metrics slightly outside the thresholds listed. The mapping rate for all 13 samples was 99.94% or higher. Only sample no. 3 had an OnTarget rate below 50%, reporting at 44%. Samples no. 8 and 9 reported a 21% duplicate rate slightly outside the threshold, possibly due to PCR or patterned flow cell. Successful sequencing was obtained for all 13 samples based on the reported sequencing metrics. These results demonstrated that a sample with greater than 83% cfDNA can lead to successful library preparation and sequencing. Samples with a %cfDNA lower than 83% were not a part of this study and have not been evaluated to determine if they are viable samples for library preparation and meaningful sequencing results. Additional samples with a wide range of %cfDNA need to be evaluated in order to provide a more robust understanding of how the %cfDNA metric relates to sequencing success. The DKFZ is continuing to fine tune protocols in order to set specifications for library preparation of cfDNA samples that would lead to successful sequencing.

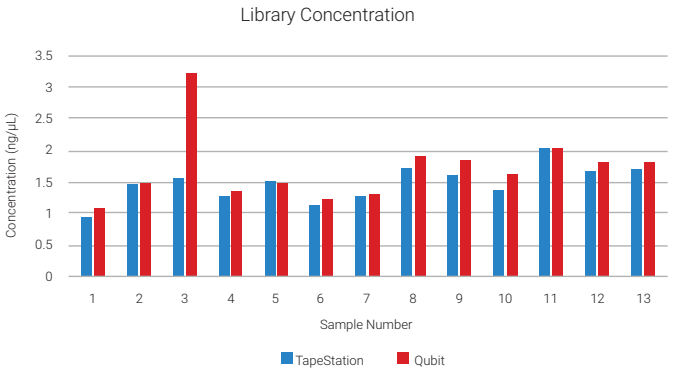

Figure 8. Library concentration was analyzed with the Agilent 4200 TapeStation system and the HS D1000 assay. The 4200 TapeStation system and Qubit concentration were very similar for all but sample no. 3.

Table 1. Sequencing quality control metrics and %cfDNA reported for all 13 samples.

Sequencing QC Metrics						
Sample Number	Mapped [%]	Duplicates [%]	On Target [%]	R2 %>=Q30 [%]	Insert size [bp]	% cfDNA
1	99.97	16.4	67.2	91.4	166	92
2	99.98	13.3	69.4	91.4	166	90
3	99.96	15.4	44.7	88.8	166	85
4	99.97	15.0	60.5	90.5	166	86
5	99.97	15.0	60.5	90.5	166	97
6	99.96	12.0	68.2	91.4	166	89
7	99.96	11.8	68.8	91.4	167	91
8	99.94	21.3	55.7	88.8	167	83
9	99.95	21.2	51.4	88.8	166	91
10	99.96	11.8	64.5	91.4	166	89
11	99.97	9.8	68.7	91.4	167	89
12	99.97	11.7	67.8	91.4	166	98

Conclusion

The TapeStation systems offer a Cell-free DNA ScreenTape specifically designed for cfDNA samples. This assay overcomes the complexity of cfDNA sample. The sensitivity of the assay allows detection of picogram levels of cfDNA with visualization of the multiple cfDNA monomers. In addition, the Cell-free DNA ScreenTape assay quantifies the cfDNA subcomponent of the sample separately from the total sample. Distinct to the Cell-free DNA ScreenTape assay is the %cfDNA metric, which aids users in determining the amount of HMW DNA in their samples and provides a screening tool for library preparation. The SPL tracked cfDNA samples from the initial QC with the %cfDNA metric, through library preparation, to the end sequencing results. All samples had a %cfDNA over 83%, with library preparation passing the size threshold of 200 to 400 bp, but reporting a concentration below 2 ng/µl, and resulting in successful sequencing.

References

1. Performance Characteristics of the Cell-free DNA ScreenTape Assay, *Agilent Technologies Technical Overview*, publication number 5994-1390EN, **2019**.
2. Cell-free DNA Quality and Quantity Assessment – A Method Comparison, *Agilent Technologies Poster*, publication number PR7000-7364, **2019**.
3. SureSelect XT Low Input Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library for Illumina Multiplexed Sequencing Platforms Protocol, *Agilent Technologies*, publication number G9703-90000, **2018**.

www.agilent.com/genomics/tapestation

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

This information is subject to change without notice.

PR7000-7442

Printed in the USA, September, 2020
5994-2284EN

Agilent

Trusted Answers