

Automated Nucleic Acid Sample Quality Control in NGS Workflows

Agilent TapeStation systems





Importance of Nucleic Acid Quality Control in Next-Generation Sequencing

As biological molecules, DNA and RNA are fragile compounds that can undergo fast degradation if not handled with precaution and stored under optimal conditions. Poor nucleic acid sample guality may have a direct impact on the results of downstream experiments including next generation sequencing (NGS). NGS library preparation is a time consuming and lengthy process, typically involving enzymatic or physical DNA fragmentation and multiple amplification and ligation reactions, followed by sample clean-up steps. Due to the labor, specialized chemistry, and analysis of large datasets, NGS remains an expensive process from start to finish. To prevent a garbage-in, garbage-out situation and avoid a loss of time and money, strict quality control (QC) steps to monitor the starting material, intermediate products, and final libraries are recommended for all protocols of major vendor library preparation kits. The ideal QC solution would be easy-to-use, economical, flexible, and sensitive, while simultaneously providing timely results that are accurate and reliable. In addition to sample concentration and purity, QC results encompass size distribution as well as degradation status and sample integrity.

The Agilent TapeStation systems, based on ScreenTape technology, offer significant advantages including exceptional ease-of-use in combination with constant cost per sample. ScreenTape devices are ready to use, credit card sized, disposable gels that allow for automated separation of up to 96 samples per run. The ability to use unused gel lanes at a later time guarantees a consistent low price per sample. These key advantages makes the TapeStation system the ideal solution for NGS sample QC.

Contents



- Monitoring Library Preparation for NGS in Systems Biology 5 **Omics Analysis**
- Quality Control in Illumina Sequencing Workflows Using 12 the TapeStation System



- Comparison of DNA Assays Using the 4200 TapeStation 16 System and 2100 Bioanalyzer System
- 21 Use of the Agilent 4200 TapeStation System at the German Cancer Research Center (DKFZ)



- 27
 - DV₂₀₀ Evaluation with RNA ScreenTape Assays
- 31 Questions and Answers: An Interview with Rainer Nitsche, PhD, Product Manager TapeStation Systems

Nucleic Acid Quality Control

Quality control steps are a requirement in most NGS workflows. These steps can include the input sample, intermediate steps through the workflow, and the final library product. Information to be gathered at these crucial checkpoints include the sizing, purity, quantification, and integrity of the sample. The TapeStation systems offer a streamlined solution for the quality control of any nucleic acid.



Nucleic Acid Sizing

Automated electrophoresis with the TapeStation systems deliver accurate measurement of DNA size. A quick check of nucleic acid size gives you the confidence you need to know that you are working with the right fragments or smears.



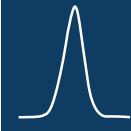
Purity

Electrophoresis separation helps to identify impurities and contaminations like PCR artifacts and primer dimers in DNA libraries or gDNA contamination in cfDNA samples. Sample QC can help you optimize your workflows to avoid these impurities.



Quantification

Electrophoresis separation followed by fluorescence detection with ScreenTape technology enables sensitive detection and quantitation of nucleic acids. Make sure your downstream calculations and dilutions are correct with quick and confident quantification.



Sample Profile/Integrity

Sample profiles give you a lot of information at a glance, allowing you to effectively qualify your nucleic acid samples before you commit to time and cost-intensive applications. Automatically calculated integrity numbers help to easily analyze DNA and RNA sample profiles.

Application Note

Monitoring Library Preparation for Next-Generation Sequencing in Systems Biology Omics Analysis

Authors

Christine Voigt Alacris Theranostics

Eva Graf Agilent Technologies, Inc.

Introduction

CanPathPro (www.canpathpro.eu), EU Horizon 2020 project, is addressing the challenge of predictive modeling of biological data. During the project, bio-informatic and experimental tools are being developed and refined for the evaluation and control of modeling predictions for systems biology. Within CanPathPro, genetically engineered mouse models are used to follow the temporal molecular and signaling changes occurring during cancer development. This includes the histology of the tumor, proteomics, as well as the genome and transcriptome using next-generation sequencing (NGS). CanPathPro aims to develop a new systems biology platform for generating and testing cancer signaling hypotheses in biomedical research. Alacris Theranostics is working in the field of precision oncology, biomarker applications, systems biology, and drug response modeling and co-ordinates the CanPathPro project.

Results and discussion

DNA library preparation

30 samples of genomic DNA (gDNA) extracted from genetically engineered mouse models were manually prepared according to the SureSelect^{XT} protocol.

Genomic DNA integrity

Quantity and integrity of gDNA were evaluated to ascertain that the starting material was suitable for library preparation. This first important QC step was performed with the 4150 TapeStation system and the Genomic DNA ScreenTape assay. Genomic DNA quality assessment based on the DNA integrity number (DIN) was previously demonstrated. The DIN algorithm is included in the TapeStation software providing an objective quality metric for standardized DNA integrity assessment. Signal distribution across the sizing range of the assay is converted into a numerical value ranging from 10 (intact DNA) to 1 (degraded DNA). Freshly extracted gDNA from mouse tissue is expected to depict high DIN values. The DIN values of the 30 gDNA samples ranged from 6.6 to 9.3 (data not shown). Since high quality could be verified for the starting material, all 30 samples were used for library preparation.

Additional Information

To read the full application note please visit agilent.com and search 5994-0946EN

Genomic DNA quantification

The SureSelect^{XT} protocol is optimized for the preparation of libraries from 3 µg gDNA starting material. The protocol recommends the Qubit system for quantification, which was used to evaluate gDNA concentration. In parallel, the samples were quantified with the 4150 TapeStation system and the Genomic DNA ScreenTape assay. DNA samples were normalized before the analysis with the Genomic DNA ScreenTape assay. The DNA quantity determined by both methods is compared in Figure 1. The concentration data evaluated with the Genomic DNA ScreenTape assay was highly comparable with the Qubit data exhibiting an R² correlation value of 98.5%. This result shows that the Genomic DNA ScreenTape assay enables reliable and accurate quantification for genomic DNA as starting material for NGS applications.

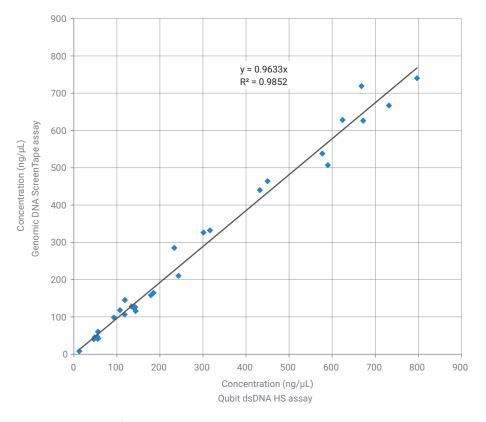


Figure 1. Comparison of gDNA concentrations measured with the Agilent Genomic DNA ScreenTape assay and with the Qubit dsDNA HS assay. The samples were diluted to match the assay ranges. The displayed concentrations represent the original sample concentrations.

Characterization of sheared DNA

The first step of the SureSelect^{XT} protocol is fragmentation of gDNA by shearing with the Covaris ultrasonicator. The shearing protocol was adjusted to produce fragments with a larger target size to allow for sequencing in 2 x 100 bp mode. Optimal shearing in NGS workflows is typically verified with the evaluation of the size distribution and electropherogram pattern of fragmented DNA samples with the DNA 1000 assay and the 2100 Bioanalyzer system. In addition, the samples were analyzed with the 4150 TapeStation system and the D1000 ScreenTape assay. Example electropherograms of sheared DNA analyzed with the 4150 TapeStation and the 2100 Bioanalyzer systems are shown in Figure 2. All samples displayed an even size distribution with no undesirable shouldering. Due to the modified shearing protocol, the expected average size range of fragmented DNA is about 100 bp larger than recommended in the protocol (between 200 and 325 bp). The fragmented DNA samples showed a maximum peak size between 260 and 310 bp on both systems, which verified optimal shearing. The size of this QC step can be used as a core value for each sample to compare with the size after adapter ligation later in the SureSelect^{XT} workflow. The sizing results of the 2100 Bioanalyzer and the 4150 TapeStation systems correlated highly with an average deviation of 2.2% (Figure 3).

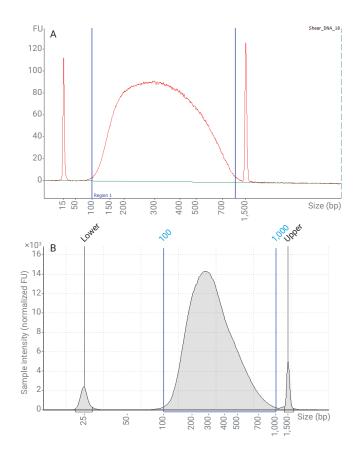


Figure 2. A fragmented DNA sample from the SureSelect^{XT} workflow showing even size distribution. A) Sheared DNA analyzed with the Agilent DNA 1000 assay, B) Electropherogram of the same sample analyzed with the Agilent D1000 ScreenTape assay.

Qualification of precapture libraries

The next steps for sheared DNA in the SureSelect^{XT} protocol are comprised of end repair, adenylation, and adapter ligation followed by amplification. To verify efficient adapter ligation, purified adapter ligated DNA samples (precapture samples) were qualified with the DNA 1000 assay (2100 Bioanalyzer system) and the D1000 ScreenTape assay (4150 TapeStation system). The precapture libraries were analyzed by region analysis with boundaries set from 200 to 700 bp. The average region size ranged from 340 to 420 bp with a maximum peak size at approximately 300 bp. All samples showed an expected size shift (approximately 30 bp) with regards to the sheared DNA samples, which is related to adapter ligation. The average size determined by both systems was highly comparable for all 30 samples with an average deviation of 2.7% (Figure 4). Hence, the average region size of libraries can be confidentially assessed using the region functionality with the DNA 1000 assay as well as with the D1000 ScreenTape assay.

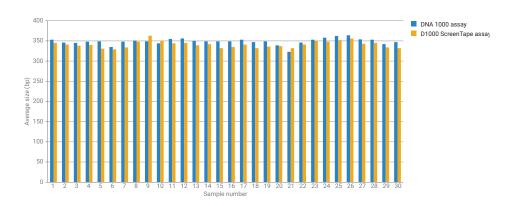


Figure 3. Average size of 30 sheared DNA samples with a region set from 100 to 1,000 bp. The average size was determined with the Agilent DNA 1000 assay and the Agilent D1000 ScreenTape assay.

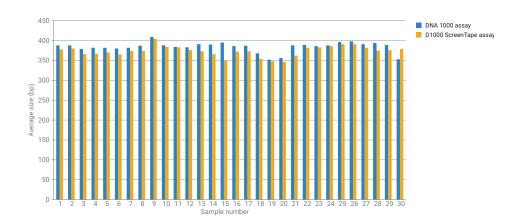


Figure 4. Comparison of the average size of 30 precapture libraries with a region set from 200 to 700 bp. The average size was determined with the Agilent DNA 1000 assay and the Agilent D1000 ScreenTape assay.

Final library QC

Adapter-ligated samples were processed according to the SureSelect^{XT} protocol, including hybridization and capturing, followed by indexing and amplification. In general, final libraries are equimolar pooled for sequencing, which requires a QC step to ensure accurate and precise evaluation of library size and quantity. The libraries were analyzed with the HS DNA assay (2100 Bioanalyzer system) and with the HS D1000 ScreenTape assay (4150 TapeStation system). Electropherogram patterns with a broad single peak and the absence of artifacts represent high-quality libraries. All samples depicted high-quality libraries, example electropherograms are shown in Figure 5.

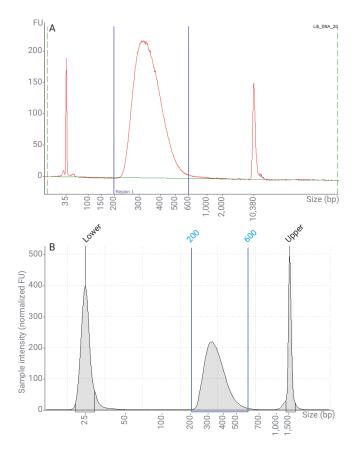


Figure 5. Example electropherograms of end products of the SureSelect^{XT} workflow. A) Final library analyzed with the Agilent HS DNA assay, B) Sample analyzed with the Agilent HS D1000 ScreenTape assay.

The maximum peak size ranged from 350 to 400 bp. Due to the increased fragment size that was aimed during shearing, the expected average size ranges from 325 to 400 bp. To determine the average size for molarity calculation, a region was set ranging from 200 to 600 bp. The sizing results of the HS D1000 ScreenTape assay were compared to the results determined by the HS DNA assay with the 2100 Bioanalyzer system. The results were highly comparable for all samples with an average deviation of 0.9% (Figure 6).

Both systems provide reliable results for final library size determination.

Alacris Theranostics uses the average size provided by the TapeStation analysis software and the concentration determined with the Qubit assay to determine molarity data. The concentration data provided by the HS D1000 ScreenTape was compared to the quantification results attained from the Qubit dsDNA HS assay (Figure 7). The quantification with HS D1000 ScreenTape assay was highly comparable to Qubit results with an average deviation of 11.1%.

The electropherogram pattern of DNA smears and libraries can differ between the 4150 TapeStation system and the 2100 Bioanalyzer system due to differences in technology. Nevertheless, resulting sizing data was highly comparable between both systems. Furthermore, the electropherogram view of

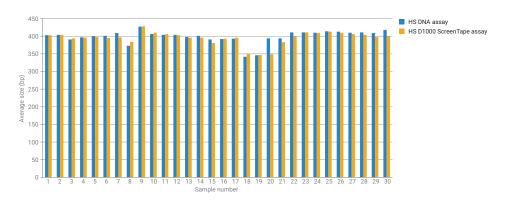


Figure 6. Sizing comparison of the final libraries generated with the Agilent SureSelect^{XT} workflow. The size was determined by region analysis ranging from 200 to 600 bp. Sizing data was determined with the Agilent HS DNA assay and compared to the Agilent HS D1000 ScreenTape assay.

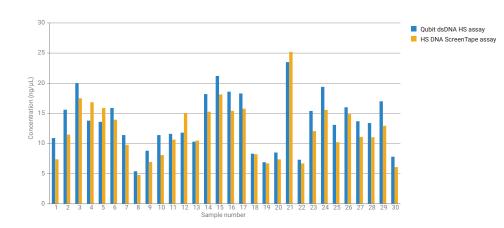


Figure 7. Quantification of SureSelect^{XT} final libraries with the Qubit dsDNA HS assay compared to the Agilent HS D1000 ScreenTape assay. The samples were diluted to match the Agilent HS D1000 ScreenTape assay range. The graph shows the original library concentration.

either system enabled visual evaluation of an even size distribution and the absence of artifacts. The quantification of final libraries with the 4150 TapeStation system provided equivalent values compared to the Qubit system. Thus, the 4150 TapeStation system together with the ScreenTape assay portfolio can be confidentially used as a QC tool during library preparation with the SureSelect^{XT} protocol.

Conclusion

Alacris Theranostics successfully evaluated the 4150 TapeStation system with the ScreenTape assay portfolio to monitor exome library preparation for systems biology omics analysis. 30 DNA samples were extracted from genetically engineered mouse models and sequenced as part of the CanPathPro project. The 4150 TapeStation system was compared with the 2100 Bioanalyzer and Qubit systems by performing quality control steps on the platforms in parallel. It was demonstrated that the 4150 TapeStation system is a reliable tool for sizing and quantification at any step of library preparation for exome sequencing. It can be used to qualify sheared DNA, precapture libraries, and final libraries with sizing results equivalent to the 2100 Bioanalyzer system. Quantitative data of all employed ScreenTape assays highly correlated with Qubit results. The ready-to-use ScreenTape devices simplify the quality control steps that are part of the library preparation workflow, and significantly reduce hands-on-time compared to the 2100 Bioanalyzer system. Moreover, the 4150 TapeStation system enables reliable quality and quantity assessment of genomic DNA with the quality metric DIN. The 4150 TapeStation system can load any sample number from 1 to 16 samples per run without wasting consumables, thus providing a constant cost-per-sample. Together with the easy switch between different assays, this gives the user high flexibility to easily accommodate to project-driven throughput.

Additional Information

To read the full application note please visit agilent.com and search 5994-0946EN

www.agilent.com/genomics/tapestation

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

This information is subject to change without notice.

© Agilent Technologies, Inc. 2019 5994-0946EN

Application Note



Quality Control in Illumina Sequencing Workflows Using the TapeStation System

Authors

Nicolle Diessl, Ute Ernst, Angela Schulz, and Stephan Wolf DKFZ Genomics and Proteomics Core Facility, High Throughput Sequencing Unit Eva Graf Agilent Technologies, Inc.

Introduction

The DKFZ is the largest biomedical research institution in Germany. The High Throughput Sequencing Unit of the DKFZ Genomics and Proteomics Core Facility provides sequencing services to national and international cancer genome projects using several Illumina paired-end and single-read sequencing platforms. The service includes QC of starting material, and library preparation including QC, clustering, sequencing, and data analysis. Various sequencing applications are provided such as whole genome sequencing (WGS), exome sequencing (WES), targeted resequencing, RNA sequencing (RNA-Seq), and sequencing of protein-binding regions (ChIP-Seq).

The selected protocols are suitable to process high-quality DNA and RNA starting material, as well as more challenging samples with low integrity or concentration. The sequencing core facility subjects all samples to an incoming QC upon receipt. During library preparation, additional QC steps are performed to monitor critical passages of the workflow. Lastly, the quality of final libraries is assessed before pooling the samples for sequencing. To verify the success of library preparation, at least one positive control is processed in parallel per batch of samples at the DKFZ sequencing core facility. This Application Note describes representative QC data of these reference samples analyzed with the 4200 TapeStation system.

Methods

Additional Information

To read the full application note please visit agilent.com and search 5994-0327EN **Custom ChIP-Seq:** The generation of adaptor-ligated libraries was performed in accordance with the NEBNext ChIP-Seq protocol. The PCR enrichment was modified, using 10 μ L of adapter-ligated DNA sample and 10 μ L of High-Fidelity 2X PCR Master Mix for the PCR reaction.

QC and sequencing of generated libraries: Final libraries were analyzed with the D1000 ScreenTape assay according to the assay guide to evaluate the library size. The molarity was calculated by the maximum peak size of the Agilent TapeStation system and the concentration measured by Qubit or Filtermax microplate fluorometer. In general, libraries of customer samples were normalized to 10 nM, equimolar pooled, transferred to a sequencing flow cell, and loaded onto the appropriate Illumina sequencer. Reference standards were used to control successful library preparation, and were only sequenced for testing, validation, or troubleshooting purposes.

Results and discussion

NEBNext ChIP-Seq: Chromatin immunoprecipitation (ChIP) sequencing is an NGS method combining ChIP with massive parallel sequencing to reveal binding sites of DNA-associated proteins. The starting material for the NEBNext ChIP-Seq workflow is chromatin immunoprecipitated DNA unlinked from protein. The workflow requires two QC steps, the first being quality assessment of the unlinked DNA, and the second the quality assessment of the final library. After incoming QC of the unlinked DNA starting material (Figure 1), the samples are end-repaired, followed by dA-tailing and adaptor ligation.

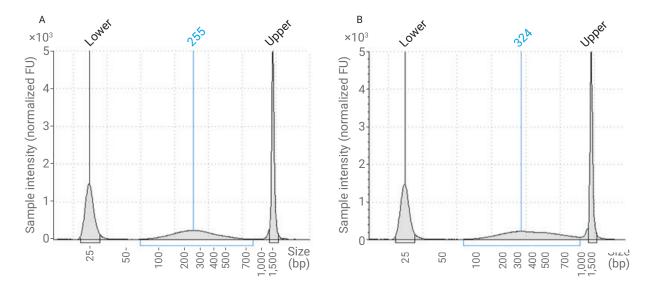


Figure 1. Size determination of starting material for ChIP sequencing with the D1000 ScreenTape assay. A) Example for a sample with an insert size of 255 bp, which was used for the protocol optimized for 250 bp. B) Example for a dsChIP-Seq DNA with an insert size of 324 bp, assigned to the protocol optimized for 300 bp.

The adapter-ligated libraries undergo a bead-based size selection. The conditions for the size selection are optimized for specific fragment lengths, which are 150, 200, 250, 300, or 400 bp. The insert size of the starting material is determined (Figure 1), and samples are collated to the closest available fragment length for the size selection step.

End products of the ChIP workflow are analyzed with the D1000 ScreenTape assay to verify the expected total library size according to the size selection step (Figure 2).

The final library includes inserts and adapters; therefore, a size shift of 120 bp for single-index adaptors, and 130 bp for dual-index adaptors is expected compared to the starting material (Figure 2A). Figure 2B shows an example of a final library with 494 bp, generated from starting material with approximately 300 bp. Larger libraries may show an increased size shift, since the bead-based size selection is not as accurate. The sizing result is used to calculate the molarity of the libraries, which are then normalized, pooled, and sequenced.

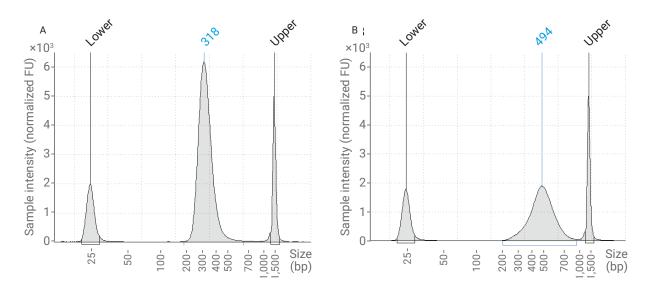


Figure 2. Final library of the ChIP-Seq workflow analyzed with the D1000 ScreenTape assay. Successful libraries show a narrow library distribution with a peak size of 120–130 bp larger than the starting material. A) Example for a final library with a maximum peak size of 318 bp, correlating to an insert size of 200 bp. B) Example of a final library with a maximum peak size of 494 bp, generated with starting material with an approximate size of 300 bp.

Conclusions

Sample QC enables monitoring of the success of library preparation for various Illumina NGS applications, minimizing the risk of producing unreliable sequencing data due to poor sample guality. The DKFZ sequencing core facility successfully implemented the 4200 TapeStation system together with the ScreenTape portfolio for quality control at key steps of multiple Illumina sequencing workflows. For the core facility as a service provider, it is essential to determine whether the input material provided by clients is fit for purpose. At the sequencing core facility, DNA or RNA starting material sample integrity is assessed with the 4200 TapeStation system, and quantification is currently performed using the Qubit assay. However, all ScreenTape assays also provide quantitative data, revealing the sample concentration within the same QC step. The intermediate QC steps are specifically useful during the establishment of a new library preparation protocol and for troubleshooting established protocols. Reference samples as positive controls are used to identify deviations, which allows for timely implementation of corrective actions in case of failure. To create optimum cluster densities during sequencing, it is crucial to accurately quantify the final libraries of all workflows. Currently, the molarity determination of final libraries is validated in the sequencing core facility with an external calculation step using the sizing results of the D1000 ScreenTape assay together with the concentration results of the Qubit fluorometer or microplate reader. However, the TapeStation Analysis software also provides molarity data using the region function, which allows to directly evaluate library size and molarity within a single QC step.

Additional Information

To read the full application note please visit agilent.com and search 5994-0327EN

www.agilent.com/genomics/tapestation

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

This information is subject to change without notice.

© Agilent Technologies, Inc., October 2018 5994-0327EN

Technical Overview

Comparison of DNA Assays Using the 4200 TapeStation System and 2100 Bioanalyzer System

Authors

Eva Graf Agilent Technologies, Inc.

Introduction

The Agilent 2100 Bioanalyzer system is a well established system for DNA quality control in multiple workflows. Specifically in NGS workflows, the sample throughput has dramatically increased, creating the need for high-throughput systems for DNA sample quality control. The Agilent 4200 TapeStation system has scalable throughput from 1 to 96 samples and walk-away operation, which are essential features for high-throughput analysis. The 4200 TapeStation system fully automates sample processing for DNA electrophoresis, including sample loading, separation, and imaging. Both platforms offer several DNA assays appropriate for a wide size and concentration range for the analysis of PCR products, fragmented DNA, and DNA libraries.

This Technical Overview compares the performance of the Agilent D1000 ScreenTape assay and Agilent High Sensitivity D1000 ScreenTape assay (HS D1000 ScreenTape assay) analyzed on the 4200 TapeStation system directly with the Agilent DNA 1000 assay and Agilent High Sensitivity DNA assay (HS DNA assay) of the 2100 Bioanalyzer system. Analytical specifications were compared by evaluating accuracy and precision of quantification and sizing with a suitable sample set of sheared DNA.

Additional Information

To read the full application note please visit agilent.com and search 5991-9093EN

Analytical Specifications

Table 1 summarizes the analytical specifications of the D1000 and HS D1000 ScreenTape assays for the 4200 TapeStation system and the specifications of the DNA 1000 and HS DNA assays for the 2100 Bioanalyzer system. The analytical specifications of the ScreenTape assays were previously systematically validated for sensitivity, sizing, quantification, and molarity.

Results and Discussion

Molarity

Sample molarity is often used to determine the load volume of NGS libraries for pooling before sequencing. The calculation of molarity is based on average size and concentration of library samples. Thus, for accurate molarity results and successful library sequencing, sizing and quantification must be accurate. Excellent assay performance for NGS library samples was shown previously for the D1000 and HS D1000 ScreenTape assays. It was also demonstrated that the assays are suitable for sample quality control in a high throughput sequencing environment. Table 1. Analytical specifications of the D1000 and HS D1000 ScreenTape assays (4200 TapeStation system) and the DNA 1000 and HSDNA assays (2100 Bioanalyzer system).

	Agilent 4200 Tap	eStation System	Agilent 2100 Bioanalyzer System		
Analytical specifications	D1000 ScreenTape Assay	High Sensitivity D1000 ScreenTape Assay	DNA 1000 Assay	High Sensitivity DNA Assay	
Sizing range	35-1,000 bp	35-1,000 bp	25-1,000 bp	50-7,000 bp	
Sizing accuracy*	±10%**	±10%**	±10%	±10%	
Sizing precision*	5% CV	5% CV	5% CV	5% CV	
Quantitative range (DNA fragments)	0.1-50 ng/µL	10-1,000 pg/µL	0.5-50 ng/µL*	5-500 pg/µL*	
Quantitative range (DNA smears)	5-100 ng/µL	0.5-15 ng/µL	-	0.1-10 ng/µL	
Quantitative accuracy	±20%***	±20%***	±20%*	±20%*	
Quantitative precision	0.1-1 ng/μL: 15% CV 1-50 ng/μL: 10% CV	15%	25-500 bp: 15 % CV* 500-1,000 bp: 5 % CV*	50-2,000 bp: ± 15%* 2,000-7,000 bp: ± 10%*	

* Determined by analyzing the respective ladder as sample.

** Accuracy of software ladder ±20 %.

*** Measured against the Agilent 2200 TapeStation system.

The lower and upper size limits of the DNA assays (Table 1) apply for maximum peak sizes of DNA fragments. DNA libraries must match the assay range with the whole sample size to fit between the lower and upper marker of an assay. Table 1 shows that the quantitative range for DNA smears differs from the quantitative range for DNA fragments.

A dilution series of sheared DNA was analyzed with the D1000 and HS D1000 ScreenTape assays on the 4200 TapeStation system. A ladder was run on each ScreenTape to ensure the most accurate sizing.

The molarity data was directly compared to the data of the same samples analyzed with the DNA 1000 and HS DNA assays of the 2100 Bioanalyzer system. The results showed excellent correlation for molarity analysis of sheared DNA with the 4200 TapeStation system compared to the 2100 Bioanalyzer system (Figure 1).

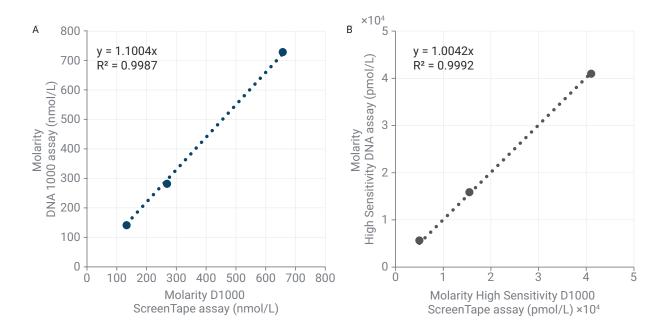


Figure 1. Correlation of molarity data of sheared DNA samples in three concentrations (n = 6). The results of the 4200 TapeStation system are plotted on the X-axis and the results obtained with the 2100 Bioanalyzer system on the Y-axis. The molarity was evaluted using region functionality. A) Sample molarities obtained with the D1000 ScreenTape assay compared with molarity data from the DNA 1000 assay. B) Correlated molarity data of sheared DNA samples analyzed with the HS D1000 ScreenTape assay and the HS DNA assay.

Figures 2 and 3 show that the electropherogram patterns of the ScreenTape assays for DNA smears are not identical to the pattern obtained with the 2100 Bioanalyzer system. This effect occurs due to different separation processes during gel electrophoresis on the Bioanalyzer chips and the ScreenTape devices. The position of the region on the X-axis depends on the size range of the assay.

Automatic integration and data processing of the TapeStation Analysis software and the 2100 Expert software generate highly comparable results for sizing, quantification, and molarity analysis of sheared DNA samples measured with the 4200 TapeStation and the 2100 Bioanalyzer systems.

Conclusion

This Technical Overview shows molarity data of DNA ScreenTape assays for the Agilent 4200 TapeStation system highly correlate with data of equivalent assays on the Agilent 2100 Bioanalyzer system. Both systems show similar performance of DNA smears. Electropherogram patterns of sheared DNA analyzed with ScreenTape assays differ slightly from patterns obtained with corresponding assays of the 2100 Bioanalyzer system. For the most accurate sizing and molarity results, it is recommended to run a ladder on each ScreenTape.

Additional Information

To read the full application note please visit aglient.com and search 5991-9093EN

www.agilent.com/genomics/tapestation

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

This information is subject to change without notice.

© Agilent Technologies, Inc., May 2018 5991-9093EN

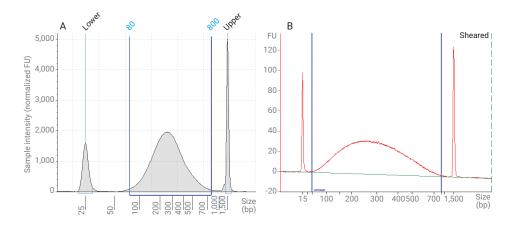


Figure 2. Electropherogram pattern of sheared DNA analyzed with the 4200 TapeStation and 2100 Bioanalyzer system. A) Electropherogram of sheared DNA separated with the D1000 ScreenTape assay. B) Electropherogram of the same sample obtained with the DNA 1000 assay.

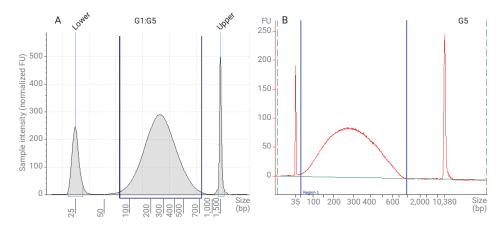


Figure 3. Electropherogram patterns of sheared DNA analyzed with the 4200 TapeStation and 2100 Bioanalyzer systems. A) Example electropherogram of a sheared DNA sample analyzed with the HS D1000 ScreenTape assay. B) The same sample analyzed with the HS DNA assay.

5 Reasons To Use the TapeStation System for Nucleic Acid Quality Control

No matter if you choose the higher throughput 4200 TapeStation system that is compatible with 96-well sample plates, or the more compact 4150 TapeStation system that allows the analysis of up to 16 samples, both platforms offer all advantages of ScreenTape technology. Complete assay compatibility guarantees seamless switching between systems.

Enjoy Exceptional Ease-of-Use	Obtain Quality Metrics	Save Samples	Save Money	Save Time
 Ready-to-use ScreenTape devices. Intuitive software for instrument operation and data analysis. 	 DIN - DNA integrity number for genomic DNA samples. RIN^e - RNA integrity number equivalent for total RNA samples. %cfDNA - to assess the quality of Cell-free DNA samples 	 Only 1-2 µL of sample required for high-sensitivity applications. Sensitive detection of nucleic acids down to 5 pg/µL. 	 ScreenTape technology offers constant cost per sample independent of throughput. Sizing, quantitation, and integrity analysis all in one step. 	 Minimize preparation: ScreenTape devices are ready-to-use. Fast analysis time: 1-2 minutes per sample. Switch applications simply by changing the ScreenTape device.
	2	3	4	5

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

© Agilent Technologies, Inc. 2019

Application Note

Use of the Agilent 4200 TapeStation System for Sample Quality Control in the Whole Exome Sequencing Workflow at the German Cancer Research Center (DKFZ)

Authors

Jana Petersen Marie Beckhaus, and Stephan Wolf DKFZ Genomics and Proteomics Core Facility, High Throughput Sequencing Unit

Elisa Viering Agilent Technologies, Inc.

Additional Information

To read the full application note please visit agilent.com and search 5991-7615EN

Introduction

The German Cancer Research Center (DKFZ) is one of the largest biomedical research institutions in Germany. The High Throughput Sequencing Unit of the DKFZ Genomics and Proteomics Core Facility provides sequencing services to national cancer genome projects for multiple applications including whole exome sequencing.

Mandatory for the experimental success of whole exome sequencing is the quality of the incoming genomic DNA (gDNA) material and the DNA samples at various stages of the library preparation workflow. The High Throughput Sequencing Unit subjects the received gDNA samples to an incoming quality control (QC), prepares the libraries with additional QC steps, and performs the sequencing. The sequencing is followed by the final data QC analysis according to the workflow shown in Figure 1.

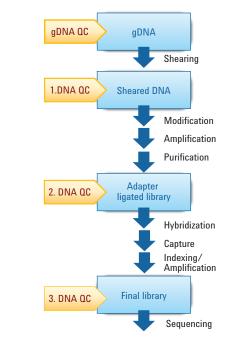


Figure 1. Workflow of the Agilent SureSelect^{XT} automated target enrichment protocol for Illumina paired-end multiplexed sequencing established at the DKFZ High Throughput Sequencing Unit. The QC steps are performed by using the Agilent 4200 TapeStation system with the Agilent Genomic DNA and Agilent D1000 ScreenTape assays.

Incoming gDNA QC

To determine if the samples were suitable for library preparation, a QC was performed at the beginning for the batch of 88 gDNA samples from FFPE tumor tissue by the High Throughput Sequencing Unit. This initial QC includes quantification and analysis with a 4200 TapeStation system and the Genomic DNA assay to determine DNA quality based on the DNA integrity number (DIN). The DIN algorithm is included in the 4200 TapeStation software, and provides a numerical assessment of the DNA sample by assigning a numerical score for DIN from 1 to 10. A high DIN indicates highly intact gDNA, and a low DIN a strongly degraded gDNA sample.

Figure 2 shows a representative subset of the data collected on the 4200 TapeStation system with the Genomic DNA ScreenTape assay. gDNA samples extracted from FPPE material are often of low DNA integrity, but can still be sufficiently intact for whole exome library preparation and successful sequencing. Depending on the quality of the DNA material, modified fragmentation protocols can be applied for optimized results.

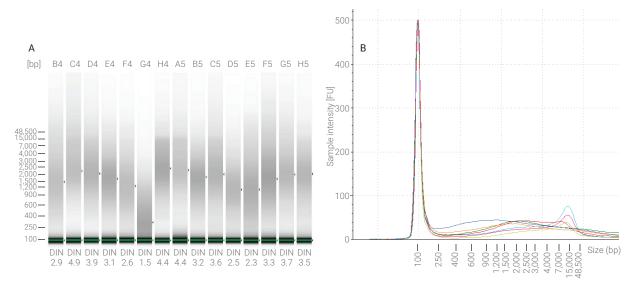


Figure 2. QC for gDNA extracted from FFPE tissue was performed by using the Agilent 4200 TapeStation system and the Agilent Genomic DNA ScreenTape assay. A) The gel image of 15 samples with the determined DIN on bottom and the well on top of each lane are shown. B) An overlay of the electropherograms of 15 samples. Both views enable comparison and visualization of the integrity of the samples.

Based on empirical evidence, the High Throughput Sequencing Unit had established a DIN threshold of 7 for their incoming gDNA samples. Figure 3A shows the DNA integrity of all FFPE samples in this study. DIN values ranged from DIN 1.3 to 6.2, and are well below the QC threshold of DIN 7. Samples with lower DIN require a special review, and are only subjected to library preparation if there was approval to proceed from the customer. Depending on the quality, modified protocols can be used. However, in these cases, successful library preparation cannot be guaranteed.

The second criterion for QC of incoming samples is the calculation of total DNA abundance. A minimum of 450 ng gDNA is required for DNA library preparation. Samples were quantified in triplicate using the Qubit system. The Qubit assay is the established quantification method of the High Throughput Sequencing Unit. The concentration analysis of the 4200 TapeStation was used as control to double check the Qubit quantification.

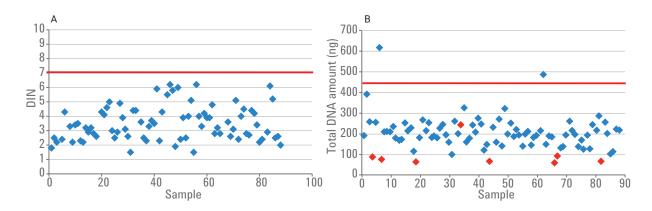


Figure 3. The incoming QC consisting of quality analysis was performed using the Agilent 4200 TapeStation system and the Agilent Genomic DNA ScreenTape assay. A) Qualification of the 88 gDNA samples with the DNA Integrity number (DIN), the red line indicates the incoming QC criteria for DNA quality (DIN \ge 7). B) Quantification of the 88 gDNA samples. The total DNA amount was calculated based on the measured concentration and the volume. The eight gDNA samples marked in red were not further processed. The red line indicates the incoming QC criteria for DNA quantity (\ge 450 ng).

To calculate the total amount of DNA present in the sample, the remaining sample volumes were determined. Based on the measured volume and concentration, the total available DNA amount was calculated (Figure 4).

The DNA amount of most of the 88 tested gDNA samples was relatively low. Despite the low gDNA quality and the low DNA amount, it was decided to proceed with the library preparation for 80 of the samples. Eight samples were not further processed after the incoming QC, and were replaced with positive and negative controls.

After initial quality control, 80 customer samples were prepared according to the Agilent Low Input Exome Sequencing Human v5 protocol using automated processing in a 96-well plate format.

Final library QC

The final libraries (Figure 4) were expected to be sized between 250 and 350 bp with a minimum concentration of 2 ng/ μ L. The 4200 TapeStation and Qubit systems were used for quantification. This study exhibited successful DNA library preparation for all 80 samples, the six positive control samples, and even the three samples that were slightly outside the QC criteria after adapter ligation. All customer samples were used for sequencing.

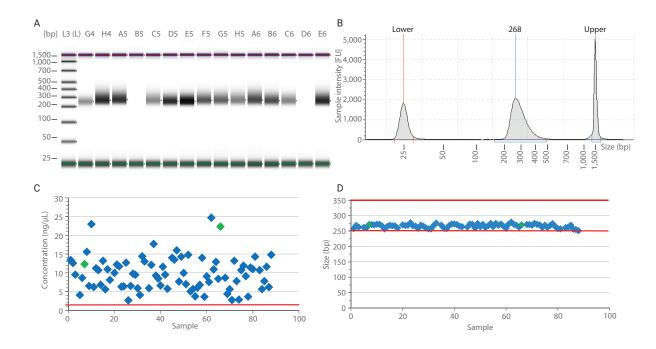


Figure 4. The sizing analysis of finished library was performed using the Agilent 4200 TapeStation system and the Agilent D1000 ScreenTape assay. A) Gel view of 15 samples, lane B5 and D6 show negative controls. B) Example of an electropherogram of one sample. The Agilent D1000 Assay includes a lower and upper marker. C) The distribution of the concentration for all 80 samples plus eight controls. The two positive controls are shown as green symbols. The red lines indicate the recommend concentration threshold (2 ng/ μ L). D) The size distribution of all 80 samples plus eight controls. The two positive controls are shown as green symbols. The red lines indicate the recommend size range (250 to 350 bp).

Sequencing

The 80 libraries were equimolar pooled; each pool was sequenced on two lanes using the Illumina HiSeq 4000 System (PE 100).

Table 1 summarizes the data QC criteria for whole exome sequencing.

Figure 5 summarizes the sequencing results. The mapping rate (dark blue) for all samples is approximately 100%. Between 55 to 70% of the mapped reads are on-target (green). The data of both sequenced lanes are merged to calculate the average target coverage (turquoise); the coverage for all samples is above 100%. The duplicate rate (red) varies between 4 and 15%. VerifyBAM FREEMIX (purple) is a tool to determine the contamination level of genomes. Only four of the 80 samples have a contamination level slightly above 2%. Structurally, changes in DNA, which occur often in tumor samples, can cause this increase of the VerifyBAM FREEMIX level.

Despite the relatively low DNA integrity and total DNA amount of the gDNA starting material, it was possible to obtain meaningful sequencing results.

QC of finished libraries assisted in predicting sequencing results. In this sample set, all libraries fulfilled the quality criteria, and all libraries were successfully sequenced. Table 1. Used data QC criteria for exome sequencing.

QC criterion	Range
Quentenon	Kange
Mapping rate	> 90%
Duplicate rate	≤ 10%
On target mapped reads	> 65%
Average target coverage	> 100x
Contamination level (VerifyBAM FREEMIX)	< 2%

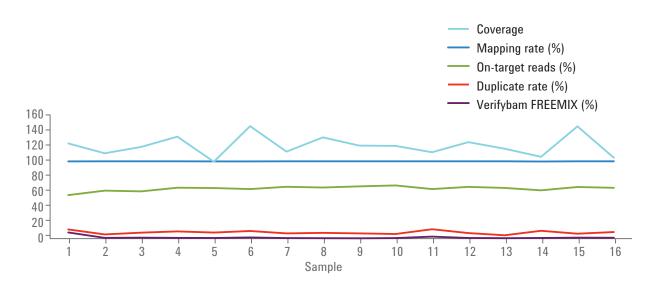


Figure 5. The QC values to determine the quality of sequencing data are mapping rate, on-target mapped reads, average coverage, duplicate rate, and the VerifyBAM FREEMIX contamination level. The graph shows an example with the sequencing quality criteria for 16 samples sequenced on one lane.

Additionally, the 4200 TapeStation system is a useful tool for troubleshooting the entire NGS library preparation workflow. Based on data obtained, steps to optimize established workflows may be taken. The low hands-on time and flexible 1-to-96-well plate format allows the 4200 TapeStation system to easily be included in any existing NGS workflow, especially for automated processes. Furthermore, a report with the results can be created for customers, and the data can be included in individual LIMS systems.

Conclusion

This Application Note demonstrates the successful integration of the Agilent Genomic DNA and Agilent D1000 ScreenTape assays using the Agilent 4200 TapeStation system as a QC tool into the whole exome sequencing workflow. The benefit of using the 4200 TapeStation system is that it enables optimization of the sequencing library preparation process from the beginning to the end with a final QC of finished libraries to ensure good sequencing results. The 4200 TapeStation system provides critical information at multiple key steps, from incoming gDNA material to predicting valuable sequencing results.

Additional Information

To read the full application note please visit agilent.com and search 5994-7615EN

www.agilent.com/genomics/tapestation

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

This information is subject to change without notice.

© Agilent Technologies, Inc., December 2016 5991-7615EN

Technical Overview

DV₂₀₀ Evaluation with RNA ScreenTape Assays

Authors

Eva Graf Agilent Technologies, Inc.

Abstract

The success of library preparation for RNA sequencing workflows depends highly on the quality of RNA starting material. The RNA integrity number equivalent (RIN^e) obtained with Agilent RNA ScreenTape assays is a reliable and reproducible metric to evaluate sample integrity for standard RNA sequencing. Library preparation of RNA samples originating from formalin-fixed paraffin-embedded (FFPE) tissue is challenging, due to RNA degradation and restricted sample volume, thus tailored protocols are recommended for FFPE RNA library preparation. The fragment size distribution of FFPE RNA samples has a major influence on the library yield, and can be represented as the percentage of RNA fragments above 200 nucleotides (nt) by the DV_{200} quality metric. The Agilent RNA ScreenTape and Agilent High Sensitivity RNA ScreenTape assays are convenient for highly reproducible DV_{200} evaluation of degraded RNA samples extracted from FFPE tissue. The Agilent TapeStation Analysis software enables simple assessment and processing of DV_{200} data by region analysis, which can be automated for repeated DV_{200} analysis.

Data Analysis

The DV_{200} represents the percentage of RNA fragments larger than 200 nucleotides with respect to all RNA fragments. TapeStation Analysis software can be used to calculate the DV_{200} by defining a region for fragments larger than 200 nt. The software provides several ways for region setup, suitable for low to high sample throughput. Consult the Agilent Information Center (AIC) for further information and how-to videos about region analysis.

Additional Information

To read the full application note please visit agilent.com and search 5991-8355EN

Region Setup for Individual Samples

To define a DV_{200} region for individual samples, click **Region**, then right click in the electropherogram to add a region. The region borders can be individually adjusted by shifting to the left or to the right.

Region Setup for Data Files

To define a DV_{200} region as default for all samples in an individual data file, click **Region**, then click **Region Settings**, and enter the lower and upper limits of the region in the dialog box (Figure 1). To identify the region in the electropherogram, type DV_{200} as region comment. To apply the region to all samples in the active file, select **This File**.

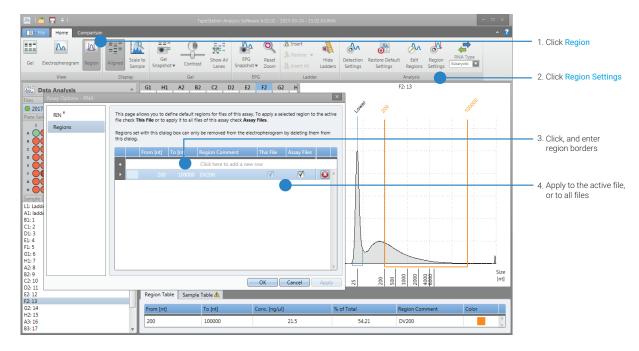


Figure 1. Setting up default regions to calculate DV_{2n0} with the Agilent TapeStation Analysis software.

Results and Discussion

Reproducibility

A set of three FFPE RNA samples, kindly provided by the tissue bank of the National Center for Tumor Diseases (NCT, Heidelberg, Germany) with five individual replicates was used to verify the reproducibility of the DV_{200} with the RNA and the High Sensitivity RNA ScreenTape assay. Both assays showed high reproducibility with a CV of less than 5% (Figure 2).

Concentration Dependence

A dilution series for three individual FFPE RNA samples was analyzed with the RNA and High Sensitivity RNA ScreenTape assays to determine the consistency of the DV_{200} analysis (Table 1). Both the RNA ScreenTape and High Sensitivity RNA ScreenTape assay yielded comparably high DV_{200} precision of less than 5% CV.

The dilution series showed that DV_{200} evaluation is most precise when measuring above the lower limit of the specified assay concentration range, data not shown.

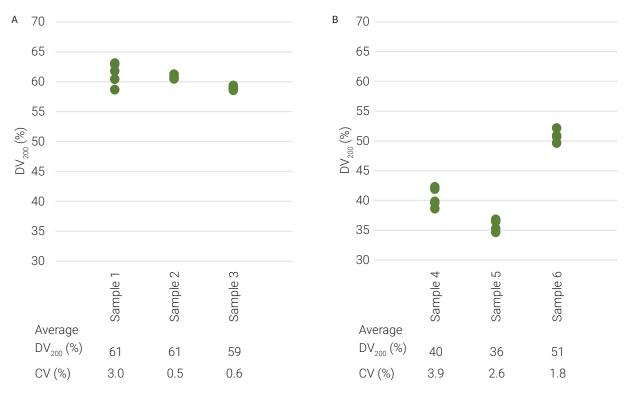


Figure 2. Reproducibility of DV₂₀₀ analysis for three individual samples with (A) the Agilent RNA ScreenTape assay and (B) the Agilent High Sensitivity RNA ScreenTape assay.

	Agilent RNA ScreenTape assay			Agilent High Sensitivity RNA ScreenTape assay		
	Sample 7	Sample 8	Sample 9	Sample 10	Sample 11	Sample 12
Average DV_{200} in %	46	52	42	57	66	64
Standard deviation	1.1	1.6	1.7	1.6	0.6	0.9
% CV	2.4	3.1	4.1	2.8	1.0	1.4

Table 1. Consistency of DV_{200} analysis with the Agilent RNA and the Agilent High Sensitivity RNA ScreenTape assay measured in dilution series for three samples per assay.

Conclusion

The Agilent RNA and Agilent High Sensitivity RNA ScreenTape assays together with the Agilent 4200 TapeStation enable fast and easy quality analysis of FFPE RNA samples. The Agilent TapeStation Analysis software displays the DV_{200} after region setup as a percentage of total. The RIN^e algorithm remains unaffected by the DV_{200} calculation. The DV_{200} evaluation with the RNA and High Sensitivity RNA ScreenTape assays is highly reproducible, and both assays yield highly comparable results. For best performance, it is recommended to run a ladder together with the samples. Within the specified assay ranges, the DV_{200} is consistent in dilution series. Highly degraded samples with insufficient marker separation should be diluted for most accurate results.

Additional Information

To read the full application note please visit agilent.com and search 5991-8355EN

www.agilent.com/genomics/tapestation

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

This information is subject to change without notice.

© Agilent Technologies, Inc., September 2017 5991-8355EN

Questions and Answers

TapeStation Systems for NGS Sample QC

An Interview with Rainer Nitsche, PhD

Q: Why should scientists perform quality control of their nucleic acid samples in the context of next generation sequencing?

A: Sample quality control is required by all major vendors of library preparation kits for next generation sequencing (NGS). NGS experiments are expensive as associated library preparations are complex, lengthy, and involve many steps. In addition, sequencers generate massive amounts of data that need to be analysed. The goal must be to avoid a garbage-in, garbage-out situation. So, it's in the best interest of scientists to perform all required and recommended nucleic acid quality control (QC) steps to prove that samples are fit-for-purpose. This includes checking for sample concentration, sizing, molarity, and integrity at multiple checkpoints, including the starting material, intermediate products, and final libraries. Low guality NGS libraries include those that are too low in concentration or contain impurities such as adapter dimers or PCR artefacts. If these are loaded onto a flow cell, the sequencing results will lead directly to poor metrics including low target coverage and mapping rates, and high duplication rates.

A variety of factors impact sample quality; for example, the nature, age, and storage conditions of the sample material as well as the extraction method. In general, DNA and RNA molecules are fragile and can undergo fast degradation if not stored and handled correctly. A common method to preserve tissue material is to fix it with formalin and embed it into paraffin. By this procedure, material can be stored at room temperature for many years, during which time degradation is possible. This formalin-fixed paraffin-embedded (FFPE) tissue represents a valuable source of genetic information, as millions of samples are stored in biobanks. Depending on the fixation method, age, and storage conditions, nucleic acids extracted from FFPE material can exhibit a large range of integrity. If degraded source material is the only option available, some library preparation methods may require protocol modifications to help obtain the best library possible. This may include increasing the input concentration if samples exceed a certain level of degradation or changing the fragmentation conditions depending upon the sizing of the sample. Thus, the quality of the starting material to be used for library preparation needs to be monitored closely.



Rainer Nitsche, PhD, Project Manager TapeStation systems, Agilent Technologies

Another example of when to QC NGS libraries is when multiple libraries are to be pooled for sequencing. In order to achieve good results for each of the libraries, the samples must be pooled at equimolar amounts before they are loaded on a sequencer. This requires the exact analysis of each library size distribution and concentration to determine molarity. Pooling the libraries at equimolar concentrations will allow optimal usage of the flow cell, leading to good sequencing results.

The method of choice for DNA and RNA sample QC is separation by electrophoresis. Agilent offers several instrument platforms that address different user needs for throughput and applications. For the integrity analysis of genomic DNA or total RNA, Agilent has developed standards that are used on the TapeStation systems: the DNA integrity number equivalent (DIN) and RNA integrity number equivalent, (RIN^e). Both quality scores range from one to 10, where 10 indicates an excellent sample integrity. The integrity numbers provide the user an independent, objective, and reproducible tool for nucleic acid quality assessment. With the help of these software provided integrity numbers, it is very easy for the user to interpret electrophoretic data and to establish sample type and application specific thresholds that need to be met. Samples that are below the threshold limit would be rejected from the workflow to save money and time by ensuring experimental success.

Q: What are the benefits of the TapeStation systems?

A: The TapeStation systems use proprietary and unique ScreenTape technology developed by Agilent, which enable some key advantages for our customers and addresses major pain points in NGS nucleic acid sample QC. The ScreenTape technology offers unmatched ease-of use, higher sensitivity, speed, and resolution over classical slab gel electrophoresis. The ScreenTape device is a credit card sized plastic disposable that enables the electrophoretic separation of up to 16 samples in individual lanes, eliminating any cross-contamination between samples. Each sample lane includes a preloaded sieving polymer matrix, electrodes, and a buffer chamber. A major benefit of this technology is that the ScreenTape device does not require any preparation. Simply remove the ScreenTape from the packaging, place it onto the instrument, and the instrument is ready to run immediately following sample preparation. There is no need to use all 16 lanes of the ScreenTape device during a single run. Only run the number of samples that you need and save the remaining lanes on the ScreenTape for use later. This enables maximum flexibility in throughput as well as constant costs per sample.

A large portfolio of different DNA and RNA ScreenTape assays is available that addresses different customer needs for separation range and sensitivity. Depending on the assay, only 1 to 2 μ l of precious sample is required, with concentrations as low as 5 pg/ μ l. Each ScreenTape device is barcoded and automatically detected by the instrument, so that the specific assay settings, such as electrophoresis parameters and data analysis configurations, are automatically loaded.

Agilent offers two instruments that are based on ScreenTape technology. The 4200 TapeStation system is compatible with 96-well plates and can run up to 96 samples unattended, while the newer 4150 TapeStation instrument is a low-throughput instrument for the analysis of one to 16 samples. Both instrument versions provide full compatibility with the existing ScreenTape application portfolio.

Q: Can you please highlight the differences between the individual TapeStation models?

A: The 4200 TapeStation system, that was introduced in 2015, provides nucleic acid sample QC solutions that meet the needs of mid-to-high throughput customers. As these customers typically process samples in well plate format, the 4200 TapeStation instrument is compatible with 96 well plates. There has been growing demand from budget restricted customers for a price-attractive instrument that offers all the advantages of the ScreenTape technology for lower throughput customers. While the 4200 TapeStation instrument can process up to 96 samples in an automated and walk-away fashion, the 4150 TapeStation instrument allows the analysis of up to 16 RNA or DNA samples. Both instruments share the same set of accessories and consumables. In addition, both systems are compatible with all DNA and RNA ScreenTape assays with the same specifications of sensitivity, sizing and quantitation accuracy and precision. The compact footprint of the 4150 TapeStation saves valuable bench space in the lab, which is highly appreciated in smaller labs. In case customers develop a need for a higher throughput solution, it is possible to seamlessly migrate to the 4200 TapeStation system. This minimizes the requirement of operator familiarization and training as well as the effort of lab method revalidation.

Q: What is the operation principle of the Agilent TapeStation instruments? How are DNA and RNA samples analysed?

A: The TapeStation instruments are gel electrophoresis systems. The separation takes place in a credit-card sized, ready to be used consumable, the ScreenTape device. The application specific ScreenTape device is placed into the instrument and detected by the barcode. Only a minimum amount of sample preparation is required. Typically, the sample is mixed in a well plate or strip tube in a given ratio with an application specific sample buffer. The buffer is available in the reagent kit or as a standalone product and includes a fluorescent dye and internal standards which are required for the calculation of sample size and quantity. To ensure a complete and homogeneous mix of sample and buffer, a dedicated vortex mixer is included as an accessory in both TapeStation systems. If the samples are RNA, they must be shortly heat denatured to remove secondary structures that might interfere with the electrophoretic separation. To reduce evaporation during the run, it is recommended to seal the well plate with a provided pierceable foil. The well plate or strip tubes are then placed onto the instrument, and the method is started. During the run, the prepared samples are automatically loaded onto the ScreenTape device using a robotic single channel pipettor. Individual sample lanes and pipette tips eliminate cross contamination between samples. The electrophoretic separation takes between 1-2 minutes

per sample. During separation, the dye intercalates with the nucleic acids, creating a fluorescent signal with the instrument LED. The imaging system takes pictures of each lane and the data is processed by the TapeStation software. After the run is finished, the ScreenTape device can be put back at 4°C and unused lanes can be used in the following two weeks. The higher throughput 4200 TapeStation system has the ability to remove the used ScreenTape device using a tape gripper, and replace it with a new one that is pre-loaded onto a dedicated rack during instrument preparation.

The TapeStation software is intuitive and user friendly and presents the data in a variety of formats. Results can be reviewed as a gel image or an electropherogram trace in which fluorescence intensity is plotted against size. In addition, calculated data like sizing, quantitation, molarity and integrity values are presented in a customizable table format. With a few mouse clicks, the software generates customizable reports that can be stored and shared with colleagues. Furthermore .xml and .csv export functionality allows the data to be exported into a customer specific lab information management system (LIMS).

Dedicated software algorithms have been implemented in the TapeStation software for the integrity analysis of genomic DNA and total RNA samples. The DNA integrity (DIN) and RNA Integrity number equivalent (RIN^e) enable an objective and reproducible assessment of sample integrity. Both integrity numbers range from 10 (complete intact sample) to one (very much degraded sample). Users need to define and validate individual integrity number thresholds for individual sample types and workflows and can thereby guarantee the success of their downstream applications.

www.agilent.com/genomics/tapestation

For Research Use Only. Not for use in diagnostic procedures. This information is subject to change without notice.

© Agilent Technologies, Inc. 2019

Learn more www.agilent.com/genomics/tapestation

U.S. and Canada 1 800 227 9770 agilent_inquiries@agilent.com

Europe UK 0800 096 Denmark 45 8830 5083 Germany 0800 180 66 78 Netherlands 0800 022 7243 Other EU 45 3136 9878 info_agilent@agilent.com

Asia Pacific China 800 820 3278 Singapore 65 6571 0888 inquiry_lsca@agilent.com

For Research Use Only. Not for use in diagnostic procedures. This information is subject to change without notice.

PR7000-7080

© Agilent Technologies, Inc. 2019 Published in the USA, November 1, 2019 5994-1591EN

