Application Note
Genomics

Agilent
Trusted Answers

Pre-Analytical RNA Integrity Assessment for Biomarker Discovery with Agilent TapeStation Systems

Authors

Whitney Pike,
Agilent Technologies

Cheryl Sesler,
Guzel Shaginurova,
Lukasz Wylezinski,
Elena Grigorenko,
Frank Cockerill III,
Charles Spurlock III,
Decode Health

Abstract

RNA sequencing (RNA-seq) enables analysis of gene expression for biomarker discovery. However, variability in sample quality and processing can compromise data reliability. Decode Health, an advanced healthcare analytics company that leverages artificial intelligence and machine learning to detect, monitor, and predict patient outcomes, has developed a robust, end-to-end quality control (QC) framework that spans pre-analytical, analytical, and postanalytical stages to enhance reproducibility in RNA-seq workflows. Central to this framework is the Agilent TapeStation system, which assesses RNA integrity and fragment size distribution, as well as final library quality. This application note demonstrates the importance of reliable pre-analytical QC to ensure sequencing success and data consistency.

Introduction

The identification of molecular biomarkers is a key objective in biological research, offering insights into cellular processes, disease mechanisms, and environmental responses. RNA sequencing (RNA-seq) has become a foundational tool for transcriptome analysis, enabling an unbiased and comprehensive examination of gene expression across diverse biological conditions. Its ability to detect novel transcripts, isoforms, alternative splicing events, and subtle expression changes makes it particularly valuable for biomarker discovery. However, RNA-seq-based biomarker discovery faces several challenges, including variability in sample preparation, sequencing, and data analysis pipelines. Addressing these issues with standardized methods is vital for ensuring highly reproducible and reliable data.

To address these challenges, Decode Health has implemented a streamlined quality control (QC) framework that spans the entire biomarker discovery RNA-seq workflow (Figure 1). The approach involves three key areas: pre-analytical, analytical, and postanalytical QC. Pre-analytical QC encompasses specimen collection, RNA isolation, DNase treatment, and library preparation. Analytical QC includes sequencing and alignment processes, while postanalytical QC involves advanced bioinformatics analyses and multimethod validation. This comprehensive approach is designed to minimize technical variability and enhance the reproducibility of RNA-seq-based biomarker discovery.[1]

A critical component of the pre-analytical QC process is the assessment of RNA integrity, which directly impacts the downstream RNA-seq data.

Decode Health uses the Agilent TapeStation system to evaluate RNA quality prior to library preparation. This automated electrophoresis platform provides fast, reproducible analysis of RNA samples, generating RNA Integrity Number equivalent (RIN^e) and DV_{200} scores that help determine sample suitability for sequencing. In addition to RNA integrity assessment, the TapeStation is also used to evaluate the quality and fragment size distribution of the final sequencing libraries, ensuring they meet the necessary criteria before proceeding to the analytical phase. This application note highlights the role of the TapeStation in pre-analytical QC within the RNA-seq pipeline, emphasizing its importance in assessing sample quality and suitability for sequencing, and ensuring successful downstream processing.

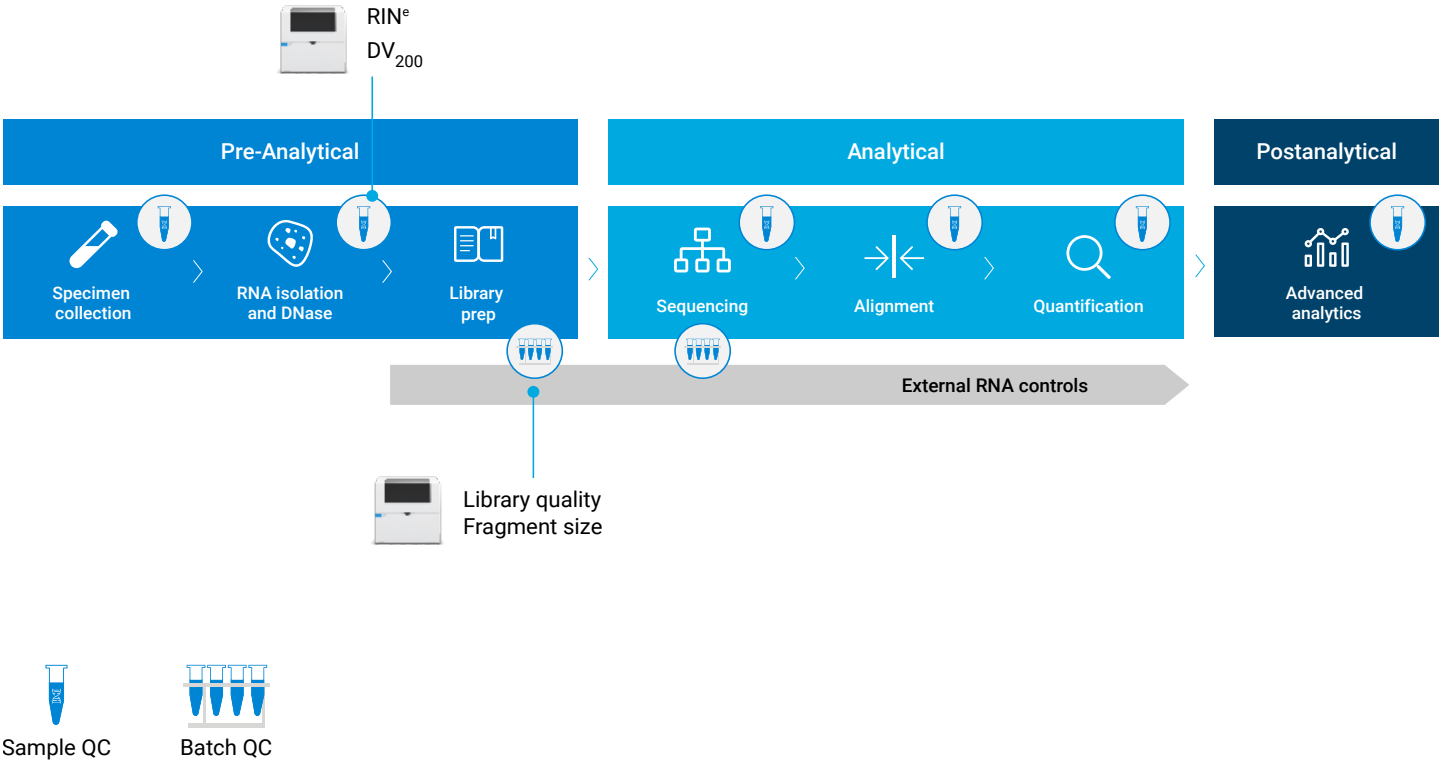

Figure 1. QC workflow used by Decode Health, detailing the pre-analytical, analytical, and postanalytical QC steps used throughout the RNA sequencing biomarker discovery process. Steps where the Agilent TapeStation systems are used for RNA integrity analysis and library quality control before sequencing are highlighted.

Methods

Peripheral whole blood was collected into PAXgene Blood RNA tubes (BD Biosciences, p/n 762165). Total RNA was isolated using the QIAcube and the PAXgene Blood RNA kit (IVD) (Qiagen, p/n 762164), according to the manufacturer's protocol. RNA and DNA concentrations were determined on the Qubit 4.0 Fluorometer (Thermo Fisher Scientific, p/n Q32850, Q32853, Q33223, Q33224). RNA integrity was assessed using the Agilent 4200 TapeStation system (p/n G2991BA) with the Agilent RNA ScreenTape (Agilent Technologies, p/n 5067-5576) and reagents (Agilent RNA ScreenTape ladder p/n 5067-5578; Agilent RNA ScreenTape sample buffer p/n 5067-5577). Agilent TapeStation analysis software assigns RNA quality metrics for integrity characterization, including the RIN^e and the DV_{200}.

A secondary DNase treatment was performed using the TURBO DNA-free Kit (Thermo Fisher Scientific, p/n AM1907), followed by RNA sample cleanup with the RNA MagClean DX kit (Aline Biosciences, p/n C1005). Libraries were prepared with the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, p/n E7760S/L), which includes random fragmentation, conversion to double-stranded cDNA, and rRNA depletion. Globin transcripts were removed using the QIAseq FastSelect-Globin Kit (Qiagen, p/n 334376). The final libraries were pooled and assessed on the TapeStation using the Agilent D1000 ScreenTape (p/n 5067-5582) and Agilent D1000 reagents (p/n 5067-5583) prior to sequencing. The Illumina NovaSeq 6000 platform was used to perform 150-bp paired-end sequencing.

Results and discussion

Pre-analytical QC

A key factor in achieving accurate NGS results is the quality of the input material. This is particularly important in RNA-seq workflows, as RNA is highly susceptible to degradation due to handling, freeze-thaw cycles, and the presence of RNases. The TapeStation provides robust and reliable QC metrics for total RNA using RNA ScreenTape assays. The system evaluates each sample and assigns an objective RIN^e score on a scale from 1 to 10, where 1 indicates highly degraded RNA, and 10 represents highly intact RNA. This user-independent metric serves as a valuable QC tool for establishing thresholds for downstream applications such as RNA-seq.

Library preparation from RNA extracted from challenging matrices, such as plasma or formalin-fixed, paraffin-embedded (FFPE) tissue, presents additional challenges due to RNA degradation and limited sample volume. Despite these limitations, samples are often irreplaceable and must be sequenced regardless of integrity. Many RNA-seq workflows can be optimized to accommodate such samples.

To address the unique challenges of certain RNA studies, the TapeStation also provides a DV_{200} score, which represents the percentage of RNA fragments longer than 200 nucleotides.[2] The quality score can also be applied to other samples to provide a measure of the fragment size distribution. The DV_{200} metric helps classify degraded RNA based on fragment size and effectively distinguishes samples suitable for NGS from those that are not.

At Decode Health, RNA quality metrics generated by the TapeStation are a critical component of the pre-analytical QC process. To help ensure reliable NGS results, Decode Health has established acceptance criteria that include a RIN^e score of ≥ 6.0 and a DV_{200} of ≥ 70%. Additional pre-analytical QC metrics include specimen weight, RNA concentration, and residual DNA content. Table 1 summarizes the QC metrics tested, acceptance criteria, and the number of samples that failed QC. An example electropherogram of a passing sample is shown in Figure 2. Notably, the failure rate for the RNA integrity scores was only 4.0% for the RIN^e and 1.8% for the DV_{200}. Only samples that meet all predefined criteria are advanced for use in biomarker discovery studies.

Table 1. Pre-analytical QC metrics used by Decode Health following RNA isolation and initial DNase treatment.

Metric	Acceptance Criteria	Number of Samples Assessed	Number Failed (%)
RNA Concentration	≥ 25 ng/µL	297	4 (1.4)
RIN^e	≥ 6.0	297	12 (4.0)
DV_{200}	≥ 70%	272	5 (1.8)
Residual DNA Content	≤ 10% of RNA concentration	234	139 (59.4)

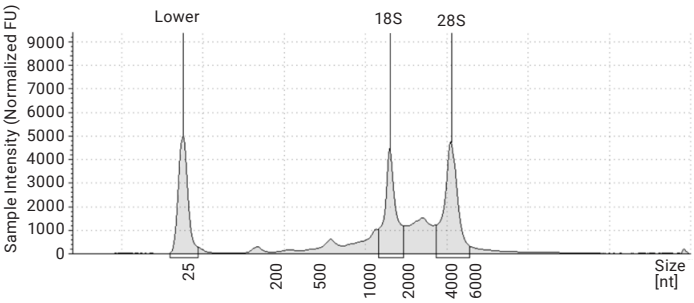

Figure 2. Representative RNA QC results from the Agilent 4200 TapeStation system for a sample with an RIN^e of 8.7.

In addition to RNA QC, the final sequencing libraries also undergo QC using the TapeStation with the D1000 ScreenTape assay. This step is essential for confirming library quality prior to sequencing. A high-quality library is typically indicated by a defined smear or a single peak without side products, with a specific shape and size distribution depending on the library preparation protocol used (Figure 3). The TapeStation helps determine whether the library falls within the expected size range, and confirms the absence of unwanted primer and dimer fragments, both of which are critical for sequencing success.

By identifying samples with insufficient RNA concentration or compromised integrity early in the workflow, QC measures help ensure that only suitable samples proceed to sequencing. The pre-analytical QC workflow improves efficiency, reduces resource expenditure, and prevents unreliable data from affecting downstream analyses or interpretation.

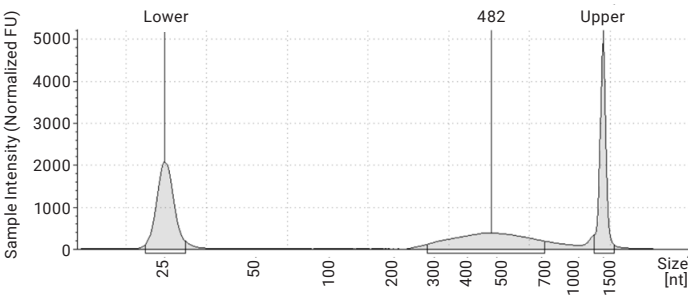

Figure 3. Representative final library QC results from the Agilent 4200 TapeStation system.

Workflow optimization

Insights from pre-analytical QC allowed for optimization of the Decode Health RNA-seq library preparation workflow. One example is the implementation of a second DNase treatment step. Initially, a single DNase treatment resulted in a failure rate of approximately 60%, primarily due to elevated genomic DNA (gDNA) contamination (Table 1). By introducing a second DNase treatment before library preparation, gDNA concentrations were reduced by over 90%, and all samples subsequently met the required QC thresholds.

While effective at reducing gDNA, DNase treatment can also impact RNA yield and integrity. To assess these effects, the TapeStation was used to monitor RNA integrity after DNase treatment. Although RIN^e scores significantly decreased after secondary treatment, the majority of samples remained within acceptable criteria, with only 3.6% falling below the RIN^e threshold (Figure 4a). Importantly, plotting RIN^e scores before and after treatment revealed no correlation (Figure 4b), indicating that initial RNA quality did not predict post-treatment integrity. Despite the modest impact on RNA integrity, most samples remained suitable for RNA-seq, supporting the inclusion of the second DNase treatment in the workflow. This example highlights the value of integrating the TapeStation at multiple points in the workflow, not only for screening input material but also to optimize workflow processes. With insights gained from pre-analytical QC, Decode Health enhanced the robustness and reproducibility of its RNA-seq pipeline.

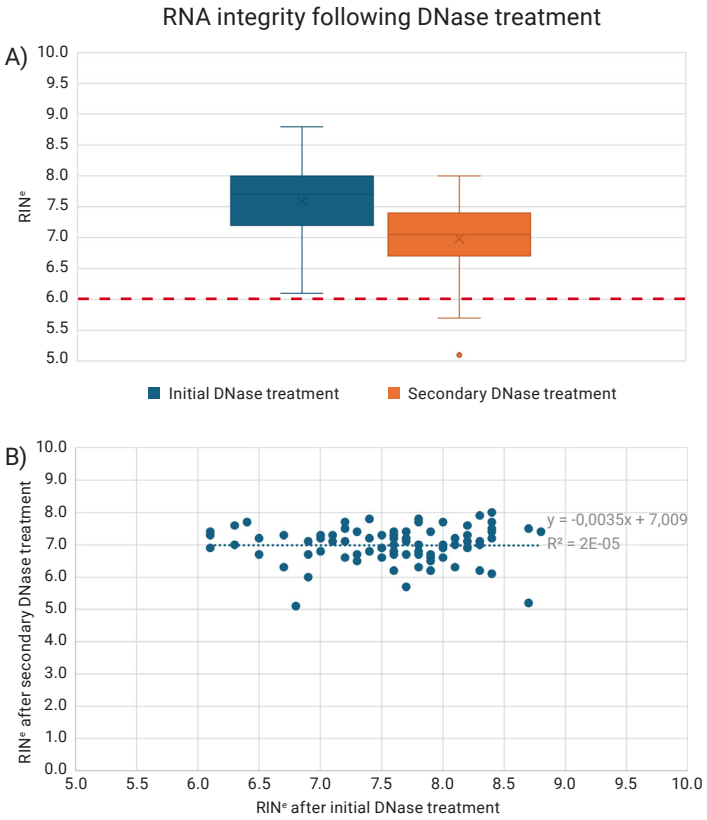

Figure 4. The RNA Integrity Number (RIN^e) for RNA samples was measured before and after secondary DNase treatment. A) Box and whisker plot of RIN^e values before and after secondary DNase treatment. The dotted line indicates passing criteria for the RIN^e score. B) Correlation plot of RIN^e values measured before and after secondary DNase treatment. n = 84. The p-value is < .00001. The result is significant at $p < 0.05$

Correlation of pre-analytical QC to sequencing results

To evaluate the predictive value of pre-analytical QC on sequencing outcomes, particularly when considering low-input or challenging sample types, Decode Health processed 297 RNA samples through its RNA-seq pipeline. Of these, 209 samples passed all pre-analytical QC criteria and were advanced to sequencing. Notably, 92.3% of the samples that passed pre-analytical QC also met all analytical QC metrics, indicating that pre-analytical QC is a strong predictor of sequencing success.

To further explore the relationship between pre-analytical and analytical QC metrics, Decode Health examined correlations between RNA integrity metrics— RIN^e and DV_{200}—and sequencing performance. While both metrics showed statistically significant correlations with analytical QC outcomes, such as mapping rate, multimapped reads, and intergenic reads, the strength of these correlations was modest (Figure 5). RIN^e demonstrated a stronger association with sequencing quality than DV_{200}.

However, the data indicate that analytical QC alone cannot reliably infer pre-analytical sample quality, highlighting the importance of maintaining rigorous pre-analytical QC, even when analytical QC metrics are available.

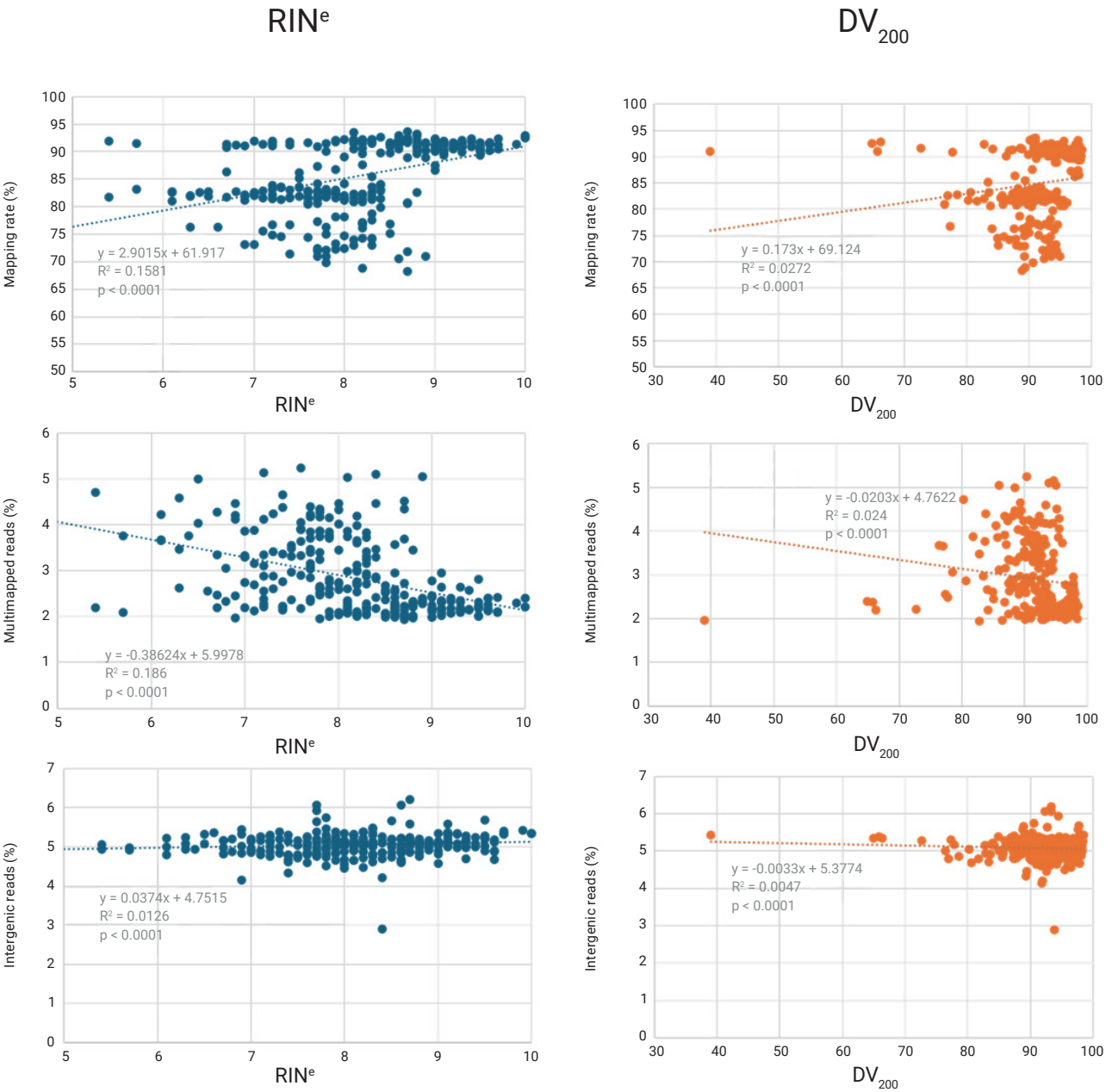

Figure 5. Correlation plots of pre-analytical QC measurements from the Agilent 4200 TapeStation system and analytical sequencing metrics. Statistical analysis of each plot is significant at $p < 0.05$. RIN^e n = 275; DV_{200} n = 250.

Conclusion

RNA sequencing (RNA-seq) has become a fundamental tool in basic biological research, allowing comprehensive analysis of gene expression across various conditions. With great utility for biomarker discovery, especially in exploring molecular mechanisms behind cellular processes and environmental responses, accurate results and interpretation depend on implementing robust QC metrics.

Using the TapeStation for pre-analytical QC provides insights into RNA integrity, final library size, and quality. This ensures that only high-quality input materials move forward in the workflow, improving data consistency and reproducibility. This application note highlights the use of the Agilent 4200 TapeStation system for QC within an RNA-seq pipeline, supporting robust biomarker discovery for clinical diagnostic applications.

References

1. Sesler, C. L.; Shaginurova, G. I.; Wylezinski, L. S.; Grigorenko, E. V.; Cockerill, F. R.; Spurlock, C. F. Development of an End-to-End Total RNA Sequencing Quality Control Framework for Blood-Based Biomarker Discovery. *J. Mol. Diagn.*, **2025**. DOI:10.1016/j.jmoldx.2025.05.008.

2. DV_{200} Evaluation with RNA ScreenTape Assays. *Agilent Technologies technical overview*, publication number 5991-8355EN, **2017**.

www.agilent.com/genomics/tapestation

PR7004-1201

Published in the USA, February 18, 2026
5994-8973EN

Agilent
Trusted Answers