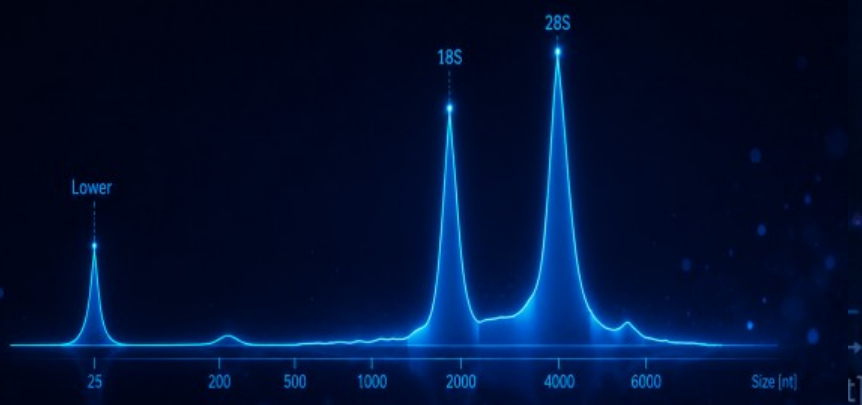


RNA-Seq 성공의 시작

Input RNA Quality Assessment와
Workflow Selection



RNA-Seq 를 위한 RNA Quality Assessment

Input RNA 품질 평가부터 Library Preparation 선택까지

RNA sequencing(RNA-Seq)는 유전자 발현을 정량적으로 분석하는 대표적인 transcriptome 분석 기술로, 기초 연구부터 질환 연구까지 다양한 분야에서 널리 활용되고 있다. 최근에는 Bulk RNA-Seq 뿐만 아니라 Single Cell RNA-Seq, Spatial Transcriptomics 등 다양한 분석 기술이 확장되면서 연구 목적과 시료 특성에 적합한 RNA-Seq workflow 를 선택하는 것이 더욱 중요해지고 있다.¹

RNA-Seq 결과의 품질은 sequencing 장비나 분석 방법뿐 아니라 Input RNA 의 품질에 크게 영향을 받는다. RNA 가 분해되어 있거나, 불순물 또는 genomic DNA(gDNA)가 잔존한 경우에는 library preparation 효율이 저하될 수 있으며, sequencing data 의 정확성과 재현성에도 영향을 줄 수 있다.²⁻⁴ 따라서 RNA-Seq 를 수행하기 전에 RNA 의 Integrity, Purity, Quantity 및 gDNA contamination 을 종합적으로 평가하는 과정이 필요하다.

또한 모든 RNA-Seq workflow 가 동일한 RNA 품질을 요구하는 것은 아니다. Library preparation 방법에 따라 요구되는 RNA 품질과 적용 가능한 시료가 달라지므로, RNA 품질과 연구 목적을 함께 고려하여 적절한 workflow 를 선택하는 것이 중요하다. 본 Tech Note 에서는 RNA Quality Assessment 의 핵심 요소와 주요 Library preparation 방법별 RNA 품질 요구사항을 살펴보고, 연구 목적과 시료 특성에 적합한 RNA-Seq workflow 선택에 도움이 되는 정보를 소개하고자 한다.⁴⁻⁸

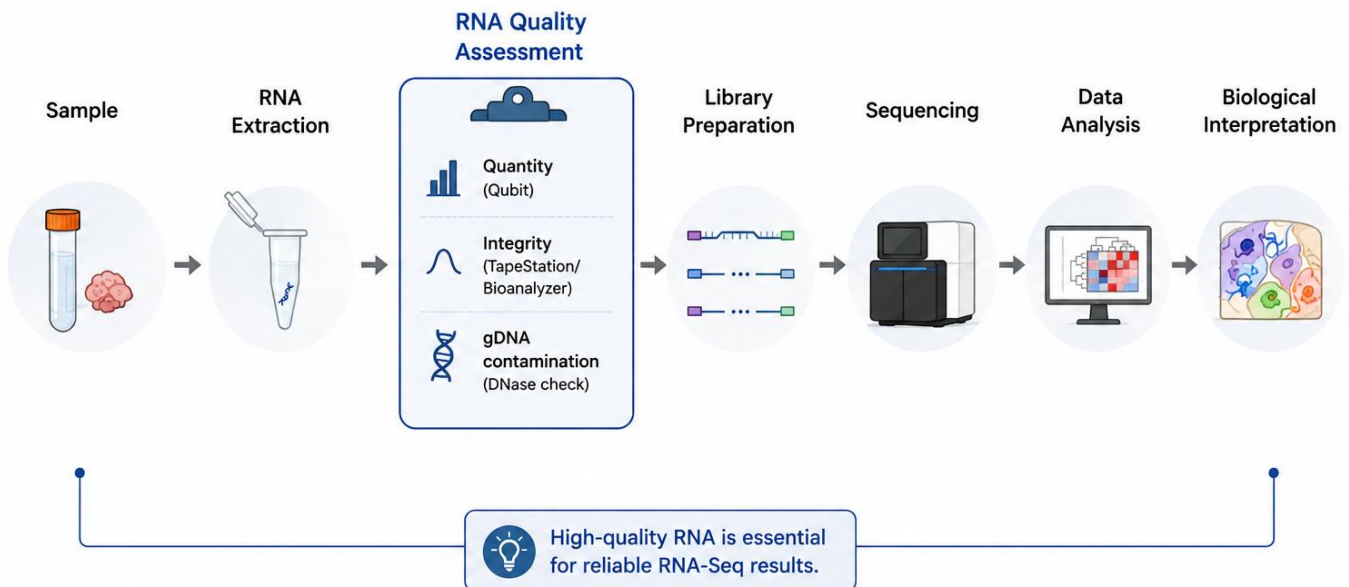


Figure 1. Overview of RNA Quality Assessment and RNA-Seq workflow.

RNA-Seq 에서 Input RNA 가 중요한 이유

RNA-Seq 는 시료에 존재하는 RNA 를 sequencing 하여 유전자 발현을 정량화하고 transcriptome 을 분석하는 기술이다. 따라서 최종 데이터의 품질은 sequencing 장비나 bioinformatics 분석뿐 아니라, **실험의 시작 단계인 Input RNA 의 품질에 크게 영향을 받는다.¹

RNA 는 DNA 보다 화학적으로 불안정하여 RNase 에 의해 쉽게 분해될 수 있으며, 조직 채취부터 RNA 추출, 보관 및 운송 과정에서도 품질이 저하될 수 있다. RNA 가 심하게 분해되면 library preparation 효율이 감소하고 transcript 전체가 균일하게 분석되지 않아 gene expression 정량, transcript isoform 및 alternative splicing 분석에도 영향을 줄 수 있다.^{2, 3}

하지만 RNA-Seq 에서 중요한 것은 고품질 RNA 를 확보하는 것만이 아니다. 최근에는 다양한 Library preparation 방법이 개발되면서 RNA 품질에 따라 적용 가능한 workflow 도 달라지고 있다. 따라서 RNA Quality Assessment 는 단순히 실험 가능 여부를 판단하는 과정이 아니라, 시료 특성에 적합한 Library preparation 방법을 선택하기 위한 과정이라고 볼 수 있다.⁴⁻⁸

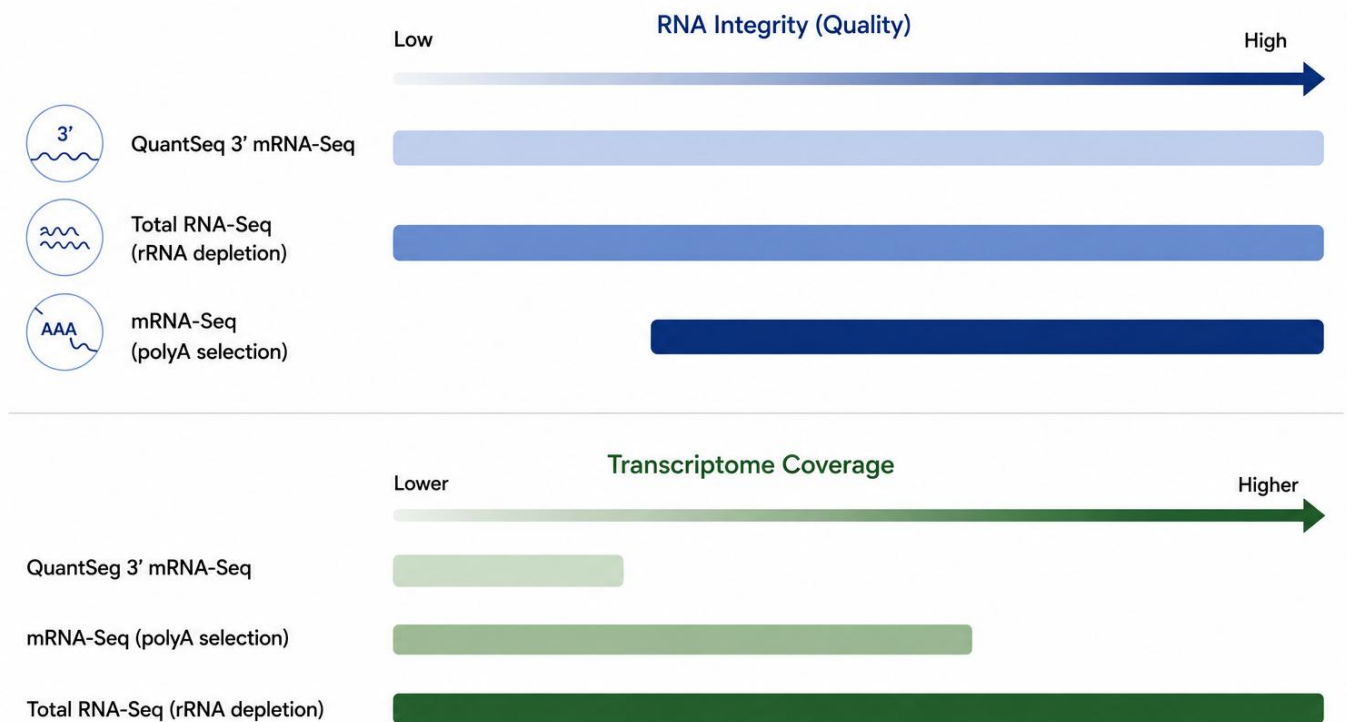


Figure 2. Selection of RNA-Seq workflows based on RNA integrity and transcriptome coverage.

RNA Quality Assessment

RNA Quality Assessment 는 RNA-Seq 실험에서 가장 기본적인 Quality Control(QC) 과정으로, Library preparation 및 sequencing 결과의 영향을 미치는 중요한 단계이다. 일반적으로 RNA 품질은 Purity, Quantity 및 Integrity 를 중심으로 평가하며, 최근에는 genomic DNA(gDNA) contamination 여부와 시료 특성까지 함께 고려하여 적절한 Library preparation workflow 를 선택하는 것이 권장된다.⁴⁻⁸

Purity 는 단백질, phenol, guanidine salt 등 불순물의 잔존 여부를 평가하는 지표로, 일반적으로 A260/A280 과 A260/A230 ratio 를 이용하여 확인한다. 일반적으로 두 ratio 모두 약 2.0 을 권장하지만, Purity 는 RNA Integrity 를 직접 반영하지 않으며 RNA-Seq workflow 를 결정하는 주요 기준은 아니다. 따라서 Purity 는 보조적인 QC 지표로 활용되며, RNA Integrity 및 gDNA contamination 과 함께 종합적으로 고려하는 것이 중요하다.^{2, 5}

RNA Quantity

RNA Quantity 는 Library preparation 에 필요한 RNA 양이 충분한지 확인하기 위한 품질 평가 항목이다. 대부분의 RNA-Seq Library preparation kit 는 권장 RNA input requirement 를 제시하고 있으며, 충분하지 않은 RNA 를 사용할 경우 library complexity 감소, PCR duplication 증가 및 sequencing bias 가 발생할 가능성이 높아질 수 있다.⁵⁻⁸

Library preparation 에 사용되는 RNA input 은 일반적으로 fluorescence 기반의 Qubit 을 이용하여 측정하는 것이 권장된다. Qubit 은 RNA 특이적 형광 염료를 사용하여 RNA 만을 선택적으로 정량할 수 있어 정확한 RNA 농도 측정이 가능하다. 반면 NanoDrop 은 핵산의 흡광도를 기반으로 RNA 농도를 추정하므로 DNA, free nucleotide 및 기타 오염물질의 영향을 받을 수 있으며, 실제 RNA 양을 과대평가할 가능성이 있다. 따라서 주요 제조사들은 RNA-Seq Library preparation 전 RNA input 은 Qubit 과 같은 fluorescence 기반 방법으로 확인할 것을 권장하고 있다.⁵

RNA Integrity

RNA Integrity 는 RNA 의 분해 정도를 평가하는 대표적인 품질 지표이며, RNA-Seq workflow 를 선택하는 중요한 기준으로 활용된다. 일반적으로 Agilent Bioanalyzer 또는 TapeStation 을 이용하여 RNA Integrity Number(RIN) 또는 RIN equivalent(RINe)를 측정하며, RIN 값은 1 에서 10 사이의 범위를 가지며 값이 높을수록 RNA 의 integrity 가 우수하고 degradation 이 적은 것으로 해석된다.²

일반적으로 Poly(A) enrichment 기반 mRNA-Seq 은 상대적으로 높은 RNA integrity 를 요구하는 반면, rRNA depletion 기반 Total RNA-Seq 및 QuantSeq 3' mRNA-Seq 은 일부 degraded RNA 에도 적용할 수 있도록 설계되어 있어 비교적 낮은 RNA integrity 에서도 library 제작이 가능하다. 다만 권장 기준은 제조사 및 Library preparation 방법에 따라 차이가 있으므로, 연구 목적과 시료 특성을 함께 고려하여 적절한 workflow 를 선택하는 것이 중요하다.⁵⁻⁸

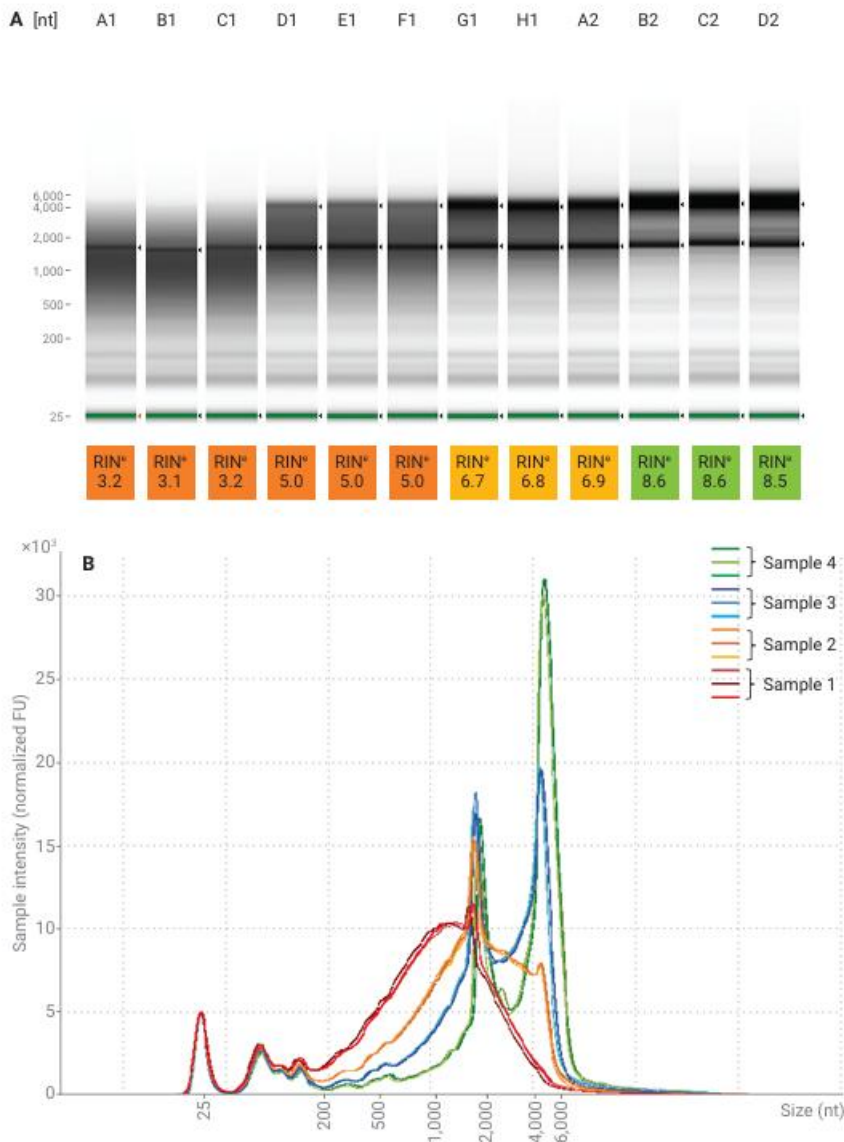


Figure 3. Representative TapeStation virtual gel images and electropherograms of RNA samples with different RINe values.²

RNA Quality Assessment 의 종합적 해석

RNA Quality Assessment 는 단일 지표만으로 RNA-Seq 의 수행 가능 여부를 판단하는 과정이 아니라, Purity, Quantity 및 Integrity 를 종합적으로 평가하여 시료에 적합한 Library preparation workflow 를 선택하는 과정이다. 따라서 RNA 품질 평가는 연구 목적과 시료 특성을 함께 고려하여 해석하는 것이 중요하다.

예를 들어 fresh tissue 에서 추출한 고품질 RNA 는 대부분의 Library preparation workflow 에 적용할 수 있지만, FFPE 조직이나 임상 시료처럼 RNA degradation 이 되었을 확률이 높은 경우에는 rRNA depletion 기반 Total RNA-Seq 또는 QuantSeq 3' mRNA-Seq 과 같이 degraded RNA 에 적합한 workflow 를 선택하는 것이 유리할 수 있다. 또한 RNA Integrity 가 우수하더라도 genomic DNA contamination 이나 화학적 오염이 존재하는 경우에는 sequencing 결과의 품질에 영향을 줄 수 있으므로, 각 품질 지표를 개별적으로 판단하기보다는 전체적인 RNA 품질을 종합적으로 해석하는 것이 바람직하다. 또한 동일한 RNA 품질이라 하더라도 연구 목적과 분석하고자 하는 transcriptome 의 범위에 따라 적합한 Library preparation workflow 는 달라질 수 있다. 다음 장에서는 각 RNA-Seq workflow 의 특징과 적용 가능한 시료 유형에 대해 살펴본다.

Sample Characteristics	Primary QC metric	Recommended library preparation	Recommended use
Intact RNA	RIN	mRNA-Seq Total RNA-Seq QuantSeq 3' mRNA-Seq	Select based on research objectives
Degraded / FFPE RNA	DV200	Total RNA-Seq QuantSeq 3' mRNA-Seq	Suitable for degraded RNA
Low-Input RNA	RNA amount (Qubit)	QuantSeq 3' mRNA-Seq	Suitable when RNA quantity is limited

Table 1. RNA Quality-Based Library preparation Guide

Library Preparation 방법에 따른 RNA Quality Requirement

RNA Quality Assessment 의 목적은 단순히 시료의 품질을 평가하는 데 그치지 않는다. 평가 결과를 바탕으로 RNA 의 상태와 연구 목적에 적합한 RNA-Seq library preparation workflow 를 선택하는 것이 더욱 중요하다. 동일한 RNA 시료라도 선택하는 library preparation 방법에 따라 요구되는 RNA 품질 기준과 생성되는 데이터의 특성이 달라질 수 있기 때문이다.

RNA-Seq 에서 널리 사용되는 library preparation workflow 는 Poly(A) enrichment, rRNA depletion, 3' mRNA Sequencing(Quantseq 3' mRNA-Seq)이며, 각 workflow 는 RNA 를 선택하는 원리와 분석 대상이 서로 다르다. 따라서 주요 제조사들은 RNA integrity 뿐만 아니라 연구 목적과 시료의 특성을 함께 고려하여 적절한 workflow 를 선택할 것을 권장하고 있다.⁵⁻⁸

왜 Library Preparation 마다 RNA 품질 기준이 다른가?

RNA-Seq library preparation 은 RNA 를 sequencing 가능한 library 로 전환하는 과정이다. 그러나 RNA 를 선택(capture)하고 cDNA 를 합성하는 방식은 workflow 마다 서로 다르며, 이러한 차이로 인해 동일한 RNA 시료라도 RNA degradation 의 영향을 받는 정도와 최종적으로 얻을 수 있는 데이터의 특성이 달라질 수 있다.

예를 들어, Poly(A) enrichment 는 mRNA 의 poly(A) tail 을 이용하여 target RNA 를 선택하므로 RNA degradation 이 심할 경우 poly(A) tail 이 유지된 transcript 의 회수 효율이 감소하여 library 제작 효율이 저하될 수 있다. 반면 rRNA depletion 은 ribosomal RNA 만 제거하고 나머지 RNA 를 분석하는 방식이므로 RNA degradation 의 영향을 상대적으로 덜 받아 일부 분해된 RNA 에서도 안정적으로 library 를 제작할 수 있다. QuantSeq 3' mRNA-Seq 는 transcript 의 3' 말단을 선택적으로 분석하는 방법으로, RNA 가 부분적으로 분해된 시료에서도 상대적으로 안정적인 유전자 발현 분석이 가능하다.⁵⁻⁸

따라서 RNA Quality Assessment 는 단순히 실험 가능 여부를 판단하기 위한 과정이 아니라, RNA 의 상태와 연구 목적에 가장 적합한 library preparation workflow 를 선택하기 위한 과정으로 이해하는 것이 바람직하다. RNA 의 상태에 따른 권장 RNA-Seq library preparation workflow 와 주요 QC 지표는 Table 1 에 요약하였다.

Poly(A) Enrichment

Poly(A) enrichment 는 진핵생물 mRNA 의 poly(A) tail 을 선택적으로 포획하여 protein-coding transcript 를 효율적으로 분석하는 가장 널리 사용되는 RNA-Seq library preparation 방법이다. Ribosomal RNA 가 대부분 제거되므로 sequencing read 를 coding transcript 에 집중할 수 있어 높은 sequencing efficiency 를 제공한다.⁵

이 방법은 RNA integrity 가 우수한 시료에서 가장 우수한 성능을 나타내며, Fresh tissue 나 cultured cell 과 같이 intact RNA 를 사용하는 연구에 적합하다. 대부분의 제조사는 높은 RNA 품질(RIN ≥ 7)을 권장하고 있으며, 고품질 RNA 에서는 균일한 transcript coverage 와 높은 library complexity 를 얻을 수 있다.^{5,6}

반면 RNA degradation 이 진행되면 poly(A) tail 이 유지된 transcript 의 회수 효율이 감소하여 transcript coverage 의 편향이 발생할 수 있으며, gene expression 정량의 정확성에도 영향을 줄 수 있다. 또한 poly(A) tail 이 없는 non-coding RNA 는 분석 대상에서 제외되므로 FFPE 시료나 non-polyadenylated RNA 를 포함한 transcriptome 연구에는 적합하지 않을 수 있다.

FEATURE	DESCRIPTION
PRINCIPLE	Selective enrichment of polyadenylated mRNA using the poly(A) tail.
RECOMMENDED SAMPLE	High-quality RNA (Fresh tissue, cultured cells, RIN ≥ 7 recommended).
RNA DEGRADATION TOLERANCE	Low
DETECTABLE RNA SPECIES	Primarily protein-coding mRNA
MAIN ADVANTAGES	High sequencing efficiency, excellent coding transcript coverage, high library complexity
MAIN LIMITATIONS	Not suitable for degraded RNA or FFPE samples; non-polyadenylated RNA is not captured
TYPICAL APPLICATIONS	Differential gene expression analysis of protein-coding genes

Table 2. Key features of the Poly(A) enrichment workflow.

rRNA Depletion

rRNA depletion은 ribosomal RNA를 선택적으로 제거한 후 남아 있는 RNA를 모두 library에 포함시키는 RNA-Seq library preparation 방법이다. Poly(A) tail의 존재 여부와 관계없이 transcript를 확보할 수 있으므로 protein-coding RNA뿐만 아니라 다양한 non-coding RNA까지 함께 분석할 수 있다.^{6,7}

RNA degradation의 영향을 Poly(A) enrichment보다 상대적으로 적게 받기 때문에 FFPE 조직, 임상 시료, archived sample 등 RNA integrity가 낮은 시료에서도 안정적으로 적용할 수 있다. 또한 long non-coding RNA(lncRNA), circular RNA(circRNA) 등 poly(A) tail이 없는 RNA를 포함한 transcriptome 분석에 적합하다.

다만 분석 대상 RNA의 범위가 넓기 때문에 동일한 sequencing depth에서도 protein-coding transcript에 할당되는 read 비율은 Poly(A) enrichment보다 낮아질 수 있다. 따라서 연구 목적에 따라 적절한 sequencing depth를 함께 고려하는 것이 필요하다.⁶

FEATURE	DESCRIPTION
PRINCIPLE	Selective removal of ribosomal RNA while retaining the remaining transcriptome.
RECOMMENDED SAMPLE	High-quality RNA and degraded RNA, including FFPE and archived samples.
RNA DEGRADATION TOLERANCE	High
DETECTABLE RNA SPECIES	Protein-coding RNA and diverse non-coding RNAs (e.g., lncRNA)
MAIN ADVANTAGES	Compatible with degraded RNA; enables comprehensive transcriptome profiling
MAIN LIMITATIONS	Lower proportion of coding reads; generally requires greater sequencing depth than Poly(A) enrichment
TYPICAL APPLICATIONS	Whole-transcriptome analysis, FFPE samples, non-coding RNA research

Table 3. Key features of the rRNA depletion workflow.

3' mRNA Sequencing (QuantSeq 3'mRNA-Seq)

QuantSeq 3' mRNA-Seq는 transcript의 3' 말단 인근에서 하나의 read를 생성하도록 설계된 RNA-Seq library preparation 방법으로, 효율적인 gene expression quantification을 목적으로 개발되었다. Transcript당 하나의 read를 생성하는 것을 기본 원리로 하기 때문에 비교적 낮은 sequencing depth에서도 안정적인 유전자 발현 정량이 가능하다.⁸

Transcript 전체를 sequencing하지 않고 3' 말단만 분석하므로 RNA degradation의 영향을 상대적으로 적게 받으며, 일부 degraded RNA나 low-input RNA에서도 적용할 수 있다. 또한 많은 샘플을 동시에 분석하거나 비용 효율적인 differential gene expression 연구를 수행할 때 유용한 대안으로 활용된다.

반면 transcript 전체의 구조 정보를 제공하지 않기 때문에 transcript isoform 분석, alternative splicing 분석, fusion gene 탐색과 같은 연구에는 적합하지 않다. 따라서 QuantSeq 3' mRNA-Seq 는 gene expression profiling 을 주된 목적으로 하는 연구에서 가장 효과적으로 활용될 수 있다.⁸

FEATURE	DESCRIPTION
PRINCIPLE	Generation of one read from the 3' end of each transcript for gene expression quantification.
RECOMMENDED SAMPLE	High-quality RNA, partially degraded RNA, and low-Input RNA.
RNA DEGRADATION TOLERANCE	Moderate to high
DETECTABLE RNA SPECIES	Polyadenylated mRNA (3' end only)
MAIN ADVANTAGES	Cost-effective gene expression profiling; suitable for low-Input and partially degraded RNA; reduced sequencing requirements
MAIN LIMITATIONS	Not suitable for transcript isoform analysis, alternative splicing analysis, or fusion gene detection
TYPICAL APPLICATIONS	Differential gene expression analysis and large-scale expression profiling

Table 4. Key features of the QuantSeq 3' mRNA-Seq workflow.

Genomic DNA Contamination 이 RNA-Seq 결과에 미치는 영향

RNA Integrity 와 RNA Purity 는 RNA 품질 평가에서 가장 널리 활용되는 지표이지만, 이러한 지표만으로 RNA-Seq 에 적합한 RNA 품질을 완전히 평가하기는 어렵다. RNA 추출 과정에서 함께 분리된 genomic DNA(gDNA)가 충분히 제거되지 않은 경우, library preparation 과정뿐 아니라 sequencing 결과와 downstream analysis 에도 영향을 줄 수 있기 때문이다.^{4, 9, 10}

Residual genomic DNA 는 대부분의 RNA 추출 과정에서 발생할 수 있으며, 조직의 종류, 세포 수, 추출 방법 및 시료 특성에 따라 그 정도가 달라질 수 있다. 이러한 이유로 주요 RNA-Seq library 제조사들은 RNA Integrity 와 함께 genomic DNA 제거 여부 역시 중요한 품질 관리 요소로 제시하고 있으며, RNA 추출 이후 적절한 DNase treatment 를 권장하고 있다.⁴⁻⁸

왜 Genomic DNA Contamination 이 중요한가?

RNA 추출 과정에서는 세포가 파쇄되면서 RNA 와 함께 genomic DNA 가 동시에 방출된다. 일반적인 RNA purification 과정을 거치더라도 모든 genomic DNA 가 완전히 제거되는 것은 아니며, 일부 DNA 는 최종 RNA 시료에 잔존할 수 있다. 특히 조직 시료나 핵산 함량이 높은 시료에서는 이러한 현상이 상대적으로 빈번하게 발생할 수 있다.^{2, 5}

RNA-Seq library 는 RNA 를 cDNA 로 역전사(reverse transcription)한 후 증폭하여 제작되지만, 제거되지 않은 genomic DNA 역시 library preparation 과정에 함께 포함될 수 있다. 이 경우 일부 sequencing read 가 실제 transcript 가 아닌 genomic DNA 에서 유래하게 되며, 결과적으로 RNA-Seq 데이터의 정확성과 해석에 영향을 줄 가능성이 있다.^{4, 9}

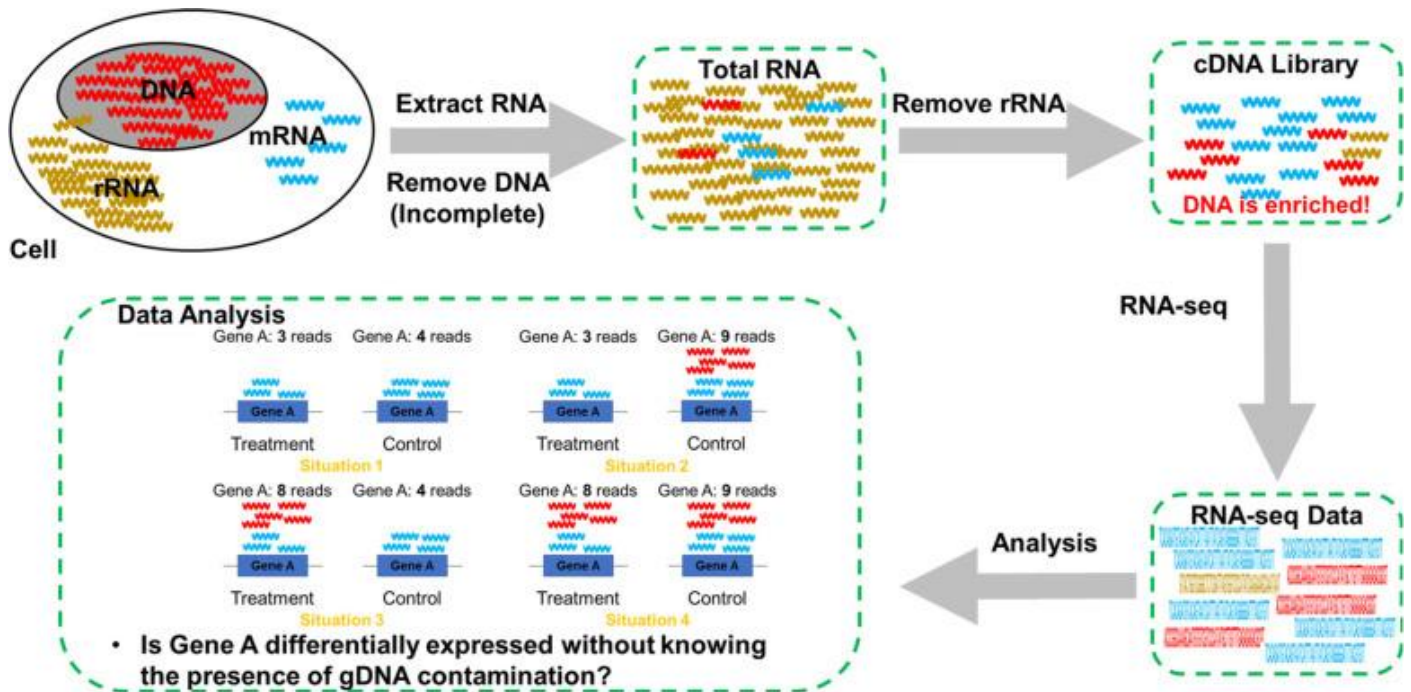


Figure 4. Overview of genomic DNA contamination during RNA-Seq library preparation and its potential impact on downstream analysis. Adapted from Ref.¹⁰

Genomic DNA Contamination 이 RNA-Seq 결과에 미치는 영향

Residual genomic DNA 는 RNA-Seq 데이터의 여러 단계에서 다양한 형태의 bias 를 유발할 수 있다.

가장 대표적인 영향은 intronic 및 intergenic read 의 증가이다. 일반적으로 RNA 에서 유래한 sequencing read 는 exon 에 주로 mapping 되지만, genomic DNA 가 함께 sequencing 될 경우 intron 또는 intergenic region 에 mapping 되는 read 의 비율이 증가할 수 있다. 이러한 변화는 실제 transcriptome 을 충분히 반영하지 못하는 sequencing 결과로 이어질 가능성이 있다.^{9,10}

또한 genomic DNA 의 잔존은 gene expression quantification 에도 영향을 줄 수 있다. 실제 발현되지 않은 genomic DNA sequence 가 함께 분석되면서 특정 유전자의 발현량이 과대 또는 과소 평가될 가능성이 있으며, 특히 저발현(low-abundance) 유전자에서는 이러한 영향이 더욱 크게 나타날 수 있다. 최근에는 이러한 contamination 이 differential gene expression 분석의 정확도를 저하시킬 수 있다는 보고도 발표되고 있다.^{9,10}

일부 제조사에서는 genomic DNA contamination 이 stranded RNA-Seq 에서 strand specificity 를 감소시키거나 background signal 을 증가시킬 수 있음을 함께 언급하고 있다.⁴

제조사는 왜 DNase Treatment 를 권장하는가?

대부분의 RNA-Seq library 제조사는 RNA 추출 후 genomic DNA 를 충분히 제거한 RNA 를 사용하는 것을 권장하며, 일부 workflow 에서는 DNase treatment 를 표준 workflow 의 일부로 포함하고 있다.⁴⁻⁸ 이는 단순히 RNA 의 순도를 높이기 위한 목적이 아니라, genomic DNA 유래 sequencing read 를 최소화하여 sequencing 결과의 정확성과 재현성을 확보하기 위한 것이다.

RNA extraction 이후 적절한 DNase treatment 와 clean-up 과정을 수행하면 residual genomic DNA 를 효과적으로 감소시킬 수 있으며, 이를 통해 library quality 와 downstream analysis 의 신뢰성을 향상시킬 수 있다. 따라서 DNase treatment 는 library preparation 성공률을 높이기 위한 과정이라기 보다, RNA-Seq 데이터가 실제 RNA 발현을 보다 정확하게 반영하도록 하기 위한 품질 관리(Quality Control)의 과정으로 이해하는 것이 적절하다.

RNA-Seq 데이터만으로 Genomic DNA Contamination 을 확인할 수 있을까?

RNA-Seq 분석 결과에서 intronic 또는 intergenic read 의 비율이 증가한 경우 genomic DNA contamination 가능성을 의심할 수 있으며, 최근에는 이러한 특징을 이용하여 contamination 여부를 추정하거나 보정하기 위한 다양한 분석 방법과 software 가 제안되고 있다.^{9,10}

그러나 이러한 지표만으로 genomic DNA contamination 을 확정하기에는 한계가 있다. Intronic read 의 증가는 미성숙 RNA(pre-mRNA), 조직 특이적인 전사 활성 또는 reference annotation 의 특성에 의해서도 나타날 수 있으며, intergenic read 역시 sequencing 특성이나 분석 조건의 영향을 받을 수 있기 때문이다.

따라서 RNA-Seq 데이터는 genomic DNA contamination 의 가능성을 평가하는 참고 자료로 활용될 수는 있지만, 이를 단독으로 contamination 의 존재 여부를 판단하는 근거로 사용하기에는 제한이 있다. 신뢰성 있는 품질 관리를 위해서는 RNA 추출 단계에서의 적절한 DNase treatment 와 함께 RNA Quality 및 Library QC 결과를 종합적으로 해석하는 것이 바람직하다.

Summary

RNA-Seq 의 성공은 sequencing 단계에서 시작되는 것이 아니라, 고품질의 Input RNA 를 확보하는 단계에서부터 시작된다. RNA 품질은 library preparation 효율과 sequencing data 의 정확성에 직접적인 영향을 미치며, 연구 목적에 적합한 workflow 를 선택하기 위한 중요한 기준이 된다.

RNA Quality Assessment 에서는 RNA Purity, RNA Quantity 및 RNA Integrity 를 함께 평가해야 하며, 각 지표는 서로 다른 품질 정보를 제공한다. 따라서 단일 지표만으로 RNA-Seq 수행 가능 여부를 판단하기보다는, 시료의 특성과 연구 목적을 종합적으로 고려하는 것이 중요하다.

또한 모든 library preparation 방법이 동일한 RNA 품질을 요구하는 것은 아니다. Poly(A) enrichment 는 높은 RNA integrity 를 가진 시료에 적합하며, rRNA depletion 과 QuantSeq 3' mRNA-Seq 는 일부 degraded RNA 에서도 적용할 수 있도록 설계되어 있다. 따라서 RNA Quality Assessment 는 단순한 품질 평가를 넘어, 시료 특성에 적합한 library preparation 전략을 선택하기 위한 과정으로 이해할 필요가 있다.

마지막으로, genomic DNA contamination 은 RNA-Seq 데이터의 정확성과 해석에 영향을 줄 수 있는 중요한 품질 요소이다. 대부분의 제조사가 DNase treatment 를 권장하는 이유 역시 sequencing 결과의 정확성과 재현성을

확보하기 위한 것이며, RNA-Seq 데이터만으로 genomic DNA contamination 을 확정하기에는 한계가 있다는 점도 함께 고려해야 한다.

결론적으로, 성공적인 RNA-Seq 연구를 위해서는 RNA Integrity, RNA Purity, RNA Quantity, genomic DNA contamination 및 연구 목적에 적합한 library preparation 전략을 종합적으로 고려하는 통합적인 품질 관리가 필수적이다.

E-biogen Insight

RNA Integrity(RIN)는 RNA 품질을 평가하는 가장 널리 사용되는 지표이지만, RIN 값만으로 RNA-Seq 의 성공 여부를 판단하기에는 한계가 있다. 실제 RNA-Seq 에서는 RNA Purity, RNA Quantity, genomic DNA contamination, 시료의 특성 및 연구 목적을 함께 고려하여 적절한 library preparation workflow 를 선택하는 것이 중요하다.

E-biogen 에서는 RNA 품질 평가 시 단일 QC 지표에 의존하기보다, 시료의 특성, 연구 목적 및 적용 가능한 library preparation workflow 를 종합적으로 검토하여 실험 가능 여부를 판단하는 것을 권장한다. 이러한 통합적인 접근은 보다 안정적이고 재현성 높은 RNA-Seq 결과를 확보하는 데 중요한 기반이 된다.

References

1. Wang Z, Gerstein M, Snyder M. *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Rev Genet. 2009.
2. Fleige S, Pfaffl MW. *RNA integrity and the effect on quantitative gene expression analysis*. Mol Aspects Med. 2006.
3. Gallego Romero I, Pai AA, Tung J, Gilad Y. *RNA-seq: impact of RNA degradation on transcript quantification*. BMC Biol. 2014.
4. Illumina. *Illumina RNA Prep with Enrichment (L) Tagmentation Reference Guide*.
5. Lexogen. *CORALL RNA-Seq Library Prep Kit User Guide*.
6. New England Biolabs. *NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® Manual*.
7. Promega. *RNA Quality and Integrity Guidelines*.
8. Lexogen. *QuantSeq 3' mRNA-Seq Library Prep User Guide*.
9. Gaidatzis D, et al. *CleanUpRNAseq: Detecting genomic DNA contamination in RNA sequencing datasets*. Bioconductor. 2024.
10. Kim J, et al. *Genes expressed at low levels raise false discovery rates in RNA samples contaminated with genomic DNA*. BMC Genomics. 2022.