

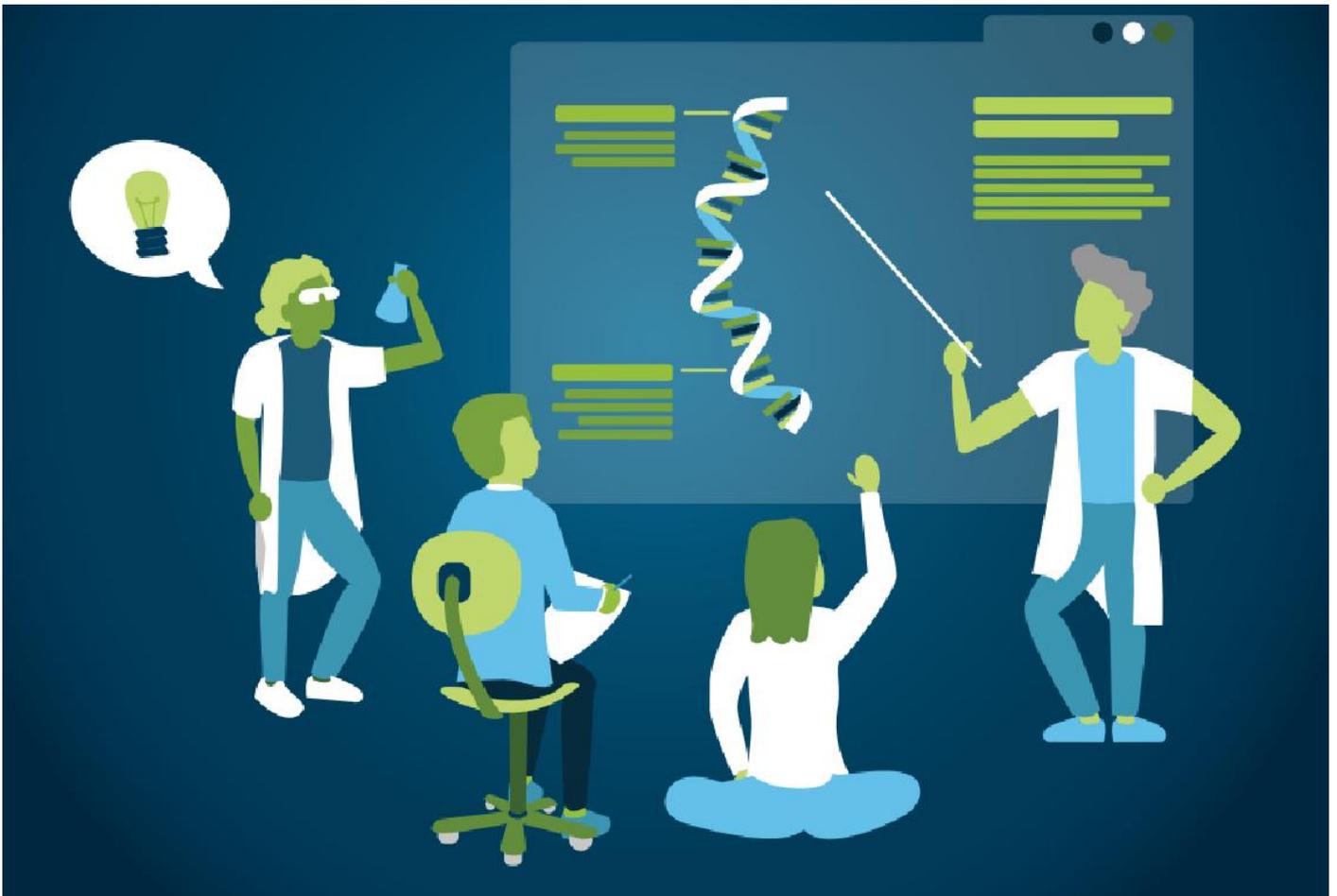
## RNA-Seq의 모든 것: 실험 설계부터 정밀한 전체 전사체 및 3' UTR RNA 분석까지

RNA Sequencing 은 유전자 발현을 정량적으로 분석하고, 새로운 전사체 구조를 규명하며, 3' 말단 특이적인 annotation 까지 수행할 수 있는 고정밀 NGS 기반 핵심 기술이다.

최근에는 품질이 낮은 시료(분해가 진행된 RNA, FFPE 등)에서도 효율적인 분석이 가능한 3' mRNA-Seq (Lexogen QuantSeq)과, 전체 전사체의 구조적 다양성과 isoform 정보를 정밀하게 포착할 수 있는 mRNA-Seq 이 함께 활용되는 추세이다.

이에 따라 연구 목적과 시료 특성에 따라 적합한 RNA-Seq 기술을 선택하는 것이 더욱 중요해지고 있다.

본 기술노트에서는 RNA-Seq 의 핵심 원리와 실험 설계 전략을 소개하고, 최신 분석 도구와 실제 적용 사례를 통해 이바이오젠의 RNA-Seq 서비스를 보다 효과적으로 활용할 수 있는 방안을 제시하고자 한다.



LEXOGEN-RNA LEXICON-eBook[1]

# 1. RNA-Seq 의 분자 생물학

RNA-Seq 은 세포 내 mRNA 의 염기서열 정보를 고속 시퀀싱하여, 유전자의 발현 양상과 전사체 구조를 정량적, 구조적으로 규명하는 차세대 전사체 분석 기술이다. 이 기술의 핵심은 세포 내에서 전사된 mRNA 를 역전사(reverse transcription)를 통해 cDNA 로 전환한 후, NGS(Next-Generation Sequencing) 플랫폼을 활용해-고해상도로 시퀀싱하여, 발현 수준의 정량화는 물론 splice variant, fusion transcript, novel transcript 등의 구조적 특성까지 함께 분석할 수 있다는 점에 있다. RNA-Seq 실험은 일반적으로 다음과 같은 주요 단계로 구성된다.

## 1.1 RNA 품질 관리

- RNA 의 무결성(RIN) 및 분해 정도(DV200)에 따라 RNA-Seq 데이터의 재현성과 신뢰도가 달라 질 수 있다.

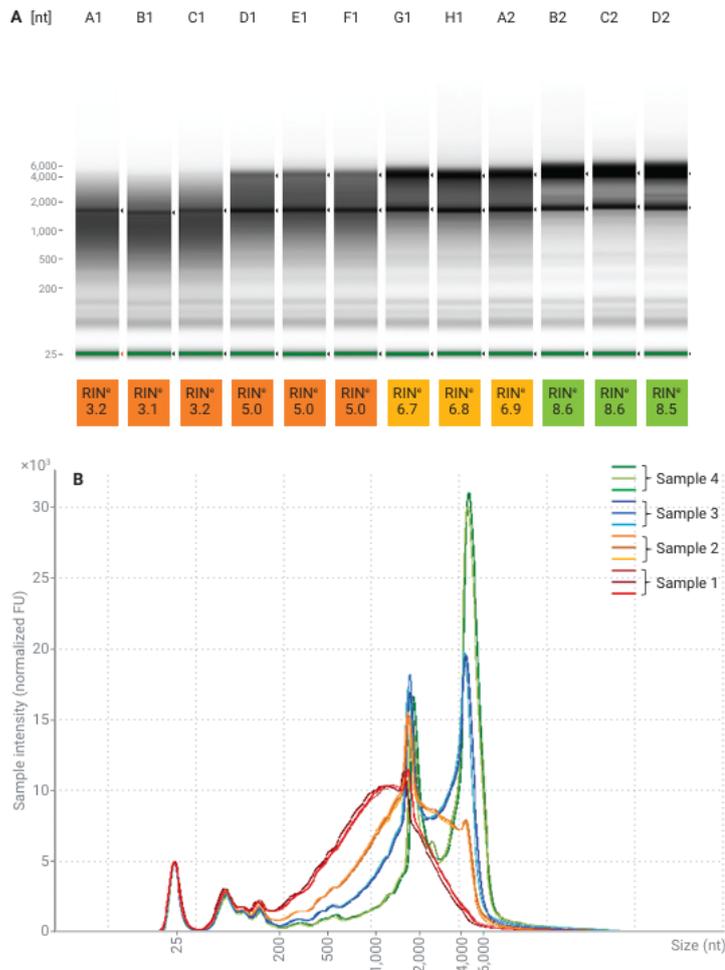
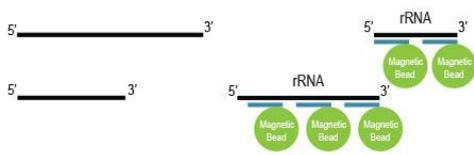


그림 1. Total RNA integrity analysis [2]

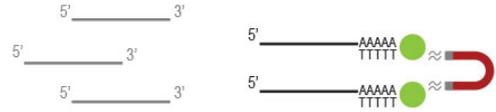
## 1.2 전처리 전략

- Poly(A) Enrichment: Eukaryotic mRNA 에서 polyadenylated transcript 를 선택적으로 포획하여 전사체 전체에 걸친 분석에 적합하다
- rRNA Depletion: ribosomal RNA 를 제거하여 non-coding RNA, lncRNA, partially degraded RNA 까지 포괄한 total RNA 기반 분석에 유용하다.

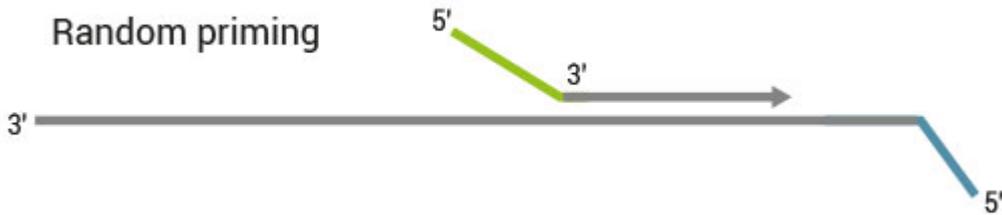
- 분해가 진행된 RNA 샘플이거나 극소량의 샘플은 전처리 과정 없이 3' UTR 영역을 선택적으로 분석하는 QuantSeq 3' mRNA-seq 분석이 적합하다.



A. rRNA Depletion Method



B. Poly(A) Enrichment Method



C. QuantSeq Method(without 전처리)

그림 2. RNA-Seq 전처리 차이점 [3]

### 1.3 cDNA 합성 및 라이브러리 제작

- 1 차 cDNA 합성은 random primer 또는 oligo(dT) primer 로 시작되며, 2 차 합성 후 adapter ligation 과 PCR 증폭이 이어진다.
- Lexogen QuantSeq 는 oligo(dT) priming 을 활용하여 3' 말단에 특이적인 short fragment 만을 수집함으로써 library 구성 및 read-to-count 매핑을 단순화시킨다.

### 1.4 Unique Molecular Identifier (UMI)의 도입

- UMI 도입을 통해, PCR amplification 에 따른 bias 를 줄이는 중복 제거(deduplication)를 수행함으로써 absolute quantification 에 근접한 정량 데이터를 얻을 수 있다.

### 1.5 Sequencing & Platform

- 대부분의 RNA-Seq 은 Illumina NovaSeq 또는 NextSeq 플랫폼을 기반으로 수행되며, 실험 목적에 따라 paired-end 또는 single-end 방식이 선택된다.
- QuantSeq 의 경우 일반적으로 single-end 100bp read 를 권장하며, mRNA-Seq 은 보다 긴 paired-end 150bp 가 권장된다.

각 단계는 실험 목표와 샘플 특성에 따라 최적화가 필요하다.

## 2. 분석 전략: 실험 설계와 라이브러리 선택

RNA-Seq 분석의 성공은 실험 목적에 부합하는 정확한 실험 설계에서 시작된다. 유전자 발현 정량, alternative splicing 탐색, non-coding RNA 발굴, 또는 FFPE 조직 분석 등 다양한 연구 목표에 따라 적절한 라이브러리 구성 전략과 분석 플랫폼이 달라져야 한다.

### 2.1 연구 목적에 따른 접근 전략

- **Differential Expression (DEG) 중심 분석:** 유전자 발현 차이를 빠르고 정량적으로 파악해야 하는 경우, QuantSeq 3' mRNA-Seq 이 비용 및 효율 측면에서 유리하다
- **전사체 구조 분석 (isoform detection, transcript discovery):** exon 간 연결이나, alternative splicing 등을 포괄적으로 확인하려면 mRNA-Seq 기반의 full-length 분석이 필수적이다
- **Low-quality sample 또는 FFPE 샘플 분석:** 분해된 RNA 기반에서는 3' end 기반 분석이 read 정렬효율 및 정량 신뢰도를 높일 수 있다.

### 2.2 플랫폼 비교: QuantSeq vs mRNA-Seq

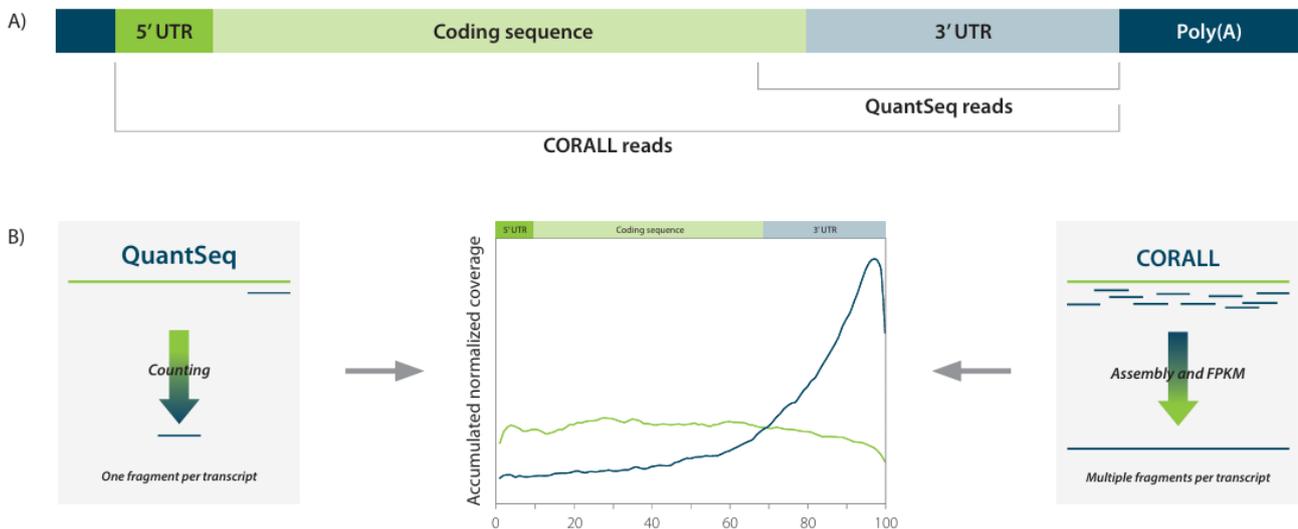


그림 3. Whole Transcriptome vs 3' End Focused: CORALL 과 QuantSeq 의 커버리지 비교 [4]

QuantSeq 와 mRNA-Seq 은 전사체를 바라보는 관점에서 근본적인 차이를 가진다. QuantSeq 는 'read-to-count' 전략, 즉 각 전사체의 3' 말단에서 단 하나의 read 만을 생성하여 정량적 유전자 발현 분석에 최적화되어 있다. 이 방식은 라이브러리 구성의 단순성과 read-to-count 매핑의 직관성을 통해 비용 효율적이고 빠른 DEG 분석이 가능하다는 장점이 있다. 반면, mRNA-Seq 는 'read-to-structure' 전략을 통해 전사체 전반에 걸쳐 다수의 read 를 생성하고 이를 바탕으로 exon-exon junction, alternative splicing, isoform 구조 등을 정밀하게 해석할 수 있는 구조 분석에 강점을 가진다. 따라서 실험 목적이 정량 분석인지, 구조 해석인지에 따라 선택해야 하며, 경우에 따라 두 방법을 병행하는 것도 고려할 수 있다.

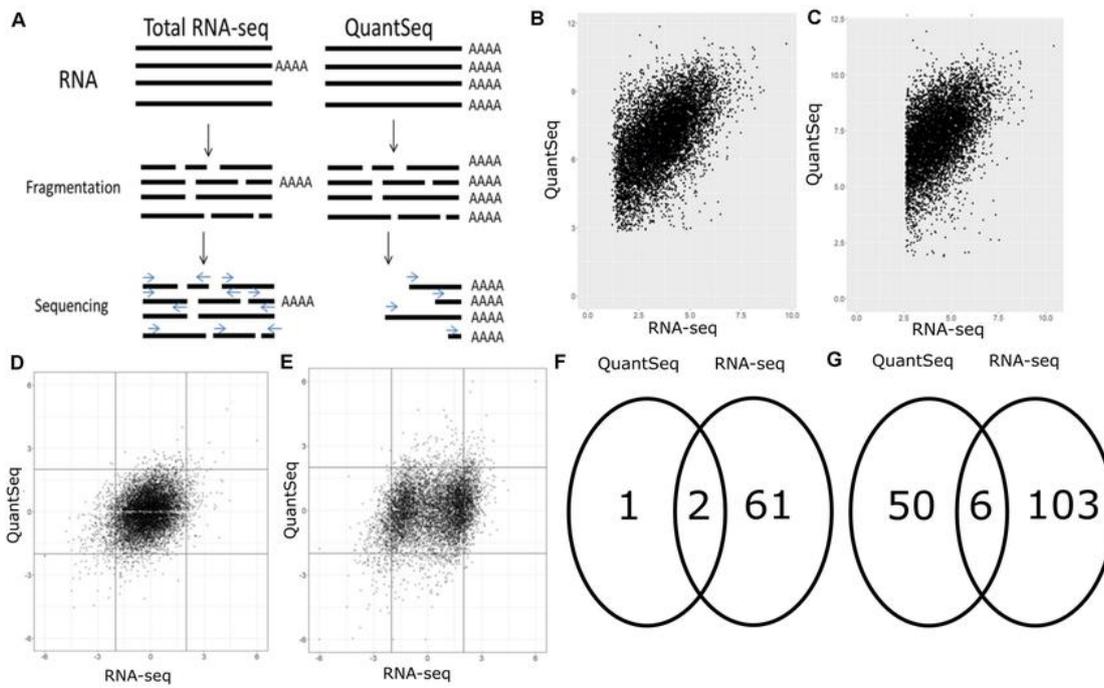


그림 4. Core differences between QuantSeq and RNA-seq [5]

그림 4A는 두 RNA-Seq 방법의 실험 원리를 비교한 그림이다. Total RNA-seq은 전체 mRNA를 fragmentation 후 시퀀싱을 수행하여, 다양한 위치에서 read를 생성한다. 반면, QuantSeq은 fragmentation 없이 3' 말단 영역을 시퀀싱하며, 전사체당 하나의 read만 생성된다. 이처럼 전통적인 RNA-Seq은 전사체 구조 분석에 강점을 가지며, QuantSeq은 정량적 유전자 발현 분석에 최적화되어 있다. 그림 4 B,C는 유전자 발현량의 상관성을 분석한 결과로, 두 플랫폼간 전반적으로 높은 상관관계를 보였다. 그림 4 D, E는 각 유전자에 대한 발현 차이의 통계적 유의성을 비교한 결과이다. 두 플랫폼 간 전체적인 분포 양상은 유사하나, 일부 유전자에서는 유의한 차이가 나타난다. 이는 실험 조건 및 분석 목적에 따라 차등 발현 유전자의 검출 민감도에 차이가 있을 수 있음을 보여준다. 그림 4 F, G는  $p_{adj} < 0.05$  기준에서 각 플랫폼 간 검출된 유의한 유전자(significant gene)의 교집합을 비교한 Venn diagram이다. 일부 조건에서는 각 플랫폼 간 검출되는 유전자 리스트의 차이를 보이는데, 이는 라이브러리 구성 방식의 차이가 영향을 주었을 수 있다. [5]

항목	QuantSeq 3' mRNA-Seq	mRNA-Seq
리드 위치	mRNA의 3' 말단만 타겟	mRNA 전체 길이에서 커버리지
라이브러리 구성	각 전사체당 1개 리드 → 정량 정확도 향상	다양한 위치에서 리드 → 구조 정보(스플라이싱 등) 확보
시퀀싱 depth 및 비용	낮은 depth(1-5M read/sample)로도 실험 가능 → 저비용·다량의 샘플 분석에 유리	높은 depth 필요 → 고비용·상대적으로 증가
분석 효율성	Isoform, splicing 등 구조 정보 추악	전체 구조 분석 가능, 다양한 분석 범위 제공
RNA 품질 민감도	degraded RNA도 분석 가능	고품질 RNA 필요

각 플랫폼의 특징은 위의 표에서 확인 할 수 있다. 성공적인 RNA-Seq 분석을 위해서는 실험 설계 및 분석 목표와 샘플의 상태(조건)을 명확히 설정하는 것이 가장 중요하다. 이바이오젠에서는 실험디자인, 플랫폼 선택, RNA QC, 데이터분석까지 전 과정에 걸쳐 전문적인 RNA-Seq 컨설팅을 제공한다.

### 3. 비교 사례

RNA-Seq 분석은 다양한 생물학적, 의학적 맥락에서 실제로 활용되고 있으며, 특히 Lexogen 기반의 mRNA-Seq 및 QuantSeq 3' mRNA-Seq 은 고품질 전사체 데이터 확보에 효과적인 전략으로 입증되고 있다.

#### FFPE 샘플에서의 유전자 발현 정량 분석 – QuantSeq 기반

2021 년에 BMC Geomics 에 발표된 "Application of the 3' mRNA-Seq using unique molecular identifiers in highly degraded RNA derived from formalin-fixed, paraffine-embedded tissue" 논문에서는 DV200 값이 30% 미만인 FFPE RNA 시료를 대상으로 Lexogen QuantSeq 3'mRNA-seq 키트를 적용하고, 이를 TruSeq mRNA-seq 및 RNA Exome Capture 방법과 비교 분석했다[6]

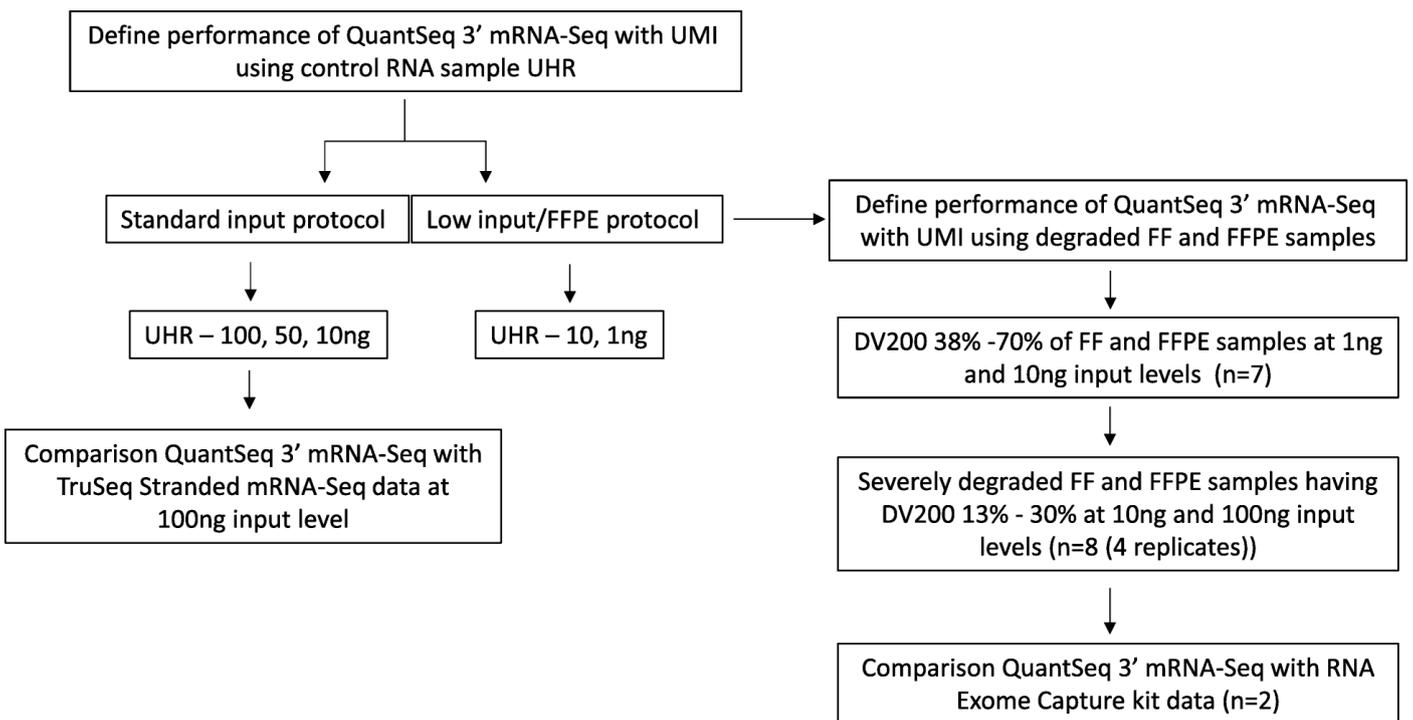


그림 5.The overall experimental design [6]

Lexogen 의 QuantSeq 3'mRNA-Seq 과 illumina TruSeq Stranded mRNA-Seq 을 비교한 결과, 서로 다른 플랫폼간 전체 발현 패턴은 유사(R=0.78)하였으나 부분적인 차이도 관찰되었다. 매핑된 리드의 분포를 살펴보았을 때, Exonic 영역에는 TruSeq 이 83%, QuantSeq 이 65%로 TruSeq 이 더 많이 매핑되었음을 확인할 수 있다. 반면, Intronic, Intergenic 영역은 QuantSeq 이 더 많은 비율로 매핑되었는데 이는 TruSeq 이 mRNA 캡처하는 방식으로 진행했기 때문에 exonic 중심으로 매핑되고 QuantSeq 은 3' UTR 영역을 중심으로 캡처하여 진행하기 때문에

비정형 RNA 영역의 포함 비율이 높다. 검출된 유전자 수는 QuantSeq 에서 22,304 개, TruSeq 에서 21,319 개의 유전자가 검출되었으며 이중 17,003 개 유전자는 두 플랫폼에서 공통적으로 검출되었다. 두 플랫폼 모두 유사한 유전자 수를 검출하며 공통적으로 검출된 유전자도 상당히 많은 것을 확인할 수 있다. QuantSeq 이 3' 말단 기반으로 제한적인 방식임에도 불구하고 TruSeq 과 유사한 수준의 커버리지를 확보할 수 있음을 보여준다. 또한, 검출된 유전자 중 단백질 코딩 유전자 비율은 QuantSeq 이 71%, TruSeq 이 77%로 두 플랫폼 모두 유사한 수준으로 검출되는 것을 확인할 수 있다. [6]

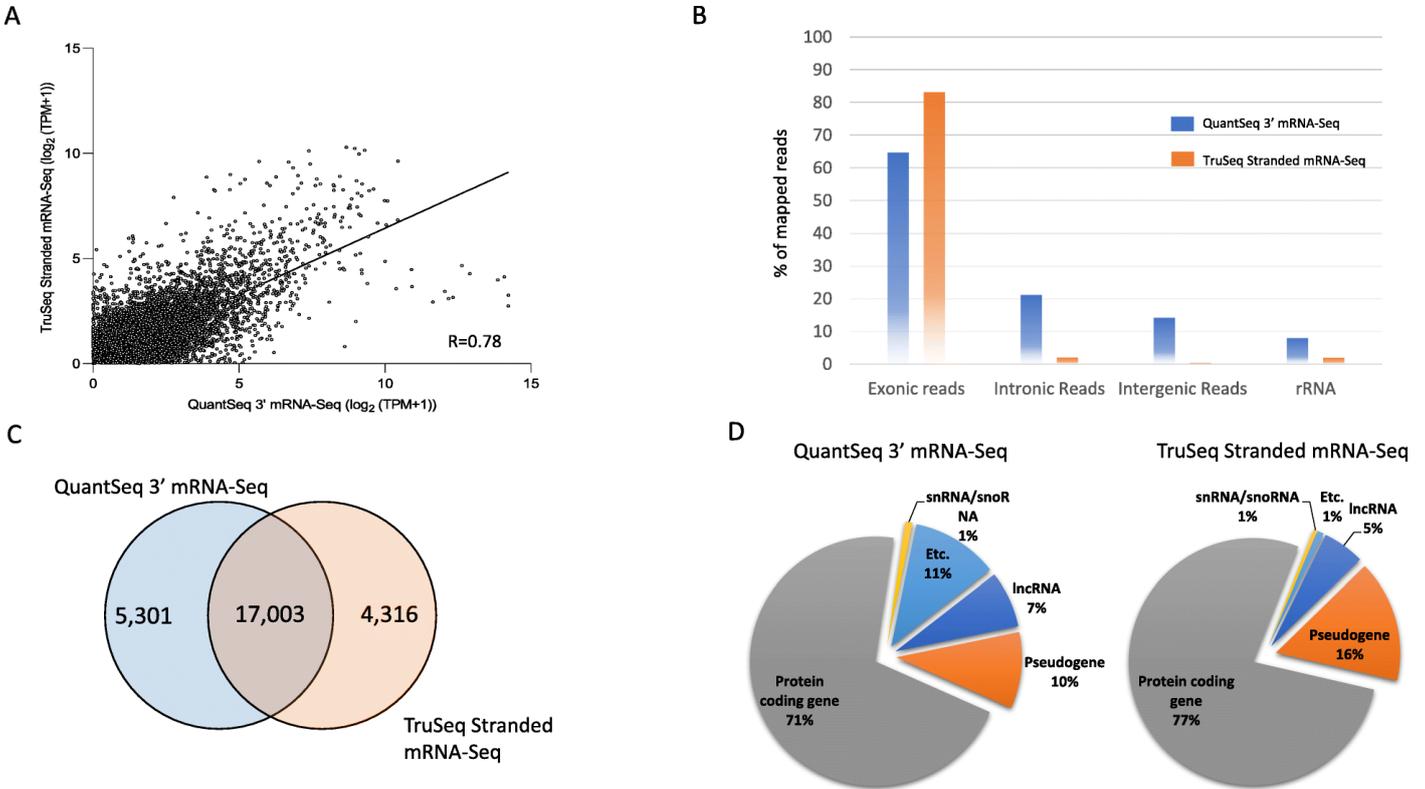


그림 6. Data comparison between the QuantSeq 3' mRNA-Seq and TruSeq Stranded mRNA-Seq kit [6]

#### 4. Ebiogen's RNA-Seq Service

RNA-Seq 은 단순한 유전자 발현 정량을 넘어서, 전사체의 구조적 다양성과 기능적 해석까지 아우를 수 있는 핵심 기술로 자리잡고 있다. 특히 mRNA-Seq 과 QuantSeq 3' mRNA-Seq 은 연구 목적과 시료 조건에 따라 상호 보완적으로 활용될 수 있는 두 축이라 할 수 있다.

mRNA-Seq 은 전사체 구조, 스플라이싱, isoform 수준의 분석에 적합하며, 고품질 RNA 가 확보된 상황에서 강력한 해석 능력을 제공한다. 반면, QuantSeq 는 적은 read 로도 안정적인 발현 정량이 가능하고, FFPE 와 같은 분해된 RNA 에서도 높은 효율을 보이며 대용량 샘플 처리나 DEG 중심 분석에 매우 적합하다.

따라서 RNA-Seq 실험의 성공을 위해서는 연구자가 분석 목표, 시료 품질, 예산 등 다양한 요소를 종합적으로 고려한 실험 디자인을 수립해야 하며, 이 과정에서 플랫폼 선택은 매우 중요한 결정 요소이다.

이바이오젠은 Lexogen 공식 파트너로서, 연구자가 실험 설계부터 데이터 분석까지 신뢰할 수 있는 가이드를 제공하며, QuantSeq 및 mRNA-Seq 전 과정에 대한 컨설팅 기반의 맞춤형 서비스를 지원한다. 또한, 자체 개발한 ExDEGA 분석툴을 통해 DEG 해석의 효율성과 직관성을 높이고 있어, RNA-Seq 을 처음 시도하는 연구자도 폭넓게 활용하실 수 있다.

## QuantSeq 3' mRNA-Seq

<b>Sample requirement<sup>+</sup></b>	>1µg total RNA (Min. >10ng total RNA)
<b>Library method</b>	QuantSeq 3' mRNA-Seq Library Prep Kit FWD
<b>NGS run format</b>	NextSeq, SE100 bp
<b>Data yield</b>	>10M read/sample
<b>Turnaround time</b>	~2 weeks after RNA QC
<b>Sample type<sup>++</sup></b>	Tissue, Cell, Whole Blood(Paxgene), Plant, Fungi, FFPE, Sorted cell, etc.

## mRNA-Seq /total RNA-Seq

	<b>total RNA-Seq</b>	<b>mRNA-Seq</b>
<b>Sample requirement</b>	>2µg total RNA	>2µg total RNA
<b>Library method</b>	RiboCop rRNA Depletion kit + CORALL RNA-Seq Library Prep kit	Poly(A) RNA Selection kit + CORALL RNA-Seq Library Prep kit
<b>NGS run format</b>	NextSeq, PE150 bp	NextSeq, PE150 bp
<b>Data yield</b>	> 9 Gb/sample	> 6 Gb/sample
<b>Turnaround time</b>	~3 weeks after RNA QC	~2 weeks after RNA QC
<b>Sample type<sup>+</sup></b>	Tissue, Cell, Whole Blood(Paxgene), Plant, Fungi, etc.	

이바이오젠에서 지원하는 분석툴인 ExDEGA (Excel based Differentially Expressed Gene Analysis)는 엑셀 기반으로 DEG 분석을 보다 쉽게 진행할 수 있다. 같이 제공하는 ExDEGA GraphicPlus는 다양한 그래프(Clustering Heatmap, Correlation, PCA, DAVID functional analysis etc)를 쉽게 만들어주는 ExDEGA 지원프로그램이다. ExDEGA 및 GraphicPlus 툴은 이바이오젠에서 제공하는 RNA-Seq 데이터에서만 사용 할 수 있다.

## 6. 참고문헌

1. **Lexogen.** LEXOGEN RNA Lexicon - A Practical Guide to RNA Sequencing. Retrieved from <https://www.lexogen.com/rna-lexicon/>
2. **Agilent Technologies.** RNA High Sensitivity ScreenTape Assay: Performance and Use Case. *Technical Overview*, Publication No. 5994-1038EN. Retrieved from <https://www.agilent.com>
3. **Lexogen.** Homepage. Retrieved from <https://www.lexogen.com/>
4. **Lexogen.** QuantSeq and CORALL: Comparison Flyer. Product Release 02.05.2023.
5. **Ren, H. et al.** A Comparison of Low Read Depth QuantSeq 3' Sequencing to Total RNA-Seq in FUS Mutant Mice. *BMC Genomics*, 22, 801 (2021).
6. **Jang, J.S., Holicky, E., Lau, J., McDonough, S., Mutawe, M., Koster, M.J., Warrington, K.J., & Cunningham, J.M.** Application of the 3' mRNA-Seq using unique molecular identifiers in highly degraded RNA derived from formalin-fixed, paraffin-embedded tissue. *BMC Genomics* 22, 769 (2021).