

Transcriptomics – Direct RNA Sequencing by Nanopore Long Read Technology

Introduction

Oxford Nanopore Technologies (ONT) real-time long-read sequencing technology and library preparation kits allow to sequence native RNA eliminating PCR bias from the data. The technology feeds a single-stranded RNA molecule through a protein nanopore and measures changes in electrical current as the molecule passes through. The single native RNA molecule long read sequencing allows for superior reading through full-length and maintains strand methylation status compared to traditional next-generation sequencing. These long reads are essential for analyzing Full-Length RNA transcripts including isoforms, structural variations, Modifications, Poly-A Tail Length. The samples are sequenced with Oxford Nanopore Technologies PromethION P24 sequencer and the resulting reads are then subjected to quality filtering, assembly, annotation, and quality checks using Oxford Nanopore Technologies EPI2ME software and the Nanopore data analysis pipeline developed by Poochon Scientific. The data generated is from 10 Gb and up to 50 Gb with 10 M reads to 30 M reads and N50 read length ~ 2 Kb.

Applications

- **Analyze full-length transcripts**
- **Accurate isoform identification**
- **Detect base modifications**
- **Measure poly(A) tails**
- **Rapidly identify and characterize RNA viruses**

Transcriptomics – Direct RNA Sequencing by Nanopore Long Read Technology

Sample Preparation and Submission

- **Total RNA from any species:** $\geq 5 \mu\text{g}$ per sample, $A_{260}/A_{280} = \sim 2.0$, $A_{260}/A_{230} = 2.0-2.2$, Average size $\sim 1.5 \text{ Kb}$, Concentration $\geq 200 \text{ ng}/\mu\text{l}$, Buffer: DEPC ddH_2O (Water), 10 mM Tris pH 8.0, or low TE ($< 0.1 \text{ mM}$ EDTA)
- **Cell pellet samples from any species:** ≥ 5 million cells per sample
- **Shipping:** Ship with ice packs or dry-ice

Service Description

Construction of an amplification-free library using ONT Direct RNA Sequencing Kit (SQK-RNA004) and PromethION™ Flow Cells RNA (FLO-PRO004RA). The library is sequenced with a PromethION P24 device. The resulting reads are subjected to quality filtering, assembly, annotation, quality checks, isoforms analysis, differential expression analysis, modification analysis (m6A), and poly(A) tail length analysis using the following software ONT EPI2ME, Medaka, QUAST, Dorado, Modkit. Data specifications: bases up to 20 Gb with 10 M reads to 20 M reads and N50 read length $\sim 2 \text{ Kb}$ per dedicated flow cell.

Deliverables

Raw data files:

- 1) Fasta.gz - a compressed file of all the raw ONT sequencing reads
- 2) BAM (optional)

Analysis report files:

- 1) Reports (html) – Transcriptome
- 2) .fasta = polished consensus sequence of the genome
- 3) VCF files
- 4) TSV files
- 5) Other QC files

Turnaround Time

7-15 business days

Workflow

Direct RNA Sequencing

