Xpore

↳ หาตำแหน่ง NRA

Problem Statement

Nanopore Sequencing คืออะไร?

Central Dogma

↳ ลำดับเบสของ DNA มี ACTG

↓ Transcription

MRNA → ตัวที่สนใจ

↓ Translation

Protein

สิ่งที่สนใจของ mRNA

- การแสดงออกของยีน ว่ามีการแสดงออกมากหรือน้อยแค่ไหน ดูลำดับเบสว่าปกติมั้ย

การทำ RNA Sepuencing สมัยก่อน

เตรียม RNA

↓

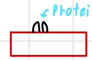
ตัดสาย RNA เป็นสั้นๆ

↓

แปลงกลับไปเป็น DNA

RNA Sepuencing

minIon



โมเลกุลทำให้เกิดกระแสไฟฟ้า แตกต่างกัน

output คือ กราฟไฟฟ้า

↓

ข้อดี

1. ขนาดเล็ก
2. ทำเรียลทามได้

• Data Collection and Preparation

สกุลไฟล์

1. Fast5 สัญญาณไฟฟ้า pA (Raw signal)
   อยู่ใน HDF5 formet (binary)
2. FastQ เก็บข้อมูลลำดับเบสในรูปแบบข้อความ
3. FastA คือไฟล์ลำดับเบส
4. BAM/SAM ไฟล์ที่จัดเรียง FastQ และ FASTA

Direct RNA sequencing
ทำ oxford Nanopore

Basecalling
Guppy (FAST5 → FASTQ)
ดูว่าสัญญาณตรงกับตัวไหน

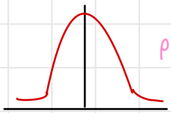
Sequence aligning
(FASTQ + FASTA → BAM) minimap2
ผล ↑ (BAM)

signal event aligning
ใช้ Nanopolish

} nf-core/nanoseq

• Bayesian [multi-sample] Gaussian mixture modelling

Gaussian



$$P(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x \mid \mu_k, \Sigma_k)$$

Gaussian mixture

Frequentis
} หา parameter
Bayasian

• Evaluation

ตัวอย่างทดลอง

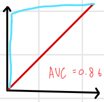
HEK293T: wild type (มี m^6A) / Ko (METTL3 -/-) } Nanopore ทำ direct RNA seq ⇒ ML ทำนาย จาก DRACH

$D \in \{A, G, T\}$

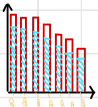
$R \in \{A, G\}$

$H \in \{A, C, T\}$

• Validation: m6A calling



- มีความแม่นยำ 86%



Top positions
- m^6ACE-seq + DRACH
มี ACC > 95%