

RNA-seq 入門

NGS の基礎から de novo 解析まで
準備編・実践編

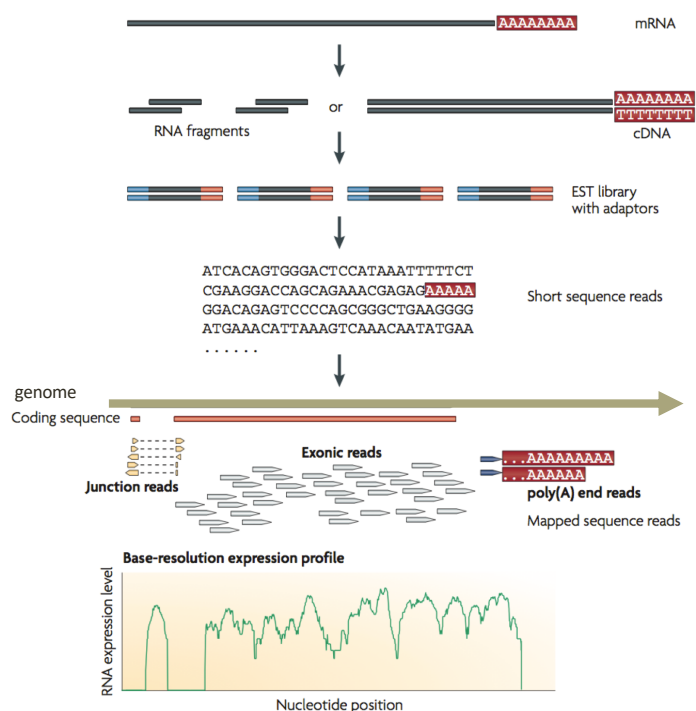
May 30-31, 2019 @ NIBB (Okazaki)

重信秀治 / Shuji Shigenobu

- サポート Wiki

<https://github.com/nibb-gitc/gitc2019may-rnaseq/wiki>

RNA-seq



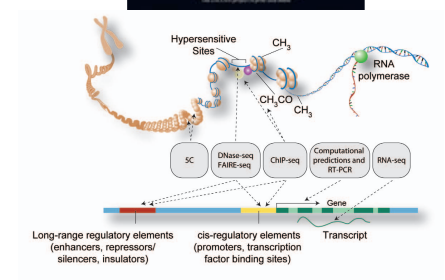
(Wang 2009 with modifications)

Two major goals

- Gene cataloguing
- Gene expression analysis

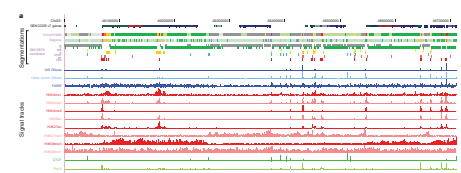
RNA-seq is unraveling complexities of eukaryotic transcriptomes in model and non-model organisms

- Gene expression analysis
- Novel gene discovery (model org.)
 - Coding and non-coding genes
- Gene cataloguing (non-model org.)
- Anti-sense transcripts
- RNA editing
- Novel splicing variants & fusion genes
- Allele-specific expression
- Single-cell RNA-seq for cell population analysis



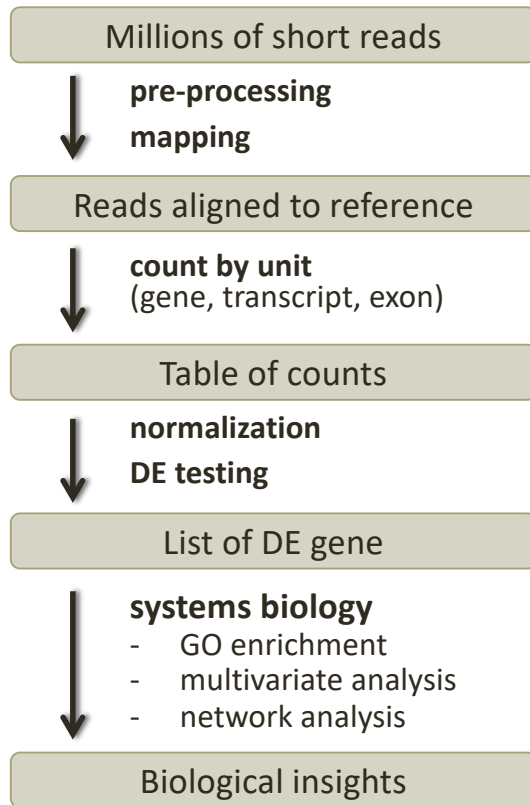
Beyond transcriptome

- DB for proteome analysis
- SNP finding
- *and more ...*



RNA-seq analysis pipeline for DE

Differential Expression analysis



解析ツールの現状: RNA-seq

- 全てのプロセスをこなせる万能ツールはない。
- それぞれのステップに特化したツール群が次々に登場している。

基本戦略

- 各ステップに最適なツールをチョイス、組み合わせた、解析パイプラインの構築。

Pipeline

- 本コースで学ぶオススの3つのパイプライン
 - Transcriptome-based: Bowtie/eXpress/edgeR
Salmon/edgeR
 - Genome-based: HISAT2/StringTie/edgeR

Biologist が身に付けるべき 6つの informatics スキル

- 初級) UNIXの基礎
- 初級) 統計的な考え方と技術
- 初級) 業界標準のツール
- 初級) データ可視化
- 中級) 初歩的なプログラミング
- 中級) データベース

統計的な考え方と技術

- 大規模な-omics データは統計的な考え方なしでは適切に扱えない。
- 統計処理やデータ可視化、データマイニングの道具に習熟する。ーRが最適
- 本コースでは統計的な考え方の基礎から勉強します。データ解析の際に有用だけでなく、実験デザインを組む際にも不可欠です。
- データの可視化も重視します。データを見ながらじっくり「考える」ことは時にブレークスルーをもたらします。

オススメ勉強法

- コマンドやプログラムは自分で試してみる。
copy & pasteでなくタイピングすること。
(熊楠メソッド)
- 気軽に質問する。講師はもちろん、隣や前後の受講生にも。その一方で、ヘルプやマニュアルドキュメントをうまく活用する。
- 自分の研究との接点を常に意識する。自分の研究に応用する。