

RNA-seq 入門

RNA-seq 解析パイプライン

生物情報学を必ずしも専門としない生物研究者が、RNA-seqで得られる網羅的な遺伝子発現データを解析し、生物学的な情報を抽出するための、基礎的技術と考え方を身に付けることを目的としたコース。

<理念>

- For biologists
- 基礎力と即戦力をバランス良く

オンライン講義・実習のための準備

サポートWiki

<https://github.com/nibb-gitc/gitc2025mar-rnaseq/wiki>

- テキスト
- 練習問題
- 参考情報
- コース終了後も公開

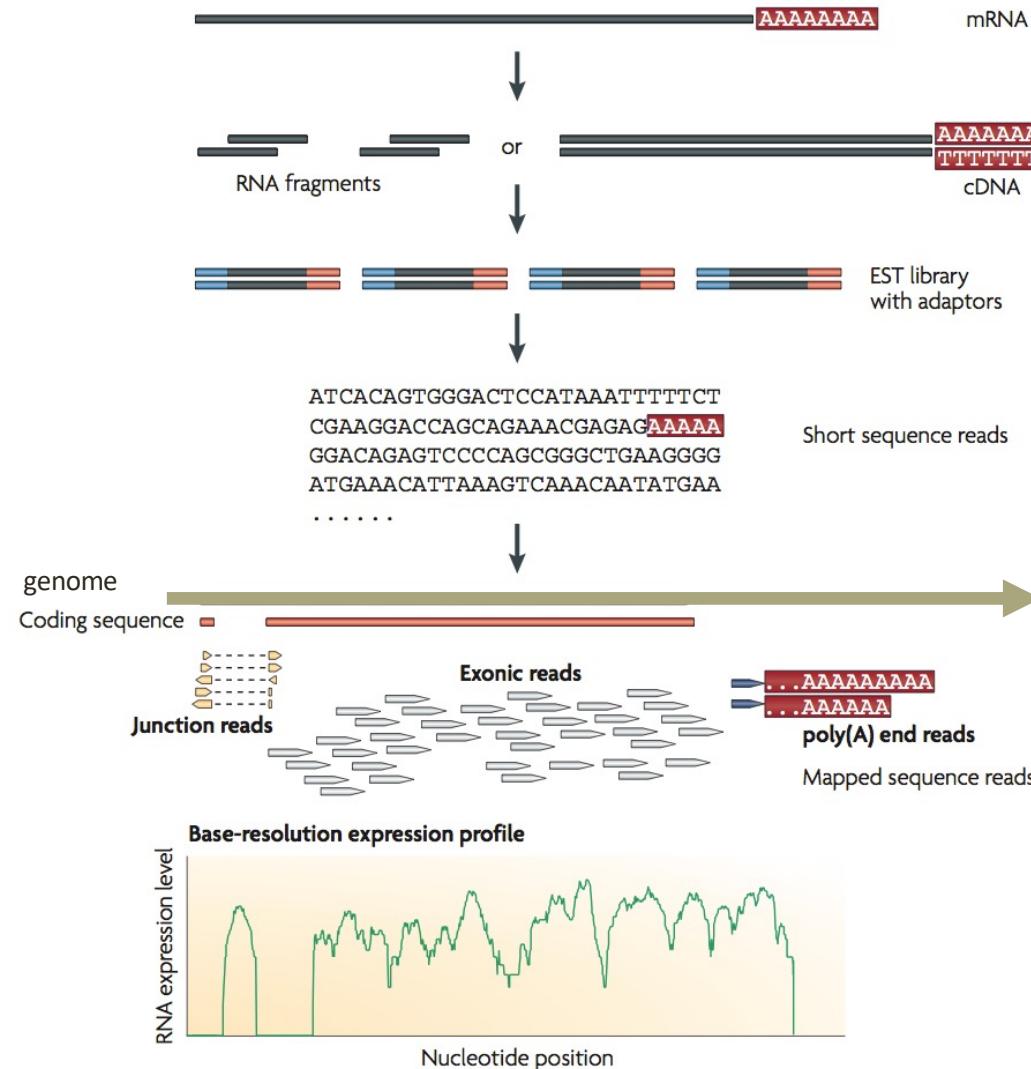
- 講義 オンサイト with Zoom (ハイブリッド)
- 実習
 - 受講生(オンサイト) : Linuxサーバ（リモート環境）と手元のMacBook Pro（ローカル環境）を併用して実習を行います。
 - 聴講生(オンライン) : ご自身の環境

RNA-seq 入門

Program

- NGS 基本フォーマット・ツール復習
- NGS データ可視化ツール
- RNA-seq パイプライン:基礎
- RNA-seq パイプライン:トランскriプトベース
- RNA-seq パイプライン:ゲノムベース
- RNA-seq パイプライン上級:de novo assembly & single-cell RNA-seq
- 多変量解析
- 機能アノテーションとGO解析
- 実践演習

RNA-seq



(Wang 2009 with modifications)

Two major goals

- Gene cataloguing
- Gene expression analysis

宿題

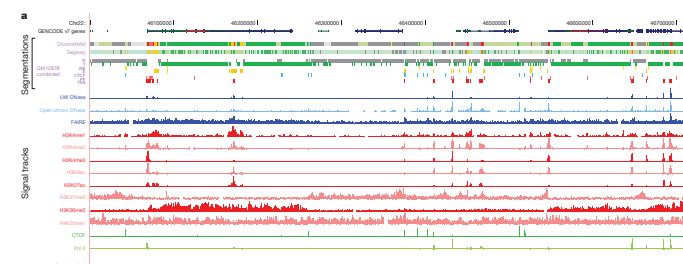
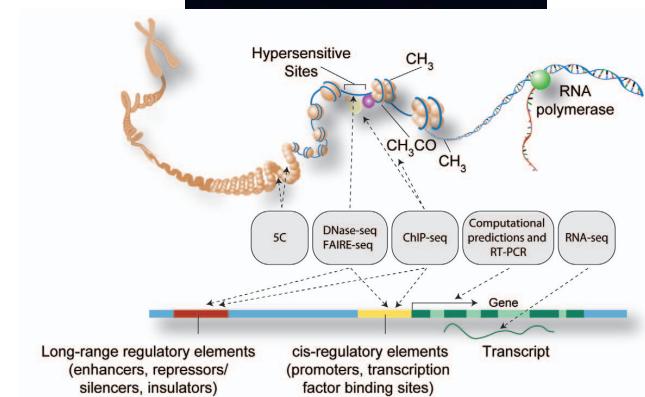
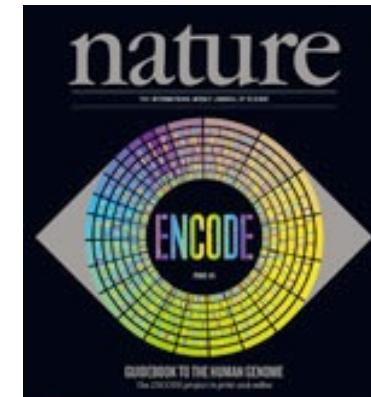
RNA-seqの主な目的を、2つ以上挙げてください。

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
- Spatial transcriptome
- Geneformer (deep learning model)

Beyond transcriptome

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



背景を理解してパイプラインを構築する

- ・バイオインフォマティクス解析は複数のプロセスから構成される
- ・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開発されている。

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

- ・料理のレシピのようなもの。
- ・正解は一つではない。でも王道はある。その上で、目的と環境に合った自分なりのパイプラインを構築。
- ・環境:UNIXコマンドライン、R、Jupyter、プログラミング言語

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多く開発されてきたが、近年成熟の域に入ってきたようだ。

基本戦略

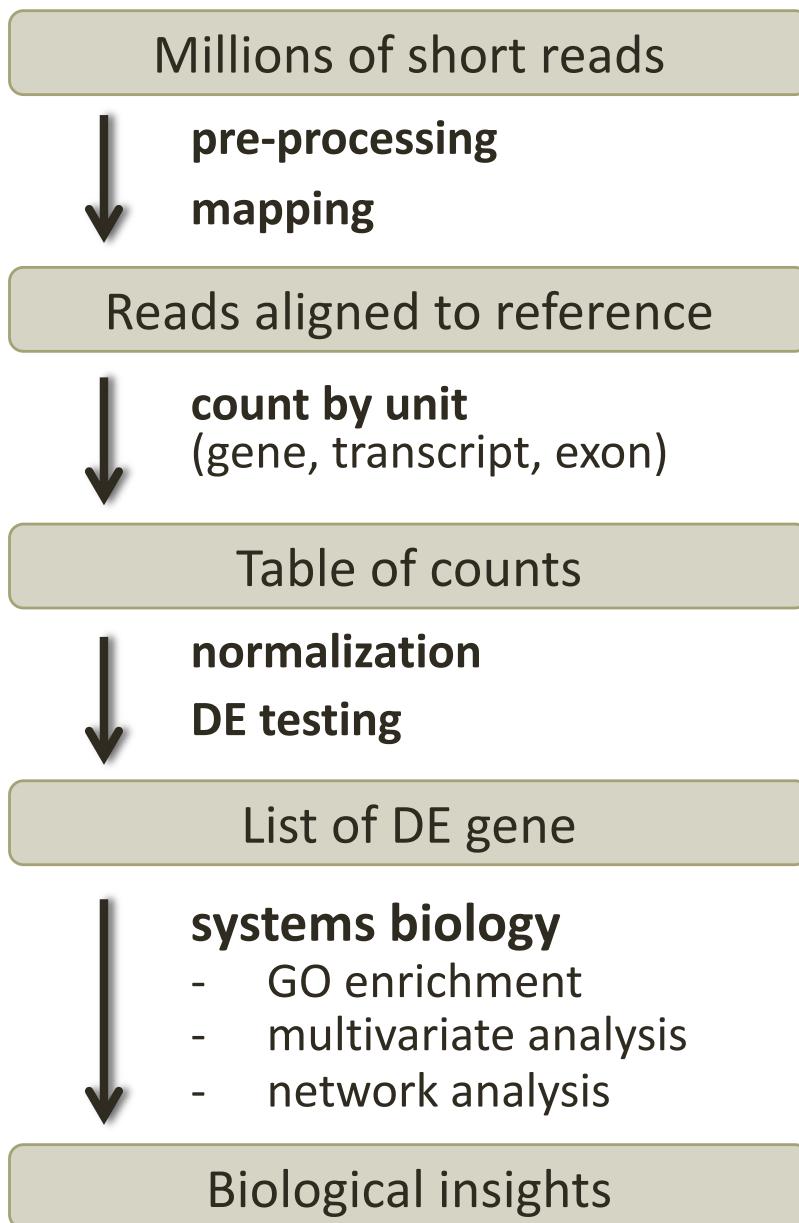
- ・各ステップに最適なツールをチョイス、組み合わせた、解析パイプラインの構築。

Pipeline

- ・本コースで学ぶオススメの2つのパイプライン
 - ・Transcriptome-based: Salmon/edgeR
 - ・Genome-based: HISAT2/StringTie/edgeR

RNA-seq analysis pipeline for DE

Differential Expression analysis



RNA-seq 入門

Program

- NGS 基本フォーマット・ツール復習
- NGS データ可視化ツール
- RNA-seq パイプライン:基礎
- RNA-seq パイプライン:トランскriプトベース
- RNA-seq パイプライン:ゲノムベース
- RNA-seq パイプライン上級:de novo assembly & single-cell RNA-seq
- 多変量解析
- 機能アノテーションとGO解析
- 実践演習

「RNA-seq入門」の位置づけ・目標

- 「NGS解析入門」で身に付けたUNIX, R, NGS等の基本スキルを基盤に、RNA-seqデータ解析の一連の基本パイプラインを自分自身で実行できるスキルを身に付ける。
- RNA-seq基本パイプラインを自分のプロジェクト向けにアレンジできる程度の応用力を身に付ける。

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、受講生どうし教えあう！。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。

- ・一定の基準を満たしたオンラインサイト受講生には修了証書を発行します。
- ・演習問題3問以上

修了証書

田中 太郎殿

あなたは GITC2022 夏 RNA-seq 入門において「RNA-seq 入門：RNA-seq 解析パイプライン」を修了したことを証します

令和 4 年 9 月 1 日



自然科学研究機構
基礎生物学研究所
重信 秀治
内山 郁夫

質問力を向上させよう

- ・研究に関する質問は通常の研究活動のそれと同じ。
- ・コンピュータに関する質問に際しては、質問者にも工夫と心構えが必要。
 - ・自分が実現したこと（目的・ゴール）を明示。
 - ・自分が実施したことの正確な伝達。例えば実行したコマンド、パラメータ詳細。問題の起きた環境（OSやソフトウェアのバージョンなど）。
 - ・何が起きたかを正確に伝える。エラーメッセージは情報の宝庫。
 - ・自分なりの原因の予想や、解決策案を提示するとなお良い。自分なりに試行錯誤したことがあればその情報も。
- ・コマンドヘルプやmanページを積極的に活用する。
- ・質問が上手な人は、成長が早いです。