上田研究室
Ueda Laboratory
http://www.biods.rcast.u-tokyo.ac.jp/

データサイエンスのテクノロジーで生命現象を読み解く

Biological Big Data to Knowledge, using Data Science

データサイエンスを用いた生命情報解析

次世代シーケンサおよび質量分析機から出力される計測データをハイスループットに解析する情報科学的手法の開発を行っています。近年、計測技術の発展により、生物学において算出される電子データは増加の一途をたどっており、大量の生物学データを従来の方法で処理することが困難な課題となっています。加えて、異なる次元のデータを統合し、従来モデル化が難しいデータに対しても関連性を見出すためには、ビッグデータ解析技術や機械学習の最新の成果（データサイエンス）を取り入れて情報解析を行うことが不可欠です。大量のゲノムデータの中から生物学的な意味や関連性を見出すには大規模にデータを集約させ、分散処理を行う必要があります。次世代シーケンサおよび質量分析機の応用範囲は多岐に渡りますが、以下のような領域で研究を行い、同時にソフトウェアを開発しています。特にエピトランスクリプトーム解析のための深層学習やクラウド技術応用、新規アルゴリズムの開発に注力しています。

研究領域の紹介

(1) nanoporeシーケンサを用いたエピトランスクリプトーム(RNA修飾)解析
(2) がんゲノミクス
(3) タンパク質の転写後修飾の解析
(4) シングルセル解析

また、本年度から「先端研Cross-disciplinary Data/AI」イニシアチブを立ち上げ、クラウド/AI技術を用いて先端研各分野のデータ利活用を進めます。

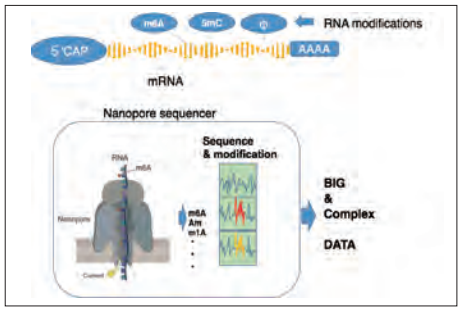
Biological Data Science

With the advancement of sequencing technologies, it has become a challenging task to process large volumes of data with conventional methods. In order to extract knowledge from biological big data, (ex. Multi-omics data) it is necessary to incorporate the latest Data Science technology, such as cloud computing and machine learning. We are developing cloud based Single Cell NGS analysis pipeline using Hadoop/Spark, (cloud computing framework) and developing the method to identify RNA modifications using deep learning method.

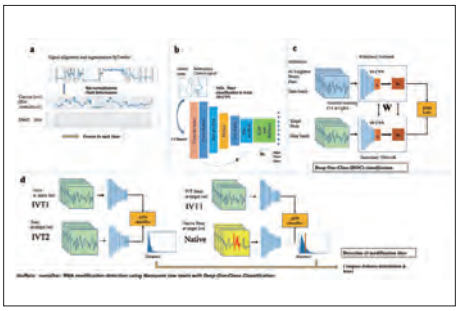
Our research include following:

(1) Epitranscriptome (RNA modifications) analysis using nanopore sequencer
(2) Cancer genomics
(3) Proteomics and post translational modification
(4) Single Cell genomics

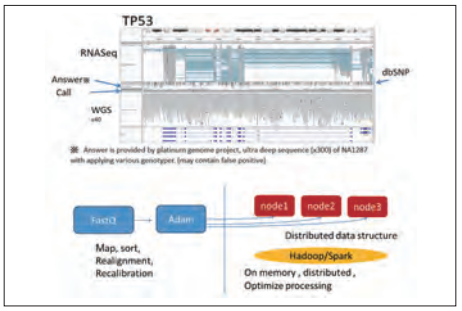
Also, from this fiscal year, the RCAST Cross-disciplinary Data/AI initiative to utilize data from each field of RCAST using cloud/AI technology to be launched, and we will be taking the main role in implementing that project.



1 ナノポアシーケンサを用いたRNA修飾の解析
RNA modification analysis using nanopore sequencer



2 1D-CNNと一クラス分類手法を用いたRNA修飾解析アルゴリズムの開発
Development of RNA modification analysis algorithm using 1D-CNN and one-class classification method



3 Hadoopを用いたRNAシーケンスおよび全ゲノムシーケンス 解析結果
RNA Sequencing and Whole genome sequencing using Hadoop

特任講師
上田 宏生
Hiroki UEDA, Project Lecturer
専門分野：情報生命科学、がんゲノミクス、機械学習
Specialized field：Computational Biology, Cancer Genomics, Machine Learning
E-mail：ueda@biods.rcast.u-tokyo.ac.jp