

進化学夏の学校

「次世代シーケンサーを用いたゲノム多様性解析入門」

日時： 2024年8月24日（土）10:00～12:30

場所： 神奈川県立生命の星・地球博物館 1階講義室

司会： 神澤 秀明（国立科学博物館）

木村 亮介

琉球大・院医

集団ゲノム学入門

An introduction for population genomics

DNA 解析技術やコンピュータ解析技術の革新に伴い、集団遺伝学の理論を用いてゲノム全域を解析する研究が行われるようになった。そうした研究分野は、理論面と実践面とが相乗的に進歩を遂げ、「集団ゲノム学 population genomics」として発展している。従来の集団遺伝学解析では、通常ミトコンドリア DNA の一部の配列など限られた遺伝子領域を PCR 増幅して、系統解析などが行われていた。しかしながら、二倍体生物においては、個々のゲノム領域が必ずしも個体間の遺伝的近縁性を反映せず、また、分岐した集団間において遺伝子の不完全な系統仕分け (incomplete lineage sorting: ILS) が生じることがあることから、少数のゲノム領域の解析から集団間の関係について言えることは多くない。集団間の正しい関係を知るためには、多くのゲノム領域を用いた解析が必須となる。また、ゲノムレベルでの多型データの解析では、予め集団を定義することなく、個体レベルでの解析を進め、解析結果から集団構造を観察し、集団を定義することが可能である。本発表では、集団ゲノム学の理論面や解析結果の解釈について解説する。

長田 直樹

北大・院情報

次世代シーケンサーを用いたゲノム多様性解析の概要

Introduction of genomic diversity studies using Next Generation Sequencers

次世代シーケンスコストの低下により、大規模なプロジェクトではなく研究室レベルの研究で生物のゲノム配列を決定することができるようになってきた。また、1個体だけでなく、1つの生物種または近縁種から多数のゲノム配列を決定し、100万年から数10万年程度の比較的最近の集団史や、そこで起こった自然選択について研究を行うことが可能になってきている。一方、膨大なデータの取り扱いや集団遺伝学の基礎知識など、これらのゲノ

ム多様性データ（集団ゲノムデータ）を扱うためには分野横断的な知識・経験が必要であり、そのハードルはまだ高いと言わざるを得ない。本講演では、これらのデータの利用法や、得られたデータの性質について、初歩的なところから解説を行い、以降の技術的な解説を行う講演の内容の理解へとつなげていきたい。

河合 洋介

国立国際医療研究センター

ゲノムデータから集団サイズの変化を推定する方法

Methods for estimating changes in population size from genomic data

集団遺伝学では対象とする集団の個体数（染色体数）を「有効な集団サイズ」であらわす。単独の遺伝子座の解析やシンプルなモデルを用いる解析では有効な集団サイズと現実の生物の個体数の間には大きな違いがあり、有効な集団サイズは解析上のモデルパラメーターとして扱われることが多かった。ところがゲノム全体を対象とする解析や現実的なモデルを用いることにより実際の個体数を反映した有効な集団サイズを推定することが可能になった。特にゲノムデータを用いて過去の集団サイズの変化を推定する手法は、生態学や人類学の研究に応用されている。本講演ではゲノム多様性情報から集団サイズを推定する原理を解説して、いくつかの解析プログラムを紹介する。

松波 雅俊

琉球大・院医

ターゲットシーケンスの原理と解析法

Introduction for target-sequencing

ゲノムを解析するにあたっては、さまざまな理由からゲノムの全ての塩基配列ではなく、その一部だけを読み取る必要性が生じることがある。ターゲットシーケンスとは、このような場合にゲノムの領域を限定して解読する手法のことである。次世代シーケンスのコストは低下したと言われているが、巨大なゲノムを持つ生物や多検体のサンプルを解析する際には、引き続き重宝されている手法である。しかし、ターゲットシーケンスにより解読されたデータは、全ゲノムデータに比べると特殊であり、独自の解析パイプラインを構築する必要がある。本講演では、代表的なターゲットシーケンシングの手法と解析法を概説し、実例として講演者がこれまでおこなってきた RAD-seq 解析を紹介する。

藤原 一道

遺伝研

動植物のゲノム解析とアセンブリ入門

近年、次世代シーケンス技術の発展により、ヒト以外の動植物のゲノム解析が急速に進展している。シーケンスに必要なコストの低下により、モデル生物以外でも全ゲノムの塩基

配列の解読が可能となり、一部の配列を用いていた従来の手法に比べ、全ゲノム情報という大量のデータを用いて詳細な研究が行えるようになった。さらに、最近では PacBio HiFi reads や Nanopore ultra-long reads の登場により、従来はアセンブリが困難であった生物のゲノム配列のアセンブリも容易になりつつある。これらのデータを用いた解析には、様々なツールを用いた手法が存在するが、現状ではどれか一つに確立した手法は存在しない。本講演では、動植物の全ゲノム配列の利用方法やゲノムアセンブリについて、入門的な内容から技術的な解説を行う。また、一つの実例として講演者が行ってきた解析についても一部紹介する予定である。