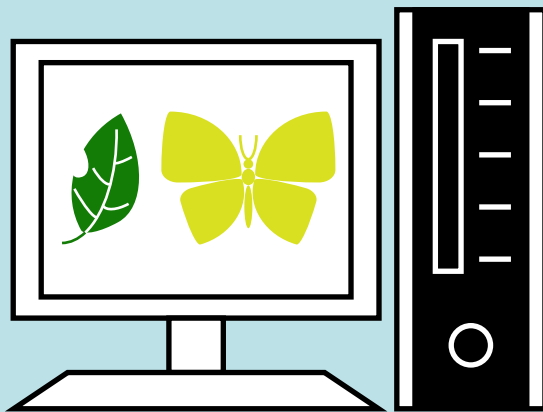
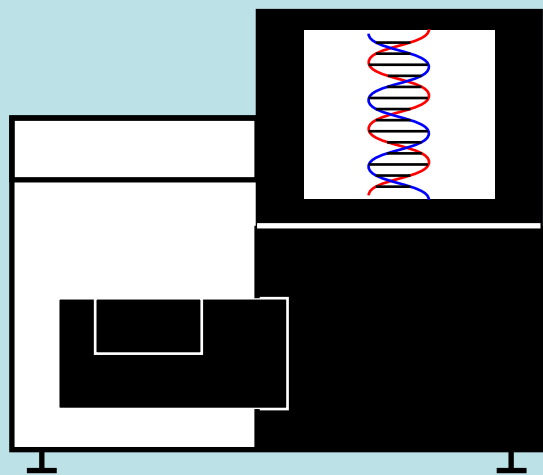


# 公共データを利用した 新知見の発掘

NGS・メタ解析・機械学習といった最新の研究手法を学ぶことで、公共データベースや文献・書籍・図鑑に登録されているデータを活用して、誰でも既存の情報の集合から新しい知識を取り出すことができます。研究成果と合わせて、解析方法とその学習方法を解説します。

7月23日(日)  
9:00 - 11:00

会場：JT 生命誌研究館



**仲里 猛留**

製品評価技術基盤機構

昆虫研究に使えるデータ  
ベースと解析方法の学び方

**母 浩平・坊農 秀雅**

広島大学・統合生命

公共 RNA シーケンスデータを利用した  
メタ解析による発現変動遺伝子の特定

**横井 翔**

農研機構・生物機能

公共データベースに存在するシーケンス  
データを利用したミツバチのメタ解析

**尾崎 克久**

JT 生命誌研究館

教師なし機械学習を用いた  
アゲハチョウ科昆虫の食性進化の解析

## 昆虫研究に使えるデータベースと解析方法の学び方

仲里 猛留（製品評価技術基盤機構）

遺伝子やゲノムのデータといえば、もともとはヒトやマウスのようなモデル生物を中心に構築されてきた。しかし、近年はNGS技術の発達などもあり多くの昆虫で遺伝子やゲノムのデータが産出されつつある。これらには既存のデータと同様に昆虫データが収載されているものから、独自のデータベースとして公開されているものまでさまざまである。今回は、どこにどのようなリソースがデータベースとして存在しているかを紹介するとともに、実際にデータ解析する際のコンピューター環境の整備などについても紹介する。

## 公共RNAシーケンスデータを利用したメタ解析による発現変動遺伝子の特定

梶 浩平（広島大学・統合生命）・坊農 秀雅（広島大学・統合生命）

発表者たちは、公共データベースに登録されているRNAシーケンスデータの統合と再解析（メタ解析）を通して、公共RNAシーケンスデータは新規の発現変動遺伝子の発見をもたらす貴重な資源となることを示してきました（Toga et al. 2022, *Insects*, <https://doi.org/10.3390/insects13100864>, Toga and Bono 2022, *bioRxiv*, <https://doi.org/10.1101/2022.11.20.516280>）。本発表では、これらの結果や手法を概説するとともに、分子生物学・生理学的な解析を主に行ってきた研究者が、公共RNAシーケンスデータの解析手法を習得できた方法を紹介したい。

## 公共データベースに存在するシーケンスデータを利用したミツバチのメタ解析

横井 翔（農研機構・生物機能）

ミツバチは社会性昆虫の代表種として長年研究材料として利用されてきた。一方、ハチミツの産生や送粉者として産業上重要な家畜昆虫である。そのため、公共データベースにはミツバチ種のゲノムやRNA-Seqデータが多数登録されている。これらのデータはそれぞれ独立した研究テーマを実施する過程で得られたデータである。メタ解析とはこれらの、複数のデータを統合的に解析して新たな知見を得ようとすることである。演者は公共データベースにあるミツバチのゲノムとRNA-Seqデータを用いてメタ解析を実施し、ミツバチ種における転移因子の比較解析とセイヨウミツバチの参照トランスクリプトームデータの確立を行ったので、解析の具体的な方法を交えながらお話したい。

## 教師なし機械学習を用いたアゲハチョウ科昆虫の食性進化の解析

尾崎 克久（JT生命誌研究館）

アゲハチョウ科の昆虫は植食性で、多くは特定の科の植物に依存している。どのチョウがどの植物を食べたかについて、多くの愛好家の多大な努力によって情報が蓄積されており、それを知識として利用できるようなデータベースとして構築している（*InsectInDB*）。文献から何らかの昆虫に対して毒性を示したという記述がある植物化学物質をデータとして収集し、教師なし機械学習を用いて植物の毒性の類似性を分類した。その結果から、アゲハチョウ科の多様化がどのようにして進行したのか考察する。今回、機械学習を利用した解析方法と、その学び方について紹介する。

*InsectInDB*: <https://ja.insect-plant.org/>