

データサイエンス講習会のお知らせ

Rによるデータ解析の基礎と応用

一般社団法人バイオ産業情報化コンソーシアム(JBIC)

東京大学大学院農学生命科学研究科・アグリバイオインフォマティクス教育研究ユニット

2025年1月9日

Rは統計やデータ解析に特化したプログラミング言語・環境です。特に学問の世界ではRによるデータ解析が主流ですので、既にPythonなど他のコンピュータ言語を使ったデータ解析を行っている人にも、R言語は身に着ける価値のあるスキルとなっています。本講習会は、これまで(2020年2月、2021-2024年3月)に実施した講習会と同様に、**データサイエンティストの裾野を広げることを目的**として実施します。本データサイエンス講習会は、**オンライン形式にて2日間開催**します。講習会では、Rを利用したことない**初心者を対象とした基礎的な内容(1日目)**と**応用的な内容(2日目)**について、Rを便利にするための開発環境であるRStudioを使用した演習形式で行います。

1日目は、R/RStudioの基本的な使い方・活用方法を学びながら統計解析、データの可視化方法に関する基本的な知識と技術を学びます。2日目は、Rを用いた遺伝子発現解析の基礎を学ぶとともに、近年、生命科学で重要性が高まっている遺伝子型情報を利用したデータ解析手法の一つであるゲノムワイド関連解析(GWAS)について学びます。2日間を通して受講することでRを用いたデータ解析を独自で行えるようになることを目指します。

本講習会では基本的に植物から得られた情報を対象とした解析を行います。受講者各々が取り扱っている生物種にも適用可能な解析方法についての解説・演習となっています。

*本講習会は、2019年2月に、経済産業省の委託事業として実施した『「産業界横断的なバイオ分野データサイエンス関連人材」のスキル標準の策定及びそれらの育成プログラムの開発/実証』バイオインフォマティクス講習会の後継として実施するものです。

URL: <https://www.learning-innovation.go.jp/verify/c0037/>

日時: 2025年3月6日(木)~7日(金)、10:00~16:00(予定) 各日の講義終了後に質問の時間を設けます(参加自由)。

会場: **Webex オンライン講義**

対象者: 講義開催日時点でJBIC会員企業に所属される方

※接続用URLは、受講申込者に、講習会前日までにメールでお知らせいたします。

※講習会終了後、講義動画をJBIC会員サイトにて公開予定です。

1日目(2025年3月6日(木)予定内容):初心者向け講習会(基礎)

・午前:[Session 1] Rを用いた統計解析・データ可視化(東京大学大学院農学生命科学研究科・准教授 大森良弘)

1日目午前は、Rを用いたデータの読み込みから整形、簡単な統計解析とグラフ化、ファイルの保存までの一連の作業を学びます。演習はデータ解析入門者向けの基礎知識の解説を織り交ぜて行います。

・午後:[Session 2] Rでインタラクティブなグラフの作成(東京大学大学院農学生命科学研究科・准教授 大森良弘)

[Session 3] R Markdownでレポートの作成(東京大学大学院農学生命科学研究科・准教授 大森良弘)

1日目午後は、Rでのインタラクティブなグラフの作成について簡単な演習を行います。データの探索や理解を向上させる高度なデータ可視化技術の基礎を学びます。また、後半には、R Markdownを利用したレポートに作成について簡単な演習を行います。Rで再現可能なデータ解析とレポート作成を実践するための技術を学びます。

2日目(2025年3月7日(金)予定内容):中級者向け講習会(応用)

・午前:[Session 4] Rを用いた遺伝子発現解析、遺伝子発現ネットワーク解析-入門 (東京大学大学院農学生命科学研究科・准教授 大森良弘)

2日目午前は、RNA シーケンスから得られた数値(発現値)情報を用いた遺伝子発現解析の演習を行います。発現変動遺伝子の検出(DEG 解析)、機能解析(GO 解析、パスウェイ解析、ネットワーク解析)の基礎を学びます。

・午後:[Session 5 Rを用いたゲノムワイド関連解析(GWAS)] (東京大学大学院農学生命科学研究科・助教 石森元幸)

2日目午後は、近年ヒトや様々な動植物集団における原因遺伝子探索手法として一般的となったゲノムワイド関連解析(GWAS)を取り上げ、R パッケージ RAINBOWR を使用して、遺伝子型情報(ゲノムワイドマーカー)を活用した解析の演習を行います。GWAS は中級者としてはやや難しい内容ですが、生命科学で扱うビッグデータの一つであるゲノム情報の解析について学びます。

受講申込:<https://forms.office.com/r/GHmR432nBN>にある参加申し込みフォームよりお申込みください。

(*本講習会はJBIC 会員企業に所属されている方に限定させていただきます。)

申込期間:2025 年 1 月 9 日(木)~3 月 3 日(月)17:00 まで

受講費用: 無料

注意事項:

1. 本講習会はデータサイエンス初心者~中級者を対象としておりますが、事前準備として、講習会で使用する PC への R と RStudio のインストールを受講者各人をお願いいたします。事前準備は 4.「令和 6 年度の講義資料」内の「day0_事前準備」の「R と RStudio のインストール」を参照して行ってください。講習会は R と Rstudio の最新版が問題なくインストールされている前提で行います。また、講義中に R のパッケージを PC にインストールすることがありますので PC の設定にご注意ください。設定の変更が難しい場合は posit Cloud 等の RStudio を Web ブラウザー上で実行できるサービスの利用をご検討ください。
2. 今回のオンライン講義は、RStudio を操作しながら行うハンズオン形式ですので、可能であれば、講義視聴用の PC と RStudio 用の PC をそれぞれ別にご用意頂き、講義はデスクトップ PC などの大きめのモニターで視聴いただくことを推奨いたします。
3. 本講習会で利用する遺伝子型情報はバリオアントコール後のファイル(Variant call format; VCF など)です。FASTQ ファイルの前処理やマッピングなどはい行いませんのでご注意ください。
4. 令和 6 年度講習会の講義資料は、以下の JBIC 会員専用サイト内に 2 月下旬頃に公開予定です。
https://www.jbic.or.jp/members_only/data_science (ID・PW が不明な場合は JBIC にご連絡ください)
5. 両日ともにレポート課題を出します。レポートの提出は必須ではありませんが、提出して合格と認められれば、「令和 6 年度 R 講習会 修了認定証(仮称)」を授与することを予定しています。

問合せ:

一般社団法人バイオ産業情報化コンソーシアム(JBIC) lecture@jbic.or.jp (片倉)

東京大学大学院農学生命科学研究科・アグリバイオ事務局 koshukai@iu.a.u-tokyo.ac.jp (大森)

以上