

はじめに

現代の生命科学研究は、まさに「データ駆動の時代」の真っ只中にあります。次世代シーケンサー（NGS）が日々生み出す膨大なゲノムやトランスクリプトームのデータ、高性能な顕微鏡が捉える精細なイメージデータ、質量分析計が網羅的に測定するプロテオームやメタボロームのデータ——私たちの周りには、生命の謎を解き明かすための貴重な情報が溢れています。本書『生命科学データ解析をはじめの前に読む本』は、これらの大規模データを自らの手で解析し、意味のある知見を引き出すスキル、つまりすべての生命科学研究者にとって不可欠な「第二の実験技術」を身につけるための一冊です。

しかし、そう言われても、「プログラミングは難しそう」「どこから手をつければいいのかわからない」「自分はコンピューターが苦手だから…」といった不安を感じる方は少なくないでしょう。研究室の先輩や同僚が黒い画面に謎のコマンドを打ち込んでいるのを見て、自分とは違う世界の話だと感じてしまった経験があるかもしれません。何を隠そう、著者自身もそうでした。入門書を買ってみたものの、最初の環境構築でつまずき、次々と表示されるエラーメッセージの意味もわからず、静かに本を閉じてしまった経験が一度ならずあります。

もしあなたがそのような経験をおもちだとしても、どうか心配しないでください。本書は、まさにそのような方々へ向け、生命科学者による、生命科学者のためのデータ解析入門書の決定版をめざして執筆しました。徹底的に生命科学者・学生の目線に立ち、専門用語を可能な限り噛み砕き、誰もが順を追って着実にステップアップできる構成を心がけています。

本書の最大の目的は、データ解析という広大な海に挑むための「考え方の地図」と「進み方のコンパス」を皆さんに提供することです。そして、皆さんが本書を通じて、データ解析の世界への第一歩を踏み出し、将来的に機械学習のような高度なトピックにも挑戦できる、普遍的で揺るぎない土台を築けるようになることをめざします。

プログラミングは決して特別な才能が必要なものではありません。外国語と同じで、最初は単語も文法もわからず戸惑うかもしれませんが、使い続けるう

ちに必ず上達します。大切なのは、完璧をめざすことではなく、まず一步踏み出し、試行錯誤を繰り返すことです。そうして学習を進めることで、本書を読み終える頃には、あなたは目の前の課題に対して「どのようなツールを使い、どう調べれば解決できるか」という道筋を自分で考え、実行できるようになっているはずです。そうなれば、黒い画面のコマンドやプログラミング言語は、もはや得体の知れない呪文ではなく、あなたの研究を力強く前に進めるための、ピペットや遠心機と同じ「頼れる道具」のひとつになっていることでしょう。

最後になりますが、本書の出版にあたり多大なるご尽力をいただいた羊土社の皆様に、この場を借りて心より御礼申し上げます。企画段階から並走し、本書の方向性を導いてくださった本多正徳様。そして、本書の「最初の読者」として、執筆の全行程において粘り強く細やかなフィードバックをくださった亀谷博就様と岩崎太郎様。お三方のプロフェッショナルな支えと、読者の視点に立った的確なご指摘がなければ、この一冊を完成させることはできませんでした。

さあ、一緒にデータ解析の世界への扉を開きましょう。この一冊が、あなたの研究生活をより豊かで実りあるものにするための一助となれば、著者としてこれ以上の喜びはありません。

2026年4月

東京科学大学 総合研究院 M&D データ科学センター

東京科学大学 大学院医歯学総合研究科

清水秀幸