

データ駆動型のがん多様性研究

がんの多様性は、腫瘍内多様性(intra-tumor)、腫瘍間多様性(inter-tumor)、患者間多様性(inter-patient)という”広がり”の多様性概念と、ゲノム変異、エピゲノム制御、遺伝子発現、そして表現型や細胞機能という”深さ”の多様性概念が複雑に絡み合ったもので、従来の仮説検証型研究も重要である一方、この複雑な課題に挑戦するためには、大量のデータから新しい発見を目指す「データ駆動型アプローチ」が必要です。私たちの研究室では、シングルセル解析とバルク解析の両方で ATAC-seq, RNA-seq, ChIP-seq, Cut&Tag などさまざまなモダリティによる臨床検体、マウスモデル、細胞株モデルの解析を進めています。それらを統合し、重要な生物学的知見にたどり着くためには、再現性の高い解析環境の整備から始まり、各種解析パイプラインの構築、そして R や Python を使用した高次解析の幅広い知識と応用力が必要です。

私たちはこれまで、Docker を用いた解析環境の整備と解析コードの一般化、より簡便に実行できる一次解析パイプラインの構築、そして複数のパッケージや独自に構築した解析コードを組み合わせ、乳がんの腫瘍間・腫瘍内多様性に迫る研究を行ってきました。その結果、多くの研究シーズが生まれており、DRY 解析で見つかった現象を、どのようにすれば実際に検証できるかを、WET 実験を担う研究者とディスカッションをしながら研究を進めています。

バイオインフォマティクス解析は、未経験者であっても、課題を自ら解決しようとする意欲と集中力があれば、必ず習得できる技能です。バイオインフォマティシャンは、その需要に供給が追いついておらず、今後もニーズが拡大し続けると言われています。WET 研究者であっても、その基礎的な手法についての理解と実践経験は、研究を進めていく上で大きなアドバンテージとなります。私たちの研究室では、ほぼ未経験から始めたメンバーが多くのデータ解析に従事しています。教育体制も整備しており、週 1 回の解析チームミーティングで、解析の妥当性や結果の解釈について少人数で議論するほか、エキスパートにいつでも質問可能な体制を整えています。

(文責：糸川昂平)