

基礎生物学における新発見にデータ同化はどう有用か？

木村 暁 (Akatsuki Kimura)

国立遺伝学研究所・細胞建築研究室、
総合研究大学院大学・遺伝学専攻

データ同化は、ある現象に対応する数理モデルにおいて、現象を再現・予測するパラメータ値などを観測データから推測する手法である。多数のパラメータからなる複雑な数理モデルに対しても、観測データの量や質に応じて、パラメータ値の推定や現象が示す挙動の予測を行うことができるため、天気予報など幅広い分野で力を発揮している。

発表者が属する基礎生物学の分野（より具体的には細胞生物学や発生生物学の分野）においては、近年、生命現象を数理モデルの枠組みで表現し、理解する機運が高まっている。これらの現象の多くは、膨大な種類の遺伝子の機能が絡み合っており、それらを表現する数理モデルも多数のパラメータからなる複雑なものになると思われる。一方で、顕微鏡技術の発展に伴い、細胞内といった微小空間の様子を高い時空間解像度で得られるようになってきて、観測データもビッグデータ化しつつある。このような状況（複雑な数理モデル、膨大な観測データ）から、基礎生物学分野においても、データ同化にかかる期待は大きい。

このような期待の一方で、基礎生物学分野とデータ同化の相性については懸念点もある。データ同化においては、数理モデルの基盤となる確固たる理論の存在が重要だが、生物学分野においてはそのような理論はほとんどない。そのような理論を発見すること自体が基礎生物学研究の大きな目的である状況の中で、データ同化をどのように基礎生物学分野における発見に活用できるかは議論が必要なポイントであろう。また、データ同化は未知の将来を予測する目的に威力を発揮するが、基礎生物学研究の現場においては、すでに再現性良く観察されている現象の説明が目的となることが多い。（医療診断の現場では予測はとても重要だが。）

以上のように、基礎生物学研究の現場において、データ同化は大きな潜在能力を有していると同時に、懸念点もある。本発表では、生物学研究者の視点から、データ同化にかかる期待の背景や、発表者らが行なっているデータ同化を取り入れた研究の試みについても紹介し、議論をしたい。