



ACIDES：タンパク質スクリーニング解析 アルゴリズムの技術革新

ヒューマン・メタバース疾患研究拠点（WPI-PRiME）

特任准教授 根本 孝裕

 https://researchmap.jp/takahiro_nemoto

研究の概要

タンパク質スクリーニングはタンパク質の創出や機能調査で使われる実験手法の一つで、創薬や病気のメカニズム解明など、医療分野の実験において頻繁に使用されている。精度の高いスクリーニングを行うためには実験結果をより正確に解析する必要があるが、次世代シーケンシング（NGS）を組み合わせた実験の統計誤差評価はまだ理解が進んでいなかった。今回我々は、フランスの Institut de la Vision の研究グループと共同で、高分散な NGS ノイズを記述する統計モデルとタンパク質スクリーニングの数理モデルを組み合わせた（ACIDES）。これによりタンパク質スクリーニング実験の統計誤差決定が、既存手法の精度を超えて可能になった。

研究の背景と結果

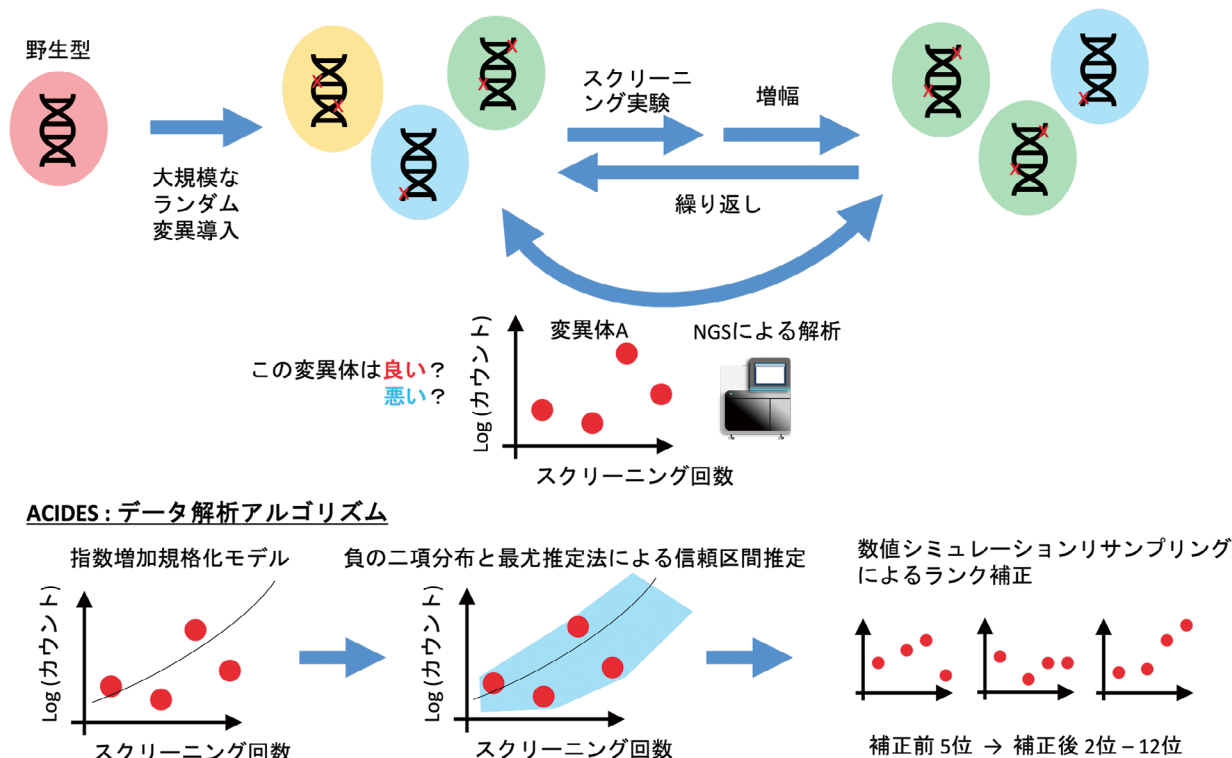
タンパク質スクリーニングとは、ランダムな遺伝的変異を与えたタンパク質をふるい分ける手法で、遺伝的変異によってタンパク質が受ける影響を調べたり、また新しいタンパク質を設計したりするのに使われる。SARS-CoV-2 受容体結合領域を調べるのにも使われた Deep Mutational Scanning 法や、2018年度のノーベル化学賞の対象にもなったタンパク質設計手法である指向性進化法（Directed Evolution）も、タンパク質スクリーニングの応用の一つである。特に、数百万もの DNA 配列を同時に決定できる次世代シーケンシング（NGS）と併用することで、スクリーニング中のタンパク質に対して大量のデータを得て、機械学習を応用する研究が最近注目され始めている。しかし、

次世代シーケンシングによる統計的な誤差は一般的に大きく、それがタンパク質スクリーニングに対して与える影響はよく知られていなかった。

本研究では、フランスの Institut de la Vision の Deniz Dalkara、Ulisse Ferrari らと共同で、次世代シーケンシングによる統計誤差がタンパク質スクリーニング実験においてどのように伝搬されるかを調べた。PCR 増幅などによって増倍される高分散な NGS ノイズは負の二項分布によって統計的に記述できることが知られていたが、その統計モデルを、タンパク質スクリーニング実験の数理モデルと最尤推定法を使って組み合わせた。既存の複数のアルゴリズムと系統的に比較したところ、今までにない精度でタンパク質スクリーニング実験の評価が出来ることが明らかになった。提唱したアルゴリズムは、ACIDES (Accurate Confidence Intervals for Directed Evolution Scores) の名の下、一般に公開されている。

研究の意義と将来展望

ACIDES を応用することで、指向性進化法の実験パイプラインをより正確に設計し、遺伝子治療分野を前進させるウイルスベクターの開発につながる事が期待される。Institut de la Vision の共同研究グループは、網膜細胞への遺伝子導入に最適化されたアデノ随伴ウイルス変異体（AAV7m8）を開発した実績も持つ。タンパク質スクリーニングを使う指向性進化法を用いてアデノ随伴ウイルスを改良する研究は盛んに行われていることから、今後 ACIDES が遺伝子治療の開発に大きな役割を果たすことが期待される。



特 許

論 文

Nemoto, Takahiro; Dalkara, Deniz; Ferrari, Ulisse et al. ACIDES: on-line monitoring of forward genetic screens for protein engineering. Nature Communications. 2023, 14, 8504. doi: 10.1038/s41467-023-43967-9

参考URL

https://resou.osaka-u.ac.jp/ja/research/2023/20231226_1

キーワード

遺伝子治療、指向性進化法、次世代シーケンシング、アルゴリズム開発