

タンパク質構造予測からの新奇酵素の探索

課題責任者

高木 善弘

海洋研究開発機構 超先鋭研究開発部門 超先鋭研究開発プログラム

著者

高木 善弘*¹, 立岡 美夏子*², 澄田 智美*²

*¹海洋研究開発機構超先鋭研究開発部門超先鋭研究開発プログラム, *²海洋研究開発機構海洋機能利用部門生命理工学センター

キーワード: 新規酵素, タンパク質立体構造, AlphaFold2

1. 背景

深海には太陽光エネルギーに直接影響されない広大な深海平原が広がっており、ここには表層からの沈降有機物が堆積し、これを栄養源とする生態系が存在する。海洋表層の植物プランクトンにより生産された有機物は、沈降過程で分解が容易な物質から優先的に消費され、難分解性物質が濃縮された状態で深海に到達する。そのため難分解性物質の資化に最適化された生存戦略は、深海の生存競争を勝ち抜く上で大変重要であり、その生物ゲノムには、これまで発見されていないユニークな資化酵素群がコードされていることが期待される。一方、深海平原のデトリタスを中心とした生態系と異なり、海底の下部からの熱エネルギーから自ら有機物を合成する生態系が存在する。深海熱水噴出孔周辺には、熱水に含まれる水素、硫化水素等の還元物質をエネルギー源とした化学合成微生物が生息し、これら微生物を体内で宿し、生育に必要な栄養を獲得する大型動物が繁茂している。この特異的な共生システムを成り立たせている遺伝子には未知な機能を備わっていると期待される。

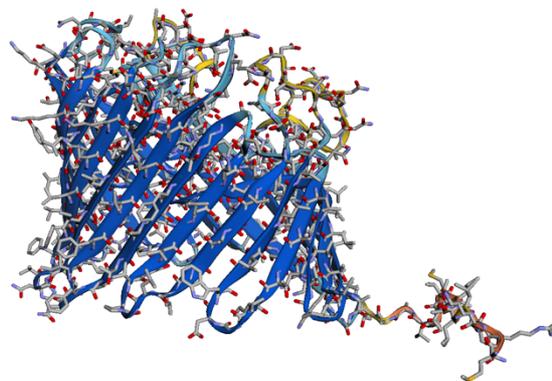
近年のゲノムシーケンス技術の進歩により、深海生物のゲノムを解読することが可能になり、膨大な遺伝子配列データベースが構築され、今なお増大し続けている。このデータベースから、これまで発見されていないユニークな資化酵素群がコードされていることが期待される。我々は、これまでの研究において、難分解性のセルロースを分解する新奇酵素を深海微生物のゲノム上に多数見出している[1]。しかしながら、機能が推定可能な遺伝子は、ほんの一部であり、半数以上が相同配列がなく機能不明である。我々は、その機能不明遺伝子の中に利用価値の高い酵素遺伝子が潜んでいると考えている。最近、高精度のタンパク質構造予測プログラム、「AlphaFold2」が開発され幅広い研究分野にて注目を集めている[2]。このプログラムは、配列ベースに相同性を示さない機能未知タンパク質でも、その構造を予測できる。タンパク質のその構造と機能に密接な関係があり、特定の構造をとることではじめて機能を持つことができる。したがって、機能未知タンパク質の機能を解明するにあたり、予測される構造は重要な手がかりになる。

2. 実施内容

本課題では、機能未知タンパク質の機能を解明するため、機構が所持する ES4 上で AlphaFold2 を使ったタンパク質立体構造予測プラットフォームを構築することを目指した。この目的のため、1) AlphaFold2 プログラムの使用法の修得、2) デスクトップサーバーを使い、AlphaFold2 のインストール、3) ES4 へのプログラムインストールと構造予測の実施、といった作業計画を立て実施した。

2.1 AlphaFold2 プログラムの使用法の修得

AlphaFold2 プログラムは、「Google Colab」といった Web サービスで利用できる。そこで、化学合成共生二枚貝における高発現している機能未知遺伝子のタンパク質立体構造予測を試みた。ここで予測された構造の一つが、ベータシートが膜貫通の筒型構造を形成するベータバレル型であった。この構造をもつ遺伝子は熱水噴出孔あるいは深海湧水周辺に生息する二枚貝、シロウリガイ、シロカイヒバリガイが宿す共生細菌ゲノムに共通に見出される遺伝子であり、トランスクリプトーム解析から、いずれの共生細菌においても高いレベルで発現していることがわかっていた[2]。このことから、この遺伝子は共生細菌が合成した有機物の宿主への運搬において重要な役割をなす示唆された。



シロウリガイ共生細菌遺伝子の構造予測
(ベータバレル構造)

Hydrothermal Field. Appl Environ Microb 88: e00758-21 (2021).

Google Colab 系の Web サービスでは、使用可能な GPU メモリ量が限られており、AlphaFold2 で予測できるアミノ酸残基の長さに制限があるため、比較的アミノ酸残基の長い遺伝子が多い表層タンパク質や多糖分解酵素には不都合が多くなる懸念された。

2.2 デスクトップサーバーへの AlphaFold2 のインストール

AlphaFold2 のインストールには、コンピューターのマシンスペックやネット環境による様々なテクニカルな問題が多く生じることが Web 上で報告されている。そのため、ES4 への本インストール前に、予め作業工程の問題点を洗い出すための工程が必要と考え、デスクトップサーバーへのインストールを実施した。AlphaFold2 プログラムは、構造予測可能な残基長が GPU メモリサイズに依存すること、予測計算に膨大なデータベースを用意しなければならないなど普通のプログラムよりインストールの敷居が高いプログラムである。まず、必要とするマシンスペック (2.5TB 以上の SSD/HDD 容量、CUDA11 に対応した NVIDIA 製 GPU、大容量の GPU メモリ) を備えたデスクトップサーバーの改良を行った。

インストールを以下の手順でおこなった。1) データベースの準備、2) Docker の環境整備、3) AlphaFold2 プログラムの導入。インストール後、実際プログラムを実行したところ実行エラーとなり結果を得る事ができなかった。原因としてメモリー不足、入力ファイルのフォーマット不備等を確認、改善したものの現段階で解決に至らなかった。

3. 今後の対応

本チャレンジ利用枠での目標であった ES4 でのプログラムインストールを達成できなかった。原因の特定には至っておらず、今後ハードおよびソフトの両面での原因究明につとめ、ES4 への本インストールへの準備を完遂させる

謝辞

本研究は ES4 チャレンジ利用枠に採択され行われました。

文献

- [1] Tsudome, M., Tachioka, M., et al. An ultrasensitive nanofiber-based assay for enzymatic hydrolysis and deep-sea microbial degradation of cellulose. *IScience* 25: 104732. (2022)
- [2] Jumper, J., Evans, R., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589 (2021).
- [3] Hirayama, H., Takaki, Y., et al. Multispecies Populations of Methanotrophic *Methyloprofundus* and Cultivation of a Likely Dominant Species from the Iheya North Deep-Sea