

2023 年度年次報告書

ゲノムスケールの DNA 設計・合成による細胞制御技術の創出

2020 年度採択研究代表者

榊原 康文

慶應義塾大学 理工学部

教授

深層学習を用いたゲノムスタイル特徴抽出と DNA 配列 de novo 設計と合成

主たる共同研究者:

片岡 正和 (信州大学 学術研究院工学系 准教授)

正木 慶昭 (東京工業大学 生命理工学院 准教授)

宮本 憲二 (慶應義塾大学 理工学部 教授)

研究成果の概要

本研究課題では、人工知能 (AI) 技術を用いて、導入する細胞のゲノムスタイルに合わせて DNA 配列を変換することで、異なる種でゲノムを活性化させるための DNA 配列を設計することを主要な目的とする。まず、榊原グループは、大規模言語モデルなどの最新の生成 AI が用いているエンコーダデコーダ方式の Transformer アーキテクチャーを使用し、そして学習データの大規模化に取り組んだ。RefSeq から 150 菌種の全 200 万対のオーソログ配列を収集して学習した結果、未知の DNA 配列の変換に関して精度の向上が確認された。次に、DNA 配列変換 AI システムを起点に、DNA の設計から PET 分解活性の検証までの一連の実験パイプラインを確立した。具体的には、PET を分解する能力を持つ *Ideonella sakaiensis* 201-F6 株の PETase を含む 14 の酵素遺伝子の DNA 配列を取得し、これを配列変換した後、片岡グループが DNA 合成と *Bacillus subtilis* RM125 に導入することで新たな株を作製した。宮本グループにより、AI システムで設計された DNA 配列の PETase に対して、PET 分解活性の評価が行われた。

大規模言語モデルに基づく DNA 配列変換の最新モデルは、モデルのパラメータ数も 2,100 万を超え、DNA 配列の変換を効率的に行うことができる。PET フィルムを用いた分解活性実験において、PETase を配列変換した遺伝子を枯草菌に導入した株が PET を有意に分解している可能性が示唆された。

代謝系遺伝子群をオペロンとして集積し、それらを枯草菌の染色体上に導入する技術を開発した。全遺伝子の枯草菌染色体への組み込みにより、順次安定性が良く、培養時に抗生物質での選択を必要としない染色体上での合成ゲノム展開を実施することが可能となった。