

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 年度採択研究者

2021 年度 年次報告書
------------------

山田 亮祐

大阪府立大学大学院工学研究科  
准教授

有用物質生産を志向した機械学習支援ゲノムデザイン

## § 1. 研究成果の概要

微生物を用いることで、グルコースを、解糖経路によりピルビン酸へと変換した後に、D-乳酸などの様々な目的物質を生産可能である。今年度は、酵母における解糖系関連酵素遺伝子の発現量と、D-乳酸生産性との関係を示すデータを用い、解糖系関連酵素遺伝子の発現量から D-乳酸生産性を予測することが可能な機械学習モデルを構築することを目指した。

機械学習モデルの構築において、約 3000 の構造を検討し、決定係数が 0.677 となる高い精度を示すモデルを見出した。本研究により、初めて機械学習の技術により、解糖系関連酵素の発現量から D-乳酸生産性を予測可能であることが示唆された。

また、微生物を用いることで、グルコースを、解糖経路によりピルビン酸へと変換した後に、更にメバロン酸経路を経由して、 $\beta$ -カロテンなどの種々の高付加価値な二次代謝産物を生産可能である。そこで今年度は、メバロン酸経路を経て生産される  $\beta$ -カロテンをモデル化合物とした、 $\beta$ -カロテン生産酵母ライブラリーを作製し、メバロン酸経路関連遺伝子の発現量から、 $\beta$ -カロテン生産性を予測するために用いる機械学習データを収集することを目指した。更に、収集したデータを用い、メバロン酸経路関連酵素遺伝子の発現量から  $\beta$ -カロテン生産性を予測する機械学習モデルを構築した。

$\beta$ -カロテン生産酵母にメバロン酸経路改変プラスミドライブラリーを導入することで、 $\beta$ -カロテン生産・メバロン酸経路改変酵母ライブラリーを作製した。この内 45 株のメバロン酸経路関連酵素遺伝子の発現量と  $\beta$ -カロテン生産性に関するデータを用い、機械学習モデルを構築した。約 3600 の構造を検討し、高い精度(決定係数:0.752)を示す機械学習モデルを見出した。本研究により、初めて機械学習の技術により、メバロン酸経路関連酵素遺伝子の発現量から  $\beta$ -カロテン生産性を予測可能であることが示唆された。