

## 「ゲノム・メタゲノムリソースからの新規ゲノム編集ツールの探査」

慶應義塾大学大学院 政策・メディア研究科 修士課程 2 年

先端生命科学 (BI)

森 秀人

### 研究概要

原核生物において反復配列は、ゲノム進化、遺伝子調節、免疫機能を含む様々な生物学的事象に関与している。反復配列の複製過程における DNA ポリメラーゼのスリッページ（横すべり）は、反復単位の挿入、欠失を誘導し、ゲノムの構造を変えるだけでなく、遺伝子およびプロモータ内の反復配列で起こると、対象の遺伝子の機能や発現を大きく変化させうる。こうした変異は、多くの原核生物において環境応答や宿主の免疫回避機構として利用されている。また、原核生物から発見された周期的な反復配列の中にはゲノム編集技術として利用されているものもあり、それらは 主に(1) タンデムな反復タンパク質配列、または (2) 短い Spacer 配列を周期的に挟んだインタースペース性の反復配列 CRISPR (clustered regularly interspaced short palindromic repeat) に分類することができる。(1)の例である TALE (transcriptional activator-like effector) は *Xanthomonas* から発見されたタンデムな反復配列を含んだタンパク質であり、反復モチーフ内の可変アミノ酸残基の組み合わせが特定の塩基を認識している。(2)の例として、CRISPR-Cas システムは原核生物において広く確認されている免疫機構であり、Repeat 配列間の領域に外敵のゲノム配列が周期的に保存されていく。このように、周期的な反復配列を探索することは、生物学的に重要な発見及び、新規ゲノム編集ツール候補の発見・開発につながる可能性があるにもかかわらず、これまでゲノム・メタゲノムリソースからの周期的な反復配列を高スループットに検出可能なソフトウェアは存在しなかった。本研究では、大規模なゲノムリソースから網羅的に周期的反復配列の検出することができるソフトウェア SPADE を開発し、原核生物における周期的反復配列の生物学的機能の推定及び、新規ゲノム編集技術として応用可能な反復配列候補の探索を行った。SPADE は 7,006 個の原核生物ゲノムからの未知のものを含む多くの周期的反復を網羅的に捕捉した。SPADE の性能は、一般的に利用されている CRISPR 予測ソフトウェアやタンデムプロテインリピートの予測ソフトウェアよりも優れていた。また、新たに検出されたタンパク質反復配列には、いくつかの新しいゲノム編集モジュール候補となり得る配列が含まれていた。遺伝子間の短いタンデムな DNA 反復配列は、遺伝子の進化およびゲノム拡張モデルを示唆した。また、異なる CRISPR のサブタイプが、反復配列モチーフのみによって分類可能であることを明らかにすると同時に、本結果を利用し、「非反復」の CRISPR 配列 (primordium CRISPR) の存在を示した。

今後、2017年5月に白鳳丸での公開にてサンプリングした海底の微生物のメタゲノムリソース等に対して、実際に SPADE を用いた探索を行い新規ゲノム編集ツールの探索をしていく予定である。

※国際論文誌投稿前のため、本研究に関する詳細な図表の報告を控えさせていただきます。ご理解よろしくお願ひ致します。