令和　元年度　　京都大学化学研究所　スーパーコンピュータシステム　利用報告書

生体分子情報データベースの開発とその応用
Development of Databases for Biomolecular Information and its Application

情報・システム研究機構　データサイエンス共同利用基盤施設
ライフサイエンス統合データベースセンター　　五斗　進

研究成果概要

本研究では、京都大学化学研究所スーパーコンピュータシステムを利用し、生体分子情報データベースおよびバイオインフォマティクス技術の開発に取り組み、その成果をゲノムネット（http://www.genome.jp/）から広く公開している。2019 年度は 2018 年度に引き続き、ゲノムネットサービスとして公開してきたデータベースと解析ツールに対し、改良を加えるとともに、ネットワークや表現型の観点から遺伝子機能推定などの応用に結びつけることを目指し、以下の成果を得た。

1. 2018 年度に基盤整備を行った、微生物の代謝機能類似性と培養培地類似性の相関を調べるためのシステムについて、培養培地データベースの拡張を行った。具体的には、理化学研究所の Japan Collection of Microorganisms と製品評価技術基盤機構の NITE Biological Resource Center が提供する 500 以上の菌株の培地情報を Resource Description Framework でデータベース化した。さらに、生物種のリストから培地を比較するためのインタフェースを構築し、KEGG のパスウェイモジュールなど代謝機能類似性と比較できるようにした（ライフサイエンス統合データベースセンターの川島秀一特任准教授、守屋勇樹特任助教らとの共同研究）。
2. KEGG で定義されたオーソログごとにプロファイル隠れマルコフモデルを定義し、データベース化した。このデータベースを用いた遺伝子機能を推定する仕組みとして、配列から KEGG のオーソログを推定するシステム KofamKOALA を開発した（京都大学化学研究所の荒巻拓哉氏、緒方博之教授らとの共同研究）。

発表論文（謝辞あり）

1. Sakurai, T., Nishiyama, H., Nagai, T, Goto, S., Ogata, H. and Kudo, M.; Deficiency of Gankyrin in the small intestine is associated with augmented colitis accompanied by altered bacterial composition of intestinal microbiota. *BMC Gastroenterology*, **20**:12 (2020).
2. Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S. and Ogata, H.; KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*, in press (2019).