

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。2018 年度は、2017 年度までに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、引き続き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た。具体的には、微生物の表現型として生育環境を反映する培養培地情報と代謝・生理機能を反映するパスウェイモジュール情報とに着目し、両者を結びつけるための基盤開発を行った。まず、嫌気性細菌 50 種、好気性細菌 50 種を幅広い系統分類群から選定し、KEGG MODULE を用いて代謝・生理機能を定量化する Module Completion Ratio をゲノム情報から計算し、微生物間の代謝機能類似性を計算した。MCR をもとに主成分分析を行ったところ、第 2 主軸により嫌気性と好気性を大まかに分類できることを確認した。次に、選定した嫌気性細菌と好気性細菌を配布しているバイオリソースセンターが提供するウェブサイトや文献から培養培地情報を抽出してデータベース化した。この情報を培地成分によるベクトルで表現し、クラスタリングすることによって、微生物間の培養培地類似性を計算した。現在、その結果を代謝・生理機能による主成分分析結果と対応付け、代謝機能類似性と培地類似性の相関について解釈を行っているところである。今後、その結果を用いてゲノムデータから培地情報を予測するシステムを構築するために、更に多くの微生物に対して培地情報や表現型の情報をデータベース化する予定である。

発表論文(謝辞あり)

1. Arai, W., Taniguchi, T., Goto, S., Moriya, Y., Uehara, H., Takemoto, K., Ogata, H. and Takami, H.; MAPLE 2.3.0: An improved system for evaluating the functionomes of genomes and metagenomes. *Biosci. Biotechnol. Biochem.*, **82**:1515-1517 (2018).
2. Yoshida, T., Nishimura, Y., Watai, H., Haruki, N., Morimoto, D., Kaneko, H., Honda, T., Yamamoto, K., Hingamp, P., Sako, Y., Goto, S. and Ogata, H.; Locality and diel cycling of viral production revealed by a 24 h time course cross-omics analysis in a coastal region of Japan. *ISME J.*, **12**:1287-1295 (2018).