

生体分子情報データベースの開発

Development of Databases for Biomolecular Information

京都大学化学研究所バイオインフォマティクスセンター化学生命科学 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/>) として広く公開している。平成 27 年度は、以下のデータベースおよびツールの開発と改良を行った。

- Virus-Host Database: RefSeq にゲノムが登録されているウイルスの宿主情報を RefSeq や UniProt から抽出するとともに、マニュアルでアノテーションしたデータベースを開発し、公開した (<http://www.genome.jp/virushostdb/>)。
- MGENES: Tara Oceans の 243 サンプルのアノテーションを公開し、機能分類情報のパイチャート表示を実装した (http://www.genome.jp/kegg/catalog/org_list3.html)。
- Ortholog Cluster (OC): KEGG GENES 4,060 生物種分のオーソログクラスターを計算し、更新した (<http://www.genome.jp/tools/oc/>)。
- RAxML / FastTree: 系統樹作成ツールをウェブ上の入力から利用できるようにし、公開した (<http://www.genome.jp/tools/raxml/>, <http://www.genome.jp/tools/fasttree/>)。
- KAAS: ゲノムの機能アノテーションツールにおいて、高速ホモロジー検索システムである GhostX と GhostZ を選択できるように改良した (<http://www.genome.jp/tools/kaas/>)。
- MAPLE: メタゲノム機能評価ツールにおいて、機能モジュールの評価方法を改良した版を公開した (<http://www.genome.jp/tools/maple/>)。
- E-zyyme2: 平成 26 年度に公開した酵素遺伝子予測システムに関する論文を発表した (<http://www.genome.jp/tools/e-zyyme/>) [1]。

発表論文(謝辞あり)

1. Moriya, Y., Yamada, T., Okuda, S., Nakagawa, Z., Kotera, M., Tokimatsu, T., Kanehisa, M. and Goto, S.; Identification of enzyme genes using chemical structure alignments of substrate-product pairs. *J. Chem. Inform. Modeling.* (in press) 2016.
DOI:10.1021/acs.jcim.5b00216

謝辞: 上記のデータベースとツールは以下の方々との共同研究で開発されています。緒方博之、Romain Blanc-Mathieu、三原知子、西村陽介、清水祐吾、吉川元貴、西山拓輝、Pascal Hingamp、金久實、中川善一(京大化研)、上原英也、大久保宏一(日本 SGI)、守屋勇樹、時松敏明(DBCLS)、山田拓司、小寺正明(東工大)、奥田修二郎(新潟大)、高見英人、荒井渉(JAMSTEC)、谷口丈晃(三菱総研)、竹本和弘(九工大)。