

令和 2 年 6 月 10 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K07260

研究課題名(和文)大規模計測データからの汎用的な天然物生成経路予測

研究課題名(英文)Generic prediction of natural product biosynthetic pathways from large-scale measurement data

研究代表者

小寺 正明 (Kotera, Masaaki)

東京大学・大学院工学系研究科(工学部)・准教授

研究者番号：90643669

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：本研究では、標的分子の構造を、生成経路の出発物質に対応する部分構造である生成単位(BU)に自動的に分解する予測ワークフローを開発した。本ワークフローでは、出発物質、主要中間体、およびそれらの誘導体を含む生成単位ライブラリ(BUL)を使用する。KEGG PATHWAY から出発物質を取得し、765個のBUをBULに登録した。次に、これらのBUの組み合わせを最適化した。提案ワークフローの性能を評価するために、KEGG二次代謝マップに含まれる943個の分子を用いた結果、約95.8%の903分子が正しく分解された。また、90.7%の分子については、1分以内に計算を終えることができた。

研究成果の学術的意義や社会的意義

本ワークフローは、正しさと計算時間の両面で天然物の効率的な分解を可能にしました。また、ユーザーが計算結果を理解しやすいように、BNに対応する部分構造を自動的に色分けして出力します。利用者は、出発分子を事前に指定する必要がなく、データベースにない分子であっても、任意のターゲット分子を入力することができます。このワークフローは、天然物の生成の理解や予測に大いに役立つと考えています。

研究成果の概要(英文)：In this study, we developed a predictive workflow named the Metabolic Disassembler that automatically disassembles the target molecule structure into relevant biosynthetic units (BUs), which are the substructures that correspond to the starting materials in the biosynthesis pathway. This workflow uses a biosynthetic unit library (BUL), which contains starting materials, key intermediates, and their derivatives. We obtained the starting materials from the KEGG PATHWAY database, and 765 BUs were registered in the BUL. We then examined the proposed workflow to optimize the combination of the BUs. To evaluate the performance of the proposed Metabolic Disassembler workflow, we used 943 molecules that are included in the secondary metabolism maps of KEGG PATHWAY. About 95.8% of them (903 molecules) were correctly disassembled by our proposed workflow. In addition, for 90.7% of molecules, our workflow finished the calculation within one minute.

研究分野：化学情報学

キーワード：代謝化合物

1. 研究開始当初の背景

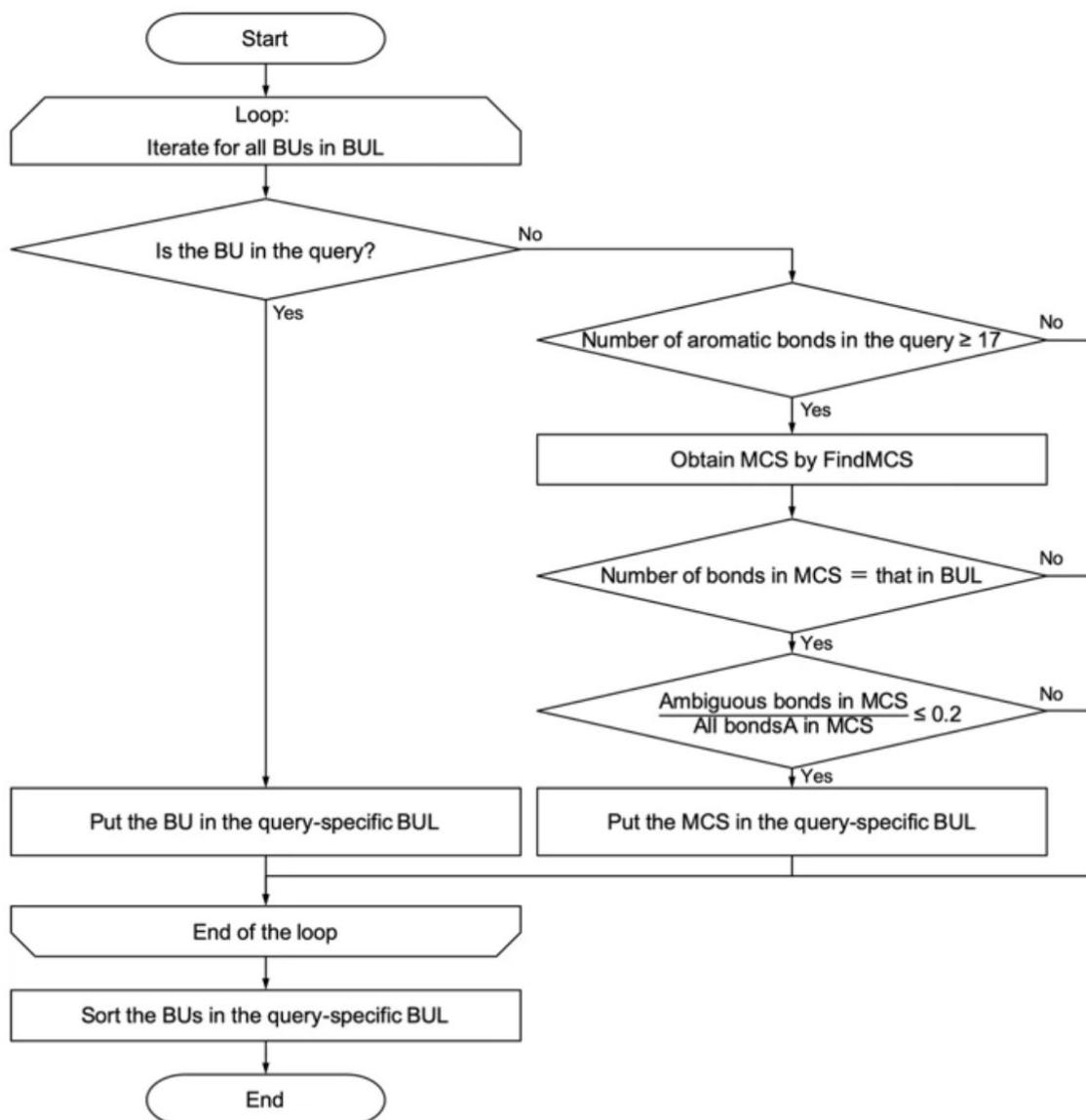
天然物は医薬品をはじめとする様々な機能性材料の源であり、その生合成経路を理解することは、合成生物学的アプローチにより、その効果的な生産に役立つ情報を提供することができる。これまでに、化学構造から生合成経路を逆合成的に予測しようとする研究が数多く行われてきたが、目的の分子から出発物質に到達しないまま計算が終了してしまうことがあった。

2. 研究の目的

この問題を解決するために、適切な出発物質を見つける方法を開発した。

3. 研究の方法

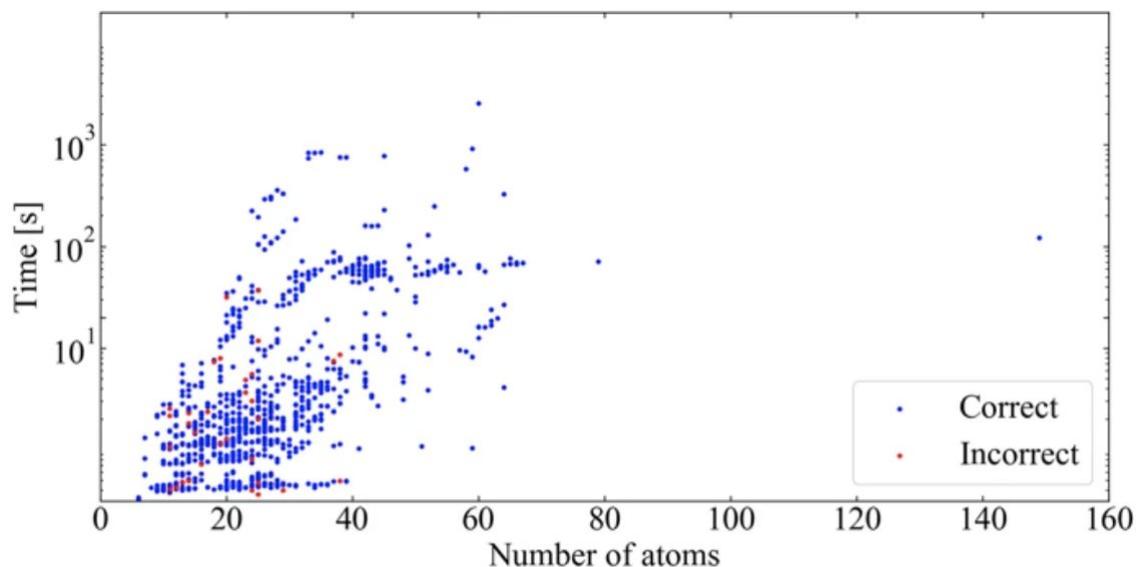
本研究では、標的分子の構造を、生合成経路の出発物質に対応する部分構造である生合成単位 (BU) に自動的に分解する予測ワークフロー「メタボリック・ディスアセンブラ」を開発した。本ワークフローでは、出発物質、主要中間体、およびそれらの誘導体を含む生合成単位ライブラリ (BUL) を使用した。



The flowchart for generating a query-specific biosynthetic unit library (BUL)

4. 研究成果

KEGG PATHWAYデータベースから出発物質を取得し、765個のBUをBULに登録した。次に、提案したワークフローを用いて、BUの組み合わせを最適化するための検討を行った。提案ワークフローの性能を評価するために、KEGG PATHWAYの二次代謝マップに含まれる943個の分子を用いた。その結果、約95.8%の903分子が正しく分解された。比較のために遺伝的アルゴリズムを用いたワークフローも実施したが、その精度は約 52.0%であった。また、90.7%の分子については、1 分以内に計算を終えることができた。



Plot of computational time for each molecule. The computational time is given using a log scale

メタボリック・ディスアセンブラは、正しさと計算時間の両面で天然物の効率的な分解を可能にした。また、ユーザーが計算結果を理解しやすいように、BNIに対応する部分構造を自動的に色分けして出力する。利用者は、出発分子を事前に指定する必要がなく、データベースにない分子であっても、任意のターゲット分子を入力することができる。このワークフローは、天然物の生合成の理解や予測に大いに役立つと考えている。

発表文献

[1] Amano, K., Matsumoto, T., Tanaka, K. et al. Metabolic disassembler for understanding and predicting the biosynthetic units of natural products. *BMC Bioinformatics* 20, 728 (2019). <https://doi.org/10.1186/s12859-019-3183-9>

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 1件/うちオープンアクセス 0件）

1. 著者名 Grisoni Francesca, Neuhaus Claudia S., Hishinuma Miyabi, Gabernet Gisela, Hiss Jan A., Kotera Masaaki, Schneider Gisbert	4. 巻 25
2. 論文標題 De novo design of anticancer peptides by ensemble artificial neural networks	5. 発行年 2019年
3. 雑誌名 Journal of Molecular Modeling	6. 最初と最後の頁 112
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.1007/s00894-019-4007-6	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Sakata Fusako, Kotera Masaaki, Tanaka Kenichi, Nakano Hiroshi, Ukita Masakazu, Shirasawa Raku, Tomiya Shigetaka, Funatsu Kimito	4. 巻 19
2. 論文標題 Developing Novel Descriptors to Predict Physical Properties of Inorganic Compounds from Compositional Formula	5. 発行年 2018年
3. 雑誌名 Journal of Computer Aided Chemistry	6. 最初と最後の頁 7~18
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.2751/jcac.19.7	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計7件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 海東 和麻、小寺 正明、船津 公人
2. 発表標題 機械学習を利用した.CYP 代謝部位検出法の開発
3. 学会等名 第62回日本薬学会関東支部大会
4. 発表年 2018年

1. 発表者名 山本博之、小寺正明
2. 発表標題 構造推定の精度向上を目指した化合物データセットの選択
3. 学会等名 第27回 環境化学討論会
4. 発表年 2018年

1. 発表者名 Miyabi Hishinuma, Francesca Grisoni, Slaudia S. Neuhaus, Gisela Gabernet Garrida, Jan A. Hiss, Masaaki Kotera, Gisbert Schneider
2. 発表標題 Designing Angi-Cancer Peptides by Counterpropagation Artificial Neural Network
3. 学会等名 CBI学会2018年大会
4. 発表年 2018年

1. 発表者名 Kohei Amano, Tsubasa Matsumoto, Miyabi Hishinuma, Kimito Funatsu, Masaaki Kotera
2. 発表標題 Prediction of biosynthetic building blocks in complicated natural products
3. 学会等名 CBI学会2018年大会
4. 発表年 2018年

1. 発表者名 Tsubasa Matsumoto, Kohei Amano, Miyabi Hishinuma, Kimito Funatsu, Masaaki Kotera
2. 発表標題 Pathway prediction of natural products by reverse synthetic analysis
3. 学会等名 CBI学会2018年大会
4. 発表年 2018年

1. 発表者名 Kohei Amano, Tsubasa Matsumoto, Miyabi Hishinuma, Masaaki Kotera
2. 発表標題 Prediction of Biosynthetic Basic Parts of Compounds with Complicated Structures
3. 学会等名 Chem-Bio Informatics Society Annual Meeting 2017
4. 発表年 2017年

1. 発表者名 Tsubasa Matsumoto, Kohei Amano, Miyabi Hishinuma, Masaaki Kotera
2. 発表標題 Prediction of natural biosynthetic pathway by reverse synthetic analysis
3. 学会等名 Chem-Bio Informatics Society Annual Meeting 2017
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----