

Title	代謝反応および生合成経路の予測に資する植物代謝産物の分類とオントロジーの構築
Author(s)	時松, 敏明
Citation	京都大学化学研究所スーパーコンピュータシステム研究成果報告書 (2014), 2013: 41-42
Issue Date	2014-03
URL	http://hdl.handle.net/2433/186397
Right	
Type	Article
Textversion	publisher

代謝反応および生合成経路の予測に資する
植物代謝産物の分類とオントロジーの構築

化学生命科学 時松 敏明

背景と目的

植物は総計100万種以上の代謝化合物を生合成すると推定されており¹⁾、その中には多数の医薬品のシード、工業原材料、栄養的、そして生理的に重要な代謝化合物が存在する。また、植物は生薬の主要な原料でもある。また、これらの代謝化合物のうち、化学構造が決定されている化合物はおよそ6万種程度にすぎず、いまだ構造が明らかになっていない有用代謝化合物がまだまだ存在するものと推測される。このため、植物メタボローム研究はますます重要な研究分野となっており、メタボロームを用いた新規化合物の発見や代謝経路決定の研究が進められている。これらの研究を進めるための基盤データとしては、高品質の代謝化合物データベースや代謝化合物と遺伝子をつなぐ代謝パスウェイデータベースの存在が重要である。当研究室では、長期にわたってこれらのデータベースを開発・維持し、他のデータベースとのリンク情報の提供を含め、統合的なデータベース環境を提供してきた。

しかしながら、植物代謝産物の分野においては、我々のデータベースが現在カバーしている範囲は非常に限定的であり、さらなる情報の充実を必要としている。このため、奈良先端大金谷研究室が開発しているKNAPSAcK データベースの情報をを用いて植物代謝産物の骨格構造分類を行い、骨格分類と代謝経路を対応付けたオントロジーを構築することにより、不足情報をカバーすることを企図した。本年度は、手始めに植物の重要な代謝化合物を多数含み、構造分類がわかりやすい TypeIII ポリケチド合成酵素(PKS) から派生する化合物群の構造による分類を行った。

検討内容

化合物データとしては、KEGG COMPOUND (17,184 化合物(2014年)2g 現在)およびKNAPSAcK(奈良先端大金谷研究室、50,897 化合物(2013年5月現在))を用いた。

まず、文献より TypeIII PKS およびそれらが生合成する化合物骨格を抽出し、TypeIII PKS 派生化合物骨格の階層型オントロジーを作成した。KNAPSAcK の各化合物について、KCF-S を用いたクラスタリングとマニュアルによるキュレーションにより、これらの化合物骨格に分類される化合物の探索を行った。

結果

文献より、16種の TypeIII PKS から派生する53の骨格による化合物分類を規定した(図1、2つの代表構造を決められない化合物分類を含む)。KEGG COMPOUND および KNAPSAcK の化合物をこれらの骨格分類にあてはめたと、現在も継続的に作業中であり暫定的な結果であるが、KEGG COMPOUND で812化合物、KNAPSAcKで10,015化合物が TypeIII PKS 派生の化合物骨格に分類された。53の骨格分類のうち8の骨格分類については、KNAPSAcK にのみ化合物が存在する骨格であった。また、Chalcone synthase (CHS)から派生するFlavonoidがKEGG COMPOUNDで638化合物(78.6%)、KNAPSAcKで8751化合物(87.3%)と TypeIII PKS 派生化合物の大半を占めていた。

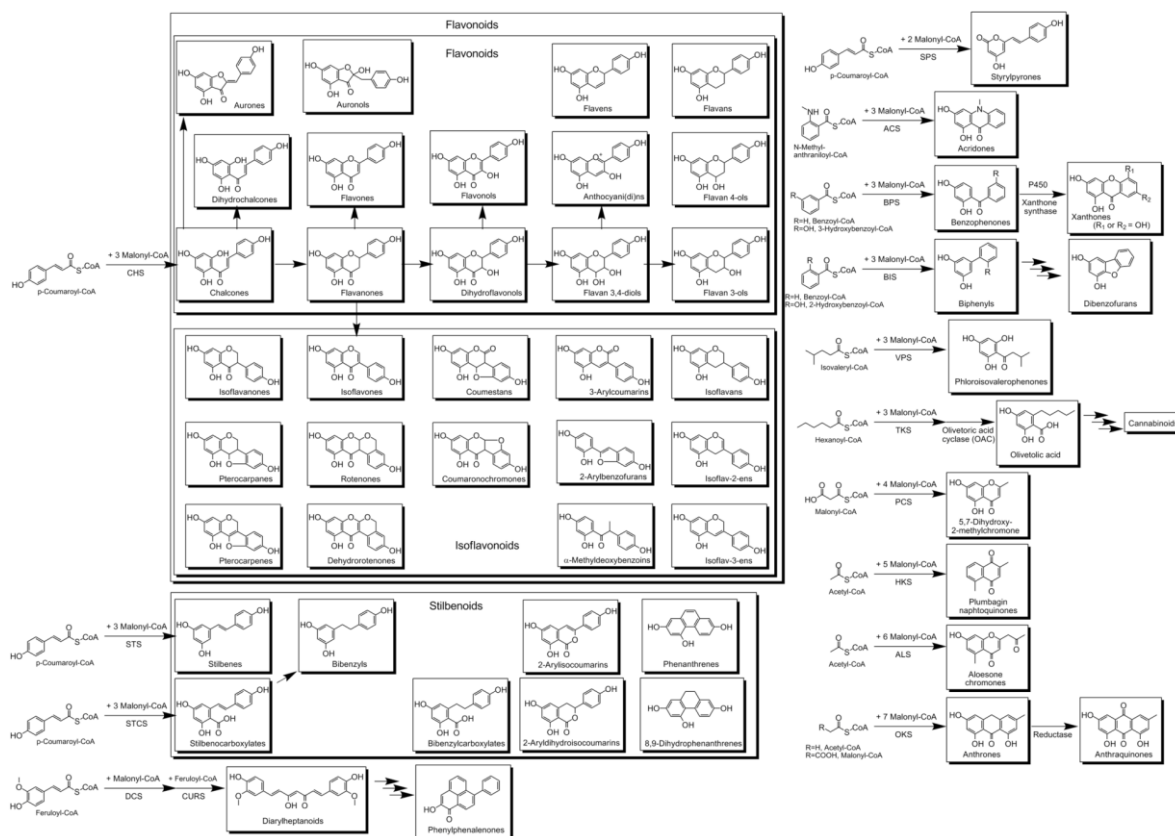


図1. TypeIII PKS およびそこから派生する化合物骨格

考察

今回、KEGG COMPOUND と KNApSACk の 2 つのデータベースの分類を行うことにより、KEGG COMPOUND ではカバーできていない骨格があることが定量的に示された。双方のデータベースの性格上、これは妥当な結果であるので、両方のデータベースを補完利用することにより、情報の充実を図ることが必要であると考えられる。

今回の分類により、代謝経路完全に決定されていない化合物について、部分的な代謝経路情報が付与できたので、今後他の化合物群においても分類を行い植物代謝産物全般にわたり有用情報の提供を可能にすることが必要であると考察される。

参考文献

[1] Afendi, F.M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., Ikeda, S., Takahashi, H., Altaf-Ul-Amin, M., Darusman, L.K., Saito, K., Kanaya, S.; KNApSACk Family Databases: Integrated Metabolite-Plant Species Databases for Multifaceted Plant Research. *Plant Cell Physiol.*, 53, e1 (2012).