

## 論文内容の要旨

博士論文題目 二次代謝物にもとづいた生物分類におけるデータ駆動型科学に関する研究

氏名 LIU KNAG

(論文内容の要旨)

植物、動物、微生物などのさまざまな生物が生合成あるいは蓄積する二次代謝物種の数は、約数百万種類と推定されている。KNApSack Core DBには、5万種の二次代謝物が、それらを生産する生物情報とともに格納されており、10万件の生物種-代謝物情報が格納されている。一次代謝物がエネルギーの産出、細胞の分裂など生命の本的機能に重要な役割を演ずるのに対して、二次代謝物は、光対策、他の生物との戦いにおいて必要とされる、いわゆる個々の生物に固有の生存戦略として機能を有し、生物固有の進化と関連してその生合成系が構築されたと考えられている。そこで、二次代謝物の類似性に基づいた生物分類と、遺伝子の進化系統樹による分類の共通性ならびに特異性について、データ駆動型サイエンスにより検討した。それぞれの生物について報告されている二次代謝物の集合間の比較から、生物種の類似性を検討したところ、代謝物集合におけるD統計量により、食用植物は、遺伝学における進化系統と大きく関連があることが示された。一方、薬用植物についてはこの傾向が極めて低い。また、代謝物集合における生物間の類似性を考慮した、それぞれの生物が所有すべき代謝物の予測における方法論を確立した。本研究では、ゲノムからではなく、メタボロミクスから生物進化を解明しようとする新規方法論ならびに生物の特異性を評価することが可能となった。

氏 名	LIU KANG
-----	----------

(論文審査結果の要旨)

平成29年12月22日に開催した公聴会の結果を参考に、平成30年2月9日に本博士論文の審査を実施した。

以下に述べる通り、本博士論文は、本学位申請者が情報科学のデータサイエンス分野で研究開発活動を続けていくために必要十分な素養を備えていることを示すものである。

LIU KANG は、二次代謝物の類似性に基づいた生物分類と、遺伝子の進化系統樹による分類の共通性ならびに特異性について、データ駆動型サイエンスにより解明した。

本論文は、情報科学とバイオインフォマティクスの境界領域をデータサイエンスによりつなぐメタボロミクス発展に大きく貢献するものである。よって、本論文は、博士（工学）の学位論文として十分な価値があるものと認める。