

(続紙 1)

京都大学	博士 (理 学)	氏名	守屋 勇樹
論文題目	ゲノムからのパスウェイ推定の為のバイオインフォマティクス研究		
(論文内容の要旨)			
<p>代謝系やシグナル伝達系などの生体内分子間相互作用ネットワーク、いわゆるパスウェイを俯瞰することは、様々な分子生物学的解析の基盤であり、生命現象を理解する上で重要な役割を担っている。そのため、KEGG、BioCyc、Reactome など多くのパスウェイデータベースが構築され、様々な生物種のパスウェイが計算機上で再現されている。例えば KEGG では、ゲノムが決定された生物種の配列データを蓄積し、各遺伝子配列から機能をアノテーションすることによって、各生物種のパスウェイを構築している。しかし、一般にパスウェイの構築には多くの専門知識と手作業による精査が必要であり、ゲノム配列の決定と、その生物の持つパスウェイの再現を直接結びつけることは容易ではない。さらに、近年の DNA シークエンス技術の発達により、大量の配列データが蓄積されてきたため、配列データからのパスウェイの自動推定が求められるようになった。本論文では3つ方向からこの問題の解決を目指している。</p> <p>第1は遺伝子機能の自動アノテーションに基づくパスウェイ推定である。新規にゲノム配列が決定された生物種の各遺伝子のオーソログ遺伝子をデータベースから探索し、その機能を割り当てた上で、リファレンスとなるパスウェイにその情報をマッピングすることによってパスウェイを自動的に推定する方法論を開発した。</p> <p>第2は未知代謝パスウェイの予測である。生物界に存在する全ての代謝パスウェイを表現するにはパスウェイデータベースの整備は未だ不十分であり、合成・分解経路が判明していない代謝化合物が多く存在する。そのため、データベースに登録されている既知の反応パターンを利用して、新規に合成・分解経路を予測することにより、データベースには登録されていない代謝パスウェイを推定する方法論を開発した。</p> <p>第3は基質と生成物の構造を用いた代謝酵素の予測である。パスウェイデータベースには代謝活性は確認されているものの、触媒酵素をコードする遺伝子が同定されていないものが多数ある。また、推定されたパスウェイ中には酵素をコードする遺伝子の同定が必要なものもある。これらの代謝反応を触媒する酵素遺伝子を予測するために、データベースに登録されている反応パターンを拡張し、代謝反応の類似性をより詳細に計算できるようにした。また類似反応を代謝する遺伝子のオーソログ及びパラログを探索することにより候補遺伝子の予測を可能にした。</p> <p>これらの研究により、パスウェイ推定の自動化が大きく進み、様々な生物の生命現象の解析に利用できるようになった。また予測されたパスウェイを実験にフィードバックすることで、より正しいパスウェイを求めるための手がかりとして用いることが可能であり、さらなる生命現象の理解への貢献が期待される。</p>			

(続紙 2)

(論文審査の結果の要旨)

申請者は、ゲノム情報や反応情報から代謝系を始めとする生体内のパスウェイを推定するための複数の方法論を開発するとともに、遺伝子機能の予測に結び付けるといふ、ユニークかつ有用な研究を行った。論文では、パスウェイ推定を(1)ゲノム中の遺伝子セットからその生物が持つパスウェイを推定する問題、(2)新規に同定された代謝化合物を合成または分解するパスウェイを推定する問題、(3)酵素活性のみが報告されており、その酵素をコードする遺伝子が未知の遺伝子を同定することによって、パスウェイ中の酵素反応と遺伝子を対応付ける問題の3つに分けて研究し、それぞれについて新しい方法論を開発・評価している。さらに、ウェブ上で誰もが利用できるように実装し公開している。

第1の問題では、ゲノムが決定された生物種の遺伝子配列情報を用いて機能予測とパスウェイ推定を行う計算機的手法を開発し、手作業で構築されたパスウェイをどの程度再現できるかによって評価している。近縁種の多くがデータベースに登録されている細菌では精度が高い一方、データベース登録種が少ない植物などでは精度が比較的低くなるが、データベースに登録される生物種が年々増えているため精度は向上している。また、本手法をウェブ上で利用できるように実装したシステムKAASは現在までに1,000以上の解析に利用されている。

第2の問題では、データベースに登録されている反応パターンを利用して新規反応経路を計算する手法を開発している。特に、細菌が環境物質を分解する経路と植物が二次代謝産物を合成する経路に特徴的に利用される反応に着目し、それらの経路を70~80%の精度で予測できるようにしている。本手法もウェブ上で利用できるように実装し、PathPredシステムとして公開している。

第3の問題では、データベースに登録されている反応パターンを拡張し、反応の類似度を定義する方法を開発している。この反応類似度から、既知の反応と酵素遺伝子との対応関係を結び付けることによって、新規反応に対応する酵素遺伝子を探索する方法を開発している。それを既知の情報を用いて評価し、妥当な結果が得られる指標を提案している。その指標を用いると、現在、対応する酵素遺伝子が見つかっていない約4,000の反応のうち約4割にあたる1,600強の反応に対して何らかの候補が割り当てられることを明らかにしている。さらに、その候補のうち4例について文献調査により検証し、手法の有効性も示している。

申請者の研究はいずれも生物学的な知見と計算機的な方法論を統合した上で新たな知見を発見するための方法を提供するという重要なものである。特に、ゲノム・遺伝子・タンパク質の機能推定に反応という観点を入れた上でパスウェイを推定するアプローチは全く新しいものである。また、現在の生物学には欠かせないオミックスデータの機能解釈を進めるための基礎を提供するため、生命機能の理解へ大きく貢献すると考えられる。よって、本論文は博士(理学)の学位論文として価値あるものと認める。また、平成29年1月17日に論文内容とそれに関連した事項について口頭試問を行った。その結果、合格と認めた。

要旨公表可能日： 平成 年 月 日以降