

戦略的国際科学技術協力推進事業（日本－英国 BBSRC 研究交流）

1. 研究課題名：「From text to pathways: text mining techniques for reconstructing signalling pathways」
2. 研究期間：平成 21 年 4 月～平成 23 年 3 月
3. 支援額： 総額 19,028,000 円
4. 主な参加研究者名：

日本側（研究代表者を含め 6 名までを記載）

	氏名	所属	役職
研究代表者	北野 宏明	システム・バイオロジー研究機構	会長
研究者	辻井 潤一	東京大学大学院	教授
研究者	松崎 拓也	東京大学大学院	助教
研究者	松岡 由希子	システム・バイオロジー研究機構	研究員
研究者	Samik Ghosh	システム・バイオロジー研究機構	研究員
研究者	Hwajung Kang	システム・バイオロジー研究機構	研究員
参加研究者 のべ 7 名			

英国側（研究代表者を含め 6 名までを記載）

	氏名	所属	役職
研究代表者	Sophia Ananiadou	University of Manchester, Computer Science	Professor
研究者	Michael White	University of Liverpool, School of Biological Sciences	Professor
研究者	Dean Jackson	University of Manchester, Life Science	Senior lecturer
研究者	Makoto Miwa	University of Manchester, Computer Science	Research associate
研究者	Sampo Pyysalo	University of Manchester, Computer Science	Research associate
研究者	Tomoko Ohta	University of Manchester, Computer Science	Research associate
参加研究者 のべ 6 名			

5. 研究・交流の目的

本研究は、生化学的パスウェイの記述とテキストマイニングの技術を統合し、システムバイオロジー研究分野の推進に役立つプラットフォームを開発することを目的とする。

本研究では、マンチェスター大学等と東京大学で研究中のテキストマイニング技術と、システム・バイオロジー研究機構（SBI）が開発中のパスウェイ作成共有技術を融合させたプラットフォームを構築する。

6. 研究・交流の成果

6-1 研究の成果

本研究・交流では、マンチェスター大学等と東京大学で研究中のテキストマイニング技術と、システム・バイオロジー研究機構（SBI）、沖縄大学院大学(OIST)が開発中のパスウェイ構築成共有技術を共有したことで、生化学パスウェイの構築と構築に必要な文献情報のテキストマイニングを統合したプラットフォーム Payao - PathTEXT を構築することができた。構築されたプラットフォームでは、生化学的パスウェイのグラフィカル記法 (SBGN=Systems Biology Graphical Notation) の 1 つである Process Description Diagram 記法から、文献検索の Query を生成し、Medie, FACTA, KLEIO 等各種テキストマイニングエンジンへ Query 結果を一括閲覧可能となった。具体的には、SBML+CellDesigner 独自の Annotation タグが

ら、**Nodes** と **Edges** の情報を抽出し、生化学的パスウェイにおける構造と文献におけるイベント構造の対応付けを行い、**Query** を定義することで、パスウェイからの文献を対象としたイベント検索を可能とした。テキストにおけるイベントの抽出では、生化学的オントロジーを使ったイベント定義なども駆使して、パスウェイとイベントの関連づけ手法を研究することができた。この技術を応用すれば、論文などテキスト知見から、**PPI** などのインタラクションの関連性だけでなく、より詳細のメカニズムをグラフィカルな記法で自動生成することが可能となる。また、パスウェイを生物学の情報統合のプラットフォームと捉えれば、最新論文の情報を半自動的に検索、データベース化していくことが可能となる。これは、大規模パスウェイ構築やキュレーションプロジェクトにおいて特に役立つことは、本研究交流で実施したベータテスト（インフルエンザ複製パスウェイ（日本 **ERATO**）、結核の代謝パスウェイ（インド、**OSDD**）、およびパーキンソン病パスウェイ（ルクセンブルグ大学））構築プロジェクトで実証されている。今後は、**Query** 結果のユーザによる評価システムの構築など、イベント検索結果の精査方法についても研究を進める予定である。また、**Garuda** プロジェクト (<http://www.garuda-alliance.org>) という汎用プラットフォーム構築プロジェクトを通じて、テキストマイニング技術を生化学的パスウェイ構築・キュレーションだけでなく、さらに、生物学全般のネットワーク、モデル化に応用するための基盤を得て、汎用的な利用に道が拓けたといえよう。

6-2 人的交流の成果

日英双方のメンバーは、主に、システムバイオロジー関連の合同ワークショップ等の場を利用した交流の実施した。今回の交流では、数ヶ月に及ぶような長期滞在型ではなかったが、年に数回と頻繁に、かつ2日から4日程度の合宿型の場を設けることで、インテンシブなディスカッションと実際の共同作業を同時に行うことが可能となり、内容の濃い効果的な交流が図れた。また、合同ワークショップ等への参加は、単にプロジェクトメンバー内の交流ではなく、広く関連分野の研究者との交流を可能にしたという点で有意義であった。

今後は、パスウェイ構築、テキストマイニングの統合から、システム・バイオロジーのプラットフォーム構築プロジェクトである **Garuda** プロジェクトに進展、これまでのメンバーの枠を超えた研究開発交流を今後も発展させる予定である。

7. 主な論文発表・特許等（5件以内）

相手国側との共著論文については、その旨を備考欄にご記載ください。

論文 or 特許	・論文の場合： 著者名、タイトル、掲載誌名、巻、号、ページ、発行年 ・特許の場合： 知的財産権の種類、発明等の名称、出願国、出願日、 出願番号、出願人、発明者等	特記事項
論文	Ghosh S, Matsuoka Y, Asai Y, Hsin KY, Kitano H., Software for systems biology: from tools to integrated platforms. Nature Reviews Genetics (2011) doi: 10.1038/nrg3096	
論文	Kitano H, Ghosh S, Matsuoka Y, Social engineering for virtual big science in systems biology, Nat Chem Biol. (2011) 7, 323-6. doi:10.1038/nchembio.574	
論文	Matsuoka, Y., Ghosh, S., Kikuchi, N., Kitano, H., Payao: A Community Platform for SBML Pathway Model Curation. Bioinformatics, 2010 [doi:10.1093/bioinformatics/btq143]	
論文	Kemper B, Matsuzaki T, Matsuoka Y, Tsuruoka Y, Kitano H, Ananiadou S, Tsujii J., PathText: a text mining integrator for biological pathway visualizations. Bioinformatics. 2010 Jun 15;26(12):i374-81.	共著
論文	Ananiadou, S., Pyysalo, S., Tsujii, J. and Kell, D. B. (2010). Event extraction for systems biology by text mining the literature. Trends in Biotechnology, 28(7), 381-390	