

課題: 動的ネットワーク抽出のためのイン・シリコパイプラインの構築

代表研究者: 東京大学 医科学研究所 宮野 悟

1. 当初計画・目標、途中計画変更等あればその旨記載

マクロファージの分化・活性化ネットワークに関する遺伝子発現情報等の大規模・総合的データから動的転写制御ネットワークを抽出し利用する計算戦略をイン・シリコパイプラインとして構築することを目的とする。そのために、まず、状態空間モデルと次元圧縮技術により転写モジュールネットワーク構造を抽出し、さらに、複数ソース情報を利用した数理統計手法により転写制御ネットワーク構造を精緻化する方式を構築する。次に、データ同化技術により時系列データを利用してネットワーク構造へダイナミクスを導入する技術及び文献キュレーション情報を用いて、動的な転写制御ネットワークを抽出するパイプラインを構築する。こうした動的転写制御ネットワークを開発・活用・発展させるためのソフトウェアプラットフォームをソフトウェア Cell Illustrator を用いて利用できるように開発する。これにより、マクロファージの分化・活性化に関して、分化のマスター遺伝子やクリティカルパスウェイのイン・シリコ探索及びシミュレーションによる分化・活性化の再現解析を可能にすることを旨とする。

2. 現状、進捗状況、成果

理化学研究所から提供された時系列データ (CAGE データ、qRT-PCR による遺伝子発現データ、マイクロアレイによる遺伝子発現データ) の詳細なサーベイを行った。その結果、CAGE データについては統計的解析手法の観点からその解析に利用可能な時点数が当初利用可と想定していた時点数よりかなり少ないことが平成 20 年 3 月の時点でわかり (14 時点から 5 時点に減)、精緻な動的ネットワークを推定するには、情報解析手法との合理性を考慮して新たなデータを取得する必要がある、現時点で適用可能な解析技術としてはクラスタリング解析適切であることが明らかとなった。また、qRT-PCR 及び Illumina マイクロアレイデータについても同様の状況 (8 時点) であることが判明した。状態空間モデルをはじめとする動的ネットワーク抽出に関連する技術については、Cell Illustrator 環境で利用可能となるように JAVA コード化を行っており、順調にすすんでいる。データ同化技術についてもそのコード化が進んでいる。また、新たな要素技術適用についても検討した。マクロファージ文献情報からの転写制御・シグナル伝達に関わる反応の抽出作業については、キュレータの訓練が終了し、そのデータベースの作成が開始され、40 編の論文からの 40 の動的パスウェイモデルの作成し、平成 19 年度に計画していた目標を達成することができた。以上、データのサーベイを完了し、イン・シリコパイプラインの設計については今後の技術展開を考慮し、30%の自由度を残して完成した。

3. 本年度計画、当初計画・目標に対して達成の見込み

イン・シリコパイプラインの実施及びチューニングを行う。ただし、CAGE データの利用については、合理的なデータが得られた場合に利用できるよう、CAGE のデータモデルを作りそれに基づいた形でネットワーク推定技術を構築する。qRT-PCR 及び Illumina データについても同様に対応する。平成19年度に引き続き、プログラム開発を継続して行うことにより、マイクロファージの分化・活性化に関わるネットワーク情報を統合するプラットフォームの構築、ネットワーク構造の抽出技術、データ同化技術によるネットワーク構造へのダイナミクスの導入技術、そしてこれら開発されたプログラム等をつなぐことにより解析プロセスを構築し、それをチューニングし、動的ネットワーク抽出のためのイン・シリコパイプラインを完成する見込みである。

4. 成果予定、論文発表、特許出願、データ産出状況等

(投稿予定論文)

Masao Nagasaki, Seiya Imoto, Andre Fujita, Teppei Shimamura, Rui Yamaguchi, Osamu Hirose, Ayumu Saito, Emi Ikeda, Kazuko Ueno, Satoru Miyano, Computational pipe line for developing dynamic networks from genome-wide omics data.

Andre Fujita, Masao Nagasaki, Seiya Imoto, Teppei Shimamura, Rui Yamaguchi, Ayumu Saito, Satoru Miyano, Comparison of different gene expression quantification methods.

5. プロジェクトの中での研究の位置づけ

ゲノムネットワークプロジェクトの目標のひとつは、転写制御ネットワークの解析手法を開発し、先導的な解析例を示すことで解析手法の有用性を証明し、さらには、将来のさまざまな生命現象における転写制御ネットワーク解析に資することである。本研究は、この目標を達成するために有用であるソフトウェアプラットフォームを開発するものと位置づけられる。

⇒本年度計画予定（別紙スケジュール表）

