

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年6月5日現在

機関番号：12601

研究種目：基盤研究（B）

研究期間：2010～2012

課題番号：22300099

研究課題名（和文）ネットワーク解析による薬剤応答パスウェイの解析法の研究

研究課題名（英文）Drug-response pathway analysis methods based on network analysis

研究代表者

宮野 悟（MIYANO SATORU）

東京大学・医科学研究所・教授

研究者番号：50128104

研究成果の概要（和文）：薬剤にตอบสนองして動く遺伝子発現の時系列データやケース・コントロールデータから、ネットワークモデルを抽象的レベルで推定する計算的・統計的方法を開発した。また、生体分子間の相互作用に関する文献情報から効率よく動的モデルを構築し、その妥当性を体系的に検証するための知識処理技術を開発した。前者にはベイジアンネットワークと状態空間モデルを用い、後者には拡張ペトリネットをネットワークモデルとして用いた。

研究成果の概要（英文）：We developed computational/statistical methods for inferring abstract network models from time-course or case-control gene expression data of drug-responses. We also developed a series of knowledge processing technologies for efficiently constructing and systematically validating dynamic models from the literature information on biomolecular interactions. Bayesian network and state space model are employed for the former and an extended Petri net is used for the latter.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	2,700,000	810,000	3,510,000
2011年度	2,000,000	600,000	2,600,000
2012年度	2,000,000	600,000	2,600,000
年度			
年度			
総計	6,700,000	2,010,000	8,710,000

研究分野：情報学

科研費の分科・細目：情報学・生体生命情報学

キーワード：遺伝子ネットワーク，薬剤応答パスウェイ，動的モデル，シミュレーション，パスウェイ解析，ベイジアンネットワーク，ペトリネット，状態空間モデル

1. 研究開始当初の背景

各種の先端計測・ハイスループット計測技術が駆使され、ゲノム、トランスクリプトーム、プロテオーム、インタラクトームなど生命システムに出現している部品群が網羅的に解析可能となった。そして、それらの部品群がどのようなシステム設計原理のもとで組み合わせられて部分的生物機能を作り出しているか、さらにそれらがシステムとなって

全体としてどのように機能を生み出しているかを解明するために、定量的・動的な解析が大規模に行われている。日本における大規模なものとして、特定領域研究「ゲノム4領域」、ゲノムネットワークプロジェクト、CREST「生命システムの動作原理と基盤技術」などにより革新的な研究が展開されてきたことは周知のとおりである。米国では、系統的なプロジェクト・グラントのほか、9つの

システム生物学の研究センターが逐次設立された。また、NCIでは2009にCenters for Cancer Systems Biologyの公募が行われ、実験生物学と数理モデリング・コンピュータシミュレーションを必須事項にしたセンターが10以上2010年からスタートした。ヨーロッパでもその仕込みが開始された。特に、イギリスでは2005年以降6つのセンターが設立されている。ドイツでは、HepatoSysという肝細胞のシステム生物学プロジェクトが既に3年間のパイロット研究を終了し、2007年から本格展開に入った。

申請者等はこうした研究に関わるなかで(特定領域研究(研究課題:イン・シリコ生命ネットワーク構築のための遺伝子ネットワーク推定とシミュレーション)及びゲノムネットワークプロジェクト(研究課題: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)),「遺伝子ノックダウンや薬剤応答をはじめとする様々のパターンに基づいたマイクロレイ遺伝子発現データから動的・静的ネットワークを抽出する技術」及び「計測データや文献情報から動的なパスウェイモデルを構築し知識化してシミュレートする技術(ソフトウェアCell Illustratorの開発及び生命システムのためのデータ同化技術)」などで世界をリードする研究を行ってきた。前述の欧米での研究においても、生命をシステムとして理解するため基盤情報技術が開発されているが、個々の要素技術開発にとどまっているものや、システムとして統合していても特定の細胞や計測技術に特化されており、汎用的に実用レベルで利用するには困難があるものがほとんどであった。一方、申請者らは、これまでゲノム特定等の研究で開発された要素技術群をシステムとして統合する形で、ゲノムネットワークプロジェクトにおいて、状態空間モデル及び非線形回帰を導入したダイナミックベイジアンネットワークを用いた時系列データからの動的ネットワーク構造の抽出プロセス、文献キュレーションによる動的パスウェイデータベース、データ同化による動的モデルの精緻化プロセスをXML形式のCell System Markup Language形式で統一し、一連のプロセスをグラフィカルにコンピュータ上で行うことができるパイプラインソフトウェアを構築してきた。こうした貢献により、一貫通貫的に、転写制御などの動的なネットワークをデータから抽出し、可視化し、シミュレーション解析することが可能になり、薬剤応答のシステム解析に展開できる情報技術が準備できた。

2. 研究の目的

薬のターゲットタンパク質が解明されていても、それに薬として機能すべき化合物が、

他の生体内分子に影響することが考えられ、それが毒性となったり、副作用として現れたりする。本研究では、抗がん剤などの薬剤がパスウェイのどこに、いつ、どれくらい、影響を与えているかを解析するために、薬剤に応答して動く生体分子の時系列データやケース・コントロールデータから、ネットワークを抽象的レベルで推定、併せて、その妥当性、仮説生成、及びその検証を体系的に行うための知識処理技術を開発することにより、薬剤応答パスウェイを解析する方法を開発することを目的とする。

3. 研究の方法

次の2項目の研究を行った。

(1) 遺伝子発現プロファイルデータからのシステム推定法を開発する。これにより薬剤などの細胞の内外からの刺激に応答して動く生体分子の時系列データやケース・コントロールデータと文献やデータベースに蓄積されているパスウェイに関する知識から、その細胞の中で機能している動的・静的ネットワークを抽象的レベルで推定し、薬剤応答のメカニズムや影響下にある遺伝子群を探索する技術を開発する。

(2) 動的モデルの妥当性、仮説生成、及びその検証を体系的に行うための知識処理技術を開発する。これにより、メカニズムが十分にわかっていないネットワークに対して、どのようなシステム摂動実験を行い、その挙動をどのように観測すればよいかの指針を与えることができる。

以上、データからのアプローチと生物知識情報からのアプローチを統合することにより、薬剤応答パスウェイの解析法を開発する。具体的には、状態空間モデルや非線形回帰ベイジアンネットワークによる時系列データからの動的ネットワーク探索法の応用、ケース・コントロールデータからの有意なパスウェイを推定法の開発、文献情報からの動的ネットワーク情報の抽出技術の形式知化、並びにシステム生物学のための計算論理を構築する。

4. 研究成果

データからのシステム推定法に関して以下の成果を得た。

1. 薬剤などを含む細胞の内外からの刺激に応答して動く生体分子の時系列データからベイジアンネットワークモデルで遺伝子ネットワークを推定する際に、隠れ交絡因子の影響を推定するための統計科学的方法を開発した(⑩)。この隠れ交絡因子は、遺伝子ネットワーク推定の際に計算コストを下げるために外された遺伝子であったり、未知の生体内分子に対応する。この隠れ交絡因子の存在はネットワ

ーク推定の精度に影響したり、ネットワーク解釈を困難にすることがある。本研究で開発した方法は、ベイジアンネットワークによって推定した局所的ネットワーク構造の不整合性に基づいて隠れ交絡因子を推定し、EM アルゴリズムを用いてこの隠れ変数をもったベイジアンネットワークに基づいて理想的なプロファイルを推定する。この方法は、遺伝子ネットワーク推定における欠損遺伝子や未知生体分子を理論的に特徴づけることを行った世界で最初の研究である。

2. DNA チップによる薬剤応答の時系列遺伝子発現プロファイルデータは、遺伝子数と時点数の間に大きなギャップがあり、そのため統計モデルはこの通常規格をはずれた状況を解決しなければならない。申請者等は、これまで、状態空間モデルにおいてこの困難を次元圧縮という方法で解決してきた。本研究では、この方法によりどのようなシステム摂動実験を行い、その挙動をどのように観測すればよいかというデータ計測設計法を提案し、薬の作用機構を推定する方式を開発した (10)。
3. 薬剤応答パスウェイモデルを網羅的にスクリーニングするための統計的計算手法を開発した (5)。この方法は、まず、文献ベースで作ったプロトタイプの動的モデルから自動的に候補となる薬剤応答パスウェイを生成する。次に、データ同化技法を用いて時系列遺伝子発現データから、モデルのパラメータを自動的に推定する。そしてベイズ情報量基準により予測能力を評価し、モデルのランク付けを行う。5つのプロトタイプモデルを用いた計算機実験を行った。これから53のシミュレーションモデルを生成し、開発した手法を、ステロイド系抗炎症剤を投与したラットの肝臓の遺伝子発現データに適用し評価した。その結果、このシステムマティクな手法の有効性を示すことができた。
4. モデル推定方では、ベイジアンネットワークについては、制約条件下で最適なベイズネットワークを探索する技法を開発した (6)。状態空間モデルについては、高並列化でソフトウェアを利用できるように改良した。

知識推論技術に関して以下の成果を得た。

1. 時系列遺伝子発現プロファイルデータか



図 1 : XiP のダウンロードページ

(<http://xip.hgc.jp>)

ら、その変動に関与するパスウェイの活性状態遷移ダイアグラムを構築して解析する新たな方法を開発した (12)。パスウェイのモデル化には高機能ペトリネットを用いた。この方法により、概日周期や薬剤などに応答して機能するキー制御遺伝子やサブネットワークを捉えることができると考えている。

2. 文献ベースで高精度の動的パスウェイモデルを作るにはマニュアルキュレーション (人出による文献読みの作業によりパスウェイ情報を抽出して動的モデルを作る) と大規模なパスウェイのバリデーションが必要である。通常のキュレーションでは、そのパスウェイに関する生物知識をもっているドメインエキスパートが必要であり、モデルのバリデーションとそのフィードバックを繰り返しながらモデルの更新を行うため、そのコストは非常に大きい。そのため、本研究では、キュレーションにおける繰り返しコストを削減するために、CSO (Cell System Ontology) validator を開発した (7, 8)。このツールによりキュレーターとドメインエキスパートとのやり取りが削減され、マクロファージ動的パスウェイモデルのデータベース構築において、より効率よく動的パスウェイモデルを文献ベースで作成することができ、その有効性が示された (9)。
3. 薬剤応答パスウェイを解析するプラットフォームの完成版として XiP を開発した (2)。これは、Java で開発されており、解析フロー作成・実行ツールで、ネットワーク解析など様々な解析コンポーネントを導入した (図 1)。

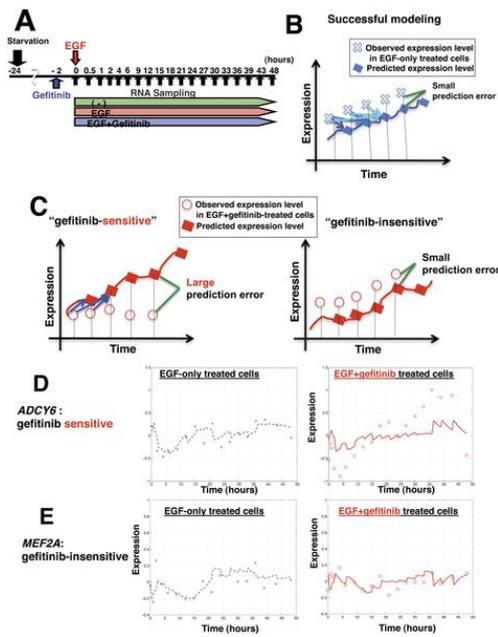


図 2 : 実験データ取得手順と状態空間モデルによるゲフィチニブ応答遺伝子セットの選択法

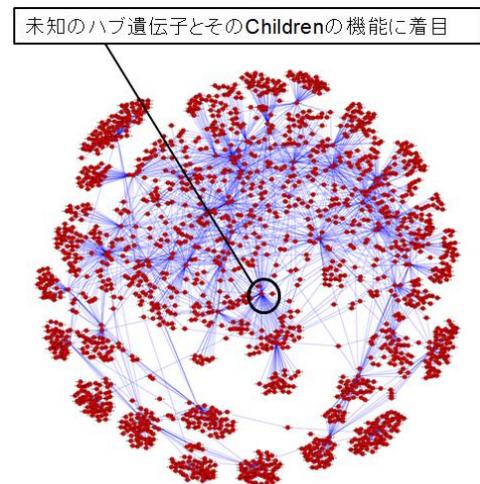


図 3 : メラノーマに関連することがわかっている 45 の転写因子 (ABL1, AKT1, CCNA2, CCNB1, CCNB2, CCND3, CDC16, CDC2, CDC25B, CDC37, CDK2, CDK4, CDK7, CDKN2C, CEBPD, CEBPZ, CHEK1, CTNNB1, ETS1, FOXM1, FOXO3A, GABARAP, HDAC2, CDAC3, HSF2, MAP2K1, MAPK1, MCM2, MITF, NCOR2, NMI, PCNA, PIAS1, PIK3CB, RB1, RBL2, RELA, SKP2, SP1, SP100, STAT1, STAT3, STAT6, TCEA1, TP53) をノックダウンした遺伝子発現プロファイルデータからヒトゲノム解析センターのスパコンで計算した。

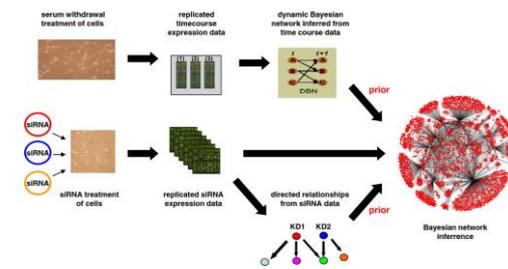


図 4 : 遺伝子ノックダウン遺伝子発現プロファイルデータからのベイジアンネットワーク推定法

以上、これまでに開発してきた方法、特に遺伝子ネットワーク推定を中心におき、統合的に活用することにより、遺伝子発現データと薬剤、及び GEO や NCI など公開されているデータを用いて薬剤がどのようにパスウェイに影響を与えているかを解析した結果、実験による検証を含め以下の一連の成果を得た。これにより本研究で開発した方法論の有効性が確認されたと考える。

- (a) 肺がん細胞株、及び肺上皮細胞に抗がん剤ゲフィチニブを投与した時系列遺伝子発現データから、状態空間モデルと次元圧縮モデルにより予測モデルを推定した。そのモデルを利用し、EGF 刺激により EGFR パスウェイの活性化によって増殖している状態の細胞のどの遺伝子にゲフィチニブが関与しているかをシミュレーションにより選定した結果、139 の遺伝子はその候補となった。がん生物学的な詰めはその系の複雑さのため困難であったが、この 139 遺伝子を用いて、予測が困難とされていたステージ I の肺腺がんの予後予測に、米国 NCI, Duke 大学、国立がん研究センターの肺腺がんのデータで成功した (④) (図 2)。
- (b) ベイジアンネットワークによるスパコン解析で、メラノーマの全遺伝子ネットワークを推定した。ネットワークの未知のハブ遺伝子とその子遺伝子の機能に着目することで、メラノーマにアポトーシスを誘導する 2 つの新規遺伝子 ELMOD1 と TMC01 の発見した (③) (図 3)。データは、メラノーマのがん細胞生物学の観点から 45 の転写因子を選び、siRNA で遺伝子ノックダウンしたメラノーマ細胞株 A375 を使い、遺伝子発現プロファイルデータをとったものを利用した。
- (c) ベイジアンネットワークによる全遺伝子ネットワーク推定により、血管内皮細胞において、遺伝子 Vaslohibin-1 がアポトーシスを誘導するキー遺伝子であることを同定した (①)。図 4 はその推定法の概観である。

5. 主な発表論文等

[雑誌論文] (計 16 件)

- ① Affara M, Sanders D, Araki H, Tamada Y, Dunmore B, Humphreys S, Imoto S, Savoie C, Miyano S, Kuhara S, Print C, Charnock-Jones DS. Vasohibin-1 is identified as a master-regulator of endothelial cell apoptosis using gene network analysis. *BMC Genomics*. 14(1):23, 2013. (査読有)
doi: 10.1186/1471-2164-14-23.
- ② Nagasaki M, Fujita A, Sekiya Y, Saito A, Ikeda E, Li C, Miyano S. XiP: a computational environment to create, extend and share workflows. *Bioinformatics*. 29(1):137-139, 2013. doi: 10.1093/bioinformatics/bts630. (査読有)
- ③ Wang L, Hurley D, Watkins W, Araki H, Tamada Y, Muthukaruppan A, Ranjard L, Derkac E, Imoto S, Miyano S, Crampin E, Print C. Cell cycle gene networks are associated with melanoma prognosis. *PLoS One*. 7(4): e34247, 2012. (査読有)
doi: 10.1371/journal.pone.0034247.
- ④ Yamauchi M, Yamaguchi R, Nakata A, Kohno T, Nagasaki M, Shimamura T, Imoto S, Saito A, Ueno K, Hatanaka Y, Yoshida R, Higuchi T, Nomura M, Beer DG, Yokota J, Miyano S, Gotoh N. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. *PLoS One*. 2012;7(9): e43923. (査読有)
doi: 10.1371/journal.pone.0043923.
- ⑤ Hasegawa T, Yamaguchi R, Nagasaki M, Imoto S, Miyano S. Comprehensive pharmacogenomic pathway screening by data assimilation. *Lecture Notes in Bioinformatics*. 6674: 160-171, 2011. doi: 10.1007/978-3-642-21260-4_18 (査読有)
- ⑥ Imoto S, Kojima K, Perrier E, Tamada Y, Miyano S. Searching optimal Bayesian network structure on constraint search space: super-structure approach. *Lecture Notes in Computer Science*. 6797: 210-218, 2011. (査読有)
doi: 10.1007/978-3-642-25655-4_19
- ⑦ Jeong E, Nagasaki M, Ueno K, Miyano S. Ontology-based instance data validation for high-quality curated biological pathways. *BMC Bioinformatics*. 12(Suppl 1): S8, 2011. doi:10.1186/1471-2105-12-S1-S8 (査読有)
- ⑧ Jeong E, Nagasaki M, Ikeda E, Saito A, Miyano S. CSO validator: improving manual curation workflow for biological pathways. *Bioinformatics*. 27(17): 2471-2472, 2011. doi: 10.1093/bioinformatics/btr395 (査読有)
- ⑨ Nagasaki M, Saito A, Fujita A, Tremmel G, Ueno K, Ikeda E, Jeong E, Miyano S. Systems biology model repository for macrophage pathway simulation. *Bioinformatics*. 27 (11): 1591-1593, 2011. doi: 10.1093/bioinformatics/btr173 (査読有)
- ⑩ Yamaguchi R, Imoto S, Miyano S. Network-based predictions and simulations by biological state space models: Search for drug mode of action. *J. Computer Science and Technology*. 25(1): 131-153, 2010. doi: 10.1007/s11390-010-9311-7 (査読有)
- ⑪ Higashigaki T, Kojima K, Yamaguchi R, Inoue M, Imoto S, Miyano S. Identifying hidden confounders in gene networks by Bayesian networks. *Proc. 10th IEEE Bioinformatics and Bioengineering*. 168-173, 2010. DOI Bookmark: <http://doi.ieeecomputersociety.org/10.1109/BIBE.2010.35> (査読有)
- ⑫ Li C, Nagasaki M, Saito A, Miyano S. Time-dependent structural transformation analysis to high-level Petri net model with active state transition diagram. *BMC Systems Biology*. 4:39, 2010. doi:10.1186/1752-0509-4-39 (査読有)

[学会発表] (計 5 件)

- ① 宮野 悟. スパコンで焙り出すがんのシステムとしての薬剤応答性と多様性. 第 71 回日本癌学会学術総会. 札幌市教育文化会館 4 階 K-7. 2012 年 09 月 19 日.
- ② Imoto S. NetworkProfiler: Uncovering Cancer Heterogeneity in Transcriptome Data. The 4th International Conference of the ERCIM WG on COMPUTING & STATISTICS. 2011 年 12 月 19 日. London, UK
- ③ Yamaguchi R. Identification of dissimilarities in high-dimensional gene regulatory systems by predicted discrepancies from state space model. Joint Meeting of the 2011 Taipei

International Statistical Symposium
and 7th Conference of the Asian
Regional Section of the IASC. 2011
年12月17日. Academia Sinica, Taipei,
Taiwan.

- ④ Miyano S. Systems Biology in
Supercomputing Environment.
International Workshop on Biological
Processes & Petri Nets (BioPPN 2011).
2011年6月20日. Thistle County Hotel,
Newcastle upon Tyne, UK.

6. 研究組織

(1) 研究代表者

宮野 悟 (MIYANO SATORU)
東京大学・医科学研究所・教授
研究者番号：50128104

(2) 研究分担者

井元 清哉 (IMOTO SEIYA)
東京大学・医科学研究所・准教授
研究者番号：10345027
山口 類 (YAMAGUCHI RUI)
東京大学・医科学研究所・講師
研究者番号：90380675
長崎 正朗 (NAGASAKI MASAO)
東京大学・医科学研究所・准教授
研究者番号：90396862
(H22～H23)

(3) 連携研究者

渡辺 賢治 (WATANABE KENJI)
慶應義塾大学・医学部・准教授
研究者番号：70191757
後藤 典子 (GOTOH NORIKO)
東京大学・医科学研究所・特任准教授
研究者番号：10251448