

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 10 日現在

機関番号：10101

研究種目：若手研究(B)

研究期間：2011～2013

課題番号：23710233

研究課題名(和文)多階層ネットワークに基づく遺伝子間の非線形相互作用のモデル化と代謝解析への応用

研究課題名(英文)Integrative modeling of genetic interactions based on intertwined multi-level networks

研究代表者

瀧川 一学(Takigawa, Ichigaku)

北海道大学・創成研究機構・特任助教

研究者番号：10374597

交付決定額(研究期間全体)：(直接経費) 3,400,000円、(間接経費) 1,020,000円

研究成果の概要(和文)：従来、遺伝子は原因因子として個別に機能や現象への関与が研究されてきた。しかし近年では遺伝子はその他の様々な遺伝子や生体内分子との相互作用単位として、特定条件下での込み入った制御を通して対象の疾患・表現型・機能と関わっていると考えられている。本研究課題では遺伝子を階層的に絡み合う様々な分子間ネットワークの構成要素と位置づけ、既知の各分子の定量的情報と各分子ネットワークの既知情報に基づき遺伝子の非線形な相互作用を解析する。特に、現在利用可能になってきた様々なデータを互いに関連づけ階層的な生体内分子ネットワークの理解へ貢献する解析手法開発を行う。

研究成果の概要(英文)：Genes had been considered to be factors that can be analyzed individually, and their effect on functions and phenotypes can be modeled linearly. However, now they are rather considered as non-linearly interacting entities to other genes and other types of biomolecules whose roles and control would be more complicated and context-dependent than we thought. This research regards genes as building blocks of intertwined multi-level networks of various biomolecules, and develops computational methods for integratively analyzing behaviors thorough the networks and for relating quantitative measurements of individual building blocks to the known structures of corresponding networks. Recent studies are producing large data such as transcriptomes, metabolomes, proteomes, and fluxomes, and many databases on biological networks are being developed. The integrative computational approaches thus need to be established to leverage these rich data on various networks and their constituents.

研究分野：情報科学

科研費の分科・細目：ゲノム科学・システムゲノム科学

キーワード：多階層ネットワーク 分子間相互作用 遺伝子 発現制御

1. 研究開始当初の背景

近年「遺伝子」という概念がますます曖昧化している。疾患や表現型の説明因子としての遺伝子は、従来、ある対象に効果を及ぼす単一独立因子として解析されてきた。例えば、ある遺伝子における変異がどの程度対象となる疾患や表現型の変化に効果を及ぼすかを定量的に分析する、といった解析である。しかし、原因因子としての遺伝子は単一効果というより、他の様々な遺伝子や生体内分子との相互作用単位として、特定条件下での込み入った制御を通して対象の疾患・表現型・機能と関わっていると考えられる。このため、現在あまり解明されていない生命現象においては「単一では弱い効果しか持たない複数の因子が多数関与し、加算的に(線形的に)その現象を成立させている」という従来型の見方ではなく、遺伝子の「非線形な相互作用」を様々な分子との相互作用ネットワークの中に位置づけて理解しようとするアプローチが重要視されるようになってきている。

加えて、近年になり、計算機技術とバイオインフォマティクス研究の進展・整備によって、生体内の様々な分子ネットワークに関する既知知見がデータベースとして整備されるようになってきた。生体内には、代謝産物と代謝反応からなる代謝反応ネットワーク、各々の代謝反応酵素をコードする遺伝子のネットワーク、各遺伝子を制御するための各種制御ネットワーク、遺伝子産物としての酵素とその他タンパクやリガンドとの相互作用ネットワーク、など「多階層の」ネットワークが存在し、各々について既存の知見がデータベースに整備されるようになってきている。また、ネットワーク情報だけでなく、各分子に付随する定量的な計測データ、構造情報、疾患関連情報、論文情報なども各種データベースに整備され、関連づけられるようになってきている。

2. 研究の目的

本研究課題では、データベースに整備されつつある生体内分子に関する様々な多階層的网络の既存知見や各種分子の情報(定量的計測データ等)を統合的に利用し、遺伝子ネットワークや遺伝子とその他の生体内分子のネットワーク様の非線形な関係を解析するためのバイオインフォマティクス手法を開発することを目的とする。

特に近年ではマルチオミクス研究も実施されるようになっており、同一細胞に対して、例えば、遺伝子発現(トランスクリプトーム)の定量、タンパク質(プロテオーム)の定量、代謝産物(メタボローム)や代謝フラックス(フラクソーム)の定量など、複数の-omics デ

ータの計測を行うことも行われている。このように多階層ネットワークの各構成要素についての網羅的定量、各層の既知ネットワークの情報、他分子間相互作用ネットワークの情報など、生体内の分子ネットワークを全体的に理解するためのデータが蓄積されはじめてある。このような背景から、これらのデータをただ蓄積し検索するだけでなく、互いに関連づけ全体的理解につなげるための情報解析技術の進展が望まれていると考える。

3. 研究の方法

(1) 代謝反応系におけるネットワーク構造と代謝産物量・酵素遺伝子発現量の解析

具体的な対象として KEGG データベース等に整備されている代謝反応ネットワーク、つまり代謝産物と代謝反応が作るネットワーク、と、各々の代謝反応の触媒酵素をコードしている酵素遺伝子の転写計測を関係づけるための分析手法を研究する。代謝ネットワーク自体はある化学反応の産物が別の科学反応の基質となって連鎖する化学反応ネットワークに過ぎないが、これらの多数の化学反応を細胞内で円滑に運用するために各触媒反応にはその触媒酵素をコードする遺伝子が関わっている。この代謝産物、代謝反応、酵素遺伝子、酵素タンパク質、という各層のネットワークと全体的な挙動の全容は未だに未知な部分が多い。代謝反応は真核生物・原核生物を問わずすべての生物に共通する非常に基礎的な機能であるにも関わらず、これだけ多数の化学反応を細胞内でどのように円滑に制御しているのかがほとんど未知というのは一つには複雑な対象に対する分析手法の不足もあると考える。多階層ネットワークを通じて生じる遺伝子間の非線形性の理解の具体的対象として、酵素遺伝子の転写レベル情報がどのようにその下層の代謝反応ネットワークと関係づけられるのかの理解を目指す。

(2) 遺伝子発現の共発現ネットワークのバイクラスタリングによる解析

複数条件下での遺伝子発現プロファイルの分析において、いくつかの遺伝子がある実験条件(アレイ条件)のみでしか共発現しないことがある。所定の条件以外では共発現しないため遺伝子発現プロファイル全体としては似ていないため通常のペアワイズな類似性解析やグループ化ではこうしたパターンは見逃されてしまう。実際の細胞内では複雑な制御によって特定の条件がそろったとき共制御される現象も知られており、この非線形な遺伝子発現関係をバイクラスタリングという問題として抽出・解析する手法の構築を

行う。バイクラスタリングは遺伝子発現に留まらず同様のプロテオームやメタボローム計測にも適用できる技術であり、単にプロファイルの類似を見るに留まらず、特定の条件下のみでの制御をデータから同定するための枠組として技術研究も期待されている。

(3) ネットワークを取り扱うための情報科学技術の基礎研究

多階層ネットワークや各層の分子の計測量を関係づけることで分子間の非線形な関係を探る、という問題を取り扱う際に付随するネットワーク解析の基礎技術研究も行う。特に、代表者の従来からの研究の展開として、ネットワーク・グラフ構造の分析手法研究、低分子薬剤と標的タンパク質の分子間相互作用ネットワーク解析などを行う。

4. 研究成果

方法で述べた3つの研究課題について以下の成果を得ることができた。今後はこれらの成果をさらに発展させたい。生体内分子およびネットワークに関する既知知見は近年、様々な角度で収集され整備されるようになっており、こうしたさらなる新しいリッチデータを活用して多階層ネットワークが遺伝子間の非線形な相互作用を生む様子の理解に貢献する分析法開発を目指していきたい。

(1) 代謝反応系におけるネットワーク構造と代謝産物量・酵素遺伝子発現量の解析

代謝系における各遺伝子の発現制御、その遺伝子産物の生成量、またその遺伝子産物が触媒する反応の産物の量、の関係は予想以上に複雑であることが近年示されてきた。この関係をより定量的に理解するために遺伝子発現量、タンパク質量、代謝産物量、代謝フラックスなど関係する量を同時に定量する研究も行われている。そこで、E. coli K12 MG1655株のストレス応答に対する遺伝子発現と代謝産物のプロファイルに基づき、代謝関連遺伝子の発現制御と各代謝反応の生成物量変化がどのように代謝ネットワーク構造と関係しているのかを定量的に理解するためのモデル化を行った。(a)代謝ネットワークにおいて発現プロファイルから見て最も連動していると考えられる代謝経路を同定する、(b)その経路上の遺伝子の発現プロファイルに基づきターゲット産物を同定する、という手順で、特定の代謝産物の量の制御に関わっている遺伝子・経路モジュールの同定・解析を行い、論文として報告した。この解析のために、細胞内で活性化されている代謝経路を代謝関連遺伝子の発現量から順位付けして出力する研究代表者が以前開発した計算法の拡張を行った。また、各代謝産

物の生成に関わっている遺伝子群を共制御関係とネットワーク構造に基づき「スコープ」として同定し、系全体の挙動を代表するような「スコープの最小セット」について解析を行った。この結果、局所的制御と同時に、代謝ネットワーク上で離れた関係にある産物まで影響が及ぶケースも発見され、遺伝子間に起こり得る複雑な相互作用の一端を明らかにした。

(2) 遺伝子発現の共発現ネットワークのバイクラスタリングによる解析

遺伝子発現プロファイルの分析において、いくつかの遺伝子がある実験条件(アレイ条件)のみでしか共発現しない場合、遺伝子ごとのプロファイルと比較する通常のクラスタリングでは十分同定することができない。従って遺伝子プロファイルの部分セットの類似性を探索する必要がある。遺伝子とアレイ条件の両方の類似性から同時にクラスタリングを行う問題はバイクラスタリングとして研究されてきたが、部分セットの取り方が限定的であるものが大部分である。本課題ではバイクラスタリング問題を頻出アイテムセット列挙の問題に帰着させることで特定の条件でのバイクラスタを完全列挙する手法を構築した。完全列挙の場合、出力が大きくなってしまいうため、出力パターン 요약のため、各バイクラスタの有意性評価、冗長なバイクラスタのマージ、バイクラスタをネットワークとして可視化する方法も併せて提案した。また、これらの手法をウェブサーバーSiBICとして実装しインターネットを通じて一般に利用できるよう公開した。

(3) ネットワークを取り扱うための情報科学技術の基礎研究

従来からの研究課題として、低分子薬剤/リガンドと標的タンパク質がなす相互作用ネットワークを対象に、機械学習を用いた予測技術およびグラフマイニング技術に関するサーベイ研究を行った。また、併せて、グラフ構造をなすデータに関する統計推論に関して基礎研究を行った。これによりネットワークのトポロジーを統計的に取り扱う上での情報科学手法に関する理解が得られた。この知見に基づいて、関連遺伝子の変異株の発現プロファイルや各種制御ネットワークも統合的に用いて、大腸菌や酵母など微生物モデル生物の生体内ネットワークに遺伝子欠損を与えたときの非線形な挙動(エピスタシス)を理解するため、現在ネットワークの到達可能性(reachability)と信頼性理論に基づく分析を行う手法研究を行っている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者に

は下線)

〔雑誌論文〕(計 6 件)

Ding H, Takigawa I, Mamitsuka H, Zhu S, Similarity-based machine learning methods for predicting drug–target interactions: a brief review. Briefings in Bioinformatics. (2013) 印刷中・先行オンライン出版済(査読有)
DOI: 10.1093/bib/bbt056

Takahashi K, Takigawa I, Mamitsuka H, SiBIC: A web server for generating gene set networks based on biclusters obtained by maximal frequent itemset mining. PLoS One. 2013;8(12) e82890. (査読有)
DOI: 10.1371/journal.pone.0082890

Takigawa I, Tsuda K, Mamitsuka H, An in silico model for interpreting polypharmacology in drug–target networks. In Silico Models for Drug Discovery (Methods in Molecular Biology) 2013; 993: 67-80. (査読有)
DOI: 10.1007/978-1-62703-342-8_5

Takigawa I, Mamitsuka H, Graph mining: procedure, application to drug discovery and recent advances. Drug Discovery Today. 2013; 18(1-2): 50-57. (査読無, Book Chapter)
DOI: 10.1016/j.drudis.2012.07.016

Hancock T, Takigawa I, Mamitsuka H, Identifying pathways of coordinated gene expression. Data Mining for Systems Biology (Methods in Molecular Biology) 2013; 939: 69-85. (査読無, Book Chapter)
DOI: 10.1007/978-1-62703-107-3_7

Hancock T, Wicker N, Takigawa I, Mamitsuka H, Identifying neighborhoods of coordinated gene expression and metabolite profiles. PLoS One, 2012; 7(2): e31345. (査読有)
DOI: 10.1371/journal.pone.0031345

〔学会発表〕(計 3 件)

Takigawa I, Finding structural patterns shared among interacting molecules. 3rd Beilstein Symposium on Glyco-Bioinformatics, Potsdam, Germany. (June 10-14, 2013)

Takigawa I, Identifying paired substructures significantly shared among interacting structure pairs. 2013 International

Workshop on Machine Learning and Applications to Biology (MLAB Sapporo 2013), Sapporo, Japan.
(August 5-6, 2013)

Takigawa I, Learning sparse linear models over subgraph indicators. 2012 Sapporo Workshop on Machine Learning and Applications to Biology (MLAB Sapporo 2012), Sapporo, Japan.
(August 6-7, 2012)

〔その他〕
ホームページ等

SiBIC:
<http://utrecht.kuicr.kyoto-u.ac.jp:8080/miami/faces/index.jsp>

6 . 研究組織

(1)研究代表者

瀧川 一学 (TAKIGAWA, Ichigaku)

北海道大学・創成研究機構・特任助教

研究者番号：10374597