

1M-5

# 相互作用ネットワーク検索によるパスウェイ情報補完の研究

関美緒<sup>†</sup> 瀬々潤<sup>†</sup>

<sup>†</sup>お茶の水女子大学 理学部情報科学科

## 1 はじめに

現在までに様々な種のゲノム配列が読まれ、遺伝子の情報が調べられてきたが、細胞内にどのようなシグナル伝達経路が存在しているかは依然としてほとんど明らかになっていない。既存のパスウェイ [1] は、人手によって整理されており、生体内のパスウェイを網羅するには膨大な労力を要する。また、文献等の整理では未知のパスウェイを推定することは難しい。本研究では、たんぱく質-たんぱく質相互作用ネットワーク情報を利用することで、新たなパスウェイの発見を目指す。具体的には、膨大な相互作用ネットワークから、着目の遺伝子群を繋ぐなるべくエッジが少ない部分グラフの抽出を行うことで、新しいパスウェイの推定を行う。我々のアルゴリズムに DNA 修復に關与する 3 遺伝子 (RFC1, RAD17, RAD53) を入力として、これらを結ぶ部分グラフを抽出したところ、新たなパスウェイ候補となるネットワークを発見した。このネットワークは KEGG [1] に登録されている DNA パスウェイと異なるものの、大きさはほぼ同一の相互作用ネットワークである。

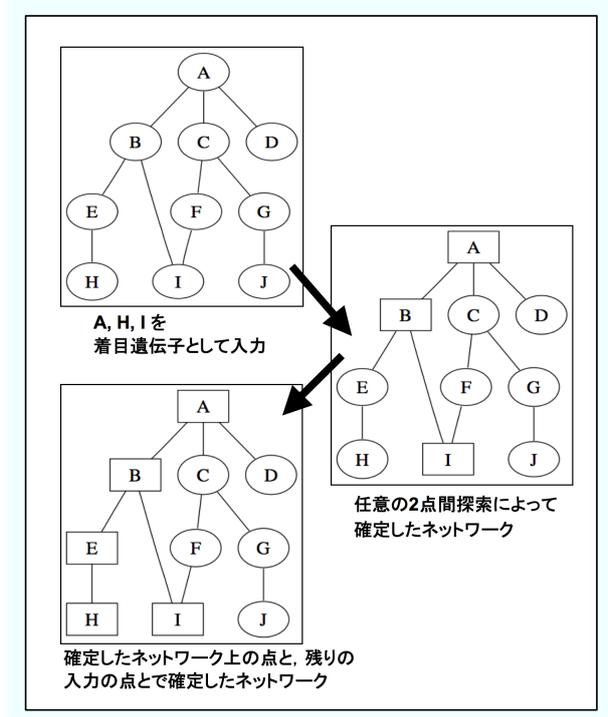


図 1: 3 点を入力した場合の探索

## 2 関連研究

たんぱく質相互作用ネットワークから、パスウェイを推定する試みとして、着目する 2 つの遺伝子を結ぶ、最短のパスをダイクストラ法によって見つける提案が為されている [2] が、パスウェイに対するシグナルの入力も出力も複数考えられるので、3 遺伝子以上でも発見する必要がある。

## 3 研究内容

シグナルの入力となる遺伝子群、及び、出力となる遺伝子群から、それらを結ぶ経路を推定する方法として、たんぱく質相互作用ネットワーク (ノードが遺伝子、エッジが相互作用に相当) 上で、入力、出力遺伝子群を結ぶ、部分グラフの抽出が考えられる。エッジは

相互作用を表しているので、求めた部分グラフのエッジ数が少ないグラフ程、シグナル伝達として利用されていると考えるのが妥当であろう。

しかし、一般に、与えられたグラフの部分グラフの内、指定した全てのノードを含む最小小さな部分グラフを見つけることは容易ではない。本研究では、着目した遺伝子群を含みなるべくエッジの本数が少ない部分グラフを現実的な時間で発見する方法を提案する。

手順

- (1) 与えられた相互作用ネットワークグラフを  $G$  とする。初期状態としてグラフ  $T$  を空のグラフとする。 $T$  を確定ネットワークと呼び、最終的に  $T$  に求めるグラフが入る。
- (2) 着目遺伝子群 (遺伝子数 2 以上)  $V$  の任意の 2 点間で最短パス  $P$  を求める。
- (3)  $T = P$  とする。
- (4)  $V \cap (G - T) = \emptyset$  となるまで (5), (6) を繰り返す。
- (5) ノード  $x \in T, y \in V \cap (G - T)$  のうち、2 点間の距離が最短となる  $x, y$  及び最短パス  $P'$  を求める。
- (6)  $T = T \cup P'$  とする。

Pathway Inference By Searching Sub-graph of Protein-Protein Interaction Network

Mio Seki<sup>†</sup>, Jun Sese<sup>†</sup>

<sup>†</sup>Dept. of Information Science, Ochanomizu University

図 1 にアルゴリズムの動作例を示す．四角のノードが確定したネットワーク T 上のノードである．

#### 相互作用ネットワークデータ

相互作用ネットワークとして Yeast two-hybrid(Y2H) データ (Ito らの Full data [3, 4] と Uetz ら [5]) 及び TAP データ (Krogan ら [6] のデータの内 Confidence Score が 0.8 以上のもの) を統合して使用した．また, これらの実験において bait, pray 間の関係と既知のシグナルが伝達される向きに有意な関係が見られなかったことから, ネットワークは無向グラフとして扱った．

## 4 実行結果

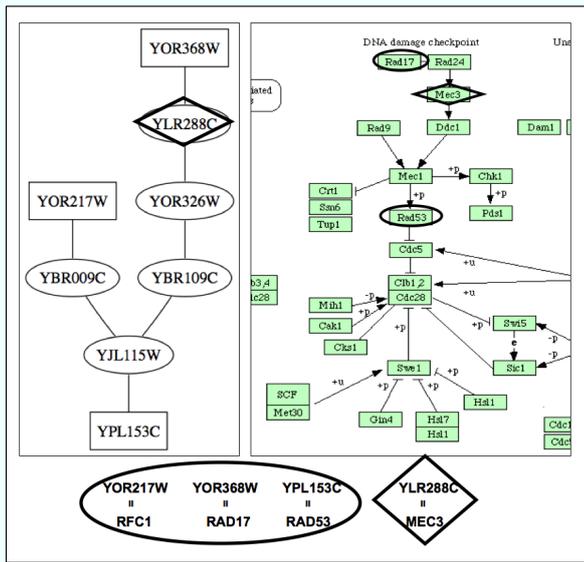


図 2: (左) 実行結果, (右) KEGG のパスウェイ

実行結果の例を図 2(左) に示す．長方形のノードが入力した遺伝子．楕円のノードは求めた部分グラフ上に現れた遺伝子である．例では, DNA 修復複製制御に関連する RFC1(YOR217W), RAD17(YOR368W), RAD53(YPL153C) の 3 遺伝子を着目遺伝子として入力した．RFC1 が関係する DNA 複製チェックポイントのシグナルと, RAD17 が関係する DNA 傷害チェックポイントのシグナルはともに合流し, RAD53 がさらに下流へ伝えていく予測となった．これは, 既知の知見とも一致する．

図 2(右) は, KEGG 上の DNA 障害チェックポイントのパスウェイの一部である．楕円で囲まれた遺伝子は, 上で述べた着目遺伝子 RAD17 と RAD53 である．ひし形で囲んだ遺伝子は図 2(左) の図のひし形と対応する．今回発見した経路は KEGG 上のパスウェイとの類似も見られつつ, 一部の経路が異なる新たなパスウェイの可能性もある．また, 実行結果上の RAD17 と RAD53 間のパスウェイの長さは KEGG 上の同一遺伝子間の経路とほぼ等しく, 推定したパスウェイも十分

存在する可能性がある．しかしながら, Y2H, TAP のみのデータ上での探索であるため, 着目遺伝子や着目シグナルによっては, 本手法では, とても大きなネットワークが求められてきてしまい, 新たなパスウェイであると推測するのは難しい結果となった．

## 5 まとめと今後の課題

本研究では, たんぱく質 - たんぱく質相互作用ネットワーク情報を利用して, 膨大なネットワーク上から, 着目遺伝子を含みなるべくエッジの本数が少ない部分グラフを探索する手法を述べた．本手法により, 着目シグナルに關与する, 存在する可能性のある新たなパスウェイを発見できることがわかった．

今後の課題としては, 得られた実行結果が有意であるか検証していくことと, 検索するネットワーク情報に, リン酸化ネットワークの実験結果 [7] などを含めることで実行結果として求めたネットワークやその探索の有意性を高め, パスウェイ情報の推定に, より有用なものとしてほしい．

## 参考文献

- [1] M. Kanehisa, and S. Goto. *KEGG: Kyoto Encyclopedia of Genes and Genomes*. Nucleic Acids Res, 28: 27-30, 2000.
- [2] H. Hermjakob, et al. *IntAct: an open source molecular interaction database*. Nucleic Acids Res, 32: D452-455, 2004.
- [3] T. Ito, et al. *Toward a protein-protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins*. Proc. Natl. Acad. Sci. 97, 1143-1147, 2000.
- [4] T. Ito, et al. *A comprehensive two-hybrid analysis to explore the yeast protein interactome*. Proc. Natl. Acad. Sci. 98, 4569-4574, 2001.
- [5] P. Uetz, et al. *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae*. Nature 403, 601-603, 2000.
- [6] N. J. Krogan, et al. *Global landscape of protein complexes in the yeast Saccharomyces cerevisiae*. Nature 440, 637-643, 2006.
- [7] J. Ptacek, et al. *Global analysis of protein phosphorylation in yeast*. Nature 438. 679-684, 2005.