

遺伝子機能間の関係を明示する 遺伝子ネットワークの束化と可視化

中澤里奈^{†1} 伊藤貴之^{†1} 瀬々潤^{†2} 寺田愛花^{†2}

遺伝子の機能や他の遺伝子との関係性を解明する情報共有手段として遺伝子ネットワークが役立てられている。このようなネットワーク構成を表示する際、複雑な接続構造のためにノードやエッジの画面上での重なりが多く生じる。この問題を解決するために、本手法では可視化手法 FRUITS Net を画面配置アルゴリズムとして適用し、ネットワークのエッジを束化することによって大量の線情報を要約し、視認性を向上させている。

Edge Bundling and Visualization for Summarization of the Relations between Gene Functions

RINA NAKAZAWA^{†1} TAKAYUKI ITOH^{†1}
JUN SESE^{†2} AIKA TERADA^{†2}

We usually use gene networks so as to reveal gene functions and relationship. However, it may be difficult to understand relations between genomic functions and gene-gene interactions because of the cluttering. In this paper, we introduce a simultaneous visualization of a large gene network and gene ontology (GO). The technique represents the functions provided by GO as colors of nodes, and bundles edges depending on the gene functions to ease visual complication of the network.

1. はじめに

遺伝子ネットワークの可視化はこれまでも多く試みられてきた。しかし、ノードやエッジ数が非常に多いため、その中で注目すべき箇所を一目で理解することは難しい。このため、ネットワーク構成の明快な表示が重要だが、複雑な接続構造のためにノードやエッジの画面上での重なりが多く生じ、機能と相互作用の視認性を妨げることがある。

この問題を解決するため、本手法では可視化手法 FRUITS Net¹⁾を画面配置アルゴリズムとして適用し、エッジで連結された遺伝子クラスタ、あるいは共通の機能を有する遺伝子クラスタを、画面上でできるだけ近く配置する。遺伝子の機能情報として、Gene Ontology(GO)²⁾を用いる。その上で、ノードやエッジの画面上での重なりを低減させ、機能の相互作用の関係性に関する視認性を向上するために、エッジの束化を適用した可視化手法を提案する。本手法では同一の2クラスタ間に接続されたエッジが一定本数以上を有する場合に、これを束として描画する。GOによって定義された機能情報を色に対応づけ、遺伝子ネットワーク上に機能情報と接続関係を同時に可視化することで、遺伝子の機能情報と関係性をより把握しやすくしている。

2. 関連研究

伊藤らは FRUITS (Framework and User Interface for Tangled Segments) Net¹⁾というネットワーク可視化手法を

提案し、遺伝子ネットワークの可視化に適用している。この手法は、ノードに付与された1個以上のアイテム情報を色で表現し、同一アイテムを共有するノードをできるだけ近くに配置する。これにより、ノードに与えられたアイテム情報とネットワーク構造の関係性を可視化することができるので、本報告の目的に向いていると考えられる。

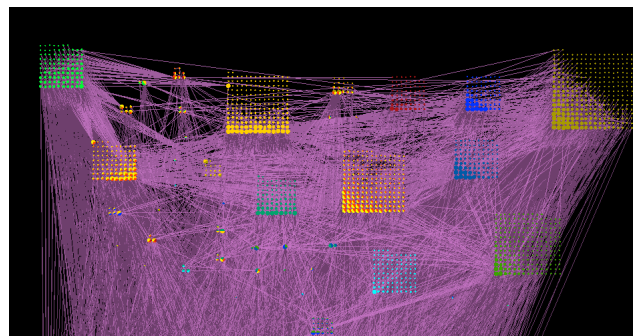


図1 FRUITS Netの可視化例

3. 提案手法

3.1 ノードクラスタリングと画面配置

本手法では FRUITS Net を用い、遺伝子をノード、遺伝子間の相互作用をエッジとして遺伝子ネットワークを可視化する。ここで GO タームをアイテム情報としてノードに付与し、ノードを色付けする。

より多くの GO 情報を可視化結果に反映するために、もともと DAG 構造をとっている GO タームを、データ構造上の距離をもとにクラスタリングし、特定の条件を満たす 10~15 種類の GO タームクラスタをアイテム情報としてノードに付与する。さらに、FRUITS Net を使用する前処理段

^{†1} お茶の水女子大学
Ochanomizu University

^{†2} 東京工業大学
Tokyo Institute of Technology

階として、ノードクラスタリングを実施する。ノードのアイテム情報である GO タームの組み合わせと、エッジの密度の両者を考慮してクラスタリングをする。そして、クラスタを単位として FRUITS Net の配置アルゴリズムを用いることでノードの配置を決定する。力学モデルによって共通のアイテム情報を持つクラスタを近くに配置し、空間充填モデルによりノードやクラスタの重なりを回避しつつ、配置画面占有面積を減らすことができる。

3.2 エッジの束化

続いて、線情報の視認性を向上するためにエッジの束化を施す。エッジの束化とは大量の線情報を束ねることで視認性を向上させる方法である。提案手法では、接続された遺伝子ペアという位相的な情報だけが重要なので、ノードと束の重なりを回避しつつ、より簡易にクラスタ間に太線を束として描くことで高速に束化を実現する。また本手法では、エッジの方向性ではなく同じ機能の組み合わせを持つクラスタを単位としてエッジの束化を適用する。

束化を適用するかどうかは、ユーザがその閾値をスライダーによって設定することができる。本数のみを閾値としてしまうと、クラスタに所属する遺伝子数が少ない場合に常に束として描かれないという問題が生じる。そこで、本手法では束に接続するノード数とそのクラスタの全ノード数に対して占めている割合を求め、その割合が一定以上ならば束化を施すことも可能にしている。その上で、エッジの束化を行うかどうかの基準をユーザが本数かノード数の割合か選べるようにしている。

4. 実行例

4.1 使用データ

我々は iRefIndex³⁾ に公開されているショウジョウバエの遺伝子ネットワークを適用し、各遺伝子をノード、タンパク質間相互作用をエッジとし、GO によって定義される遺伝子機能情報をノード付与して可視化を試みた。ノード数は 8945、エッジ数 32703 である。

4.2 実行結果

本手法の実行結果を図 2 に示す。ここで、(a)から(e)までの遺伝子クラスタに付与されている主要な GO タームは次の通りである。

- (a): protein-DNA complex (GO: 0032993)
- (b): intracellular organelle part (GO: 0044446),
non_membrane_bounded organelle (GO: 0043228),
organelle part (GO: 0044422)
- (c): non_membrane_bounded organelle (GO: 0043228)
- (d): cell projection (GO: 0042995)
- (e): intracellular organelle part (GO: 0044446),
organelle part (GO: 0044422)

図 2 から機能間の主要な関係性が、特に(a)と(b)、(a)と(c)、(b)と(e)、(d)と(e)の各遺伝子クラスタ間において、より多

くの相互作用が見られることがわかった。これらの GO タームと実行結果から得られた束との意味合いを見てみると、以下のように生物学的な知見との一致が見られた。

- ・(b)-(c)の束： 細胞器内と細胞質は情報伝達の役割を果たすものが相互作用する。
- ・(b)-(e),(c)-(e)の束： (b)と(c)はそれぞれ膜タンパクには結合する。
- ・(d)-(e)の束： 膜タンパクである(e)を支える骨格のように(d)が結合する。
- ・(a)-(b),(a)-(c)の束： 膜タンパクが外側にシグナルを受け取った時に反応して転写因子が活性化する。

今回の結果から、遺伝子機能情報と接続関係を同時に可視化しつつ、エッジの束化を適用することで、エッジが煩雑で見えなかった同じ遺伝子クラスタ内のノード同士の接続性を把握することができた。

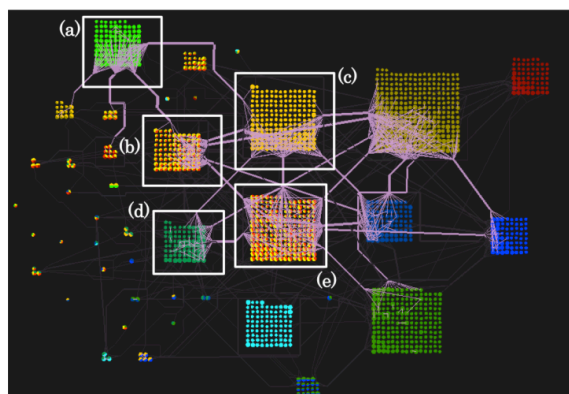


図 2 本手法の可視化結果

5. おわりに

本報告では、エッジの束化によって遺伝子ネットワークと遺伝子機能間の関係を可視化する手法を提案し、その実行例を示した。今回の実行結果から、機能間の繋がりに関して既に明らかになっている生物学的な知見との一致が見られた。このことから本手法は、例えば、たくさんの知見を一画面に全部示すことで利用者間での同意を手早く得たい場面に有効であると言える。今後の課題としては、Cytoscape⁴⁾といった既存ツールとの連携なども考えていきたい。

謝辞 貴重なご意見をいただきました、産業技術総合研究所 油谷幸代様に感謝いたします。また本研究の一部は、日本学術振興会科学研究費補助金の助成に関するものです。

参考文献

- 1) T. Itoh, C. Muelder, K. Ma, J. Sese: A Hybrid Space-Filling and Force-Directed Layout Method for Visualizing Multiple-Category Graphs, IEEE Pacific Visualization Symposium, 121-128, 2009.
- 2) The Gene Ontology, <http://www.geneontology.org/>
- 3) iRefIndex, <http://irefindex.uio.no/wiki/iRefIndex>
- 4) Cytoscape, <http://www.cytoscape.org/>