

進化アルゴリズムを用いた分子間相互作用ネットワークの効率的な推定手法の開発

古森, 朝子

<https://doi.org/10.15017/1654677>

出版情報：九州大学, 2015, 博士（システム生命科学）, 課程博士
バージョン：
権利関係：全文ファイル公表済

氏 名 : 古森 朝子

論 文 名 : 進化アルゴリズムを用いた分子間相互作用ネットワークの
効率的な推定手法の開発

区 分 : 甲

論 文 内 容 の 要 旨

生物をシステムレベルで理解することを目標としたシステム生物学分野において、個々の生体分子がどのように相互作用し合っているか知ることは非常に重要である。近年、各生体分子のオミックスデータから、分子間相互作用ネットワークを再構築する研究が盛んに行われている。

分子間相互作用ネットワーク推定手法の一つに、実験から得られた生体分子の濃度のタイムコースデータを、微分方程式モデルである **S-system** を用いて再現し、**S-system** のシステムパラメータから分子間相互作用ネットワーク構造を推定する手法がある。この手法では、フィードバックなど、ループ構造を含む複雑な構造も推定でき、推定されたネットワークの各システムパラメータの感度解析が行えるなどの利点を持つ。一方、規模の大きな相互作用ネットワークを推定する際には、膨大な数のシステムパラメータを最適化せねばならず、この問題に対処するため、これまで先行研究において、実数値遺伝的アルゴリズム (実数値 GA) を用いた手法が開発されていた。しかし、従来の実数値 GA では最適化可能なパラメータ総数は 2 桁に留まっていた。そこで、申請者らは、**AREX+JGG** 法およびハイブリッド法 (**AREX+JGG** 法+修正 **Powell** 法) という新たな推定手法 (提案法) を開発してきた。これまで、提案法が、従来法に比べ推定速度では約 20 倍速く、また推定可能なパラメータ数では約 15 倍の規模を推定することができる有力な手法であることを示してきた。しかしそれでも推定可能なパラメータ総数は未だ 300 程度に留まっており、実用化のためにはさらなる工夫が求められる。本論文は、実数値 GA という進化アルゴリズムを用いた分子間相互作用ネットワークの効率的な推定手法の開発を取りまとめたものである。

推定可能なパラメータ数が限られているような手法を大規模な相互作用ネットワークに適用する方法の一つとして、相互作用ネットワークをクラスタリングする方法が挙げられる。この方法は、各クラスタに提案法を適用することで元の相互作用ネットワークの再構築を目指すものである。しかし、適切なクラスタリングのためには、事前に各構成要素の濃度のタイムコースの形状や、各要素に関するデータベース情報等が求められる。そこで、本研究では元の相互作用ネットワークを構成する要素から、ランダムに要素を選択し、選ばれた要素で構成されたサブネットワークを推定することで元の相互作用ネットワークを再構築する **Random selection method (RA 法)** を開発した。具体的には、ランダムに推定する要素を選択し、選択した要素間の推定に **AREX+JGG** 法またはハイブリッド法を適用して、その後、推定した要素間の共通構造を重ねあわせるという手法である。

本論文では、まず RA 法の有効性を検証するために、30 要素ネットワークモデルを作成した。このネットワークモデルは **AREX+JGG** 法及びハイブリッド法単独の手法では推定が困難な問題である。作成するサブネットワーク数を区切り、30 要素ネットワークモデルを推定したところ、RA 法はネットワークモデルに含まれる相互作用の内、約 84%を推定することができ、ネットワーク構造

のほとんど再構築することに成功した。作成するサブネットワーク数を増やすことで、推定される相互作用数も増加すると考えられる。

次に、RA法のノイズ耐性を検証した。一般に、実用化の際に入力データとして用いるタイムコースデータには、測定誤差や生物学的ノイズなどの誤差が必然的に含まれることから、推定手法にはノイズへの耐性が求められる。まず、RA法の有効性を示す前段階として、従来法とのノイズ耐性を比較検証した。ケーススタディーとして4要素および7要素ネットワークモデルを用意した。Langevin方程式を用いてノイズを含ませた各要素のタイムコースデータを作成し、推定したところ、従来法では7要素ネットワークモデルを一つも推定することができなかったが、RA法ではより優れた結果を示した。特に、ハイブリッド法を用いたRA法ではすべての相互作用を推定することに成功した（許容誤差は、タイムコースデータに含まれる相対誤差で換算すると30%程度）。さらに、ノイズを含む30要素ネットワークモデルを用いRA法のノイズ耐性を検証したところ、ノイズを含まないデータを用いた場合と同等の推定能力を有していた。したがって、RA法は誤差を含むタイムコースデータに対しても有効であると考えられた。

さらに、RA法が様々な特徴を持つような相互作用ネットワークに対しても有効であるか検証した。検証には、複雑ネットワークの代表的なモデルであるランダムネットワークおよびスケールフリーネットワークの特徴を有する2つの100要素ネットワークモデルを用いた。その結果、RA法はどちらのネットワークモデルにおいても相互作用を推定する能力を有していた。現在のところ、推定された相互作用は全体の25~35%程度であるが、推定率は試行回数に対して単調増加しており、今後、試行回数を増やすことでTP率はさらに上昇すると考えられる。

以上の点から、RA法はノイズ耐性を有し、様々なネットワーク構造に対応可能な手法であることが明らかになり、実問題においても優れた性能を示すと考えられる。今後、DNAマクロアレイデータを用いた解析や、階層間相互作用ネットワーク推定への応用が期待される。