

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 30 日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25870396

研究課題名(和文)細胞間の多様性に基づいた新しい細胞内モデリング手法 - その発展と応用 -

研究課題名(英文)On the novel mathematical modeling framework of cells based on their heterogeneity

研究代表者

寺口 俊介 (Teraguchi, Shunsuke)

大阪大学・免疫学フロンティア研究センター・特任助教

研究者番号：00467276

交付決定額(研究期間全体)：(直接経費) 1,000,000円

研究成果の概要(和文)：本研究では、細胞のシミュレーションを行うための新しい手法を数学的に発展させることで、この手法をより多くの生命現象に応用可能にしました。また、新たな近似法と高速シミュレーターを用いることで、実際のデータからシミュレーションモデルを自動的に構築するシステムを開発しました。また、この手法を応用して、免疫寛容と自己免疫応答のバランスを理解するための免疫細胞の抗原提示の数理モデルを構築しました。

研究成果の概要(英文)：In this research, we have mathematically developed new methodology to perform simulation of cells to increase the applicability of our method. With the help of newly developed approximation and a fast simulator, we constructed a system to identify simulation models from real data automatically. In addition, by applying our framework, we developed a mathematical model of antigen presentation of immune cells to understand the balance between immune tolerance and autoimmune reaction.

研究分野：システム生物学

キーワード：数理モデル 機械学習 シミュレーション 免疫システム

1. 研究開始当初の背景

現在の生物学の主流となっている分子生物学的アプローチは、生命現象の因果関係を各遺伝子の機能に還元することで大きな成功を収めてきた。このような遺伝子/蛋白質分子に関する個別の知見は既に膨大なものであるが、一方で、これらの因果関係が全体としてどのように統合され、実際の生命現象を形作っているのかという問題は、かなりの部分が手つかずのまま残されている。このような状況の中で、システム生物学と呼ばれる、定量的なデータに基づき生命現象をシステムとして理解することを目指した生物学のアプローチの必要性が強く認識されてきた。システム生物学の重要な方向性として、数理モデルや計算機シミュレーションを用いて細胞システムの定量的な再構築を行うことが挙げられる。それにより、システムの設計原理を理解し、予測や制御までを可能にすることが期待されている。

このように細胞レベルでの生命現象を数理的に記述するに当たっては、生化学反応の方程式系をベースにした数理モデル/シミュレーションが標準的なアプローチであると言える。このような数理モデルでは個々の細胞間の差異は一旦無視し、平均的な振る舞いをする理想的な細胞のモデルを構築することを出発点にすることが多い。しかし、実際には、同一の細胞種であっても外部シグナルに対する個々の細胞の応答は同一とは限らない。特に多様な病原体を検出する必要がある免疫システムでは、このような細胞間の差異も重要な役割を果たしていると考えられる。そこで、我々は、このような細胞応答の多様性・不均一性を積極的に利用し、細胞集団における細胞内現象を記述する新しい手法である“Stochastic Binary Modeling”(SBM)を開発、提案していた(Teraguchi S, 他 4 名、Physical Review E, 2011)。

2. 研究の目的

本研究の目的は、細胞間の多様性に基づく新しいモデリング手法をさらに発展させ、多様な生命現象に応用可能にすることであった。また、研究期間中に一定のめどが立ったため、当初の研究計画を拡張し、実験データからのモデル推定を行うためのシステムの開発も本研究の重要な目的の一つと位置付けチャレンジすることとした。

3. 研究の方法

(1) 新しい近似手法の開発

SBM では通常モデリング手法では扱いづらい細胞間の多様性を直接モデルに組み込むことができる。しかしながら、その代償として、力学変数の数がモデル化するシステムの複雑さに対して指数関数的に増加してしまうという欠点がある。この問題を回避し大

規模なシステムのシミュレーションを行う一つの方法としてはモンテカルロシミュレーションに基づく確率的な近似手法がある。しかし、この方法では数値的な解析しか行えないこと、また、毎回確率的に異なる結果が得られてしまうことから、目的によっては必ずしも便利な近似法ではない。そのため、SBM の特徴を生かした新しい近似法の開発が望まれていた。

(2) 高速シミュレータの開発

上述のように、特に大規模なシステムではSBM のシミュレーションの速度は研究のボトルネックになり得る。特に、後述する機械学習システムにおいては、パラメータを変化させながら多数のシミュレーションを繰り返す必要があるため、少しでも高速なシミュレータを用いることは重要である。また、近似法を開発する際にも、その妥当性は数値的により厳密な解との比較を行うことでチェックする必要があるため、直接計算を行うシミュレータの高速化は本研究において必須であった。

(3) 機械学習によるシステム同定手法の開発

SBM でのモデリング自体は非常に簡便で、各構成要素(遺伝子/蛋白質分子)の間の相互作用に大まかな時定数を与えるだけで、システムのシミュレーションが可能になる。しかし、そのような相互作用の存在とパラメータの選択を客観的な形で行うのは容易ではない。この問題に対する一つの回答は、実験で得られた時系列データから自動的にシステム同定を行うデータ駆動型のアプローチである。このアプローチは実際の生命現象へSBM を応用していくために非常に有用であると考えられるものの、当初はこの方向での問題解決の目処が立っていなかったため研究計画に含めていなかった。初年度を終えた段階でこのアプローチに一定の目処が立ったため、新たに機械学習システムによるシステム同定手法の開発を本研究の大きな柱に据えた。

(4) 免疫システムにおける一細胞レベル実験データの解析

もともと SBM というモデリング手法を着想したのは、免疫系の実験研究者との議論において、実際の細胞では細胞間の不均一性が高く、全ての細胞が同じように振る舞うという通常近似では実際の免疫細胞の生命現象を捉えることは難しいという指摘を受けたところであった。細胞間の不均一性が実際どの程度のものであり、どのようなシステムがSBM のようなモデリング手法との相性がよいのかを知るためには、一細胞レベルでの実験データが必要となる。

そのため、本研究では実験研究者との共同研究を通じて以下のような実際の実験データ

の解析を行った。

蛍光蛋白質を用いた遺伝子発現の時系列データの解析

一細胞 qPCR を用いた多数の遺伝子発現データの解析

(5) 免疫システムに対する数理モデルの構築

我々のモデリング手法の妥当性を示す最も効果的な方法は実際の生命現象に対して応用を進めることである。上記(4)にある遺伝子発現の径時変化のデータを用いてそのような数理モデルの構築を進めることができれば理想的であったが、本研究時点では複数遺伝子発現をモニターするためのレポーターシステムの構築が難航し、このデータの直接の利用は難しかった。

そこで、今回は SBM のフレームワークを拡張することで、理論的なアプローチから適応免疫システムの数理モデリングを行った。

4. 研究成果

(1) 新しい近似手法の開発

繰り込みの手法を用いたアプローチ

SBM は物理学における量子力学の定式化と数学的な類似性があり、そこでの代表的な近似手法である摂動論が応用できる可能性があった。通常、細胞の数理モデル研究は数値的なシミュレーションに頼る必要があるが、摂動論では近似的ながらも解析的な計算が可能になるメリットがある。しかし、量子力学の場合と異なり、SBM では相互作用を行う期間と行わない期間を分離できず、時間の経過とともに摂動論が破綻してしまうため、残念ながら実用的な利用は行えなかった。今回、同じく物理学においてよく利用される繰り込みの手法を用いてこの問題を回避/緩和できないか研究を行い、実際、繰り込みの手法を応用することが可能であることがわかった。しかし、同時に、繰り込みの手法で解消できる摂動論の破綻の程度はモデルに大きく依存してしまうことも明らかになったため、残念ながら現時点では具体的な応用は難しい状況である。

確率の独立性を仮定した近似法の開発

SBM のシステム方程式を近似計算するためのもう一つのアプローチとして、システム方程式に現れる条件付き確率を条件のない確率で置き換える近似法を開発した。この近似法を用いると SBM のシステム方程式に現れる高次の同時確率を低次の同時確率で表現し、システム方程式を任意の複雑さで切断することが可能になる。この近似法を数値シミュレーションで検証したところ、この近似は多くの場合高い精度で成り立ち (Fig.1)、興味のある観測量が低次の同時確率で表現できる状況では非常に有用な手法であること

が分かった。

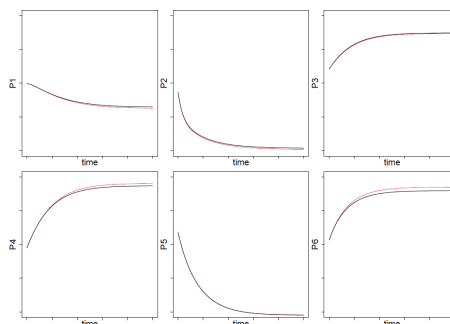


Fig.1 確率の独立性を仮定した近似解の例。横軸：時間、縦軸：各ノードがオンである確率。黒線：厳密解、赤線：近似解。

(2) 高速シミュレータの開発

以前の研究で SBM のシミュレータは構築していたものの、このシミュレータはシステム同定のような応用を視野に入れておらず、そのシミュレータの実行速度は十分なものではなかった。そのため、本研究では高速なシミュレータを新たに開発し直した。以前のシミュレータと比べて、高速な低レベル言語の採用、スパース行列の利用、計算アルゴリズムの改良を行い、大幅な高速化を果たした。

(3) 機械学習によるシステム同定手法の開発

本研究での大きなチャレンジが、元々の研究計画には含めていなかった機械学習システムの開発であった。この方向にかじを取る大きな契機となったのは "Continuous Time Bayesian Network" (CTBN) と呼ばれるネットワーク推定法の生物学への応用研究を行っている海外研究者との交流によるものだった。このネットワーク推定と SBM のフレームワークには大きな類似性があり、計算コストの高い SBM のシミュレーションを行わなくとも、比較的高速なネットワークの推定が可能であると考えられた。この手法を SBM のシステム推定の前段階として用い、それを初期値としたデータ同化の手法を用いることで、効率の良いシステム同定ができるのではというのが当初のもくろみであった。

実際、上述の研究者の協力を得ることで、CTBN によるネットワーク推定手法が利用可能となった。しかし、一方で、生物学で利用できる通常の実験データからでは、CTBN のネットワーク推定の精度は高くなく、さらにデータ同化の手法を組み合わせたとしても、複雑なシステムでは、実用的な精度でのネットワーク推定は極めて難しいことがわかってきた。特に、どうしても問題になるのは、複雑なシステムにおける観測量の数とシステムのパラメータ数とのギャップである。事前情報がないシステム同定においては、観測量が増えるにつれ、あり得る相互作用パラメータの数はその数の二乗に比例して増えてしまうため、多数の構成要素を持つ複雑な

システムではパラメータの同定が困難になってしまう。

そこで、今回の研究では一旦 CTBN の利用を諦め、システム同定に利用する実験データとして Flow cytometry のような一細胞レベルの解像度を持つ実験データに特化したシステム同定の手法を新たに開発することとした。この手法の最大のメリットは、システムの構成要素の数が増えるにつれて、同定すべきパラメータの数を超えて観測量の数が増えるところである。標準的な細胞の数理モデルの手法では個々の細胞間の差異を一旦無視するため、このようなデータの利用法は難しいのに対し、SBM のシステムではこのような一細胞のデータと非常に相性がよい。結果的に、少なくとも *in silico* の実験では、システムの複雑さが増しても推定精度を維持できるシステム同定手法の開発に成功した (Fig.2)。現在、この手法の実際の実験データへの適用を試みている。

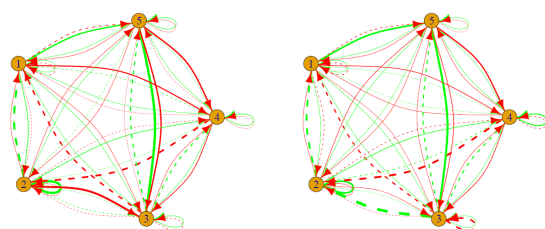


Fig.2 今回開発したシステム同定手法によるパラメータ推定の実行例

左：推定されるモデルのパラメータ。右：推定されたモデルのパラメータ。緑（赤）：ノードのオン（オフ）。実（破）線：相互作用時の親ノードのオン（オフ）。線の幅が相互作用の強さに対応する。

(4) 免疫システムにおける一細胞レベル実験データの解析

蛍光蛋白質を用いた遺伝子発現の時系列データの解析

私たちの研究グループでは、遺伝子改変マウス由来の免疫細胞を顕微鏡下で経時的に観察することによる一細胞レベルでの遺伝子発現解析実験を進めている。この実験は必ずしも本研究のために計画されたものではないが、そこで得られる観測データの SBM での再利用を念頭に、時系列イメージからの発現データの抽出を目指したアルゴリズムの開発を行った。明視野イメージからの各細胞の境界検出と、蛍光イメージからの対応する細胞の発現量の検出に成功した。SBM で利用するには、さらにトラッキングを行い、各時点での細胞を同定する必要があるが、残念ながら今回の実験条件下では、細胞の数と時間間隔の両方が大きく、トラッキングまで完了するには至らなかった。実験条件の再検討も含めてこれは今後の課題である。

一細胞 qPCR を用いた多数の遺伝子発現

データの解析

細胞間の多様性に注目することで定式化した SBM を実際の生命現象に応用していく上で、実際の細胞システムの多様性の程度を把握しておくことは欠かせない。本研究は理論研究であり、実際の細胞を直接観測する機会を得るのは難しいが、実験研究者との共同研究を通じて 1 細胞レベルの遺伝子発現データを解析する機会に恵まれた。この研究では、光照射により波長が変わる特殊な蛍光色素を組み込んだ遺伝子改変マウスから得られた免疫細胞の遺伝子発現を、マイクロ流路デバイスを用いて一細胞レベルで測定した。この実験データの解析から、一見同じように見える制御性 T 細胞でも、様々な転写因子の発現が細胞間で異なっており、その傾向は細胞動態の履歴によってバイアスされることを示していた。この共同研究ではさらに高機能な制御性 T 細胞を見分けるマーカー遺伝子を同定することができた。

(5) 免疫システムに対する数理モデルの構築

本研究では具体的な応用先として免疫システムを主な対象としている。本研究では SBM で導入された細胞応答の多様性・不均一性の考え方をさらに拡張し、免疫システム等で重要な細胞間の相互作用を導入した。また、理論的な考察に基づくトップダウンの免疫システムのモデルも構築した。この数理モデルでは、自然免疫システムと獲得免疫システムをつなぐ鍵である抗原提示細胞による T 細胞への抗原提示のプロセスに焦点を当て、免疫学の主要なテーマである免疫抑制と自己免疫反応のバランスを定量的に議論できる。このモデルでは SBM で導入した各細胞の非一様性を、各抗原提示細胞と相互作用する T 細胞の非一様性に拡張し、T 細胞と抗原提示細胞の相互作用、抗原提示細胞上の T 細胞の分布に応じた T 細胞のプライミングを確率的に扱った。さらに、これらのイベントを免疫システム全体での T 細胞のポピュレーション動態にフィードバックすることで免疫抑制と自己免疫反応という相反するリスクを、システム全体として自然に定義することが可能になった。

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 0 件)

〔学会発表〕(計 3 件)

Shunsuke Teraguchi, Stochastic Binary Modeling and Network Motifs, The International Conference on Systems Biology 2013, 2013 年 8 月 30 日 ~ 9 月 3 日, Tivoli Congress Center, Copenhagen,

Denmark

Shunsuke Teraguchi, Stochastic Binary Modeling and Network Motifs, EMBO Workshop, Complex Systems in Immunology, 2013年12月2日~12月4日、Matrix Building, Biopolis, Singapore

寺口 俊介, Stochastic Binary Modeling and Network Motifs, 定量生物学の会 第六回年会, 2013年11月22日~11月24日、大阪大学銀杏会館

〔図書〕(計 0件)

〔産業財産権〕

出願状況(計 0件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計 0件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

寺口 俊介 (TERAGUCHI, Shunsuke)
大阪大学・免疫学フロンティア研究センター・特任助教
研究者番号：00467276

(2) 研究分担者

なし

(3) 連携研究者

なし