

生命情報ネットワークの構造および動的挙動の数理解析

●阿久津 達也¹⁾ ◆五斗 進¹⁾ ◆望月 敦史²⁾ ◆時田 恵一郎³⁾

1) 京都大学化学研究所バイオインフォマティクスセンター 2) 理化学研究所 3) 大阪大学サイバーメディアセンター

<研究の目的と進め方>

生体内では遺伝子、タンパク質、化合物など様々な物質が相互作用することにより生命活動を維持している。それらの相互作用は一種のネットワークをなし、それらの構造の推定が実験技術および情報技術を両輪として進展しつつある。しかしながら、単に物質間の接続関係や制御関係が明らかになるだけでは不十分であり、なぜ、そのような接続関係で、生物や細胞が安定性と環境変化に対する適応性を両立させているかを解明する必要がある。

本研究では、代謝ネットワーク、遺伝子ネットワーク、類似遺伝子配列のなすネットワークなど様々な生命情報ネットワークを対象として、グラフ理論、数理生物学、複雑系物理学など多様な観点からネットワーク構造の特徴づけを行い、安定性と適応性が保たれている理由を数理的に解明することを目指す。また、その解明のために、ネットワーク比較、ネットワークからの特徴抽出、ネットワークの安定性や動的挙動の解析などの情報解析ツールも開発し、得られた知識のデータベース化も図る。

<2008 年度の研究の当初計画>

本研究では(1) ネットワーク構造生成、(2) ネットワーク構造解析、(3) ネットワーク動的挙動解析の3課題を中心に行う。特に阿久津は(1)を、五斗は(2)を、望月、時田は(3)を中心に取り組むが、互いに方法、結果、データをフィードバックしながら行う。

(1) ネットワーク構造生成に関しては、相互作用の種類を分類するなどにより、タンパク質相互作用ネットワークのより精緻なモデル化を行い、それを説明する数理モデルを構築する。

(2) ネットワーク構造解析に関しては、昨年度に引き続き、代謝制御ネットワークやタンパク質のドメイン組み合わせネットワークを主な対象として、ネットワークの構造的な安定性や機能の多様性を獲得してきた仕組みを解析する。さらに、昨年度に開発した代謝ネットワークの安定性計算アルゴリズムの改良を図るとともに、その拡張や応用を試みる。

(3) ネットワーク動的挙動解析に関しては、ネットワークの構造と多様性についてのこれまでの理論的成果をもとにホヤの遺伝子ネットワークの推定を行う。ホヤの初期発生において主要な転写因子の発現パターンは全て調べられているが、遺伝子間相互作用はまだ十分分かっていないので、理論をもとに未知の遺伝子間相互作用の推定を試みる。また、これまでに行ってきたネットワークの定常状態解析や制御に関する理論研究を深化させるとともに、その発現データ解析などへの適用法の検討を継続して行う。複雑な相互作用をもつ大規模進化動力学モデルに対する平衡統計力学的解析手法を用いた研究についてはほぼ解決しつつあり、今年度は、周期解やカオスを含む本質的に動的な非平衡状態の解析と実証研究者との共同研究を進める。

<2008 年度の成果>

2008 年度は計画とは一部異なる展開をみせたものの、着実に成果を得ることができた。主要な成果は以下のとおりである。

(1) ネットワーク構造生成に関する研究

タンパク質相互作用ネットワークのより精緻な分析のために、ドメインとタンパク質の包含関係を調べ、ドメインを含むタンパク質の個数分布はスケールフリーとなるが、個々のタンパク質に含まれるドメインの種類の分布はポアソン分布になることを見出し、それを説明する数理モデルを構築した。また、2007 年度に着手した原核生物の生育温度と代謝ネットワーク構造の関連性を説明する数理モデルを完成させ、論文にまとめた。

(2) ネットワーク構造解析に関する研究

(2-1) 複数の神経変性疾患に関わるタンパク質の相互作用ネットワークを構築し、ネットワーク比較による疾患間の共通性・特異性を明らかにした。また、KEGG のパスウェイ情報を統合してグローバルなネットワーク解析ができるウェブサーバーを開発した。

(2-2) 代謝ネットワークの遺伝子変異に対するロバスト性解析の研究を進め、標的化合物を生成不可能にするために不活性化することが必要な反応の最小数を整数計画法を用いて計算するアルゴリズムを開発した。さらに、KEGG データベースより取得したデータに適用し、有効性を確認した。

(3) ネットワーク動的挙動解析に関する研究

(3-1) 新たな理論的取り組みとして、遺伝子制御ネットワークと遺伝子発現パターンのデータから、遺伝子発現ダイナミクスの吸引領域を推測する解析法を開発した。これにより、発現ダイナミクスの断片的情報(スナップショット)である発現パターンの情報から、遺伝子発現の時間変化の概要を予測できると分かった。

(3-2) 遺伝子ネットワークの確率的な離散モデル上で定常状態を計算するアルゴリズムを開発し、また、その平均的な個数の評価を行った。さらに、(確率的でない) 離散モデルにおける定常状態の分布と確率的な離散モデルにおける定常確率分布との関係について、基礎的な場合に関して関係式を導いた。

(3-3) 一般的な進化動力学モデルに対する統計力学的研究を進め、系の安定性や個体数、蛋白量、代謝物質量などの様々な生物量の分布を与える理論を構築した。細胞内で特徴的な生物量の有限性の効果や触媒反応などの3体相互作用の影響も明らかになった。

<国内外での成果の位置づけ>

2007 年度も個々の成果は、バイオインフォマティクス、統計物理学、数理生物学などの主要英文論文誌に発表されており、また、学会誌などでの解説記事掲載や国内外での招待講演なども行われるなど高いレベルを維持していると考えられる。

タンパク質相互作用ネットワークや代謝ネットワークの構造生

成の数値モデルに関する研究は国際的にも一段落したと思われるが、我々は地道に研究を続けており、国内外からも一定の評価を受けていると考えられる。特にオープンアクセス論文誌に掲載された生育温度と代謝ネットワーク構造の関連性を説明する数値モデルに関する論文は多くのアクセスを受けており、多くの研究者の関心を集めていると考えられる。

神経変成疾患に関わるタンパク質の相互作用ネットワーク解析は、アルツハイマー病の研究をされている実験系の研究者との共同研究である。また、文献を中心に手作業で集めたデータを用いており、他にはないユニークなものである。なお、現在はKEGGの一部として公開されている。

以前の研究成果である神経細胞樹状突起形態形成の数値モデルに対しては、発生生物学者、神経生物学者を中心に高い評価を得ている。また「ネットワークの構造と発現多様性の関係」に関する研究への評価は年々高まっている。生物学者に加えて、数理科学者から新たな手法としての期待を受けている。

複雑な相互作用をもつ大規模進化動力学モデルの量の分布の理論研究については、国内外でもほとんど行われておらず、統計力学を用いた生命システム研究の先駆けである。また、量の有限性や3体相互作用は通常の物理系では扱われないため、統計力学自体の発展にも寄与するものである。

<達成できなかったこと、予想外の困難、その理由>

ネットワーク構造生成についてタンパク質とドメインの関連については以前より精緻な解析を行うことができたが、相互作用ネットワークのより精緻な解析には至らなかった。今回得た結果を相互作用ネットワークの解析につなげる必要がある。

代謝ネットワークのロバスト性解析については研究を開始した後にフラックスモード解析によるものなど数多くの関連研究があることがわかった。しかしながら、その結果、既存手法の問題点も明らかになりつつあり、独自の改良につながるとの感触を得ることができた。

代謝制御ネットワークのモチーフ抽出に関しては、現在、パスウェイを比較してモジュールを抽出するパイプラインを構築中であり、進化的な解析には至らなかったが、引き続きこのパイプラインを用いて生物種間比較に基づいた解析を進める予定である。

今回の新たな成果である遺伝子ダイナミクスの吸引領域の決定法の開発は、ホヤの遺伝子制御ネットワークを解析する過程で、得たものである。ホヤの遺伝子制御の実験データは一見膨大な量に見えるが、実は遺伝子発現ダイナミクスを再現するには不十分であり、新たな推測法を開発する必要があった。

生物量の有限性の効果については、(物理の分野においても数学的取り扱いが困難なため)現状では数値シミュレーションの結果のみであり、数学的に厳密な解析はできていない。

<今後の課題>

(1) ネットワーク構造生成に関する課題

ネットワーク構造生成の数値モデルについては分野全体として行き詰まり感があるとは思われるが、より精緻な解析を行うことにより新たな構造的な特徴や数値モデルを見出すことができた。そこで、今後タンパク質相互作用ネットワークを中心により精緻な解析を進めるとともに、遺伝子ネットワークや代謝ネットワークの数値モデルに関しても新たな展開を図りたい。

(2) ネットワーク構造解析に関する課題

パスウェイを比較してモジュールを抽出する枠組みを完成させ、引き続き、機能的に意味のあるネットワーク構造を抽出する。また、生物種間のネットワーク構造比較を進め、ネットワーク構造生成の解析につなげる。

代謝ネットワークのロバスト性解析については他手法との比較を行うとともに、より多くのネットワークデータについて計算機実験を行い、方法の改良を試みる。また、これまでに開発した離散モデルに基づく方式とフラックスモードに基づく方式には何らかの関連性があると思われるので、それを明らかにしたい。さらに、代謝ネットワークのみならずシグナル伝達ネットワーク解析への拡張に取り組みたい。

(3) ネットワーク動的挙動解析に関する課題

今回得られた遺伝子ダイナミクスの吸引領域を推測する手法を、実際にホヤの初期発生の遺伝子制御ネットワークに適用することにより、発生に関わる遺伝子機構の推測に取り組む。例えば発生の系譜上のどこで細胞運命が決定されるのか、といったことを予測できる可能性がある。

複雑な相互作用をもつ大規模進化動力学モデルの挙動については、非線形性や量の有限性の影響の研究は物理学においても発展途上であり、より独創的な研究を行う必要がある。

<成果公表リスト>

1) 論文/プロシーディング

- 0812201541
Song, J., Tan, H., Takemoto, K., and Akutsu, T.: HSEpred: predict half-sphere exposure from protein sequences, *Bioinformatics*, 24, 1489-1497 (2008).
- 0812201546
Takemoto, K. and Akutsu, T.: Origin of structural difference in metabolic networks with respect to temperature, *BMC Systems Biology*, 2, 82 (2008).
- 0812201551
Hayashida, M., Sun, F., Aburatani, S., Horimoto, K., and Akutsu, T.: Integer programming-based approach to allocation of reporter genes for cell array analysis, *International Journal of Bioinformatics Research and Applications*, 4, 385-399 (2008).
- 0901090527
Hayes, C. N., Diez, D., Joannin, N., Honda, W., Kanehisa, M., Wahlgren, M., Wheelock, C. E. and Goto, S.: varDB: a pathogen-specific database of protein families involved in antigenic variation, *Bioinformatics*, 24, 2564-2565 (2008).
- 0901061733
Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., Goto, S. and Kanehisa, M.: KEGG Atlas mapping for global analysis of metabolic pathways, *Nucleic Acids Research*, 36, W423-W426 (2008).
- 0805141120
Mochizuki, A.: Structure of regulatory networks and diversity of gene expression patterns, *Journal of Theoretical Biology*, 250, 307-321 (2008).
- 0805141120
Yoshino, Y., Galla, T., and Tokita, K.: Rank abundance relations in evolutionary dynamics of random replicators, *Physical Review E*, 78, 031924 (2008).