

生物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>

生体や細胞中では、遺伝子、タンパク質、化合物など様々な物質が相互しあう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構造をなし、その構造の推定が、実験技術および情報技術を両輪として進展しつつある。その進展により、近い将来、より多くのネットワーク構造が明らかになることが期待される。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物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、また、それにより得られた知識のデータベース化も図る。

具体的には (1) ネットワーク構造生成のための数理モデル開発、(2) ネットワーク構造解析のためのアルゴリズム開発、(3) ネットワーク動的挙動解析のための数理モデル開発、の3課題を中心に研究を行う。(1)では、代謝や相互作用といった種類による構造の違い、また、生物種によるネットワーク構造の違いを調べ、その違いを表現可能とするネットワーク構造生成のための数理モデルを開発する。(2)では、生物種の違いなどによるネットワーク構造の共通点および相違点を検出するためのネットワーク比較アルゴリズム・ツールを開発し、インターネット上で公開する。(3)では、生物情報ネットワークの持つ、安定性と環境の変化に対する適応性の両立という性質を統一的に解明するための数理モデルを開発する。また、その数理モデルに基づき、様々な生命現象を数理的に特徴づける。

<研究開始時の研究計画>

本研究では(1) ネットワーク構造生成のための数理モデル開発、(2) ネットワーク構造解析のためのアルゴリズム開発、(3) ネットワーク動的挙動解析のための数理モデル開発、の3課題を中心に行う。これらは全く独立に行うわけではなく、互いに方法、結果、データをフィードバックしながら行う。各課題の年度ごとの計画は以下のとおりである。

(1) ネットワーク構造生成のための数理モデル開発

本テーマは阿久津が中心になり以下の予定で行うが、データ収集、解析、生物学的事項に関する助言に関して五斗が協力する。

[17年度]タンパク質相互作用やタンパク質のドメイン構成に関するスケールフリー性を説明可能な数理モデルを、タンパク質の進化に基づいて開発する。

[18年度]遺伝子制御ネットワーク構造や代謝ネットワーク構造の生成モデルの開発を試みる。

[19年度]種によるネットワーク構造の違いを再現可能な数理モデルの開発を試み、そのツール化を図る。

[20年度]異なる種類のネットワークを生成することのできる数理モデルの開発を試み、ツール化を図る。

[21年度]これまでのモデルを統合し、実際のいくつかの生物種におけるネットワーク構造と似た構造を生成可能か検証する。さらに、生物進化における既知の事実と照合して、その妥当性や問題点を明らかにする。また、開発したモデルのツール化を進め、公開を試みる。

(2) ネットワーク構造解析のためのアルゴリズム開発

本研究は五斗が中心になり以下の予定で行うが、アルゴリズム開発において阿久津が協力する。

[17年度]生物種の違いなどによるネットワーク構造の共通点および相違点の検出法を開発する。

[18年度]代謝反応式からの知識抽出アルゴリズムやネットワークからのモジュール検出法を開発する。

[19年度]配列やタンパク質立体構造とネットワークの関連性を解析するアルゴリズムを開発する。

[20年度]開発した手法をツール化し、WEBサーバーを通じてのインターネット上での公開を試みる。

[21年度]開発したツールを用いて複数の生物種の解析を行い、得られた知識のデータベース化を図る。

(3) ネットワーク動的挙動解析のための数理モデル開発

本研究は望月および時田が中心になって行うが、離散的な挙動解析手法については阿久津も分担する。

[17年度]遺伝子発現量の動的変化に関する微分方程式モデルおよび確率モデルを開発する。また、時田が以前に開発した多種進出力学系モデルの遺伝子発現量解析への適用法を検討する。

[18年度]前年度に開発したモデルに基づき、遺伝子ネットワークにおける安定性の解析や、安定性と細胞の異常化(癌化など)との関連性を調べる。

[19年度]遺伝子ネットワークにおいて安定性が失われた際の再安定化手法の開発を試みる。

[20年度]これまでに得られた手法の他のネットワーク解析への適用を試み、環境変化に各種ネットワークがどのように適用するかについて定性的および定量的な解析を行う。

[21年度]これまでのモデルに基づき、実際の実験により検証可能な新たな生命現状の予測を試みる。

<研究期間の成果>

(1) ネットワーク構造生成のための数理モデル開発

本課題では、タンパク質ドメイン構成、タンパク質相互作用ネットワーク、代謝ネットワークについては計画以上に精緻なモデル化を達成することができたが、遺伝子ネットワークや統合モデルについては計画どおりとはいかず課題として残された。

タンパク質ドメイン構成の進化については、まず、単数ドメインだけを持つタンパク質を対象に、ドメインの重複と突然変異に

基づく進化の数理モデルを構築した。その結果、同じドメインを持つタンパク質の個数の分布がべき乗則に従うことを示すことができた。そして、べき指数が、重複と突然変異の比の関数として導かれることを示した。さらに、同じドメインを持つタンパク質を枝で結んだタンパク質ドメインネットワークについても解析を行い、このネットワークが、大域的には他のスケールフリーネットワーク同様に負のべき乗則に従うが、局所的には正のべき乗則に従うという興味深い性質を見出した。この性質を数理的に示すとともに、データベース解析によっても確認した(0606171412)。

次に、上記で示したタンパク質のドメイン構成の進化モデルをもとに、ドメイン間の相互作用の有無が一様ランダムに決定されるという仮定を追加して、タンパク質相互作用ネットワークの数理モデルを構築した。その結果、相互作用する確率が低い場合には次数分布がべき乗則にほぼ従うネットワークが得られるが、相互作用する確率が高い場合には次数分布が指数分布に近いネットワークが得られるとの興味深い性質を見出した。この性質を数理的に導くとともに、データベース解析と計算機シミュレーションを組み合わせた解析によっても確認した(0902211520)。

これまでに述べてきたモデルでは数理解析の簡単化のために、各タンパク質は1個のドメインからなるとの仮定を置いてきた。しかしながら実際には複数のドメインからなるタンパク質が多数存在する。そこで、複数ドメインからなるタンパク質の進化モデルを構築した。複数ドメインの進化には様々な可能性があるため、2種類の数理モデルを構築した。一つは、タンパク質に対応する頂点とドメインに対応する頂点からなる二部グラフを用いた数理モデルであり、もう一つは、前述の重複と突然変異に基づく進化モデルにドメインの融合や内部重複を加えた進化モデルである。いずれにおいても、ドメインの出現頻度はべき乗則に従うが、個々のタンパク質に含まれるドメインの種類は指数分布に従うという興味深い性質が得られた。この性質はデータベース解析により知られていたものであったが、本研究により理論的にも導くことができた(0912021613)。

代謝ネットワークに関しては種による違いをモデル化するために、まず、原核生物における生育温度とネットワーク構造の関連性を調べた。そのために113種類の生物を対象にKEGGデータベースなどを用いて、辺密度、次数分布、クラスター係数、部分グラフの分布の解析を行った。その結果、生育温度が低くなるにつれ構造化やモジュール化が進展することを見出した(0710261720)。さらに、その事を説明する数理モデルを構築した。このモデルでは、新規化合物を合成する反応と既存のパスをショートカットする反応の2種類の反応の追加に基づいており、その割合の違いにより生育温度とネットワーク構造の関係を説明することに成功した(0812201546)。[阿久津]

(2) ネットワーク構造解析のためのアルゴリズム開発

本課題では、実施年度が若干前後することはあったものの、全体的には当初予定した研究計画をほぼ達成することができた。

生物種の違いなどによるネットワーク構造の共通点および相違点の検出法開発については、主に代謝ネットワークとタンパク質のドメイン組合せネットワークについて研究を進めた。代謝ネットワークでは従来の酵素反応のネットワークだけでなくフィードバック阻害やフィードフォワード活性化にも着目し、酵素の阻害因子や活性化因子など制御因子の情報を収集してネットワークを構築した。このネットワークから特徴的な制御因子やネットワークモチーフを検出するために、複数の生物種でネットワークを比較する方法を開発し、制御因子として利用される化合物の種類という観点からヒトと大腸菌の違いを明らかにした(0704300718)。

もう一方のネットワーク構造解析のターゲットであるタンパク質のドメイン組合せネットワークについては、全ゲノムが決定された真核生物について、Pfamドメインを全タンパク質に割り当て、同じタンパク質に存在するドメイン同士をつないだネットワークとして構築した。各ドメインの古さを推定するアルゴリズムを開発し、ドメインの組合せが進化の過程でどのように変化してきたかを解析できるようにした。その結果、多様なシグナル伝達系を持つ生物種群や免疫系を持つ生物種群に特有なドメイン組合せ、それらの祖先系で出現してきた痕跡を見いだすことができ、機能との相関を明らかにした(0708061727)。

代謝反応式からの知識抽出アルゴリズムやネットワークからのモジュール検出法の開発については、KEGG LIGAND データベースやPATHWAY データベースを用いて解析する手法を開発した。代謝反応式からの知識抽出では、KEGG LIGAND に登録されている反応から基質と生成物の二項関係を定義して、両者の化学構造変化パターンを抽出する手法を開発した。抽出した変化パターンは、マニュアルでチェックしたのち、KEGG LIGAND の反応パターンデータベース RPAIR として公開している。さらに、反応パターンと代謝ネットワークとの関連を解析し、分解系によく使われるパターンを用いた化合物分解経路予測システムを開発した(0911261849)。

代謝ネットワークやタンパク質間相互作用ネットワークを生物種間で比較すると、そのノードである遺伝子やタンパク質の進化系統的な保存度(系統プロファイル)によりサブネットワークごとにばらつきがある。本研究では系統プロファイルに基づいたネットワークからのモジュール検出法を開発した。これをゲノムが決定された生物種の代謝ネットワークに適用して、機能的に関連づけられたモジュールが抽出できることを示した(0606170656)。

配列やタンパク質立体構造とネットワークの関連性を解析するアルゴリズムについては、配列や立体構造に限らず、上記の系統プロファイルや遺伝子発現など、様々な尺度に基づく遺伝子間の類似度と機能の類似度とを関連づけるアルゴリズムを開発し、代謝ネットワークの再構築に応用した(0704300708)。また、タンパク質間相互作用ネットワークと配列モチーフの関連性を解析し、神経変性疾患に関わるタンパク質の特徴抽出に応用した(0710231055)。

代謝ネットワークでは化合物の構造類似性も重要な要素であるため、低分子化合物の構造類似性と代謝ネットワークとの関連性を解析するアルゴリズムを開発し、環境ホルモンの分解系における毒性変化の予測に応用した(091126859)。

開発した手法のツール化とWEBサーバーでの公開については、以下の4つのツールを公開した。

- ① 様々な類似度を基にした遺伝子間の機能相関ネットワーク予測法を実装したWEBサーバー GENIES を公開した。
 - ② KEGG の代謝パスウェイマップ群を一つに統合した全体像として、その上に様々な情報をマッピングするためのWEBサーバー KEGG Atlas を公開した(0901061733)。
 - ③ 同様に、様々なオミックスデータをネットワーク解析するためのJavaアプリケーションである KegArray を開発して公開した。KegArray の特徴はマイクロアレイなどの遺伝子発現情報とメタボロームなどの化合物の情報とを同時に扱い、KEGG のパスウェイにマッピングできることである(0911262114)。
 - ④ 反応パターンデータベースを応用して新規反応の反応パターンを抽出し、酵素反応分類番号を割り当てるWEBサーバー E-zyme を公開した(09011262107)。
- 開発したツールを用いた複数生物種の解析と得られた知識の

データベース化については、真核生物における糖転移酵素の遺伝子ファミリーを糖鎖合成経路と関連づけた解析を行い、各生物種で合成可能な糖鎖構造の予測につなげた (0911262100)。また、同様の解析を脂肪酸合成経路に関しても行った。ここで得られた糖転移酵素や脂肪酸合成酵素の遺伝子ファミリーの情報は KEGG データベースの一部として公開されている。

ネットワーク解析は疾患の解析にも応用し始めている。既に述べた神経変性疾患以外にも、循環器系の疾患であるアテローム性動脈硬化症において、発症頻度の男女差を説明する遺伝的マーカーの探索に遺伝子発現パターン解析とネットワーク解析を組み合わせて応用している (0911262122)。[五斗]

代謝ネットワークの遺伝子変異に対する頑健性解析にも取り組み、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、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 (0912021146)。また、様々な外部刺激に対し顕著に反応する部分ネットワークを遺伝子発現データを用いて同定する方法を開発し、酵母のネットワーク解析に適用して有効性を確認した (0701221322)。[阿久津]

(3) ネットワーク動的挙動解析のための数理モデル開発

ネットワーク動的挙動解析のための理論構築が目標の一つであったが、これに関して「ネットワーク構造と力学挙動の関係」を示す一般的な理論を打ち立てることに成功した (0601270049, 0805141120)。

基本アイデアは簡単であり、「各生体分子の活性ダイナミクスは、それを制御する因子の活性状態の関数である」、という自明の論理だけを用いる。この考えには二つの側面があり、生体分子活性状態の「不和合性 (incompatibility)」、及び「独立性 (independency)」と名づけた。前者の「不和合性」の性質によって、活性状態の定常状態の可能性を絞り込み、可能な状態数の上限を決定できる。一方で後者の「独立性」から、分子の活性状態の可能な組み合わせについての条件を導くことができる。例えば、ウニの初期発生にかかわる遺伝子ネットワークを解析したところ、多数の遺伝子を含む制御ネットワークの中から、遺伝子発現多様性に重要な、ごく少数の遺伝子を抽出できた。さらに遺伝子発現パターンの情報を取り入れることで、当時未発見の遺伝子発現制御を予測できた。この予測された制御は、後の実験で存在が確認されたものであった。

この理論を展開することで、制御ネットワークの構造と生体分子の振る舞いに関して、幾つかの法則を導くことができた。例えば遺伝子発現パターンの多様性は、細胞分化の多様性の起源だと考えられており、遺伝子制御ネットワークからの理解が試みられてきた。この理論によれば、遺伝子の数や遺伝子間の相互作用の数は、遺伝子発現の多様性とは無関係であり、相互作用ネットワークにおけるループ構造 (とそれに相当する構造) のみが重要であることが分かる。また絡まりあったループ構造に対して、それを重要な構造に分解することもできる (0805141120)。

この研究は実験と組み合わせることで、力を発揮する。理論はモデルの数式にほとんど依存しないため、遺伝子制御、シグナルパスウェイなどの対象を区別せず、ほとんど全ての制御ネットワークに対して適用できる。また既知の情報だけから未知の制御の予測も可能であり、現在様々な生体分子制御ネットワークを対象に、実験生物学者と検証的共同研究を行っている。

上記以外に制御ネットワークに限らず、様々な生命現象に対して、数理モデルを用いてその動態を理解する研究を、実験生物学者との共同研究として行った。

例えば、シアノバクテリア概日リズムについては、概日リズム

を制御するタンパク質 KaiC のリン酸化振動に対して、数理的解析を行い、KaiC が複数の異なるリン酸化状態をもつことを予測した (0701111521, 0608060705)。

また、神経細胞形態形成を数理的に研究した。ある種の神経細胞は、樹状突起を曲面上で、むら無く一様に分布させることから、space filling type と呼ばれる。この神経細胞の空間秩序の原理を理解する数理モデルを構築し解析した。解析により、樹状突起がもつ3つの空間制御：一様分布、空間分割、再生が同一の原理によって理解できることを示した (0801211521)。

さらに、線虫の温度走性 (0801211508)、双子葉植物の葉脈の多様性に関する研究 (0701111509, 0701111503) や、細胞内極性の自律的創出過程 (0801211514, 0801211508)、制限酵素の進化 (0608060649) などについても研究を行った。[望月]

生態系における種、細胞内の代謝物質、タンパク質、mRNA などの生物量の分布と Zipf の法則との関係を導き、非平衡統計力学的手法を用いた解析を行なった (0608062027)。さらに、HIV などのウイルスと免疫系などの大規模進化動力学モデルの解析を行った。また、HIV などのウイルスと免疫系のモデルにおける、AIDS 発症まで無症候期間の長さが交差免疫の強さに依存することなどを理論的に明らかにした。

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手法を用いた研究を行い、生物量の分布 (abundance distributions) を初めて解析的に導いた (0608062015)。

そして、この大規模進化動力学モデルに対する非平衡統計力学的手法を発展させ、代謝反応などに特有な多体相互作用をもつ系や、周期解やカオスを含む本質的に動的な非平衡系の解析を行い、ダイナミクスが平衡状態に収束するパラメータ領域に対しては、系の安定性や外界から流入する資源量に対する依存性を解析的に得た (0710242350)。さらに、非平衡系に対して、生物量の分布を与える理論を構築した (0901080102)。その結果、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。[時田]

遺伝子ネットワークの離散数理モデルであるブーリアンネットワークの解析および制御にも取り組み、多くの理論的成果をあげることができた。具体的には、定常状態を高速に計算する一連のアルゴリズムの開発 (0912021118, 0911201303, 0911201141, 0904221538, 0708061456, 0708061445)、ブーリアンネットワークの制御問題の定式化とその計算複雑度の解析 (0702111145)、さらに、定常状態の制御問題の定式化とアルゴリズム開発 (0812201537) といった成果をあげた。また、既存知識から得たネットワークを実験データに合わせて修正するという問題をブーリアンネットワークを用いて「ネットワーク補完」という名称で定式化し、その計算複雑度を明らかにした (0912021123)。[阿久津]

<国内外での成果の位置づけ>

個々の成果はバイオインフォマティクス、数理生物学、発生生物学、分子生物学、生態学、物理学などの主要英文誌において発表されており高いレベルを維持することができたと考えられる。また、これらの成果は、国内外での招待講演や国内学会誌および商業誌における解説記事執筆などにも反映されている。以下、個別の研究課題について記載する。

(1) ネットワーク構造生成に関する成果の位置づけ

ネットワークのスケールフリー性に関する数理モデルについては着実に新規な成果をあげ、国内外で認知されていると考えられる。しかしながらスケールフリーネットワークに関しては数多くの研究がなされており、ブレイクスルーとなるほどの成果をあげるには至らなかった。なお、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の関係性の論文は Highly accessed とマークされ、関心の高さが示された。[阿久津]

(2) ネットワーク構造解析に関する成果の位置づけ

構造解析に関しては有用なデータベースやアルゴリズムが得られ、国内外に広く認知される成果が得られたと考えられる。

オープンアクセス論文誌に掲載された論文のうち、代謝制御ネットワーク解析に関するものとタンパク質ドメイン組合せネットワーク解析に関するものは Highly accessed とマークされ、関心の高さが示された。

代謝ネットワークの解析については、クラミジアゲノム、酢酸菌ゲノム、ゼニゴケ EST など、国内のいくつかのゲノムプロジェクトとの共同研究に発展した。さらに、疾患に関するネットワーク解析については海外との共同研究にも発展した。現在進行中のアテローム性動脈硬化症のネットワーク解析はスウェーデンのカロリンスカ研究所が持つサンプルデータを共同で解析しているものである。また、GENIES をはじめとする機能相関ネットワーク解析アルゴリズムの開発は、今年度から始まったフランスのキュリー研究所との癌の共同研究へと発展した。[五斗]

(3) ネットワーク動的挙動解析に関する成果の位置づけ

様々な生命現象において、生体分子が複雑な制御ネットワークのシステムを構築していることが明らかにされており、それら分子レベルの情報は、現在ますます増加しつつある。一方で相互作用ネットワークの構造が、遺伝子発現や酵素活性などのダイナミクスに対して、どの様な影響を与えるのか、或いはどの様な意味を持っているのか、といった理解は、これまでほとんど進んでいなかった。本研究で提案した理論は、制御の構造とダイナミクスとの間を結びつける初めてのものである。また、次のような優れた特色を持っている。(i) 理論の一般性が高く、対象を選ばず適用できる。(ii) 理論と実験情報を組み合わせることで、既知の情報から未知の情報を引き出せる。このような利点から、複雑な制御システムに取り組む生物学者から高い評価を受けている。現在、複数のグループとの共同研究が複数進んでいる。具体的には、EGF 受容体下流のシグナル伝達系について、理化学研究所の佐甲靖志主任研究員、岡田眞里子チームリーダーとの共同研究、線条体可塑性シグナル伝達系について、大阪大学の館野高准教授と共同研究を進めている。また制御ネットワークの動態を制御構造から理解する研究は、アイデア自体が画期的であり、数理科学的にも大きな評価と機体を得ている。世界の第一級の数学者から、共同研究の申し出を受けている。

また、個別の生命現象を対象とした数理モデル研究については、それらの分野の実験生物学者から高い評価を得た。シアノバクテリア概日リズムについては、予測に対応するような事実が実験的に明らかにされつつあり、実験・理論の研究者から多くの注目を集めた。神経細胞樹状突起形態形成の数理モデルに対しては、発生物学者、神経生物学者を中心に高い評価を得ている。[望月]

複雑な相互作用をもつ大規模進化動力学モデルにおける、個体数、蛋白量、代謝物質質量など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、それらの量の有限性や、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。上記の研究結果は著書にもまとめられ(060862009, 0702111244, 0807181749, 0911301456)、さらに、数多くの国際会議において招待講演を行ったことから国内外から非常に高く評価されていることがわかる。[時田]

ブーリアンネットワークに関する研究は研究者人口が少ないこともあり、特にブーリアンネットワークのアルゴリズムという観

点からは世界でもトップレベルに位置していると考えられる。国内においてはほとんど評価されていないが、海外からはチュートリアル講演や本の一部の章の執筆を求められるなど、高く評価されていると考えられる。特に制御問題に関しては本研究における定式化を基に、中国科学院、カーネギーメロン大学、北陸先端大・東工大の研究グループなどが研究を進めており、今後、さらに広がることが期待できる。[阿久津]

<達成できなかったこと、予想外の困難、その理由>

個々にはレベルの高い成果を数多く出すことができ研究計画の多くを達成することができた。また、代表者と分担者間では学会への招待なども含め何度も意見を交換しあう機会を持つことにより貴重なフィードバックを得ることができた。しかしながら、個々の研究や成果を連携させ全体としてまとまった成果を出すということはできなかった。理論的研究が中心であるという研究の性格上、個別に研究が進むのは仕方のない面もあるが、代表者がメンバー間の共同研究をより積極的に促すべきであったかもしれない。

本研究の目標の一つとして、生物学的実験により検証可能な仮説の提唱、および、実験系研究者との共同研究によるその検証があった。一部の分担者は共同研究を円滑に行い生物学的に有用な成果を得ることができたが、グループ全体としては多くを達成することはできなかった。しかしながら、いくつかの共同研究が進行中であるので、本研究プロジェクトがそのきっかけとなったということは評価できる。

以下では全体としてではなく個々の課題の問題点について記述する。

(1) ネットワーク構造生成における困難や理由

本研究ではスケールフリーネットワークの次となる概念の提案を目指したが、その達成には至らなかった。しかしながら、この困難は本グループだけではなく、分野全体が直面している困難であると考えられる。新規概念は一朝一夕に出てくるわけではなく、地道な研究の積み重ねの上に出てくるものである。地道な研究という面では配列進化に基づく精緻な数理モデルを構築でき、有用な成果が得られたと考えられる。

生物種ごとのネットワークの違いを表現可能なモデルの構築に関しては、代謝ネットワークに関してはある程度達成できたが、タンパク質相互作用ネットワークや遺伝子ネットワークについては達成することができなかった。その理由として、代謝ネットワークと異なり、これらのネットワークにおいては種ごとの違いまで適切に表現された信頼性の高いデータが存在しなかった、ということがあげられる。また、数理モデル構築に時間をとられてしまったこと、および、現段階でのツール化の意義があまり感じられなかったこともあり、構造生成モデルのツール化は達成できなかった。[阿久津]

(2) ネットワーク構造解析における困難や理由

ネットワーク解析手法の確立や、いくつかのターゲットに対するネットワーク解析による機能解明については、十分に研究計画を達成できたが、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普遍的なメカニズムを解明するには至らなかった。また、ネットワーク構造解析に関しては当初の目的を達成できたと考えているが、構造生成や動的挙動との連携までは手が回らなかった部分がある。普遍的なメカニズムの解明には、両者の更なる連携が必要になると考えられるが、解析できるデータの充実化も必要だと思われる。[五斗]

(3) ネットワーク動的挙動解析における問題点

ネットワークの構造と発現多様性に関する理論を実際のデータ

に適用する上で、最初に選んだ遺伝子制御ネットワークは、理論の対象としては、必ずしも適切ではなかった。解析を進めた結果、未知の制御として膨大な情報が導かれてしまい、実験に対する予測としては、実利が乏しいことが分かった。これは対象とする制御ネットワークを変えることで解決できた。上記の遺伝子ネットワークに対しても、実験データがもう少し蓄積した段階で実利的な適用ができると考えている。

神経細胞樹状突起の形態形成研究に関しては、今回計算したものは表皮近くの神経細胞など、2次元的に樹状突起を張り巡らせるものであった。さらに3次元空間における樹状突起パターンを計算しようとしたが、思った以上に難しく、時間がかかった。形態形成の数理的研究として、3次元の形状を扱えるようになることは一般的に必要なので、この問題を解決したい。[望月]

HIVと免疫系の進化動力学モデルは典型的な非線形非平衡系であるため、これまで解析的な取り扱いが行われておらず、本研究においてもシミュレーション研究が主で、統計力学的な解析には至っていない

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。

非平衡系の解析については、ダイナミクスが平衡状態および特殊な定常状態に収束する解に対しては解析的な取り扱いが可能になったが、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はその数学的取り扱いの困難さから未だに成功していない。

実証研究者との共同研究により、生物量の分布の理論の実証を予定していたが、実証データの収集の困難さから未だに共同研究途上にある。[時田]

<今後の課題、展望>

(1) ネットワーク構造生成における課題と展望

タンパク質ドメインの進化モデルに関してはマルチドメインタンパク質の進化までは拡張することができたが、その進化モデルを用いたタンパク質相互作用ネットワークのモデル化と解析が今後の課題として残されている。

種ごとの違いを生成可能な数理モデルについては、生育温度以外の特徴を説明可能な数理モデルの構築、タンパク質相互作用ネットワークや遺伝子ネットワークに適用可能な数理モデルの構築が今後の課題として残されている。さらに、正常細胞と腫瘍細胞の違いを説明可能な生成モデルの構築も今後の課題となっている。[阿久津]

(2) ネットワーク構造解析における課題と展望

ネットワークを解析する上で利用可能なデータは今後ますます増えると考えられる。例えば、次世代シーケンサーを用いたメタゲノムやパーソナルゲノムなど新しい技術により、より詳細な条件の違いを解析できるデータが取得できるようになる。また、脂質合成系や糖鎖合成系に関しては、メタボローム解析から発展したりビドーム、グライコームプロジェクトも出てきている。これらのプロジェクトから、環境変化や安定性に関する量的なデータが取得でき、動的挙動解析との連携を進めることができれば、細胞の安定性、適応性メカニズムの解明に近づくことができると考えられる。[五斗]

(3) ネットワーク動的挙動解析における課題と展望

「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。今後は様々な生命システムに適用を繰り返し、生命システム理解の方法として、実用性と信頼性を高めていく予定である。また生物学者に加えて、数理科学者から新たな手法としての期待を受けている。数理科学的にも展開すること

で、生命科学と数理科学とのさらなる連携に寄与することを考えている。[望月]

HIVと免疫系の進化動力学モデルの非平衡ダイナミクスの数学的取り扱いについては、海外の有力研究グループとの共同研究を通じてブレークスルーを見だしつつある。

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は、生物量の有限性の効果以外についてはほぼ解決したといえるが、現実の生物ネットワークにおいては生物量は有限であるため、より定量的、現実的な理論の構築のための研究の発展が望まれる。

周期解やカオスを含む本質的に非平衡動力学的な挙動に関しては物理学においても発展途上の領域であり、摂動論などを用いた、平衡領域から非平衡領域への非平衡相転移の理論の構築が必要である。

生態学、細胞生物学における実証研究の発展により、定量的な実証データが蓄積しつつあるので、それらによる理論の検証、さらには理論予測に基づく実証研究など、理論と実証研究の更なる密な連携が期待される。[時田]

生物情報ネットワークの制御についてはブーリアンネットワークを用いた定式化を行ったが、ブーリアンネットワークだけではダイナミクスを十分に記述できない場合があるため、他のシステムとの融合が必要である。シミュレーションであれば拡張は比較的容易であるが、制御を対象とすると慎重に拡張することが必要である。まずは制御理論分野において理論基盤が確立している線形制御理論との融合が必要であると考えられる。また、ブーリアンネットワークの制御だけにおいても計算複雑度の壁があり、大規模なネットワークの制御が難しい。計算複雑度に関する精緻な解析とともに、実用的な制御系列計算手法の開発が今後の課題となっている。制御のためにはダイナミクスも含めたネットワークの正確なデータが必要であるが、実際にはそのようなデータは利用不可能である。そこで、ネットワークの詳細が不明でもシステム同定を行いつつ制御する手法や、ロバストな制御方法の開発が必要である。[阿久津]

<研究期間の全成果公表リスト>

1) 論文/プロシーディング

1.0912021118

Hayashida, M., Tamura, T., Akutsu, T., Ching, W-K. and Cong, Y.: Distribution and enumeration of attractors in probabilistic Boolean networks, IET Systems Biology, 3, 465-474 (2009).

2.0902211520

Nacher, J. C., Hayashida, M. and Akutsu, T.: Emergence of scale-free distribution in protein-protein interaction networks based on random selection of interacting domain pairs, BioSystems, 95, 155-159 (2009).

3.0911262114

Wheelock, C. E., Wheelock, A. M., Kawashima, S., Diez, D., Kanehisa, M., van Erk, M., Kleemann, R., Haeggstrom, J. Z. and Goto, S.: Systems biology approaches and tools for investigating cardiovascular disease, Molecular Biosystems, 5, 588-602 (2009).

4.0911262208

Nakazato, K. and Mochizuki, A.: Steepness of thermal gradient is essential to obtain a unified view of thermotaxis in *C. elegans*, Journal of Theoretical Biology, 260, 56-65 (2009)

5.0812201546

Takemoto, K. and Akutsu, T.: Origin of structural difference

- in metabolic networks with respect to temperature, *BMC Systems Biology*, 2, 82 (2008).
- 6.0901061733
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).
- 7.0805141120
Mochizuki, A.: Structure of regulatory networks and diversity of gene expression patterns, *Journal of Theoretical Biology*, 250, 307-321 (2008).
- 8.0901080102
Yoshino, Y., Galla, T. and Tokita, K.: Rank abundance relations in evolutionary dynamics of random replicators, *Physical Review E*, 78, 031924 (2008).
- 9.0708061456
Ching, W-K., Zhang, S-Q., Ng, M. K. and Akutsu, T.: An approximation method for solving the steady-state probability distribution of probabilistic Boolean networks, *Bioinformatics*, 23, 1511-1518 (2007).
- 10.0702111145
Akutsu, T., Hayashida, M., Ching, W-K. and Ng, M. K.: Control of Boolean networks: hardness results and algorithms for tree structured networks, *Journal of Theoretical Biology*, 24, 670-679 (2007).
- 11.0708061727
Itoh, M., Nacher, J. C., Kuma, K., Goto, S. and Kanehisa, M.: Evolutionary history and functional implication of protein domains and their combinations in eukaryotes, *Genome Biology*, 8, R212 (2007).
- 12.0710242350
Yoshino, Y., Galla, T. and Tokita, K.: Statistical mechanics and stability of a random eco-system, *Journal of Statistical Mechanics*, P09003 (2007).
- 13.0704300708
Yamanishi, Y., Mihara, H., Osaki, M., Muramatsu, H., Esaki, N., Sato, T., Hizukuri, Y., Goto, S. and Kanehisa, M.: Prediction of missing enzyme genes in bacterial metabolic network: A reconstruction of lysine degradation pathway of *Pseudomonas aeruginosa*, *FEBS Journal*, 274, 2262-2273 (2007).
- 14.0801211508
Ishihara, S., Otsuji, M., and Mochizuki, A.: Transient and steady state of mass-conserved reaction-diffusion systems, *Physical Review E*, 75, 015203 (2007).
- 15.0801211514
Otsuji, M., Ishihara, S., Co, C., Kaibuchi, K., Mochizuki, A. and Kuroda, K.: A mass conserved reaction-diffusion system captures properties of cell polarity, *PLoS Computational Biology*, 3, 1040-1054 (2007).
- 16.0606171412
Nacher, J.C., Hayashida, M. and Akutsu, T.: Protein domain networks: scale-free mixing of positive and negative exponents, *Physica A*, 367, 538-552 (2006).
- 17.0606170656
Yamada, T., Kanehisa, M. and Goto, S.: Extraction of phylogenetic network modules from the metabolic network, *BMC Bioinformatics*, 7, 130 (2006).
- 18.0608062015
Tokita, K.: Statistical mechanics of relative species abundance, *Ecological Informatics*, 1, 315-324 (2006).
- 19.0601261956
Nacher, J.C., Ueda, N., Kanehisa, M., and Akutsu, T.: Flexible construction of hierarchical scale-free networks with general exponent, *Physical Review E*, 71, 036132 (2005).
- 20.0601270049
Mochizuki, A.: An analytical study of the number of steady states in gene regulatory networks, *Journal of Theoretical Biology*, 236, 291-310 (2005).
- 番号のみのリスト (上記以外)
- [2009 年度]
- 0912021613, 0912021146, 0912021123, 0911262122, 0911262107, 0911262100, 0911201312, 0911201303
- [2008 年度]
- 0911201141, 0905151534, 0904221538, 0904221533, 0902211515, 0901090527, 0812201557, 0812201551, 0812201541, 0812201537
- [2007 年度]
- 0911261859, 0911261849, 0812201532, 0801211521, 0710261720, 0710261717, 0710261713, 0710231055, 0708061451
- [2006 年度]
- 0708061445, 0704300731, 0704300718, 0701231025, 0701221338, 0701221329, 0701221322, 0701111521, 0701111509, 0701111503, 0608060705, 0608060649, 0606171406, 0606170733, 0606140017
- [2005 年度]
- 0606171401, 0601270533, 0601270524, 0601262030, 0601262022, 0601262009, 0601262002
- 2) 図書
- 0912021538, 0912021535, 0911301456, 0807181749, 0608062009
- 3) データベース/ソフトウェア
- 1.0702120524 (ソフトウェア)
GENIES: gene network inference engine based on supervised analysis
<http://www.genome.jp/tools/genies/>
- 2.0702120531 (データベース)
ODB: a database of operons accumulating known operons across multiple genomes
<http://odb.kuicr.kyoto-u.ac.jp/>
- 3.0507070117
LIGAND: database of chemical compounds and reaction in biological pathways
<http://www.genome.jp/ligand/>