

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年5月31日現在

機関番号：32661

研究種目：若手研究（B）

研究期間：2011～2012

課題番号：23700352

研究課題名（和文） 転写調節ネットワークにおけるモジュール性の出現解明

研究課題名（英文） Elucidation of modularity in transcriptional regulatory networks

研究代表者

ホセ ナチエル（Jose Nacher）

東邦大学・理学部・准教授

研究者番号：60452984

研究成果の概要（和文）：本研究では、パイオ macromolecules (PRM) と遺伝子調節ネットワークを関連する非コード RNA (ncRNA) とタンパク質間の機能的な相互作用によって定義されたネットワークのコミュニティ構造を比較する。コミュニティサイズの分布は、複雑系構造を明らかにするスケールリング則を示した。また、K-クリークサブグラフと重なった ncRNA 分子の識別コミュニティが、潜在的に特定の機能に関連付けることができることを示している。最後に、我々はまた、数学的にのみ二部ネットワーク情報を使用して投影されたネットワーク構造を予測する方程式を導出した。

研究成果の概要（英文）：In this research, we investigate and compare the community structure of the network defined by functional interactions between non-coding RNAs (ncRNAs) and proteins related bio-macromolecules (PRMs) and the transcriptional regulatory network. Our findings show that the distributions of community sizes followed a scaling-law revealing a complex structure. In addition, the k -clique sub-graphs and overlaps show that the identified communities of the ncRNA molecules could potentially be associated with certain functions. Finally, we also derived mathematically the equations that predict projected network structure using only bipartite network information.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	1,400,000	420,000	1,820,000

研究分野：総合領域

科研費の分科・細目：情報学・生体生命情報学

キーワード：転写調節ネットワーク、ncRNA、モジュール性、ネットワーク解析

1. 研究開始当初の背景

遺伝情報が mRNA の分子を用いて DNA からタンパク質へと流れることはよく確立されている。結果として、それは遺伝

子やタンパク質との間の1対1の対応を考察することができる。これは単純な原核生物においては正確であると思われるが、真核生物では、起源と非タンパク

質コード分子の高い豊富な役割はまだ不明である。急速な技術進歩は、非タンパク質コード遺伝子の割合は、ヒトにおける98.5%に上昇することが示されている。この結果は、遺伝情報の処理に関する現在の知識は、主に不完全であることを示唆している。ncRNA機能との相互作用の構造に光を当てるためには、ネットワーク分析を使用し、遺伝子調節ネットワークから得られた結果と比較した。

2. 研究の目的

この研究では、転写調節ネットワークとnon-coding RNAs (ncRNA)-タンパク質相互作用におけるモジュール性や複雑なネットワーク構造を比較して分析するのが目的である。特に、二部ネットワークとして表される生物学的ネットワークのモジュールの構造の研究を行った。また、二部ネットワークからの予測ネットワーク構造の数学的予測に関する研究も、同様に二部表現を用いた遺伝子制御やncRNAのネットワークの両方のコミュニティと制御機能として、行われた。

3. 研究の方法

データベース

大腸菌のための標的遺伝子を調節する転写因子 (TF) によって定義される転写調節相互作用がRegulonDBデータベースから収集した。S. セレビスエ生物に対応するデータセットは、[7]で最初に利用したウリアロンのウェブサイトからダウンロードした。ncRNA-PRM用ペアワイズ相互作用ネットワークはNPInterデータベースから得た。このような大腸菌、S. cerevisiae、Caenorhabditis虫、ショウジョウバエ、ハツカネズミとホモサピエンスなど6モデル生物からのデータセットがダウンロードされ、6部グラフを構築した。データ分析の統計を高めるために、我々は、すべての利用可能な生物間の相互作用が含まれる1つ以上のネットワークを構築した。これから我々はすべてのncRNAとタンパク質として、このネットワークを参照した。700個収集された相互作用の合計数は、98個のncRNA分子および425個の蛋白質との間に構築した。PRMのほとんどがタンパク質であるが、mRNAおよびゲノムDNAと関連分子をコードするタンパク質を伴ういくつかの相互作用が

ある。一方、人間のncRNA媒介性相互作用を収集するデータセットは、34個のncRNAおよび154個のタンパク質分子(図1参照)からなっている。

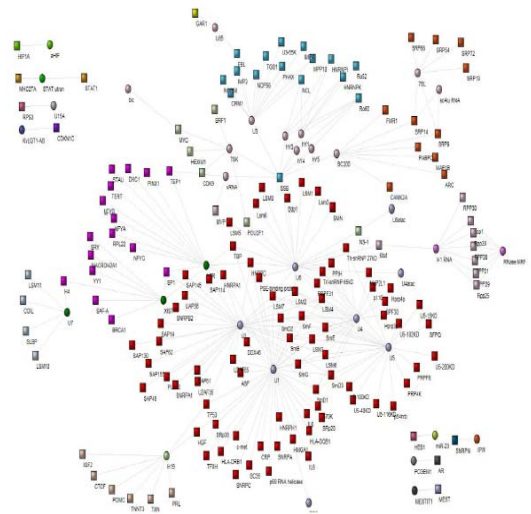


図1. ヒトのncRNA-タンパク質ネットワークのモジュール性

モジュール性の計算と解析

2つのノードのタイプで構成される二部ネットワークは、元の二部ネットワークの投影と呼ばれる、二つのネットワーク上に(つまり変換)に投影することができる。各投影されたネットワークノードは、その後、一種類で構成されている。TF/ncRNAと標的遺伝子/タンパク質に対する二部グラフは、正式には、 $G = (T, P, E)$ と定義される。TはTF/ncRNA分子の集合であり、Pは標的遺伝子のセット/タンパク質、およびEはそのTとPからの2つのノードのリンクの辺の集合として定義することができる。二部グラフの投影が大幅エッジの数を増加させ、境界ノードにおいて、正確な結果にはつながらない場合があるので、我々が直接二部ネットワーク自体のモジュール性を計算する。我々は二部構造のモジュール化機能を最大限にシミュレーテッドアニーリングベースのアルゴリズムを使用した。以前のアルゴリズムを使用して、我々は、モジュールを識別することができた。そこで、遺伝子調節とncRNA相互作用ネットワークの両方でこれらのモジュールの大きさの分布を計算することができる。

数学的解析と投影ネットワークの予測

平均場理論近似を用いることにより、我々は二部ネットワーク情報を利用して投影されたネットワークの構造を予測する方程式を導き出すことができる。この方法は、遺伝子制御ネットワークを含む、実際の生体データ情報に適用した。さらに、我々は二部ネットワークにおける分析転写およびncRNAの相互作用ネットワーク上で制御機能を解析できた。

k-クリークの解析

相補的なコミュニティの定義は、コミュニティ内の特定の分子は多くの他の分子に接続することができるという観察に基づいている。つまり、共有するノードとそのサブグラフ(完全に接続されている)の小さい完全な組成物としてコミュニティを見出すことができます。これらの完全な部分グラフは、kを部分グラフ内のノードの数を示したkクリーク、と呼ばれている。我々は、ncRNAの相互作用の両方のk-クリークを識別するためにクリーク浸出法を用いた。

4. 研究成果

今回の知見は、ncRNA - タンパク質相互作用が高いモジュール性と普及群集構造を持っていることを示している。大腸菌および S. セレビスエに対応する転写調節ネットワークのモジュールレベルも計算され、ncRNA を (図 2 参照) を用いて算出されたものと比較される。S. セレビスエと大腸菌用転写因子 (標的遺伝子) のためのモジュールは、それぞれ 0.604 (0.777) と 0.362 (0.445) である。それぞれの ncRNA (タンパク質) が 0.578 のモジュールスコア (0.598) に H. サピエンスリードにおけるモジュール性を検出した

(図 2 参照)。すべての生物のすべての可能な相互作用を用いたモジュールの計算は ncRNA (タンパク質) 分子に対して 0.720 (0.752) の値を与える。高い値では稀であるが、良いモジュールの値は、通常、0.3-0.7 の間にある。非常に低いモジュール性 (<0.3)、複雑なモジュール構造の欠如を示している。さらに、二部グラフのモジュールに基づく薬剤 - タンパク質複合体上の我々の分析は、新しい polypharmacological 特性を明らかにした。

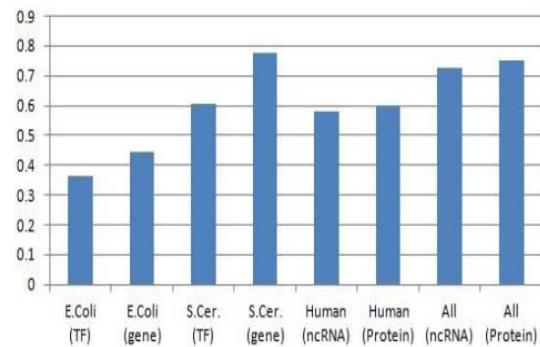


図 2. コミュニティのモジュール性のスコア

転写調節ネットワーク (TR ネットワーク) と、ncRNA タンパク質ネットワーク (ncRNA ネットワーク) 次数分布は、共通のスケールフリー構造を示す。しかし、本研究で我々は、ネットワークのモジュール性を計算することによって、新しいグローバルプロパティを発見した。その結果は、ネットワークが高いモジュール性のスコアを持っていることを示している。さらに重要なこととして、コミュニティのサイズの分布は、スケールフリー関数であることを示した。このプロパティは、異種スケールの相互接続されたモジュールによって形成された階層的な組織化されたシステムを強調している。図は、指定されたサイズのコミュニティを発見する確率が転写との ncRNA-相互作用の両方のための 2 に近い指数でべき乗則として、徐々に減衰することを示している。第二に、K-クリークサブグラフと重なって H. サピエンスの ncRNA 分子の識別コミュニティが潜在的に特定の機能に関連付けることができることを示している。真核生物において、スプライソソームは、イントロン配列を除去するための必須のスプライシング反応を行う。主要スプライソソームの ncRNA コンポーネントは U1, U2, U4, U5, U6 とされている。興味深いことに、これらの分子の k クリークコミュニティは明らかに我々の分析で識別される。理論的な面では、我々は唯一の二部ネットワーク情報を使用して投影されたネットワーク構造 (べき乗則次数分布と度指数) を予測する方程式を導き出すことができる。また、TF-遺伝子ネットワークを制御するノード数の最小値 (57%) は、

ヒトタンパク質の ncRNA ネットワーク (71%) からのものよりもわずかに低いように思われる。全6生物はの ncRNA ネットワークで考慮されている場合、我々は 68% を取得した。

この結果は、ncRNA の相互作用の複雑なモジュール構造、TR ネットワークからの類似性、細胞における可能な制御関係に焦点を当てることができた。最後に、相補的な分析が ncRNA の発現データを用いて提示され、ネットワーク上のマッピングを行うことによって行うことができる。現在収集された ncRNA データセットは、既存のすべての相互作用の小部分を表すという事実にもかかわらず、これらの知見は有望であり、非コード媒介相互作用の実験的分析および分類をしておくことが望まれる。

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 5 件)

(1) J.C. Nacher and T. Akutsu, Structural controllability of unidirectional bipartite networks, Scientific Reports, 査読有, **3**, 2013, 1647 (8 pp) doi:10.1038/srep01647 .

(2) J.C. Nacher, Community structure of non-coding RNA interaction network Journal of Integrative Bioinformatics, 査読有, **10**, 2013, 217-229 doi:10.2390/biecoll-jib-2013-217

(3) J.C. Nacher and T. Akutsu, Dominating scale-free networks with variable exponents: heterogeneous networks are difficult to control, New Journal of Physics, 査読有, **14**, 2012, 073005 (24 pp.) doi:10.1088/1367-2630/14/7/073005J

(4) J.C. Nacher and J.-M. Schwartz, Modularity in protein complex and drug interactions reveals new polypharmacological properties, PLoS One, 査読有, **7**(1), e30028, 2012, 1-13. doi:10.1371/journal.pone.0030028

(5) J.C. Nacher and T. Akutsu, On the degree distribution of projected networks mapped from bipartite networks, Physica A, 査読有, **390**, 2011, 4636-4651 doi:10.1016/j.physa.2011.06.073

[学会発表] (計 2 件)

(1) J.C. Nacher Community structure of non-coding RNA interaction network IB 2013 9th International Symposium on Integrative Bioinformatics, IPK-Gatersleben, Germany

(2) J.C. Nacher Modular organization of non-coding RNA-protein networks 8th International Workshop on Computational Systems Biology, WCSB 2011. In Proc. TIC Signal Processing series, **64**, pp. 141-147. Zurich, Switzerland

6. 研究組織

(1) 研究代表者

ホセ ナチエル (Jose Nacher)
東邦大学・理学部・准教授
研究者番号 : 60452984

(2) 研究分担者 なし

(3) 連携研究者 なし