COMMENTARY

酵母タンパク質相互作用ネットワークの構造的特徴と
その生物学的な特徴The Specificity of Topology of the Protein-Protein Interactions in
Yeast and Its Biological Characteristics鈴木泰博^{*a), b)}・中川 草^{b)}・長谷武志^{b)}・荻島創一^{b)}・田中 博^{b)}Yasuhiro SUZUKI, So NAKAGAWA, Takeshi HASE,
Sohichi OGISHIMA, and Hiroshi TANAKA

(Received December 3, 2004; Accepted March 15, 2005)

The specificity in topology of the protein network of yeast was investigated by analysis of the number of the protein-protein interactions. In order to examine the specificity of the protein network, the proteins was classified into three layers; the low, middle, and high with respect to the number of the protein-protein interactions. In the middle layer, proteins were mainly interconnected to each other, so proteins in the high layer were forced to be wired to those in the low layer. Hence, the results indicates that the proteins in the high layer connected between the low layer and the middle layer and integrated the network as a whole. Characterization of each layer was carried out biologically. It seems that in the past, studies have been focussed on the proteins in the high layer (highly connected proteins), but in the future, studies on the proteins in the middle layer will be more significant topologically and biologically.

1. 背 景

「どうぞ、お座りください。」途中から乗り込んできた紳士に席を譲ろうとすると、「大丈夫です。ありがとうございます。私は昭和7年生まれですが、同年代とお見受けしますが何年生まれでいらっしゃいますか?」「私も昭和7年です」。すると隣に立っていた紳士が「おや、私も昭和7年生まれです」¹⁾。われわれはよくこうした現象に出くわす。こうした現象についての社会学的な研究としてミルグラム (Milgram) の実験²⁾が知られている。

この実験の目的は、米国内に住むランダムに選ばれた2名の間の社会的な距離—何人知人を介して両者はつながるのか?—であった。そのため、マサチューセッツ州シャロン在住の神学部大学院生の妻と、ボストンに住む株式仲買人の2名を“ゴール”として選び、そして、カンザス州ウィチターとネブラスカ州オマハの住民160名をランダムに選択、彼らの知人を介した手紙が何人を経て“ゴール”

の2名に到着するかが調査された。その結果、160通のうち42通が平均5.5名を介して“ゴール”に到着した。この実験以降、社会的、人工的そして生物的なネットワークについて同様の調査が行われ、いずれの場合でもネットワーク内部の任意点間の最短経路はかなり短いことがわかってきた。巨大なネットワークでもわずかなステップ内にどのノードも結合されていることから、このようなネットワークの性質はスモールワールド (small world) と呼ばれるようになった³⁾。

その一方で数学の分野でもネットワークの研究は行われてきた。その先鞭となったのがエルデシュ (P. Erdos) とレーニイ (A. Renyi) が1960年代の初頭に提案した⁴⁾ランダムグラフ (random graph) である。ランダムグラフとは単純な確率モデルで、たとえるなら“ランダムにばらまかれた n 個のボタンを、目隠しをされた人がランダムに二つ選び、それらを1本の糸でつなぐ (糸は全部で m 本)”試行を繰り返すことにより生成されるネットワークである。この試行を繰り返して生成されるネットワークでは各々の“ボタン”につながっている“糸”の本数の度数分布が正規分布 (様) となり、ボタンの総数と糸の総数が同数になるとクラスターが出現する^{*1)}。

こうした背景をふまえ1990年代の後半から物性物理や

^{*a)} 名古屋大学大学院情報科学研究科複雑系科学専攻 (☎464-8601 名古屋市中種区不老町)

Department of Complex Systems Science, Graduate School of Information Science, Nagoya University (Furocho, Chikusa, Nagoya 464-8601, Japan)

^{b)} 東京医科歯科大学難治疾患研究所生命情報学 (☎113-8510 東京都文京区湯島 1-5-45)

Bioinformatics, Medical Research Institute, Tokyo Medical and Dental University (1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan)

^{*1)} ボタンの総数よりも糸の本数が少ない場合は、糸でつながれたボタンが散在するだけであるが、糸の総数がボタンの総数と等しくなると相転移的に糸でつながれたボタンの塊 (クラスター) が出現する。

複雑系の分野で急速に発展したのが複雑ネットワーク (complex networks) の現象論とその数理である^{3), 5)~7)}。研究の当初はランダムグラフをベースにしたさまざまな研究が行われた⁸⁾。その中で数理的な定義が曖昧であったスモールワールドを有するネットワークについて、

- ネットワーク内の2点間の最短パスがランダムグラフに比べて短い
- ネットワークがランダムグラフに比べて密に結合されている (ネットワークのトポロジーで“三角形”を形成している割合が高い)

の二つの性質を満たすネットワークとの定義が提唱された³⁾ (以下ではこれらの性質をスモールワールド性 (small worldness) と呼ぶことにする)。また、その一方でインターネットなどの大規模な複雑ネットワークの解析が行われた。その結果、ほとんどのネットワークはランダムグラフ (結合数の度数分布が正規分布) ではないことが明らかになった。それらネットワークで特徴的なことは、平均的な結合数の数百倍以上の結合数を有するノードがごく少数ではあるが存在することである。そのため、ランダムグラフのようにネットワークを特徴づける平均的な結合数の尺度 (スケール) が存在しないため、このようなネットワークはスケールフリーネットワーク (scale free network) と呼ばれる。スケールフリーネットワークとして知られているものには、例えば、食物連鎖、論文の共著者そしてタンパク質相互作用ネットワークなどがある。スケールフリーネットワークの発見はランダムグラフをベースにしてきたそれまでのネットワークの理論的研究が必ずしも現実のネットワークを反映したものではないことを示すこととなった⁵⁾。

1.1 計算生物学との関連

かかるネットワークの研究は、生物系、特にタンパク質相互作用ネットワークの解析に応用され、2000年をピークとして欧米でブームの様相を呈するが、数理的にはともかくとして、生物学的にはそれほど重要な指摘や発見はなされなかった。その後ブームは去り計算生物学的にはスケールフリー性などの概念は単なる常識の一つとなったが、実際の生物ネットワークではスケールフリー性は一般的な性質ではないという批判的な研究もある⁹⁾。その一方で複雑ネットワークの数理解析手法とその生物学への応用は、計算生物学 (バイオインフォマティクス (bioinformatics) やシステム生物学 (systems biology)) の中心的な研究課題となりつつある。例えば、Alonらは解析の対象となるネットワークと、当該ネットワークでの結合をランダム化した場合と比較することで特徴的なネットワークの部分構造、ネットワークモチーフ (network motif) を抽出する方法を提案している¹⁰⁾。ネットワークモチーフは当初、大腸菌の遺伝子ネットワークの記述のために提案されたが、その後さまざまな生物ネットワークの記述と解析に応用されている。

2. 出芽酵母のタンパク質相互作用ネットワークの解析

タンパク質相互作用ネットワークについて、従来から最もよく調べられている生物種は出芽酵母である。この出芽酵母のタンパク質相互作用ネットワークもスケールフリーネットワークであり、かつ、スモールワールド性を有する。複数のグループにより酵母2ハイブリッド法 (yeast two hybrid) を用いた大規模なスクリーニングが行われ^{11)~13)}、その数理的特徴^{6), 11), 14)}、タンパク質の機能推定¹³⁾などさまざまな研究が行われてきた。本論文ではタンパク質相互作用のデータとして Munich Information Center for Protein Sequences (MIPS) の Comprehensive Yeast Genome Database の出芽酵母のデータ (17 November, 2003) を用い、細胞内局在のデータとして Yeast GFP Fusion Localization Database を用い解析を行った結果について述べる¹⁵⁾。なお、酵母^{*2}ハイブリッド法の実験結果には相互作用の方向2に関する情報が含まれるが、その情報の生物学的な意味はほとんどないため⁴⁾、無向として扱い^{*3}また自らへの結合 (self-interaction) はすべて除外した。

2.1 ネットワークの全体的な特徴

まず、ネットワークの全体的な構造の特徴について調査した。各々のタンパク質の相互作用数とネットワークの構造との関係に着目し、当該ネットワークと相互作用をランダム化したネットワークと比較することにより、特定の相互作用数をもつタンパク質間に特徴的な構造があるかを調査した (Fig. 1)。相互作用のランダム化は、解析している当該ネットワークと同じタンパク質数、相互作用数、および結合の頻度分布を用い、ランダムなネットワークを1,000生成しその平均をとることにより行った。以下では

Fig. 1. Protein-protein interactions in the network. $C = I_{(i,j)}^{obs} / I_{(i,j)}^{mean}$

^{*2} bait タンパク質を A, prey タンパク質を B とし、実験の結果、両者が結合することが判明した場合にはその結果は A→B と表記される。

^{*3} 文献 14 はわれわれと同様の解析を行っているが相互作用の向きに関する情報を用いている。

Table 1. The Ratio of Interconnected Links within the Layer, Where the Number of Interconnected Links within the Layer is Divided by the Total Number of Links of the Layer^{a)}

Layer	Real network	Randomized network
Low	0.4779	0.4877 ± 0.0698
Middle	0.4921	0.3962 ± 0.0629
High	0.0114	0.0868 ± 0.0295

^{a)} In order to calculate randomized case, 1,000 null models that have the same connectivity as the original data of each layer was linked randomly.

このネットワークをランダム化ネットワークと呼ぶことにする。特定の相互作用数をもつタンパク質間に特徴的な構造があるかを調査するため、相互作用数が i と j のタンパク質の間で相互作用をしている数を、ランダム化する前と後で比較した。ここで、ランダム化する前の相互作用数 i と j のタンパク質間での相互作用の回数を $I_{(i,j)}^{\text{obs}}$ 、ランダム化した後を $I_{(i,j)}^{\text{mean}}$ と表し、両者の比較は $C = I_{(i,j)}^{\text{obs}} / I_{(i,j)}^{\text{mean}}$ を計測することにより行った。 $C > 1$ の場合は相互作用数 i, j のタンパク質間の相互作用が実際のネットワークではランダム化した場合よりも多い (enhanced) ことを示し、 $C < 1$ では少なく (suppressed), $C = 1$ では両者に違いはないことをそれぞれ示す。この解析の結果、ネットワークは階層的に構成されており、特に相互作用数が中程度 (7~40) のタンパク質間同士の相互作用は、ランダム化した場合に比べて多く ($1.0 < C < 1.8$) なっていることが示された。その一方で、この中程度に相互作用をしているタンパク質群とそれ以外 (相互作用数が多い、あるいは、少ないタンパク質) との相互作用はランダム化した場合に比べて少なくなっていることも示された ($0.8 < C < 1.0$)。そこで、以下では相互作用数が 7 以下のタンパク質群を低層 (low layer), 7 から 40 を中層 (middle), そして 40 以上を高層 (high) と呼ぶことにする¹⁵⁾。

Maslov らはかかるネットワークについて高層と低層の^{*4}相互作用がランダム化した場合に比べて多いことを重要な構造的特徴であると指摘しているが¹⁴⁾、このような特徴は中層内部での相互作用が多いため、高層のタンパク質が中層とあまり相互作用を行えず、低層と相互作用することを余儀なくされるために生じたものと考えられる¹⁵⁾。Table 1 は、各々の階層における層内部での相互作用数のランダム化ネットワークとの比較である。中層は他の階層に比べ層内部で相互作用する傾向が高いことが示されている。

以上の全体的特徴についてまとめると、出芽酵母のタンパク質相互作用ネットワークは相互作用数に応じ低層、中層、そして高層の 3 層構造となっており、特に中層では層内部の相互作用が多いために、高層と低層間の相互作用がランダム化した場合に比べて多くなる傾向にある。 Fig.

Fig. 2. A snapshot of the protein-protein interaction network related to SRP1 protein (centered in the figure). SRP1 is classified in the high layer.

2 は高層に属するタンパク質である SRP1 から派生する相互作用ネットワークを切り出したものであるが、SRP1 と相互作用するタンパク質は中層よりも低層のタンパク質が多く、また中層-低層間の相互作用は極めて少ない様子が示されている。

2.2 中層内の構造について

タンパク質相互作用ネットワークでは密に相互作用している中層が重要な意味をもつことが示されてきたが、中層の内部構造を調査した結果、中層は単に密な相互作用しているのではなく相互作用が疎密に分かれていることが明らかになった¹⁵⁾。ここで疎密とは、あるタンパク質と相互作用しているタンパク質同士が、さらにどの程度互いに相互作用しているかを計測し、その程度が高いものを“密な相互作用をしているタンパク質”、低いものを“疎な相互作用をしているタンパク質”として分類した。その結果、密な相互作用をしているタンパク質と、疎な相互作用のタンパク質はそれぞれ約 270 程度存在していた。そして、密な相互作用をしているタンパク質の相互作用は約 40% が層外部との相互作用であるのに対し、疎な相互作用をしているタンパク質の相互作用は 60% が層外部との相互作用であった。この結果から、疎な相互作用をしているタンパク質は層内部の密結合しているタンパク質群を束ねるとともに、層外部との“コネクタ”となっていることが示唆される。また、密に相互作用をしているタンパク質群^{*5}は、複合体を形成するなど機能単位となっている場合も多く見受けられた。

3. 中層の生物学的特徴

次に、出芽酵母の中層に属するタンパク質の生物学的な

^{*4} 文献 14 では低層、中層、そして高層の 3 層構造については言及していない。

^{*5} 最短経路 (shortest path length) を用いクラスター分析を行ったところ、中層のタンパク質は 14 のクラスターに分割された。

Table 2. The Number of Protein Complexes Forming Interactions of Each Layer^{a)}

Cellular localization	Layer	The number of proteins	The number of interactions forming protein complex	Rate
Cytoplasm	Low	266	60	0.23
	Middle	150	109	0.73
	High	4	2	0.5
Nucleus	Low	277	91	0.33
	Middle	142	160	1.13
	High	6	1	0.17
Golgi component	Low	18	5	0.28
	Middle	11	15	1.36
	High	0	—	—
Nucleolus	Low	22	17	0.77
	Middle	3	1	0.33
	High	0	—	—
Actin	Low	5	2	0.4
	Middle	13	22	1.69
	High	0	—	—
Cell periphery	Low	7	4	0.57
	Middle	17	9	0.53
	High	0	—	—
Yeast as a whole	Low	2,619	290	0.11
	Middle	540	530	0.98
	High	16	5	0.31

^{a)} There were the largest numbers of interactions in cytoplasm and nucleus. The proteins classified as the middle layer had the highest relevance for forming protein complexes.

特徴について調査を行った。その結果、全般的に中層に属するタンパク質は複合体形成に関与している割合が高いことが明らかになった¹⁵⁾。低層に属するタンパク質の総数は2,619、中層は540、高層は16であったが、各々の層において複合体形成の相互作用に関与しているタンパク質の割合は低層が11%、中層が98%、そして高層が31%と中層が高い割合を示した (Table 2)。

また細胞局在に関するデータ¹¹⁾を用い各局在について調べてみたが、同様に中層に属するタンパク質が複合体形成に関与している割合が高かった。

3.1 相互作用数とタンパク質複合体の大きさとの相関

しかしながら、個別に見ると核小体 (nucleolus) では、中層 (33%) よりも低層 (77%) のほうが複合体形成に関与している割合が大幅に高かった。生物学的には、核小体は小さな複合体が自己会合して構成されていることが知られている。そこで、タンパク質の相互作用数とそれらが構成する複合体の大きさとの関連について調査した (Fig. 3)。ここでは、複合体の大きさ (size) をタンパク質複合体を構成するタンパク質の総数と定義した。調査の結果、その両者には正の相関があることが明らかになった。つまり、相互作用数が少ないタンパク質は小さなタンパク質複合体を構成する傾向にあり、相互作用数が多くなるにつれて構成する複合体の大きさも大きくなる傾向にある。したがって、核小体の場合は小さな複合体が自己会合しているために、複合体形成に関与するタンパク質の相互作用数も少なかったと考えられる。以上の傾向より細胞表面 (cell periphery) に局在するタンパク質が構成する複合体は、小さ

Fig. 3. The correlation between the size of a protein complex and the average of maximal numbers of interactions of each protein in the complex.

いものと中程度のものが半分ずつ程度あることが推測される (Table 2)。

以上より、中層に属するタンパク質はその多くが複合体形成に関与していることが示された。また、一般にあるタンパク質が複合体を形成する場合、その相互作用数は形成する複合体の大きさをほぼ反映していることが明らかになった¹⁵⁾。

4. おわりに

出芽酵母のタンパク質相互作用ネットワークでは中層が構造的にも生物学的にも重要な意味をもつことが示されてきたが、この構造が他の生物種でも保存されているかを調べるため、線虫とショウジョウバエのタンパク質相互作用

ネットワークの解析を行い比較を行った。その結果、出芽酵母と同様にいずれの場合も中層内で相互作用数が多くなる構造が保存されていることが明らかになった^{*6, 15)}。よって、タンパク質相互作用ネットワークにおいて中層によって特徴づけられる階層構造は構造的、生物学的そして進化的にも重要な構造であることが示唆された。今後はかかる階層構造の構成原理やその生物学的な意味について考察を深めたい。

以上、複雑ネットワークの生物学への応用研究について述べてきたが、サイエンスとしてのみならず医療や製薬との観点からも生物をシステムとしてとらえる方向での研究が近い将来重要となることは明らかである。そのための“ツール”の一つとして、複雑ネットワークの手法は有望であると思われる。

謝 辞 本研究は文部科学省学術フロンティア（知能情報・同志社大学）から補助を受けた。

文 献

- 1) 朝日新聞朝刊, 2004 年 11 月 5 日 (2004).
- 2) S. Milgram, *Physiology Today*, **2**, 60, (1967).
- 3) D. J. Watts, “Small Worlds,” Princeton Univ. Press (1998).
- 4) A. Barabasi (青木 薫訳), “新ネットワーク思考”, NHK 出版 (2002).
- 5) Erdos and Rery, *Publ. Math. Debrecen*, **6**, 290 (1959).
- 6) I.Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki, *PNAS*, **98**(8), 4569 (2001).
- 7) H. Jeong, S. P. Mason, A.-L. Barabasi, and Z. N. Oltvai, *Nature*, **411**, 41 (2001).
- 8) A. Vazquez, A. Flammini, A. Maritan, and Al. Vespignani, *Nature Biotech.*, **21**, 697 (2003).
- 9) M. Arita, *PNAS*, **101**(6), 1543 (2004).
- 10) R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii and U. Alon, *Science*, **298**, 824 (2002).
- 11) W. Huh *et al.*, *Nature*, **425**, 686 (2003).
- 12) A. Tong *et al.*, *Science*, **294**, 2364 (2001).
- 13) P. Uetz *et al.*, *Nature*, **403**, 623 (2000).
- 14) S. Maslov and K. Sneppen, *Science*, **296**, 910 (2002).
- 15) Y. Suzuki *et al.* (submitted).

Keywords: Protein-protein interaction of yeast, Complex networks, Small worlds, Scale free

^{*6} しかし、ショウジョウバエでは高層がほとんどなくなり低層と中層の 2 層構造に近い構造へと変化していることが示された¹⁵⁾。