

3 疾患を理解するためのネットワーク的アプローチ

菊地 正隆

新潟大学脳研究所 遺伝子機能解析学分野

Network Approaches for Understanding of Diseases

Masataka KIKUCHI

Department of Molecular Genetics, Brain Research Institute,
Niigata University

要 旨

オミックス解析技術の進展に伴い、従来のように個々の遺伝子レベルから現象を明らかにするトップダウン型研究に加え、網羅的解析により得られたデータから関連する遺伝子を探索するというボトムアップ型の研究も多く行われるようになってきた。これにより我々は生命システムを遺伝子やタンパク質から成るネットワークとして捉えることができるようになり、そのネットワーク異常として疾患を再定義できるようになりつつある。ではこれまで報告されているような疾患関連遺伝子はネットワークレベルでどのような特徴があるのか、また疾患を克服するためにはどのような遺伝子をターゲットとして薬剤を設計するべきか、これまで得られた知見について述べていく。

キーワード：タンパク質間相互作用ネットワーク、ハブタンパク、モジュール、アルツハイマー病

はじめに

1990年代後半のマイクロアレイの開発から近年の次世代シーケンサーの目覚ましい発展に伴い、分子生物学研究はもはや大規模データの解析なしには成り立たない。大規模データの解析により、我々は約2万3,000のヒト全遺伝子から疾患などの表現型に関連する遺伝子群を探しだすことが可能となった。しかし、個々の遺伝子は単独で機能するわけではない。疾患の発症機序を含め、生命現象を本当に理解するためには、遺伝子群がどのように影響しあい表現型を決定するのかが知ることが必須である。そのためには遺伝子やタンパク質といった生体分子の間の関係性を明らかにし、ネットワークとして理解する必要がある。で

は、これまでにどのような手法で生体分子がネットワークとして表現され、解析が行われてきたのか。本稿では網羅的ネットワーク解析により新たに分かってきた遺伝子やタンパク質の特徴や、それを応用した疾患候補遺伝子の同定や薬剤標的分子の探索を行った研究を概説するとともに、我々が行ったアルツハイマー病におけるネットワーク解析を紹介する。

生体分子間の相互作用

これまで試みられてきた生体分子間のネットワーク表現として遺伝子共発現ネットワークとタンパク質間相互作用ネットワークを紹介する。

遺伝子共発現ネットワークはマイクロアレイデー

Reprint requests to: Masataka KIKUCHI
Department of Molecular Genetics, Brain
Research Institute, Niigata University,
1-757 Asahimachi-dori, Chuo-ku,
Niigata 951-8585, Japan.

別刷請求先：〒951-8585 新潟市中央区旭町通1-757
新潟大学脳研究所遺伝子機能解析学分野 菊地 正隆

タなどにより出力された個々の遺伝子発現量(シグナル強度)を元に発現パターンが似ている遺伝子同士をつなげたネットワークである。例えば時系列のマイクロアレイデータより構築した共発現ネットワークでは、相互作用する遺伝子はお互いに同時期に発現しており、いくつかのマスタレギュレーター遺伝子により調節されている可能性が考えられる。発現パターンの類似性についてはピアソンの相関係数を用いる手段が最もシンプルな方法であり、任意の2つの遺伝子の発現量の相関係数が一定以上であった場合にそれらをつなげていく。ピアソンの相関係数は線形な関係しか検出できないが、非線形な関係を検出する相互情報量を用いた方法としてARACNEアルゴリズムもよく用いられる方法の一つである¹⁾。遺伝子共発現ネットワークの解析ツールとしてWeighted gene co-expression network analysis(WGCNA)が知られており、発現データを入力することで共発現ネットワークを構築してくれる²⁾。

次にタンパク質間相互作用ネットワークはタンパク質同士の物理的な結合によるネットワークである。これらのデータはハイスループットなyeast two-hybrid法やアフィニティ精製と質量分析法を組みあわせるなどの実験的な方法により同定されたもので、BioGRIDやHPRD、DIPなど様々なデータベースに登録されており利用可能である。最近ではThe International Molecular Exchange(IMEx)コンソーシアム(<http://www.imexconsortium.org/>)が立ち上がり複数のデータベースを統合する動きもある。

生体分子ネットワークの構造的特徴

網羅的に同定された生体分子ネットワークについて、まずその大域的な構造が解析された。各生体分子がいくつかの生体分子と相互作用しているかを示す指標を結合次数という。統計物理学者であるAlbert-László Barabásiらは大腸菌の代謝ネットワークにおいて結合次数の分布がスケールフリー性に則ることを明らかにした³⁾。これは大半の分子は少数の相手とのみ相互作用するが、ごく少

数の分子は非常に多くの分子と相互作用するハブとして機能することを意味している。さらにいくつかの網羅的なネットワーク(ウェブサイト間のつながりをみたWWWネットワークや、フェイスブックやmixiなどで見られる人と人の間の知人関係をみたソーシャルネットワークなど)も同様にスケールフリー性を示していたのである。それまでは網羅的な情報が手に入りやすく、多くのネットワークはランダムネットワーク(結合次数分布が正規分布)だと考えられていたため、これらの結果は複雑ネットワーク科学という新たな学問分野を切り開く一端となるほどの大発見であった。その後、酵母(*Saccharomyces cerevisiae*)、ショウジョウバエ(*Drosophila melanogaster*)、ヒト(*Homo sapiens*)などの生物で網羅的 yeast two-hybrid法等によりタンパク質間相互作用が同定され、それらネットワークでもスケールフリー性が確認された。

ハブタンパクの生物学的特徴

ネットワークはノード(遺伝子共発現ネットワークにおける遺伝子、タンパク質間相互作用ネットワークにおけるタンパク質)とノード間のリンクにより構成される。スケールフリーネットワークの特徴はランダムにノードを削除してもネットワークの構造にさして影響はなく頑強であるが、多くのノードと結合するハブの削除には脆弱で大規模に構造が崩壊することである。つまり、多くのノードは少数のノードとしかつながっていないため、それらノードを一つ外したとしてもネットワーク構造に影響を与えない。しかし、ハブを選択的に攻撃した場合には多くのノードとの結合が失われ、ネットワーク構造が保たれなくなるということである。このように、ハブはスケールフリーネットワークにおいて重要な役割を占めていると考えられる。そこで生体分子ネットワークにおいてハブの生物学的特徴が研究され始めた。タンパク質間作用ネットワークを用いた研究でこれまで解明されたいくつかのハブタンパクの特徴を挙げる。

1. ハブタンパクはアミノ酸配列に影響を与えるような変異が入りにくく進化的拘束が強い。

酵母 (*Saccharomyces cerevisiae*) と線虫 (*Caenorhabditis elegans*) のオーソログについて、酵母タンパク質間作用ネットワークを用いた研究では、結合次数が1アミノ酸サイト当たりの置換数と負の相関をもつことが示された⁴⁾。

2. ハブタンパクは多くの組織で発現している。

ヒトのデータを用いた解析では、結合次数とそのタンパクが発現している組織数のあいだには正の相関がある⁵⁾。上に述べたようにハブタンパクは進化的拘束が強く、(少なくとも繁殖期が終わるまでは) 生命維持の中心的な役割を担っていることがわかる。同様に多くのハブタンパクは幅広い組織で発現するハウスキーピングである傾向があることから、細胞レベルで主要な機能を果たしていると考えられる。

3. ハブタンパクは致死性タンパクや疾患関連遺伝子の割合が高い。

酵母やヒトのデータを用いたところ (ヒトの致死性遺伝子はマウスで致死性だったオーソログ遺伝子より推定)、結合次数とともにそれらの割合が多くなることが報告されている^{6)–8)}。例えばガン抑制遺伝子 p53 は生体分子相互作用データベース BioGRID (ver3.2.106) では 548 の遺伝子/タンパク質と相互作用する。結合次数の平均値及び中央値はそれぞれ 15.6 と 4 であることから考えてもスーパーハブである。

ネットワークから考える創薬戦略

ハブタンパクは疾患に関連することを示したが、では薬剤標的遺伝子としてハブタンパクは有望な候補となり得るのだろうか? ハブタンパクは致死性遺伝子の割合も多いことから答えはノーである。ハブタンパクを狙ったのでは副作用が強すぎるのである。かといって、低結合次数のタンパクを狙ったのでは効果は期待できないだろう。実際にアメリカ食品医薬品局 (FDA) で認可または

試験中の薬剤がどのような結合次数をもつタンパクを多くターゲットにしているかを調べてみると、なんと高くも低くもない中程度の結合次数をもつ遺伝子/タンパク質が最も多かったのである⁸⁾。実はアルツハイマー病の原因遺伝子の一つである APP は BioGRID では 1,968 の遺伝子/タンパク質と相互作用する。これは BioGRID の中でも 2 番目に多い (ちなみに 1 番はユビキチン C で 9,721)。アルツハイマー病の根本治療薬の開発に向けては APP やその周辺の遺伝子をターゲットとしているが、重篤な副作用や認知機能の改善が診られないなど難航している。このようなネットワークの視点から、中程度の結合次数をもつ分子をターゲットにすると、もしかすると有力な候補遺伝子を絞り込むことができるかもしれない。

ネットワーク異常としての疾患

疾患は遺伝子転写異常や有害な変異、コピー数異常など様々な要因により引き起こされる。このような個々の遺伝子に対する異常が周辺の遺伝子にどのような影響を与えるのかを分析する際にもネットワーク解析は有用である。

乳がんにおいて予後経過が良好だった患者群と予後が悪化した患者群それぞれの発現データを用い、原因遺伝子の一つである BRCA1 とその周辺のタンパク質間相互作用するタンパク質とのあいだの共発現関係に注目したところ、予後良好群では共発現していたが、予後悪化群では協調関係が失われていた⁹⁾。この事実は遺伝子発現データやタンパク質間相互作用データを組み合わせることで未知の予後マーカー探索の可能性を示唆する。

疾患の中には p53 のように 1 遺伝子で観察された複数の異なる変異が数十種類のガンの原因になる場合もあれば、QT 延長症候群のように 12 個の異なる遺伝子に入った変異がそれぞれ同一の疾患を招く場合もある。

このような疾患の多様性を説明するために、Haiyuan Yu らのグループはタンパク質同士の結合がドメインを介して実現することに注目した。彼らはタンパク質同士の共結晶構造データよりタ

ンパク質ドメイン間の相互作用を推定し、これまでに報告された疾患に関与するアミノ酸変異がどのような場所に多いかを観察した¹⁰⁾。その結果、それら変異はタンパク結合ドメインの外に入る割合よりも、ドメイン内に入る割合の方が高いことを示した。さらに1遺伝子で観察された各々異なる疾患を引き起こす変異の局在を調べたところ、興味深いことに同一のドメインに局在するよりも、それぞれ異なるドメインに局在する割合の方が高かった。これはあるドメインAに入った変異はその場所で結合していたタンパクとの結合を阻害し、また別のドメインBに入った変異はドメインAで結合していたタンパクとは別のタンパクとの結合を阻害することを示唆する。実際にWiskott-Aldrich症候群、X連鎖性血小板減少症、X連鎖性好中球減少症を引き起こす変異をもつWASPと呼ばれる遺伝子はVASPと結合するWH1ドメインとCDC42と結合するPBDドメインをもつが、X連鎖性血小板減少症を引き起こす変異はWH1ドメインに、X連鎖性好中球減少症を引き起こす変異はPBDドメインに入っており、これらの変異がそれぞれのタンパクとの結合が阻害されていたことが実験的に示された。これにより1つの遺伝子に対する変異であっても異なるパートナーとの結合が阻害されるため多様な疾患を呈す

ることを説明できる。ネットワーク内の一部の相互作用の欠損などにより障害を受けることを“Edgetic perturbation”と呼び¹¹⁾、今後ネットワークとして疾患を理解するためのモデルとして注目を集めつつある。

アルツハイマー病の進行に伴う ネットワークの崩壊

著者らはヒトのタンパク質間相互作用ネットワークとアルツハイマー病の遺伝子発現データを組み合わせることで異なる脳部位（嗅内皮質、海馬、上前頭回）において、アルツハイマー病の進行ステージ（Braak stage）ごとに発現するタンパク質間のネットワークを同定し、それらがどのように変動するかを調べた¹²⁾。解析の結果、アルツハイマー病初期でダメージを受ける嗅内皮質ではその他の脳部位と比べ、タンパク質間の相互作用が著しく失われており、ネットワークが崩壊していた。さらにネットワークをモジュールと呼ばれる部分ネットワークに分割することで、ステージの進行とともに多くの相互作用が失われるモジュールを同定した。その結果、ヒストンアセチルトランスフェラーゼ（HAT）に関連するモジュールはBraak stageの進行とともにモジュール内の相

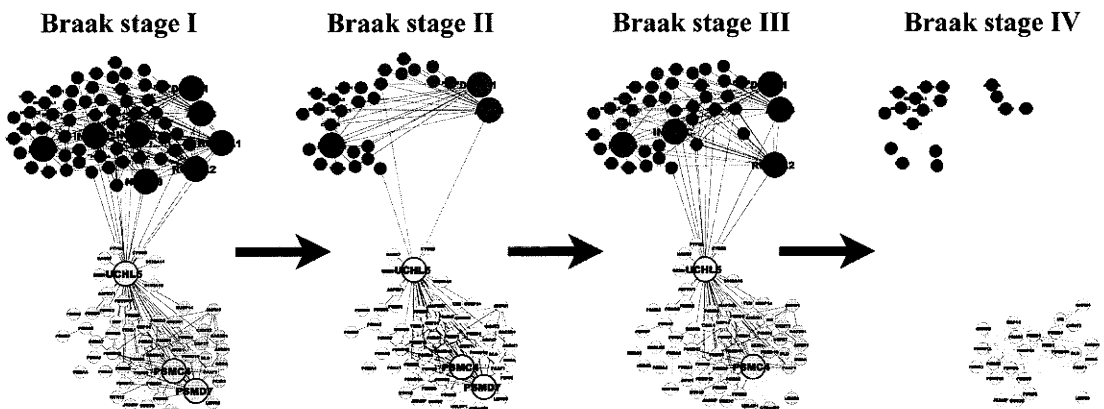


図1 Braak stageの進行に伴うHATモジュール（上）とプロテアソームモジュール（下）の崩壊。

相互作用が最も多く消失していた。さらに HAT モジュールを構成する多くのタンパクはプロテアソームに関連するモジュール内の脱ユビキチン化酵素 UCHL5 と相互作用しているが、Braak stage IV において UCHL5 が発現しなくなることで HAT-プロテアソーム間の連携が完全に絶たれることが観察された (図 1)。このことから UCHL5 はネットワーク崩壊の責任遺伝子の一つではないかと考えられる。

おわりに

疾患の発症機序を理解し、効率的に薬剤標的分子を探るために、これまでの分子生物学的な手法に加え、生体分子間の関係をネットワークとして記述することは有効な手段の一つである。先行研究では概ね疾患群を一括りにし、静的な事象を扱っている。しかし、多くの疾患は連続的に発症し、時系列の進行に伴い細胞内のネットワークもダイナミックに変動しているはずである。著者らが研究しているアルツハイマー病の最大のリスクは老化と言われている。今後は、遺伝子発現データやネットワーク解析を通じ、老化が時間の流れとともにシステムとしてどのように変化しているのか明らかにし、どのような異常がアルツハイマー病という状態を形成するのかを解明していきたい。

引用文献

- 1) Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R and Califano A: ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 7 Suppl 1, S7, 2006.
- 2) Langfelder P and Horvath S: WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559, 2008.
- 3) Jeong H, Tombor B, Albert R, Oltvai ZN and Barabási AL: The large - scale organization of metabolic networks. *Nature* 407: 651 - 654, 2000.
- 4) Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C and Feldman MW: Evolutionary rate in the protein interaction network. *Science* 296: 750 - 752, 2002.
- 5) Bossi A and Lehner B: Tissue specificity and the human protein interaction network. *Mol Syst Biol* 5: 260, 2009.
- 6) Jeong H, Mason SP, Barabási AL and Oltvai ZN: Lethality and centrality in protein networks. *Nature* 411: 41 - 42, 2001.
- 7) Goh KI, Cusick ME, Valle D, Childs B, Vidal M and Barabási AL: The human disease network. *Proc Natl Acad Sci USA* 104: 8685 - 8690, 2007.
- 8) Hase T, Tanaka H, Suzuki Y, Nakagawa S and Kitano H: Structure of protein interaction networks and their implications on drug design. *PLoS Comput Biol* 5, e1000550, 2009.
- 9) Taylor IW, Linding R, Warde - Farley D, Liu Y, Pesquita C, Faria D, Bull S, Pawson T, Morris Q and Wrana JL: Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol* 27: 199 - 204, 2009.
- 10) Wang X, Wei X, Thijssen B, Das J, Lipkin SM and Yu H: Three - dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat Biotechnol* 30: 159 - 164, 2012.
- 11) Zhong Q, Simonis N, Li QR, Charlotiaux B, Heuze F, Klitgord N, Tam S, Yu H, Venkatesan K, Mou D, Swearingen V, Yildirim MA, Yan H, Dricot A, Szeto D, Lin C, Hao T, Fan C, Milstein S, Dupuy D, Bresseur R, Hill DE, Cusick ME and Vidal M: Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* 5: 321, 2009.
- 12) Kikuchi M, Ogishima S, Miyamoto T, Miyashita A, Kuwano R, Nakaya J and Tanaka H: Identification of unstable network modules reveals disease modules associated with the progression of Alzheimer's disease. *PLoS One*, in press, 2013.