

合成生物学によるモノづくり微生物のデザインに向けて

石井 純¹・荒木 通啓¹・中津井雅彦²・崎濱 由梨³・
 柘植 陽太¹・蓮沼 誠久¹・近藤 昭彦^{2*}

はじめに

近年のバイオテクノロジーの進展により、微生物を用いたモノづくりに注目が集まっている。たとえば、微生物の代謝経路や酵素反応を利用することできわめて多段階の変換反応や複雑な構造をもつ化合物変換を選択的、かつより温和な条件で行うことが可能であり、乳酸やエタノール、1,3-プロパンジオールなど、さまざまな化学品原料の生産に利用されている。こうした微生物を用いた化学品原料の生産では、CO₂削減や持続可能な社会の構築に向けたCSR (corporate social responsibility) の観点から、再生可能なバイオマスを原料とする研究開発を中・長期経営計画にあげて取り組む企業が増えてきている。

自然界に存在する微生物の代謝経路を構成する酵素やタンパク質をコードする遺伝子をクローニングして、大腸菌や酵母・コリネ型細菌などの遺伝子組換えや培養が容易な宿主に導入することで化学品原料などのさまざまな化合物を生産する微生物を作出する遺伝子工学的もしくは代謝工学的なアプローチは20年以上前から行われており、目的の化合物やタンパク質を高生産する微生物が多数開発されてきた。こうしたアプローチは主にボトムアップ的な要素還元論的手法¹⁾をもとに発展してきており、実績も豊富で非常に有効な方法である。しかし、還元論的手法では、複雑な対象も要素の集まりからできていると考えて、対象を単一の要素に分解して個別に理解することで全体を再現・改善しようとするのが一般的であり、対象を構成する要素が線形な集まりとして表現できることを仮定しているため、これらの要素間に動的で非線形な関係が存在する場合には問題を扱うことができない²⁾。つまり、複雑な細胞現象をタンパク質などのパーツに分解して考えるため、微視的な現象にとらわれてシステム全体を捉えることが困難な場合が多い¹⁾。そのため、近年ではシステム全体を考慮したトップダウン的なモデル構成論的手法¹⁾が提唱され、「つくることによって理解する」「合成生物学」として生命科学研究全般に新たなパラダイムシフトを起こしつつある。

生命をより全体論的に理解したり、生命のパーツを組

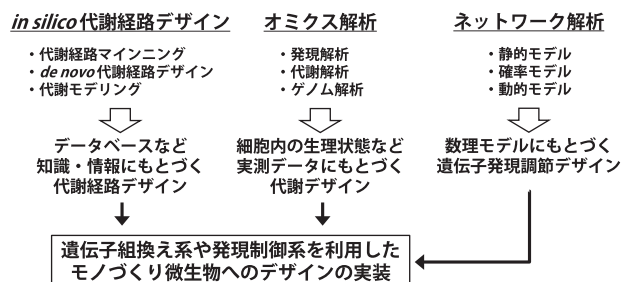


図1. 合成生物学によるモノづくり微生物の創出スキーム

み合わせて生命現象を再現しようとする合成生物学において、とりわけ新しい生命機能やシステムをデザインして組み立てて新たな価値を生み出す工学的な分野を切り出して、我々は“合成生物学”と呼んでいる³⁾。たとえば、合成生物学では微生物の代謝機能や遺伝子発現を合理的にデザインすることできわめてハイスペックなモノづくり微生物を創出することを目指している。こうしたトップダウン的なモデル構成を実現するためには、代謝経路を*in silico*においてデザインするツールや細胞内代謝・遺伝子発現の状態を正しく把握するための分析技術に加え、デザインしたシステムを実装するための遺伝子組換え系や発現制御系、さらにはネットワーク解析技術など、さまざまな技術開発が必要になる(図1)。

本稿では、合成生物学における技術について紹介し、モノづくり微生物をデザインするための展望を述べる。

*in silico*代謝経路デザイン

合成生物学のコンテキストでモノづくり微生物の代謝機能を合理的にデザインしていくためには、デザイン(作業仮説)の根拠となる知識・情報ベースとそれらを利用した代謝経路マイニングが必要となる。一般に、代謝関連情報は代謝物と酵素反応・遺伝子情報により構成されており、これまでの知識・情報を集積したデータベースとしてKEGG⁴⁾、BRENDA⁵⁾、BioCyc⁶⁾などが提供されている。特にKEGGデータベースは、独自のPATHWAYデータベースとして各種データを統合・視覚化した情報基盤を提供しており、代謝パスウェイ解析において標準的に利用されている。こうしたデータベースはウェブイ

* 著者紹介 ¹ 神戸大学自然科学系先端融合研究環重点研究部 E-mail: akondo@kobe-u.ac.jp

² 神戸大学大学院工学研究科応用化学専攻 ³ 高機能遺伝子デザイン技術研究組合

ンターフェイスも同時に提供されており、属人によるボトムアップ的な代謝デザインの一助になる一方で、集積されたデータを活用したデータマイニングにより、トップダウン的な代謝経路デザインにも有効に活用することができる。

代謝関連データベースを基にした代謝経路デザインでは、すでにさまざまなデザインツールが開発されているが、そのアプローチの違いで大きく分けて2つに分類できる。1つは、代謝関連データベース中の酵素反応・代謝物情報のみを利用して、目的の物質生産に必要な代謝パスウェイをデザインしていく方法である。この方法では、目的とする代謝経路に必要とされる異種宿主由来の酵素遺伝子群を特定し、モデル微生物にて代謝パスウェイを再構築していくことが主たる目的となる。近年のデータ蓄積の増大に伴い、FMM⁷⁾、MRSD⁸⁾、DESHARKY⁹⁾といった新たなアルゴリズムによるデザインツールも次々と開発されてきており、それぞれ利用しているデータベース、データの質・量、代謝パスウェイ再構築の方法などが違う。一方、これらのデザインツールは、既存データベースの情報にのみ依存するため、まったく新しい代謝パスウェイの予測というのは困難である。

このため、新規代謝経路をデザインする方法として、別のアプローチによるデザインツールも開発されている。BNICE¹⁰⁾では、ゲノムの由来を特定せずに酵素反応の特徴を定義することで、新たな酵素反応予測を行うアルゴリズムが開発されている。同様のアプローチで、KEGG特有の酵素反応情報を利用した代謝パスウェイ予測ツールであるPathPred¹¹⁾、酵素反応比較に特化したEC-BLAST¹²⁾なども開発されている。一方で、こうしたアプローチでは計算負荷が非常に高いため、我々は網羅的な酵素反応データをより効率良く処理し、代謝経路デザインを行うツールを開発し、グラフィカルに分かりやすく使いやすいウェブツール (M-path) も同時に開発している (図2)¹³⁾。こうしたアプローチは、前者を含めたモデル微生物における *de novo* 代謝経路のデザインを行う上で不可欠な技術となりつつある。

また、上述の代謝経路デザインとは別に、代謝モデルを定量的に扱うことで、より精度の高い代謝モデリングを行う方法も開発されてきている。Palssonらは、既存の代謝パスウェイ情報をもとに代謝モデル (Genome Scale Model) を構築し、定常状態における代謝フラックスの解析基盤を開発することで、目的物質の生産性向上のための代謝モデリングを行っている¹⁴⁾。また、Genome Scale Modelをもとに、遺伝子のノックアウト

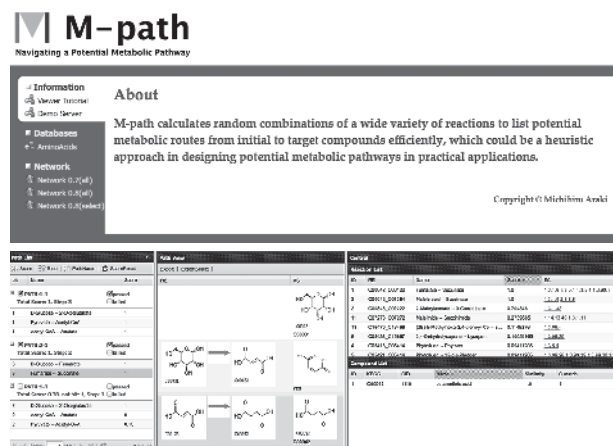


図2. M-path (代謝経路デザインのためのウェブツール) <http://bp.scitec.kobe-u.ac.jp/m-path/aa/>

モデルを生成するためのツールとして、OptKnock¹⁵⁾といった遺伝子設計のためのツールも開発されており、こうしたツールは代謝経路デザインにおいても今後さらに活用が進んでいくものと考えられる。

オミクス解析

モノづくり微生物における代謝の合理的なデザインのためには、データベースや代謝モデルを利用した *in silico* の予測手法とともに、細胞内の実際の生理状態を把握した上で代謝フラックスの律速点やカギとなる遺伝子制御ポイントを推定するアプローチも重要となる。こうした実測データを基にしたデザインのためには、トランスクリプトームやメタボローム解析など、細胞内の遺伝子発現や代謝産物を網羅的に分析する手法が有効である^{16,17)}。マイクロアレイによる発現解析はさまざまな宿主において利用できるようになっており、比較的容易に遺伝子発現量の変動を把握することが可能である。また、質量分析装置を用いたメタボローム解析についてもさまざまな手法が開発されており、細胞内代謝中間体の蓄積量や変換量 (代謝フラックス) を測定することができる。さらに最近では、次世代シーケンサー (NGS) の台頭により、微生物の全ゲノム配列を個人の研究者レベルで解読できるようになった。NGSを活用することで、これまで一般的に使われていた大腸菌や出芽酵母、コリネ型細菌など微生物以外でもゲノムを解読できることから、モノづくりの宿主として利用できる微生物の幅は格段に広がっている。また、進化工学的手法により取得した高生産株について、リファレンス株との比較で塩基配列の相違点をゲノムレベルで抽出することも可能となった。こうした各種のオミクス解析手法は、実測データを

基にしていることから、合理的なデザインを決定するための有効かつ実践的な手段として活用されている。

遺伝子組換え・発現制御系

*in silico*やオミクス解析により設計された代謝デザインを宿主微生物に実装するためには、遺伝子組換えの技術が必須である。こうした代謝デザインでは複数箇所の遺伝子の発現や破壊が必要になることがほとんどである。大腸菌などのバクテリアではオペロンを組ませることで1つのプロモーターで複数の遺伝子を発現できるが、酵母などの真核生物は原則的に1対1対応で遺伝子に対してプロモーターを準備しなければならない。こうしたデザインを実装するために真核生物用に複数の遺伝子発現を可能にする遺伝子組換え手法¹⁸⁾やベクター開発¹⁹⁾が進んでいる。また、遺伝子破壊については部位特異的組換え酵素 (FlpやCre) を用いた手法²⁰⁾が一般的であるが、近年ではCRISPR-Cas9などのゲノム編集技術を用いた複数遺伝子の同時破壊の技術も開発されている²¹⁾。さらに、単純な遺伝子の強発現や破壊だけではなく、siRNA²²⁾や遺伝子スイッチ²³⁾など遺伝子発現を調節する制御系も利用され始めており、より複雑かつ高度な代謝デザインも実装することが可能となりつつある。

遺伝子ネットワーク解析

各遺伝子の発現には、転写や翻訳・翻訳後修飾、代謝ネットワークなどを隠的に内包する、動的かつ非線形の相互作用が存在する。これらの相互作用の集合である遺伝子ネットワークを数理モデルとして記述することにより、シミュレーションを通して細胞の挙動を解析し、相互作用に対して改変を加えた場合の挙動を予測することができる。また、感度解析を行うことで、相互作用のどの部分を改変すればモノづくりにおける生産性を効率的に向上できるかの予測につなげることもできる。

遺伝子ネットワークを記述するための数理モデルは、主に、静的モデル、確率モデル、動的モデルの三種類に分類される。ブーリアンネットワーク²⁴⁾は静的モデルの代表であり、網羅的に測定された1遺伝子破壊株から²⁵⁾、または時系列の発現プロファイルから²⁶⁾、遺伝子間の関係性を論理関数によって記述する。確率モデルとしては、遺伝子間の制御関係を条件付確率によって記述するベイジアンネットワーク^{27,28)}があげられる。ブーリアンネットワーク、ベイジアンネットワークともに、時系列データを扱うことが可能であるが^{26,29)}、より自然に、かつ詳細に動的ネットワークの挙動を解析するためには、各遺伝子の発現量の変化量を直接記述する微分方程式モ

デル^{30,31)}が適している。一般化された非線形の微分方程式モデルとしては質量作用則を一般化した一般質量作用則 (GMA)³²⁾や、GMAをもとにした概念モデルであるS-system³³⁻³⁵⁾などがあげられる。

遺伝子ネットワークの解析は、一般的に、遺伝子間の関係性を記述するための一般化された数理モデルを用いし、破壊株や、時系列でサンプリングされた発現プロファイルなどの実験値をもとに、数理モデルの内部パラメータを推定する。このとき、シミュレーション値と実験値との間の誤差関数を最小化するパラメータ値を求めるのが一般的であるが、特に微分方程式モデルの自由度の高さと得られる実験値の量との間に大幅な乖離がある場合は、一意にネットワークを推定することは困難であり、ネットワークの規模に伴って計算負荷も指数関数的に増大する。一方で、遺伝子ネットワークに関する既知の知見や、遺伝子ネットワークの構造の類型化に関する研究³⁶⁻³⁸⁾、前述の静的・確率的な手法による予測、などにより、動的パラメータ値の有無にかかわらず、ネットワークの構造に関する情報が得られる場合がある。我々は、これらネットワーク構造に関する既知情報を活用することにより、フルスクラッチからネットワーク構造と動的パラメータを同時に推定する従来の手法よりも効率的に遺伝子ネットワークの推定を行う手法を開発している。また、これらの遺伝子ネットワーク解析手法を応用し、酵母の好気・嫌気条件における遺伝子ネットワークの変化についての解析を行っているところである。

近年の実験技術の飛躍的な発展による測定データ量の増大と、シミュレーションやネットワーク推定技術の進展、さらには計算機パワーの増大も相まって、動的かつ非線形の遺伝子ネットワーク解析が現実味を帯びてきた。これらの手法は、遺伝子発現の調節やデザインにおいて、今後活用が進むものと考えられる。

おわりに

本稿で紹介した各種のデザインツールは膨大なデータベースの情報や生体内のオミクス解析の網羅的な実測データから客観的な判断によって合理的な代謝経路の設計を提案してくれる。こうしたツールはデータの蓄積により信頼性が高まるため、ウェットとドライの研究者が強力に連携して融合研究を進めることが重要であり、合成生物学においてトップダウン的なモデル構成を実現するために解決しなければならない課題である。遺伝子組換え系や発現制御系の発展に伴って提示された代謝経路デザインを実装することは容易になりつつあり、モデルやアルゴリズムの信頼性が高まることで、高性能なモ

ノづくり微生物を短期間かつ容易に創出できるようになるであろう。バイオインフォマティクスとバイオロジーの両方において優秀な研究者が多数存在する我が国において、合成生物学におけるモノづくり微生物開発の機運が高まることを願っている。

謝 辞

本研究は、経済産業省委託事業「革新的バイオマテリアル実現のための高機能化ゲノムデザイン技術開発」の一部として実施されました。

文 献

- 1) 小松崎民樹：生体の科学, **65**, 444 (2014).
- 2) 北 英輔：情報処理学会論文誌 数理モデル化と応用, **47**, i (2006).
- 3) Kondo, A. *et al.*: *J. Biotechnol.*, **163**, 204 (2013).
- 4) Kanehisa, M. *et al.*: *Nucleic Acids Res.*, **36**, D480 (2008).
- 5) Schomburg, I. *et al.*: *Nucleic Acids Res.*, **41**, D764 (2013).
- 6) Karp, P. D. *et al.*: *Nucleic Acids Res.*, **33**, 6083 (2005).
- 7) Chou, C. H. *et al.*: *Nucleic Acids Res.*, **37**, W129 (2009).
- 8) Xia, D. *et al.*: *Bioinformatics*, **27**, 1581 (2011).
- 9) Rodrigo, G. *et al.*: *Bioinformatics*, **24**, 2554 (2008).
- 10) Hatzimanikatis, V. *et al.*: *Bioinformatics*, **21**, 1603 (2005).
- 11) Moriya, Y. *et al.*: *Nucleic Acids Res.*, **38**, W138 (2010).
- 12) Rahman, S. A. *et al.*: *Nat. Methods*, **11**, 171 (2014).
- 13) Araki, M. *et al.*: *Bioinformatics*, **31**, 905 (2015).
- 14) Bordbar, A. *et al.*: *Nat. Rev. Genet.*, **15**, 107 (2014).
- 15) Burgard, A. P. *et al.*: *Biotechnol. Bioeng.*, **84**, 647 (2003).
- 16) Hasunuma, T. *et al.*: *Appl. Microbiol. Biotechnol.*, **90**, 997 (2011).
- 17) Hasunuma, T. *et al.*: *Microb. Cell Fact.*, **10**, 2 (2011).
- 18) Chang, J. J. *et al.*: *Biotechnol. Biofuels*, **5**, 53 (2012).
- 19) Ishii, J. *et al.*: *FEMS Yeast Res.*, **14**, 399 (2014).
- 20) 森 浩禎：生物学, **90**, 293 (2012).
- 21) Jakočiūnas, T. *et al.*: *Metab. Eng.*, **28**, 213 (2015).
- 22) Na, D. *et al.*: *Nat. Biotechnol.*, **31**, 170 (2013).
- 23) Tominaga, M. *et al.*: *PLoS One*, **10**, e0120243 (2015).
- 24) Kauffman, S. *et al.*: *Nature*, **224**, 177 (1969).
- 25) Akutsu, T. *et al.*: *Proc. 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, **98**, 695 (1998).
- 26) Liang, S. *et al.*: *Pac. Symp. Biocomput.*, 18 (1998).
- 27) Pearl, J.: *Proceedings of the 7th Conference of the Cognitive Science Society*, 329 (1985).
- 28) Friedman, N. *et al.*: *Proceedings of the Fourteenth conference on Uncertainty in Artificial Intelligence*, 139 (1998).
- 29) Murphy, K. and Mian, S.: Modelling gene expression data using dynamic Bayesian networks. *Technical report, Computer Science Division, University of California, Berkeley, CA* (1999).
- 30) Chen, T. *et al.*: *Pac. Symp. Biocomput.*, **4**, 29 (1999).
- 31) De Hoon, M. *et al.*: *Pac. Symp. Biocomput.*, 17 (2002).
- 32) Kacser, H. and Burns, J. A.: *Symp. Soc. Exp. Biol.*, **27**, 65 (1973).
- 33) Savageau M. A.: *Biochemical System Analysis: A Study of Function and Design in Molecular Biology*, Addison-Wesley, (1976).
- 34) Maki, Y. *et al.*: *J. Bioinform. Comput. Biol.*, **2**, 533 (2004).
- 35) Nakatsui, M. *et al.*: *Math. Biosci.*, **215**, 105 (2008).
- 36) Herrgård, M. J. *et al.*: *Genome Res.*, **13**, 2423 (2003).
- 37) Shen-Orr, S. S. *et al.*: *Nat. Genet.*, **31**, 64 (2002).
- 38) Lee, T. I. *et al.*: *Science*, **298**, 799 (2002).