



バイオインフォマティクスによる バイオものづくり

情報科学が加速するバイオものづくりの新展開

石谷孔司 Koji ISHIYA

「バイオものづくり」は、産業構造や社会構造のパラダイムシフトを引き起こす原動力の1つとして期待されている。その中で、生物情報科学（バイオインフォマティクス）の果たす役割は大きくなりつつある。本稿では、バイオものづくりでのバイオインフォマティクスの利用についてその背景を含めて解説する。

バイオものづくりとは？

経済協力開発機構（OECD）によるレポート¹⁾によれば、2030年にはバイオエコノミー分野の経済規模は、約1.6兆ドルにも達するとされている。これを受け、日本・米国・欧州等をはじめとし、各国で持続可能な社会の実現に向けた関連政策が策定されている。その中でも「バイオものづくり」はCO₂排出量の削減と経済成長や国際競争力強化を両立させる働きを期待されている（図1）。本稿では、今般注目の集まる「バイオものづくり」とは何か、その中でバイオインフォマティクスがどういった役割を期待されているのかをそれぞれ概説したい。

昨今では、第四次産業革命とも呼ばれるビッグデータ・AI等とロボティクス・IoTの融合が産業・社会構造に大きな変革をもたらしている。同様に、国際規模での環境・エネルギー・材料・食糧分野の問題が深刻化する中で、バイオテクノロジーと情報科学の融合は「第五次産業革命」²⁾と呼ぶに値する、産業構造や社会構造のパラダイムシフトを引き起こす原動力の1つとして期待されている。生物によるものづくりは、我々の身近に存在し、長らくその利用は続いてきた。例え

いしや・こうじ

国立研究開発法人産業技術総合研究所生命工学領域 主任研究員

〔経歴〕2018年東京大学大学院理学系研究科博士課程修了、博士（理学）。同年産業技術総合研究所に入所し、バイオものづくり、生物資源探索、ゲノム進化等のバイオインフォマティクス研究に取り組んでいる。〔専門〕バイオインフォマティクス、ゲノム進化学。

E-mail: koji.ishiya@aist.go.jp



図1 バイオものづくりによる成長戦略イメージ

ば、日本においても古来より微生物による発酵・醸造の利用は継続的に行われてきた。その長い歴史の産物として醤油・味噌・酒・漬物等といった食品類は、我々日本人にとっても非常に身近なものとなっている。こうした微生物（バイオ）の活用は、昔から食品の風味や栄養価を向上させるために重要な役割を果たしてきた。現在では、バイオファウンダリをはじめとし、微生物を機能性物質の生産プラットフォームや化学合成プロセスの代替プラットフォームとして捉えた研究開発や設備投資が積極的に進められており、その役割は大きな変革期を迎えている。

微生物の育種や改変は、従来まで専門家が過去の経験や知識に基づき独自のプロトコルや手法を元に開発を行い、その過程には職人技とも言える暗黙知の要素を多分に含んでいた。そのため、目的物質を生産する育種開発は高度な専門知識や技術、さらには最適な条件等を見つけ出すための時間も多く必要とされた。また、技術や経験を他者に正確に伝えることも困難であった。特に、将来的な産業界への橋渡しを考えた場合、バイオものづくりの合理性・網羅性・効率性の3点は極めて重要な要素となる。これらを満たすための

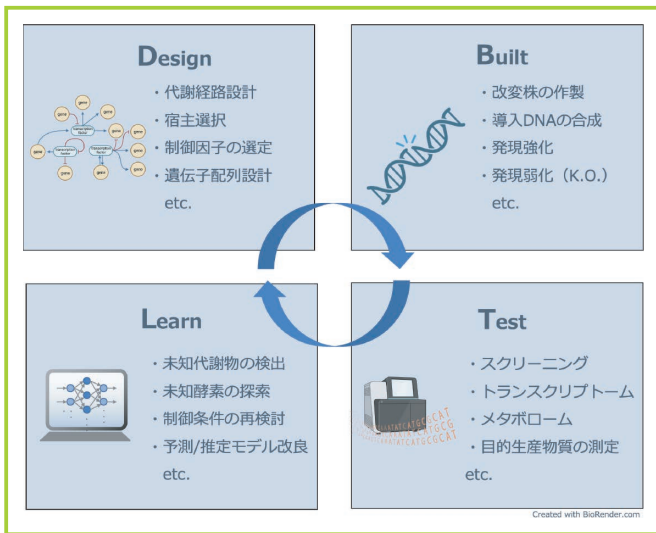


図2 DBTL サイクルにおける育種開発プロセス

アプローチの1つとして、バイオインフォマティクスが注目されている。バイオインフォマティクスの利用によって、育種開発における事前の予測や網羅的かつ合理的に原因や関連因子を絞り込んでいくことができるようになり、飛躍的にその開発効率を上げている。

最近では、高出力型のDNAシーケンサーや自動化された実験前処理装置等に代表される大規模に生体データを生み出すための計測機器や関連機器が多数登場しており、生物学分野においても大きな変革期を迎えている。バイオインフォマティクスでは、大量に生み出された生体ビッグデータを効率的に処理することももちろんのこと、巨大なデータから生命現象を統合的に理解し、改良や改善に向けて人間の経験や知識によらない合理的な解決策を提示できる点に強みを持っている。そのため、ゲノムやタンパクといった生物の構成要素を工学的に理解し、設計する合成生物学³⁾(Synthetic Biology)との組み合わせによって、バイオインフォマティクスは、これまでにないデータ駆動型の合理的な生物生産系の構築に積極的に利用されている。合成生物学を主軸としたバイオものづくりにおいては、品質管理など業務管理の継続的な改善方法であるPDCAサイクル(Plan→Do→Check→Action)のように、DBTLサイクル(Design→Build→Test→Learn)を効果的に回すことで迅速かつ効率的に育種開発を行うことができる(図2)。日本国内では、このような育種開発技術を適用し、物質生産能を最大限に高めた微生物(細胞)であるスマートセル⁴⁾の開発が積極的に進められている。

バイオものづくりにおけるバイオインフォマティクス

DBTLサイクルの中でも特にDesign(設計)とLearn

(学習)におけるバイオインフォマティクスの役割は大きい。先述した高出力型のDNAシーケンサーは、自然環境中の未活用であった微生物資源を探し出す上では極めて強力なツールとなっている。地球上には1兆種もの生物が存在しているとされているが⁵⁾、そのうち、99.99%以上の生物がいまだ認識されておらず、その中には新規物質生産系の宿主微生物となるものや、バイオものづくりにおいて有用な遺伝子を有する微生物が存在していると考えられる。ゲノム構築⁶⁾や遺伝子機能予測⁷⁾等のバイオインフォマティクス技術によって、大量に読まれた核酸配列から、これまで未知であった微生物の存在を明らかにし、そこから有用な遺伝子や代謝経路を有しているかを判断できるようになっている。例えば、最近では、自然環境から複数の微生物由来DNAを取得し、ここから微生物(系統)単位でゲノムを再構成するMetagenome Assembled Genome(MAG)の手法⁸⁾が注目されている。得られたMAGは完全なゲノム情報ではないものの、新たな遺伝子や酵素・代謝経路を一度に得ることができ、自然環境から生物資源の網羅的な探索を可能とした。

また、バイオものづくりでは、微生物の改変をいかに合理的に行うかが重要であるが、そのためにはゲノム情報や代謝反応データベースからの代謝反応モデル(Genome Scale Model; GSM)⁹⁾の構築、そのモデルに基づいた代謝フラックスバランス解析(Flux Balance Analysis; FBA)¹⁰⁾がそれぞれ行われる。通常、ゲノムスケールでは数千以上の代謝反応が含まれることから、手作業による構築には多くの時間と労力を要する。昨今では、GSMに関するデータベースも拡充されてきており、ゲノム情報からGSMを自動構築する技術⁹⁾や人工的な代謝経路の設計技術^{11,12)}等と合わせて宿主微生物開発の効率化が図られている。GSMの構築プロセスでは、ゲノム情報から配列の相同性やモチーフ検索により酵素等の遺伝子機能を予測し、その結果をKEGG(<https://www.kegg.jp/>)やBRENDA(<https://www.brenda-enzymes.org/>)といった代謝反応・酵素反応のデータベースに対して検索することで、関連する代謝反応情報を取得する。その後、基質や中間物質、細胞内外の物質輸送、代謝産物の取り込みや排出などがモデルに組み込まれ、ゲノムスケールでの代謝経路がモデル化される。代謝フラックスバランス解析(FBA)では、上記で設計された代謝モデルを含む反応を化学量論式として束ねた上で代謝フラックスが計算される。計算の際には、条件等に応じてフラックスの上限・下限、利用する栄養源、実測値への固定といった各種条件を設

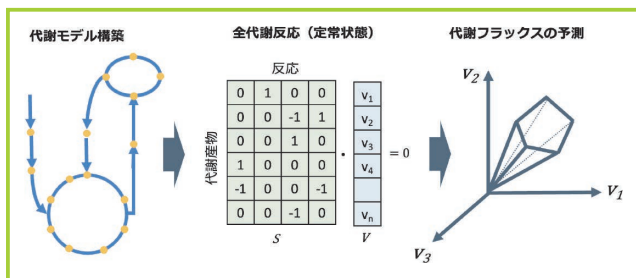


図3 代謝モデル構築からFBAまでの流れ

定し、細胞の増殖速度等を目的関数にとった線形最適化問題として解かれる。FBAによって、特定の条件下でどの代謝反応が活性化され、どの代謝物が生産されるか等をシミュレーションできるため、微生物改変の影響や効果を事前に予測することができる。このようにGSM設計とFBAによって、微生物育種における改変候補の策定を合理的に行うことができる(図3)。

また、実際の微生物設計では、目的物質生産時に機能する関連遺伝子や転写因子等を探し出し、それらがどのように協同して働いているかを調べるのが重要となる。このアプローチの1つとして、遺伝子発現制御ネットワーク技術が利用されている^{13,14}。遺伝子発現制御ネットワークは、主に「遺伝子選択」と「ネットワーク構造推定」の2つのプロセスからなる。遺伝子選択では、主に高生産株と低生産株のような物質生産量の違いに応じて統計的に有意な変化を示す遺伝子や物質生産量と有意に相関を示す遺伝子を候補として選択する。その後、これらの候補遺伝子による遺伝子発現の関係性をベイジアンネットワーク等による数理モデルによって初期モデルとして構築し、構造方程式モデリング(SEM)を基盤とするネットワーク構造推定によってターゲット遺伝子等に対する制御候補の遺伝子を探し出し、改変候補として提案する。

Learn(学習)プロセスでは、Test(検証)プロセスで得られた実際の計測データのフィードバックを受けることになる。例えば、代謝反応における学習としては、Designで予測された代謝フラックスとTestで計測した代謝情報のギャップを解析する。ギャップとしては、予測時には想定していない未知の酵素や競合阻害物質、代謝物、副次的な代謝物の存在がないか等を実際の計測データから検討し、これらを次のDesignプロセスにおいて、新たな特徴量や構築モデルへの制約条件として設定する。さらに最近では、新たな代謝や酵素反応の探索に向けたアプローチとして、文献情報や基質・生成物の化学構造、酵素アミノ酸配列等を学習したモデルによって新たな基質・生成物と酵素アミノ

酸配列の組み合わせを探し出すこと等も行われている^{15,16}。

おわりに

培養工学や発酵工学を基に発展してきたバイオものづくりでは、これまで職人技とも言える、専門家による豊富な経験や高度な知識を元に育種開発が行われてきた。それに対し、昨今のバイオインフォマティクスによる高度に合理化されたバイオものづくり技術の登場は、育種開発のコンセプトを変えたエポックメイキングな出来事の1つと言える。最近では、培地成分の設計¹⁷や、培養過程の遺伝子発現量変化をモデル化して解釈する¹⁸等、バイオものづくり分野において機械学習やAI技術を取り入れたバイオインフォマティクスの研究開発が積極的に進められている。

しかしながら、その活用にはまだ解決すべき課題もある。例えば、複数の微生物や異なる物質生産系に適用可能な汎化性に優れた機械学習・AIモデルの構築やその予測はまだまだ困難である。この理由としては、データセット数がまだ限られていることに加え、条件の不均一性、バッチ間差や計測ノイズが潜在的に存在していること等が考えられる。より効率的な育種開発を行うためには、統一化されたデータセットを作り出すための計画的な実験デザインや実験の自動化技術等も重要なことは言うまでもないが、生成AIやゼロショット学習等の限られたデータセット数で適用可能な手法の活用も今後重要になるだろう。

- 1) OECD, The Bioeconomy to 2030: Designing a Policy Agenda, 2009.
- 2) バイオ小委員会・経済産業省, 「バイオテクノロジーが拓く第五次産業革命」, 2021.
- 3) S. A. Benner, A. M. Sismour, *Nat. Rev. Genet.* **2005**, 6, 533.
- 4) <https://www.nature.com/collections/aehfjihbj>
- 5) K. J. Locey, J. T. Lennon, *Proc. Nat. Acad. Sci. USA* **2016**, 113, 5970.
- 6) P. E. Compeau, P. A. Pevzner, G. Tesler, *Nat. Biotechnol.* **2011**, 29, 987.
- 7) L. Stein, *Nat. Rev. Genet.* **2001**, 2, 493.
- 8) J. C. Setubal, *Biophysical Reviews* **2021**, 13, 905.
- 9) I. Thiele, B. Ø. Palsson, *Nat. Protoc.* **2010**, 5, 93.
- 10) J. D. Orth, I. Thiele, B. Ø. Palsson, *Nat. Biotechnol.* **2010**, 28, 245.
- 11) https://www.jba.or.jp/nedo_smartcell/theme/06.php
- 12) M. Araki, R. S. Cox III, H. Makiguchi, T. Ogawa, T. Taniguchi, K. Miyaoku, K. Miyaoku, M. Nakatsui, Y. K. Hara, A. Kondo, *Bioinformatics* **2015**, 31, 905.
- 13) S. Aburatani, *Gene Regul. Syst. Biol.* **2011**, 5, 75.
- 14) S. Aburatani, H. Toh, *IGI Global* **2015**, 458.
- 15) N. Watanabe, M. Murata, T. Ogawa, J. C. Vavricka, A. Kondo, C. Ogino, M. Araki, *J. Chem. Inf. Model.* **2020**, 60, 1833.
- 16) T. Fuji, S. Nakazawa, K. Ito, *Bioinformatics* **2020**, i1770.
- 17) K. Yoshida, K. Watanabe, T. Y. Chiou, M. Konishi, *J. Biosci. Bioeng.* **2023**, 135, 127.
- 18) K. Ishiya, T. Taniguchi, *AIP Conf. Proc.* **2023**, 2872.