

海洋における環境ゲノミクス

Environmental genomic approach in marine ecosystems

木暮 一啓*

Kazuhiro KOGURE*

東京大学大気海洋研究所

Atmosphere and Ocean Research Institute, The University of Tokyo

摘 要

近年の遺伝子解析技術の発展は、海洋環境中の生物群集の遺伝子を直接抽出、解析することを可能にした。とりわけ、いわゆる次世代シーケンサーの開発により、微生物群集を中心に短時間に飛躍的に多量の塩基配列データを得ることが可能になった。これにより、従来見落とされていた未知の微生物群集の存在、あるいはプロテオロドプシンに代表されるような新たな遺伝子の存在が確認されてきた。さらに、遺伝子の発現解析により、それらの機能についての考察が可能になりつつある。しかし環境中の膨大な数の未知の微生物および遺伝子の存在は、それらを分離培養し、その遺伝子構造と機能との関わりを丹念に記述していく作業の必要性を示している。今後、遺伝子情報と環境情報の総合的な把握とそれらの多量のデータを扱うバイオインフォマティクスの導入により、新たな海洋生命観が生まれてくることが期待される。

キーワード：遺伝子解析、遺伝子発現、機能、群集構造、微生物

Key words：genetic analyses, gene expression, function, community structure, microorganisms

1. はじめに

近年、メタゲノムという用語がしばしば用いられる。ゲノムとは遺伝子(Gene)と集合(-ome)を合わせた言葉で、ある生物の持つ遺伝子全体を示す。メタとはギリシャ語起源で、複数の事象の上位あるいは統合的な意味を示す。つまり、メタゲノムとは、生物の集合体からなる環境試料から直接得られた遺伝子を対象として扱う研究領域を示す。本来全ての生物群が対象となるが、ここではほぼ微生物を対象に考える。というのは、海洋では微生物は数的にも生物量的にも大型の動植物を圧倒するとともに、動植物は個々の生物体を個別に扱うことができるので、多数の個体をまとめて解析するような方法論に意味がないためである。なお、本稿での環境ゲノミクスはメタゲノムと同義として扱う。

微生物は真核性の単細胞生物と原核生物からなる。本稿で扱う対象微生物は、特に断りがない限り、真正細菌 *Bacteria* と古細菌 *Archaea* からなる原核生物である。その第一の理由は、海水中にはさまざまな真核性の微生物も含まれるが、原核生物の細胞数が相対的に圧倒的に多いためである。海洋表層付近での原核生物の細胞密度はおおむね 10^6 /mL である。これに対して植物プランクトンや原生動物の細

胞密度は通常その2桁から3桁下なので、単細胞生物の遺伝子を扱う場合には必然的に原核生物のそれが主要な部分を占める。ちなみに、全海洋に生息する原核生物の数は 10^{29} と推定されている¹⁾。対象を原核生物に絞る第二の理由は、それらの大部分が培養できない、つまり人工的な環境下で選択的にその細胞数を増加させることができないためである。さらに、個々の細胞サイズが小さく、個々の細胞を扱うことが極めて困難である。このためその機能や分類群、群集構造などの推定には直接その遺伝子を対象にしたアプローチが有効となる。

通常、一般動植物に対する遺伝子レベルの解析は、特定種の全ゲノム解析を除けば目的に応じてその対象種、さらに対象遺伝子を決めた上で解析に着手するのが普通である。しかし、環境ゲノミクスのアプローチは、膨大な数の未知の微生物種とそれらが持っている膨大な数の遺伝子の混合物を対象とする。例えば、海洋表層の1 mLの海水中に 10^6 の微生物がおり、それぞれが3,000の遺伝子を持つと仮定するならば、そこには 3×10^9 の遺伝子が存在することになる。現状ではこの 10^6 が何種を含むのかわからないので、実際に何種類の遺伝子がそこにあるかはわからない。しかしわれわれはその解析の手法を手に入れつつあると同時にそれを軸にした新しい研

受付：2011年1月20日、受理：2011年2月14日

* 〒277-8564 千葉県柏市柏の葉5-1-5, e-mail: kogure@aori.u-tokyo.ac.jp

究分野が広がりつつある。本稿ではそうした手法について簡単に紹介するとともに、そこから生まれてきた研究結果について、群集構造、特定遺伝子、発現解析の3点に絞って概観する。

2. 遺伝子解析技術の発展

新しいアプローチを可能にしたのは、近年の急速な分子生物学的な技術の発展である。これらの技術が海洋微生物群集に適用され始めたのは1990年前後。メタゲノムというアプローチがとられ始めたのは2000年前後。そして2010年の現在では、これに新たな塩基配列解析技術が加わって膨大な量のデータが急速に集約されつつある。これらの技術の詳細については関連する専門書や技術書を見ていただきたいが、本稿に関連することについてのみ簡単に記述しておく。

2.1 特定遺伝子の検出

環境ゲノミクスの出発点は環境中の特定の遺伝子をPCRによって増幅し、その存在を明らかにするとともに、遺伝子の塩基配列を求めてデータベースと照合し、系統的な位置づけを解明するアプローチである。これは対象遺伝子の塩基配列が既知であり、その遺伝子増幅のために適当なプライマー、つまり保存性の高い領域が複数あることが条件になる。それぞれの位置が遺伝子の両端近くに近いほど、PCRによって増幅される範囲が広がる。

このアプローチの先駆けとなったのは、1990年 Giovannoni ら²⁾によって発表された、サルガッソー海の微生物群集の16S RNA 遺伝子の解析結果である。彼らはサルガッソー海の表層水を孔径0.1 μmのフィルターにてろ過後、フィルター上に集積した細胞を壊してDNAを抽出し、そのDNA中から16S rDNAをPCR増幅し、クローニング後、ダイデオキシ法という手法でその塩基配列を求めた。その結果、計12のクローンを解析して9の新たな配列を見出し、検討の結果、その一部はラン色細菌のそれに相当すること、さらにSAR11というクラスターがあることを示した。

その後の技術の進展により塩基配列の決定法は大きく変わりつつあるが、基本的なアプローチは現在でも同じである。この手法が最も有効に使われているのは、上記のような16S rDNAを対象とした群集構造解析および、後述するように、特定の機能遺伝子を対象にした一連の研究である。

2.2 未知の遺伝子の検出

天然の遺伝子配列の中で、その機能がわかっているものは実はごく一部にしか過ぎない。つまりPCRによらずに環境中に存在する遺伝子全体を明らかにしていく方法、つまり、環境ゲノミクスのための手法が必要である。最も広く用いられてきたのは、環境中のDNAを抽出後、制限酵素で分解した

後、プラスミッドベクター、大腸菌人工染色体ベクター(BAC: Bacterial artificial chromosome)、あるいはfosmidベクター等に入れて遺伝子ライブラリーを構築し、その塩基配列を解析していく方法である。このアプローチの威力を示したのが、Venterらによるサルガッソー海表層域の環境ゲノミクスである³⁾。彼らはサルガッソー海の表層から200 Lの水を採取し、DNAを抽出後、プラスミッドベクターを用いてライブラリーを作り、最終的に約10億塩基対の配列を読んだ。そこに、少なくとも148の新種と思われる配列を含む1,800種の細菌の配列、120万の新たな遺伝子の存在を確認している。

現在、このアプローチはさらにmRNAに適用され、いわゆるトランスクリプトーム解析として遺伝子発現を網羅的に調べていく研究に繋がっている(後述)。

2.3 次世代型シーケンサ

2006年、従来の機器とは測定原理を異にするいわゆる次世代シーケンサーが欧米3社(アプライドバイオシステム社、イルミナ社、ロシュ社)から発売され、遺伝子解析は新たな時代に入った。微小なビーズ上でDNAの伸長反応を行い、塩基に応じた発光色を検出して配列を決めるもので、多量のビーズを並行して解析することにより、以前の機器と比較すると解析速度が約2桁向上している。これらの機種は、微生物さらにはウイルスを対象とした環境ゲノミクスにも急速に使われつつある。ただし、いずれの機器も一度に長い配列を読めないことが大きな欠点である。例えばアプライドバイオシステム社 SOLiD4 では50塩基、イルミナ社 Genome Analyzer IIx では75塩基、ロシュ社の454 FLX Titanium では約400塩基程度である。16S rDNAが約1,500塩基であることを考えれば、その解像力には限界があることを認識した上でデータを扱う必要がある⁴⁾。

2.4 特定微生物種的全ゲノム

環境中から得られたDNA試料を分析し、そこにある遺伝子の全容を明らかにしたとしても、その大部分の機能は未知のままであるし、それらがどのような微生物由来なのかも明確ではない。そこで、さまざまな微生物がどのような遺伝子構造を持ち、それらの遺伝子がどのような機能を果たしているのかが知る必要がある。1995年、最初の原核生物であるインフルエンザ菌 *Haemophilus influenzae* のゲノムが明らかにされて以来⁵⁾、急速に解析が進んでいる。2010年12月段階で、NCBI(National Center for Biotechnology Information)には1,379種の原核生物のゲノムが登録されており⁶⁾、しかも3,000以上の解析が進行中である。今後、環境からの情報とこれらの情報を統合していくことが期待される。ただし、分離株のゲノム配列がわかったとしても、それらの遺伝子の機能や発現を制御している

環境要因の多くはまだ知られていないことに留意する必要がある。

3. 環境ゲノミクスの新しい流れ

上記のような新たな解析手法の導入はどのような知見を明らかにしてきたのだろうか。ここでは群集構造、特定遺伝子、遺伝子の発現解析の三つに分けて要点を記述する。

3.1 群集構造

海洋から直接核酸を抽出し、解析するアプローチは Giovannoni ら²⁾によって1990年に始められたが、その後、多くの研究者によってさまざまな海域に適用されてきた。その主要な知見を以下にまとめる。

- ① 図1は分子的な手法によって存在が確認された主な系統群を示す⁷⁾。個々の枝によって真正細菌では11、古細菌では2のグループが示されているが、これらはいずれも広い分布を持つ。この中で、それまで培養法によって存在が確認されていたグループはこれらのごく一部を形成するに過ぎない。つまり、海洋微生物の中で培養できるものはごく一部であること、海洋にはそれまで未知であった原核生物が多量に存在することが明らかになった。
- ② 従来、古細菌の分布は高熱、塩田、嫌気的環境など、いわば特殊な環境に限定されるものと考えられていた。しかし、Fuhrman ら⁸⁾は太平洋水深100mおよび500mから得られた16S rDNA塩基配列の中から、それらのグループとは異なる古細菌の一群の存在を見出し、深海にかなり広範に分布している可能性を指摘した。同年、DeLong⁹⁾は古細菌に特異的なプライマーを用いて沿岸表層域の微生物群集を解析し、古細菌が全菌の2%程度を占めることを報告して

いる。その後、Karner ら¹⁰⁾はハワイのTime series station (HOT)にてFISH (Fluorescent in situ Hybridization)法によって古細菌の量的分布を表層から水深4,750mまで連続的に観察した。その結果、深度に応じて古細菌の相対的な割合が増え、1,000m以深では真正細菌とほぼ同数にまでなること、一般に表層では古細菌の中のEuryarchaeotaという系統群が、深層ではCrenarchaeotaという系統群がより多くなることを示した。これらの研究によって、古細菌が外洋深層に広く分布し、海洋における原核生物の約30%程度を占めることが明らかになった。これらの古細菌の分離培養はまったくできなかったが、最近、Inoue ら¹¹⁾が好塩菌に近いグループを外洋水界中から分離し、系統的には好塩性の古細菌に近いことを明らかにした。

- ③ Giovannoni の見出したいわゆる SAR11 というクラスターは、 α プロテオバクテリアという系統群に属す。多くの海域のとりわけ表層域に共通して出現するグループで、海域に応じて全菌数の4分の1から3分の1程度を占めることが明らかになった¹²⁾。このグループは表層のみならず深層にも分布し、16S rDNAの配列から、かなり多様な一群と見られる¹³⁾。なお、SAR11は2002年に初めて海洋から分離され¹⁴⁾、*Candidatus Pelagibacter ubique*と名付けられ、後にそのゲノムも解明されている¹⁵⁾。

では、いわゆる次世代シーケンサを利用した研究成果は、群集構造解析にどのようなインパクトを与えてきたのだろうか。Sogin ら¹⁶⁾はロシュ社の454 Life Sciencesを用いて、大西洋深層水などの群集構造を解析した。対象としたのが16S rDNAの中で比較的変異が大きいとされるV6 regionという領域だが、塩基数としては80しかないため、系統解析手法としての信頼性には限界がある。しかし、こ

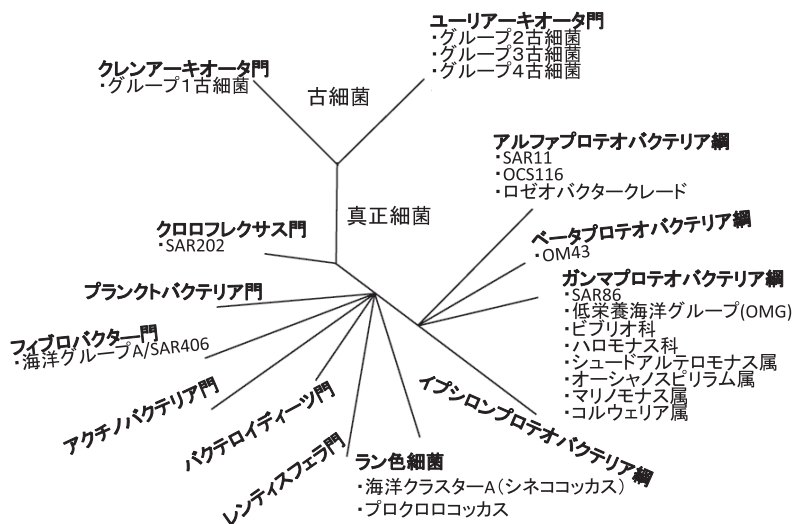


図1 海洋微生物の主要系統群。
Giovannoni and Stingl⁷⁾を改変。

のアプローチはその後 International Census of Marine Microbes (ICoMM, <http://icomm.mbl.edu/>) というプロジェクトとして展開され、約 70 カ国から世界の海域でのサンプルを集め 18 万以上のデータ解析が行われた。その結果、14 万を超える真正細菌および 2,000 を超える古細菌の OTU (Operational Taxonomy Unit) を見出している (<http://vamms.mbl.edu/>)。このシーケンサの特徴は、1 サンプルについて、万単位の配列を読めることである。以前のいわば“古典的”分子生物学的手法ではせいぜい数百程度しか得られなかったことと比較すれば、データ量としては 2 桁程度増えている。

では、こうした多量のデータに基づく解析から何ができてきたのだろうか。その概念を図 2 に示す。この図の中で、横軸にはさまざまな種(厳密には種に相当すると考えられるタクソン)が並んでおり、縦軸はそれぞれの個体数を示す。左方に位置する群は多数の個体数を持ついわばコスモポリタン種、それに対し、右方に向かうにつれて、個体数が減少する。

第一に、次世代シーケンサは、図 1 に示されるようなコスモポリタンなグループの存在を再確認した。図 2 の左方にあるグループはそれに相当し、種数では少ないが、通常個体数全体の 7~8 割を占める。もし 1 サンプルあたりの解析配列数が少ないと、それらばかりが見えてくることになる。第二に、膨大な数の“希少グループ”の存在である。この図の右に位置する半数以上の種はこのサンプルの中では 1 回しか出現しない種(シングルトン)で、その左に連なるのは 2 回出現する種である。シングルTON は多くの海域で、全体の種数の半分から 3 分の 2 程度を占めることが明らかになってきた。第三に、希少グループの多くが個々のサンプルや海域に特異的な傾向を見せる。逆に言えば、探索の対象海域を増やせば増やすほど、そうした希少種の数次第に増えていく。つまり、海洋にはおそらくまだ膨大な数

の種が記載されずに残されており、現在のところその総数がどのくらいになるかは不明、ということになる。

なお、海洋に(あるいは地球上に)どのくらいの数の微生物種が存在するかは興味深い課題である。環境ゲノミクスで 16S rDNA に関する情報を拾い上げていく努力をする一方で、なんらかのモデルを立ててそれを推定する試みが続けられている。一般に微生物群集はその個体数が多いごくわずかな種、その個体数が極めて少ないわずかな種、そしてその中間に位置する種からなると考えられる。それを図示したのが図 3 である。横軸は個体数で、対数で与えられる。縦軸は種数である。例えばこの曲線のより右に位置する種はその個体数がより大きくなる。図の一番右に位置し、海洋で最大の個体数を示すのは SAR11 と予想される。一方、最も種数が多いのは中央のピークに相当する、“多からず少なからず”存在する種群、ということになる。この曲線の形と何点かにおける具体的な数値が出れば、曲線に囲まれた中央部分の量、すなわち総種数が出る¹⁷⁾。まだこの曲線の形状と具体的な値を議論するには早い段階であるが、次世代シーケンサの導入はそうした推定を初めて可能にしつつある。その結果が地球上の生物種全体の推定に必須であることは言うまでもない。

3.2 特定遺伝子の検出と動態

微生物の群集構造はいわば登場人物の特定化である。例えばその登場人物が光合成生物ならば、一次生産に関与することは予想がつくが、図 1 に示されている微生物の大部分は従属栄養細菌であり、彼らが何をやっているのかはわからない。そのような状況下で、遺伝子解析技術の発展は海洋に分布するさまざまな機能遺伝子の実態を明らかにし始めた。例えば DeLong ら¹⁸⁾はハワイの ALOHA Station において、鉛直的に 7 つの層からサンプリングを行い、約 6,400 万塩基対の配列を読んで得られた配列を Kyoto Encyclopedia of Genes and Genomes (KEGG)、

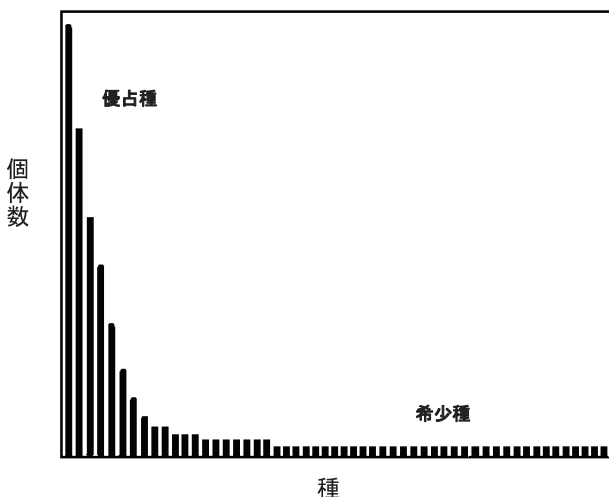


図 2 種とそれぞれの個体数との関係。

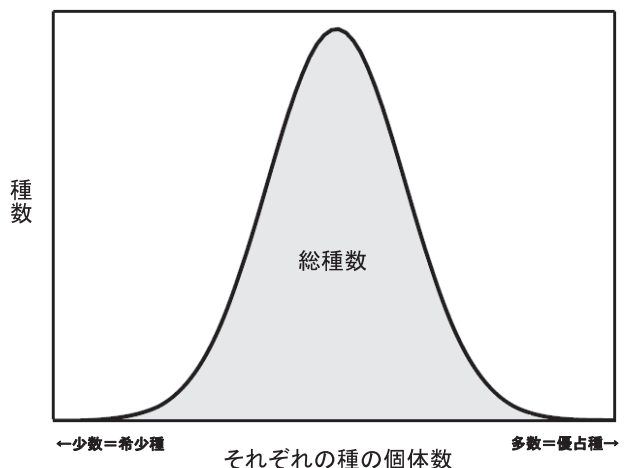


図 3 個体数と種数との関係。

および Clusters of Orthologous Groups (COG) のデータベースを使って検討している。その結果、多数の機能遺伝子の鉛直的な分布が明らかにされており、例えば有光層には光合成関連遺伝子、プロテオロドプシン遺伝子(後述)があると同時に、べん毛合成や化学走性、タイプⅢ分泌系に関わる遺伝子が見られる一方、深層ではピリ合成系、多糖合成系、抗生物質合成系、タイプⅡ分泌系などが見られた、と報告している。なお、遺伝子の存在自体は機能の発揮を意味するわけではない。しかしある遺伝子が不必要ならば次第に淘汰されて見えてこないだろう。さらに、その遺伝子の分布がより広い、あるいはより多様化しているならば、その遺伝子がさまざまな微生物に存在し機能していることを示唆する。表1にそうした視点から研究されてきた機能遺伝子の例を挙げる。ここではそれぞれについて解説するスペースはないので、一例としてプロテオロドプシンの例を挙げる。

1990年代後半に環境ゲノミクスの手法が外洋の海洋微生物群集に適用された際、SAR86(図1)の16S rDNAを含む遺伝子断片の中に、ロドプシン遺伝子に似た配列が見いだされた¹⁹⁾。ロドプシンとはヒトの網膜の中にあって光受容に関わる色素で、オプシンと呼ばれるタンパクとレチナル(ビタミンA関連物質)からなる。従来、塩田を赤くしている古細菌がバクテリオロドプシンという化合物を持ち、それが光エネルギーを利用してプロトンを出し、ATP合成をする機能があることが知られていた。しかし、外洋にはそれらのグループは存在しないと考えられているし、そもそも SAR86 は γ プロテオバクテリアに属す真正細菌である。解析の結果、このロドプシン様の遺伝子はいわゆる7回膜貫通型の構造を持ち、520 nm をピークとする光吸収を持つ物質をコードしていることが確認された。さらに大腸菌の中でその遺伝子発現を行った結果、プロトン排出能を持つポンプであることがわかった。排出されたプロトンは細胞膜内外に勾配を作る。この勾配に応じてプロトンが細胞内に流入する際に ATP 合成や能動輸送などにエネルギーを供給する。これらの結果から、SAR86には光を使ってエネルギー形成を行う新しいタイプのロドプシンがあることがわかり、プロテオロドプシンと名付けられた²⁰⁾。

現在までのところ、プロテオロドプシン遺伝子(*pr*)は SAR86 に加えて SAR11²¹⁾、SAR92²²⁾、Proteobacteria の α 、 β 、 γ グループ、Bacteroidetes とりわけその中の Flavobacteria²³⁾、Polaribacter²⁴⁾ などのグループに、広く分布することが確認されている^{25, 26)}。量的には方法論に依存して大きく異なるが、外洋表層の数%から時に70%~80%の微生物が *pr* を持つ可能性がある²⁷⁾。この見積もりは今後検討の余地があるが、*pr* は少なくとも表層域の微生物群集の間に広く分布しかつ発現されている²⁸⁾と考えられて

表1 海洋微生物群集の機能遺伝子例. Moran 2008²⁹⁾を改変.

<u>光合成系</u>	
<i>pufM</i>	酸素非発生型光合成
<i>rbcl</i>	RuBisCo
<i>pr</i>	プロテオロドプシン
<i>psbA</i>	光化学系Ⅱ
<u>窒素代謝系</u>	
<i>narG</i>	硝酸還元
<i>nirS</i>	亜硝酸還元
<i>nirK</i>	亜硝酸還元
<i>norB</i>	一酸化窒素(nitric oxide)還元
<i>nosZ</i>	亜酸化窒素(nitrous oxide)還元
<i>amoA</i>	アンモニアモノオキシゲナーゼ
<i>hao</i>	ヒドロキシルアミン酸化還元
<i>nifH</i>	窒素固定(ニトロゲナーゼ)
<u>イオウ、鉄代謝</u>	
<i>dsr</i>	硫酸還元
<i>aps</i>	APS(Adenosine phosphosulfate)還元
<i>sox</i>	イオウ酸化
<i>fox</i>	ferrous iron oxidation
<u>酸化、分解関連</u>	
<i>xsc</i>	タウリン分解
<i>aphA</i>	ポリアミン分解
<i>dmdA</i>	DSMP 分解
<i>dmgdh</i>	グリシン-ベタイン分解
<i>chiC</i>	キチナーゼ
<i>pcaH</i>	芳香族化合物分解
<i>boxA</i>	芳香族化合物分解
<i>nahH</i>	カテコール分解
<i>coxL</i>	CO 脱水素酵素
<u>栄養獲得系</u>	
<i>amt</i>	アンモニア輸送
<i>napA</i>	亜硝酸輸送
<i>narB</i>	硝酸還元
<i>nasA</i>	同化的硝酸還元
<i>nirA/B</i>	亜硝酸還元
<i>ureC</i>	ウレアーゼ
<i>phnC</i>	ホスホネート輸送
<i>phoA</i>	リン酸分解
<i>pstS</i>	リン酸取り込み
<u>その他</u>	
<i>hsp</i>	ヒートショック
<i>atpB-H</i>	ATP 合成系
<i>ftsZ</i>	細胞分裂
<i>dna B/X</i>	DNA ポリメラーゼ
<i>rpo</i>	RNA ポリメラーゼ

いる。

もしプロテオロドプシンが細胞内で実際にプロトン排出を行い、プロトン勾配形成を行っているならば、エネルギー的により有利になるはずである。それは例えば光の増殖促進などに表れてくると予想される。その視点に立って、分離株を用いた検討が行われてきたが、奇妙なことに *pr* は環境中には極めて広く分布しているのに、*pr* を保持する分離株の数は 2010 年秋の時点で 10 株あまりしかない^{26, 27)}。さらにそれらの中で光の明確な増殖効果を確認できるのは、*Dokdonia* sp. strain MED134 のみである³⁰⁾。このため、プロテオロドプシンは例えば飢餓状態での生残などにおいてその機能が発揮されるか³¹⁾あるいは別の機能(光センサー)を持つと予想される。*pr* がなんらかの機能を発揮していることは確かと考えられるが、その具体像や海洋生態系へのエネルギー供給量の推定などは今後の検討にゆだねられているのである。

このように、プロテオロドプシンの研究は、環境ゲノミクスによる遺伝子の発見⇒バクテリオロドプシンの知見に基づいたその機能推定⇒遺伝子を保持する微生物の特定化⇒培養株を用いた機能の測定、という順序を経て進んできた。従来の研究は、天然での機能の検出⇒関与する微生物の特定化⇒関与する遺伝子解析、という形で行われてきた。このように、環境ゲノミクスによる新しいアプローチは従来のその逆であり、Reverse Biogeochemistry と称されている²⁹⁾。ここでは、これを“逆システム解析”³²⁾と称すことにする。

3.3 遺伝子の発現解析

逆システム解析によって、今後も新たに未知の遺伝子が見出されていくだろう。しかし、遺伝子の存在は必ずしもその機能の発揮を意味しない。例えば大腸菌の約 4,000 余りの遺伝子のうち、通常の実験室的条件の中で発現されている遺伝子はそのごく一部である。海洋生態系における実際の生物機能を探っていくには、少なくともそれらの遺伝子が発現していることを確認する必要がある。これは伝統的にはなんらかの生化学的反応として測定されてきた。例えば蛍光標識された人工基質で酵素反応を測定する、RI でラベルされた物質の取り込みを見る、同位体を与えてその行方を追う、などである。これに対し、近年いわゆるトランスクリプトームと称されるアプローチ、すなわち遺伝子の発現を網羅的に調べる手法が開発され、海洋微生物群集にも導入されてきた³³⁾。

Frias-Lopez ら³⁴⁾はハワイの ALOHA 測点の水深 75 m から採取した海水からまず DNA を抽出し、次世代シーケンサーを用いてそれらの遺伝子を網羅的に解析し、データベース (COG) と照合してどのようなタンパクに相当する遺伝子が存在するかを調べた。並行して、RNA を抽出し、3' 末端にいわゆ

るポリ A テールをつけた後、逆転写して cDNA ライブラリーを得、次世代シーケンサーでそれらの塩基配列を調べた。さらに、それぞれのライブラリーから得られたタンパクの配列を機能に応じたクラスターに分け、どのクラスターに属すタンパクがよく発現されているかを定量的に検討した。これは“よく使われている遺伝子”を抽出する作業と言える。その結果、mRNA 由来として確認された遺伝子の総数は、全 DNA のライブラリーのそれの約 1/8 に相当した。つまり一部の遺伝子のみが発現している。さらに、mRNA から作られた cDNA ライブラリーのうち、約 4 割の配列が全 DNA ライブラリーには見出されなかった。これは、見落とされている遺伝子が発現されていることを示す。この解析の場合、全 DNA ライブラリー構築のための解析遺伝子数を少なくとも数倍から一桁程度増やさないと全容がつかめないことになる。

環境ゲノミクスとトランスクリプトームを統合した逆システム解析は今後大きな潮流となっていくと予想されるが、以下の理由でその解釈は慎重に行わなければならない。

第一に、遺伝子情報はその転写 (mRNA 合成) ⇒ 翻訳 (タンパク合成) ⇒ 活性発現、という順序を経て機能に至る。遺伝子の存在を生態系の機能に結びつけるには、トランスクリプトームに加えて、さらにプロテオーム (全タンパクの解析) さらにその活性測定を行うことが必要である。しかし多くの場合、活性の測定には感度や手法上の問題がある。活性が測定されていない段階では、「mRNA が合成されている以上、おそらく機能しうる」とまでしか言えない。

第二に、われわれはまだ生物界に広がるさまざまな遺伝子のごく一部しか知らない。例えば、Gilbert ら³⁵⁾は環境中から得られた遺伝子配列を SEED というデータベース (http://www.theseed.org/wiki/Home_of_the_SEED) と照合した結果、その機能が推定されるものは約 30% (20% ~ 46%) 程度であり、残りはどんな遺伝子かわからなかった、と報告している。現在、最も解析が進んでいる生物であると見なされる大腸菌でも、その約 3 分の 1 の遺伝子の機能は推定段階にあるか不明である³⁶⁾。そんな状況下で、環境中から新たな遺伝子配列が蓄積していても、不明遺伝子が増える一方になりかねない。つまり、さまざまな遺伝子の基礎的情報なしには環境ゲノミクスは発展しえない。

第三に、近年の分子生物学では遺伝子を理解すれば生命のメカニズムを理解できる、という考えが主流であった。つまり個々の遺伝子がコードするタンパクの機能一つ一つとその発現メカニズムを重ね合わせていけば生命の基本的なメカニズムを知ることができることになる。例えば遺伝病の克服にはその遺伝子を特定化して遺伝子治療を施せば解決される、と期待される。しかし解析が進むにつれて、生

体はある遺伝子が機能しなくなれば、別の関連遺伝子が動き出して補完的な働きをする、すなわち遺伝子の集合体がネットワークとして機能していることが明らかになってきた。微生物でも同様に、例えば微生物の教科書にはほぼ必ずグルコースとラクトース混合培養系の記述がある。すなわち、この二つの基質を同時に培地に入れて細菌を増殖させると、グルコースが β ガラクトシダーゼの発現を抑えるため、ラクトースは利用されずにグルコースのみが使われ続ける。グルコースが枯渇するとその発現抑制が消えて、 β ガラクトシダーゼの生合成が始まり、それによってラクトースが分解、利用される。しかし、この著名な現象はグルコースの濃度を次第に下げていくと見えなくなる³⁷⁾。また、Liuら³⁸⁾は、グルコースで増殖させた菌体を次いで酢酸、あるいはプロリンの培地に移すと、増殖速度が低下するとともに、それぞれ200以上の遺伝子の発現が増加することを見出した。つまり、グルコースからより使いにくい別の基質が与えられると、大腸菌はより広い種類の基質を利用できるようにその代謝系を変化させる。となると、個々の遺伝子発現をばらばらに解析するのではなく、ネットワークとして捉える必要がある。“逆システム”という和語は、そもそも制御系のネットワーク構造を想定して提案された用語であり³²⁾、ここでもそのような視点で捉えることが必要と考える。

4. 今後の展開

さて、環境ゲノミクスのゴールは単にどんな遺伝子がどこにいるのかを明らかにすることではなく、それらが実際にどのような機能に結びついているのかどうか、そしてそれがどのようなネットワークによって制御されているのかを知ることであろう。逆システム解析とは、それを遺伝子レベルで、その発現解析と環境データとを総合させて解析するアプローチのことである。今後これが大きな潮流となっていくことは間違いない。しかし、ネットワークと言っても、個々の生物個体内のネットワーク、生物間ネットワーク、生物-環境間のネットワークが相互に階層的な構造を作っており、際立って複雑である。一細胞内の代謝ネットワークの理解もままならない状況下で、生態系全体のネットワーク構造が解明されるのだろうか。計算機の能力の格段の向上とアルゴリズムの開発、革新的なアプローチとそれらに必要な莫大な研究費があったとしても、現状ではそれは不可能と見なしておくべき、と筆者は考えている。したがって個々の研究者は、それぞれが例えば特定の遺伝子や生物群に着目するような独自の切り口でシステムのどこか一部を解明していくこと、そしてそれが全体の理解にどのように結びついていくのかを常に意識していくことが求められる。関連して若

干繰り返しになるが、今後の方向性としていくつか指摘しておきたい。

第一に、環境中にあると想定される未知の膨大な遺伝子群に比較して、われわれの持つ既知情報はあまりにも少ない。海洋の微生物の新たな分離培養を試みる、それらのゲノム解析を行う、そしてそれらの生理特性の解析から機能遺伝子の役割を明らかにしていく作業が求められている。その際、これまで一般的に行われてきた富裕な栄養状態と好適な物理化学的環境下での知見のみではなく、より天然環境に近い条件下での挙動を見ることが必要である。

第二に、われわれは遺伝子発現メカニズムについてももっと知る必要がある。大腸菌はあらゆるモデル生物の中で最も解析が進んだ種で、基本的には余計な遺伝子をそぎ落とす方向で進化してきた³⁹⁾。しかしこの種でさえ、異なる代謝系が相互にどのようなネットワークでつながれ、それらがどのように制御されているのかがやと垣間見え始めた段階である。前述したように、Liuら³⁸⁾によれば、使いにくい栄養基質が与えられると、菌はいろいろな基質を利用できるように、200以上の新たな遺伝子を発現させる。それらの遺伝子がコードするタンパクが合成されたとしても、実際には使われないかもしれない。つまり、転写は必ずしも機能とリンクしないことになる。このように、複数の遺伝子間の制御メカニズムの理解は、環境ゲノミクスのトランスクリプトームの正確な解釈に必須である。

第三に、多くの微生物学者は環境から遺伝子を抽出し、特定の遺伝子の塩基配列を決め、データベースと照合してその機能や系統的な位置付けを明らかにする技術を持っている。しかし環境ゲノミクスが対象にしている遺伝子情報はその量も質も桁違いに大きく、従来の技術では対応できない。このため、いわゆるバイオインフォマティクスの専門家との連携が必須になりつつあるが、この分野の人材は生物系の研究者の数と比較すればごくわずかである。その人材養成が、環境ゲノミクスのデータを客観的かつ有効に活用するには必須となりつつある。

環境ゲノミクスは、個々の生物の個体レベルの知見から、群集内での相互作用、さらには環境との関わりなどを総合的に扱うとともに、さらにはバイオインフォマティクスという新たな学問を包括する壮大な研究領域である。極めて魅力的である一方で、何を目的にしてどこからどのような切り口を狙うのかを明確にしないと、データの山に埋もれて何も言えないことになりかねない。ともあれ、この拙文が環境ゲノミクスの理解、発展とそれに基づく新たな海洋の生命観形成に役立てば幸いである。

引用文献

- 1) Whitman, W. B., D. C. Coleman and W. J. Wiebe

- (1998) Prokaryotes: The unseen majority. *Proceedings of National Academy of Science of the USA*, 95, 6578-6583.
- 2) Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, 344, 60-63.
 - 3) Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Neelson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers and H. O. Smith (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304, 66-74.
 - 4) Wommack, K. E., J. Bhavsar and J. Ravel (2008) Metagenomics: Read length matters. *Applied and Environmental Microbiology*, 74, 1453-1463.
 - 5) Fleischmann, R. D. *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 269, 496-512.
 - 6) National Center for Biotechnology Information (NCBI)
<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>
 - 7) Giovannoni, S. J. and U. Stingl (2005) Molecular diversity and ecology of microbial plankton. *Nature*, 437, 343-348.
 - 8) Fuhrman, J. A., K. McCallum and A. A. Davis (1993) Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. *Applied and Environmental Microbiology*, 59, 1294-1302.
 - 9) DeLong, E. F. (1992) Archaea in coastal marine environments. *Proceedings of National Academy of Science of the USA*, 89, 5685-5689,
 - 10) Karner, M. B., E. F. DeLong and D. M. Karl (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*, 409, 507-510.
 - 11) Inoue, K., T. Itoh, M. Ohkuma and K. Kogure (2011) *Halomarina orientalis* gen. nov., sp. nov., a halophilic archaeon isolated from a seawater aquarium. *International Journal of Systematic and Evolutionary Microbiology*. (in press)
doi: 10.1099/ijs.0.020677-0
 - 12) Morris, R. M., M. S. Rappe, S. A. Connon, K. L. Vergin, W. A. Siebold, C. A. Carlson and S. J. Giovannoni (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*, 420, 806-810.
 - 13) Rusch, D. B. *et al.* (2007) The Sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern tropical Pacific. *Plos Biology*, 5, 0398-0431.
 - 14) Rappe, M. S., S. A. Connon, K. L. Vergin and S. J. Giovannoni (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature*, 418: 630-633.
 - 15) Giovannoni, S. J., H. J. Trip, S. Givan, M. Podar, K. L. Vergin, D. Baptista, L. Bibbs, J. Eads, T. H. Richardson, M. Noordewier, M. S. Rappe, J. M. Short, J. C. Carrington and E. J. Mathur (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, 309, 1242-1245.
 - 16) Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta and G. J. Herndl (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of National Academy of Science of the USA*, 103, 12115-12120.
 - 17) Curtis, T. P., I. M. Head, M. Lunn, S. Woodcock, P. D. Schloss and W. T. Sloan (2006) What is the extent of prokaryotic diversity? *Philosophical Transactions of the Royal Society B*. 361, 2023-2037.
 - 18) DeLong, E. F., C. M. Preston, T. Mincer, V. Rich, S. J. Hallam, N. -U. Frigaard, A. Martinez, M. B. Sullivan, R. Edwards, B. R. Brito, S. W. Chisholm and D. M. Karl (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science*, 311, 496-503.
 - 19) Beja, O., L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich and E. F. DeLong (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science*, 289, 1902-1906.
 - 20) Beja, O., E. N. Spudich, J. L. Spudich, M. Leclerc and E. F. DeLong (2001) Proteorhodopsin phototrophy in the ocean. *Nature*, 411, 786-789.
 - 21) Giovannoni, S. J., L. Bibbs, J-C. Cho, M. D. Stapels, R. Desiderio, K. L. Vergin, M. S. Rappe, S. Laney, L. J. Wilhelm, H. J. Tripp, E. J. Mathur and D. F. Barofsky (2005) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature*, 438, 82-85.
 - 22) Stingl, U., R. A. Desiderio, J-C. Cho, K. L. Vergin and S. J. Giovannoni (2007) The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Applied and Environmental Microbiology*, 73, 2290-2296.
 - 23) Zhao, M., F. Chen and N. Jiao (2009) Genetic diversity and abundance of Flavobacterial proteorhodopsin in China Seas. *Applied and Environmental Microbiology*, 75, 529-533.
 - 24) Gonzalez, J. M. *et al.* (2008) Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proceedings of National Academy of Science of the USA*, 105,

- 8724-8729.
- 25) McCarre, J. and E. F. DeLong (2007) Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environmental Microbiology*, 9, 846-858.
- 26) DeLong, E. F. and O. Beja (2010) The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *Plos Biology*, 8, 1-5.
- 27) Fuhrman, J. A., M. S. Schwalbach and U. Stingl (2008) Proteorhodopsins: an array of physiological roles? *Nature Review Microbiology*, 6, 488-494.
- 28) Riedel, T., J. Tomasch, I. Buchholz, J. Jacobs, M. Kollenberg, G. Gerdts, A. Wichels, T. Brinkhoff, H. Cypionka and I. Wagneer-Dobler (2010) Constitutive expression of proteorhodopsin gene by a Flavobacterium strain representative of the proteorhodopsin-producing microbial community in the North Sea. *Applied and Environmental Microbiology*, 76, 3187-3197.
- 29) Moran, M. A. (2008) Genomics and metagenomics of marine prokaryotes. In: D. L. Kirchman ed, *Microbial Ecology of the Oceans* (2nd ed.), John Wiley & Sons Inc., New York, 91-129.
- 30) Gomez-Consarnau, L., J. M. Gonzalez, M. Coll-Llado, P. Gourdon, T. Pascher, R. Neutze, C. Pedros-Alio and J. Pinhassi (2007) Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature*, 445, 210-213.
- 31) Gomez-Consarnau, L., N. Akram, K. Lindell, A. Pedersen, R. Neutze, D. L. Milton, J. M. Gonzalez and J. Pinhassi (2010) Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. *Plos Biology*, 8, 1-10.
- 32) 金子勝・児玉龍彦 (2004) 逆システム学, 岩波書店.
- 33) Poretsky, R. S., N. Bano, A. Buchan, G. LeClerc, J. Kleikemper, M. Pickering, W. M. Pate, M. A. Moran and J. T. Hollibaugh (2005) Analysis of microbial gene transcripts in environmental samples. *Applied and Environmental Microbiology*, 71, 4121-4126.
- 34) Frias-Lopez, J., Y. Shi, G. W. Tyson, M. L. Coleman, S. C. Schuster, S. W. Chisholm and E. F. DeLong (2008) Microbial community gene expression in ocean surface waters. *Proceedings of National Academy of Science of the USA*, 105, 3805-3810.
- 35) Gilbert, J. A. *et al.* (2010) The taxonomic and functional diversity of microbes at a temperate coastal site: A 'multi-omic' study of seasonal and diel temporal variation. *PLoS One*, 5 (11), e15545.
doi: 10.1371/journal.pone.0015545
http://www.theseed.org/wiki/Home_of_the_SEED
- 36) Riley, M. *et al.* (2006) *Escherichia coli* K-12: a cooperatively developed annotation snapshot-2005. *Nucleic Acids Research*, 34, 1-9.
- 37) Lendenmann, U. and T. Egli (1995) Is *Escherichia coli* growing in glucose-limited chemostat culture able to utilize other sugars without lag? *Microbiology*, 141, 71-78.
- 38) Liu, M., T. Durfee, J. E. Cabrera, K. Zhao, D. J. Jin and F. B. Blattner (2005) Global transcriptional programs reveal a carbon source foraging strategy by *Escherichia coli*. *The Journal of Biological Chemistry*, 280, 15921-15927.
- 39) Iwasaki, W. and T. Takagi (2007) Reconstruction of highly heterogeneous gene-content evolution across the three domains of life. *Bioinformatics*, 23, i230-i239.



木暮 一啓

Kazuhiro KOGURE

1975年に東京大学農学部水産学科卒業、1980年に同海洋研究所微生物部門にて大学院を修了。農学博士。日本学術振興会奨励研究員、アメリカメリーランド大学ポストドクトラルフェローを経て、1983年に海洋研究所助手、1993年に同助教授、2002年に教授に就任。現在、東京大学大気海洋研究所、地球表層圏変動研究センター、生物遺伝子変動分野教授。

