

メタゲノミクスを用いた大規模海洋調査

遠藤 寿*

1. はじめに

人類の調査が及んでいる範囲に限れば、海洋は地球上で最大の規模を誇る生物居住空間である。中でも、太陽光が届く水深 200 メートル以浅は生物活動が特に活発であり、海洋の生物生産の大半を占める。我々に馴染みの深い陸域生態系と比較して海洋が独特なのは、光合成を行う基礎生産者が樹木などの高等植物でなく、微小な植物プランクトンであることだろう。森林などの陸域生態系を考える場合、通常は高等植物から始まる食物網を広大な空間スケールで考慮しなければならない。一方、海洋ではコップ一杯程度の海水をすくえばそこにはすでに基礎的な生態系構造（生産、消費、分解を介した物質やエネルギーの流れ）が存在している。この微小な生態系が地表の約 7 割を覆う海洋全体に広がっており、赤道域から極域に至る地理環境や季節の違いに応じてその構成生物や相互作用の形態は様々である。前世紀末には、この海洋微生物生態系が地球上の光合成生産量の約半分を占めるとい見積もりが提示され (Field et al., 1998)、海洋生態系の実像やその変化が気候変動の文脈でも活発に議論されるようになった。しかし、この生態系を構成する微細な生物群の全容が詳細に測られるようになったのは比較的最近のことである。つまり、我々は海洋微生物生態系というシステムの「機能」を測る技術を先に獲得し、その構成要素、すなわち「部品」に関しては長らくブラックボックスとなっていたのである。そして、この「部品」たる微生物群集の実体把握を可能にしたのが、次世代シーケンシング (NGS) と呼ばれる革新的な塩基配列解析技術の登場であった。この技術は、今日に至るまで海洋微生物学に

おける古典的な疑問の解決に大いに貢献するだけでなく、これまで考えられてもいなかった海洋像を次々と生み出している。本記事では、海洋微生物生態系の理解を飛躍的に発展させたメタゲノム解析の概念とその代表的な成果を、著者らが取り組んできた研究成果を交えて紹介する。

2. 次世代シーケンサーによる大規模配列解析

メタゲノムとは群集中に存在するゲノムの総和を意味し、メタゲノム解析は環境サンプル中の DNA 配列を網羅的に決定することを指す。典型的な外洋の表層海水には、1 ミリリットルの海水中に真核生物が少なくとも数千～数万細胞、原核生物は数十万～数百万細胞、ウイルスは数千万～数億粒子が存在している。つまり、海水サンプルの DNA を「網羅的」に取得しようとした場合、きわめて大量の配列情報を機械的に解読する技術が不可欠である。実際には同じゲノム配列を持つ個体数が多数存在するため固有ゲノム数はずっと小さくなるが、これを考慮しても要求される塩基配列情報は膨大である。真に網羅的な配列を得るのは現在の技術を持ってしても容易ではないが、それに迫る量の配列を得ることは NGS によって可能である。この一連の配列解析技術が実用化されたのは、2005 年に Roche 社から発売された GS FLX+ (通称 454) までさかのぼる。454 の配列解析能力、すなわち一度の稼働で得られる配列の数は現在の基準で言えば特別に高いとは言えないが、それまで主流であったサンガー法と比べると画期的であった。著者は大学院に在籍していた 2008 年頃から遺伝子配列解析を始めたが、当時はまだ海洋学分野で NGS は普及しておらず、変

*京都大学化学研究所附属バイオインフォマティクスセンター助教

性剤濃度勾配電気泳動法 (DGGE) や TA クローニングで単離した配列をサンガーシーケンサーで1本ずつ読み取っていた。これらの手法では、複数回の稼働を経て1つのサンプルから数十からせいぜい百程度の配列 (リード) が得られるに過ぎず、シーケンスの前処理も非常に手間がかかった。これが、454 の登場によって一度に数千のリードが得られるようになり、その後登場する様々な原理に基づく NGS 機器によって出力される配列情報量は指数関数的に増加していった。なお、この最初の NGS である 454 は 2016 年に発売を終了し、実質的に稼働を終えている。さらに進化した NGS が続々と登場し、コストや性能面で競争力を失ったのが原因であろう。現在では、NGS の性質に応じてリード数では 3 億以上、リード長では数十万塩基の出力性能を持つ機器が発売されている。したがって、これらの機器をひとまとめにして「次世代シーケンサー」と呼ぶことに違和感を覚える研究者は少なくない。近年では、第 2 世代、第 3 世代といったように基盤技術に応じて通称を区分する場合も多い。各 NGS の原理や性能は多くの文献やインターネットサイトで紹介しているので、本稿では割愛する。

3. 配列情報に基づく海洋微生物の調査

塩基配列情報が環境調査にどのように役立つのだろうか。言うまでもなく、我々人間を含む生物はそれぞれ種レベル、あるいは個体レベルで固有のゲノムを持っており、それは親から子へと受け継がれる。ゲノムは塩基配列による一連の遺伝子情報を含むため生命の設計図にも例えられるが、この設計図は地球上のすべての生物において起源を共有すると考えられる、よく似通った領域を含んでいる。例えば、すべての細胞性生物はリボソーム RNA (rRNA) 遺伝子という共通の遺伝子セットをゲノム内に有する。この遺伝子はタンパク質翻訳に関わるリボソームを構成する RNA をコードするが、その配列の違いに応じて系統を分類し、またそれらの近縁関係を特定するのに役

立つ。これは、進化の過程でゲノム内に生じた変異が蓄積されたため、近縁の種間ほど似た配列を持ち、遠縁の種間ほど異なる配列を持つという事実に基づいている。微生物の群集構造を幅広い系統群にわたって調べる場合、程よく変異が蓄積されている rRNA の小サブユニット (真核生物では 18S rRNA、細菌・古細菌では 16S rRNA が該当する) を指標とすることが多い。GenBank と呼ばれる塩基配列データベースには、膨大な数の rRNA 遺伝子配列が系統と関連付けられて蓄積されており、環境から得た配列をデータベースに照会することで系統群の同定が可能となる。また、環境中の DNA から微生物の多様性解析を行う場合、試料中に含まれる rRNA 配列も一定閾値以上の類似度をもとに客観的に分類し、その分類群の数を多様性の指標とすることができる。このようにして得た分類単位は、種や属と行った階層的な分類とは基準が異なるため、操作的な分類単位 (Operational Taxonomic Unit, OTU) と呼び区別されている。

環境 DNA サンプルから rRNA 遺伝子を網羅的に取得する方法は大きく分けて 2 つある。1 つ目は、抽出した全量 DNA から rRNA 遺伝子のみをポリメラーゼ連鎖反応 (PCR) で増幅し、その産物をシーケンスする方法 (アンプリコンシーケンス)、もう一つは、PCR 増幅を用いずに断片化したゲノムをショットガンシーケンスで包括的に取得し、その配列から rRNA 遺伝子のリストを再構築する手法である。後者は前者に比べて配列取得効率が劣るが、PCR プライマーのミスマッチや増幅に起因するバイアスを避け、幅広い遺伝子解析に利用可能なメタゲノムライブラリを得ることができる。実際には、分析対象や分解能に応じ、rRNA 遺伝子の他にも多様なマーカー遺伝子が群集構造解析に用いられている。

4. 海洋植物プランクトンの多様性と現存量

大規模配列解析技術が海洋学分野に導入されてから研究者がまず興味を持ったのは、海水中にど

のくらいの OTU 数が存在しているのかということであった。動物や植物などの野外調査では、一般に、ある生息域で発見される生物種数は観測を重ねるにつれ増加するが、その積算値は徐々に収束し、実在する生物の多様性に近づいていく。これは目視観測が不可能な微生物群集にも当てはまり、海水中から得られた標本数、あるいはリード数が多ければ多いほど、真の値に近い多様性が得られる。NGS 以前から用いられてきた顕微鏡観察や小規模の配列解析を用いた場合、その処理能力の低さ、あるいは形態が酷似した異種を分類することができないといった技術的制約から、実際の多様性は過小評価される傾向にあった。そんな中、海洋微生物の多様性研究において一つのブレイクスルーとなったのが、2015 年に科学誌 *Science* に掲載された国際海洋探査プロジェクト *Tara Oceans* の一連の研究成果である。このプロジェクトでは、南大洋、インド洋、太平洋、大西洋など 47 の観測点から得た DNA サンプルを分析し、数億配列に及ぶ 18S rRNA 遺伝子断片を取得した。それらを解析することで、各観測点で数千から数万、全球積算では 10 万を超える真核微生物由来の OTU が報告された (de Vargas et al., 2015)。この 10 万という数字は OTU の定義に依存するため絶対的な意味を持たないが、重要なことは、系統間あるいは地点間で比較可能なデータセットが大規模かつ高精度で得られたということである。中でも、植物プランクトンに属する系統のみに着目すれば、渦鞭毛藻類、珪藻類、そしてハプト藻類が特に高い多様性を有していることが示された。この成果によって、極域や深海などの特殊な環境を除き、海洋全域における真核微生物の多様性や系統の全体像が浮き彫りとなった。その後、海洋微生物生態学者の主な興味は、それら微生物群集の構造を決定する内外的要因や、群集構造や多様性が生態系機能にもたらす影響といったものに移行しつつある。

その一例として、著者らが行った研究を紹介する。本研究では、海洋における主要な植物プラン

クトン系統である珪藻類とハプト藻類の群集組成と多様性を評価し、藻類の多様性がどのような機構に支配されているかを調査した (Endo et al., 2018)。用いた DNA 試料は、海洋研究開発機構 / 東京大学大気海洋研究所の学術研究船白鳳丸 KH-12-3, KH-13-7, KH-14-3 の 3 航海で表層および亜表層クロロフィル極大層から採取した。同航海では、太平洋の東経 160 度線上の北緯 5–47 度、および西経 170 度線上の南緯 40 度から北緯 68 度にかけて計 35 の観測点で試料を採取した。採取した試料から DNA を抽出・精製した後、珪藻類とハプト藻類に特異的なプライマーを用いたアンプリコンシーケンシングにより 18S rRNA 遺伝子断片を取得し、群集解析をおこなった。

各藻類の多様性分布を比較した結果、珪藻類は緯度毎に多様度の変化が大きく、単一の属の優占によって個々の群集において多様度が低下する傾向にあった。一方で、ハプト藻類では高緯度を除く観測点で複数の系統が同程度の寄与率を有しており、多様度が均一かつ高く保たれていた。このような珪藻類とハプト藻類の群集構造および多様度の地理分布の違いは、両系統群の生態生理学的な特性を反映している可能性が考えられる。すなわち、珪藻類は大半の種が光合成に特化した生態戦略を取っており、生理機能の重複によって競争的排除が生じた結果、多様性が低下したと考えられる。このような種は生態学的には *r* 選択型と呼ばれ、高い増殖速度を持つ代償として環境中に安定的に生存することが困難であり、各々の環境に最も適応した種のみが優占する傾向がある。一方のハプト藻類は、光合成による有機物合成に加えて、細菌捕食により増殖資源を獲得できる系統が普遍的に存在しており、著者らの観測でもそれらの系統が常に安定した生物量や多様性を有していた。つまり、ハプト藻類は群集内で生理生態機能を多様化させることで種間競争を避け、多種共存により多様性を確保していたと考えられる。本研究により、これまで主に環境要因によって説明が試みられてきた藻類の多様性形成メカニズムの解

明に、生態戦略という新しい切り口が導入された。

著者らはさらに、同様の試料に対して異なるプライマーを設計し、珪藻類とハプト藻類の rRNA 遺伝子コピー数の空間変化を評価した。この分析には、DNA を PCR で増幅する際のサイクル数と増幅シグナルとをモニタリングすることで、元の試料中の遺伝子コピー数を定量する定量 PCR という手法を用いた。なお、藻類の単位細胞内における rRNA 遺伝子コピー数は種によってばらつきが大きい、細胞サイズとコピー数との間に強い正の相関があることが分かっている (Zhu et al., 2005)。つまり、海水中の rRNA 遺伝子コピー数は、その生物群のバイオマスの指標として用いることができる。太平洋における珪藻類とハプト藻類の現存量分布は、両系統群ともに亜熱帯域で低いが、緯度とともに増加し、亜寒帯域付近で極大となった。これは、海水中の主要栄養塩濃度が高緯度域で高いことと関連しており、過去に現場観測や人工衛星観測で推定された分布と同調的であった。また、高緯度域では概して珪藻類がハプト藻類の生物量を大きく凌駕していたが、鉄欠乏海域においてはハプト藻類が優占することが示された。これは、珪藻類の方が細胞内の鉄要求量が高いことに加え、先程論じたように栄養資源の獲得様式が制限されていたためと考えられる。本研究の結果は、海洋植物プランクトンの分布や多様性、現存量が主要・微量栄養塩の利用可能性といった外的要因や、生態戦略といった内在的因子との相互作用によって形成されていることを示している。

5. 海洋巨大ウイルスの多様性

海洋植物プランクトンの現存量や群集構造に影響を与える他の要因として、ウイルス感染が挙げられる。真核植物プランクトンに感染するウイルスの存在は古くから知られていたが (Cottrell and Suttle, 1991; Milligan and Cosper, 1994; Tarutani et al., 2001)、それらの多様性や系統関係はやはりメタゲノム解析の登場まで長らく謎で

あった。海洋植物プランクトンに感染する種が多く見つかっているウイルス系統群の一つに、核細胞質性大型 DNA ウイルス (Nucleocytoplasmic Large DNA virus, NCLDV) がある。このウイルス群が感染する植物プランクトン宿主としては渦鞭毛藻類、ハプト藻類、緑藻類、ラフィド藻類などが報告されており、そのいくつかに関しては実際の海洋環境においても大増殖した植物プランクトンを死滅させる効果が報告されている (Nagasaki et al., 1994; Brussaard et al., 1996; Jacquet et al., 2002)。しかしながら、NCLDV の現存量や群集構造の定量的な観測は非常に困難であったため、それらの生態学的役割に関してはほぼ明らかとなっていない。

ウイルスが生物か否かの議論はまだ終焉を迎えていないが、分子生態学の観点からすると、ウイルスの分布や多様性は細胞性生物と同様に遺伝子指標を用いて評価が可能である。なぜなら、彼らは DNA ウイルスにしる RNA ウイルスにしる、個体群の「識別コード」である遺伝情報を有して環境に存在しているからである。ウイルスはわずかな例外を除いてリボソーム RNA 遺伝子をゲノム内に持たないため、系統群を跨いで保存されている機能遺伝子をマーカーとして用いるのが一般的である。全種に普遍的に存在する遺伝子が存在するかは疑問だが、NCLDV の解析においてよく用いられるのは、ウイルスゲノムを取り囲むカプシド蛋白をコードする MCP 遺伝子や、DNA ポリメラーゼ B ファミリーをコードする polB 遺伝子などである。

海洋の NCLDV をメタゲノム手法で解析した代表的な研究としては、Hingamp et al. (2013) が挙げられる。彼らは、*Tara Oceans* で大西洋、地中海、紅海、インド洋から採取した 17 の DNA 試料からメタゲノムライブラリを作成し、多様性や分布を評価した。このメタゲノムライブラリから NCLDV 由来の polB 遺伝子が 264 同定され、その多くはメガウイルス科とフィコドナウイルス科の植物プランクトンを宿主に持つ既存の種と近

縁であった。

Tara Oceans 関連プロジェクトの観測地点とサンプル数はその後大幅に拡張しており（現在では太平洋、南極海、北極海からも試料が得られている）、NGS技術の進歩によってメタゲノム配列の情報量も大きく増大した。著者らが最新の *Tara Oceans* メタゲノムデータを用いて再度、*polB* 遺伝子を指標とした多様性解析を行ったところ、全球積算で6,000を超えるNCLDVの系統型が検出され、地理的要因による群集構造の違いも明らかとなってきた。同研究は未発表のデータを含むため本稿では詳しく述べないが、今後、NCLDV群集と真核生物群集との関連を広域スケールで解析することで、ウイルスが海洋微生物群集、ひいては海洋生態系機能に果たす役割が定量的に解明されることが期待される。

謝辞

本稿で紹介した成果のうち、著者が携わった研究を行うにあたり、多くの先生方、先輩方、同輩の皆さんに多大なご支援をいただいた。植物プランクトンの地理分布に関する研究は、北海道大学博士研究員時代に、大学院在籍時の指導教員でもあった鈴木光次教授のプロジェクトのもとで行ったものである。また、巨大ウイルスに関する研究は、京都大学緒方博之教授の研究室に着任後、携わる機会を頂いた。両教授には特に多くのご指導、ご協力を賜ってきた。ここに厚くお礼申し上げる。

引用文献

- 1) Brussaard, C. P. D., Kempers, R. S., Kop, A. J., Riegman, R., Heldal, M. (1996). Virus-like particles in a summer bloom of *Emiliana huxleyi* in the North Sea. *Aquat. Microb. Ecol.*, 10(2), 105-113.
- 2) Cottrell, M. T., Suttle, C. A. (1991). Widespread occurrence and clonal variation in viruses which cause lysis of a cosmopolitan,

eukaryotic marine phytoplankter, *Micromonas pusilla*. *Mar Ecol. Prog. Ser.*, 78, 1-9.

- 3) De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., *et. al.* (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348(6237), 1261605.
- 4) Endo, H., Ogata, H., Suzuki, K. (2018). Contrasting biogeography and diversity patterns between diatoms and haptophytes in the central Pacific Ocean. *Sci. Rep.*, 8, 10916.
- 5) Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281(5374), 237-240.
- 6) Hingamp, P., Grimsley, N., Acinas, S. G., Clerissi, C., Subirana, L., Poulain, J., *et al.* (2013). Exploring nucleo-cytoplasmic large DNA viruses in *Tara Oceans* microbial metagenomes. *ISME J.*, 7(9), 1678.
- 7) Nagasaki, K., Ando, M., Itakura, S., Imai, I., Ishida, Y. (1994). Viral mortality in the final stage of *Heterosigma akashiwo* (Raphidophyceae) red tide. *J. Plankton Res.*, 16(11), 1595-1599.
- 8) Tarutani, K., Nagasaki, K., Itakura, S., Yamaguchi, M. (2001). Isolation of a virus infecting the novel shellfish-killing dinoflagellate *Heterocapsa circularisquama*. *Aquat. Microb. Ecol.*, 23(2), 103-111.
- 9) Zhu, F., Massana, R., Not, F., Marie, D., Vaultot, D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.*, 52(1), 79-92.