

序文：「拡がる環境ゲノム研究」発刊にあたり

中嶋 信美

(国立研究開発法人 国立環境研究所 生物多様性領域 (環境ゲノム研究推進室))

DNAの塩基配列を読み取る技術(シーケンシング技術)が開発されてからまもなく50年を迎える。この技術が開発された1977年当時は1回の分析で100塩基読み取るのが精一杯であったが、1980年代後半にDNAシーケンサーが開発され、20世紀末には1回の分析で50万塩基読み取れるようになった。しかし、20世紀末まではゲノム情報を利用した研究は、ヒト、マウス、イネ、シロイヌナズナなどのモデル生物に限定されていた。ところが、2005年に次世代シーケンサー(Next Generation Sequencer; NGS)が開発され、現在は1回の分析で6兆塩基(ヒトゲノム2,000人分に相当)読み取ることができる。NGSが開発されたことにより、モデル生物以外の生物種についても全ゲノム解読が進み、2024年3月現在、NCBI(The National Center for Biotechnology Information; <https://www.ncbi.nlm.nih.gov/genome/>)には30,530種の真核生物、567,228種の原核生物、66,429種のウイルスの全ゲノム配列が登録されていたが、2024年12月では46,000種の真核生物、2,400,500種の原核生物、205,000種のウイルスの全ゲノム配列が登録されており、わずか9ヶ月で登録数が飛躍的に増加している。このようなことから、21世紀中には現在確認されている全ての生物種の全ゲノム解読が実現する可能性が高く、21世紀は「ゲノム科学の時代」と言っても過言ではない。このように全ゲノムデータベースが充実するに従い、研究方法も大きく変わってきている。これまでは個々の遺伝子の働きを調べ、部分的な情報から全体での働きを推測する方法しか取れなかった研究が、全遺伝子や全一塩基多型情報(SNPs)といった全体情報から必要な情報を抽出し、その細部を解明する研究方法に移行しつつある。そのため、生物を研究対象とする研究者は、研究対象生物のゲノム情報を取得してから研究を開始するのが当たり前になった。

また、種を同定するためのバーコード配列のデータベースも急速に充実してきている。国際バーコード配列データベースであるBOLD(BARCODE OF LIFE DATA SYSTEM; <https://www.boldsystems.org>)に登録された種数は2024年12月現在、動物261,000種、植物72,000種、菌類など25,000種となっている。バーコード配列データベースが充実するに伴い、形態観察では種同定が難しかった種についても、バーコード配列を調べることで、簡単に種同定ができるようになった。また、生物を捕獲せずに排泄物や体毛や皮膚の一部あるいは生息地域の水や土壌を採取して、そこに存在するDNA塩基配列情報をNGSにより解析して、バーコードデータベースと参照することで、調査地に生息している生物種を網羅的に捉えることができるようになりつつある。これにより生物の分布調査にかかるコストが大幅に削減でき、調査地点も増やすことができるため、各種調査において取り入れられつつある。さらに、環境中のRNAを分析することで、その地域に生息している生物の生理状態を捉える試みも始まっている。

一方、環境中の微生物で培養可能な種は全体の1%以下と言われていて、かつてはその全体像をつかむことは不可能だったが、バーコード配列のPCRによる増幅とNGSによる塩基配列解析により、培養できない種についても存在を確認できるようになり、網羅的な微生物生態の解析が可能となった。

以上のように急速なゲノム研究の進展により、環境研究においても、環境汚染物質の影響評価、環境リスク評価、生物多様性評価など多くの分野でゲノム情報を利用した新たな研究が進みつつある。この特集ではゲノム情報の利用によって、大きな進歩が見られた分野や、解析方法に大きな変革がもたらされた分野を中心に、最先端で活躍する研究者にその分野の現状と実際の解析例を紹介していただいた。ゲノム研究の進歩は非常に速いため、この特集が刊行されるころには内容が古くなってしまわないか心配しているほどである。この特集が皆さんのお手元に届いたら、できるだけ早く読破していただけることを願います。

この序文を入稿したあとに、AlphaFold2(武田 一貴氏の総説参照)の開発者が2024年のノーベル化学賞を受賞した。21世紀中には現在確認されている全ての生物種の全ゲノム解読が実現する可能性が高いだけでなく、ゲノムにコードされている全てのタンパク質の機能や役割についても解明される可能性が高いことを示唆するニュースである。