

ゲノム,ゲノム,ゲノム

ゲノムという 闇夜のなかの光明

本雑誌が休刊することになり、この連載も今回が最終回となった。そこで、この10年ほどのあいだ自身の研究の中心としている「ゲノム」について語ってみたい。

私の2冊目の単著「ゲノムと進化」⁽¹⁾は、「ゲノムとは、現代生物学にとって、混沌とした現実の生命界という漆黒の闇夜を抜けだすのに必須の光明である。」という文章から始まっている。この表現に、ピンとこない読者もおられるだろう。生命界が漆黒の闇夜だって?! 生命は生きているのだ。むしろ明るいのではないか。しかし、生命はきわめて多様ではないか。この膨大な多様性という混沌は、単純な論理では到底把握することはできない。この状態を、思い切って「漆黒の闇夜」と表現したのである。

では、なぜゲノムはこの闇夜を抜けだすのに必須の光明なのか? それは、ゲノムの塩基配列という客観的なパターンが、人間の世界認識メカニズムに適合しているからだ。DNAは物質だが、塩基配列は情報とみなすことができる。しかも、4種類の

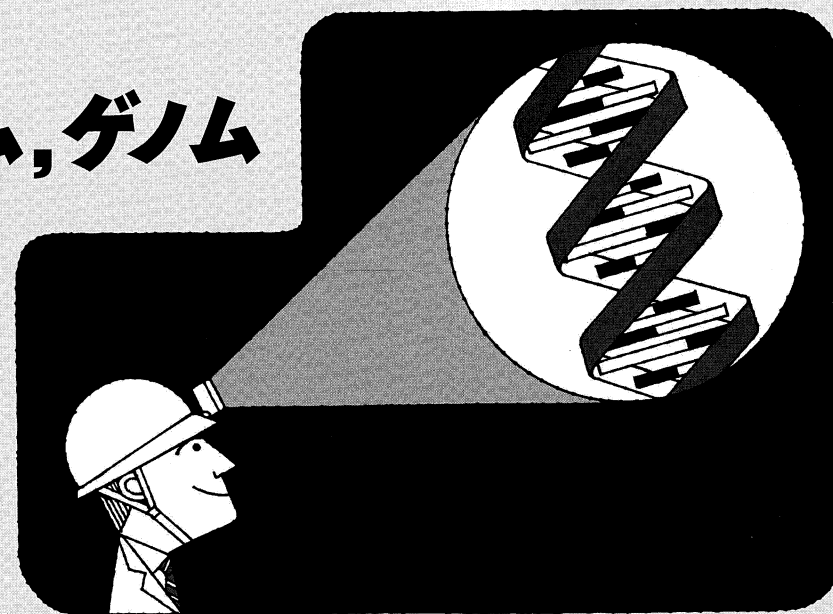


Illustration / Masaaki Hosoda

塩基が一列に並んだ、単純で美しい構造である。この美しさ、明快さにわれわれ人間はすがりついてしまうのである。

遺伝子が生物の物質交代全体をすべて決定すると考える「遺伝子決定論」も、人間の同じような傾向にもつづいている。DNAから出発して考えると、生命現象を理解しやすいからだ。実際の生物、実際の細胞では、DNAもタンパク質も糖も脂質も浑然一体となつて存在しており、DNAにしても自分たちが司令塔だとは思っていないだろう。もちろんDNAに心はないのだが、DNAからRNAを経てタンパク質に情報が流れていくという、古典的な「分子生物学のドグマ」も、人間の認識にすぎない。そのような抑制の効いた認識のもとではあるが、生物をまるごと理解する第一歩として、ゲノム塩基配列はやはり重要である。

斎藤成也

(さいとう・なるや)1957年福井県生まれ。1979年東京大学理学部生物学科人類学課程卒業、1986年テキサス大学ヒューストン校生物学大学院修了(Ph.D.)。1989年東京大学理学部助手、1991年国立遺伝学研究所助教授、2002年同教授。総合研究大学院大学遺伝学専攻、東京大学大学院生物科学専攻教授を兼任。日本学術会議会員。専門分野はゲノム進化、人類進化。

サンガー本人から 始まった比較ゲノム研究

昨今は、いわゆる「第二世代シーケンサー」が流行しているが、つい最近までは第一世代のサンガー法(ジデオキシ法、1977年発表)による塩基配列決定が一般的だった。現在でもサンガー法を用いたキャピラリーシーケンサーを使っている方も多いだろう。この方法を発明したフレデリック・サンガー自身が「ゲノム比較」の草分けであることは、『ゲノム進化学入門』⁽²⁾でも言及した。彼のグループは、ヒトのミトコンドリアDNAのゲノム配列を1981年に決定した後、翌1982年にはウシのミトコンドリアDNAについてもゲノム配列を決定したのである。これら2種の比較から、タンパク質やRNAの遺伝子が高く進化的

*1 全ゲノムショットガン塩基配列決定法

500塩基程度の短い配列を多数ランダムに配列決定し、それらをコンピュータでつなぎ合わせる手法。

*2 ゲノム生態学

ある環境に存在するすべての微生物のゲノムを調べる環境ゲノム学をさらに発展させて、微生物であるかどうかにかかわらず、ある生態系に存在するすべての生物のゲノム配列を決定し、それら膨大なデータから生物間の相互作用を考察する新分野。ecogenomicsという英語もある。

*3

たとえば、増山和花がおこなった系統特異的ドメイン組合せの解析、鈴木留美子がおこなった同義置換の生じない領域の解析、Tim Jinamが進めているゲノム規模SNPデータによる人類集団の解析など。

に保存されていることがわかった。

哺乳類のミトコンドリアDNAゲノムは、1万6000塩基程度の小さなものだが、バクテリアゲノムは数百万塩基の規模である。つぎの大きな目標として、日本の研究者グループが世界に先駆けて1990年ごろに大腸菌全ゲノム解読計画を開始した。しかしさまざまな理由からなかなか進展しなかった。結局、世界初のバクテリア全ゲノム解読は、大腸菌よりもずっとゲノムサイズが小さいヘモフィルス・インフルエンザ菌で、全ゲノムショットガン決定法^{*1}を駆使したアメリカのクレイグ・ベンターたちにさらわれた⁽³⁾。1995年のことである。

それから15年以上が経ち、比較可能なデータは桁違いに増えた。日本DNAデータバンク(DDBJ)のGIB(Genome Information Broker データベース)⁽⁴⁾によれば、2011年2月現在、1371個のバクテリア完全ゲノム配列が公開されている。おそらく今年末までには2000個以上のバクテリアゲノム配列が公開されるだろう。

塩基配列が決定されたバクテリアゲノムは数は多いが、配列としてみれば、ヒトゲノムは標準的なバクテリアゲノムのおよそ千倍である。現在でも、哺乳類はDDBJ/EMBL/GenBank国際塩基配列データベースにおいて、ベスト10のうちの5種(ヒト、マウス、ラット、ウシ、ブタ)を占めている。しかし、おそらくここ数年のうちに大変動がおこるであろう。「メタゲノム」というまったく新しい枠組みが登場してきたからだ。メタゲノムとは、ある環境に生存するすべての生物のゲノムのことである。これらを決定し解析することにより、「ゲノム生態学」^{*2}が始まっている。

斎藤研究室のゲノム研究

私の研究室は、類人猿ゲノム計画の一環として、理化学研究所が中心となった5か国の国際共同研究である「チンパンジー第22番染色体の決定とその解析」⁽⁵⁾に加わり、比較ゲノム研究に参入した。ここ数年は、もっぱら自分たちでゲノムの塩基配列を決定するのではなく、どんどん増えるゲノムデータを取得してそれらを解析することを中心としている。江澤潔は、哺乳類ゲノムの重複遺伝子間における遺伝子変換がひんぱんに生じることを示し^(6,7)、現在これら重複遺伝子の出現パターンを調べている。高橋真保子はバクテリアゲノムのオリゴヌクレオチド頻度情報にもとづく系統樹を作成し⁽⁸⁾、近縁種間ではかなり正確に系統関係を与えることを示した。現在は哺乳類内の系統特異的な非コード領域高度保存配列の解析をおこなっている。大田聡史はヒトSNPデータから哺乳類のGC含量構造の進化を解析し、突然変異パターンの微妙な変化だけで複雑な構造が生まれうることを示した⁽⁹⁾。Kirill Kryukovが開発したソフトウェア「MISHIMA」⁽¹⁰⁾は、バクテリアの多数の近縁なゲノムを高速に多重配列することができ、このほかにも、現在続々とゲノム全体から出発した解析を進めている^{*3}。乞うご期待!

生物学が中心か、 情報科学が中心か

大規模なゲノムデータをコンピュータで解析することが中心である「比較ゲノム学」は、ともすると情報科学の専門家の独壇場であるように思われるかもしれない。理想

的なのは、生物学の素養をもち、コンピュータサイエンスの面にも強いことだ。これら二つの分野のあいだのバランスが必要だろう。また、技術の発達やデータ量の増大速度によって、これらの分野のどちらかがより比重が大きくなる場合がありうる。現在のように、ゲノムデータが急速に増えていくときには、まだ解析ツールが十分普及していないために、データ解析能力が必要とされるが、ある程度落ち着いてくると、どのような切り口でデータを解析するべきかを考える、生物学者の視点が必要となる。

生物学出身研究者に期待したいのは、コンピュータが苦手だと思わず、どんどんプログラム言語を習得したり、積極的に新しい技法を学んでほしいということだ。コンピュータサイエンス出身の研究者には、いろいろな生物学の教科書を読んで、生物学的な思考法を学んでほしい。数学や物理学などで身につけた論理的思考は、かならずやゲノム解析で生きてくるはずである。

私は、バイオインフォマティクスの専門家とみなされることが時々あるが、出身は生物学である。しかし、人類学という風変わりな分野だったので、骨などの多量の数値データを処理するための多変量解析を身につけることが学部時代から要求され、統計学とコンピュータ言語(当時はフォートランだった)が必修だった。この出自が、現在のゲノム配列大規模解析への傾倒を深めた大きな要因だと思う。

結局のところ、新しい時代を切り拓いてゆくのは、既存の分野に属さない、境界にいる人間なのである。そのような若手研究者が陸続と誕生してゆくことを願って、本連載を終える。

参考文献

- [1] 斎藤成也:「ゲノムと進化—ゲノムから立ち昇る生命—」新曜社(2004)
- [2] 斎藤成也:「ゲノム進化学入門」共立出版(2007)
- [3] Fleischmann R D et al: Science 269(1995)496-512
- [4] <http://gib.genes.nig.ac.jp/>
- [5] The International Chimpanzee Chromosome 22 Consortium: Nature 429(2004)382-388
- [6] Ezawa K, Oota S & Saitou N: Mol Biol Evol 23(2006)927-940
- [7] Ezawa K, Ikeo K, Gojobori T & Saitou N: Mol Biol Evol 27(2010)2152-2171
- [8] Takahashi M, Krukov K & Saitou N: Genomics 93(2009)525-533
- [9] Oota S, Kawamura K, Kawai Y & Saitou N: Genome Biol Evol 2(2010)558-571
- [10] Kryukov K & Saitou N: BMC Bioinformatics 11(2010)142