

HUMAN GENOME

ヒトゲノム計画はいま

21世紀はバイオの世紀なのか？

SAITOU NARUYA
齋藤 成也

「ゲノム」とは？

◆本来の機能定義から構造定義へ

「ゲノム」とは、かつては“生命体の生活に必須な最小の遺伝子セット”という意味だったが、今では単に“生命体のもつ遺伝子セットの全体”を指す言葉になっている。「ゲノム」はドイツ語風の発音だが、これはこの概念が誕生した1930年代に、ドイツが世界の生物学をリードしていたなごりである。英語(genome)では「ジノーム」に近い発音になる。

ゲノムの概念が誕生した時代には、DNAが遺伝子の本体であることがまだわかっていなかった。では、どのようなものを調べていたかというと、“染色体”である。長大なDNAからなりたっている染色体は、解像度の高い電子顕微鏡を用いなくても光学顕微鏡で見ることができる。染色体の種類や数を調べることによって、ゲノム構造の違いをおおまかに分析することができるのだ。当時の日本は、小麦の染色体を調べてゲノムの進化を研究する「ゲノム解析」がさかんに行われていた。その中心人物だった木原均氏は次のような言葉を残している。「地球の歴史は地層に刻まれている。生命の歴史は染色体に刻まれている」。この言葉は、遺伝子の物質的本体であるDNAを直接調べることができるようになった現在、ますます輝いている。生物の遺伝子を比較して過去を復元することが、いまの生物学ではごく普通に行われているからだ。



齋藤成也

'57年福井県生まれ、'79年東京大学理学部生物学科卒。現在、国立遺伝学研究所進化遺伝研究部門助教授。専門は人類進化、遺伝子進化。

◆さまざまな生命の
地球上の生物はそれ
が生物種に多様性を与
大きさ(総塩基数)だが、
ゲノムは約30億個の塩基
ジヨウバエ(昆虫)のゲ
る大腸菌のゲノムは40
伝子の数が少ないので
(RNAのゲノムをもつ)
ゲノムをもつ)でも24万
に大きなゲノムをもつ
億個(ヒトゲノムの35倍
の違いは、塩基数の変
いったものである。

一般には、遺伝子の
をつくることができな
ことが期待される。タ
の数種類しかもてない
たとえアミノ酸が突然
ひらの上で動いている
が増加しなければ、ウ
この意味で、ゲノムサ
要因である。しかし、
いDNAを多数抱えてい
い生物には、こうした
ぶ)が多数存在してい
子の種類数が多い複雑
さて、ゲノムの大き
突然変異が起こる必要
伝子重複)、DNA配列
ンによる「遺伝子転移
る「ゲノム重複」があ
る現象だ。たとえば
が4倍体、ナガイモ
生類では倍数体化の
でも、爬虫類・鳥類

◆さまざまな生命のゲノムサイズ

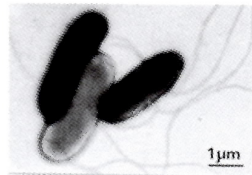
地球上の生物はそれぞれ独自のゲノムをもっているが、ゲノムの多様性が生物種に多様性を与えている。この多様性を示す一つの指標はゲノムの大きさ(総塩基数)だが、それは生物によって大きく異なっている。ヒトゲノムは約30億個の塩基からなるが、遺伝学の研究でよく用いられるショウジョウバエ(昆虫)のゲノムは1億8000万個でヒトの6%、バクテリアである大腸菌のゲノムは400万個で、ヒトの0.1%ほどしかない。ウイルスは遺伝子の数が少ないので、ゲノムのサイズもA型インフルエンザ・ウイルス(RNAのゲノムをもつ)で1万8000個、最大のポックス・ウイルス(DNAのゲノムをもつ)でも24万個である。一方、生物によってはヒトよりもはるかに大きなゲノムをもつものがある。肺魚の仲間には、ゲノムサイズが1100億個(ヒトゲノムの35倍)のものが知られている。このようなゲノムサイズの違いは、塩基数の変化する突然変異が長い進化の過程で蓄積して生じていったものである。

一般には、遺伝子の種類がある程度なければ複雑な物質交代のシステムをつくることができないので、そのような生物ほどゲノムサイズの大きいことが期待される。タンパク質の種類数がいくら多くても、それらのうちの数種類しかもてないウイルスでは、その進化の可能性は限られている。たとえアミノ酸が突然変異でくるくる変化していても、お釈迦さんの手のひらの上で動いているようなものである。何か特別の原因で遺伝子の種類が増加しなければ、ウイルスから別の生命体に進化することはできない。この意味で、ゲノムサイズの増大は生物の多様性を生み出すために必須の要因である。しかし、生物はゲノムのなかに、生存に必ずしも必要ではないDNAを多数抱えているのである。とくにゲノムの大きさがきわめて大きい生物には、こうしたあまり役に立っていないDNA(「がらくたDNA」と呼ぶ)が多数存在している。だからゲノムサイズが大きいからといって、遺伝子の種類数が多い複雑な生物だとは限らない。

さて、ゲノムの大きさを増大させるためには、塩基数が大きく変化する突然変異が起こる必要がある。それには、遺伝子全体のコピーが生ずる「遺伝子重複」、DNA配列の中をあちこち飛び回ることのできるトランスポゾンによる「遺伝子転移」などがある。さらに、ゲノム全体がまるごと倍増する「ゲノム重複」がある。これは倍数体化とも呼ばれ、植物ではよく見られる現象だ。たとえばトコロイモの仲間では、トコロが2倍体、ヤマノイモが4倍体、ナガイモにいたっては14倍体である。また、動物でも魚類や両生類では倍数体化の起こることが知られている。ただし、脊椎動物のなかでも、爬虫類・鳥類・哺乳類では、性染色体の種類によって生まれてくる



ショウジョウバエ



大腸菌

個体の雄雌を決定するメカニズムが確立しているため、倍数体化した個体の子供はすべて雄になってしまう。このため、将来ともゲノム重複を引き起こすことはできないと考えられている。

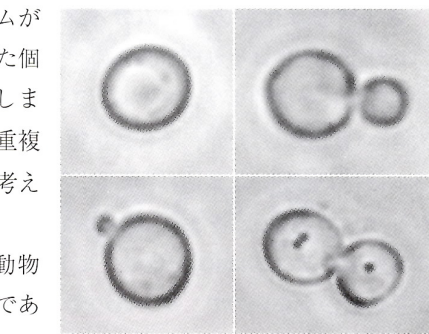
動物のなかでもとくに脊椎動物は、ゲノム構成から見ると特殊であり、たとえば形態形成に重要である *Hox* 遺伝子がじゅつつながりになった

Hox クラスターが、ショウジョウバエなど無脊椎動物ではこれまで調べられた生物すべてに1個しかないのに対して、ヒトやマウスはこのクラスターが4本の別の染色体に散らばって存在する。同様に、主要組織適合性複合体(MHC)遺伝子クラスターも、同じ構造が四つ別べつに存在することが最近発見されている。これらは、脊椎動物が誕生する直前に、ゲノムの倍増が2回あった可能性を示している。

◆さまざまな生物のゲノム計画

ゲノムサイズの小さい生物では全ゲノムの塩基配列が決定されている。バクテリアでは、2000年11月の段階で、すでに36種類の生物(大腸菌、枯草菌、ヘリコバクター・ピロリ、インフルエンザ菌、シアノバクテリアなど)で全ゲノムの塩基配列が公表されている。一方、民間で配列決定がなされているが公表されていないものもかなりあるだろうし、ゲノム解読が進行中のバクテリアはおそらく100種類を超えるだろう。また、単細胞ではあるがバクテリアよりもヒトにずっと近いパン酵母(いわゆる「イースト」)は16本の染色体をもち、それらすべての塩基配列が日本の研究者を含む国際協力によって1996年に決定されている。これらの公開データに興味のある方は、たとえば国立遺伝学研究所のGenome Information Broker (<http://mol.genes.nig.ac.jp/gib/>) をご覧いただきたい。

ここまでは単細胞生物の話だが、多細胞生物で初めて、線虫 *C. エレガンス* (腸内寄生虫であるギョウチュウや回虫の仲間、体長1 mm程度の小さい虫)の全ゲノムが1998年に欧米の研究グループによって決定されている。さらに1999年、アメリカのセラ・ジェノミクス社と公的研究機関は



酵母



線虫 *C. エレガンス*

キイロショウジョウバエ。このほかにも、多数ある。おもしろいものがある。フグの仲間で選ばれたのである。日本で取られたものが用いられているゼブラ魚を見ると、ゲノムサイズに進み、2000年末まで一つであるナズナの仲間で、遺伝学の研究にもサイズはもっと大きい研究所を中心に行われてゲノムの全配列決定を1かもしたが、その後、には着手していないように列も急ピッチで読み進めよう。

ヒトゲノム

それでは、人間のゲノムを構成する細胞は非常に多細胞に起源をもち、それら。人間の細胞は通分裂を経て生じる精子のような1セットが「

人間のゲノムは塩基を解読しようという1セットのゲノムは22対



ゼブラフィッシュ

キイロショウジョウバエのほとんどの遺伝子を含むゲノム配列を決定した。

このほかにも、多数の多細胞生物において、ゲノム計画が進められている。おもしろいものとして、イギリスを中心とするトラフグのゲノム計画がある。フグの仲間のゲノムは比較的小さいので、脊椎動物のモデルとして選ばれたのである。ちなみに、DNAを抽出する材料であるトラフグは、日本で取られたものが使われている。また魚類では、発生学の材料として用いられているゼブラフィッシュのゲノム配列決定も進んでいる。植物を見ると、ゲノムサイズが比較的小さいシロイヌナズナのゲノム計画が急速に進み、2000年末までに全塩基配列が決定された。この植物は春の七草の一つであるナズナの仲間だが、10日間で種が親になってまた種をつくるので、遺伝学の研究にも適している。一方、経済的な重要性から、ゲノムサイズはもっと大きい、イネゲノムの配列決定が、日本の農林水産省の研究所を中心に行われている。昨年、先ほども登場したセレーラ社はイネゲノムの全配列決定を1年足らずで終わらせるという計画を発表して物議をかもしたが、その後、彼らはヒトゲノムに集中したため、まだイネゲノムには着手していないようである。なお、マウス(はつかねずみ)のゲノム配列も急ピッチで読み進められており、2001年中には大部分が決定されるだろう。

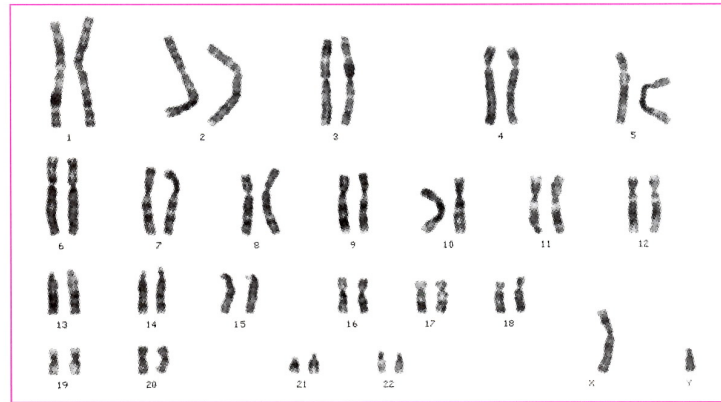


シロイヌナズナ

ヒトゲノム計画の提唱から現在までの小史

それでは、人間のゲノムはどうなっているのだろうか。一人の人間を構成する細胞は非常に数が多いが、それらはすべてただ一つの受精卵という細胞に起源をもち、その全遺伝子は精子と卵に由来する2セットで構成される。人間の細胞は通常2セットの遺伝子をもっている2倍体だが、減数分裂を経て生じる精子と卵は1セットしかもっていない1倍体である。このような1セットが「ゲノム」に対応する。

人間のゲノムは塩基数にして約30億個あるといわれており、これら全部を解読しようというのが、「ヒトゲノム計画」である。ヒトでは、これら2セットのゲノムは22対の常染色体と一对の性染色体にわかれて収められて



ヒトの23対の染色体

いる。ヒトの染色体数を46本ということが多いが、これは常染色体44本(22×2)と性染色体2本(女性はXX, 男性はXY)を合計した数である。各染色体は1本のきわめて長いDNAのひもでなりたっており、その全長は、1個の細胞にあるDNAだけでも1mを超える。細胞内ではこれらの長いDNAが次つぎとくるまって、コンパクトな構造になっている。

1970年代になって、塩基配列の簡便な二種類の決定法がウォルター・ギルバートらとフレデリック・サンガーらによって開発された。しかし、現在大部分の場合に用いられているのは、二つのうちでもより簡便なサンガーらの方法である。またおもしろいのが、サンガー自身、ゲノム構造の決定に関係してきたということである。彼のグループは、まず ϕ X174というファージ(バクテリアにとりつくウイルス)のゲノム配列5375塩基対(base pair, 以下bpと略)を1977年に発表した後、4年後にはその3倍以上の大きさのヒトミトコンドリアゲノム16,500bpを決定している。現在、イギリス、ケンブリッジの近郊にあるヒンクストン・コートに、彼の名前を記念して「サンガー・センター」という研究所があるが、ここは線虫C.エレガンスの全ゲノム配列決定に中心的な働きをした後、現在はヒトゲノムの配列決定の主要センターの一つである。

このようないびつななかで、ヒトゲノムの全塩基配列を決定するという計画が複数の人びとから提唱されるのは、時間の問題だったといえよう。私は1980年代前半にアメリカに留学していたが、当時のヒトゲノム計画への熱狂をにがにがしく思っていた人類遺伝学者が多数いた。その理由は、自分たちの研究費が削られるかもしれないという現実的・政治的観点からである。

その後、“がらくたDNA”を多数含むヒトゲノムの塩基配列を片っ端から決定するのは時間と研究費の浪費だという観点から、cDNA(mRNAに相補

的に対応するDNA)の (expressed sequence t んどん決定するという かし、ゲノムを構造と の塩基配列が解読さお 言えないのである。

ヒトゲノム計画が立 は、古典的な遺伝学の 図(遺伝子と遺伝子が るのかを示したもの て、少しずつ細かくし 一つずつ塩基配列を決 部のゲノム配列を決め た。

ところが、逆転の発 うに多数決定して(鏡 かし、その名もショット てつないでゆくとい いたのでは大変なの (bacterial artificial cl ここでも、これらBA 再びつなげるだけで つまり、従来の発 のに對し、ここでは、 は、遺伝子の本体で てデジタルに記述で い方法を武器にして 究費を用いる民間の く、ありとあらゆる

米・欧・日

現在、ヒトゲノム ヨーロッパではイギ 本であり、主要なゲ 科学総合研究センタ ター(榎佳之教授)、

的に対応するDNA)の配列決定という計画も立ちあがった。とくにEST (expressed sequence tag)と呼ばれる、cDNAの部分配列(200~300bp)をど んどん決定するという計画も、広い意味でヒトゲノム計画に含まれる。し かし、ゲノムを構造ととらえると、最終的には染色体の上のひとつながりの塩基配列が解読されなければ、本来の意味でゲノムを明らかにしたとは 言えないのである。

ヒトゲノム計画が立ちあがった1980年代後半から1990年代前半にかけて は、古典的な遺伝学の考え方を用いて塩基配列を決める前に、まず連鎖地 図(遺伝子と遺伝子が同一の染色体の上でどのような順序・間隔で並んでい るのかを示したものを)、いろいろな制限酵素でDNAを切断するなどし て、少しずつ細かくして調べてゆき、そうしてつなげたDNA断片について 一つずつ塩基配列を決定してゆく、という戦略がとられた。この場合、全 部のゲノム配列を決定するには、30年くらいかかるだろうと推定されてい た。

ところが、逆転の発想が生まれた。まずDNAの塩基配列をめぐらめっば うに多数決定して(銃の所持がめずらしくないアメリカで開発されたせい か、その名もショットガン法と呼ばれる)、それらをコンピュータのソフト でつないでゆくというものである。もちろんヒトゲノム全体を相手にして いたのでは大変なので、最初に100kbp(1kbpは1000bp)ほどのBAC (bacterial artificial chromosome)というまとまりに分断しておく。ただし ここでも、これらBACのあいだの関係は問わないのである。後でそれらを 再びつなげるだけでよいのだ。

つまり、従来の発想がトップダウンでだんだん細かくしていこうとした のに対し、ここでは、ボトムアップでゲノム配列を決めるのである。これ は、遺伝子の本体であるDNAがA、C、G、Tという四種類の文字列とし てデジタルに記述できるという特性があるからできるのである。この新し い方法を武器にして、今や政府の研究費を用いる公的研究機関と私的な研 究費を用いる民間の研究機関がしのぎを削って、ヒトゲノム計画だけでな く、ありとあらゆる生物のゲノム計画を進めている。

米・欧・日を中心とするヒトゲノム計画の現状

現在、ヒトゲノム計画はアメリカとイギリスを中心に行われている。 ヨーロッパではイギリス以外の寄与は低いようである。それに続くのが日 本であり、主要なゲノム配列決定センターとしては、理化学研究所ゲノム 科学総合研究センターおよび東京大学医科学研究所ヒトゲノム解析セン ター(榎佳之教授)、慶応大学医学部分子生物学研究室(清水信義教授)、東

表1 国内外のおもなヒトゲノムプロジェクトに関するウェブサイト

国内	
◆ DDBJ [国立遺伝学研究所生命情報研究センター] (http://www.ddbj.nig.ac.jp/)	
◆ 東京大学医科学研究所ヒトゲノム解析センター (http://www.hgc.ims.u-tokyo.ac.jp/)	
◆ 理化学研究所ゲノム科学総合研究センター (http://www.gsc.riken.go.jp/)	
◆ BODYMAP [大阪大学 細胞生体工学センター] (http://bodymap.ims.u-tokyo.ac.jp/)	
◆ HUGE [かずさDNA研究所] (http://www.kazusa.or.jp/huge/)	
◆ GENOTK [大塚GEN研究所] (http://genotk.genome.ad.jp/)	
◆ ヘリックス研究所 (http://www.hri.co.jp/)	
◆ ゲノムネット [京都大学化学研究所] (http://www.genome.ad.jp/)	
国外	
◆ GenBank/NCBI (http://www.ncbi.nlm.nih.gov/)	
◆ EBI/EMBL (http://www.ebi.ac.uk/)	
◆ SWISS-PROT (http://www.expasy.ch/sprot/sprot-top.html)	
◆ PIR (http://www-nbrf.georgetown.edu/pir/)	
◆ Sanger Centre (http://www.sanger.ac.uk/)	
◆ WashU Genome Sequencing Center (http://genome.wustl.edu/gsc/)	
◆ MIT Center for Genome Research (http://www-genome.wi.mit.edu/)	
◆ Celera Genomics (http://www.celera.com/)	
◆ Incyte Pharmaceuticals (http://www.incyte.com/)	

海大学医学部分子生命科学(猪子英俊教授)がある。一方、cDNA(短いESTおよび全長)については、日本でも大阪大学細胞生体工学センターのBodyMap(大久保公策教授)をはじめとして、さまざまな計画が進行中である。これら国内外の主要なゲノム関連計画のウェブサイトについては、表1を参照されたい。

ヒトゲノムに限らず、すべての生物の塩基配列は、最終的に「DDBJ/EMBL/GenBank国際塩基配列データベース」に登録されることが推奨されている。このデータベースは、日・欧・米の3データバンク(DDBJ, EMBL, GenBank)が国際協調を通して共同で構築しているものである。日本の塩基配列データバンクであるDDBJ(DNA Data Bank of Japan; 日本DNAデータバンク; 図1)は、静岡県三島市にある国立遺伝学研究所の生命情報研究センターで運営されており、筆者もかかわっている。2000年11



月現在で、ヒトゲノムデータベースなどで公開されているところでも登場したセクシーなものを合わせて90%以上だが、それらの配列に決定しているアメリカの巨額の契約金を支払うの塩基配列閲覧を

現在、ヒトゲノムになっているのは、公で研究している機関得られた成果を広くの義務はない。この立に行われると、前密裏に自分たちのデータさせることができる

ショウジョウバエ交換して共同で完成ように経済波及効果共同作業が可能かと世界だが、情報は秘まるとしたら、不幸でほしいものだ。ヒ



図1 日本DNAデータバンクのウェブサイトのトップページ
(<http://www.ddbj.nig.ac.jp/Welcome-j.html>)

月現在で、ヒトゲノム配列全体の90%以上はすでにこの国際塩基配列データベースなどで公開されている。ところが、ショウジョウバエゲノムのところでも登場したセラ社は、すでに彼らだけで80%を決定し、公開されたものを含めて90%以上がすでに決定されたと2000年の1月に発表した。ただし、それらの配列データは公表されていない。また、cDNAを大量に決定しているアメリカのインサイト・ファーマシューティカルズ社は、巨額の契約金を支払った少数の製薬会社を中心とする民間企業にのみ、彼らの塩基配列閲覧を許可しているとのことである。

現在、ヒトゲノム計画を最大の焦点として多くのゲノム計画で問題となっているのは、公的資金を得て研究している機関と民間の私的資金のみで研究している機関との確執である。公的資金を得た場合には、基本的に得られた成果を広く社会に公表する義務がある。しかし私企業の場合、その義務はない。このため、同じ生物のゲノム計画が公的機関と私企業で独立に行われると、前者が塩基配列データを公開するのに対して、後者は秘密裏に自分たちのデータと公開データの双方を利用してゲノム配列を完成させることができる。

ショウジョウバエゲノムの場合は、これら公的機関と私企業がデータを交換して共同で完成をめざすということが行われているが、ヒトゲノムのように経済波及効果が潜在的に大きいと考えられる場合には、そのような共同作業が可能かどうかは不明である。このあたりは科学ではなく政治の世界だが、情報は秘密にして握ったほうが勝ちになるという図式があてはまるとしたら、不幸なことである。科学上の発見はなんでもオープンにしてほしいものだ。ヒトであれハエであれバクテリアであれ、すべての生物

は数十億年の「進化」のたまものである。これらの塩基配列は、人類文明の共通財産として、ぜひとも公開してほしい。

このような「公」と「私」の対立とは少し観点が異なるが、共通の根をもつものとして特許の問題がある。ある生物の塩基配列を決定するという事は、一種の「発見」であり、発明に関する特許がなぜ絡んでくるのか不思議だが、明らかにされた塩基配列の情報をもとに新しい薬品が開発されると、どうやら特許問題がでてくるらしい。どの時点で特許を認めるかまだはっきりしない点があり、10年近く前から論争が続いている。

また、遺伝性の病気をより多く発見しようとするため、最近SNP (single nucleotide polymorphism; 一塩基多型)の研究が、ゲノム多様性という新しいかたちでのゲノム計画として進められている。個人ごとのDNAの遺伝的な違いなどを発見することで、遺伝病の予防や治療につながる可能性があるため、現在、世界中でSNP発見競争が繰り広げられようとしている。ここでも特許の問題がいずれでてくるだろう。「特許」というシステムを資本主義社会での必要悪と考えれば仕方がないことなのかもしれないが、私自身はこのあたりの経済的仕組みがよくわからないので、今ひとつ釈然としない。

ポストゲノムもゲノム配列決定から

われわれ生物学の研究者にとって、ゲノム計画は現代生物学の大きなうねりであるが、当然ながらその後の動きとして、ひっくるめて「ポスト・ヒトゲノム計画」と呼ばれるさまざまな計画が進められている。主として遺伝

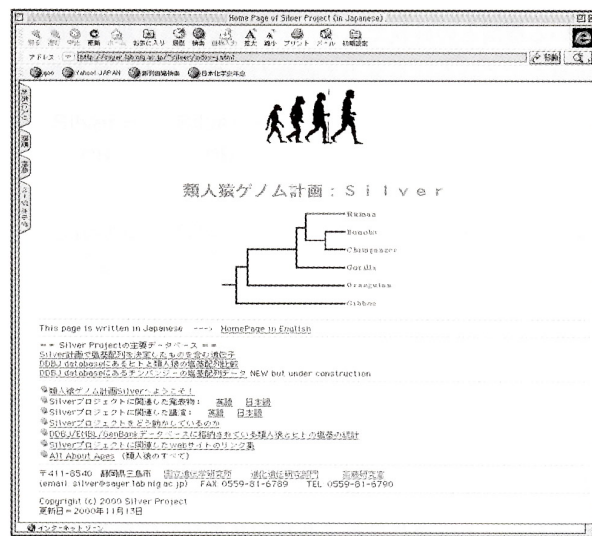


図2 類人猿ゲノム計画Silverのウェブサイトのトップページ (<http://sayer.lab.nig.ac.jp/~silver/index-j.html>)



子の発見パターンを材料から決めるといった、垣根を離れてゆこうとしている主旨のものである。発見されたときに、その発見を比較するという、進化の過程で生じたな塩基配列レベルで発見からだ。

この観点から、筆者らの計画の一つとして、「類人猿立ちあげている (http://sayer.lab.nig.ac.jp/~silver/index-j.html) が進化の過程でどの生物である類人猿(チンパンジー)と比較する必要がある。ゲノム配列を決定し、その比較を行うのである。ゲノム配列を決定するように直結している。21世紀の生物学は、ゲノム解析が主流とな



チンパンジー

子の発現パターンを枚挙的に調べたり、タンパク質の立体構造を片っ端から決めるといった、塩基配列データを基礎にしながら、なるべくそこから離れてゆこうとしているが、結局はこれまでの手法を大規模に展開するという主旨のものである。この意味で、ある一つの生物のゲノム配列が決定されたときに、その近縁種についても同様にゲノム配列を決定し、それらと比較するという、進化的観点から見たゲノム生物学も重要になってくる。結局のところ、遺伝子の発現パターンが生物によって異なるのは、進化の過程で生じたなんらかの突然変異に依存しており、それら突然変異を塩基配列レベルで発見できれば、遺伝子の機能により近づくことができるからだ。

この観点から、筆者の研究室では、1999年から近縁生物の比較ゲノム計画の一つとして、「類人猿ゲノム計画」を開始した。すでにホームページも立ちあげている (<http://sayer.lab.nig.ac.jp/~silver/>; 図2)。人間の独自性が進化の過程でどのように生じてきたのかを明らかにするには、近縁な生物である類人猿(チンパンジー、ゴリラ、オランウータン、テナガザル)と比較する必要がある。ヒトゲノムの塩基配列を参考にして多数のPCRプライマーを作成し、それらを用いて類人猿ゲノムのDNAを増幅し、塩基配列を決定する。こうして得られた塩基配列を多重整列し、分子進化学的解析を行うのである。ゲノム計画は、実験とコンピュータを用いた解析がこのように直結しているという意味で、これまでの生物学とは異なっている。21世紀の生物学は、ゲノム計画の大規模性を模範として、どの分野でも大規模解析が主流となってゆくだろう。

写真提供：ショウジョウバエ：林 茂生、大腸菌：西村昭子、酵母：荒木弘之、上村陽一郎、線虫C.エレガンス：石原健、ゼブラフィッシュ：武田洋幸(以上、国立遺伝学研究所)、シロイヌナズナ：坂本 亘(岡山大学資源生物科学研究所)、ヒト染色体：沼部博直(東京医科大学病院)。