

## 生物におけるゲノム研究—その現状と明日への展望

農業生物資源研究所 佐々木卓治

はじめに：

地球上に存在するあらゆる生物は遺伝情報を DNA に書き込んで、ゲノムとよばれるひとまとめのセットにして子々孫々伝えている。DNA の、情報の運び屋としての起源には諸説あるが、実に巧妙な仕組みが成り立ったものである。この仕組みの解明は、人類が自分たちの起源を宗教ではなく科学として理解しようと考え始めた時から、長い間諸説が現われ、消えていった。その中で有名なのは前成説と後成説であり、現在の我々の知識からみれば正しく理解されていない部分があるのは当然としても、それぞれの説には直感的に真理をつかんでいた部分もあった。例えば前成説では、「入れ子」の存在を仮定しているが、これを遺伝暗号を背負った DNA と見れば、前成説もあながち間違っていたとはいえない。しかし、「入れ子」の隠喩が DNA であることが明らかされるまでには、それぞれの時代に主流であった学説との戦いと、それぞれの時代が蓄えた、底辺を支える科学知識の総結集によるブレークスルーが必要なのであった。メンデルによる遺伝の基本法則の発見も、当時修道院が社会でどんな役割を果たし、彼が居住したブルノという地域が農工業の中心地であり育種素材に関する豊富な知識の集積が社会全体で共有できたことなどを知れば、彼が大発見を成し得たのもなるほどと思えるのである。二重らせん構造の発見ではご本人が著わした本でも明らかなように、シャルガフによる化学分析、フランクリンによる結晶回折写真など、果実は2人による収穫を待っていたのである。

現代の多様な生物を対象にしたゲノム解析は、この遺伝学の流れの中から生まれた新たな支流であり、遺伝情報とそれが具現するものがどのようにしてオンラインになっているのかを知る研究である。では、このラインを理解するための環境は整っているのだろうか。メンデルやワトソンとクリックが成し遂げた、時代を代表するブレークスルーはもうすぐそこまで近づいているのだろうか。本稿ではゲノム時代の生物学の展開を解説するとともに、そこから遺伝で支配される生物の仕組みの何がわかるのか、わかることが期待されるのか、述べてみたい。

### ゲノム解析とは？

ゲノムが生物の前遺伝情報の一揃いだとすると、その解析とは何を指すのか？一般の方々にはゲノム解析とは即ちゲノム塩基配列を多くの自動解析装置を用いて行うことだと思われるようである。この推察の一部は正しく、一部は間違っている。間違いをもたらす原因はゲノム解析を行っている側にある。それは成果を発表する際に塩基配列の解析であれば解析した塩基数をいえば、その大きさと他と較べるのが容易だからであり、マスメディアもこれを歓迎する。しかし、ゲノム解析の正しい意味は塩基配列情報を拠り所とした生物の遺伝現象の分子レベルでの解析であり、塩基配列の解析はそのほんの一部なのである。ただ、この「ほんの一部」が解析対象を真核生物とする場合、たいへん巨額の資金を必要とするために、目立ってしまうのである。とはいえ、全ゲノム塩基配列情報は従来の遺伝学あるいは生物学における研究手法や生物そのものの見方を大幅に変え、生命の仕組

みの理解にメンデルやワトソン・クリックの発見が与えたと同等の影響を及ぼすもの、といっても過言ではない。ただし、情報は正しくないと間違った結論を与えてしまう。特にヒト医療につながる場合には深刻な結果と多大の浪費を招くことになる。この点が最近の「早さが一番、正確さは二の次」という、科学者としては受入れられない、投資効率論が優位となっている政府の考え方が10年、20年後に国民にどのような成果を遺すのか、疑念を抱かざるを得ない。

#### ゲノム解析の足跡

メンデルがエンドウマメを用いて7個の表現形質について顕性・隠性に従って分離して遺伝する法則を発見したときに、彼は遺伝を支配する物質の存在も同時に予言していた。しかし、われわれがそれが何であるのかを知るには長い年月が必要であった。ヒトの遺伝病は米国 Johns Hopkins 大学が管理している著名なデータベース Online Mendelian Inheritance in Man(OMIM)によると、現在マップされているだけで8309座ある。しかし、遺伝病の遺伝子座が DNA マーカーによってその存在を明確にされるようになったのは、1983年に J.Gusella らによって Huntington disease 遺伝子と密接に連鎖する DNA の配列多型が発見されて以来である。塩基配列上の多型が表現形質の多型と同様にメンデルの法則に従って遺伝し、メンデルが気づかなかった連鎖現象を利用することで再現性のよい遺伝地図が作成できることがわかった時点で Huntington disease 遺伝子自身がどのようなものか明らかにされることが約束されたのである。ヒトゲノム解析においてはその後、全塩基配列解析に向けて米国と英国を中心に巨額の公的資金の投入が開始され、自動塩基配列解析装置の開発と相俟って、巨大な組織となった解読チームが世界の各所に出現した。この動きは解読戦略の改善を招き、大容量の遺伝情報を処理するコンピューター解析技術の発展とともに民間会社に利益をもたらす産業としてのゲノム解析が認知されたのである。と同時に科学なのかビジネスなのか研究者に選択を迫る分野ともなっている。

これまでにゲノム全塩基配列が解読されたとして論文発表された生物種で最も多いのは微生物である。自動配列解析装置とコンピューターの発展と歩調を合わせて開発された「全ゲノムショットガン法」により1995年に H.influenza の配列解読(1.83Mb)が報告され、その後この程度のゲノムサイズの多くの微生物で試された結果、よほど繰り返し構造に富んでいない限り本法が短期間で正確に配列情報が得られることが示されている。現在公的データベースで情報が公開されただけでも104種類の微生物の全塩基配列が解読されている。民間等では非公開ながら世界中で相当数の微生物種が解析されていると推察されている。

有核生物の場合、正確なゲノム上の位置を確認した DNA 断片を対象にショットガン法で配列を決め、DNA 断片毎の配列を重ね合わせて染色体毎の配列を決めるのが通常の方法である。この方法で決定された生物種として、出芽酵母、分裂酵母、線虫、マラリア原虫、シロイヌナズナ、イネおよびヒトがあげられる。一方、微生物で成功を収めた全ゲノムショットガン法を有核生物、特にヒトゲノム解析に利用する戦略のもとで、まずショウジョウバエが試金石に選ばれた。この理由はショウジョウバエが遺伝解析の進んだ生物であり、また、部分的に BAC などの DNA 断片によるゲノム解析が行われており、全ゲノムショットガン法のみでは不足するデータをこれらから得ることを期待したからであっ

た。この生物には特有の繰り返し配列が多数存在するために、2000年に配列解読結果が発表された時点では、一応終えたとはいえ、まだ隙間も多く、一体どの程度本当のショウジョウバエゲノムの塩基配列を再現しているか疑問が投げかけられた。同様の疑問はこの方法が民間会社によりヒトゲノムの解読に用いられた際にも生じている。とはいえ、この方法は多少のいい加減さは残しても、手っ取り早くデータが得られることから、最近では流行の方法となっている。マラリア蚊、フグ、ホヤなど最近報じられた解読例はこの方法による。

いずれにしても、我々が研究者として心に留めなくてはいけないことは、塩基配列情報を使ってどのような研究を行い、その時の情報の精度が研究成果にどのような影響を与えるのかを見極めることである。DNA断片をきちんと並べてそれらを配列解読する方法であっても、全ゲノムショットガン法であっても、配列を得る過程でどのような工夫がなされ、配列のゲノム上の位置情報を確認しているかを検証することが大切になる。イネの場合、国際コンソーシアムによる従来からの丁寧な方法、シンジェンタ社による一部地図情報を取り込んだ全ゲノムショットガン法、中国北京グループによるコンピュータープログラム処理技術に依存した全くの全ゲノムショットガン法の3様のデータが出揃っているので、相互に比較が可能である。

#### イネのゲノム解析の進展

上述したDNAの多型を遺伝解析に利用する考え方は直ちにヒト以外の生物にも広がり、イネにおいても最初の地図が報告されたのは1988年であった。イネにおいては異数染色体をもつイネを利用して12本の染色体と連鎖群の番号の統一が行われたことは、その後のゲノム解析にとって幸いであった。イネはわが国において明治以来、最重要穀物として政府管掌による育種と品種管理および販売管理が徹底された歴史があり、遺伝学的な基礎研究は大学が中心であった。連鎖地図に関しては、自然あるいは放射線による突然変異形質間の連鎖関係を利用して現在までに約200個の遺伝子座が12本の染色体上にマップされている。これらは各形質のゲノム中での存在箇所を知るには欠かせないが、ゲノム全体の解析には適当ではない。1991年に農林水産省がイネゲノム解析を開始した時点で、まず目標としたのはDNA多型利用による汎用性のある連鎖地図であったのは当然の選択といえる。この分子マーカーによる遺伝地図は、その後今日に至るまで、イネゲノム解析研究の基盤となっている。ゲノム解析の対象となった生物でこういった分子マーカーが乏しい生物では、仮にゲノム塩基配列情報が得られても、遺伝子機能解析においてその利用場面は限られてくる。尤も、その配列情報を利用して、分子マーカーを作出する努力を行うのであればこの限りではない。

わが国ではイネ遺伝地図はイネ品種「日本晴」と「カサラス」のF2集団186個体に対して主にイネの発現遺伝子(cDNA)をRFLP検出プローブに用いて作成され、1994年と1998年に論文として報告された。その後もマッピングは継続され、結果はインターネットで公開されている。マーカー数は3000を越えている。現在ではRFLPマーカーからPCR利用可能で安価なマーカーへの転換が図られている。イネではこの他に、最近ゲノム中の繰り返し配列に見られる繰り返し単位の多型に着目したSSRマーカーを2200個マップした地図が公開された。SSRマーカーは特に種間での多型頻度が

高いので種子遺伝資源の評価に有効である。また、遺伝解析によるマッピングではないものの、上記 RFLP マーカーをアンカーにして作成した DNA 断片 (YAC、酵母人工染色体にクローニングしたもの) の整列化地図上に、6500個のイネ cDNA を PCR 増幅の有無によりマップした地図も作成されており、イネにおいては他の生物にはみられないほどマーカーが充実している。これらは表現形質の DNA マーカーによる標識付けに有効に利用され、DNA マーカーを目印にした選抜育種あるいは地図を利用した遺伝子単離はその代表的利用場面である。

遺伝子の単離に欠かせないのは突然変異体である。突然変異はゲノム塩基配列上の変異であり、この変異が表現形の変異として観察者に認知されれば突然変異体として登録される。変異は自然に起こることが知られているが、それが表現形に変異をもたらす確率は低く、多数の遺伝子の機能を網羅的に知るには塩基配列中に人工的に変異を導入して、その変異と表現形の相関を利用することが行われる。この考え方は広く実験生物に適用され、生物的、化学的、あるいは物理的な因子・手段が用いられる。化学的には塩基に変異をもたらす薬剤、例えばニトロソグアニジンやエチルメタンスルホン酸などが、物理的にはエックス線やガンマ線などの放射線がよく用いられる。引き起こされる変異にはそれぞれ特徴があり、薬剤は塩基置換を、放射線は長い領域での塩基の欠落を引き起こす傾向が知られている。変異が起きたゲノム上の場所の特定は、PCR や最近ではマイクロアレーを利用して以前よりは迅速に行える例も示されてはいるが、容易ではない。これらに比べ、生物的方法は利用する因子自身が DNA であり、対象生物のゲノム塩基配列情報が入手できる状況であれば挿入による変異箇所を特定することは比較的容易である。植物においては、歴史的には、まず、それまでの常識を覆したトウモロコシにおける転移因子 (トランスポゾン) の利用が、続いて植物に腫瘍を形成するアグロバクテリウムによる DNA 断片 (T-DNA) の導入が動物とは異なる変異導入の方法として開発されている。異種生物から遺伝子を導入した場合、成果物は遺伝子組換え生物となり、環境保護のために栽培や保存には法的規制がある。シロイヌナズナのように小さな植物の場合は隔離栽培が容易であるが、イネのような大型の穀類の場合は大規模な施設が必要となり、遺伝学的研究は困難になる。しかし、イネに固有のトランスポゾン、Tos17 の発見がこの状況を変え、この因子による変異体作成が大規模に行われている。Tos17 は組織培養後の再生体で元の存在箇所から新たな別の箇所に転移することで変異を誘起する。従ってこの変異イネは組換え体ではないために、通常の圃場で栽培が可能である。この変異体は表現形質の変異から遺伝子の変異箇所を明らかにする遺伝学的手段の対象になると同時に、最初に Tos17 が挿入された塩基配列を探し、変異遺伝子を突き止めた後に表現形質の変異との相関を知る、逆遺伝学的手段の対象としても利用でき、わが国で開発された、今後の遺伝子機能解析の重要な資源である。

#### 遺伝学の温故知新

メンデル以降、遺伝学は理論的考察の対象としても進展した。その中には、多数の遺伝子の相互作用による効果、いわゆる量的形質 (QTL) の解析がある。QTL の存在箇所はインターバルマッピング法で求められるが、確率論に基づくために解析対象試料数が限られることもあり、幅広い領域に分散してしまう。このことから従来は QTL に対応する

遺伝子の実在すら正しいのかどうか確かめる手段がなかった。しかし、ゲノム解析が進み、多数の DNA マーカーや塩基配列情報が入手可能になったことにより、量的形質のように複数の遺伝子座が関与する表現形質に関与する遺伝子の実体が解析可能となった。ヒトの場合、糖尿病や高血圧などの生活習慣病はほとんどが QTL（複雑形質）であり、従来は確実な対応遺伝子は見つかっていなかったが、最近のゲノム解析情報を多数の検体検索に用いることで、例えば高血圧関連の遺伝子が確定されている。ただし、ヒトの場合は実験的確認が不可能なので、マウスやラットなどの実験動物での確認が必要である。また、QTL 関連の遺伝子が同定されるに従って、類似の表現形質に関して変異が 1 遺伝子の変異でもたらされる場合と QTL で決まる場合の分子レベルでの関係も徐々に理解できるようになった。これはイネとシロイヌナズナの開花に関する日長感受性の研究など、植物を用いた解析によるところが大きい。

イネの開花期（出穂期）はイネ栽培上、最も重要な形質のひとつである。わが国は南北に長い国であり、日長も気温も幅広い分布があり、出穂期のコントロールは育種上大切な課題である。出穂期は多数の遺伝子により決まる QTL（複雑形質）であり、例えば「日本晴」と「カサラス」の組合せだけでも 10ヶ所以上の QTL が認められている。これらの遺伝子単離を目指して、各 QTL を「日本晴」を背景とした準同一遺伝子系統を作成し、1 遺伝子としての単離戦略がとられた。その結果、これまでに 3 種類の QTL に対応する遺伝子が単離され、シロイヌナズナにおいて開花期に関連する 1 遺伝子の変異として単離された、いくつかの遺伝子と比較された。その結果、イネで単離された 3 種の遺伝子はシロイヌナズナで開花期変異遺伝子として単離された遺伝子の中に全て含まれていた。ただし、イネが短日植物、シロイヌナズナが長日植物であることから、これらが日長に関して異なる反応を示すためには、まだ未同定の遺伝子座の中に時計に相当する遺伝子が存在するに違いない。この例で示されたように、QTL とはいえ、それらに含まれる遺伝子の変異は対応する 1 遺伝子の変異として検出されることがわかってきた。しかし、逆にいうと、1 遺伝子の変異情報のみからだけでは、それらの間の相互作用は理解できず、QTL 解析との併用が必要になる。実際、シロイヌナズナにおいては、1 遺伝子変異による日長感受性遺伝子検出の限界からイネと同様に QTL 解析も行われ始めている。QTL 解析と 1 遺伝子変異解析の接点を求める試みは日長感受性以外にはイネの草丈、いもち病抵抗性、あるいはトマトの果実重量などにおいて行われつつある。

QTL 遺伝子の解析結果からわかってきたことは、アリの多様性が何によってもたらされているのか、ということである。1 遺伝子の変異が表現形質に劇的な変異をもたらすのはその遺伝子のいわば心臓部分に変異が生じた場合である。しかし、同じ表現形質に対して「ゆるやかな」変異をもたらすのは心臓部分以外の変異なのである。こうでなければ農業には利用できない。わたしたちが実際の圃場で目にするのは相互に比較して初めてわかる違いであり、この違いをもたらすアリは対応する遺伝子の心臓部分には存在しない。それは、心臓部分を制御する部分であったり、同一複雑形質に関わる QTL の産物と相互作用する部分であったりするのである。これらの部分の変異はイネのような栽培植物の場合、進化と選択の結果、各栽培品種あるいはその起源である近縁野生種において、一般的に起きており、この変異の性格や効果を研究することが今後の育種につながっていく重要な課題と考えられる。

## 農業生物のゲノム解析

イネ以外の農業生物におけるゲノム研究の現状はどうであろうか。1993年より米国 USDA が中心になって毎年開催している「国際植物・動物ゲノム会議」で2003年にポスター発表等でとりあげられた作物は35種、動物は6種にのぼる。これらのゲノム解析としてのレベルには当然いろんな段階がある。作物の中ではトウモロコシとコムギがイネと並んで最もゲノム解析が進んでいる植物である。この理由は、この2種がイネと並んで世界3大穀類であることと同時に、イネゲノム解析結果がこの2種の穀類ゲノム解析に利用できることから、イネでの進展がこの2穀類の研究を刺激したことにある。トウモロコシは米国において食料というよりもむしろ飼料作物および澱粉産業の基盤として広範囲な産業を支えている穀類として、我々が考えているよりもはるかに重要な作物なのである。米国では民間企業と公的研究機関が協力と競争の関係を保ちつつ、多数の変異体資源が保持され、DNA多型に基づく精密な遺伝地図が作成されている。トランスポゾンを利用した人工変異体の作成も行われ、遺伝子単離に利用されている。2002年から2年間の試験期間でゲノム塩基配列解読の可能性をいくつかの方法で行っており、今後いずれかの方法でトウモロコシゲノム塩基配列解析が大規模に開始されるものと思われる。

一方、コムギについては最も利用規模の大きいパンコムギが6倍体であり、ゲノムサイズがイネの40倍もあるために、まず、国際協力のもとにESTの蓄積を協力して行っているのが現状である。また、DNA多型を利用した遺伝地図も作成されているが、コムギの場合、DNAの多型頻度がイネやトウモロコシに比べて低いこともあって、遺伝的マッピングのみでは十分な密度が得られない。しかし、コムギでは染色体欠失系統が充実しているので、これと併用することで地図の充実が図られている。ゲノム塩基配列解析についてもこの国際協力体制で行う方向で話し合いが開始された。

野菜では、ダイズ、トマト、ナス、レタスなど、その他ではワタ、サトウキビ、オオムギ、ソルガムなど日本ではあまりなじみのない作物がワークショップで取り扱われている。また、果実では柑橘類、バラ科果実が多くの研究者の注目を集めている。ダイズは従来考えられていたよりもゲノム構造が複雑であるために、マメ科植物のモデルとして、ミヤコグサとアルファルファの一種の2種が選ばれ、遺伝地図作成やゲノム塩基配列解析が精力的に行われている。マメ科植物の最大の特徴である根粒菌との共生メカニズムの解明が最大の目的である。

家畜動物ではウシ、ブタ、ニワトリがワークショップで取り扱われている。それぞれ遺伝解析と並んで、ホールゲノムショットガン法によるゲノム塩基配列解析が俎上に上っているようである。ただし、いずれも低いカバー量しかデータを集めないのが、いかにヒトゲノムとのシンテニーを利用するとはいえ、今後の利用場面でどこまで有効か、疑問は残る。また、魚類ゲノムもこの会議ではワークショップが毎回開催されている。

昆虫関係では、遺伝資源やこれまでの知識の蓄積から判断して、カイコが唯一ゲノム解析対象として深い解析が可能だと考えられる。わが国では産業としてのカイコの利用価値は下がる一方であるが、生糸は依然繊維としての優位性は認められており、また医療面での利用も考えられてきているので、良質な生糸生産昆虫としての改良ニーズは続くであろう。また、単一の生糸タンパク質を大量に生合成する機能を利用してヒトの医療に有効

なタンパク質成分を大量に合成させることも考えられている。その他、農業上の深刻な害虫である鱗翅目昆虫のモデルとしての研究対象となり、ゲノム情報を利用して化学農薬に依存しない退治法の開発が期待されている。これまでにゲノム解析が行われた昆虫はショウジョウバエとマラリア蚊であり、いずれもカイコからは遠い。今後のカイコゲノム解析の進展に注目したい。

おわりに：

ゲノム解析は遺伝学を基本とする新たな学問分野である。世間ではゲノム塩基配列解読の競争がマスメディアを通じて喧伝され、どんな品質であれ、最初に「完成」と言った方が勝ち、という判断がされている。この報道結果が科学の正しい発展にどんな影響をもたらすのか、マスメディアは責任を持ってはくれない。また、科学政策担当者もマスメディアの報道に迎合する傾向があり、先端科学の評価の困難さ、あるいは将来への見通しの大切さを感じているのか否か、疑問を抱かざるを得ない場面が見られる。科学は正しい判断の積み重ねによって発展する。これは歴史が教えてくれている事実である。現在行われている多くの生物を対象としたゲノム解析の成果のうち、どれだけの部分が積み石になるのだろうか、あるいは捨て去られる石になるのだろうか。研究者は多くの情報の中から、科学的論理による判断のもとに正しいものを選び出し、自分の成果を付加し、後世に伝えていく責務がある。この努力により、ゲノム解析研究が科学分野の新たな一員として認められ、今後の広い応用場面が開拓できるのである。