出芽酵母 (*Saccharomyces cerevisiae*)

第III 染色体の構造と機能の解析

特定の生物の染色体 DNA の塩基配列の全てを決定しようという試みがヒトをふくめた多くの生物で行われている。酵母は醸造の主役として我々にはなじみ深い微生物であるが、真核生物のモデルとしても多くの基礎的研究がなされている。昨年、真核生物の染色体として はじめて酵母の第 III 染色体の全塩基配列が決定され話題となった。このプロジェクトに日本から参加した筆者に、染色体の全塩基配列の解析の結果について解説していただいた。

田 中 誠 司・磯 野 克 己

1. はじめに

生物は多数の遺伝子をもっており、それらをうまく組み合わせて発現させることにより、変転する環境に適応して生育するとともに多様な生物環境を作り上げ、種分化を重ねて進化を遂げている。したがって、生物は多くの遺伝情報の集積体であるということができ、その全体像を明らかにし、生命の理解を深めるためには、生命のブループリントともいうべきゲノム^{註1}の全構造を解明する必要がある。さまざまな生物でこのような解析が進めば、現在の生物のゲノムがどのように設計されて進化してきたのかを知る上での重要な手掛かりが得られるものと期待される。このような考え方のもとに生物の「ゲノム解析」が遺伝学の新しいアプローチのひとつとして最近注目を集めており、大腸菌からヒトに至るまでの各種の生物を対象として開始されている。本稿ではそのようなゲノム解析の一例として、真核生物の染色体ではじめてその全塩基配列が決定された出芽酵母 (*Saccharomyces cerevisiae*) の第III染色体について、全塩基配列データから得られた知見を含めて、筆者らの研究室で行った解析を中心にまとめた。

註1：本稿ではゲノムという語をその生物のもつ遺伝情報の総体という意味で用いる。

2. 第III染色体

出芽酵母は極端に非対称な分裂である「出芽」によって増殖する単細胞の子囊菌であり、古くから、主としてその産業的な応用価値の高さから多くの研究が行われて

きた。また、近年の分子遺伝学や分子生物学的手法の発展に伴い、細胞周期や細胞増殖あるいは遺伝子発現の調節機構などに関する研究が広く行われており、真核細胞のモデルとしても出芽酵母はますます重要性を増している。出芽酵母のゲノムは大きさが約 14 Mb (14×10^6 塩基対)であり、長さの異なる 16 本の染色体よりなっている¹⁾。第1図に示したようにこれらの染色体はパルスフィールドゲル電気泳動法により、その大きさに従って分離することが可能である²⁾。パルスフィールドゲル電気泳動で下から3番目のバンドとして分離される第III染色体は、次に述べる理由により 16 本の酵母染色体のうちで最も解析が進んでいる染色体である。すなわち、(1) この染色体の大きさが約 300~360 kb でしかない(正確な大きさは株によって異なる³⁾) こと、(2) この染色体上には酵母の接合型を支配する3つの遺伝子座、*MAT*, *HML*, および *HMR* が存在すること、(3) *HML-MAT* 間での相同組換えを利用して、‘small-ring’ とよばれる環状の第III染色体が得られること⁴⁾、(4) この染色体上には詳細に解析された遺伝子 (*LEU 2*, *HIS 4*, *PGK 1* 等) や酵母のトランスポゾンである Ty 因子の転移するホットスポットがあること⁵⁾ などである。

このような背景を踏まえ、筆者らの研究室をはじめ他のいくつかの研究室で、この染色体を対象とした染色体レベルでの構造と機能の解析が行われた。こうした解析を行うにあたっては、染色体をそのまま解析するのは難しい。そこでまず染色体を断片化した上でクローン化し、得られた各クローンについて制限酵素地図を作製することなどにより染色体の端に位置するものから順番に

Analysis of the Structure and Function of Chromosome III of the Yeast, *Saccharomyces cerevisiae*
Seiji TANAKA and Katsumi ISONO (Department of Biology, Faculty of Science, Kobe University)

第1図 OFAGE 法による出芽酵母 *S. cerevisiae* の染色体の分離

DC5 ρ^0 株および YAT ρ^0 株の染色体 DNA を OFAGE (orthogonal-field-alternation gel electrophoresis) 法 (パルスフィールドゲル電気泳動の一種) で分離したもの。左側に染色体番号を示す。

整列させ、従来の遺伝子地図とも対応付けたフェージやプラスミドのクローンセット (整列クローンバンク) を作製し、それを用いて解析を進めるという方法がとられた。このような第Ⅲ染色体の整列クローンバンクの作製とそれを用いた染色体の解析について述べる。

3. 整列クローンバンクの作製とそれを用いた解析

筆者らの研究室では、酵母ゲノムの構造と機能を明らかにすることを目標として、まず大きさの小さな5本の染色体、すなわち第Ⅰ、Ⅲ、Ⅴ、ⅥならびにⅧ染色体について整列クローンバンクの作製を行い、これらの染色体をほぼ完全に網羅する整列クローンバンクを得ている⁶⁻⁹⁾。ここでは、第Ⅲ染色体の整列クローンバンクの作製ならびにそれを用いた解析⁷⁾について述べる。整列クローンバンクの作製にあたっては、小原らによる大腸菌の整列クローンバンク作製に用いられた方法¹⁰⁾と基本的に同じ方法が用いられた。まず、出発材料となる第Ⅲ染色体の断片を含むクローンを得るために、パルスフィールドゲル電気泳動を繰り返して数十 μ g の第Ⅲ染色体 DNA を調製し、得られた DNA から 344 個の λ フェージクローンを作製した。次にこれらのクローンについて、8種類の制限酵素、すなわち、*Bam*H I、*Bgl* I、*Eco*R I、*Hind* III、*Kpn* I、*Pst* I、*Pvu* II および *Xho* I を用いて制限酵素地図作製を行い、得られた地図をコンピュータを用いて比較し、同時にブラックハイブリダイゼーションを高密度かつ高解像度で行えるように改良した packed array hybridization 法でクローンの重複を

調べつつ整列化を進めた。その結果、1つのギャップを残して第Ⅲ染色体の右末端 (テロメア) 部から左末端部付近までのほぼ全領域を網羅する整列クローンバンクが得られた (第2図)。このバンクの物理的な長さは 334.6 kb であり、解析に用いた DC5 ρ^0 株の電気泳動的に測定した第Ⅲ染色体の長さが 340 kb であることからすると、残されたギャップは長くても数 kb であろうと考えられた。

つぎに第Ⅲ染色体の機能的・構造的特徴を明らかにするために、作製した整列バンクから第Ⅲ染色体の全領域を網羅する大小 167 個のプロンプを調製し、酵母の poly(A)⁺RNA に対しノーザンブロットを行うことによって、この染色体上で発現している遺伝子の総数およびその発現レベルを解析した。このような解析は染色体全領域の整列クローンバンクが完成されてはじめて可能になる試みである。それまでにこの染色体の遺伝子地図には tRNA を含む 34 の遺伝子がマップされていた¹¹⁾が、この数は推定される酵母の遺伝子の総数から考えるとはるかに少ないものである。その結果、第Ⅲ染色体にはその全領域にわたって 156 個の遺伝子が存在していることがわかり、それらの正確な大きさのみならず、その存在位置も決定することができた。各転写産物の大きさは 0.35 kb から 6.5 kb と様々であり、1.00 kb から 1.20 kb のものが最も多かった。このことは酵母細胞内で最も数の多い蛋白質が 40 kDa から 50 kDa であることとよく対応している。また、染色体の末端に近い部分で発現している遺伝子は発現量が低く、発現量の高い遺伝子は、染色体右腕、左腕の中ほどにある傾向が見られた。

一方 NEWLON らは、環状の第Ⅲ染色体をもとにして整列クローンバンクを作製した。環状の第Ⅲ染色体はパルスフィールドゲル電気泳動を行わなくても他の染色体から分離できる (NEWLON らの整列クローンバンク作製はパルスフィールドゲル電気泳動が開発される以前から始められていた)。彼らのバンクは YIp5 プラスミドと λ フェージのクローンからなっており、染色体ウォーキングによりクローンの整列化が行われた¹¹⁾。前述したように、彼らが用いた環状染色体は、第Ⅲ染色体左腕の末端付近に存在する *HML* から右腕中ほどに存在する *MAT* までであるので、得られた整列クローンバンクは第Ⅲ染色体の約 60% に相当する 210 kb の領域を含んでいる (第2図)。彼らは、こうして得られたクローンに自律複製因子 (ARS) をもつものがあるかどうかを網羅的に解析した。すなわち、調べたいプラスミドクローンの DNA 断片上に ARS 活性があれば、そのプラスミドは酵母細胞内で複製されるので、ARS 活性のない DNA 断片を持つプラスミドに比べて 100~1,000 倍高い形質

Chromosome III

第2図 第Ⅲ染色体の遺伝子地図と物理地図の対応

左側に第Ⅲ染色体の遺伝子地図（文献1より改変）、中央に吉川らによる物理地図（文献7より改変）、右側に NEWLON らによる物理地図（文献11より改変）を示す。記号は□はテロメアを、○はセントロメアを、●はARSをあらわす。

転換頻度が得られる。したがってプラスミド上の選択マーカーの活性を指標として追跡することにより、比較的容易にARS活性を持つクローンを同定することができるわけである。このようにして彼らは環状の第Ⅲ染色体210kbの領域内に13個のARSを同定した。さらに彼らは、これらのARSが実際の染色体上での複製の過程でどのように機能しているかなどについての解析を進めている¹²⁻¹⁵⁾。

またOLSONらは、出芽酵母全染色体由来のDNA断片を持つ約5,000個のλファージのバンクを作製し、各クローンのDNAを制限酵素EcoRIとHindIIIで完全消化し、電気泳動で分離した後400bp~7.5kbの大

きさの断片を測定し、コンピュータを用いて各クローンのデータを比較することにより整理化させるという方法で、出芽酵母の全ゲノムの整理クローンバンクを作製している¹⁶⁾。第Ⅲ染色体についても、かなりの部分を網羅するクローンのグループができているようである¹⁾が詳細についての報告はない。

4. 全塩基配列の決定

上述の整理クローン（主としてNEWLONらの作製した環状染色体由来のクローン）を用いて、第Ⅲ染色体の全塩基配列がヨーロッパの35個の研究室を中心とする共同作業により決定された¹⁷⁾。これまでに、ウイルスや、

第 3 図 第三染色体の転写産物と ORF の対応付けの例

吉川らによる第三染色体 15.6 kb から 65.6 kb までの領域の転写産物地図 (上側) とそこに対応する OLIVER によって決められた ORF の地図 (下側, 17 kb から 67 kb の領域, EMBL データベースの塩基配列をもとに作製) を示す (文献 18 より転載)。

葉緑体やミトコンドリアなどの細胞内オルガネラのゲノムの全塩基配列を決定した例は存在するが、真核生物の染色体の全塩基配列が決定されたのはこれが最初である。この塩基配列の長さは 315 kb であり、EMBL の塩基配列データベースに SCCHR III (accession no. X 59720) として登録されている。この塩基配列決定では複数の整列クローンバンクを用いざるを得なかったために、結果的に数種類の株の第Ⅲ染色体についてのデータが混在している。しかしこの点については、Ty 因子の数や存在位置に関する多形性を除いて、恐らくあまり問題にはならないと思われる（詳細は文献 18 を参照）。

この第Ⅲ染色体の全塩基配列中には、長さが 100 アミノ酸残基以上の蛋白質をコードする 182 の ORF が存在した（さらにそれらの ORF 内に完全に含まれる 33 の ORF が存在するが、この数には含めていない。しかし、互いに他と重複している ORF は含めている）。現在までに 100 アミノ酸残基以下の長さの蛋白質をコードする酵母の遺伝子のはんの数例しか知られていないので、この基準は一応妥当なものと考えられる。これら 182 の ORF のうち 37 はこれまでに知られている遺伝子と同一であった。残りの 145 の新しく見つかった ORF については、データベースを用いての既知の遺伝子や蛋白質との相同性検索の結果、15 の ORF がデータベースに収録されている *S. cerevisiae* 以外の生物の蛋白質と、また、別の 15 の ORF はデータベースに収録されている *S. cerevisiae* の蛋白質と高い類似性（相同性検索プログラムとして FASTA を用いたときに、そのスコアが 200 以上）を示した。前者のうちには予想だに示なかったような蛋白質と相同性を示した ORF がある。例えば、YCL 17 c という ORF はバクテリアの窒素固定に関与する遺伝子である *nifS* の遺伝子産物と高い相同性を示した。酵母は気体窒素の固定は行わないが、この遺伝子は生育に必須であることが遺伝子破壊実験によって示されている。したがってこの遺伝子が酵母の生育過程でどのような機能を果たしているのかについての今後の解析が待たれる。また、YCL 74 w という ORF はキショウウバエ (*Drosophila melanogaster*) のレトロトランスポゾン *copia* と高い相同性を示した。このことは、これまでに知られている 4 種の Ty 因子以外の新しい型のトランスポゾンが出芽酵母に存在している可能性を示唆している。さらに、これら以外の新たに見つかった ORF のうち、55 の ORF について遺伝子破壊実験が行われた結果、3 ORF は生存に必須であること、必須でない ORF のうち 42 のものについてさらに調べたところ、14 の ORF については破壊した場合には表現型になんらかの差異が見られたが、残りの 28 の ORF

に関しては破壊した場合でも明白な表現型の差異は見られないこと、などが報告されている。蛋白質をコードする ORF 以外には、既に第Ⅲ染色体にマップされていた 4 個の tRNA 遺伝子に加え、新たに 6 個の tRNA 遺伝子が見いだされた。

これらの結果は既に OLIVER らにより報告されている¹⁷⁾が、最近、報告されたデータの細かい修正や、別の方法による相同性検索で得られた結果などを加えてデータベース化されたものが「出芽酵母第Ⅲ染色体データベース」として SLONIMSKI と BROUILLET により作製された（私信）。この第Ⅲ染色体の全塩基配列決定プロジェクトは筆者らの研究室を含む多くの研究室による共同作業であり、巨大プロジェクトにおける国際協力の例としても非常に意義深いものである。現在、他の染色体についての全塩基配列決定プロジェクトも進行中である。

5. 転写産物と ORF の対応付け

上述したように、筆者らの研究室ではすでに第Ⅲ染色体の整列クローンバンクを用いて、この染色体上に存在する遺伝子の転写産物の同定をノーザンブロット法を用いて行いその結果を発表している⁷⁾。そこで決定された第Ⅲ染色体の全塩基配列から予想される 215 個の ORF が実際に発現している遺伝子かどうかを調べるために、第Ⅲ染色体由来の 155 個の転写産物と ORF との対応付けを行った¹⁸⁾。まず対応付けに先立ち、実験的に作製した第Ⅲ染色体の制限酵素地図⁷⁾と、塩基配列¹⁷⁾から導かれる制限酵素地図を比較したところ、これらの地図は非常によく一致することがわかった（第 3 図）。そこでこれらの地図の一致する部位を基準として、転写産物の存在範囲と ORF の地図上での位置が一致するかどうかを目安にして両者の対応付けを行った。例えば、第 3 図で、YCL 51 w という ORF は転写産物 #22 の存在領域内に存在するので、#22 は YCL 51 w の転写産物である可能性がある。しかし、この ORF は両隣の転写産物 #21 と #23 の存在領域には含まれないので、これらの転写産物とは対応しないと考えられる。このようにして対応付けを進めた結果、89 の ORF については転写産物と一対一の対応付けを行うことができた。また残りの ORF のうち、101 の ORF についてはこのような一対一の対応ではなく複数の転写産物とではあるが、対応させることができた。1つの ORF から複数の転写産物が作られる例はこれまでに多くの遺伝子で知られており、これらの ORF の中にはそのようなものもある可能性がある。残りの 20 の ORF ならびに 14 の転写産物についてはそれぞれ対応の付くものは見いだされなかった（対応付けの結果の一部を第 1 表に示してある）。

第1表 転写産物と ORF の対応付け

ORF (length)	Gene	Transcript No. (length)
YCL63w (387)	<i>DTP 1</i>	11 (6.10), 13 (1.25), 14 (1.45)
YCL62w (588)		11 (6.10), 14 (1.45)
YCL61c (990)		11 (6.10), 15 (3.45)
YCL60c (954)		11 (6.10), 15 (3.45)
YCL59c (951)		15 (3.45)
YCLX1w (399)		15 (3.45)
YCL58c (459)		15 (3.45)
YCL57w (2139)		16 (2.30)
YCL56c (435)		16 (2.30), 17 (1.15), 18 (0.55)
YCL55w (1008)		19 (1.28)
YCL54w (2175)		20 (2.71)
YCLX2c (405)		19 (1.28), 20 (2.71)
YCL53c (450)		20 (2.71)
YCL52c (1251)		21 (1.55)
YCL51w (1761)		22 (2.25)
YCL50c (966)		22 (2.25), 23 (1.20)
YCL49c (939)		23 (1.20)
YCL48w (1392)		24 (2.50), 25 (1.80)
YCLX3c (384)		24 (2.50), 25 (1.80)
YCL47c (777)		24 (2.50)
YCL47c (777)		? 25 (1.80)
YCL46w (324)		26 (1.60)
YCL45c (2283)		/
YCL44c (1254)		26 (1.60)
YCLX4w (447)		26 (1.60)
YCL43c (1569)	<i>PDI 1</i>	/
YCL42w (357)	<i>GLK 1</i>	27 (2.25)
YCL41c (495)		27 (2.25)
YCL40w (1503)		27 (2.25)
YCL39w (2280)		28 (2.43)
/	<i>HIS 4</i>	29 (1.82)
/		30 (0.80)
YCL38c (1587)		/
YCL37c (1401)		31 (1.82), 32 (2.10)
YCL36w (1701)		/
YCL35c (333)		33 (1.28), 34 (1.12), 35 (0.55)
YCL34w (639)		34 (1.12)
YCLX5c (318)		34 (1.12), 35 (0.55)
YCL33c (507)		34 (1.12), 35 (0.55)
YCL32w (1041)		34 (1.12)
YCL31c (894)	<i>HIS 4</i>	/
YCL30c (2400)		36 (4.30), 38 (2.58)
/		37 (1.05)

第3図に示した範囲の転写産物と ORF を対応付けたもの (文献 18 より抜粋)。

実験的に第Ⅲ染色体転写産物を同定した実験⁷⁾では、対数増殖期の細胞から poly(A)⁺RNA を調製しているので、対応の付くものがない ORF の中には対数増殖期以外で発現する遺伝子が含まれる可能性もあると考えられる。

以上のように、出芽酵母の第Ⅲ染色体は最も解析の進

んだ真核生物の染色体であり、ここで述べたような解析は、今後他の染色体の塩基配列決定が進めば当然行われると考えられる。このような解析が進めば、ゲノムレベルで解析結果を議論できるようになり、その意義は図り知れない。また、現在の知見をもとにして、第Ⅲ染色体自身の解析もさらに進み、真核生物の染色体のモデルとして役立つことが期待される。

<神戸大学大学院自然科学研究科・神戸大学理学部>

文 献

- 1) R. K. MORTIMER, D. SCHILD, C. R. CONTOUPOULOU and J. A. KANS : *Yeast*, 5, 321 (1989)
- 2) G. F. CARLE and M. V. OLSON : *Proc. Natl. Acad. Sci. USA*, 82, 3576 (1985)
- 3) P. DE JONGE, F. C. M. DE JONGH, S. MEIJERS, H. YDE STEENSMA and A. W. SCHEFFERS : *Yeast*, 2, 193 (1986)
- 4) J. N. STRATHERN, C. S. NEWLON, I. HERSKOWITZ and J. B. HICKS : *Cell*, 18, 309 (1979)
- 5) J. R. WARMINGTON, R. P. GREEN, C. S. NEWLON and S. G. OLIVER : *Nucl. Acids Res.*, 15, 8963 (1987)
- 6) S. TANAKA, A. YOSHIKAWA and K. ISONO : *J. Bacteriol.*, 174, 5985 (1992)
- 7) A. YOSHIKAWA and K. ISONO : *Yeast*, 6, 383 (1990)
- 8) S. TANAKA and K. ISONO : *Nucl. Acids Res.*, 20, 3011 (1992)
- 9) A. YOSHIKAWA and K. ISONO : *Nucl. Acids Res.*, 19, 1189 (1987)
- 10) Y. KOHARA, K. AKIYAMA and K. ISONO : *Cell*, 50, 495 (1987)
- 11) C. S. NEWLON, L. R. LIPCHITZ, I. COLLINS, A. DESHPANDE, R. J. DEVENISH, R. P. GREEN, H. L. KLEIN, T. G. PALZKILL, R. REN, S. SYN and S. T. WOOCY : *Genetics*, 129, 343 (1991)
- 12) A. E. REYNOLDS, R. M. MCCARROLL, C. S. NEWLON and W. L. FANGMAN : *Mol. Cell. Biol.*, 9, 4488 (1989)
- 13) P. G. PALZKILL, S. G. OLIVER and C. S. NEWLON : *Nucl. Acids Res.*, 14, 6247 (1986)
- 14) P. G. PALZKILL and C. S. NEWLON : *Cell*, 53, 441 (1988)
- 15) A. L. DESHPANDE and C. S. NEWLON : *Mol. Cell. Biol.*, 12, 4305 (1992)
- 16) M. V. OLSON, J. E. DUTCHIK, M. Y. GRAHAM, G. M. BRODEUR, C. HELMS, M. FRANK, M. MACCOLLIN, R. SHEINMAN and T. FRANK : *Proc. Natl. Acad. Sci. USA*, 83, 7826 (1986)
- 17) S. G. OLIVER *et al.* : *Nature*, 357, 38 (1992)
- 18) S. TANAKA and K. ISONO : *Nucl. Acids Res.*, 21, in press (1993)