清酒酵母のゲノム解析：その現状と展望

清酒酵母きょうかい7号 (K7) の全ゲノム解析の終了によって、清酒酵母研究は新たな段階に入ろうとしている。ここでは先ずゲノム解析結果にもとづいて K7 と実験室酵母とを詳細に比較することによって明らかとなった清酒酵母の特徴を中心に紹介し、さらに、ゲノム情報の清酒酵母研究や新規株開発への寄与について展望を述べていただいた。ゲノム情報の活用はまだ緒についたばかりであるが、他の酵母にはない清酒酵母に共通の特徴とさらには清酒酵母の各菌株固有の特徴がゲノムレベルで検討されることにより、系統発生により各清酒酵母が生まれた経緯が明らかとなり、さらに応用面では清酒製造技術にとって新機軸が打ち出されることを期待する。

赤 尾 健

1. はじめに

清酒酵母とは、言うまでもなく酵母 *Saccharomyces cerevisiae* のなかでも、特に清酒醸造に適した性質を有する菌株群のことである。清酒酵母は、実験室酵母と比較して、(1) 清酒もろみのような低温 (< 15℃) でよく発酵する、(2) 20%以上の高いアルコール生成能を有する、(3) 香味成分をバランス良く生成し出来た清酒の香味が優れているなど、清酒醸造にとって他の菌株では得難い特徴を備えている^{1,2)}。また、低孢子形成性や清酒もろみでの高泡形成性、ビオチン合成能など特有の性質も備えている。特に、清酒醸造では酵母が香味成分の多くを生成し品質を特徴づけるため、香味への寄与は本質的である。

では、このような清酒酵母に特有の「清酒酵母らしさ」の原因はどこに由来するのか。通常、それぞれの生物の性質はゲノムに記録された遺伝情報、すなわちゲノム情報によって決定される。また、ゲノム情報は、全体としては生物種を規定するものの、同じ種の中でも菌株や系統ごとに少しずつ違いがあり、それが菌株や系統ごとの性質の違いにつながっている。結局のところ、清酒酵母の特徴の原因もゲノム情報に潜んでいると考えられる。実際、清酒酵母の特徴である高泡形成性とビオチン合成能については、それぞれ清酒酵母

に特異的な遺伝子である *AWA1* と *BIO6* の関与が明らかにされている^{3,4)}。しかし、発酵能や香味成分の生成など多くの特徴について、それらが起因する遺伝的背景はほとんど何もわかっていない。そこで、清酒酵母のゲノム情報を解読して、既によく研究されている実験室酵母のゲノム情報との違いの部分に着目して醸造特性などとの関係を調べていけば、清酒酵母の清酒酵母たる所以に迫ることができると期待できる。

そのような発想から、我々は、清酒酵母の特徴をゲノムレベルで解明するための足がかりとして、代表的な清酒酵母菌株であるきょうかい7号 (K7) の全ゲノム解析を実施し、得られたゲノム情報について、実験室酵母 S288C ゲノムとの詳細な比較を行った^{5,6)}。本稿では、これまでに明らかになった K7 のゲノムの特徴を紹介し、併せて K7 ゲノム情報を基にした清酒酵母ゲノミクスの今後について、現状に加えて若干の展望を述べることにしたい。

2. ゲノム解析の実施体制

本プロジェクトは、酒類総合研究所を中心とする産学官の研究機関によるコンソーシアム及び国内の代表的なゲノムシーケンズセンターである製品評価技術基盤機構 (NITE) による密接な連携の下で実施された。まず NITE において塩基配列決定及びアセンブ

ル（染色体構造の再現）を実施し、引き続いて、コンソーシアム各機関で分担して個々の遺伝子のアノテーション（注釈付け）を行った。また、酒類総合研究所でアノテーションシステムとデータベースを作成し、プロジェクト全体のとりまとめを行った（第1図）。

3. K7 ゲノムの染色体構造

ゲノム解読の方法としては、全ゲノムショットガン法を用いた。すなわち、K7のゲノムDNAをバラバラにしてライブラリー化し、そこから得られたクローンをランダムに配列決定して、一倍体ゲノム（約

第1図 清酒酵母ゲノム解析の実施体制

12Mb: 1200 万塩基) の 9 倍強をカバーする量の塩基配列に相当するリード (生の配列データ) を得た。続いて、リード間の塩基配列がオーバーラップする部分を利用してコンティグを作成し、それらのコンティグについて、各クローンのリンク情報や特殊な方法で作成したゲノム全体の制限酵素地図などを利用して相互関係を明らかにして結合させ、17 本のスーパーコンティグを得た。スーパーコンティグの総延長は 12Mb 弱であり、推定ゲノム全長の 98% に相当した (第 1 表)。なお、残りの 2% は今回の解析では配列が得られなかった 180 箇所あまりのギャップ領域であり、データベース上のコンセンサス配列においては、便宜上、それぞれが 100 個の N に置き換えてある。これらのスーパーコンティグは 16 本の染色体とミトコンドリアに相当すると考えられた。実際に、それぞれは S288C の 16 本の染色体セット及びミトコンドリア DNA のいずれかと対応し、また対応する染色体ペアの配列は全域にわたりよく対応していた (第 2 図)。

ただし例外として、各染色体の末端近傍領域 (テロ

メア〜サブテロメア領域) では、染色体間の組換えによる再配置が多数見られた。これらは、染色体末端では同種間であっても比較的頻繁に観察される現象である。また、第 5 及び第 14 染色体の内部に、約 100kb と約 30kb の逆位領域がそれぞれ存在した (第 2 図)。第 14 染色体の逆位は他のゲノム解析菌株にも存在した⁷⁾が、第 5 染色体については、K7 で初めて見つかったものである。こうした逆位は、同一染色体上の逆向きの相同領域間での組換えが原因と考えられるが、K7 では第 5 染色体の逆位の両端の境界領域には、レトロポゾンである Ty 因子の一種である Ty2 の挿入による長い逆向きの相同配列がいずれも存在していた。ゲノムのわかっている他の菌株の該当する位置にはこのような構造は見られないことから、この逆位は K7 に特徴的な現象と考えられた。

ゲノムがわかっている他の菌株 (S288C, ワイン酵母 EC1118, 臨床分離株 YJM789) と K7 との大きな系統的な関係を見積もるために、4 株のゲノム全体を比較して、相互の小規模な多型 (置換・挿入・欠失された塩基) の数を調べた^{5,6,7,8)}。K7 とその他 3 株との間のそれぞれの多型の数は、その他 3 株同士の間での多型の数よりも多く、K7 はこれら 3 株間の関係と比較して、系統的にやや離れていることが示唆された (第 3 図)。

4. 相同染色体間のヘテロザイゴシティ (ヘテロ接合性)

ここまで触れてこなかったが、K7 は、異なる遺伝子型を有する 2 つの一倍体が接合したヘテロザイガスな二倍体である⁹⁾。つまり、2 本の相同染色体の間には、ヘテロザイゴシティ (対応する位置であっても塩基が異なる部位) が存在することを意味する。全ゲノムショットガン法の場合、2 本の相同染色体を区別してアセンブルを行うことは不可能である。したがって、ヘテロザイガスな部位においては、2 本の相同染色体のどちらかの塩基がランダムに選ばれて、その位置のコンセンサス配列を構成している。つまり、ここで得られているゲノム配列というのは、2 本の相同染色体がキメラに合成された擬似一倍体様の配列であると言える。そこで、K7 のヘテロザイゴシティの程度を明らかにするために、アセンブリー時に得られた塩基アラインメントの同一位置での塩基の不一致情報や、生

第 1 表 決定された K7 ゲノム配列の概要

リード (生の配列) の数	149,011
リードの平均の長さ	778.4 bp
得られたコンティグの数	706
コンティグの平均冗長度 *	9.24
スーパーコンティグの数	17
スーパーコンティグに含まれるコンティグの数	200
GC 含量	38.50%
スーパーコンティグの総延長 (ギャップ領域を除く)	11.85 Mb
ギャップ領域の推定の長さ	0.23 Mb
ゲノムの全長の推定値 (ギャップ領域を含む)	12.08 Mb
スーパーコンティグのゲノム全長に対するカバー率	98.10%

*) 冗長度: コンティグのある位置について、重複しているリードの数
 ・K7 は二倍体の株なので、スーパーコンティグのコンセンサス配列は 2 つの相同染色体の配列を代表している。
 ・706 から 200 を引いた 506 本のコンティグは、おそらくは無視できない相同染色体間の多型や配列の繰り返しによって、スーパーコンティグに含まれなかったと考えられた。
 ・K7 の 17 本のスーパーコンティグのセットは、S288C の 16 本の染色体及びミトコンドリア DNA に対応していた。

第2図 K7 - S288C 間の染色体構造の対応

K7 と S288C の対応する染色体間で塩基配列の相同性が高い領域をドットプロットで表した。個々のグラフの数字は染色体番号である。第5及び第14染色体の波線で囲んだ部分は逆位領域を示している。ここではグラフの長さは染色体の長さにかかわらず同一としている。

第3図 K7と各種菌株のゲノム間の変異塩基の個数

実験室株 S288C, ワイン酵母 EC1118, 臨床分離株 YJM789 の各染色体の塩基配列を, 対応する K7 の染色体 (= スーパーコンティグ) の配列と比較し, K7 に対する変異塩基の数を求めた。その他 3 株間も含めて比較の方向性については, 図中の矢印で表記している。なお, 異なる染色体間の組換えが起こった領域や繰返し配列が存在する領域については, ここでは考慮していない。

のシーケンスの波形データを参考にして, 相同染色体間でヘテロザイガスな位置を検索したところ, 1347 個が抽出された。つづいて, これらの染色体座標上での分布を調べた。すると, 大変興味深いことにヘテロザイゴシティはゲノム上に一様に存在するのではなく, 一部の領域に偏ってブロック状の分布を形成していることが明らかとなった (第4図)。

では, このような分布はどのような原因によって生じたのだろうか。まず, 単純な点突然変異が蓄積した可能性であるが, 点突然変異はゲノム上でおおむね一様な確率で生じると考えられ, このようなパターンを全く説明できない (ごく一部のヘテロザイガスな塩基に寄与している可能性はある)。では反対に, よりヘテロ接合度の高い状態から, 部分的なホモザイガス化が徐々に進行した結果, ヘテロザイガスな領域がブロック状に残った可能性はどうだろうか。ゲノムの部分的なホモザイガス化に寄与する要因としては, ヘテロ接合性の喪失 (Loss of Heterozygosity : LOH) と交配が考えられる。LOH は, 染色体の損傷の修復など

の際に起こる有糸分裂組換えによって, 一方の相同染色体のヘテロザイガスな領域の一部が, もう一方の相同染色体の対応する領域によって置換されることや, 染色体分配異常の結果として片方の相同染色体がもう一方により置換されることで生じる現象である (第5図)。LOH が起こるごとにゲノム全体のホモザイガス化が進むことになる。LOH は一定の確率で生じる現象と考えられ, 特定の遺伝子座での頻度は 10^{-5} 程度とされる。これをゲノム全体に拡張して考えると, 世代ごとにゲノム上のどこかで LOH (あるいはその原因となるイベント) が起こる確率は比較的高いと予想される。そこで, 最後の交配によって K7 の祖先となるヘテロ二倍体株が形成された後, 現在まで世代を重ねる間に, ゲノム全体で LOH が繰返し起こることで部分的なホモザイガス化が徐々に進行し, 結果的に虫食いの跡のようにヘテロザイガスな領域が残ったと考えると, このような一部に偏った分布についてうまく説明できる。ただし, もうひとつの可能な要因である交配の寄与についても考慮しなければならない (現

在の K7 は孢子形成能をほぼ失っているが、二倍体である以上、祖先株は交配によって生じたのは事実である)。まず、最後の交配が異系交配だったとすると、

その場合には K7 のヘテロザイゴシティの分布に対する LOH の寄与は決定的といえる。一方で、それが同系交配だった可能性も十分に考えられることである。

その場合、K7 の祖先の二倍体のゲノムは、形成された時点で既に部分的にホモザイガスな領域を含んだ状態だったことになる。その時点でのヘテロザイゴシティの分布がどのようなものだったかを知ることは不可能だが、推定される LOH の頻度と、人間が K7 の系統を飼ひ慣らしてきた期間の世代数を考えると、現在の K7 のヘテロザイゴシティの分布は、K7 の祖先の二倍体が生じた時点で既に形成されていたものではなく、その時の状態から、その後の世代で LOH によるゲノムのホモザイガス化が徐々に進行することで形成されたものであると考えて問題ないだろう。

ここで重要な点は、LOH は細胞単位で生じる現象であり、また LOH が起こった領域で、ハプロタイプのいずれかの選択が伴うことである。ハプロタイプとは、過去に接合した際の親である 2 つの一倍体に由来する遺伝子型のことであり、LOH が起こった染色体領域では、その位置の遺伝情報のうち、一方の一倍体に由来するものが失われもう一方のものに置き換わる。そのとき、どちらのハプロタイプが選択されるかにより、結果として生じるディプロタイプも異なる。そこが発酵力や香味成分の生成に関与する遺伝子座であった場合は、菌株の醸造特性にも変化が生じることもあるだろう。こうしたことなどから、我々は、清酒酵母のように有性生殖環をほとんど失ってしまった二倍体酵母では、交配に代わって、LOH が菌株の小進化（多様化や分化）を促す大きな推進力になっているのではないかと考えている。また、ヘテロザイガスな二倍体酵母は、LOH が起こるごとに、ホモザイガス化が進行していく。ヘテロシス（雑種強勢）の考え方によれば、このとき、生物としては適応度が低下していくことになるが、清酒酵母の場合は、曝される環境が清酒醸造という限られた条件である、既にその環境に適応したものとして選抜されている、あるいは適応力が低下しても有用形質を獲得すれば人間によって保護されるといった事情があるので、それほど不利は生じていないのかもしれない。

第4図 K7 相同染色体間のヘテロザイゴシティの分布
K7 の相同染色体間のヘテロザイガス(両者の対応する位置の塩基が異なる)な部位をアセンブリー情報から抽出し、染色体領域 10kb あたりの数を染色体座標上にプロットした。矢頭はセントロメアを示す。

第2表 K7 と S288C の遺伝子の数

遺伝子の種別	K7	S288C
ORFs *)	5,815	5,748
tRNA 遺伝子	296	299
その他の RNA 遺伝子 (リボソーム RNA 遺伝子を除く) **)	93	93

*) Ty 因子に含まれる ORF とテロメア因子上に存在する遺伝子は、含んでいない

**) K7 の第 12 染色体の配列には、rRNA 遺伝子 (RDN37, RDN5) 及び非転写スペーサー配列に対応する rDNA ユニットを 1 コピー含んでいる。

5. 遺伝子領域の予測と機能アノテーション

K7 のゲノム中の遺伝子位置の予測とその機能アノテーションを行った。K7 のほとんどの遺伝子については、S288C との間でオルソログ (相互に対応する遺伝子) の関係にあると考えられたので、まず S288C の遺伝子との相同性をベースとして、ソフトウェアも利用した自動アノテーションを行った後に、コンソーシアムメンバーが分担し、遺伝子の位置と機能の確認と修正を行った。結果的に、5815 個の ORF が予測された (テロメアやレトロポゾン Ty とオーバーラップ

第6図 K7 と S288C との間の ORF 相同性の分布

K7 の各 ORF を、S288C の全 ORF を対象に Blast 検索を行って、もっとも相同性が高かった ORF との相同性の値の分布を示した。

する ORF, Dubious ORF は除く)。各種の RNA 遺伝子については、K7 での tRNA 遺伝子 3 個の欠失を例外として、S288C と同一のセットを有していた (第 2 表)。予想されたとおり K7 の ORF と S288C の ORF リストの相同性は非常に高く、アミノ酸レベルで ORF 全体の 30% が 100% 一致、50% が 99% 一致であった。相同性が 97% 以上のもので全体の 90% 以上を占めた (第 6 図)。その一方で相同性が 90% よりも低い ORF も全体の数%程度存在した (第 6 図)。こうした ORF には、第 6 項で述べるような S288C にはオルソログが存在しない ORF や、第 7 項で述べるようなフレームシフト変異やナンセンス変異などの原因によって双方の領域の末端が一致しない ORF などが含まれている。

6. K7 と S288C の間で差のある遺伝子

K7 と S288C の遺伝子の対応付けを行った結果、両者のどちらか一方にしか存在しない遺伝子、すなわち K7 にあって S288C にないものが 48 個（ギャップにより断片化しているものはそれぞれを 1 つとして数えた）、S288C にあって K7 にないものが 49 個、合計で 97 個が存在することが判明した。これらの遺伝子のほとんどは、いずれか一方で遺伝子領域そのものが存在しなかった。これらの多くは、ゲノム中でも可塑性が高いサブテロメア領域に位置しており、ここでの染色体構造の再配置などによって欠失または獲得が起こったと考えられた。一方、染色体内部の遺伝子の多くについては、遺伝子重複やリピータ配列などが関与することで差が生じたと考えられた。

(1) K7 遺伝子のうち、S288C には存在しないもの

これらの遺伝子のいくつかについては第 3 表に例示した。この中で *AWA1*, *BIO6*, *BIO1*, *KHR1* は、これまでに報告がなされているものである^{3,4,9,10}。*KHR1* は K7 では染色体の内部領域に、Ty 因子の末端に由来する 2 つのリピータ配列（LTR と呼ばれる）に両側をはさまれる形で存在していた。対応する領域の配列の比較から、S288C では、これらの 2 つの LTR 間の組換えにより *KHR1* が欠失した推測された。

また、本解析で明らかになった *EHL1/2/3* については、パラログ関係にあり互いにほぼ同一の配列であった。これらは、環状エーテルであるエポキシド化合物の分解酵素をコードすると考えられ、エポキシド化合物で転写誘導された。しかし、K7 の 3 重破壊株で清酒小仕込を行っても表現型は観察されず、真の基質や清酒醸造での役割は不明である。このオルソログは真核生物では *S. cerevisiae* の一部の菌株や近縁の *S. paradoxus* でわずかに見出されているだけだが、バクテリアでは多くの種で見ついている。したがって、この遺伝子は、これらの酵母の共通祖先において何らかのバクテリアから水平伝播により獲得され、その後染色体末端の組換えなどにより、多くの酵母菌株で失われたものと考えられる。

(2) S288C 遺伝子のうち、K7 には存在しないもの

(1) と同様に、該当する遺伝子のうちいくつかを第 3 表に例示した。このうち、*MPH2/3*, *SOR1/2*, *HXT15/16* については、いずれも糖質資化に関与す

ると推察されるが、ゲノム中ではサブテロメア領域に *MPH* ~ *SOR* ~ *HXT* のようなクラスター構造をなしており、これが S288C では 2 コピー存在するのに対し、K7 には 1 コピーもない。これらの遺伝子の具体的な機能や役割はそれぞれよくわかっていない。ただ、他の菌株もよく調べてみないと何とも言えないが、種内の進化の過程における染色体間の移動・重複・欠失がクラスター単位で起こっているとすれば、もしかすると相互に何らかの関連が存在する可能性もあるかもしれない。

PHO3 については、S288C では染色体の内部領域において、パラログである *PHO5* と隣接して存在し、いずれも細胞外の酸性ホスファターゼをコードする。*PHO5* がリン酸非存在条件でのみ活性化されるのに対して、*PHO3* は構成的に活性化されている。以前

第 3 表 K7 と S288C のいずれか一方にのみ存在する遺伝子の例

(1) K7 遺伝子のうち、S288C には存在しないもの

遺伝子名	遺伝子産物
<i>AWA1</i>	清酒もろみの高泡形成に関わる細胞壁タンパク質
<i>BIO1</i>	ビオチン合成経路上のピメロイル-CoA 合成酵素
<i>BIO6</i>	ビオチン合成経路上の KAPA 合成酵素
<i>EHL1/2/3</i>	推定エポキシド加水分解酵素
<i>KHR1</i>	耐熱性キラー因子

(2) S288C 遺伝子のうち、K7 に存在しないもの

遺伝子名	遺伝子産物
<i>MPH2/3</i>	α -グリコシド輸送体
<i>SOR1/2</i>	ソルビトールデヒドロゲナーゼ
<i>HXT15/16</i>	ヘキソース輸送体ファミリー
<i>PHO3</i>	構成的な酸性ホスファターゼ
<i>ENA2/5</i>	P 型 ATP 分解酵素
<i>ASP3</i>	細胞壁 L-アスパラギナーゼ
<i>PPT1</i>	タンパク質 S/T ホスファターゼ
<i>ADH7</i>	NADH 依存性中鎖アルコールデヒドロゲナーゼ
<i>RDS1</i>	シクロヘキシミド耐性に関与する亜鉛クラスタータンパク質
<i>AIF1</i>	ミトコンドリア細胞死エフェクター
<i>CWP1</i>	細胞壁マンノプロテイン

・機能がわかっているか推定できるものを中心に例示した。
・下線は染色体の内部に位置する遺伝子である。

より K7 を含む優良清酒酵母の細胞外の酸性ホスファターゼ活性の発現は構成的ではなく、リン酸存在下で抑制されることが知られていたが¹²⁾、*PHO3* の欠失がその原因と考えられた。

ENA2/5 は、P 型 ATP 分解酵素をコードし、S288C では、遺伝子重複に起因すると推測される *ENA1* ~ *ENA2* ~ *ENA5* のパラログスなクラスターを形成されている。K7 や他の菌株には、いずれか 1 コピーしかで存在しない。クラスター状に 3 コピーも有するのはむしろ S288C に特有な現象と言える。また、*ASP3* は細胞壁アスパラギナーゼをコードするが、S288C ではリボソーム DNA (rDNA) の 100 コピー以上に及ぶアレイ領域に隣接して 4 コピーがタンデムに存在するが、S288C 以外では一部の実験室株でしか見つかっていない。rDNA アレイのコピー数は、条件によって増減することが知られているので、*ASP3* の存否はそのことと関連があるのかもしれない。

PPT1 は、S288C では第 7 染色体の内部に位置するが、K7 の対応する位置においては、*PPT1* を含む領域は、Ty 因子の一種である Ty2 によって置換されて失われていた。このような構造はまれな現象の結果であると考えられるので、第 8 項で改めて取り扱う。また、*PPT1* はセリン / スレオニンホスファターゼをコードするが、最近の研究により、この遺伝子の欠失が K7 の高い発酵力の原因のひとつであることが明らかとなっている¹³⁾。

このほか、(1) 及び (2) の遺伝子のいずれも第 3 表に示していないものについては、機能はわかっていないものが多いが、これらの中に清酒酵母の醸造特性に関与するものも含まれる可能性も考えられる。紙面の都合ですべてをここでは紹介できないため、文献 5 の補足データを参照していただきたい。また、サブテロメア領域の多コピー遺伝子 (*PAU*, *COS*, *DAN*, *SNO* など) については対応関係が複雑であり、本研究では考慮していない。

7. K7 と S288C の間で、構造が異なるオルソログペア

K7 と S288C 間でオルソログ関係にある遺伝子のうち、約 200 個については ORF 構造の多型、すなわち、ORF の分断、融合、末端の不一致などが見出された。これらのほとんどは、フレームシフト変異やナンセン

ス変異などの点変異が原因と考えられ、中には K7 型か S288C 型のいずれかで遺伝子産物の機能が変化、あるいは欠失しているであろうものも多く存在したが、これらの中には清酒酵母の特徴的な性質に関与するものも含まれると考えられる。例えば、ガラクトースの資化に関わる *GAL3*, *GAL4* は、K7 ではいずれもフレームシフト変異により分断されていたが、このことは既に知られている K7 のガラクトース資化能の低さとの関連があるかもしれない¹⁴⁾。

また、*MSN4* は、ストレス応答に関わる転写因子をコードするが、K7 では一塩基置換によって、DNA 結合ドメインを含む ORF の C 末端側の領域を失っていた。最近の研究では、*Msn4* の機能が実際に失われ、このことが K7 におけるストレス応答の欠損を引き起こすとともに、結果的に K7 の高発酵性の一因であることが明らかにされている¹⁵⁾。

ところで、酵母 *S. cerevisiae* には、組換えによる接合型変換機構が存在し、更には、この機構が正常に機能する菌株 (ホモタリク) としない菌株 (ヘテロタリク) とが存在する。K7 はヘテロタリクな菌株である。この接合型変換過程には *HO* 遺伝子が必須である。K7 では末端の不一致や分断などの顕著な多型ではなかったが、機能欠失を引き起こす既知の点変異が見出され¹⁶⁾、また内部に 37 アミノ酸残基相当の欠失が存在した。したがって、K7 がヘテロタリクであるのは、これらの変異により *HO* が正常に機能できないことが大きな原因と推定された。

ここでは、以上の 3 例を挙げるにとどめるが、その他にも清酒酵母の特徴的な性質に関与する多型が多数存在する可能性が大きいものと期待している。

8. レトロトランスポゾン Ty 因子

酵母のゲノム中には、可動性遺伝因子であるレトロトランスポゾン Ty 因子 (Ty1 ~ Ty5) が存在する。Ty 因子は自分自身を元の位置に維持したまま、転写、逆転写、新たな染色体領域への挿入というサイクルで転位しながらコピー数を増していく。また、Ty 因子はその両末端に存在するリピータ配列である LTR 領域内での組換えにより切り出されて失われることがある。その場合は、Ty 因子の痕跡として 1 コピーの LTR が残る。また、転位現象は細胞ごとに独立して起こるので、過去の転位の履歴は系統ごとに異なるが、

第4表 K7とS288Cのゲノム中に存在するTy因子

A. Ty因子とその末端リピート(LTR)の数

Ty及びLTRの種類別	K7	S288C
Ty1	11	31
Ty1delta	220	264
Ty2	22	13
Ty2delta	61	36
Ty3	2	2
Ty3sigma	38	42
Ty4	4	3
Ty4tau	35	35
Ty5	1	1
Ty5omega	11	8

・delta, sigma, tau, omegaは、それぞれLTRの名称である。

B. Ty因子の転位イベントの推定回数

Tyの種類別	K7とS288Cで 共通のもの①	K7に特有なもの②	K7の全ての転位①+②	S288Cに特有なもの③	S288Cの全ての 転位①+③
Ty1	150	53	203	84	234
Ty2	5	41	46	18	23
Ty3	19	9	28	21	40
Ty4	19	13	32	12	31
Ty5	5	5	10	2	7
合計	198	121	319	137	335

菌株によるTyの分布の違いについて詳細に調べた例はこれまでにほとんどない。

そこで、K7とS288Cのゲノム中に存在するTy因子とLTRについて比較した(第4表)。数については、Ty因子で40～50コピー、LTR(Tyを伴うものとsoloの両方)では400コピー弱であり、両株とも同程度であった。しかし、Ty因子の種類を考慮すると、K7ではS288Cと比較してTy1が少なく、Ty2が多かった(LTRも含む)(第4表A)。また、ゲノム中のそれぞれのTyやsoloLTRの側方配列を元に、両株での挿入位置と回数を見積もったところ、共通祖先での挿入は198回であり、両株が系統的に分岐した後の挿入は、K7で121回、S288Cで137回と大差はなかった(第4表B)。また、LTRでの組換えが起こらずに残存するTy因子は、ほとんどが両株の分岐後に挿入されたものであり(第4表B)、挿入からの時間経過とTy因子の切り出し頻度の間には一定の傾向があるのかもしれない。

Ty因子は転位の際に、まれではあるが遺伝子のプ

ロモーターやORF領域に挿入し、遺伝子の発現や機能に異常をもたらすことが知られている。例えばS288CではTy1の挿入によって、エルゴステロール合成に関与する*HAPI*が正常に機能していない¹⁷⁾。K7ではこのような遺伝子機能に異常を生じるような挿入は見出されなかった。

ただし、Ty因子の挙動により生じた興味深い構造として、K7では、第6項(2)で述べたように、*PPT1*を含む領域がTy2による置換で失われていた。すなわち、S288Cの第7染色体で*PPT1*と1個のtRNA遺伝子を含む約2.6kbの領域が、K7ではTy2によって置換された珍しい構造となっていた(第7図A)。

染色体上のTy因子が転位する際には、まずTy領域全体がRNAに転写された後に、2本鎖cDNAに変換され、染色体上の新たな位置に挿入される。位置としてはtRNA遺伝子の上流領域のことが多い。しかし挿入の競合過程として、染色体上のLTRを介した組換え置換による染色体への導入も可能である(第8

図)。S288C の *PPT1* 周辺領域の tRNA 遺伝子の分布を考慮すると、K7 の祖先において、*PPT1* を挟んで存在したであろう 2 つの染色体上の LTR と Ty2cDNA の 2 つの LTR との間の組換えの結果として置換（遺伝子変換）が起こり、それにより K7 の *PPT1* を含む領域が失われたと考え、この構造を非常によく説明できる。K7 では同様の構造が第 1 染色体でも見出されており、ここでは S288C に存在する 2 個の tRNA 遺伝子が失われていた（第 7 図 B）。ここで興味深いのは、Ty 因子の転位における染色体への導入に際しては、実際には挿入の方が置換よりもはるかに

優勢なことである¹⁸⁾。このため、Ty 因子の LTR を介した置換が観察されたのは、これまで実験室条件で人為的に挿入をブロックした場合に限られていた。遺伝的な改良を受けていない野生型の酵母で、このような現象が自発的に起こった痕跡が見出された例としては、おそらくこれが初めてである。しかし、こうした現象は低いながらも一定の確率で起こるものと考えられるので、今後様々な菌株で詳細な比較解析が進行すれば、他にも同様の痕跡が見つかるだろう。

第 7 図 レトロポゾン Ty 因子の置換による染色体領域の欠失

S288C の染色体上の波線で囲まれた領域が、K7 では Ty2 によって置換されていた。

A: K7 の第 7 染色体では、*PPT1* と tRNA 遺伝子が欠失していた。

B: K7 の第 1 染色体では、tRNA 遺伝子 2 つが欠失していた。

A Ty因子の挿入（本来の転位様式）

B Ty因子による置換（稀な現象）

第8図 Ty因子の染色体へのインテグレーション

A： 本来の Ty 因子の複製サイクルにおいては、Ty 因子の cDNA は Ty 因子自身のコードするインテグラーゼによって染色体に挿入される。これにより周囲の領域が欠失することはない。

B： まれにはではあるが、挿入の競合過程として、Ty 因子の cDNA の LTR と染色体上の LTR との間の組換えが起こり、結果的に染色体上の領域が Ty 因子に置換されることがある。このとき、置換された染色体領域は失われる。

9. K7 ゲノムからの展開

優良清酒酵母としては、K7 以外にも K6, K9, K10 及びこれらに由来する菌株が酒造現場で広く実用に供されている。分類的には K7 グループと呼ばれるこれらの菌株は、いずれも清酒製造環境での低温発酵能や高いアルコール生産力、清酒らしい香味生成能など、「清酒酵母らしさ」の根本においては K7 と共通している。実際、これらの菌株は他の菌株と比較して互いに近縁関係にあることも知られている¹⁹⁾。一方で、菌株ごとの個性、つまり特徴的な醸造特性が備わっているのは、それぞれに遺伝的な差異が存在するためと考えられる。

こうした事柄を考慮に入れ、K7 のゲノム情報を基にした今後の清酒酵母ゲノミクスの展開を考えると、大きく二つの方向性が浮かび上がってくる。一つ目としては、S288C などその他の菌株との比較解析であり、現在の清酒酵母の清酒酵母らしさの解明、すなわち清酒酵母に共通した特性の原因である清酒酵母に共通したゲノム上の特徴の解明、が主な目的である。二つ目としては、清酒酵母菌株同士の比較ゲノム解析であり、清酒酵母のそれぞれの菌株らしさの解明、すなわち菌株ごとの特性の原因である菌株ごとに固有なゲノム上の特徴の解明が主な目的である。これら二つの方向性

は、それぞれに醸造特性の原因、系統進化、識別など、清酒酵母の総合的な理解だけでなく、これからの酒造技術を考える上で欠かせない様々な側面を含んでいるが、いずれも清酒酵母の菌株間のゲノム情報の共通性と特異性を知ることが起点となる。

10. K7 に特徴的なゲノム構造の各種菌株での分布

そこでまず、今回明らかになった K7 と S288C の間のゲノム構造上の比較的大きな特徴、すなわちいずれか一方にしか存在しない遺伝子や染色体の逆位領域について、様々な菌株間での分布をサザン解析や PCR により調べた（第5表）。その結果、K7 グループの菌株はすべての項目で共通して K7 型の属性を示し、検討した範囲内ではまったく識別することができなかった（Bu9-7 はウイスキー酵母とされるが、元は K7 グループ由来と考えられる）。その他の菌株では、系統により若干の傾向はあったが、項目ごとに属性が異なっていた。K7 グループの菌株は、非常に近縁な関係にあることがゲノムの特徴の共通性からも改めて示されたことになる。ところで、K8 については、元々は K6 の変異体とされていたが、醸造特性の面では、K6 を含む K7 グループの株とは趣が異なる印象を与えていた。ここでの結果は、K8 が K6 の変異体ではなかったことをゲノムレベルで明確に示している。

第5表 K7 のゲノム構造の特徴の各種菌株での分布

菌株名及び用途		S288C がない K7 遺伝子						K7 がない S288C 遺伝子				逆位	
		K7_00005	K7_02212	K7_02354	K7_03550	K7_06167	K7_06168	<i>PHO3</i>	<i>ASP3</i>	<i>PPT1</i>	<i>AIF1</i>	第5染色体 (約 100kb)	第14染色体 (約 30kb)
きょうかい 1 号 (K1)	清酒	○	○	○	○	○	○	○	×	○	○	S	K/S
きょうかい 2 号 (K2)	清酒	○	○	○	○	○	○	×	×	○	○	S	K
きょうかい 3 号 (K3)	清酒	○	○	○	○	○	○	○/×	×	○	×	S	K
きょうかい 4 号 (K4)	清酒	○	○	○	○	○	○	○	×	○	×	S	K/S
きょうかい 5 号 (K5)	清酒	○	○	○	○	○	○	○	×	○	×	S	K
きょうかい 6 号 (K6)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 7 号 (K7)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 8 号 (K8)	清酒	○	○	○	○	○	○	○	×	○	×	S	K/S
きょうかい 9 号 (K9)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 10 号 (K10)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 11 号 (K11)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 12 号 (K12)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 13 号 (K13)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 14 号 (K14)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
きょうかい 1501 号 (K1501)	清酒	○	○	○	○	○	○	×	×	×	×	K	K
S. sake yabe-kozai (RIB0001)	清酒	○	○	○	○	○	○	○	×	○	○	S	K/S
きょうかい焼酎 2 号 (S-2)	焼酎	○	○	○	○	×	×	○	×	○	○	S	K
SH-4 (RIB1019)	焼酎	○	○	○	○	×	×	○	×	○	○	S	K
Bu9-7 (RIB1020)	ウイスキー	○	○	○	○	×	×	×	×	×	×	K	K
S-3 (RIB1021)	紹興酒	×	○	○	○	×	×	○	×	○	×	S	K
Lao-ron yeast (RIB1022)	スピリッツ	×	×	×	×	○	○	○/×	○	○	○	S	K
American Alcohol (RIB1023)	アルコール	×	×	×	○	×	×	○/×	×	○	○	S	K
Alcohol beverage 389 (RIB1024)	アルコール	×	×	×	×	○	○	○	○	○	○	S	K
EC-1118 (RIB1050)	ワイン	×	×	×	○	×	×	○	×	○	○	S	K
Montrachet (RIB1052)	ワイン	×	×	×	○	○	○	○	×	○	○	S	K
OC-2 (RIB1057)	ワイン	×	×	×	○	×	×	○	×	○	○	S	K
NCYC 240	エールビール	×	×	○	○	×	×	○/×	×	○	×	S	K/S
NCYC 1026	エールビール	×	×	○	○	×	×	○	×	○	○	S	K/S
X2180 (ATCC26109) *	実験室	×	×	×	×	×	×	○	○	○	○	S	S
S1278b (ATCC42800) **	実験室	×	×	×	×	×	×	×	×	○	×	S	S
W303-1A (ATCC208256) **	実験室	×	×	×	×	×	×	○	○	○	○	S	S
A364A (ATCC28383) **	実験室	×	×	×	×	×	×	○	○	○	×	S	S

* X2180: S288C と同一遺伝子型の二倍体

** いずれも一倍体株 (MATa)

・分布調査を行った遺伝子の産物は次のとおり。K7_00005: 機能未知, K7_02212: 機能未知 (*S. mikatae* に類似の ORF あり), K7_02354: ヒストン Aセチル転位酵素ホモログ, K7_03550: 耐熱性キラー因子, K7_06167: YDR384W に類似の推定膜タンパク, K7_06168: アミダーゼ, *PHO3*: 構成的酸性ホスファターゼ, *ASP3*: 細胞壁 L-アスパラギナーゼ, *PPT1*: タンパク質セリン / スレオニンホスファターゼ, *AIF1*: ミトコンドリア細胞死エフェクター

・遺伝子の存否については, ○が存在の場合を, ×は欠失の場合を示す。逆位領域については, K は K7 と同じ方向の場合を, S は S288C と同じ方向の場合を示す。○/×, K/S は, それぞれ両方のタイプの遺伝子座の混在を示し, おそらくはヘテロザイガスな構造を反映している。太線で囲んだ菌株は, 調査した全項目について K7 と同じ属性であったものを示す。また, K7 と同じ属性の部分については, 太字で表している。

11. 清酒酵母の識別技術

ここで、菌株の識別についても少し述べておきたい。第5染色体の逆位や *PPT1* の欠失は、K7 グループに特異的であり、その他の菌株では見られなかった（第5表）。加えて第7項で述べた *MSN4* 遺伝子の発酵力に関わる変異も、実は K7 グループに共通かつ限定して分布することが、その後の解析で明らかになっている¹⁵⁾。これらは、系統の異なる菌株間で独立して生じる確率が極めて低いまれな構造であり、K7 グループの系統をゲノムレベルで簡便に識別するために有効なマーカーとなり得るものである。最近、自然界から酵母を分離・選抜し、実用に供しようとする試みが多くなされている。自然界には K7 グループのエスケープが存在しても不思議ではない。分離されてきた候補菌株が K7 グループの系統なのか否かを、まずはこのようなマーカーを利用して識別を行えば、スクリーニングの方針として、新しい遺伝子資源の探索を優先する場合、安全醸造を優先する場合などのいずれであっても、選抜過程の効率化に有効である。識別精度は現状でも十分に高いと考えているが、今後マーカーの種類や数を増やしていくことで、更に高いものとなるはずである。また、マーカーの識別レベルの細分化や、マーカーと醸造特性との関連付けが進めば、単なる系統の識別にとどまらず、菌株の遺伝子判定、すなわちその菌株がどのような特性を備えているか、ある程度予測も可能となるだろう。

12. 清酒酵母ゲノミクスの今後

K7 グループと S288C のゲノム情報の差としては、遺伝子そのものの有無、ORF の多型以外にも微細なアミノ酸置換が多くの ORF 中に存在していた。これらの差と清酒酵母の特徴的な醸造特性との間の関連性について、明らかになった点はほんのわずかである。このような課題への取組みにおいては、手法的な困難さも依然として残されているが、清酒酵母を理解する上で不可避な問題であり、我々も様々な観点から検討を行っている。

K7 グループの優良清酒酵母同士は非常に近縁ではあるが、それぞれに個性があり、それゆえに現場では造りの種類に応じて使い分けられている。清酒醸造の現場にとっては、清酒酵母全体の特徴よりも、むしろ

菌株ごとの個性がどのような遺伝的要因により生じているかの方がより身近で興味深い問題であろう。おそらくは、主にそれぞれのゲノム間の SNP レベルの微細な差が原因であろうと予想される。

あるいは、K7 グループに比較的近縁なことが示されている明治～大正期に分離された古い清酒酵母菌株や焼酎酵母菌株¹⁹⁾と K7 グループとの違いも、同様に興味深い問題である。古い清酒酵母は K7 グループとどう違うのか、使われなくなっていく背景にはどのような遺伝的要因があるのか、焼酎酵母のどこが焼酎醸造に適しているのか、など様々な課題が手つかずで残されている。これらについても SNP レベルの差異が大きく効いているはずである。

このような多数の菌株のゲノム間の微細な差については、ここ数年の間に急速に発展したいわゆる次世代 DNA シークエンサーを用いたゲノム変異解析により、従来よりも格段に容易に調べることが出来るようになった。現在我々も、K7 のゲノム配列をリファレンスとして、清酒酵母を中心とした菌株を対象にこうした解析を網羅的に進めている。近い将来には清酒酵母の菌株間の塩基レベルでの差違がつまびらかになり、そこから優良菌株や個々の菌株を特徴づける情報をはじめとして、菌株の識別に有効な情報、菌株の系統分化に示唆的な情報など多くの知見が得られると予想される。それらを十分に活用しながら、個別の遺伝子レベルやゲノムワイドな解析などを多面的に進めることで、清酒酵母の総合的な理解が深まり、また醸造の現場での利用の高度化が進んでいくことを期待したい。

なお、K7 のゲノム情報については、酒類総合研究所と NITE が、インターネット上でそれぞれ運営する以下のデータベースで閲覧できるので、積極的にご利用いただければ幸いである。

・酒類総合研究所：清酒酵母ゲノムデータベース (SYGD)

<http://www.nrib.go.jp/ken/sygd.htm>

・NITE：NITE ゲノム情報データベース (DOGAN)

<http://www.bio.nite.go.jp/dogan/top>

最後になりましたが、K7 のゲノム解析については、多くの皆様のご尽力により成し遂げられたものです。この場をお借りしまして関係各位に深謝いたします。

(独立行政法人酒類総合研究所)

文献

- 1) Wu, H., Watanabe, T., Araki, Y., Kitagaki, H., Akao, T., Takagi, H., Shimoi, H.: *J. Biosci. Bioeng.*, **107**, 636-640 (2009)
- 2) Kodama, K.: Sake-brewing yeasts. In: Rose A. H. and Harrison J. S. (eds) *The Yeasts* volume 5, 2nd Ed, Academic Press, London, pp. 129-168 (1993)
- 3) Shimoi, H., Sakamoto, K., Okuda, M., Atthi, R., Iwashita, K., Ito, K.: *Appl. Environ. Microbiol.*, **68**, 1932-1937 (2002)
- 4) Wu, H., Ito, K., Shimoi, H.: *Appl. Environ. Microbiol.*, **71**, 6845-6855 (2005)
- 5) Akao, T., Yashiro, I., Hosoyama, A., Kitagaki, H., Horikawa, H., Watanabe, D., Akada, R., Ando, Y., Harashima, S., Inoue, T., Inoue, Y., Kajiwara, S., Kitamoto, K., Kitamoto, N., Kobayashi, O., Kuhara, S., Masubuchi, T., Mizoguchi, H., Nakao, Y., Nakazato, A., Namise, M., Oba, T., Ogata, T., Ohta, A., Sato, M., Shibasaki, S., Takatsume, Y., Tanimoto, S., Tsuboi, H., Nishimura, A., Yoda, K., Ishikawa, T., Iwashita, K., Fujita, N., Shimoi, H.: *DNA res.* **18**, 423-434 (2011)
- 6) Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S. G.: *Science*, **274**, 546-567 (1996)
- 7) Wei, W., McCusker, J. H., Hyman, R. W., Jones, T., Ning, Y., Cao, Z., Gu, Z., Bruno, D., Miranda, M., Nguyen, M., Wilhelmy, J., Komp, C., Tamse, R., Wang, X., Jia, P., Luedi, P., Oefner, P. J., David, L., Dietrich, F. S., Li, Y., Davis, R. W. Steinmetz, L. M.: *Proc. Natl. Acad. Sci. USA*, **104**, 12825-12830 (2007)
- 8) Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavorry, F., Mallet, S., Cambon, B., Legras, J.L., Wincker, P., Casaregola, S., Dequin, S.: *Proc. Natl. Acad. Sci. USA*, **106**, 16333-16338 (2009)
- 9) Katou, T., Kitagaki, H., Akao, T., Shimoi, H.: *Yeast*, **25**, 799-807 (2008)
- 10) Hall, C. and Dietrich, F. S.: *Genetics*, **177**, 2293-2307 (2007)
- 11) Goto, K., Iwatsuki, Y., Kitano, K., Obata, T., Hara, S.: *Agric. Biol. Chem.*, **54**, 979-984 (1990)
- 12) 溝口晴彦, 藤田栄信: 醸協, **77**, 361-364 (1982)
- 13) Noguchi, C., Watanabe, D., Zhou, Y., Akao, T., Shimoi, H.: *Appl. Environ. Microbiol.*, **78**, 385-392 (2012)
- 14) 向井信彦, 岡田明彦, 鈴木昭紀, 高橋利郎: 醸協, **93**, 967-975 (1998)
- 15) Watanabe, D., Wu, H., Noguchi, C., Zhou, Y., Akao, T., Shimoi, H.: *Appl. Environ. Microbiol.*, **77**, 934-941 (2011)
- 16) Meiron, H., Nahon, E., Raveh, D.: *Curr. Genet.*, **28**, 367-373 (1995)
- 17) Tamura, K., Gu, Y., Wang, Q., Yamada, T., Ito, K., Shimoi, H.: *J. Biosci. Bioeng.*, **98**, 159-166 (2004)
- 18) Sharon, G., Burkett, T. J., Garfinkel, D. J.: *Mol. Cell. Biol.*, **14**, 6540-6551 (1994)
- 19) Azumi, M. and Goto-Yamamoto, N.: *Yeast*, **18**, 1145-1154 (2001)