

酵母菌的载体系统研究进展

张博润 何秀萍 陈玉梅

(中国科学院微生物研究所 北京 100080)

由于酵母菌特别是酿酒酵母是单细胞生物,具生长快、易于遗传操作、能对外源蛋白进行翻译后加工和修饰、不产生有毒产物等优点,被认为是表达外源蛋白的最适受体菌。随着重组DNA技术的发展,酵母菌的转化和载体系统的研究已取得很大的进展,已有许多细菌、真菌和高等动植物的基因在酿酒酵母中克隆和表达,已应用酵母工程菌来生产一些重要的外源蛋白,并显示出巨大的优越性。本文着重介绍酵母菌载体系统的研究进展。

根据其在酵母菌遗传工程中的应用,可将酵母菌载体系统分为普通的克隆载体和特殊用途的载体两大类。

1 酵母菌的克隆载体

依据酵母菌质粒载体的构成和复制方式,可分为整合载体(YIp)、附加体质粒载体(YEp)、复制载体(YRp)、着丝粒载体(YCp)、线性载体(YLp)和酵母菌的人工染色体(YAC)六类。

1.1 酵母菌整合载体(YIp)

因这类载体只含 *E. coli* 的复制起点,不能在酵母菌中自律复制,而是通过同源重组,以低频率整合到细胞染色体中^[1]。整合位点数取决于载体中的互补基因组序列数。它在基因组内通常以单拷贝存在,但也可能发生多位点整合。如果在YIp内插入一个不完全基因,则可用以创造基因破坏或突变。大部分YIp质粒含酵母菌选择标记(如HIS3、LEU2、TRP1及URA3等)。这种质粒的转化子很稳定,可在无选择条件下培养很多代而不丢失,但其转化频率很低($1 \sim 10 / \mu\text{gDNA}$),使YIp线性化可明显增加转化频率 $10 \sim 100$ 倍^[2]。

1.2 酵母菌附加体质粒载体(YEp)

这类质粒载体以酵母菌内源 $2\mu\text{m}$ 质粒为基础, $2\mu\text{m}$ 质粒是一种小分子双链环状DNA,存在于酵母属的大多数菌株中,每个细胞内有 $50 \sim 100$ 拷贝。它由2个各599bp的反相重复分隔成2个区域,含有4个可读框(ORFs)和一个复制起点。其中REP1和REP2编码的产物与细胞分裂时的质粒平均分配有关,所有YEp载体都含 $2\mu\text{m}$ 质粒的REP3位点,它是保证质粒稳定性必要

的因子;此外还含 $2\mu\text{m}$ 质粒的ori、选择标记和细菌的有关DNA序列。YEp载体的稳定性比较好,是酵母遗传工程中应用最广泛的载体系统,常用于酵母菌中的一般克隆和基因表达研究。

1.3 酵母菌复制载体(YRp)

YRp含有自律复制序列(ARS),因而可作为染色体外因子复制并保留下来。这种序列来自酵母菌的基因组,当然也可能来自其它生物。这种载体转化酵母的频率很高,可达 $10^3 \sim 10^5 / \mu\text{g DNA}$,且在细胞群体内拷贝数很高,但由于减数分裂和有丝分裂时的不稳定而使群体内的拷贝数变化很大,平均拷贝数为 $1 \sim 10 / \text{细胞}$ 。上述现象是由于细胞分裂时质粒的不对称分离造成的。YRp可用作一般克隆目的,但因其不稳定性 and 不对称分离而不适于基因调节机制的研究和应用。

1.4 酵母菌着丝粒载体(YCp)

该载体除含复制起点外,还含有酵母菌染色体的着丝粒(CEN),因而能在染色体外自我复制。在细胞分裂时新复制的质粒均等分离,每一个子细胞得到 $1 \sim 3$ 个质粒,故拷贝数很低,但很稳定^[3]。由于YCp的高稳定和低拷贝数,因而适合作亚克隆载体和构建酵母菌基因组DNA文库,可用于检测有丝分裂中染色体倍性变化和分析鉴定酵母菌基因突变^[4]。

1.5 酵母菌线性载体(YLp)

YLp除含自律复制所必需的因子外,还含端粒(TEL)和选择标记。这类载体能在酵母菌中复制和保留下来,因而已被用于一般克隆目的。不少种酵母菌含有杀伤因子(Killer),Fukuhara^[5]调查了800株酵母菌,发现28株含有线性质粒,其结构与乳酸克鲁维酵母的Killer质粒相似。乳酸克鲁维酵母含有2种双链DNA质粒(pGKL1和pGKL2),两者都和Killer毒素有关。pGKL1复制需有pGKL2存在,但pGKL2可单独复制,说明后者含有质粒复制所必需的DNA-和RNA-聚合酶^[6]。Schickel等^[7]构建成一系列重组质粒载体(pMS260~269)以及线性杂种质粒(pRKL150~159),

并证明这些线性质粒可转化并保留在不同属的酵母菌中。

1.6 酵母菌的人工染色体 (YAC)

YAC 是人工构建的具有酵母菌染色体功能的人工染色体载体。Burke 等^[8]构建了第一代 YAC 载体系统。YAC 除含 Sup4(*tRNA*^{Trp} 赉石校正基因)外,还含有染色体的其它功能元件:如 ARS1、CEN4、TEL 及选择标记 (TRP1 和 URA3)。目前已根据不同用途构建了一系列 YAC 载体,其各成员间的主要差别是克隆位点不同。YAC 可以环形式在大肠杆菌中繁殖,在转化酵母后变成线状微型染色体,并较稳定地保留下来。YAC 可将其它高等生物的染色体 DNA 大片断 (50~2000kb) 导入酵母菌并进行遗传操作,已用于克隆各种生物基因组大片断,已成为构建高等生物基因库的首选载体^[9]。

2 酵母菌的特殊质粒载体

2.1 酵母菌的表达载体 (YXp)

外源基因在酵母菌中表达要比在其它生物(如细菌、昆虫及高等动植物细胞培养物)中表达更为有利,因而特别受重视,并已成功地发展了一系列酵母菌表达载体。使外源基因表达的基本框架为:启动子-异源基因-终止子。大部分表达载体都以酵母菌的有效基因启动子为基础。酵母菌的启动子至少含有 3 个成分:上游激活序列 (UASs)、TATA 序列和起始子序列,有的还含有阻遏因子。UASs 类似于哺乳动物的增强子,它通过与转录激活因子(如 GAL4 和 GCN4)特异性结合并在离起始位点不同距离处决定启动子活性和调节作用。TATA (或 TATAA) 因子位于起始位点上游 40~120bp 处,提供发生起始的窗口,起始因子是很不确定的,它在相邻位点指导 mRNA 合成。此外,酵母启动子也可能很复杂,如含多个 UASs、调节位点和多个与起始位点结合的 TATA 因子。

2.1.1 酵母菌启动子的分类及特性:酵母启动子分两大类:第一类是来自编码解糖酶基因的解糖启动子,如 ADHI、PGK、GAP、ENO1 等。这些启动子在早期被看作是组成性的,能在酵母菌中高水平组成性表达,可用于克隆哺乳动物基因,但后来发现能被葡萄糖诱导,例如用 PGK 启动子表达 α -干扰素时,在醋酸盐作碳源的培养基中加葡萄糖可诱导 20~30 倍^[10]。这类启动子已广泛应用于实验室,也有用于工业的。第二类是强调节启动子,其中 GAL1、GAL7 和 GAL10 是半乳糖调节启动子,已用作真核生物转录调节的关键模式系统。

2.1.2 酵母菌的终止子:原核和高等真核基因的转录终止子通常在酵母菌中没有活性。有效终止转录是大量表达所必需的。酵母菌的转录终止子位于启动子和克隆

位点下游,象高等真核生物一样,形成 3' 末端聚腺苷化转录物,使 mRNA 转录终止。例如 CYC1 基因终止序列 3' 缺失产生较长的 mRNA,并明显降低 mRNA 水平。转录终止子的存在可使总信息量和蛋白表达量增加。许多基因的终止子,例如 TRP1 (pYE4)、ADHI (pAAH5)、GAP、MF α 1、CYC1 (pMAC561)、pGK (pMA91) 终止子等都已用于表达载体^[11]。

2.2 酵母菌的分泌载体 (YSp)

酵母菌的分泌载体中,以高拷贝 2 μ m 质粒为基础的载体最常用,但某些整合载体可产生比附加载体更高的分泌产物。Sakai 等^[12]比较了 20 拷贝 Ty δ 整合载体和 2 μ m 载体的人神经生长因子的分泌水平,发现 Ty δ 整合载体的分泌水平高 3~4 倍,达 3~4mg/L。酵母菌的分泌载体除必须含有表达载体的启动子和终止子外,还需有分泌信号序列,经常选用相同基因的启动子和分泌信号,最广泛应用的信号序列是 α 因子 (MF α 1) 的 prepro 区。此外,酵母菌还能利用某些外源蛋白自身的信号序列,例如 *E. coli* 的 β -内酰胺酶,人 α -和 β -干扰素等,但表达水平和分泌水平都很低。酵母可利用泡盛曲霉糖化酶、大麦 α -淀粉酶和人血清白蛋白的信号序列有效分泌。此外,可以通过分离筛选超分泌型的酵母菌株来改变特殊蛋白的分泌^[13]。

2.3 酵母菌的启动子载体 (YPp)

YPp 含有易分析和可记录的报告基因,可与含转录或转译启动子序列的启动子融合,用于控制基因表达调节序列的研究。最常用的报告基因为 *E. coli* 的 β -半乳糖苷酶基因 (lacZ)。它的优点是不需要特别的寄主菌株,缺点是酶为大的四聚体,存在胞液内,用 X-gal 检测需时较长。PHO5 基因也可用作报告基因,其主要优点是产生的酶易分泌至周质内。用 SUC2 基因作报告基因的优点是产物为胞外酶,容易进行分析。酵母的 URA3 报告基因系统比其它系统好,可用于分析启动子序列和融合蛋白。此外,细菌的 galK (半乳糖激酶基因)、抗生素抗性基因 cat (氯霉素乙酰转移酶基因) 亦可用作报告基因^[14]。

2.4 酵母菌的诱变质粒载体 (YMp)

近年来已发展了一些将转座子插入酵母 DNA 序列的诱变系统。每一系统都含有在酵母中选择的标记转座子。转座子的插入会造成基因结构的破坏而发生突变,在转化酵母后,可以通过表型分析确定突变的发生情况。例如 mini-Tn3^[15]和 Tn10-LUK 系统^[16]已用于克隆在质粒载体上的基因的转座诱变。mini-Tn3 系统含有 Amp^r 基因,酵母菌 DNA 被克隆到含 *E. coli* Km^r 基因的靶子载体如 pHSS4 或 pHSS6 中,转化酵母后可进行

表型研究。Tn10-LUK系统含 lacZ 报告基因,酵母 URA3 标记和 *E. coli* Km^r 基因,可用于分离和鉴定大部分标准多拷贝穿梭载体上的酵母基因内的 mini-Tn10-LUK 插入。mini-Tn10 转座子系统已用于噬菌体克隆的诱变和 gtl1 克隆内抗原编码序列的快速鉴定和作图^[17]。由于转座子含有酵母选择标记,可通过一步基因易位把靶基因的插入突变引入酵母基因组。

酵母菌的转座子 Ty 是一种分立的 DNA 因子,它可由供体位点移到靶子位点而本身并不从供体位点丧失。它的转座与其所在的基因组的复制无关。Ty 可调节基因组重组,包括倒位、缺失和转座。在 DEL1 菌株中, Ty-Ty 重组调节的缺失发生频率相当高。Garfinkel 等^[18]用 Ty 进行了酵母菌的转座诱变,他们构建了含 GAL1p-Ty 因子的多拷贝质粒(如 pGTyH3HIS3、pGTy917ND、pGTy917N1),都有体内转座和突变靶子序列的能力。除上述外,还有 hisG::URA3::hisG 系统、YRp14 系统和 YRp15 系统的诱变载体^[13]。

3. 其它酵母菌的载体系统

酿酒酵母没有强启动子,须流加分批培养获得高细胞密度,可产生超糖基化,因而被认为不是大规模生产外源蛋白的最好寄主,需要发展其它更好的受体载体系统,已报道的有粟酒裂殖酵母、乳酸克鲁维酵母、解脂耶鲁酵母、施旺酵母、巴斯德毕赤酵母和多型性汉逊酵母等系统,现分别介绍如下:

3.1 乳酸克鲁维酵母(*K. lactis*)

由于 *K. lactis* 能利用廉价底物如乳糖和乳清发酵,尤其适于作生产低值异源蛋白的受体菌,已用于生产多种人用蛋白。由于酿酒酵母的 ARS1 或 2 μ ori 都不能在这种菌中复制,故首先要分离 *K. lactis* 的 ARS 序列。但是由 *K. lactis* 的 ARS 构建的载体很不稳定,限制了它的应用。构建载体的选择标记有 TRP1、URA3、LAC4 以及 G418 抗性基因。*K. lactis* 菌株中存在的线性质粒 K1 (8.9kb) 和 K2 (13.4kb) 可用于构建载体。这些质粒拷贝数很高(100~200)且稳定。缺点是不能识别核启动子。例如由 pGKL1 构建的载体 pYGU1 具有广谱性,可在 *K. lactis*、*S. cerevisiae* 和 *S. pombe* 中复制。*K. drosophilarius* 的质粒 pKDI 在结构上类似于 2 μ 质粒,能在 *K. lactis* 中复制,已构建了以 pKDI 为基础的类似于以 2 μ 为基础的高拷贝表达载体^[19]。含整个 pKDI 质粒的载体在无内源质粒菌株(pKDI⁻)中很稳定,因而很适于表达外源基因。

3.2 粟酒裂殖酵母(*S. pombe*)

S. pombe 的表达载体通常含 2 μ 质粒或 *S. pombe* ars 衍生序列。在 *S. pombe* 中起自律复制功能的 2 μ 序

列不包括整个 2 μ ori,也不依赖 2 μ 的编码功能。这类载体的拷贝数较低(5~10/细胞)且有丝分裂不稳定。而 ars1 载体则具有较高拷贝数(30/细胞),加入 *S. pombe* 的 stb 序列可大大提高载体的稳定性和拷贝数。例如 ars1/stb 载体 pFL20 的拷贝数可达 80/细胞,每代丢失率由 30~45% 降到 13%,高频率转化载体 pDB248 及其衍生载体对构建 *S. pombe* 基因库非常有用。

3.3 巴斯德毕赤酵母(*P. pastoris*)

P. pastoris 已用于表达多种外源蛋白。由于该菌没有天然质粒,所以设计表达载体偏向于染色体整合。已经构建了一些含 *P. pastoris* ars 序列: PARS1 和 PARS2 的多拷贝附加载体。编码酒精氧化酶的 AOX1 基因启动子能强力而有效地调节外源基因的表达,已用于生产多种外源蛋白。常用的选择标记为 HIS4。此外,由于该菌不能利用蔗糖,亦可用 *S. cerevisiae* 的 SUC2 基因作选择标记。

许多因子可影响 *P. pastoris* 整合载体的表达水平,首先是 AOX1 的破坏,当诱导 aox1 时,转化子不能同时高水平产生酒精氧化酶和外源蛋白,与野生型 Mut⁺ 比较,生长也较慢,对氧的要求亦较低。其次是整合拷贝数的影响,利用多拷贝整合可以增加外源蛋白的产量。第三是基因和其产物的性质。第四是分泌水平。已经证明 *P. pastoris* 能高水平分泌牛溶菌酶(0.55g/L)、HSA(3g/L)、aprotinin(0.93g/L)、hEGF(0.4g/L)、mEGF(0.45g/L)以及 β -葡聚糖酶等^[20]。分泌信号可用 *S. cerevisiae* 的 α -因子的 prepro 前导序列。利用 *P. pastoris* 系统的最大优点是便于进行高密度培养和放大培养规模。

3.4 多形汉逊酵母(*H. polymorpha*)

多形汉逊酵母编码甲醇氧化酶(MOX)基因的启动子已用于外源基因表达。这个基因也是强调节基因,在葡萄糖限制或缺乏条件下,基因表达呈明显解阻遏。常用的选择标记为 *S. cerevisiae* 的 LEU2 和 URA3。含 LEU2 的载体如 YEp13,转化频率和有丝分裂稳定性都低,说明与 *S. cerevisiae* 同源性不高。由于不能识别 2 μ ori,某些细菌质粒序列可能促进自律复制;含 URA3 的载体如 YIp5 也是低转化率的,转化子也不稳定,但在有染色体 DNA 序列 HARS1 和 HARS2 时,可明显增加转化频率。在非选择性培养基上延长培养可以分离到含多拷贝串联整合的稳定转化子。这个系统已用于有效表达含 preS2 序列的 HBsAg 颗粒,其中半数分泌至周质,用 β -葡聚糖诱导细胞时,这些颗粒可分泌至培养基中。Janowicz 等^[21]用甲醇调节甲酸脱氢酶基

因 FMX 和 MOX 基因启动子与 preS1-S2HBsAg 和 HBsAg 共表达, 用 URA3 及 G418 为选择标记, 得到了混合颗粒。Rodrigues 等^[22]将 *S. cerevisiae* 的 SUC2 基因克隆到 *H. polymorpha* 中, 在 AOX1 启动子控制下, 能分泌蔗糖酶至周质, 在流加分批发酵中, 得到 1.5×10^3 u / ml 酶活, 分泌的酶是糖蛋白。与 *P. pastoris* 一样, *H. polymorpha* 可以高密度培养 (100~300g / L), 可以得到很高的单位体积产量。

3.5 解脂耶氏酵母 (*Y. lipolytica*)

解脂耶氏酵母是一种重要工业菌株, 用于多种产品的生产。Fournier 等^[23]分离到 *Y. lipolytica* 的 ARS 成分 ARS18 和 ARS68。含上述成分的载体只有 1~3 拷贝 / 细胞, 但很稳定。Nicaud 等^[24]报道了用 ARS18 为基础的载体使 *Y. lipolytica* 分泌 α -干扰素终产量增加 2~3 倍。Nicaud 等用 SUC2 基因作选择标记, 使它与碱性胞外蛋白酶 (AEP) 基因 XPR2 启动子和分泌信号融合, 转化后可在蔗糖培养基上直接选择转化子。XPR2 是受 pH 和碳、氮源调节的启动子, 它的 prepro 区是充分分泌 AEP 所必需的成分, 已用于指定牛凝乳酶原和猪干扰素分泌。

本文简要综述了酵母菌载体系统近年来的研究进展, 简要总结了酵母菌载体系统构建、结构、应用和存在的问题, 讨论了受体菌株和外界条件对外源基因表达的影响。随着对酵母菌载体系统的深入研究, 必将建立更好更多的酵母载体系统。

参 考 文 献

- [1] Hinnen A, Hicks JB, Fink GR. Proc Natl Acad Sci, USA, 1978, 75:1929.
- [2] Stearns T, Ma H, Botstein D. Methods in Enzymology, 1990, 185:280.
- [3] Hartwell LH, Dutcher SK, Wood JS, et al. Recent Advances Yeast Molecular Biology, 1982, 1:28.
- [4] Rose MD, Novick P, Thomas JH, et al. Gene, 1987, 60:237.
- [5] Fukuhara H. FEMS Microbiol Lett, 1995, 131(1):1~9.
- [6] Tommasino M, Ricci S, Galeotti CL. Nucl Acids Res, 1988, 16:5863~5878.
- [7] Schickel J, Helmg C, Meinhardt F. Nucl Acids Res, 1996, 24(10):1879~1886.
- [8] Burke DT, Carle GF, Olson MV. Science, 1987, 236: 806~812.
- [9] Larin Z, Taylor SS, Smith CT. Nucl Acids Res, 1996, 24(21):4192~4196.
- [10] Tuite MF, Dobson MJ, Roberts NA. EMBO J., 1982, 1:603~608.
- [11] Romanos MA, Scorer CA, Clare SJ. Yeast, 1992, 8: 423~488.
- [12] Sakai A, Ozawa F, Higashizaki T, et al. Bio / Technol, 1991, 9:1382~1385.
- [13] Parent SA, Bostian KA. In the Yeast vol6, Yeast Genetics, 2nd ed. Rose AH, Wheals AE, Harrison JS(eds), 1995, pp.121~178, London.
- [14] Altherr MR, Rodriguez RL, in "Vectors: A Survey of Molecular Cloning Vectors and Their Uses"(Rodriguez RL, Denhardt DT, (eds), 1988, pp.405~417, Butterworths, Boston.
- [15] Seifert HS, Chen EY, Heffron F. Proc Natl Acad Sci USA, 1986, 83:735.
- [16] Huisman O, Raymond W, Froehlich K-U, et al. Genetics, 1987, 116:191.
- [17] Snyder M, Elledge S, Sweetser D, et al. Methods in Enzymology, 1987, 154:107.
- [18] Garfinkel DJ, Strathern JN. Methods in Enzymology, 1991, 194:342.
- [19] Bianchi MM, Falcone C, Jie CX, et al. Curr Genet, 1987, 12:185~192.
- [20] Roca H, Garcia B, Rodriguez E, et al. Yeast, 1996, 12:1187~1200.
- [21] Janowicz ZA, Melber K, Merckelbach A, et al. Yeast, 1991, 7:431~443.
- [22] Rodriguez L, Narciandi RE, Roca H, et al. Yeast, 1996, 12:815~822.
- [23] Fournier P, Guyaneux L, Charles M, et al. Yeast, 1991, 7:25~36.
- [24] Nicaud JM, Fournier P, La BC, et al. J Biotechnol, 1991, 10:259~270.