

甲醇酵母表达系统高拷贝数整合型表达载体的构建

李黄金 龙繁新 林志雄* 涂桂洪* 王珣章

(中山大学生物防治国家重点实验室 广州 510275)

(广州华生基因工程有限公司 广州 510275)*

摘 要 以甲醇酵母(*Hansenula polymorpha*)表达系统的 YIp 型表达载体 pJF5-1 为基础,通过引入宿主来源的一个 rDNA 随机顺序和由 SV40 早期启动子控制的 G418 抗性基因(*neo*)及对野生型标记基因 *Hpleu₂* 启动子大部分上游序列的缺失,构建了一个高拷贝整合型表达载体 pMIRH。有关转化、筛选和拷贝分析等试验结果表明,由 pMIRH 可产生含高拷贝数载体的转化子,并可通过由 *leu₂⁺* 筛选和 G418 抗性筛选所组成的二级筛选方法富集这些含高拷贝数载体的转化子。有关完整 DNA 和消化 DNA Southern 杂交分析结果进一步表明,上述高拷贝数的载体以串联重复排列的方式整合于宿主基因组。

关键词 *Hansenula polymorpha* 表达载体 高拷贝数整合 筛选 串联重复

学科分类号 Q 789

甲醇酵母是指可利用甲醇作为单一碳源的一类酵母,其中的 *H. polymorpha*、*Pichia pastoris* 和 *Candida bodinii* 等已被成功地发展成为外源基因表达系统^[1]。该类系统利用了非常强的甲醇氧化酶基因(Methanol Oxidase, MOX 或 Alcohol Oxidase, AOX)的启动子,且具有高密度发酵和分泌表达的能力,比传统的酿酒酵母(*Saccharomyces cerevisiae*)系统更具产业化发展潜力。由于没有稳定的自主复制型表达载体,所以目前在甲醇酵母系统中一般采用单拷贝或低拷贝数整合的 YIp 型载体。尽管单拷贝的野生型 MOX 或 AOX 可被诱导表达至占细胞总蛋白的 30% 以上,但当由 MOX 或 AOX 启动子驱动外源基因表达时则需要多拷贝的表达盒方可获得“超表达”^[2,3]。引入重复顺序是提高 YIp 型载体整合拷贝数的有效途径之一,Lopes 等^[4]曾通过引入核糖体 DNA(Ribosomal DNA, rDNA)这一重复顺序成功地酿酒酵母系统构建了可高拷贝数整合的 pMIRY(Multiple integration into rDNA of yeast)型载体,但类似的构建在甲醇酵母系统中尚未见报道。本文报道了 *H. polymorpha* 表达系统 pMIRY 型表达载体的构建工作。

1 材料与方法

1.1 材料

1.1.1 菌株与质粒: *H. polymorpha* A16(*leu₂*)及 pJF5-1 由 P. Sudbery 教授惠赠,含 *H. wiginii* 全长 rDNA 的 pHWR1 由 J. Venema 博士赠送。YIp5、pUC18 和受体菌 *E. coli* DH5 α 由本室保存。pBluescript-d, 含 *neo* 基因的 pUC-*neo*-d 及含 MOX5' 端序列的 MOX5'-pUC18 由本室构建。

1.2 方法

1.2.1 酵母总 DNA 制备：参照 Philippsen 等^[5]方法进行。

1.2.2 DNA Southern 杂交分析：参照 Sambrook 等^[6]方法进行。以地高辛标记与检测试剂盒(Boehringer mannheim 产品)进行探针标记和杂交检测(方法参照有关厂家说明进行)。

1.2.3 随机缺失：用嵌套缺失试剂盒(Boehringer mannheim 产品)参照厂家说明进行。

1.2.4 酵母转化：参照 Faber 等^[7]方法进行电激法转化。

1.2.5 酵母转化子 G418 抗性筛选：参照 Scorer 等^[8]方法进行。

1.2.6 酵母转化子 DNA 点杂交分析：参照 Clare 等^[9]方法进行。用上述地高辛标记与检测试剂盒进行探针标记与检测。

1.2.7 载体整合拷贝数测定：参照 Clare 等^[10]方法进行。

1.2.8 DNA 重组及其它分子生物学操作：参照 Sambrook 等^[6]方法进行。

2 结果与讨论

2.1 rDNA 随机顺序克隆

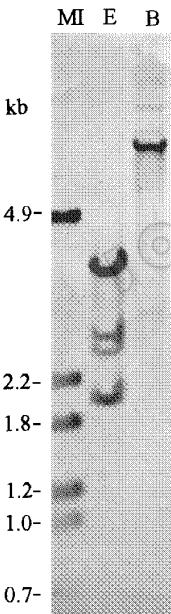


图 1 *H. polymorpha* A16(*leu2*) 总 DNA Southern 杂交图谱

Fig. 1 Southern spectra of total DNA from *H. polymorpha* A16(*leu2*)

MI. pBR328/BamHI + BglII + Hinf ;
E. EcoRI ; B. BglII

每个酵母基因组含有 100~120 拷贝左右呈串联重复排列的 rDNA 单元^[5],所以 rDNA 顺序是构建高拷贝数整合型载体较为理想的重复顺序。虽然 rDNA 是较为保守的序列,但考虑到宿主基因组来源的 rDNA 顺序更利于载体的稳定整合,所以我们从宿主基因组克隆了一个 rDNA 随机顺序。首先以限制酶 EcoRI 和 BglII 分别对 *H. polymorpha* 基因组 DNA 进行酶切,然后以含 *H. wiginii* 全长 rDNA 单元的质粒 pHWR1 为探针与之进行 Southern 杂交。结果发现,*H. polymorpha* 基因组的 rDNA 可被 BglII 切割成 1 个 9 kb 大小的可分离片段,而被 EcoRI 切割成 4 个大小分别为 4.0kb、2.7kb、2.5kb 和 2.0kb 左右的片段(图 1),其中 2.7kb 和 2.5kb 片段的杂交信号相对较弱,可能是由位于 2.7kb 片段上的一个较难切动的 EcoRI 位点不完全切割所造成的。我们对以较为清晰的条带呈现的 4kb EcoRI 片段进行了回收,并将其克隆于 pUC18 得到重组子 R-pUC18 Southern 杂交分析证实重组子 R-pUC18 插入有上述 4kb rDNA 片段。由于每个 rDNA 单元长达 9 kb 左右,由 25S、18S、5.8S 和 5S rRNA 的转录区及一些非转录区组成^[5],所以引入 rDNA 单元中哪个区段的顺序及引入多长的顺序对载体构建似乎至关重要。但关于 pMIRY 型载体整合机制研究^[11]的结果表明,高拷贝数整合是载体在宿主 rDNA 区域少数几个位点上大量“堆积”的结果。

被引入载体中的 rDNA 序列无论来自 rDNA 单元的哪个区段,只要其长度足够满足同源重组的需要,即可介导这种载体“堆积”,且不会严重影响宿主的生长。因此,我们认为上述从宿主基因组克隆的 rDNA 随机顺序应可用于载体构建。

2.2 载体构建

H. polymorpha 高拷贝数整合型表达载体 pMIRH 的构建及其主要结构如图 2 所示。

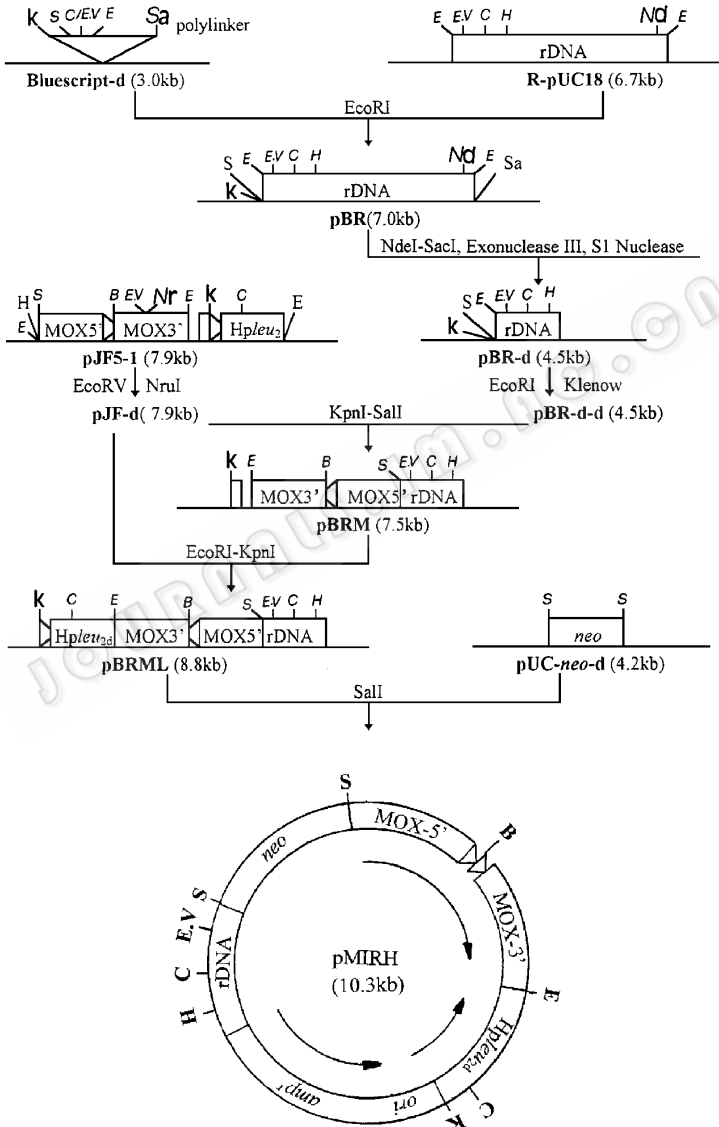


图 2 *H. polymorpha* A16 (*leu2*) 高拷贝数整合型表达载体 pMIRH 的构建

Fig. 2 Construction of pMIRH for multiple integration into chromosome of *H. polymorpha* A16 (*leu2*)
Abbreviations: B. BglII; C. ClaI; E. EcoRI; E. V. EcoRV; H. HindIII; K. KpnI; Nd. NdeI; Nr. NruI; S. SalI; Sa. SacI, respectively

其中 Bluescript-d、R-pUC18、pJF5-1 和 pUC-*neo*-d 等 4 个为出发质粒。Bluescript-d 来自 Bluescript M13(+), 其多克隆位点处的 ClaI、EcoRV 和 HindIII 等位点已被消除。R-pUC18 为上述含 4kb rDNA 片段的质粒, 对 4kb rDNA 片段的限制酶识别位点分析结果表明 除了两端的 EcoRI 位点外, 该片段还含有 EcoRV、ClaI、HindIII 和 NdeI 的单一切点, 而位于 HindIII 至 NdeI 之间长达 2.5kb 的序列中未发现常见的限制酶识别位点。pJF5-1 为 YIp 型表达载体, 含有一个由 MOX 的上、下游序列组成的表达盒(克隆位点为 BglIII) 和一个野生型标记基因 *Hpleu₂*。pUC-*neo*-d 含有一个由 SV40 早期启动子控制的 G418 抗性标记基因, 位于启动子末端的一个 HindIII 位点已被消除。构建成的表达载体 pMIRH 除含有 MOX 表达盒外, 还包含了供高拷贝数整合和转化子筛选用的 3 个关键元件, 即 rDNA 顺序、缺陷型的标记基因 *Hpleu_{2d}* 和 G418 抗性基因 *neo*。为了尽量减小 rDNA 片段在最终载体中的长度, 我们将 4kb rDNA 片段从 R-pUC18 中转移至 Bluescript-d 后, 用随机缺失法将 NdeI 端的 2.5kb 左右的序列缺失掉。在长约 1.5kb 的 rDNA 保留序列中, EcoRV 和 HindIII 位点离其最近的 2 个末端均约为 0.3kb 左右, 可用于载体线性化, 以使载体定向整合至宿主基因组 rDNA 位点。除了 rDNA 随机顺序外, 缺陷型标记基因(Deficient marker gene) 是 pMIRY 型载体实现高拷贝数整合所需的 2 个关键元件之一, 其作用在于为载体在宿主基因组 rDNA 位点大量“堆积”提供一种强大的、来自细胞内部的选择压力^[11]。本文构建的 pMIRH 的缺陷型标记基因 *Hpleu_{2d}* 来自 pJF5-1 中的野生型标记基因 *Hpleu₂*, *Hpleu₂* 的上游序列全长原为 124bp, 在 -32 位处有一个 KpnI 识别位点, 经图中所示的重组步骤后, -124 至 -32 间的序列被缺失, 以满足上述 pMIRY 型载体高拷贝数整合对标记基因的要求。以 pMIRH 和 pJF5-1 的 BglIII 位点线性化 DNA 分别转化宿主 *H. polymorpha* A16(*leu2*) 的实验表明, pMIRH 可有效转化宿主, 但转化频率比 pJF5-1 的约低 20 倍左右(25/ μ g DNA 对 510/ μ g DNA)。考虑到上述转化试验中的 pMIRH 和 pJF5-1 均是向宿主基因组的同一位点(MOX 位点) 整合的, 而二者的大小差异(分别为 7.9kb 和 10.3kb) 不至于引起如此大的转化频率差异, 所以对上述转化试验结果的唯一合理解释只能是, pMIRH 中的 *Hpleu_{2d}* 已成为启动子功能高度缺陷的、但仍具有一定表达能力的标记基因。

除了上述 2 个供载体高拷贝数整合用的关键元件外, 我们亦为最终载体引入了 G418 抗性基因 *neo*, 以便建立高拷贝整合子的简易筛选方法。

2.3 转化与筛选

为了进一步考察 pMIRH 是否可引起高拷贝数整合及 G418 抗性筛选的可行性, 我们进行了一系列的转化与筛选试验。考虑到高拷贝数整合需要较高浓度的载体, 我们以 4 μ g HindIII 位点线性化 pMIRH 转化宿主(向基因组的 rDNA 位点定向整合), 并获得 98 个 *leu2* 阳性转化子。将这些转化子混合在一起并适当稀释后, 铺于多个含不同浓度 G418 的 YEPD 平板上。3d 后在多个高于宿主最低敏感浓度(75 μ g/ml) 的 G418 平板上均有菌落出现, 且随着 G418 浓度的增加, 菌落数目逐渐减少(照片略)。这一结果表明, 由 SV40 早期启动子控制的 *neo* 基因在 *H. polymorpha* A16(*leu2*) 中能有效表达, 且上述转化可能产生了高拷贝数整合子。我们以含 *ura3* 和 *amp^r* 序列的 YIp5 作探针, 以宿主菌(基因组含单拷贝 *ura3*) 为对照, 对从 75 μ g/ml 和 5.0 mg/ml G418 平板上随机挑选的

多个转化子进行了 DNA 点杂交分析(图 3)。从 75 μ g/ml G418 平板所获的 22 个转化子中仅有少数几个呈现了较强的杂交信号,而从 5.0mg/ml G418 平板所获的转化子则几乎全部呈现出强信号,其中只有 1 个转化子呈现了较弱的杂交信号,这可能是点杂交分析前的非选择培养引起载体丢失所致。我们从上述来自 5.0mg/ml G418 平板的、具有较强 DNA 点杂交信号的转化子中随机挑选了 8 个利用 Clare 等⁹的方法进行载体拷贝数测定(表 1),结果发现各转化子所含载体均在 50 拷贝以上,最高的达到 100 拷贝左右。

表 1 高 G418 抗性转化子所含 pMIRH 的拷贝数

Table 1 Copy number of pMIRH in transformants with high G418-resistance

Transformant	H03	H08	H11	H13	H15	H16	H19	H22
Copy number	58	95	76	106	50	92	88	96

2.4 转化子 Southern 杂交分析

为了进一步考察载体在宿主中的存在状态,我们对上述表 1 中含高拷贝数载体的转化子 H11 和 H13 及上述来自高浓度 G418 平板的、但在点杂交分析中呈现较弱杂交信号的转化子 S15 的总 DNA,分别进行了完整(未经消化)DNA 及消化 DNA 的 Southern 杂交分析(图 4)。在完整 DNA 的 Southern 杂交分析(图 4A)中,以完整的宿主 A16 总 DNA 和完整载体 pMIRH 分别作阴、阳性对照,载体检测探针为 pUC18。结果显示,在较高分子量(大于 21.0kb)的完整总 DNA 的相应位置,转化子 H11 和 H13 的总 DNA 样品均呈现了较强的阳性杂交信号,而转化子 S15 及宿主 A16 的则未见显著阳性;在较低分子量(小于 21.0kb)处,上述转化子和宿主 A16 的总 DNA 样品均呈现了一些较为明显的杂交信号,这些杂交信号显然属于假阳性信号,因为宿主的 DNA 样品亦呈现了这些信号,这可能是由与 pUC18 或分子量标记探针 DNA 间具有一定同源性的序列所造成的。由于完整总 DNA 在印迹时的转膜效率较低,而较低分子量 DNA 的转膜效率则几近 100%,所以这些假阳性信号的相对强度不能真实反映其所代表的 DNA 在总 DNA 中的相对含量。因此,上述分析结果可以表明,转化子 H11 和 H13 含有高拷贝数的载体,且这些载体是整合于宿主基因组的,而经高浓度 G418 筛选的转化子 S15 所含载体可能在非选择培养(供总 DNA 制备)时被丢失,暗示并非所有整合子均可在非选择条件下稳定传代。图 4B 为上述宿主和转化子的总 DNA 经 HindIII 消化后的 Southern 杂交分析结果,其中以带有 MOX-5'端序列的 MOX5'-pUC18 作探针来同时检测载体和宿主基因组中的单拷贝基因 MOX,以便进一步进行整合载体拷贝数的测定。在线性化质粒 pMIRH 的相应位置(10.3kb),转化子 H11 和 H13 的 DNA 样品均呈现了极强的阳性杂交信号,而宿主和转化子 S15 的则只有很弱的信号,在 3.8kb 位置,上述宿主和转化子的 DNA 样品均呈现了较弱的,但比宿主 A16 和

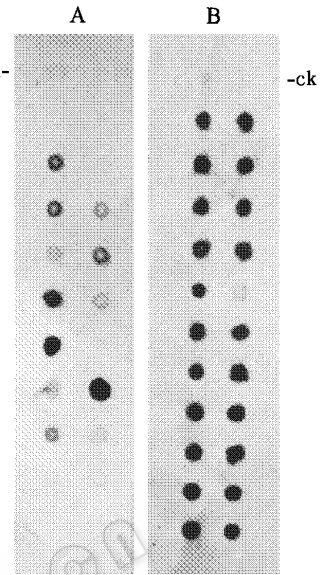


图 3 *H. polymorpha* A16(*leu2*) 转化子 DNA 点杂交分析

Fig. 3 Dot hybridization of DNA from transformants of *H. polymorpha* A16(*leu2*) A. Transformants from 75 μ g/ml G418 plate; B. Transformants from 5.0mg/ml G418 plate; ck. Host cells

转化子 S15 的 10.3kb 带信号更强的信号带。根据上述信号带的大小及其信号的相对强度可以知道,上述出现在转化子 H11 和 H13 DNA 样品中的 10.3kb 极强信号带为大量整合化载体的 HindIII 消化产物,在宿主 A16 和转化子 S15 DNA 的相应位置所呈现的较弱杂交信号属于假阳性结果,而在上述宿主和转化子 DNA 中均呈显著阳性的 3.8kb 带则为基因组单拷贝 MOX 的 HindIII 消化产物。我们对上述转化子 H11 和 H13 DNA 的整合化载体信号带和 MOX 信号带进行了光密度扫描分析(结果未呈),根据同一样品中 2 条杂交带的光密度之比推算出转化子 H11 和 H13 所含载体的拷贝数分别为 81 和 98,由于目前尚缺乏同源顺序长度与杂交信号强度相关性方面的资料,所以在拷贝数计算中没有排除因整合化载体(10.3kb)和基因组 MOX(3.8kb)间的长度差异所引起的误差,而使得拷贝数计算结果可能比实际值偏高,但所得拷贝数与上述点杂交分析法测得的结果相近。由于载体 pMIRH 中只含有 HindIII 的单一切点,而上述转化子 H11 和 H13 DNA 的杂交图谱中均只有一条均一的载体杂交带,所以这些整合载体应该是以头-尾相接的方式串联重复排列于宿主基因组的(图 4C)。该推断与 *S. cerevisiae*^[11]、*Candida utilis*^[12]、*Kluyveromyces lactis*^[13]及 *Phaffia rhodozyma*^[14]等酵母系统的类似研究结果相吻合。

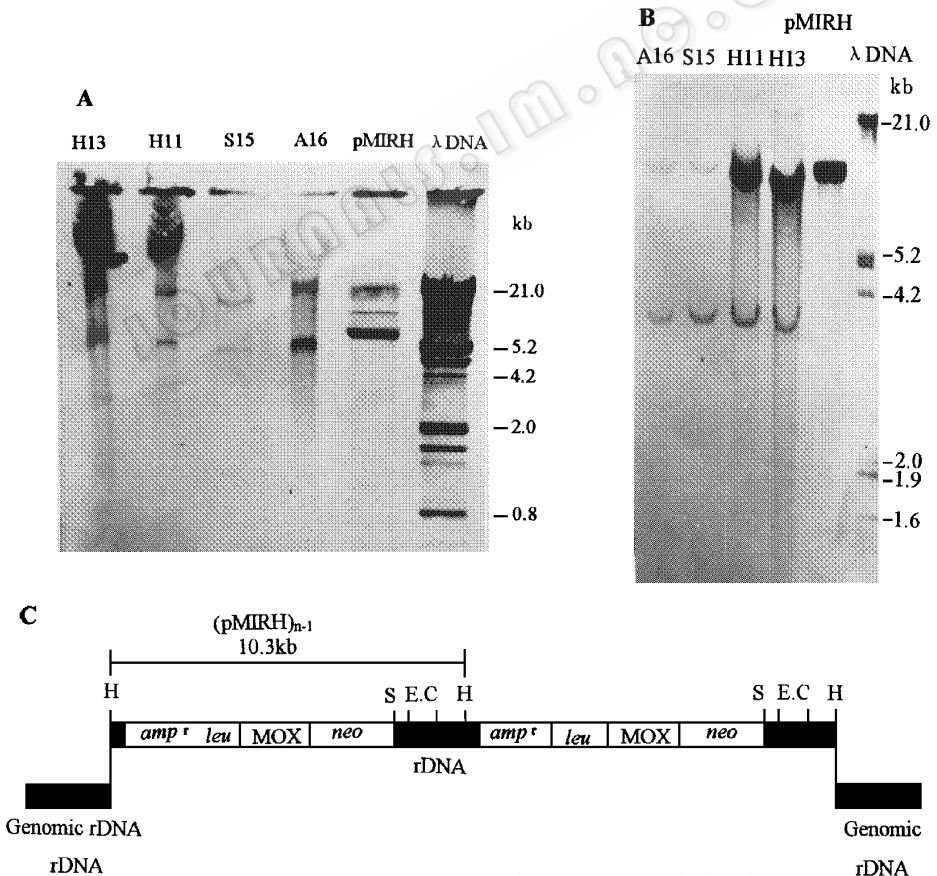


图 4 *H. polymorpha* A16(*leu2*)转化子 Southern 杂交分析

Fig. 4 Southern analysis on the transformants of *H. polymorpha* A16(*leu2*)

Part A. Spectra of intact DNA; Part B. Spectra of DNA digested by HindIII; Part C. The predicted arrange-

ment of the vector DNAs integrated at rDNA locus; C, ClaI; E, EcoRV; H, HindIII; S, SalI

致 谢 P. Sudbery 教授与 J. Venema 博士为本研究提供了关键实验材料和有益的建议, 在此一并致谢。

参 考 文 献

- 1 Sudbery P E , Gleeson M A G . Molecular and Cell Biology of Yeasts , London : Blackie Press , 1989 , 304~325
- 2 Clare J J , Romanos M A , Rayment F B *et al.* . Gene , 1991 , **105** : 205~212
- 3 Clare J J , Rayment F B , Ballantine S P *et al.* . Bio/Technology , 1991 , **9** : 455~462
- 4 Lopes T S , Klootwijk J , Veenstra A E *et al.* . Gene , 1989 , **79** : 199~206
- 5 Philippsen P , Stotz A , Scherf C . Methods in Enzymology , San Diego : Academic Press , 1991 , **194** : 169~181
- 6 Sambrook J , Fritsch E F , Maniatis T . Molecular Cloning , A Laboratory Manual , 2nd Ed , Cold Spring Harbor : Cold Spring Harbor Press , 1989
- 7 Faber K N , Haima P , Harder W *et al.* . Current Genetics , 1994 , **25** : 305~310
- 8 Scorer C S , Clare J J , McComble W R *et al.* . Bio/Technology , 1994 , **12** : 181~184
- 9 Clare J J , Romanos M A , Rayment F B *et al.* . Gene , 1991 , **105** : 205~212
- 10 Clare J J , Rayment F B , Ballantine S P *et al.* . Bio/Technology , 1991 , **9** : 455~462
- 11 Lopes T S , Hakkaart G A J , Koerts B L *et al.* . Gene , 1991 , **105** : 83~90
- 12 Kondo K , Saito T , Kajiwara S *et al.* . Journal of Bacteriology , 1995 , **177** : 7171~7177
- 13 Rossolini G M , Riccio ML , Gallo E *et al.* . Gene , 1992 , **119** : 75~81
- 14 Wery J , Gutker D , Renniers ACHM *et al.* . Gene , 1997 , **184** : 89~97

Construction of Expression Vector for High-copy-number Integration of the Methylophilic Yeast *Hansenula polymorpha*

Li Huangjin Long Qingxin Lin Zhixiong* Tu Guihong* Wang Xunzhang
(State Key Laboratory for Biological Control , Zhongshan University , Guangzhou 510275)
(Guangzhou Watson Genetech . Ltd . , Guangzhou 510275)*

Abstract Based on the YIp-type expression vector pJF5-1 of the methylophilic yeast *Hansenula polymorpha* , a novel vector , pMIRH , for high-copy-number integration was constructed , which contains a random sequence from the ribosomal DNA of the host chromosome , a deficient marker gene *Hpleu_{2d}* derived from the wild gene with a partial upstream sequence deleted and a G418 resistance gene(*neo*) controlled by the SV40 early promoter. Results from experiments of transformation , screening and Dot and Southern hybridization showed that pMIRH can result in transformants with high-copy-number vectors , which can be screened by a secondary G418-resistance selection , and these vectors were integrated into the chromosome in a tandem arrangement.

Key words Methylophilic yeast , expression vector , high-copy-number integration , screening , tandem repeat