论著

文章编号:1000-5404(2006)13-1384-04

菌丝相白假丝酵母菌长片段基因表达系列分析文库的构建及结果分析

周村建,邓永键,郝 飞,钟白玉,王 鲁,李永平 (第三军医大学西南医院皮肤科,重庆市皮肤性病研究所,重庆 400038)

提 要:目的 构建菌丝相白假丝酵母菌长片段基因表达系列分析文库并探讨其基因表达特点。**方法** 采用长片段基因表达系列分析技术(LongSAGE)构建菌丝相白假丝酵母菌LongSAGE文库,从文库中挑取阳性克隆测序,用SAGEsoftware对测序结果进行分析和检索,从而了解菌丝相白假丝酵母菌基因表达情况。**结果** 在菌丝相白假丝酵母菌的1 000个测序文件中,共获得标签17 552个,代表单一转录本6 627种,其中单一匹配标签占20.5%,多重匹配标签占6.64%,假设蛋白标签为18.47%,新标签为53.07%。标签拷贝数超过35拷贝的36个高表达标签中有15个可能属于新的表达序列。**结论** 用LongSAGE技术获得了菌丝相白假丝酵母菌基因表达谱及其丰度的量化信息,可以发现新的表达序列,为进一步研究提供了基础。

关键词: 白假丝酵母菌;二态性;长片段基因表达系列分析
中图法分类号: Q755; Q786; R379 **文献标识码:** A

Long serial analysis of gene expression in hyphal cells of *Candida albicans*

ZHOU Cun-jian, DENG Yong-jian, HAO Fei, ZHONG Bai-yu, WANG Lu, LI Yong-ping (Department of Dermatology, Southwest Hospital, Third Military Medical University, Chongqing 400038, China)

Abstract: Objective To observe the characteristics of gene expression in the hyphal cells of *Candida albicans*. Methods A LongSAGE library was established with long serial analysis of gene expression (LongSAGE). After the clones from the LongSAGE library were sequenced, the results were analyzed by SAGE software and compared with the SAGE map at NCBI. Results A total of 17 552 tags, representing 6 627 genes or EST, were extracted from the 1 000 sequence files of the hyphal cells of *Candida albicans*. Fifteen of thirty-six tags which expression frequency exceeded thirty-five may be new expressed sequences. Conclusion LongSAGE technology can be used not only in understanding the quantity information of gene expression, but also in finding new genes.

Key words: *Candida albicans*; dimorphism; long serial analysis of gene expression

白假丝酵母菌是重要的条件致病真菌,对其毒力因子的研究集中于3个方面,即黏附、分泌酶类和菌相转换[1]。白假丝酵母菌,转向菌丝相生长是其感染转变的过程,菌丝多则感染率高,有研究证实天然突变或人工中断某些基因导致白假丝酵母菌不能形成菌丝时,则其在动物实验中表现出毒力的降低或丧失[2],白假丝酵母菌菌丝的形成与发病有密切的关系[3]。

在前期获得白假丝酵母菌酵母相细胞LongSAGE标签库的基础上,为更全面的寻找白假丝酵母菌菌丝相细胞相关的致病基因及与酵母相间的差异表达基因,本研究应用LongSAGE技术,构建了白假丝酵母菌菌丝相细胞的LongSAGE标签库,通过克隆测序和初步数据分析,定性定量检测了白假丝酵母菌菌丝相细胞的表达基因谱。

1 材料与方法

1.1 菌株

白假丝酵母菌标准菌株ATCC90028,为本科室保存菌株。

1.2 主要试剂

RNA缓冲液,含0.5 mol/L NaCl、200 mmol/L Tris · Cl、10 mmol/L EDTA,pH调至7.5;0.45~0.55 mm酸洗玻璃珠;RPMI1640培养基,按说明书配制,将pH值调至7.0;I-SAGE Long kit(Invitrogen公司);DNA polymerases(Takara公司)。

基金项目:国家自然科学基金资助项目(30200013)
Supported by the National Natural Science Foundation of China (30200013)
作者简介:周村建(1970-),男,重庆市梁平县人,博士,主治医师,讲师,主要从事医学真菌方面的研究。电话:(023)68754416, E-mail: zhoucj@sina.com.cn
通讯作者:郝 飞,电话:(023)68754416, E-mail: haofei@mail.tmmu.com.cn
收稿日期:2005-06-11;修回日期:2006-03-24

1.3　菌丝相白假丝酵母菌诱导培养

参照文献[4]进行。将 YPD 斜面生长 48 h 的 ATCC90028 转种于 YPD 液体培养基,细胞调至 5.0×10^6/ml,置摇床25 ℃、150 r/min 振荡培养,48 h 后移至离心管,3 000 r/min 离心 15 min,弃上清,将细胞转种入 pH 7.0 的 RPMI1640 培养基,细胞浓度为 5.0×10^6/ml,置 37 ℃、150 r/min 振荡培养,3 h显微镜下观察超过 95% 的细胞被诱导形成芽管。

1.4　菌丝相白假丝酵母菌总 RNA 提取

以玻璃珠法制备总 RNA,RNA 缓冲液重悬白假丝酵母菌芽管,加入酸洗玻璃珠及酚/氯仿/异戊醇,拧紧瓶盖,高速振荡 5 min,室温离心,取上清液,再加入酚/氯仿/异戊醇抽提 1 次,最后用乙醇沉淀、洗涤 RNA,DEPC 水重溶,－20 ℃贮存,总 RNA 的完整性用 1.2% 琼脂糖凝胶电泳进行检测。

1.5　菌丝相白假丝酵母菌 LongSAGE 文库的构建

主要按 I-SAGE Long kit 说明书进行,并根据白假丝酵母菌酵母相细胞 LongSAGE 文库构建的经验作了部分调整,简述如下:①mRNA 与 Oligo (dT) 磁珠结合,转录合成 cDNA 第 1 链和 cDNA 第 2 链。②*Nla* Ⅲ酶酶切双链 cDNA,产生以 CATG 为粘性末端的 cDNA 片段。③cDNA 酶解片段等分成 2 份,分别与连接子 A、B 连接。④*Mme* Ⅰ酶酶切连接子与 cDNA 结合的序列,与磁珠分离,获得带有连接子的单标签。⑤T_4 DNA 连接酶连接单标签,形成 130 bp 的双标签,批量 PCR 扩增以获得足够量的双标签。⑥PAGE 纯化批量 PCR 产物。⑦*Nla* Ⅲ酶切 130 bp 的双标签,去除连接子,得到 34 bp 的双标签,PAGE 纯化。⑧T_4 DNA 连接酶连接 34 bp 的双标签形成串联体,收集串联体并克隆到线性化质粒 pZero-1 内,电转化 One Shot® TOP10 后,以抗生素 Zeocin 筛选、扩增阳性克隆。⑨PCR 筛选、验证转化阳性克隆后,收集纯化质粒,送公司行 DNA 测序,测序仪为 ABI PRISM 377 DNA 测序仪。

在文库构建过程中进行了部分步骤的改进:①延长磁珠在磁场中的放置时间。②减少连接子用量。③重复洗脱凝胶。在文库构建过程中每步均有相应的质量控制,见结果部分。

1.6　统计学分析

将所有 DNA 序列的测序文件输入计算机,用 SAGEsoftware 2000 v4.5 软件分析,并提取标签序列,将所提取的标签与 NCBI 的 SAGEmap 和 UNIGENE 中的序列进行比较获得标签的信息。

2　结果

2.1　菌丝相白假丝酵母菌的诱导培养

以 RPMI1640 为诱导培养基,将 pH 调为 7.0 及培养温度为 37 ℃作为诱导白假丝酵母菌菌丝形成的条件,实验表明只需 3 h即可诱导出纯度 >95% 的菌丝相白假丝酵母菌,此时可见菌丝分隔。

2.2　总 RNA 电泳

菌丝相白假丝酵母菌总 RNA 电泳结果见图 1。28S、18S rRNA 条带亮度比约为 2∶1,说明 RNA 纯度高。

2.3　菌丝相白假丝酵母菌 LongSAGE 文库的构建及校验

从总 RNA 提取到 LongSAGE 文库构建完成,步骤复杂且具有连贯性,任何一步的失误将导致文库构建的失败,关键步骤包括:①双链 cDNA 合成、*Nla* Ⅲ酶切验证;②连接子连接验证;③*Mme* Ⅰ酶酶切获得单标签;④双标签的连接;⑤34 bp 双标签的酶切释放,结果见图 2。34 bp 带:不带连接子的双标签;50 bp带:连接子;100 bp 带:*Nla* Ⅲ部分酶切片段;130 bp 带:带连接子的双标签。

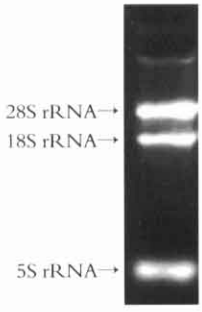
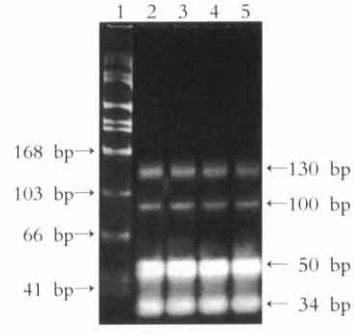


图 1　菌丝相白假丝酵母菌总 RNA 电泳图



1:分子质量标准;2－5:8% PAGE 纯化 *Nla* Ⅲ酶切产物

图 2　8%PAGE 纯化获得 34 bp 双标签

2.4　改进步骤结果

参照白假丝酵母菌酵母相细胞 LongSAGE 文库构建的经验,经相应步骤改进后,保证了文库构建的顺利进行,包括:①延长磁珠在磁场中的放置时间,使磁珠有足够时间贴壁,磁珠丢失明显减少。②减少连接子的用量,防止了过多的连接子相互连接,提高了 130 bp 条带的扩增效率。③重复洗脱凝胶,可明显提高凝胶中 DNA 的回收量。

2.5　测序结果及初步数据分析

2.5.1　标签分布一般情况　　本实验成功构建了菌丝相白假丝酵母菌 LongSAGE 标签文库,从文库中共挑取阳性克隆1 000 个进行测序。"CATGGTTGGGACACCAGTTCTTCTGTTGTGGAGCTTCCATGAAGGCATACAAAACGAGCAACAACTCTCAAACCCATG"为一阳性筛选克隆测序结果的一部分,测序结果含有 LongSAGE 的序列特征,即每 2 个双标签之间含有一锚定切点 CATG 作为标签的识别、分隔位点。测序共获得标签17 552个,代表单一转录本6 627种,其中单一匹配标签占 20%,多重匹配标签占 6.64%,假设蛋白标签为 18.47%,新标签为 53.07%。标签频率及百分比见表 1。

表1 菌丝相白假丝酵母菌 LongSAGE 标签文库的分布状态

标签类别	标签量(%)
总标签量	17 552
独特标签种类	6 627
无基因匹配标签	3 517(53.07)
有基因匹配标签	3 110(46.93)
单基因匹配标签	1 329(20.05)
多基因匹配标签	440(6.64)
假设蛋白标签	1 224(18.47)
基因组及粘粒等标签	117(1.77)
标签表达丰度	
1	4 674(70.53)
2~4	1 379(20.81)
5~9	336(5.07)
10~99	225(3.40)
≥100	13(0.20)

2.5.2 高表达标签检索结果 对拷贝数超过35的标签进行数据库检索,结果见表2。在拷贝数超过35的标签中,主要的表达基因为与细胞蛋白合成及能量代谢有关的基因,如核糖体RNA、苹果酸合酶、乙醇脱氢酶、甘油醛磷酸脱氢酶、烯醇酶等;更为重要的是,1/4(9/36)的标签找不到与之理想匹配的已知基因或表达序列,1/6(6/36)的标签为假设蛋白,提示其可能为新基因的表达序列。

表2 高表达量标签检索结果

LongSAGE 标签	表达丰度	代表基因描述
CCTGTTTGAGCGTCGTT	402	18S、5.8S、28S rRNA、ITS1、ITS2
GGTTTGAGAGTTGTTGC	177	乙醇脱氢酶
AAGCACTTCCAATTAAT	158	铜抵抗相关金属硫因
GTTTCCCACAGATCCGG	142	烯醇酶
GGGTAAATCTGGTGTTG	129	甘油醛-3-磷酸脱氢酶
GAATAATAGAATAGGAC	119	18S rRNA
CATATTGAAAGGTCTAT	105	无
TAAGAGAGTTGTCCTCA	91	无
ACAGTTTTGACAAAAAT	80	无
GGATATCACATTGAGTA	73	假设蛋白 CaJ7.0300;Tca2 retrotransposon gag polyprotein
GCCGTTCTTAGTTGGTG	61	无
CTGGTGCTAGACCTGCT	52	超氧化物歧化酶
GTTGTATATGTGTATTT	47	无
CCGATGGTTGTTTGTTG	45	假设蛋白:CaO19_11811
AGAGATAAGTGGAGAAA	37	无
CTTCCAACTATGGATGC	37	丙酮酸脱羧酶
AGAGTTGCTGTAGTTAA	36	脂肪酸去饱和酶
AAATATAGGAAGTTATA	35	苹果酸盐合酶
TGGTGGTTATAAAAGAT	209	假设蛋白:CaO19_7324
GTTTCAACGGGTAACGG	167	18S rRNA
TTAAATGTATAGAAAGG	155	无
CTGTTGCTTTATGGCTG	135	铜抵抗相关金属硫因
CGCGGGGATAAATCCTT	123	25S rRNA
GGAGCCACTATCATTCC	111	假设蛋白:CaO19_6805
GTACAAGGGTTGGGAAA	96	TEF-2,TEF-1
GTCGGGCCAGAAATAAG	87	无
GAATTTTAGATTTGATT	75	丁二醇脱氢酶
CCTCATACTACTCAAAG	64	Tca2 retrotransposon gag polyprotein
GTTTTAGTAAAGAAGAA	57	假设蛋白:CaO19.8934
CAGACAACAGTTAAGTA	50	无
TCTTTATATTTTACTTC	47	cosmid Ca41C10
GTGAACCAGAAAGAAAC	44	蛋氨酸合酶
GGTTTCGGTTTATTCGA	37	假设蛋白:CaJ7.0449
GCATTTGCTGCTACGCC	36	假设蛋白:CaO19_7676
GCTCTGGCGTATTGAAC	35	乙醇脱氢酶
TACGATGCTTGTTCCAC	35	线粒体酮醇酸还原异构酶

3 讨论

本实验成功构建了菌丝相白假丝酵母菌 LongSAGE 标签文库,共得到175 52个标签,代表单一转录本6 627种,其中2拷贝以上的转录本约占30%,基本反映了白假丝酵母菌的基因表达状态。

SAGE 是由 Velculescu 等[5]发展起来的用于大规模研究基因表达的方法,能在较短时间内检测细胞内几乎所有的 mRNA,并且能对这些 mRNA 的表达拷贝数进行定量分析,同时可以发现潜在的未知基因[6]。该方法主要基于以下3条原理:①一个14个核苷酸组成的标签(tag)从理论上可代表该转录本的足够信息。②将14 bp 的 SAGE 标签串联在一起,连接成一个串联体(concatemer),这样一次即可完成这些标签对应的转录本的检测。③根据对不同转录本的 SAGE 标签进行计数统计,就可以定量地评价相应转录本的表达水平。LongSAGE[7]是在 SAGE 基础上改良的方法,由于其获得的标签碱基数(21 bp)明显地较 SAGE 方法(14 bp)长,克服了 SAGE 方法得到标签长度太短、不便于进行基因确认的缺陷。

应用 LongSAGE 研究,检索标签有3种结果:①唯一匹配,即可以检索到唯一与之匹配的基因或 EST,这些标签一般不须进行其他检测。如标签"CATGGGGTAAATCTGGTGTTG"唯一匹配于甘油醛-3-磷酸脱氢酶基因。②多重匹配,即部分标签在现有数据中检索到多个与之匹配的表达序列。而这些表达序列除了具有相同的 LongSAGE 标签序列以外,无其他相似性,也就是说对于某一特定的组织或某种功能状态下的细胞,根据这些 LongSAGE 标签很难确定哪个基因为该组织或细胞的表达基因,需另行相关检测。如"CATGGTACAAGGGTTGGGAAA"匹配于"TEF-1、TEF-2"两个基因,在白假丝酵母菌菌丝相时表达的基因可能为其中的一个,或两个均有表达,只是其表达量不同,只有进行进一步的研究才能确认,如进行 GLGI 技术[8][generation of longer cDNA fragments from serial analysis of gene expression (SAGE) tags for gene identification]。③未知标签,即现有数据库中无与之匹配的基因或 EST,可能为潜在的新基因,但必须使用相应手段对其进行鉴定和验证,如 GLGI、3′RACE 及 Northern 杂交等。另外,对单基因匹配标签进行分析显示,大部分基因对应于单个标签,但也存在单一基因有多个不同的表达标签,即多个不同的表达标签代表着同一个基因,如 CUP1 有2个匹配标签(CATGAAGCACTTCCAATTAAT,CATGCTGTTGCTTTATGGCTG),这与转录产物的替代性剪切有关,也可能与转录产物含有多个 *Nla*

Ⅲ酶切位点有关，当*Nla*Ⅲ酶切cDNA时可在不同的位点进行酶切，以致产生不同的代表同一基因的表达标签，这种现象不影响LongSAGE标签库的构建，且由于单条转录本只能产生单个LongSAGE标签，因此不对标签量所代表的基因表达丰度产生影响。本研究表明，对LongSAGE库中拷贝数超过35拷贝的36个表达标签进行检索，发现有9个不能找到与之匹配基因的表达标签，对这些新的表达标签的进一步研究有可能发现新的基因。

参考文献：

[1] CALDERONE R A, FONZI W A. Virulence factors of Candida albicans[J]. Trends Microbiol, 2001, 9(7): 327-335.
[2] LO H J, KOHLER J R, DIDOMENICO B, *et al*. Nonfilamentous C. albicans mutants are avirulent[J]. Cell, 1997, 90(5): 939-949.
[3] MITCHELL A P. Dimorphism and virulence in Candida albicans[J]. Curr Opin Microbiol, 1998, 1(6): 687-692.
[4] 周村建，郝 飞，王 鲁，等. 白念珠菌菌丝的诱导培养[J]. 第三军医大学学报，2004，26(9)：832-833.
[5] VELCULESCU V E, ZHANG L, VOGELSTEIN B, *et al*. Serial analysis of gene expression[J]. Science, 1995, 270(5235): 484-487.
[6] CHEN J, SUN M, LEE S, *et al*. Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags[J]. Proc Natl Acad Sci U S A, 2002, 99(19): 12257-12262.
[7] SAHA S, SPARKS A B, RAGO C, *et al*. Using the transcriptome to annotate the genome[J]. Nat Biotechnol, 2002, 20(5): 508-512.
[8] CHEN J J, ROWLEY J D, WANG S M. Generation of longer cDNA fragments from serial analysis of gene expression tags for gene identification[J]. Proc Natl Acad Sci U S A, 2000, 97(1): 349-353.

（编辑 汪勤俭）

文章编号：1000-5404(2006)13-1387-01

个案与短篇

下咽癌组织中FHIT基因表达的研究

陈锦屏 （郑州市第五人民医院耳鼻咽喉科，郑州 450003）

脆性组氨酸三联体（FHIT）基因是1996年由Ohta等[1]克隆的一个位于3p14.2区域的候选抑癌基因。研究表明，FHIT在人类许多肿瘤组织或细胞系中呈现高频率的异常，与肿瘤的发生发展有关[2]。本研究采用免疫组化法，检测FHIT在下咽癌组织中蛋白的表达，并就其与临床病理指标的可能关系进行了研究。

1 材料与方法

1.1 临床资料

39例下咽癌标本选自我院耳鼻咽喉科2003年4月至2004年10月手术切除并经病理证实的组织标本，均为鳞状细胞癌，术前未行任何化疗、放疗。其中男性27例，女性12例，年龄37～79岁，中位年龄54岁。梨状窝癌13例，环状软骨后癌11例，咽后壁癌15例。临床分期：Ⅰ期5例，Ⅱ期9例，Ⅲ期17例，Ⅳ期8例。其中颈部淋巴结转移27例，无淋巴结转移12例。按组织学分级为：高分化9例，中分化11例，低分化19例；另取癌旁2 cm外经病理证实的正常下咽部上皮组织作为正常对照。

1.2 免疫组织化学

组织切片分别进行HE染色和免疫组化S-P法染色。兔抗人FHIT多克隆抗体及S-P试剂盒购自北京中山生物技术有限公司，免疫组化染色步骤按试剂盒说明进行。用PBS代替一抗作为阳性对照。

1.3 结果判断标准

参考文献[3]，观察FHIT蛋白的分布、阳性强度和阳性率。先按染色强度评分：0分为无色，1分为浅黄色，2分为棕黄色，3分为棕褐色；再按阳性细胞所占百分比评分：0分为阴性，1分为阳性细胞≤10%，2分为11%～50%，3分为51%～75%，4分为>75%，染色强度与阳性细胞百分比的乘积≥2分为免疫反应阳性（+），余为阴性（-）。

1.4 统计学处理

采用SPSS 10.0统计软件包进行统计分析。计数资料的比较采用χ^2检验及Fisher精确概率法，相关分析采用Spearman相关分析。

2 结果

2.1 FHIT在下咽癌组织和癌旁正常下咽组织中的表达

FHIT特异性着色主要位于细胞质内，部分细胞核也有表达。下咽癌组织中FHIT呈点片状或区域性分布，且多为弱阳性表达（83.3%，15/18），阳性表达率为46.2%（18/39）；正常咽上皮组织中均见FHIT阳性表达，呈弥漫性分布，且多为强阳性表达（71.4%，10/14），阳性表达率为100%（14/14）；组间差异有显著性（$\chi^2=27.459$，$P<0.05$）。

2.2 FHIT蛋白表达和下咽癌临床病理特征的关系

39例下咽癌组织中，FHIT阳性表达率在组织学分级中为：高/中分化鳞癌55.0%（11/20），低分化鳞癌36.8%（7/19），差异有显著性（$P<0.05$）。淋巴结转移组FHIT蛋白阳性率为40.7%（11/27），无淋巴结转移组为58.3%（7/12），差异有显著性（$P<0.05$）。根据临床病理分期，Ⅰ+Ⅱ期组与Ⅲ+Ⅳ期组之间FHIT蛋白阳性表达率分别为57.1%和40.0%，差异有统计学意义（$P<0.05$），见表1。

（下转1392页）

作者简介：陈锦屏（1967-），女，陕西省汉中市人，主治医师，主要从事耳鼻咽喉肿瘤方面的研究。电话：(0371)63858247

收稿日期：2005-09-29；修回日期：2006-03-21