论著

文章编号:1000-5404(2005)20-2028-03

利用SAGE标签产生长片段cDNA应用于白念珠菌基因识别

李永平,周村建,郝 飞 (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 鉴定本实验室建立的白念珠菌LongSAGE文库。**方法** 运用利用SAGE标签产生长片段cDNA应用于基因识别(GLGI)技术将白念珠菌的酵母相和菌丝相17 bp的LongSAGE标签扩增至相应的3′ cDNA端,获得数百个碱基片段,并进行序列分析。**结果** 20个标签中有15个获得稳定有效的扩增,长度在200~1 000 bp之间,获得的长片段进行序列分析后,有超过95%的碱基与已知基因相同,并且包括原来17 bp的SAGE标签,表明获得的序列为此已知基因。**结论** 研究表明,用SAGE标签作为引物,可有效地识别基因。利用SAGE-GLGI技术可以用来说明更多的白念珠菌的表达基因序列和其它真核生物的基因组。

关键词:白念珠菌;SAGE;GLGI

中图法分类号:R378.993;R394-33;R394.3 **文献标识码**:A

Identification of *Candida albicans* genes with cDNA technique from LongSAGE tags

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

Abstract: Objective To identify the LongSAGE tags from yeast LongSAGE library and hyphal LongSAGE library of *Candida albicans*. Methods The technique called the generation of longer cDNA fragments from serial analysis of gene expression (SAGE) tags was used for gene identification (GLGI). The SAGE tags of 17 bases were converted into their corresponding 3′ cDNA fragments covering hundred bases, and the sequence was analyzed. Results Fifteen out of twenty tags were stably and efficiently extended with the base size between 200-1 000 bp. The sequence was considered to represent a known gene since it matched a full-length transcript sequence with over 95% similarity in the same orientation, including the same 17-bp SAGE tag sequence. Conclusion By using novel SAGE tags as primer, we can identify efficiently many novel transcripts. A combined application of SAGE-GLGI can be used to define the boundary of expressed genes in the genomic sequences in *Candida albicans* and in other eukaryotic genomes.

Key words: *Candida albicans*; LongSAGE; GLGI

白念珠菌是人体体表及胃肠道的常居真菌,是一种致病芽殖酵母,能进行酵母和菌丝形态的转变,它的致病性与菌丝的形成密切相关[1]。为深入研究白念珠菌相转换机制,本实验运用"利用SAGE标签产生长片cDNA段应用于基因识别[generation of longer cDNA fragments from serial analysis of gene expression (SAGE) tags for gene identification GLGI]GLGI技术"[2,3],获得差异表达LongSAGE标签所代表的从标签至3′ cDNA端的基因序列,验证了LongSAGE文库的建立成功,对发现在菌相转换中起关键作用的调控基因和临床筛选可调控菌相转换生长的药物具有积极意义。

1 材料与方法

1.1 材料

1.1.1 试剂 AMV逆转录酶、EX *Tap*酶均购自大连宝生物生物技术公司;引物由北京赛百胜生物技术公司合成,TA克隆试剂盒、一步法感受态制备试剂盒购自上海生工生物技术公司;胶回收试剂盒等购自北京鼎国生物技术公司。测序仪为ABI PRISM 377 DNA测序仪,测序由公司完成。

1.1.2 菌种 白念珠菌ATCC90028. *E. coli*DH5α为本实验室保存菌种。

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:李永平(1974-),女,新疆维吾尔族自治区米泉县人,硕士研究生,主治医师,主要从事深部真菌病治疗方面的研究。电话:(023)68754419,E-mail: liuheroliuhero@163.com.cn

通讯作者:郝 飞,E-mail: haofei@mail.tmmu.com.cn

收稿日期:2005-06-30;**修回日期**:2005-09-10

1.1.3　LongSAGE 标签　　17 bp 的 LongSAGE 标签选自第三军医大学西南医院皮肤科实验室建立的白念珠菌酵母相和菌丝相的 LongSAGE 标签文库，检索证实为高表达多重匹配标签，并符合 PCR 引物设计要求。

1.2　方法

1.2.1　RNA 提取、cDNA 合成及验证　　根据参考文献[4]用玻璃珠法制备酵母相和菌丝相白念珠菌的总 RNA。逆转录反应采用 TAKARA 公司的 AMV 逆转录试剂盒，以 Oligo(dT)起始。根据白念珠菌热休克蛋白 90(HSP90)[5]和菌丝相菌丝壁蛋白 1(HWP1)[6]的 mRNA 序列用 primer5.0 设计引物，验证各相 RNA 的纯度。

1.2.2　PCR 条件扩增 3′端所缺失片段　　以 RT-PCR 产物为模板，标签为上游引物(5′-CATGxxxxxxxxxxxxxxxx-3′)，OligoA/C/G 分别为下游引物进行 PCR 扩增，PCR 反应体系(25 μl 总反应)：10 × buffer 2.5 μl，Mg^{2+} 2 μl，10 mmol/ml dNTP 1.25 μl，引物 P1(或 P2)各 1 μl，模板 cDNA 1 μl，*Taq* 酶 0.125 μl，DMSO 1.25 μl 补水至 25 μl。PCR 反应参数为：94 ℃预变性 2 min，先进行 94 ℃变性 20 s，根据标签不同退火温度为49 ℃ ~ 50 ℃，退火 20 s，72 ℃ 20 s 构成的 5 个循环，然后进行 94 ℃ 20 s，60 ℃ 20 s，72 ℃ 20 s 构成的 20 ~ 30 个循环，最后 72 ℃ 延伸 5 min。PCR 产物经 2% 琼脂糖凝胶电泳确认 PCR 产物大小。EB 染色，紫外灯下观察结果。

1.2.3　DNA 克隆、鉴定及测序　　PCR 产物回收后进行 TA 克隆。试剂加入量：新鲜的 PCR 产物(1 d 内)2 μl；10 × ligationg buffer 1 μl；pUCm-Tvector(25 ng/μl) 2 μl；STerial H_2O 4 μl；T_4 DNA ligase(3.0Wess unit/Ul)1 μl；总体积 10 μl ；在 PCR 仪上 22 ℃孵育 1.5 h，4 ℃过夜。转化大肠杆菌 DH5α 感受态，转化产物涂布于氨苄青霉素抗性平板，37 ℃培养过夜，用 PCR 法筛选阳性克隆[7]，获得阳性克隆后增菌、抽提质粒并测序。

所有测序获得的结果经检索进行同源性分析。

2　结果

2.1　PCR 反应条件

分别筛选了本实验室建立的白念珠菌酵母相和菌丝相的多重匹配 LongSAGE 标签各 10 个作为上游引物。酵母相有 8 个标签，菌丝相有 7 个标签扩增出片段，长度为 200 ~ 1 000 bp。5 个未获得稳定有效的扩增。在 PCR 体系和参数适合时，仅有 Oligo(dG)获得扩增，获得的长片段进行序列分析后，分别和基因 AY497754[8]，基因 L28817[9]有 >95% 的碱基相同，并且包括原来 17 bp 的 SAGE 标签，表明获得的序列为此已知基因(图 1)。

2.2　检测结果

通过 GLGI-SAGE 技术，获得的序列均在 200 bp 以上，其中包括原来的 21 bp 的 SAGE 标签。是原来 LongSAGE 标签长度的 100 倍以上。在白念珠菌的酵母相及菌丝相的 LongSAGE 标签中可寻找到与白念珠菌形态相关基因，并通过 GLGI 技术扩增得到 200 ~ 1 000 bp的片段，可以进行更多及有价值的生物信息学分析(表 1)。

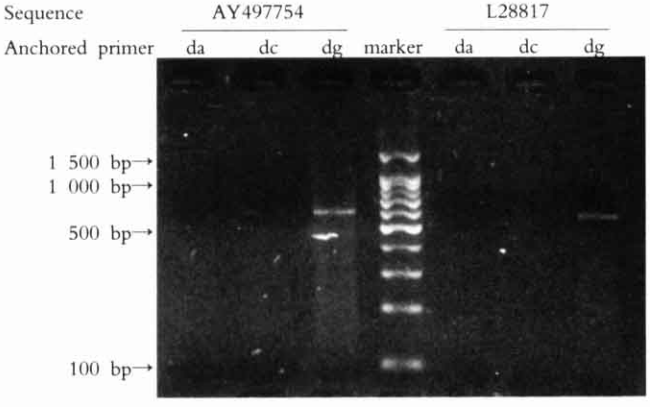


图 1　标签 RT-PCR 结果

Fig 1　The specific SAGE tag sequence by GLGI

表 1　GLGI 技术获得的部分标签鉴定结果

Tab 1　GLGI results from SAGE tags

SAGE tags(17 bases)	Amplified by anchored oligo (dT)	Length of sequence bases	Unigene identification No. hyphal
CTTAAAATCCCGACTGT	dC	300 bp	AY497754
CCTCATACTACTCAAAG	dA	300 bp	AF007776
TCAAGTTCAGAAATTTC	dG	900 bp	AF051313
AGAAGAGTTACTTACGC	dC	400 bp	AY497649
CTAGAACGCGATGATTT	dG	469 bp	L28817
CCAAGTGGTATTTTATT	dG	800 bp	AF28526
AAAGTGTGGCCTATCGA	dC	700 bp	X70659
GTACAAGGGTTGGGAAA	—	—	—
GACGTTATGTGGTACCC	—	—	—
CCCAAAGAGTCATATTC	—	—	—
Yeast			
AGAAGAGTTACTTACGC	dG	600 bp	AY497649
TACTGTGAACGACGCT	dG	457 bp	AF068866
TTCAAGAAAAGCTGGTT	dG	700 bp	U61975
AAGAGTGTATGAAAATT	dA	800 bp	AF247190
CCTGTTTGAGCGTCGTT	dG	600 bp	AY672930
CCTCATACTACTCAAAG	dC	300 bp	AF007776
CTTAAAATCCCGACTGT	dG	521 bp	AY497754
ACCTTCCCTTCGTTATC	dC	300 bp	AF050215
TCATCATCGGATTCTTC	—	—	—
GTACAAGGGTTGGGAAA	—	—	—

3　讨论

基因表达系列分析(serial analysis of gene expression, SAGE)是一种研究真核细胞表达基因信息的高通量检测技术，它能对细胞内所有的表达基因进行定性与定量分析[10]。

SAGE 标签存在 3 种结果：①唯一匹配：可以检索到唯一与之匹配的基因或标签。②多重匹配：部分标签在现有数据中检索到多个与之匹配的表达序列，需

另行相关检测。③未知标签:现有数据库中无匹配基因,可能为潜在的新基因,必须对其进行鉴定。后二者是SAGE数据处理与分析过程中的两大困难。Chen等[2,3]利用GLGI技术将包括10 bp或17 bp的SAGE标签前面加识别的碱基碱基位点CATG,作为5′端引物,3′端使用的是单一碱基dA、dG或dC与Oligo-dT结合的锚定引物进行第1次PCR扩增,只有与SAGE标签相匹配的序列才能退火并扩增。在第2次循环中,只有5′端Oligo-dA的前一碱基与锚定的Oligo-dT相匹配的序列才能扩增,而单纯的Oligo-dA与Oligo-dT结合的片段因处于锚定状态则受抑制,因而最终得到的是含有SAGE标签的长片段DNA扩增产物,产物长度为200~1 000 bp。GLGI技术提供了几种可能的结果。首先,它提供了为SAGE技术的发展策略,此次结合应用SAGE/GLGI技术可以作为人或其它物种的基因的选择和表达。此外,它可以鉴定任何基因的3′端序列的外显子,可以确定人或其它物种的3′端在基因家族中所表达的基因3′末端。

GLGI技术也有其缺陷,其突出问题是进一步增加了表达基因确认的工作量,因为在SAGE结果分析以后,大约2/3的标签[11]为未知标签或为多重匹配标签,GLGI得到的可能为几百甚至数以千计的DNA片段。GLGI实际应用困难较大,只有在拥有大规模测序分析的实验室内才可以进行,因而仅选择了其中20个多重匹配标签进行了GLGI鉴定分析并最终确认。验证了LongSAGE-GLGI技术的可行性,并为发现新基因奠定了基础。

通过GLGI技术验证了实验室建立的LongSAGE文库。实验获得的阳性结果中,基因AF247190为PHR1,是一个表达受pH值调控的基因,该基因的敲除白念珠菌在pH为7的条件下不能形成菌丝[12]。AF051313为ALS8,U87956为ALS3,同为ALS基因家族(凝集素样基因家族),其中ALS3[13]其表达受菌丝形态的调控。其余LongSAGE标签与白念珠菌结构蛋白或延长因子相关。实验中25%的标签未获得稳定有效的PCR扩增,无法进行进一步的分析。其原因为①LongSAGE实验过程中的误差;②SAGE文库测序中存在一定的错误;③LongSAGE标签自身存在引物二聚体或与下游引物Tm温度相差大(超过10℃),使得PCR扩增困难;④PCR循环参数不合理不能扩增出片段。在PCR循环体系中,Mg^{2+}在PCR产物的多少和特异性中具有重要的作用。此外,PCR的循环次数在产物的特异性中也是一个重要的因素,通过减少循环参数可以确定锚定的Oligo(dT),过多的循环次数增加错配的几率。

参考文献:

[1] Lo H J, Kohler J R, DiDomenico B, *et al*. Nonfilamentous *C. albicans* mutants are avirulent[J]. Cell, 1997, 90(5): 939-949.

[2] 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.

[3] Chen J, Lee S, Zhou G, *et al*. High-throughput GLGI procedure for converting a large number of serial analysis of gene expression tag sequences into 3′ complementary DNAs[J]. Genes Chromosomes Cancer, 2002, 33(3): 252-261.

[4] 周村建, 郝 飞, 王 鲁, 等. 白念珠菌总RNA提取的实验研究[J]. 中华皮肤科杂志, 2004, 37(8): 494-495.

[5] Staab J F, Ferrer C A, Sundstrom P. Developmental expression of a tandemly repeated, proline-and glutamine-rich amino acid motif on hyphal surfaces on *Candida albicans*[J]. J Biol Chem, 1996, 271(11): 6298-6305.

[6] Swoboda R K, Bertram G, Budge S, *et al*. Structure and regulation of the HSP90 gene from the pathogenic fungus *Candida albicans*[J]. Infect Immun, 1995, 63(11): 4506-4514.

[7] 王 安, 赵 明, 丁广治. PCR简便快速筛选阳性克隆[J]. 生物技术通讯, 2000, 11(1): 34-35.

[8] Diezmann S, Cox C J, Schonian G, *et al*. Phylogeny and evolution of medical species of Candida and related taxa: a multigenic analysis[J]. J Clin Microbiol, 2004, 42 (12), 5624-5635.

[9] Srikantha T, Gutell R R, Morrow B, *et al*. Partial nucleotide sequence of a single ribosomal RNA coding region and secondary structure of the large subunit 25 s rRNA of Candida albicans[J]. Curr Genet, 1994, 26(4): 321-328.

[10] Velculescu V E, Zhang L, Vogelstein B, *et al*. Serial analysis of gene expression[J]. Science, 1995, 270(5235): 484-487.

[11] Velculescu V E, Zhang L, Zhou W, *et al*. Characterization of the yeast transcriptome[J]. Cell, 1997, 88(2): 243-251.

[12] Yesland K, Fonzi W A. Allele-specific gene targeting in *Candida albicans* results from heterology between alleles [J]. Microbiology, 2000, 146 (Pt 9): 2097-2104.

[13] Zhao X, Oh S H, Cheng G, *et al*. ALS3 and ALS8 represent a single locus that encodes a *Candida albicans* adhesin; functional comparisons between Als3p and Als1p[J]. Microbiology, 2004, 150(Pt 7): 2415-2428.

(编辑 龙 亮)