

## 肠道微生物宏基因组文库构建及酯解酶基因筛选

段嘉<sup>1,2</sup> 程功<sup>1,2</sup> 胡永飞<sup>1</sup> 李晶<sup>1</sup> 律娜<sup>1</sup> 朱宝利<sup>1\*</sup>

(1. 中国科学院微生物研究所/中国科学院病原微生物与免疫学重点实验室,北京 100101;

2. 中国科学院研究生院,北京 100049)

**摘要** 利用宏基因组学方法,以人肠道微生物样品为原材料,构建了1个约30 000个克隆的fosmid文库。以三丁酸甘油酯为底物,通过功能筛选,获得1个酯解酶阳性克隆。对该阳性克隆构建亚克隆,挑选具有酯解酶活性的阳性亚克隆进行测序分析,最终获得1个肠道微生物来源的酯解酶基因(GenBank登录号:JQ972699)。结果表明,文库克隆的平均插入片段约为40 kb,没有重复插入片段克隆。获得的酯解酶基因推演蛋白与*Pyramidobacter piscolens* W5455的patatin样磷脂酶同源性最高,氨基酸一致性为95%。生物信息学分析结果表明该基因可能通过Vd型分泌方式进行分泌并发挥功能。本研究是通过构建人肠道微生物宏基因组大片段文库并结合重组子功能筛选获得酯解酶的首次报道,可为食品工业提供新的酯解酶来源和筛选方法。

**关键词** 宏基因组学; 肠道微生物; patatin样磷脂酶

中图分类号 Q 933

文章编号 1007-4333(2013)01-0010-06

文献标志码 A

## Screening of lipolytic genes from the metagenomic library of human intestine microbes

DUAN Jia<sup>1,2</sup>, CHENG Gong<sup>1,2</sup>, HU Yong-fei<sup>1</sup>, LI Jing<sup>1</sup>, LÜ Na<sup>1</sup>, ZHU Bao-li<sup>1\*</sup>

(1. Key Laboratory of Pathogenic Microbiology and Immunology/Institute of Microbiology,

Chinese Academy of Sciences, Beijing 100101, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100101, China)

**Abstract** Metagenomic DNA was extracted from a human fecal sample and DNA fragments around 40 kb was isolated and purified. A fosmid library containing approximately 30 000 clones was created. Restriction analysis by *Eco*R I revealed a high diversity of the cloned DNA fragments and the estimated average insert size of the library was 40 kb. Tributyrin was used as substrate for activity-based screening of the library and one lipolytic enzyme positive fosmid clone was obtained. The lipolytic gene sequence was obtained after subcloning and sequencing, which showed high homology to the patatin-like phospholipase from *Pyramidobacter piscolens* W5455 with an amino acid identity of 95%. This is the first report of function-based metagenomics on screening of lipolytic genes from human gut microbiota. It also provides food industry with a new source of lipolytic genes and screening method.

**Key words** metagenomics; gut microbiota; patatin-like phospholipase

酯解酶是所有能够催化酯类水解的酶类的总称。根据作用底物的不同,酯解酶分为3大类,即脂肪酶、酯酶及磷脂酶。脂肪酶水解脂肪酸侧链长度大于10碳的甘油三酯;酯酶水解脂肪酸侧链小于10碳的甘油三酯;而磷脂酶的水解底物为不同类型

的磷脂<sup>[1]</sup>。酯解酶可以催化不同酯类的水解合成反应,在食品工业中具有非常广泛而且重要的应用价值。在油脂加工中,酯解酶可以作为生物催化剂催化植物油中含有的某些磷脂成分,从而起到植物油脱胶的作用<sup>[2]</sup>。在乳制品加工中,酯解酶的水解催

收稿日期:2012-05-29

基金项目:国家重点基础研究发展计划(973计划)项目(2009CB522605)

第一作者:段嘉,硕士研究生,E-mail:duanj@im.ac.cn

通讯作者:朱宝利,研究员,主要从事微生物比较基因组学研究,E-mail:zhubaoli@im.ac.cn

化作用能够在不改变原有奶酪质量的同时提高产量<sup>[3]</sup>。在烘焙工业中,酯解酶可以替代或者增强传统的乳化剂,通过对面粉的极性脂类物质的酶解作用提高产品的乳化质量<sup>[4]</sup>。除此之外,某些农业昆虫能够通过体内的酯解酶对施用的杀虫剂发生作用,而使其获得一定的抗药性,降低杀虫剂的效果<sup>[5]</sup>。

环境中的微生物是酯解酶基因的重要来源。然而其中大部分是不可培养的,严重限制了酯解酶的开发和利用<sup>[6]</sup>。宏基因组学是一种不依赖于传统培养、能够直接对环境中的DNA进行研究的技术。目前已经被广泛应用于各种环境中酯解酶的筛选;Hu等2010年以fosmid载体构建了海底沉积物微生物宏基因组文库,并筛选出新酯解酶FLS18C和FLS18D<sup>[7]</sup>;2011年Nacke等以fosmid载体和TOPO载体构建森林和草原土壤微生物宏基因组文库,从中得到了37个酯解酶,其中2种不属于任何已知酯解酶家族<sup>[8]</sup>;2009年Liu等以BAC载体构建了中国荷斯坦奶牛的瘤胃微生物宏基因组文库,并筛选到了2个新酯解酶RlipE1和RlipE2<sup>[9]</sup>。

目前,有关人体肠道微生物群落中(尤其是未培养微生物中)酯解酶的分布及功能的研究仍十分缺乏,而构建宏基因组大片段文库,结合重组子功能筛选研究人体肠道微生物酯解酶基因未见报道。本研究通过构建人肠道微生物宏基因组文库,旨在利用酯解酶选择性培养基从文库中筛选出具有活性的酯解酶基因,并对其进行分析。

## 1 材料与方法

### 1.1 菌株及试剂

大肠杆菌(*E. coli*)DH5 $\alpha$ 购自Transgen公司;pUC118载体购自Takara公司;T4 DNA连接酶、*EcoR* I、*Sau3A* I等购自NEB公司;质粒提取及胶回收试剂盒购自OMEGA公司;三丁酸甘油酯(tributyrin)购自Acros Organics公司;Fosmid文库构建试剂盒(CopyControl™ Fosmid Library Production Kit)购自Epicentre公司。

### 1.2 样品采集及宏基因组DNA提取

采集健康人粪便样品,参考Zhou等<sup>[10]</sup>的方法提取宏基因组DNA。具体方法如下:称取5g粪便样品与13.5mL DNA提取缓冲液(100mmol/L Tris-HCl[pH 8.0],100mmol/L EDTA[pH 8.0],100mmol/L 磷酸钠[pH 8.0],1.5mol/L NaCl,1% CTAB)及100 $\mu$ L蛋白酶K(10mg/mL)在离心

管中混匀后,置摇床37 $^{\circ}$ C孵育30min。加入20% SDS 1.5mL,65 $^{\circ}$ C水浴2h,每20min轻轻上下颠倒几次。室温下6000g离心10min后转移上清至一新的离心管中。加入等体积的氯仿/异戊醇(24:1)后6000g离心10min。吸取水相,加入0.6倍体积异丙醇轻柔混匀后室温静置1h。室温下16000g离心20min后弃上清,用预冷的70%乙醇洗涤沉淀,自然晾干后用加入500 $\mu$ L灭菌去离子水重溶。

### 1.3 Fosmid文库构建及质量评估

脉冲场凝胶电泳分离宏基因组DNA,切取48kb上下的DNA片段(共计约6mm凝胶),经电洗脱后利用透析膜浓缩至50 $\mu$ L。Fosmid文库构建方法参考试剂盒说明书。浓缩后的DNA经末端补平后连接至载体pCC2FOS,利用包装蛋白包装成噬菌体颗粒后转染至试剂盒中提供的宿主菌*E. coli* EPI300。吸取100 $\mu$ L转染后的菌液铺氯霉素(12.5 $\mu$ g/mL)平板,次日计数,计算出文库的总克隆数。随机挑取20个菌落接入5mL液体培养基,过夜培养后使用质粒提取试剂盒提取质粒并使用*EcoR* I酶切后电泳,以判定文库插入片段的大小及多样性。

### 1.4 酯解酶活性克隆筛选及亚克隆

先配制100mL三丁酸甘油酯乳化液(含有3g聚乙烯醇及10mL三丁酸甘油酯)。加10mL乳化液于90mL LB中,配制成固体筛选培养基(12.5 $\mu$ g/mL氯霉素)。每块筛选平板上铺约包含1000个Fosmid克隆的文库菌液,37 $^{\circ}$ C培养过夜,肉眼观察菌落周围透明圈的形成,依此挑出酯解酶阳性克隆菌落。阳性克隆菌落经过夜培养后提取质粒,利用*Sau3A* I(1:1000稀释)进行部分酶切,电泳后切胶回收1~5kb大小的DNA片段,连接至载体pUC118并转化至*E. coli* DH5 $\alpha$ 。使用相同的底物筛选平板(100 $\mu$ g/mL氨苄西林)筛选阳性亚克隆。

### 1.5 亚克隆序列测定及生物信息学分析

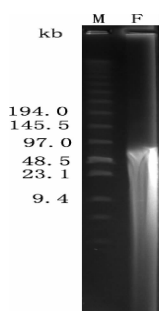
使用载体pUC118上自带的M13引物序列进行双末端测序,并进一步设计引物进行步移,直至序列测通。开放阅读框(open reading frame,ORF)预测使用在线软件ORF finder(<http://www.ncbi.nlm.nih.gov/projects/gorf/>)。使用BlastP程序检索预测蛋白的同源序列(<http://blast.ncbi.nlm.nih.gov/>)。多序列比对使用Clustal X2<sup>[11]</sup>。序列进化树构建使用MEGA5<sup>[12]</sup>。使用SignalP4.0

(<http://www.cbs.dtu.dk/services/SignalP/>)预测蛋白的信号肽序列;TMHMM 2.0(<http://www.cbs.dtu.dk/services/TMHMM/>)预测蛋白的跨膜区。保守结构域检索使用 Batch CD-Search<sup>[13]</sup>。蛋白二级结构预测使用 PSIPRED3.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>)。

## 2 结果与分析

### 2.1 人体肠道微生物宏基因组文库构建

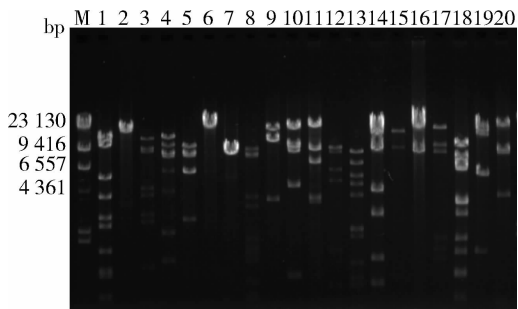
将提取的肠道微生物宏基因组 DNA 进行脉冲场凝胶电泳。结果显示提取的宏基因组 DNA 主要集中在 40 kb 附近,可以不经打碎而直接构建 Fosmid 文库(图 1)。构建的文库经平板计数估算总克隆数约为 30 000 个。*EcoR* I 酶切电泳结果显示,文库 DNA 插入片段具有丰富的多样性,平均插入片段大小约为 40 kb,涵盖约 1.2 Gb 的外源 DNA(图 2)。假设细菌基因组平均长度为 5 Mb,则文库



M:Low range PFG Marker;F:肠道微生物宏基因组 Human fecal Metagenomic DNA.

图 1 人体肠道微生物宏基因组 DNA 脉冲场凝胶电泳

Fig. 1 Pulsed Field Gel Electrophoresis of human gut microbiota DNA



M: $\lambda$ -Hind III DNA digest Marker;1~20:酶切产物电泳图谱 Gel electrophoresis of digested fosmids

图 2 Fosmid 质粒 *EcoR* I 酶切电泳图谱

Fig. 2 Gel electrophoresis of digested fosmids with *EcoR* I

约涵盖 240 个细菌基因组 DNA 信息。

### 2.2 酯解酶活性克隆筛选及亚克隆

以三丁酸甘油酯为底物对文库进行功能筛选。所有文库克隆共计铺 30 块平板(约 1 000 克隆/板),最终获得一个酯解酶活性克隆。为获得酯解酶基因的序列信息,首先使用 *Sau*3A I 对该克隆的 Fosmid 质粒进行部分酶切,连接至氨苄西林抗性载体 pUC118 上,形成可用于直接测序的小片段文库后,再次使用相同的底物筛选获得活性亚克隆。挑取其中的 1 个活性亚克隆测序。

### 2.3 酯解酶活性亚克隆序列测定及分析

经过双末端测序、步移及拼接,获得长度为 3 829 bp 的亚克隆序列(GenBank 登录号:JQ972699)。这段序列上包含 2 个完整的 ORF,其中,ORF1 长度为 1 437 bp,ORF2 长度为 2 088 bp。同源检索结果显示,ORF1 编码蛋白与数据库中的 proline-tRNA ligase 同源,而 ORF2 编码蛋白与 patatin 样磷脂酶同源。将 ORF2 基因命名为 *plp*,其编码的蛋白相应命名为 PLP。信号肽预测结果显示,PLP 蛋白 N-端前 21 个氨基酸为信号肽序列;保守结构域检索显示,在 PLP 蛋白的 N-端序列(27~233 aa)处存在一个保守的 patatin 样磷脂酶结构域。据此推断 ORF2 编码的蛋白 PLP 是产生酯解活性的酶。

### 2.4 酯解酶活性蛋白 PLP 进化分析

选取同源性最高的 50 条蛋白序列,去除其中功能注释为假想蛋白的序列,剩余的序列与 PLP 蛋白一起构建进化树(图 3)。结果显示,PLP 蛋白与来源于互养菌门(Synergistetes)的 8 株细菌的同源蛋白位于 1 个独立的进化分支上。其中,与菌株 *Pyramidobacter piscolens* W5455 的 patatin 样磷脂酶同源性最高,氨基酸一致性(Identity)为 95%;与其他菌株的同源蛋白序列一致性均低于 45%。以上结果提示,通过功能筛选获得的酯解酶基因可能来自互养菌门细菌,并与菌株 *Pyramidobacter piscolens* W5455 来自相同种属。

### 2.5 酯解酶活性蛋白 PLP 二级结构预测

保守结构域分析结果显示,仅在 PLP 蛋白的 N-端含有 1 个保守的 patatin 样磷脂酶结构域,而 C-端的约 400 个氨基酸的结构及功能仍然未知。对 PLP 的二级结构进行预测,结果发现,在 C-端存在 18 个连续的  $\beta$ -片层结构(图 4)。Salacha 等<sup>[14]</sup>对铜绿假单胞菌(*Pseudomonas aeruginosa*) 1 个 patatin

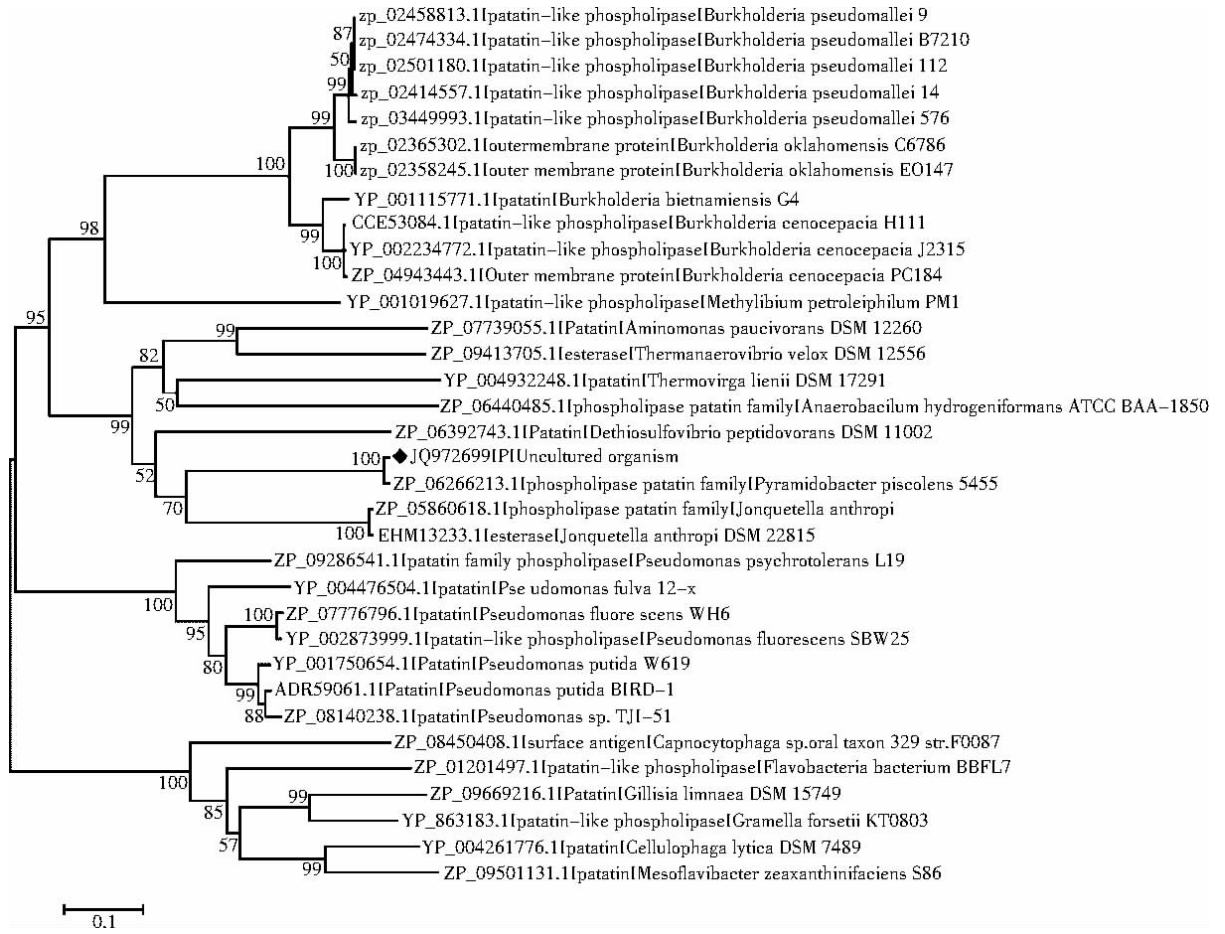


图 3 酯解酶蛋白 PLP 及其同源蛋白序列进化树

Fig. 3 Homology analysis of lipolytic protein PLP and closely related proteins

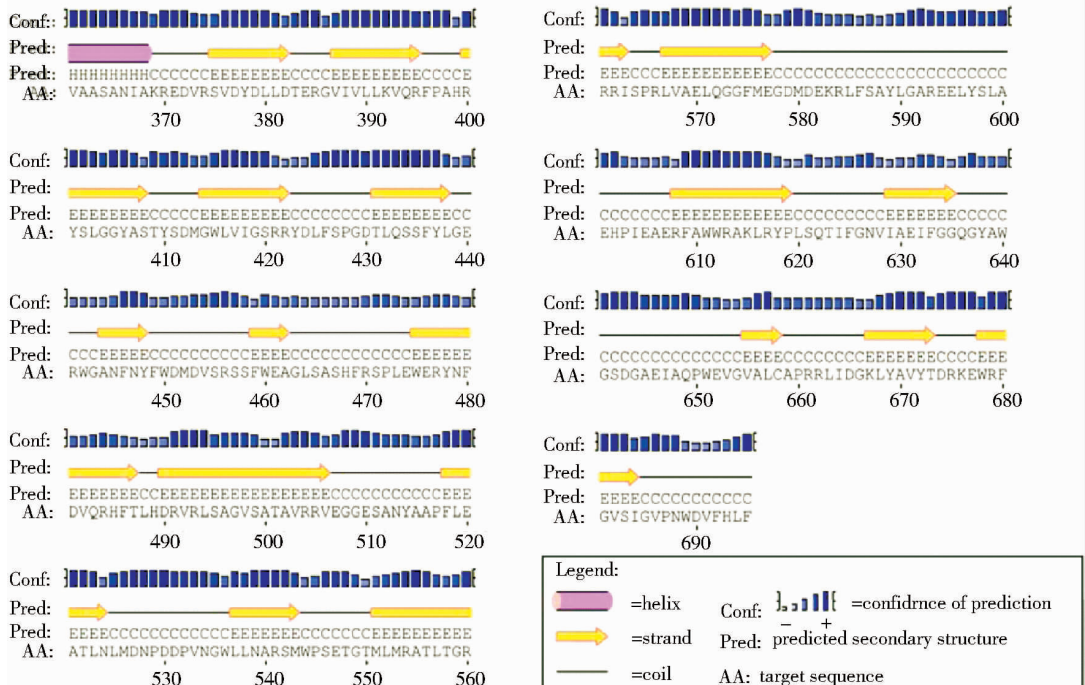


图 4 酯解酶蛋白 PLP(C-端) 二级结构预测结果

Fig. 4 C-terminal secondary structure prediction of lipolytic protein PLP

家族磷脂酶的研究结果显示,在该蛋白的 C-端存在 1 个膜定位  $\beta$ -桶结构,揭示该蛋白通过一种新的分泌机制(Vd 型分泌系统)将 N-端活性部分分泌至胞外。本研究获得的 PLP 蛋白的 C-端  $\beta$ -片层很可能与上述铜绿假单胞菌中的 patatin 家族磷脂酶类似,构成 1 个  $\beta$ -桶结构,并利用相似的分泌机制将 N-端活性区域转运至胞外。

### 3 讨论

根据蛋白序列及生物学特征的不同,Arpigny 等<sup>[1]</sup>将细菌酯解酶分成 8 个蛋白家族。后来,在铜绿假单胞菌(*P. aeruginosa*)中发现 1 个具有磷脂酶活性的毒力因子 ExoU,与马铃薯储存蛋白 patatin 存在较高的同源性<sup>[15]</sup>。Banerji 等<sup>[16]</sup>通过序列分析发现,除 ExoU 外,在细菌的基因组中还存在大量 patatin 的同源蛋白编码基因序列,提示细菌中存在 1 个新的酯解酶蛋白家族,即 patatin 样磷脂酶家族。本研究通过功能筛选的方法从人体肠道微生物宏基因组文库中获得 1 个酯解酶活性基因 *plp*。生物信息学分析结果显示,该基因编码的蛋白 PLP 属于 patatin 样磷脂酶家族。与其同源性最高的蛋白来自口腔分离菌株 *Pyramidobacter piscolens* W5455,为基因组测序后预测的产物,酯解功能尚未鉴定。

目前,针对 patatin 样磷脂酶家族蛋白的研究仍然较少,且主要集中在病原菌。除铜绿假单胞菌(*P. aeruginosa*)的毒力因子 ExoU 外,该家族已经鉴定功能的蛋白还包括来自嗜肺军团菌(*Legionella pneumophila*)的 VipD、铜绿假单胞菌(*P. aeruginosa*)的 PlpD、伤寒立克次氏体(*Rickettsia typhi*)的 RT0522 及枯草芽孢杆菌(*Bacillus subtilis*)的 YvdO<sup>[14-15,17-19]</sup>。这些蛋白可通过不同的方式分泌至胞外,如 ExoU 通过 IV 型分泌系统,VipD 通过 III 型分泌系统,PlpD 通过 Vd 型分泌系统,而 RT0522 及 YvdO 的分泌方式尚未确定。针对 PLP 的二级结构预测结果显示,在其 C-端存在 18 个  $\beta$ -片层结构,可能形成与 PlpD 一样的  $\beta$ -桶状结构,从而通过 Vd 型分泌方式将 N-端活性部位转运至胞外。

宏基因组学通过非培养方法对微生物群落进行研究,无需依赖微生物的纯培养<sup>[20]</sup>。基于序列水平的研究极大拓展了对不同环境中微生物种属及遗传多样性的认识;而基于功能的研究则为微生物资源

的开发及应用开辟了新的途径。然而,基于功能的宏基因组学研究方法也存在一些不足之处。比如,不同载体的克隆效率及插入片段大小存在较大差异。相对于普通的质粒载体,本研究使用的 Fosmid 载体具有克隆效率高、稳定及插入片段大(约 40 kb)等优势,而一般质粒载体的插入片段均低于 10 kb,涵盖的信息量及克隆效率相对较低。

本研究筛选约 30 000 个克隆,仅获得 1 个酯解酶活性克隆,效率相对较低,推测主要有以下两方面原因:1)人体肠道原始粪便样品微生物中酯解酶基因的丰度并不高;2)*E. coli* 宿主可能无法将所有酯解酶基因表达出具有活性的产物,包括启动子识别及翻译过程的偏差,不具备有效的分泌系统等。对上述问题的解决,一方面可以通过预先富集具有酯解酶活性的微生物群体后再利用功能宏基因组学进行筛选;另一方面,使用 *E. coli* 以外的宿主菌及自带启动子的表达载体等方法将有助于提高筛选效率。

### 4 结论

利用 pCC2FOS fosmid 克隆载体构建了人体肠道微生物宏基因组文库,并通过功能筛选的方法获得一个属于 patatin 样磷脂酶家族的新的酯解酶 PLP。二级结构预测及分析结果显示,它可能通过 Vd 型方式分泌至胞外。该蛋白的具体功能及酶学性质等有待于进一步研究。这是通过构建人肠道宏基因组大片段文库,结合重组子功能筛选获得酯解酶的首次报道,可为食品工业提供了新的酯解酶来源和筛选方法。

### 参 考 文 献

- [1] Arpigny J L, Jaeger K E. Bacterial lipolytic enzymes: classification and properties[J]. *Biochem J*, 1999, 343 Pt 1: 177-183
- [2] Jiang F, Wang J, Kaleem I, et al. Degumming of vegetable oils by a novel phospholipase B from *Pseudomonas fluorescens* BIT-18[J]. *Bioresource Technol*, 2011, 102(17): 8052-8056
- [3] Lilbaek H M, Broe M L, Høier E, et al. Improving the yield of Mozzarella cheese by phospholipase treatment of milk[J]. *J Dairy Sci*, 2006, 89(11): 4114-4125
- [4] Kirk O, Borchert T V, Fuglsang C C. Industrial enzyme applications[J]. *Curr Opin Biotech*, 2002, 13(4): 345-251
- [5] Ishaaya I. Insect detoxifying enzymes: Their importance in pesticide synergism and resistance[J]. *Arch Insect Biochem*,

- 1993,22(1-2):263-276
- [6] Amann R I, Ludwig W, Schleifer K H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation[J]. *Microbiol Rev*,1995,59(1):143-169
- [7] Hu Y, Fu C, Huang Y, et al. Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment[J]. *FEMS Microbiol Ecol*,2010,72(2):228-237
- [8] Nacke H, Will C, Herzog S, et al. Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories[J]. *FEMS Microbiol Ecol*,2011,78(1):188-201
- [9] Liu K, Wang J, Bu D, et al. Isolation and biochemical characterization of two lipases from a metagenomic library of China Holstein cow rumen[J]. *Biochem Bioph Res Co*,2009,385(4):605-611
- [10] Zhou J, Bruns M A, Tiedje J M. DNA recovery from soils of diverse composition[J]. *Appl Environ Microb*,1996,62(2):316-322
- [11] Larkin M A, Blackshields G, Brown N P, et al. Clustal W and Clustal X version 2.0[J]. *Bioinformatics*,2007,23(21):2947-2948
- [12] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. *Mol Biol Evol*,2011,28(10):2731-2739
- [13] Marchler-Bauer A, Lu S, Anderson J B, et al. CDD: a Conserved Domain Database for the functional annotation of proteins[J]. *Nucleic Acids Res*,2011,39(Database issue):D225-229
- [14] Salacha R, Kovacic F, Brochier-Armanet C, et al. The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system[J]. *Environ Microbiol*,2010,12(6):1498-1512
- [15] Sato H, Frank D W, Hillard C J, et al. The mechanism of action of the *Pseudomonas aeruginosa*-encoded type III cytotoxin, ExoU[J]. *EMBO J*,2003,22(12):2959-2969
- [16] Banerji S, Flieger A. Patatin-like proteins: a new family of lipolytic enzymes present in bacteria? [J]. *Microbiol-SGM*,2004,150(Pt 3):522-525
- [17] Kato S, Yoshimura T, Hemmi H, et al. Biochemical analysis of a novel lipolytic enzyme YvdO from *Bacillus subtilis* 168[J]. *Biosci Biotechl Bioch*,2010,74(4):701-706
- [18] Rahman M S, Ammerman N C, Sears K T, et al. Functional characterization of a phospholipase A (2) homolog from *Rickettsia typhi*[J]. *J Bacteriol*,2010,192(13):3294-3303
- [19] Shohdy N, Efe J A, Emr S D, et al. Pathogen effector protein screening in yeast identifies *Legionella* factors that interfere with membrane trafficking[J]. *P Natl Acad Sci USA*,2005,102(13):4866-4871
- [20] Daniel R. The metagenomics of soil[J]. *Nat Rev Microbiol*,2005,3(6):470-478

责任编辑: 袁文业