

造纸废水纸浆沉淀物中未培养微生物纤维素酶基因的克隆和鉴定

许跃强^{1,2} 段承杰^{1,2} 周权能² 唐纪良^{1,2} 冯家勋^{1,2*}

(广西大学¹ 广西亚热带生物资源保护利用重点实验室² 生命科学与技术学院 南宁 530005)

摘 要 提取纯化造纸废水纸浆沉淀物的宏基因组 DNA 并构建 16S rDNA 文库,系统发育分析显示该环境中存在大量的未培养细菌且具有种类的多样性。以柯斯质粒为载体构建了 1 个含 10000 个克隆的宏基因组文库,文库容量为 3.53×10^8 bp。筛选文库得到 2 个表达内切葡聚糖酶活性的克隆、3 个表达外切葡聚糖酶活性的克隆和 2 个表达 β -葡萄糖苷酶活性的克隆。从表达不同活性的克隆中分别挑选活性最强的进行鉴定,得到 3 个新的纤维素酶基因 *umcel5L*、*umcel5M* 和 *umbgl3D*。*umcel5L*、*umcel5M* 和 *umbgl3D* 分别编码产生内切葡聚糖酶、纤维糊精酶和 β -葡萄糖苷酶,其编码产物与已报道的纤维素酶一致性最高的分别为 43%、48% 和 46%。这是第一次采用未培养方法对造纸废水纸浆沉淀物中的细菌多样性进行分析并从中克隆纤维素酶基因的报道。

关键词: 纸浆沉淀物 细菌多样性 未培养微生物 宏基因组文库 纤维素酶

中图分类号: Q78 文献标识码: A 文章编号: 0001-6209(2006)05-0783-06

纤维素是由吡喃型 D-葡萄糖以 β -1,4-糖苷键连接而成的直链状大分子,是地球上最丰富的生物质(biomass)资源。纤维素的降解涉及到一组复合的纤维素酶系:内切葡聚糖酶(endo-1,4- β -D-glucanase, EC3.2.1.4)、外切葡聚糖酶(包括纤维二糖水解酶(1,4- β -D-glucan cellobiohydrolase, EC3.2.1.91)和纤维糊精酶(cellulodextrinase, EC3.2.1.74))及 β -葡萄糖苷酶(β -glucosidase, EC3.2.1.21)^[1]。在这些酶的协同作用下,纤维素最终被转化为葡萄糖,葡萄糖经发酵生成的酒精可用作燃料^[2],这对于解决世界能源危机及环境污染等问题具有重要意义。但目前人们所发现的纤维素酶存在效率低、成本高的问题^[3],无法在生产上大规模应用于对木质纤维素的降解转化,因此需要寻找高效的纤维素酶。

可在实验室条件下培养的微生物只占自然界中微生物种类的 1% 甚至更少^[4],占微生物种类 99% 以上的未培养微生物(uncultured microorganisms)中蕴藏了丰富的基因资源。宏基因组文库(metagenomic library)是指对环境样品中存在的所有微生物的混合 DNA 进行提取并将其克隆到合适的宿主菌中构建的文库^[5]。通过构建各种环境样品的宏基因组文库已筛选得到大量新的基因,如编码几丁质酶、蛋白酶、脂肪酶、淀粉酶、纤维素酶、抗生素等活性物质的基

因^[6]。因此,构建纤维素被活跃降解环境的宏基因组文库将有可能克隆到性能优良的纤维素酶基因。

造纸废水中富含纸浆,纸浆在废水排放过程中会在排放出口处逐渐沉淀下来并富集。纸浆的主要成分是纤维素,我们推测该环境中大量的微生物在进行纤维素的分解活动,这些微生物中可能含有丰富的纤维素酶基因资源。本工作的目的是采用未培养的手段对造纸废水纸浆沉淀物的细菌多样性进行分析,并构建宏基因组文库对该环境中的纤维素酶基因资源进行挖掘。

1 材料和方法

1.1 材料

1.1.1 样品采集 样品取自广西壮族自治区南宁市某造纸厂的排污口,主要成分是造纸废水中纸浆的沉淀物。在不同的采样点垂直取 1~15cm 深的纸浆沉淀物,每个点采样量大体一致。混合均匀后装入无菌保鲜袋,带回实验室 -80℃ 保存。

1.1.2 主要试剂:文库构建试剂盒 pWEB::TNC™ Cosmid Cloning Kit 购自 Epicentre 公司;pGEM-T easy 载体试剂盒购自 Promega 公司;限制性内切酶、Ex Taq DNA 聚合酶及 T4 DNA 连接酶购自 TaKaRa 公司;Sephadex G200 购自原 Pharmacia 公司;聚乙烯聚

基金项目:国家“863 计划”(2004AA214140);“国际科技合作重点项目计划”课题(2002AA217121)

* 通讯作者。Tel 86-771-3239401;Fax 86-771-3239413;E-mail:feng@public.nn.gx.cn

作者简介:许跃强(1981-),男,河北井陉人,硕士研究生,主要从事未培养微生物的基因克隆研究。E-mail:xuyq1981@sohu.com

其他作者:庞浩,封毅,莫新春,郭鸿,张鹏

收稿日期:2006-04-12;接受日期:2006-05-29;修回日期:2006-06-28

吡咯烷酮 (PVPP, PolyvinylPolypyrrolidone) 七叶苷 (Esculin hydrate) 柠檬酸高铁铵 (Ammonium iron(III) citrate) 4-MUC (4-methylumbelliferyl-beta-D-cellobioside) 及羧甲基纤维素钠 (Carboxymethyl cellulose sodium salt, CMC) 购自 Sigma 公司; 其它化学试剂购自上海生工生物工程有限公司。

1.2 宏基因组 DNA 的提取与纯化

宏基因组 DNA 的提取参照 Zhou 等^[7]的方法并做适当改进, 不同之处在于细胞裂解过程中将环境样品于液氮和 65°C 水浴中反复冻融 3 次。采用含 2% 酸洗 PVPP 的 Sephadex G200 凝胶柱纯化环境样品 DNA^[8], 再结合电洗脱回收 DNA 的方法获取造纸废水纸浆沉淀物的宏基因组 DNA。

1.3 造纸废水纸浆沉淀物中细菌的多样性分析

对细菌的 16S rDNA 的序列进行比较可以得知细菌之间的亲缘关系, 采用文献中广泛使用的根据 16S rDNA 的保守序列设计的引物 27f: 5'-AGAGTTTGATCATGGCTCAG-3' 和 1492r: 5'-TACGTTACCTTGTACGACTT-3'^[9] 扩增细菌的 16S rDNA。采用常规的 PCR 反应体系, PCR 反应条件为 96°C 2min, 94°C 40s, 68°C 30s, 72°C 1.5min, 共 30 个循环, 72°C 10min。

回收 PCR 产物, 以 pGEM-T Easy 为载体构建 16S rDNA 文库。随机挑取 16S rDNA 文库的克隆提取质粒测序, 用 NCBI (<http://www.ncbi.nlm.nih.gov/>) 的 BlastN 搜索 GenBank, 使用 DNAMAN Version 5.2.9 软件对所测序的 16S rDNA 序列及代表性细菌的 16S rDNA 序列进行多序列匹配排列, 以 NJ 法 (Neighbor-joining) 构建进化树, 并对所构建的进化树进行 Bootstrap 分析 (1000 个样本)。

1.4 宏基因组文库的构建

回收 30 ~ 50kb 的宏基因组 DNA, 参照 Epicentre 公司的 pWEB : :TNC™ Cosmid Cloning Kit 产品说明书进行末端补平后连接到柯斯载体 pWEB : :TNC 上, 侵染 *E. coli* EPH100 构建造纸废水纸浆沉淀物的宏基因组文库。

1.5 宏基因组文库中表达纤维素酶活性克隆的筛选

在对文库中的克隆进行筛选时, 采用氯仿熏蒸的方法对文库克隆进行破壁处理, 以使大肠杆菌中的蛋白质释放, 然后进行活性筛选。

内切葡聚糖酶活性克隆的筛选采用刚果红染色法^[10] 将文库克隆影印到含有 0.5% 羧甲基纤维素钠的 LA 平板上于 37°C 培养 24h, 用氯仿对文库克隆

熏蒸 15min, 37°C 培养箱放置 2h 后用 0.5% 刚果红溶液染色 15min, 再用 1mol/L 的 NaCl 溶液脱色, 菌落周围出现水解圈的即为阳性克隆。

β -葡萄糖苷酶活性克隆的筛选参考 Kwon^[11] 的方法, 将文库克隆影印到含 0.1% 七叶苷和 0.25% 柠檬酸高铁铵的 LA 平板上于 37°C 培养 12 ~ 16h, 用氯仿对文库克隆熏蒸 15min, 37°C 放置 2h, 菌落周围显黑色即为阳性克隆。

外切葡聚糖酶活性克隆的筛选参考 Heptinstall^[12] 的方法, 将文库克隆影印到 LA 平板上于 37°C 培养 12 ~ 16h 后, 用氯仿对文库克隆熏蒸 15min, 用含有 0.04% 4-MUC 的琼脂糖溶液 (0.6%) 倒上层平板, 37°C 培养 1h 后, 紫外灯下观察到菌落周围显荧光者为阳性克隆。挑选阳性克隆检测其是否具有 β -葡萄糖苷酶活性, 如果没有则可能为表达外切葡聚糖酶活性的克隆。

提取阳性克隆的质粒重新转化 *E. coli* EPH100, 以证实该基因表达的酶活确实来自于克隆的外源 DNA。

1.6 文库中纤维素酶基因的鉴定

在表达 3 种不同活性的克隆中分别选择活性最强的克隆进行亚克隆、测序。用 NCBI 的 ORF Finder 工具预测 DNA 片段上的开放阅读框, 用 BlastP 搜寻 GenBank 数据库, 将开放阅读框编码的蛋白质序列提交 SMART (<http://smart.embl-heidelberg.de/>) 进行结构域预测。

PCR 扩增纤维素酶基因, 扩增基因 *umcel5L* 的引物为 *umcel5L-F* (5'-AGCGGATCCCTGACACCGGAA GCACAGAC-3') 和 *umcel5L-R* (5'-GCCAAGCTTCTATT GCCCATATTCAGCA-3'); 扩增基因 *umcel5M* 的引物为 *umcel5M-F* (5'-AGCGGATCCATGAAGGGTGTCAACC TTGG-3') 和 *umcel5M-R* (5'-ACCAAGCTTTTCAGTGG TTTTCGACGAATT-3'); 扩增基因 *umbgl3D* 的引物为 *umbgl3D-F* (5'-AGCGAATTCCTGATCAAGGACGATGCC TATTACC-3') 和 *umbgl3D-R* (5'-GCCAAGCTTCCGGG TGATGCTAGTACGCTG-3')。在 *umcel5L* 和 *umcel5M* 的正向引物 5' 端加入 *Bam*H I 位点, 在 *umbgl3D* 的正向引物 5' 端加入 *Eco*R I 位点, 在 3 个基因的反向引物 5' 端均加入 *Hind*III 位点。

将相应的基因在 *E. coli* BL21 (DE3) pLys 或 *E. coli* Rosetta (DE3) 中表达并进行活性检测。

1.7 GenBank 索引号

16S rDNA 序列索引号为 DQ447303 ~ DQ447317 及 DQ483054, 3 个纤维素酶基因的索引号为:

umcel5L(DQ447318), *umcel5M*(DQ447319), *umbgl3D*(DQ447320)。

2 结果和分析

2.1 造纸废水纸浆沉淀物中宏基因组 DNA 的提取和纯化

采用本文所述的 DNA 提取方法, 1g 环境样品可得到 2μg 左右的粗提 DNA, 再使用含 2% 酸洗 PVPP 的 Sephadex G200 凝胶柱纯化后, DNA 损失量不到 20%, 并且很好地除去了可降解 DNA 的杂质。纯化后的 DNA 经琼脂糖凝胶电泳后回收 30kb ~ 50kb 的片段, 回收得率约为 50%, 得到的 DNA 可以满足构建柯斯文库的需要。

2.2 环境中细菌的多样性分析

构建的 16S rDNA 文库共含有 1863 个克隆, 随机挑选 18 个克隆提取质粒 DNA 并测其一端的序列, 结果得到 16 种不同的序列。经 BlastN 对比发现其中的 14 种序列与未培养细菌的 16S rDNA 具有较高的同源性, 2 种序列与未鉴定细菌的 16S rDNA 具

有较高的同源性(表 1)。选取代表性细菌的 16S rDNA 与本文所测得的 16S rDNA 一起构建系统进化树(图 1), 从进化树上可以看出这 16 种序列可被分为 4 个簇, 即螺旋体(Spirochaetes)、变形杆菌(Proteobacteria)、拟杆菌(Bacteroidetes)及厚壁菌(Firmicutes)。这些未培养细菌位于进化树不同的分支上, 说明该环境中的细菌具有种类的多样性。

2.3 宏基因组文库的构建及表达纤维素酶活性克隆的筛选

以 pWEB::TNC Cosmid Cloning vector 为载体构建了造纸废水纸浆沉淀物的宏基因组文库。该文库含有 10000 个克隆, 随机提取文库中的 12 个克隆的质粒用 BamHI 酶切, 结果表明所有的重组质粒都含有外源 DNA 片段, 最小为 20.7kb, 最大为 45.4kb, 平均大小为 35.3kb(图 2)。这些外源 DNA 片段的酶切带型都各不相同, 说明该宏基因组文库中克隆的外源 DNA 具有很好的随机性。该文库中外源 DNA 的克隆容量约为 3.53×10^8 bp。

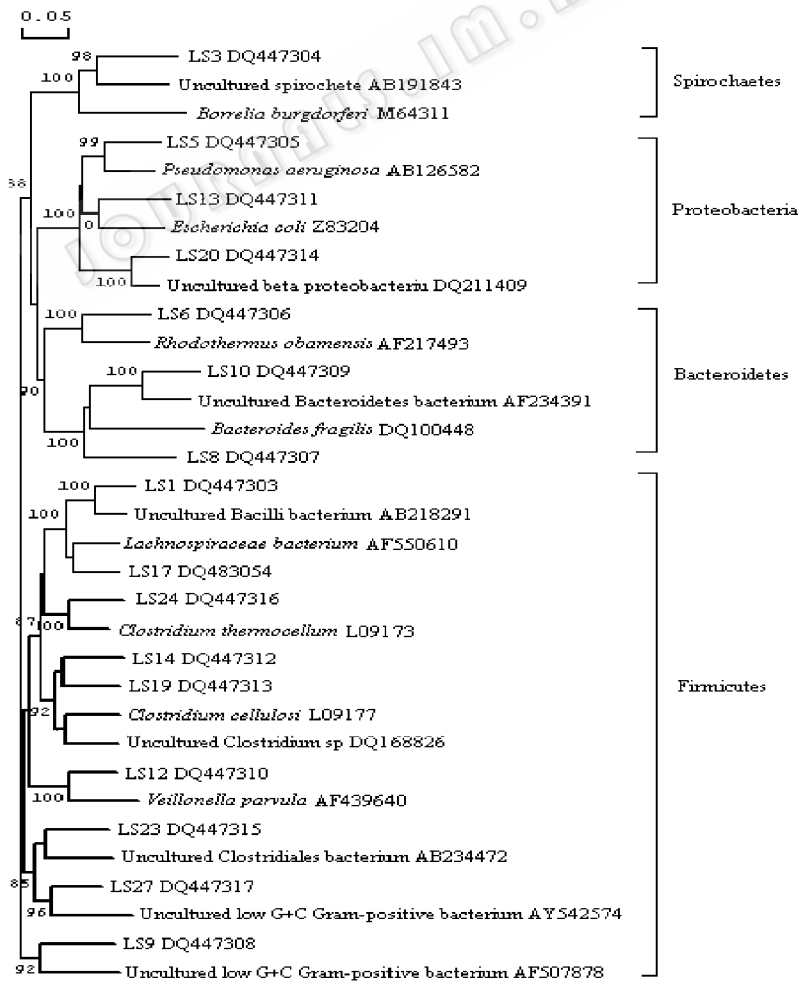


图 1 造纸废水纸浆沉淀物中的细菌与相近细菌的系统发育树

表 1 造纸废水纸浆沉淀物中细菌的 16S rDNA 序列分析结果

Table 1 Phylogenetic affiliations of 16S rDNA sequences of bacteria in pulp sediments from paper mill effluent

Representative clones	Nearest phylogenetic neighbor	16S rDNA Identity(%)
LS1(DQ447303)	Uncultured bacterium from slurry-compost (AB241562)	93
LS3(DQ447304)	Uncultured bacterium clone SIMO-2161 from salt marsh (AY711527)	88
LS5(DQ447305)	Uncultured bacterium clone CEB1 from swiss chard rhizoplane (AF392667)	95
LS6(DQ447306)	Uncultured bacterium clone AKIW1091 from urban aerosol (DQ129321)	94
LS8(DQ447307)	Uncultured anaerobic bacterium clone B-1BG from anaerobic swine lagoon (AY953168)	94
LS9(DQ447308)	Uncultured bacterium clone AKAU3609 from uranium contaminated soil (DQ125604)	98
LS10(DQ447309)	Uncultured bacterium clone COB P4-1 from termite intestinal tract (AY160848)	90
LS12(DQ447310)	Uncultured bacterium clone copi29 from copiotrophic artificial wastewater (AY563456)	99
LS13(DQ447311)	Uncultured Comamonadaceae bacterium clone from mangrove (DQ234109)	99
LS14(DQ447312)	Unidentified rumen bacterium RFN10 (AB009169)	91
LS17(DQ483054)	Unidentified eubacterium clone vadinHA42 from anaerobic digester fed by vinasses (UEU81762)	93
LS19(DQ447313)	Uncultured bacterium clone C735 from human stool sample (AY916345)	93
LS20(DQ447314)	Uncultured bacterium clone EV818BHEB5102702SA55 from subsurface water (DQ256346)	93
LS23(DQ447315)	Uncultured bacterium from bacterial consortia removing chlorine substituents from chlorobenzenes (UBA488074)	95
LS24(DQ447316)	Uncultured Clostridia bacterium clone X9Ba93 from anoxic rice field soil (AY607223)	93
LS27(DQ447317)	Uncultured bacterium HB69 from anaerobic digester (AF129867)	91

经过活性筛选得到 2 个表达内切葡聚糖酶活性的克隆, 3 个表达外切葡聚糖酶活性的克隆及 2 个表达 β -葡萄糖苷酶活性的克隆。提取这些克隆的质粒经 *Bam*H I 酶切分析表明这些克隆均为独立的重组子(数据未显示)。将这些克隆的质粒再次转化 *E. coli* EPI100 转化子在相应的底物上仍然表现出活性, 说明克隆的外源 DNA 使大肠杆菌表达了纤维素酶活性。

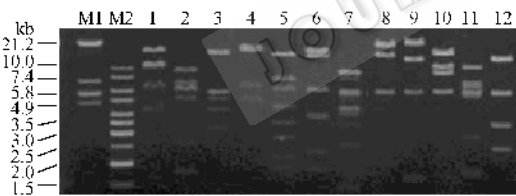


图 2 造纸废水纸浆沉淀物宏基因组文库中重组质粒的 *Bam*H I 酶切分析

Fig.2 Recombinant plasmids from the metagenomic library digested with *Bam*H I. M1 : *Eco*R I ; M2 : 1kb DNA ladder ; 1 ~ 12 : Recombinant plasmids digested with *Bam*H I .

2.4 纤维素酶基因的鉴定

2.4.1 内切葡聚糖酶基因 *umcel5L* 的鉴定

:挑选一个表达内切葡聚糖酶活性最强的克隆进行亚克隆, 将可能编码产生内切葡聚糖酶的基因定位在 6kb 的 *Hinc* II 片段上。序列分析发现该 DNA 片段上一个 1521bp 的开放阅读框的编码产物与大豆慢生根瘤菌(*Bradyrhizobium japonicum*)的一个内切葡聚糖酶(BAC48632)的同源性最高, 具有 43% 的一致性和 59% 的相似性。SMART 分析表明该多肽的第 1 ~ 25 个氨基酸残基为信号肽, 第 36 ~ 318 个氨基酸残基

为糖基水解酶家族 5 的催化功能域, 第 350 ~ 504 个氨基酸残基是一个碳水化合物结合组件(CBM, Carbohydrate Binding Module)。该基因在 *E. coli* Rosetta(DE3)中表达后使大肠杆菌表达有内切葡聚糖酶活性(图版 I -A-a), 从而证实该基因确实编码一个内切葡聚糖酶, 将该基因命名为 *umcel5L*。

2.4.2 纤维糊精酶基因 *umcel5M* 的鉴定

:挑选一个表达外切葡聚糖酶活性最强的克隆进行亚克隆, 将可能编码产生外切葡聚糖酶的基因定位在 4kb 的 *Bam*H I - *Hind* III 片段上。序列分析显示该 DNA 片段含有的一个 1053bp 的开放阅读框的编码产物与产琥珀酸丝状杆菌(*Fibrobacter succinogenes*)的一个纤维糊精酶(AAA50210)的同源性最高, 具有 48% 的一致性和 69% 的相似性。SMART 分析表明该多肽的第 12 ~ 320 个氨基酸残基为糖基水解酶家族 5 的催化功能域。该基因在 *E. coli* Rosetta(DE3)中表达后使大肠杆菌表达有外切葡聚糖酶活性(图版 I -A-b), 从而认为该基因编码一个纤维糊精酶, 将该基因命名为 *umcel5M*。

2.4.3 β -葡萄糖苷酶基因 *umbgl3D* 的鉴定

:挑选一个表达 β -葡萄糖苷酶活性最强的克隆进行亚克隆, 将可能编码 β -葡萄糖苷酶的基因定位在 2.5kb 的 *Pst* I 片段上。序列分析表明该 DNA 片段上一个 2382bp 的开放阅读框的编码产物与海栖热袍菌(*Thermotoga maritima*)的一个 β -葡萄糖苷酶(AAD35119)的同源性最高, 具有 46% 的一致性和 61% 的相似性。SMART 分析表明该多肽从第 1 ~ 19 个氨基酸残基为信号肽, 第 89 ~ 311 个氨基酸残基

为糖基水解酶家族 3 的催化功能域,第 375 ~ 643 个氨基酸残基为糖基水解酶家族 3C 的结构域。该基因在 *E. coli* BL21(DE3)_{pLys} 中表达后使大肠杆菌表达有 β -葡萄糖苷酶活性(图版 I-A-c),从而证实该基因确实编码一个 β -葡萄糖苷酶,将该基因命名为 *umbgl3D*。

3 讨论

本文采用未培养的手段对造纸废水纸浆沉淀物中细菌的多样性进行了初步研究。构建 16S rDNA 文库并测序,结果显示得到的 16 种独立的 16S rDNA 序列中有 14 种与 GenBank 中未培养细菌的 16S rDNA 的一致性最高,2 种与未鉴定细菌的 16S rDNA 的一致性最高,这 16 种 16S rDNA 在系统发育树上可分为螺旋体(Spirochaetes)、变形杆菌(Proteobacteria)、拟杆菌(Bacteroidetes)及厚壁菌(Firmicutes)4 个簇,这表明该环境中存在大量的未培养细菌且具有种类的多样性。Rappe 等^[13]指出未培养细菌大多属于新的细菌门,这些门中没有典型的可培养细菌,而本研究中的未培养细菌在系统发育树上分属于有典型可培养细菌的门,这说明未培养细菌不仅属于新的细菌门,而且存在于含可培养细菌的分类单元中。Hugenholz^[14]的研究显示大部分的可培养细菌属于变形杆菌、放线菌、厚壁菌、拟杆菌等四个门,从本文所构建的系统发育树看,该环境中没有放线菌。其原因可能有两点:(1)放线菌多为好氧菌,而造纸废水纸浆沉淀物的环境是缺氧的,该环境可能不适于放线菌的生存;(2)本工作只从 16S rDNA 文库中随机挑选了 18 个克隆进行测序,数量较少,有可能会漏掉该类细菌。另外,不同造纸厂所采用造纸原料、生产工艺的不同以及季节和地域的差异,其造纸废水纸浆沉淀物中细菌的多样性也可能不同。Yu 等^[15]曾对造纸废水处理过程中不同阶段的细菌多样性进行研究,其结果表明造纸废水中的优势细菌类群为变形杆菌和拟杆菌而没有厚壁菌,这表明造纸废水与造纸废水纸浆沉淀物中可能存在不同的细菌多样性。

当前在纤维素酶基因的克隆和鉴定方面,国内外主要是从木质纤维素被活跃降解的环境中分离具有纤维素酶活性的微生物,然后对获得的菌株进行鉴定并研究其产酶条件和纤维素酶的酶学特性,再进一步用分子生物学方法对纤维素酶基因进行克隆、鉴定和表达。随着人类对微生物认识的不断提高,从未培养微生物中直接克隆相关基因成为获取

各种活性物质的有效手段。有人从厌氧富集培养物^[16]、碱湖样品^[17]、土壤^[18]、鼠大肠^[19]及牛瘤胃^[20]的宏基因组文库克隆到新的纤维素酶基因,但还未见构建造纸废水纸浆沉淀物宏基因组文库并获得纤维素酶基因的报道。本研究从造纸废水纸浆沉淀物中提取宏基因组 DNA 并构建文库,经活性筛选得到 2 个表达内切葡聚糖酶活性的克隆,3 个表达外切葡聚糖酶活性的克隆及 2 个表达 β -葡萄糖苷酶活性的克隆。从中挑选活性克隆经鉴定得到 3 个新的纤维素酶基因,它们的编码产物与已报道纤维素酶的一致性都较低,这表明通过构建该环境样品的宏基因组文库能获得新的纤维素酶基因。

参 考 文 献

- [1] Lynd LR, Weimer PJ, Van Zyl WH, et al. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev*, 2002, **66**(3): 506 - 577.
- [2] Sun Y, Cheng J. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresour Technol*, 2002, **83**(1): 1 - 11.
- [3] Zhang YH, Lynd LR. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol Bioeng*, 2004, **88**(7): 797 - 824.
- [4] Amann RL, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev*, 1995, **59**(1): 143 - 169.
- [5] Handelsman J, Rondon MR, Brady SF, et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol*, 1998, **5**(10): R245 - 249.
- [6] Streit WR, Daniel R, Jaeger KE. Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Curr Opin Biotechnol*, 2004, **15**(4): 285 - 290.
- [7] Zhou J, Bruns MA, Tiedje JM. DNA recovery from soils of diverse composition. *Appl Environ Microbiol*, 1996, **62**(2): 316 - 322.
- [8] Wilson IG. Inhibition and facilitation of nucleic acid amplification. *Appl Environ Microbiol*, 1997, **63**(10): 3741 - 3751.
- [9] Medlin L, Elwood HJ, Stickel S, et al. The characterization of enzymatically amplified eukaryotic 16S-like ribosomal RNA-coding regions. *Gene*, 1988, **71**(2): 491 - 500.
- [10] Teather RM, Wood PJ. Use of Congo red-polysaccharide interactions in enumeration and characterisation of cellulolytic bacteria from the bovine rumen. *Appl Environ Microbiol*, 1982, **43**(4): 777 - 780.
- [11] Kwon Ki-sun, Lee Jaehoon, Kang Hyung-gyoo, et al. Detection of β -Glucosidase activity in polyacrylamide gels with esculin as substrate. *Appl Environ Microbiol*, 1994, **60**(12): 4584 - 4586.
- [12] Heptinstall J, Stewart JC, Seras M. Fluorimetric estimation of exo-cellobiohydrolase and β -D-glucosidase activities in cellulase from *Aspergillus fumigatus* Fresenius. *Enzyme Microb Technol*, 1986, **8**(2): 70 - 74.
- [13] Rappe MS, Giovannoni SJ. The uncultured microbial majority. *Annu Rev Microbiol*, 2003, **57**: 369 - 394.

- [14] Hugenholtz P. Exploring prokaryotic diversity in the genomic era. *Genome Biol* , 2002 , 3(2) : Reviews0003.1 – 0003.8.
- [15] Yu ZT , Mohn W. Bacterial diversity and community structure in an aerated lagoon revealed by ribosomal intergenic spacer analyses and 16S ribosomal DNA sequencing. *Appl Environ Microbiol* , 2001 , 67(4) : 1565 – 1574.
- [16] Healy FG , Ray RM , Aldrich HC , *et al.* Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic , anaerobic digester maintained on lingo-cellulose. *Appl Microbiol Biotechnol* , 1995 , 43(4) : 667 – 674.
- [17] Rees H , Grant S , Jones B , *et al.* Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* , 2003 , 7(5) : 415 – 421.
- [18] Voget S , Leggewie C , Uesbeck A , *et al.* Prospecting for novel biocatalysts in a soil metagenome. *Appl Environ Microbiol* , 2003 , 69(10) : 6235 – 6242.
- [19] Walter J , Mangold M , Tannock GW. Construction , analysis , and beta-glucanase screening of a bacterial artificial chromosome library from the large-bowel microbiota of mice. *Appl Environ Microbiol* , 2005 , 71(5) : 2347 – 2354.
- [20] Ferrer M , Golyshina OV , Chernikova TN , *et al.* Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ Microbiol* , 2005 , 7(12) : 1996 – 2010.

Cloning and identification of cellulase genes from uncultured microorganisms in pulp sediments from paper mill effluent

XU Yue-qiang^{1 2} , DUAN Cheng-jie^{1 2} , ZHOU Quan-neng² , TANG Ji-liang^{1 2} , FENG Jia-xun^{1 2*}

¹ Guangxi Key Laboratory of Subtropical Bioresources Conservation and Utilization ,

² College of Life Science and Technology , Guangxi University , Nanning 530005 , China)

Abstract : The metagenomic DNA of pulp sediments from paper mill effluent was extracted and purified. The 16S rDNA was amplified using the purified metagenomic DNA as template and a 16S rDNA library was prepared. Sequence analysis of 16S rDNA clones showed that diverse of uncultured bacteria inhabit in this environment , which can be classified into 4 clusters as *Spirochaetes* , *Proteobacteria* , *Bacteroidetes* and *Firmicutes* . A metagenomic library containing 10000 clones was constructed into cosmid vector , and the capacity of inserted DNA of which was 3.53×10^8 bp. Functional screening of the library resulted in isolation of two independent clones expressing endoglucanase activity , three independent clones expressing exoglucanase activity and two independent clones expressing β -glucosidase activity. One clone expressing strongest enzyme activity from each activity category was chosen to be further analyzed. Three novel cellulase genes designated as *umcel5L* , *umcel5M* and *umbgl3D* were identified by subcloning , sequencing and expression. The *umcel5L* encodes an endoglucanase belonging to glycosyl hydrolase family 5 , which is most related to an endoglucanase from *Bradyrhizobium japonicum* at 43% identity and 59% similarity. The *umcel5M* encodes a cellodextrinase belonging to glycosyl hydrolase family 5 , which is most similar to a cellodextrinase from *Fibrobacter succinogenes* at 48% identity and 69% similarity. The *umbgl3D* encodes a putative β -glucosidase belonging to glycosyl hydrolase family 3 , which shares highest homology with a β -glucosidase from *Thermotoga maritima* at 46% identity and 61% similarity. It is the first time to reveal the bacterial diversity of pulp sediments from paper mill effluent and clone novel cellulase genes from the bacteria by culture-independent method.

Keywords : Pulp sediments ; Bacterial diversity ; Uncultured microorganisms ; Metagenomic library ; Cellulase

Foundation item : National High Technology Research and Development Program of China (2004AA214140) ; International Science and Technology Collaborating Key Project Program of China (2002AA217121)

* Corresponding author. Tel 86-771-3239401 ; Fax 86-771-3239413 ; E-mail : feng@public.nn.gx.cn

Other authors : PANG Hao , FENG Yi , MO Xin-chun , GUO Hong , ZHANG Peng

Received : 12 April 2006 / Accepted 29 May 2006 / Revised : 28 June 2006