

微生物学报 *Acta Microbiologica Sinica*
54(1): 53–61; 4 January 2014
ISSN 0001–6209; CN 11–1995/Q
<http://journals.im.ac.cn/actamicrocn>
doi: 10.13343/j.cnki.wsxb.2014.01.007

瘤胃细菌 GH48 家族糖苷水解酶基因多样性

李劲亭^{1,2}, 苏小运¹, 田彦³, 董志扬¹, 胡松年³, 黄力¹, 戴欣^{1*}

¹中国科学院微生物研究所 微生物资源前期开发国家重点实验室 北京 100101

²中国科学院大学 北京 100101

³中国科学院北京基因组研究所 北京 100101

摘要 【目的】了解瘤胃细菌第 48 家族糖苷水解酶基因(GH48)多样性,为木质纤维素高效降解提供新的基因资源。【方法】通过基因序列比对,设计 *gh48* 的简并引物;同时提取两个瘤胃样品的总 DNA 和总 RNA,并将总 RNA 逆转录成 cDNA。以总 DNA 和 cDNA 为模板,通过 PCR 扩增分别建立克隆文库并对克隆文库进行测序;对所得序列进行 out 种类划分和聚类分析。【结果】本研究共得到了 455 条编码 GH48 家族蛋白的基因序列,核苷酸序列之间的相似性为 58.65%–100%。对序列的进一步分析表明,89% 可以作为区分其种类的界定标准。以此为依据确定所得到的基因序列分别编码 66 种不同的 GH48 家族蛋白,分别聚为 5 个相对独立的类群,其中新类群 C 中 OTU65 所代表的序列是 cDNA 克隆文库中的优势序列,分别占两个 cDNA 克隆文库的 36.4% 和 19.5%。我们的结果揭示瘤胃细菌 *gh48* 基因具有丰富的多样性,同时,其中也存在优势表达的 GH48 家族蛋白。

关键词: 瘤胃细菌,木质纤维素降解,GH48,糖苷水解酶,多样性

中图分类号: Q936 文献标识码: A 文章编号: 0001-6209(2014)01-0053-09

糖苷水解酶第 48 家族(Glycoside Hydrolase family 48, GH48)是一类重要的糖苷水解酶。该家族成员是纤维素外切酶的重要来源之一,可以通过和其它糖苷水解酶形成纤维小体或者以自由酶体系两种方式来降解纤维素^[1–2]。基因敲除实验证明, GH48 家族蛋白中的纤维素酶对热解纤维梭菌(*Clostridium thermocellum*)和白色瘤胃球菌(*Ruminococcus albus*)的纤维素降解活性十分重要:将 GH48 敲除以后前者的纤维素降解能力严重下降,而后者的纤维素降解能力几乎完全丧失^[1–3]。CAZy(Carbohydrate-Active EnZymes)数据库([http://](http://www.cazy.org)

www.cazy.org)显示,目前已知的 GH48 序列全部来自 4 个原核生物类群(放线菌 Actinobacteria、厚壁菌 Firmicutes、绿弯菌 Chloroflexi 和变形菌 Proteobacteria)和 2 个真核生物类群(厌氧真菌 Fungi 和节肢动物 Arthropoda), Sukharnikov 等的研究发现 Actinobacteria、Firmicutes 和 Chloroflexi 类群的 *gh48* 基因可能起源于一个共同的祖先,而 Proteobacteria、Fungi 和 Arthropoda 类群的 *gh48* 基因则是通过基因水平转移从 Actinobacteria 和 Firmicutes 获得的^[4]。

瘤胃是自然界中高效的纤维素降解系统之一,

基金项目: 国家“973 计划”(2011CB100804); 国家自然科学基金项目(31240050)

* 通信作者: Tel: +86-10-64807418/64807430; Fax: +86-10-64807429; E-mail: daixin@im.ac.cn

作者简介: 李劲亭(1985–),男,山东枣庄人,硕士研究生,从事极端环境微生物分子生态学研究。E-mail: ltting@163.com

收稿日期: 2013-04-27; 修回日期: 2013-06-07

其中蕴含了丰富的纤维素降解微生物及其编码的纤维素酶,但是目前瘤胃环境中仅有少量 *gh48* 基因序列被报道。2009 年 Brulc 等从 1025709 条瘤胃微生物元基因组序列中仅发现了 1 条 *gh48* 基因序列^[5]。另外 2011 年 Hess 等从奶牛瘤胃微生物 268G 元基因组数据仅发现了 3 条 *gh48* 基因序列^[6]。我们也曾经在 2012 年报道过类似的结果^[7]。迄今为止,只有在 2011 年一篇关于瘤胃真核微生物元转录组的文章中报道发现了较大量的 *gh48* 基因,在 19M 拼接后的 Illumina 测序数据中有 54 条 *gh48* 基因序列,但它们均来自瘤胃中的厌氧真菌^[8],而厌氧真菌在瘤胃总生物量中的比例是很低的^[9]。瘤胃纯培养微生物中,报道有 GH48 的细菌仅有白色瘤胃球菌 (*Ruminococcus albus*)、黄色瘤胃球菌 (*Ruminococcus flavefaciens*)、*Cellulosilyticum ruminicola*、丙酮丁酸梭菌 (*Clostridium acetobutylicum*), 厌氧真菌有 *Neocallimastix patriciarum*、*Piromyces equi*、*Piromyces* sp. E2, 而 GenBank 数据库也只有 13 条来源于这些微生物的 *gh48* 基因序列。这些都表明目前对瘤胃中第 48 家族糖苷水解酶的了解还十分有限。本实验通过分析 *gh48* 基因序列,设计了 *gh48* 基因简并引物并以此作为工具分析瘤胃微生物中 *gh48* 基因的多样性,为木质纤维素高效降解提供新的基因资源。

1 材料和方法

1.1 材料

1.1.1 主要试剂和仪器: 6870 Large Freezer/Mill 液氮冷冻研磨机,购自美国 SPEX SamplePrep 公司; G20 电动组织研磨仪,购自北京金银杏公司; RNA 提取与纯化试剂盒 Rneasy Plant Mini Kit,购自 QIAGEN 公司; *Premix Ex Taq* 和 Recombinant Dnase

I (Rnase-free), 购自 TaKaRa 公司; 逆转录试剂盒 ProtoScript M-MuLV *Taq* RT-PCR Kit, 购自 NEB 公司; KODFX DNA 聚合酶, 购自 TOYOBO 公司; 胶回收试剂盒 Gel-Extraction Kit, 购自 OMEGA 公司; 连接载体 *pEASY-Blunt* Simple Cloning vector 和大肠杆菌 (*Escherichia coli*) DH5 α 感受态细胞, 购自北京全式金公司。

1.1.2 瘤胃样品: 瘤胃样品采集自中国农业科学院北京畜牧兽医研究所昌平畜禽试验基地的瘦管奶牛。两头健康成年荷斯坦奶牛 (CS9007 和 MF9015) 分别饲喂玉米秸秆 (CS) 和混合饲料 (MF), 分别取其晨饲后 1h 的瘤胃样品, 使用 4 层纱布过滤瘤胃液, 获得瘤胃固相组分, 迅速放入液氮罐冷冻, 带回实验室 -80 $^{\circ}$ C 冰箱保存。

1.2 引物设计

从 NCBI (<http://www.ncbi.nlm.nih.gov/>) GenBank 数据库中获得 6 条瘤胃 *Ruminococcus* 的 *gh48* 基因序列, 其中 *R. flavefaciens* 1 条, *R. albus* 5 条, 使用 ClustalX 1.83 软件^[10] 进行序列比对, 选取它们序列的保守位点设计简并引物 RuGH48F/R (表 1); 其它引物见表 1。

1.3 瘤胃总 DNA 和总 RNA 提取及 cDNA 合成

1.3.1 总 DNA 提取: 采用 CTAB 法提取瘤胃总 DNA, 具体参见文献 [13]。

1.3.2 总 RNA 提取: 利用改进的 Trizol 法^[14] 提取总 RNA。即 Rneasy Plant Mini Kit 纯化 Trizol 法提取获得的总 RNA, 并于 -80 $^{\circ}$ C 冰箱保存纯化后的 RNA 样品。

1.3.3 cDNA 的合成: 细菌 16S rRNA 通用引物 (表 1, Universal primers) PCR 验证 RNA 溶液中没有基因组 DNA 残留后, 使用逆转录试剂盒 (ProtoScript M-MuLV *Taq* RT-PCR Kit) 的随机引物, 按照试剂盒操作说明进行逆转录反应。

表 1. 本研究所用的 PCR 引物

Table 1. The PCR primers used in this study

| primers | sequences(5'→3') | T_m / $^{\circ}$ C | reference |
|---|---|----------------------|------------|
| RuGH48F/R | F: GARGTCCTGACTACGGTCAYG R: TACCDRTCCAGTGGTTKGAACC | 52 $^{\circ}$ C | This study |
| ClGH48F/R | F: GGAATTCCKTAYCAYTCHATWGAAA R: CATGCCTGVAWWCCRAACCA | 52 $^{\circ}$ C | [11] |
| Universal primers of bacterial 16S rRNA | 27F: AGAGTTTGATCCTGGCTCAG 1492r: ACGGTTACCTTGTTACGACTT | 55 $^{\circ}$ C | [12] |

1.4 PCR 扩增和克隆文库的构建

1.4.1 PCR 扩增: 分别以两个瘤胃样品的总 DNA 和逆转录得到的 cDNA 为模板, 使用 KODFX 聚合酶进行 PCR 扩增。扩增程序: 94℃ 5 min; 94℃ 20 s, 52℃ 40 s, 68℃ 1 min 20 s, 35 个循环; 68℃ 10 min。4℃ 保存。

1.4.2 克隆文库构建: 对 PCR 产物进行琼脂糖凝胶电泳和胶回收试剂盒(Gel-Extraction Kit)回收 800 - 1500 bp 左右的条带。用 pEASY-Blunt Simple Cloning vector 载体 25℃ 连接 15 min, 转化 DH5 α 感受态细胞, 通过蓝白斑筛选, 挑白色菌落进行菌落 PCR 验证, 把有 1000 - 1200 bp 之间条带的克隆送去测序。

1.5 序列分析

1.5.1 有效序列的确定和相似性分析: 对测序获得的序列, 首先使用 NCBI 的 blastx (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 进行分析, 选取含有 GH48 保守结构域的序列作为有效序列, 使用 BioEdit 软件^[15]进行序列编辑和相似性分析。

1.5.2 种类区分标准的界定: 根据目前已知的厚壁菌中不同种内及种间 *gh48* 基因序列的相似性来确定本研究所获得序列的分类标准。从 NCBI (<http://www.ncbi.nlm.nih.gov/>) GenBank 下载了 25 条厚壁菌门不同属来源的 *gh48* 基因序列, 用 DNAMAN(LynnonBiosoft, USA) 进行同一属内种间和种内序列同源性比对, 确定区分该基因种类的标准。而后使用 ClustalX 软件进行序列比对, 将结果保存为 FASTA 格式, 用 CD-HIT 的在线分析程序 CD-HIT-EST (http://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi?cmd=cd-hit-est) 进行序列聚类分析。

1.5.3 基于 *gh48* 基因的系统聚类分析: 使用 MEGA5.1 (Molecular Evolutionary Genetics Analysis 5.1)^[16] 软件构建聚类树。

2 结果和分析

2.1 *gh48* 简并引物设计及 PCR 扩增

使用 Izquierdo 等根据梭菌 *gh48* 基因序列设计的保守引物 ClGH48F/R (表 1)^[11] 对瘤胃总 DNA 和 cDNA 样品进行 PCR 扩增, 尽管在 800 bp 至 1000 bp 之间有两条产物带 (图 1-A), 但是经过克隆测序发

现它们长度为 866 bp 和 1001 bp, 分别编码含 PIN domain (DNA polymerase I) 和 Capsule assembly protein Wzi domain (Gene involved in capsule biosynthesis) 的蛋白。阳性对照热纤维梭菌 (*Clostridium thermocellum* CGMCC1.2288) 获得 1200 bp 的 PCR 扩增条带 (图 1-B) 经克隆测序长度为 1167 bp, 所编码蛋白和 *C. thermocellum* GH48 蛋白同源性 100%, 推测瘤胃环境中已知的梭菌来源 *gh48* 基因数量较少, 无法通过 PCR 扩增反应获得。

已有的研究显示目前 *Ruminococcus* 是瘤胃中主要的纤维降解菌之一^[17-18], 并且全基因组测序数据显示这两种菌中都有单拷贝的 *gh48* 基因序列^[18-19], 因此我们利用已知的 6 条瘤胃来源 *Ruminococcus* 的 *gh48* 基因序列设计了一对保守引物 RuGH48F/R (表 1), 该引物对应 *R. albus* 8 的 *gh48* 基因 (GI: 325678899) 第 271 到 1399 位碱基, 预计扩增片段长度约 1200 bp。利用该引物对瘤胃总 DNA 和 cDNA 样品进行 PCR 扩增, 结果显示在 1200 bp 左右有单一的扩增条带 (图 1-C), 经过克隆测序证实这些序列编码含有 GH48 结构域的蛋白。

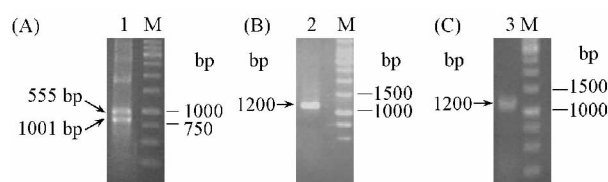


图 1. 不同引物 PCR 扩增产物电泳图

Figure 1. Analysis of PCR products using different primers by agarose gel electrophoresis. M, DNA Marker; lane 1, PCR products of primer ClGH48F/R using rumen total DNA as template; lane 2, PCR products of primer ClGH48F/R using total DNA of *C. thermocellum* CGMCC1.2288 as template; lane 3, PCR products of primer RuGH48F/R using rumen total DNA as template.

2.2 *gh48* 基因克隆文库的建立及序列测定

以瘤胃样品 CS9007 和 MF9015 的总 DNA 和 cDNA 为模板分别进行 PCR 扩增, 建立 *gh48* 基因的 DNA 克隆文库和 cDNA 克隆文库, 4 个文库一共获得了 455 条含 GH48 结构域 (Pfam02011) 的序列, 其中 CS9007 DNA 文库 94 条, CS9007 cDNA 文库 107 条, MF9015 DNA 文库 131 条, MF9015 cDNA 文库 123 条。

由于 *gh48* 基因是功能基因, 为了对其进行种类区分, 我们分析了所有厚壁菌中梭菌属 (*Clostridium*)、瘤胃球菌属 (*Ruminococcus*)、芽胞杆

菌属(*Bacillus*)和类芽胞杆菌属(*Paenibacillus*)属内种间及种内 $gh48$ 基因核苷酸序列的相似性(只统计基因组中有唯一 $gh48$ 基因序列的物种)。如图2所示,来自同一细菌种内不同菌株间的20条 $gh48$ 基因序列之间相似性都在89%以上,甚至达到

100%;而来自同属不同种的41条 $gh48$ 基因序列中40条序列之间相似性都小于89%,仅溶纤维梭菌(*Clostridium cellulolyticum* H10)和*Clostridium josui* JCM17888相似性为92.19%。据此把序列相似性 $\geq 89\%$ 定义为同一种GH48家族基因的标准。

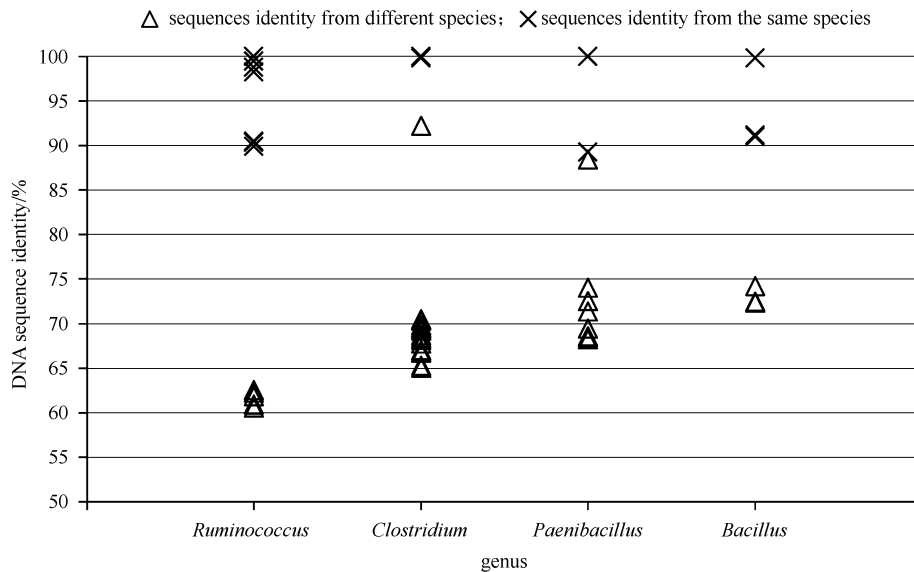


图2. 厚壁菌类群中不同属种来源的 $gh48$ 基因序列的相似性分布

Figure 2. The distribution of identities of $gh48$ gene sequences from different species/stains of Firmicutes. The species and their $gh48$ GI number: *Ruminococcus*: *R. flavefaciens* FD-1 (268611633), *R. albus* 8 (37703325), *R. albus* 8 (325678899), *R. albus* 20 (213032416), *R. albus* 7 (317057248); *Clostridium*: *C. acetobutylicum* EA 2018 (325508326), *C. acetobutylicum* DSM 1731 (336292067), *C. acetobutylicum* ATCC824 (15023811), *C. cellulolyticum* H10 (219998507), *C. cellulovorans* 743B (302578507), *C. cellulovorans* ATCC35296 (5705874), *C. josui* JCM17888 (3445478), *C. lentocellum* DSM5427 (326543492), *C. phytofermentans* ISDg (160430158), *C. saccharoperbutylacetonicum* N1-4 (451784660); *Paenibacillus*: *P. barcinonensis* BP-23 (21449824), *P. mucilaginosus* KNP414 (336301605), *P. terrae* HPL-003 (357203298), *P. polymyxa* SC2 (309246934), *P. polymyxa* M1 (343097005), *P. polymyxa* E681 (305858347); *Bacillus*: *B. pumilus* SAFR-032 (157681097), *B. licheniformis* DSM 13 (52003474), *B. licheniformis* SVD1 (371940202), *B. licheniformis* MSB03 (264670908).

以89%作为 $gh48$ 基因种类区分的标准,我们共获得66个不同种的基因序列,其中CS9007 DNA文库34种,CS9007 cDNA文库33种,MF9015 gDNA文库42种,MF9015 cDNA文库32种。4个克隆文库的稀释度曲线(Rarefaction curves)见图3。从稀释度曲线趋势可以看出,MF9015 gDNA文库(131条序列)和MF9015 cDNA文库(123条序列)趋于饱和但尚未饱和,可见瘤胃微生物中 $gh48$ 基因序列多样性非常丰富。

2.3 瘤胃细菌 $gh48$ 基因序列分析

4个克隆文库共获得了455条 $gh48$ 基因序列,这些基因序列之间的同源性从58.65%到100%,它们与已知的 $gh48$ 基因序列同源性从60.38%–99.03%,455条序列分别属于66个不

同的种,每种选1个代表序列,和其他类群来源的其他微生物 $gh48$ 基因序列的系统聚类分析结果见图4。

聚类分析结果显示,66种 $gh48$ 基因分为5个类群,类群A和D分别和*R. albus*和*R. flavefaciens*聚在一起,B、C、E则相对独立。表2的统计结果表明,*R. flavefaciens*代表的类群A中共有28种 $gh48$,多样性最丰富,其在DNA克隆文库中的序列数(44和77)明显高于cDNA克隆文库获得的序列数(18和25);以*R. albus*代表的类群D的15种 $gh48$,虽然cDNA克隆文库中的序列数(29和44)高于DNA克隆文库数(23和24),但cDNA序列中大部分来自OTU36,均为其文库克隆总数的48%(14/29和21/44),表明瘤胃中大部分已知的*R. flavefaciens*和*R.*

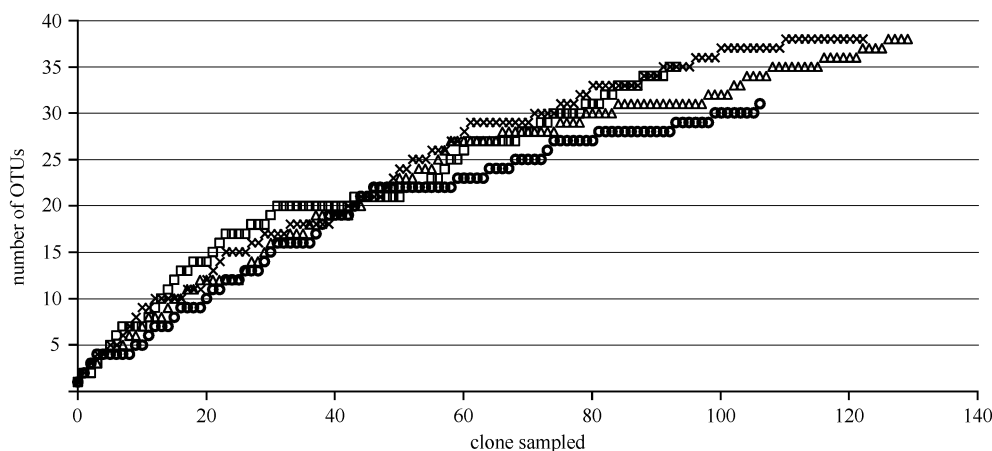


图 3. 克隆文库的稀释度曲线

Figure 3. Rarefaction curves for clone libraries of CS9007 DNA(open circle) , CS9007cDNA(open triangle) , MF9015 DNA (plus sign) , MF9015cDNA(times sign) .

albus 来源的 *gh48* 基因数量及多样性丰富但其相对表达量并不占优势; 而类群 C 尽管只有 3 种 *gh48* 基因, 但其 cDNA 克隆文库中的数量(40 和 25) 明显高于 DNA 克隆文库(12 和 14) , 分别占 2 个 cDNA 克隆文库的 37% 和 20% 左右, 表明它们代表瘤胃中较高表达的新的 *gh48* 基因类群。

在 CS9007 DNA 文库中, OTU35 (12. 87%) 、 OTU65(11. 71%) 、 OTU36(9. 57%) 代表的序列所占的比例最高, 在 MF9015 DNA 文库中, OTU18 (22. 9%) 、 OTU65(9. 92%) 、 OTU4 (6. 1%) 所占的比例最高, 表明 OTU65 在两个不同饲喂条件的健康瘤胃环境微生物 DNA 中均具有较高的丰度。对应 CS9007 cDNA 文库中, OTU65 (36. 45%) 、 OTU36 (13. 1%) 、 OTU64 (4. 67%) 、 OTU38 (3. 73%) 代表的序列所占的比例最高, MF9015 cDNA 文库中 OTU65 (19. 5%) 、 OTU36 (17. 07%) 、 OTU38 (10. 57%) 最高, 可见尽管在 DNA 水平不同的样品高丰度的 OTUs 类群不同, 但是 cDNA 水平上高丰度的 OTUs 类群非常接近。而 OTU65 代表的克隆序列无论是 DNA 还是 cDNA 中都有较高的丰度, 位于相对独立的分支 C 中, 与白色瘤胃球菌和黄色瘤胃球菌距离均较远, 提示该基因可能在纤维素降解中发挥重要的作用。

此外, OTU36 代表的克隆在两个样品的 cDNA 文库中都有较高的丰度, 并且它们在 cDNA 中的丰度明显高于其在 DNA 中的丰度, 提示该基因存在比较明显的高表达现象。

3 讨论

我们的研究结果表明瘤胃中有大量 *Ruminococcus* 来源的 *gh48* 基因序列, 并且这些基因有着丰富的多样性。那么为什么之前的瘤胃元基因组数据却只发现了很少的 *gh48* 基因序列呢(尤其是瘤胃细菌来源的)? 我们推测原因可能由于 *Ruminococcus* 仅占瘤胃细菌总数的约 0. 2%^[20-22], 并且 *gh48* 基因在 *Ruminococcus* 基因组中又是以单拷贝的形式存在^[18-19]。

已有的报道显示, CAZY 中大部分 *gh48* 是来自厌氧环境中的梭菌, 而梭菌亦是瘤胃微生物中较优势的类群^[20-23], 因此我们使用 Izquierdo 等根据梭菌来源的 *gh48* 基因序列设计的保守引物 CIGH48F/R 对瘤胃样品总 DNA 和 cDNA 样品进行扩增, 但并未得到 *gh48* 基因序列。事实上, 目前用纯培养手段从瘤胃中获得的 *Clostridium* 主要有: *Clostridiumkluyveri*^[24]; *Clostridiumlundense*^[25]; *Clostridium cellobioparum*, *Clostridiumlongisporum*, *Clostridiumlochheadii*, *Clostridiumbifermentans*^[26]; *Clostridiumbeijerinckii*, *Clostridiumacetobutylicum*, *Clostridium butyricum*^[27]; *Clostridiumproteoclasticum*^[28], 但是 GenBank 和 CAZY 数据库显示这些瘤胃来源 *Clostridium* 中有 GH48 编码基因的只有 *C. acetobutylicum*。我们的瘤胃元转录组研究数据也显示, 占总数 75% 以上的元转录组

表 2. 不同克隆文库中 *gh48* 基因类群数量分布
Table 2. Clone numbers of five *gh48* gene clusters
in different clone libraries

| phylogenetic clusters | species | CS9007 | | MF9015 | |
|-----------------------|---------|--------|------|--------|------|
| | | DNA | cDNA | DNA | cDNA |
| A | 28 | 43 | 18 | 77 | 25 |
| B | 15 | 14 | 15 | 14 | 25 |
| C | 3 | 12 | 40 | 14 | 25 |
| D | 15 | 23 | 29 | 24 | 44 |
| E | 5 | 2 | 5 | 2 | 4 |
| Total | 66 | 94 | 107 | 131 | 123 |

gh48 基因片段,其 BlastX 的最佳比对结果都是 *Ruminococcus* 来源的 GH48,这说明 *Ruminococcus* 来源的 GH48 在瘤胃中是占优势的 GH48 类群,因此,本研究所获得 *gh48* 基因序列的多样性在很大程度上反映了瘤胃细菌 *gh48* 基因序列的多样性。此外,Qi 等报道瘤胃中也有大量的厌氧真菌来源的 *gh48* 基因^[8]。

我们一共获得了 455 条 *gh48* 基因序列,而目前 CAZy 数据库中只有 166 条 *gh48* 基因序列,尽管我们获得的只是 *gh48* 基因序列的片段(约占 *gh48* 基因全长的 50%),但这是对 *gh48* 基因数据库一个很好的补充。并且我们设计的引物也可以尝试用于其他环境样品的研究,寻找更多的 *gh48* 基因资源。

通过对所获得的序列的分析,我们发现 OTU65 所代表的序列无论是在 DNA 水平还是 cDNA 水平都有较高的丰度,并且其在两个样品 cDNA 中丰度是 DNA 中丰度的 3.11 倍和 1.96 倍,即存在明显的高转录现象,并且其它几个 cDNA 中高丰度的 OTU 也存在同样的现象,已知 *gh48* 基因在 *Ruminococcus* 的基因组中单拷贝存在^[18-19],表明 GH48 家族糖苷水解酶需要高转录水平来发挥它的纤维素降解功能,同时也说明 OTU65 所代表的 *gh48* 基因序列在瘤胃纤维素降解体系中应该发挥着重要的作用,后续的工作我们希望通过反向 PCR 等其他分子生物学手段获得该 OTU65 所代表的 *gh48* 基因的全长序列,通过体外克隆表达分析它的纤维降解功能特性。

致谢:感谢中国农业科学院北京畜牧兽医研究所反刍动物营养研究室团队的所有成员,特别感谢王加启研究员、卜登攀副研究员和赵圣国博士在实验中给予的大力支持和帮助;感谢中国科学院北京基因组研究所骆迎峰副研究员和中国科学院微生物研究

所王雪薇副研究员、刘力博士和杨顶同学在实验中给予的大力支持和帮助。

参考文献

- [1] Olson DG, Tripathi SA, Giannone RJ, Lo J, Caiazza NC, Hogsett DA, Hettich RL, Guss AM, Dubrovsky G, Lynd LR. Deletion of the Cel48S cellulase from *Clostridium thermocellum*. *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(41): 17727-17732.
- [2] Steenbakkers PJ, Freelove A, Van Cranenbroek B, Sweegers BM, Harhangi HR, Vogels GD, Hazlewood GP, Gilbert HJ, Op den Camp HJ. The major component of the cellulosomes of anaerobic fungi from the genus *Piromyces* is a family 48 glycoside hydrolase. *DNA sequence: the Journal of DNA Sequencing and Mapping*, 2002, 13(6): 313-320.
- [3] Devillard E, Goodheart DB, Karnati SK, Bayer EA, Lamed R, Miron J, Nelson KE, Morrison M. *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. *Journal of Bacteriology*, 2004, 186(1): 136-145.
- [4] Sukharnikov LO, Alahuhta M, Brunecky R, Upadhyay A, Himmel ME, Lunin VV, Zhulin IB. Sequence, structure, and evolution of cellulases in glycoside hydrolase family 48. *The Journal of Biological Chemistry*, 2012, 287(49): 41068-41077.
- [5] Brulc JM, Antonopoulos DA, Miller ME, Wilson MK, Yannarell AC, Dinsdale EA, Edwards RE, Frank ED, Emerson JB, Wacklin P, Coutinho PM, Henrissat B, Nelson KE, White BA. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(6): 1948-1953.
- [6] Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z, Rubin EM. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science*, 2011, 331(6016): 463-467.
- [7] Dai X, Zhu Y, Luo Y, Song L, Liu D, Liu L, Chen F, Wang M, Li J, Zeng X, Dong Z, Hu S, Li L, Xu J, Huang L, Dong X. Metagenomic insights into the

- fibrolytic microbiome in yak rumen. *PLoS One* ,2012 ,7 (7) : e40430.
- [8] Qi M , Wang P , O'Toole N , Barboza PS , Ungerfeld E , Leigh MB , Selinger LB , Butler G , Tsang A , McAllister TA , Forster RJ. Snapshot of the eukaryotic gene expression in muskoxen rumen—a metatranscriptomic approach. *PLoS One* ,2011 ,6(5) : e20521.
- [9] Lin C , Raskin L , Stahl DA. Microbial community structure in gastrointestinal tracts of domestic animals: comparative analyses using rRNA-targeted oligonucleotide probes. *FEMS Microbiology Ecology* ,1997 ,22(4) : 281–294.
- [10] Thompson JD , Gibson TJ , Plewniak F , Jeanmougin F , Higgins DG. The CLUSTAL _ X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* , 1997 , 25(24) : 4876–4882.
- [11] Izquierdo JA , Sizova MV , Lynd LR. Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Applied and Environmental Microbiology* , 2010 , 76(11) : 3545–3553.
- [12] Ying YX , Ding WL , Li Y. Characterization of soil bacterial communities in rhizospheric and nonrhizospheric soil of *Panax ginseng*. *Biochemical Genetics* , 2012 , 50 (11–12) : 848–859.
- [13] Minas K , McEwan NR , Newbold CJ , Scott KP. Optimization of a high-throughput CTAB-based protocol for the extraction of qPCR-grade DNA from rumen fluid , plant and bacterial pure cultures. *FEMS Microbiology Letters* ,2011 ,325(2) : 162–169.
- [14] Kang S , Denman SE , Morrison M , Yu Z , McSweeney CS. An efficient RNA extraction method for estimating gut microbial diversity by polymerase chain reaction. *Current Microbiology* ,2009 ,58(5) : 464–471.
- [15] Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/ 98/NT. *Nucleic Acids Symposium Series* , 1999 , 41: 95–98.
- [16] Kumar S , Tamura K , Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics* ,2004 ,5 (2) : 150–163.
- [17] Flint HJ , Bayer EA , Rincon MT , Lamed R , White BA. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nature Reviews. Microbiology* ,2008 ,6(2) : 121–131.
- [18] Suen G , Stevenson DM , Bruce DC , Chertkov O , Copeland A , Cheng JF , Detter C , Detter JC , Goodwin LA , Han CS , Hauser LJ , Ivanova NN , Kyrpides NC , Land ML , Lapidus A , Lucas S , Ovchinnikova G , Pitluck S , Tapia R , Woyke T , Boyum J , Mead D , Weimer PJ. Complete genome of the cellulolytic ruminal bacterium *Ruminococcus albus* 7. *Journal of Bacteriology* , 2011 , 193(19) : 5574–5575.
- [19] Berg Miller ME , Antonopoulos DA , Rincon MT , Band M , Bari A , Akraiko T , Hernandez A , Thimmapuram J , Henrissat B , Coutinho PM , Borovok I , Jindou S , Lamed R , Flint HJ , Bayer EA , White BA. Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of *Ruminococcus flavefaciens* FD-1. *PLoS One* ,2009 ,4(8) : e6650.
- [20] Sirohi SK , Singh N , Dagar SS , Puniya AK. Molecular tools for deciphering the microbial community structure and diversity in rumen ecosystem. *Applied Microbiology and Biotechnology* ,2012 ,95(5) : 1135–1154.
- [21] Stevenson DM , Weimer PJ. Dominance of *Prevotella* and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. *Applied Microbiology and Biotechnology* , 2007 , 75 (1) : 165–174.
- [22] Weimer PJ , Waghorn GC , Odt CL , Mertens DR. Effect of diet on populations of three species of ruminal cellulolytic bacteria in lactating dairy cows. *Journal of Dairy Science* ,1999 ,82(1) : 122–134.
- [23] Boguhn J , Strobel E , Witzig M , Tebbe CC , Rodehutschord M. Description of the structural diversity of rumen microbial communities in vitro using single-strand conformation polymorphism profiles. *Archives of Animal Nutrition* ,2008 ,62(6) : 454–467.
- [24] Weimer PJ , Stevenson DM. Isolation , characterization , and quantification of *Clostridium kluyveri* from the bovine rumen. *Applied Microbiology and Biotechnology* , 2012 , 94(2) : 461–466.
- [25] Cirne DG , Delgado OD , Marichamy S , Mattiasson B. *Clostridium lundense* sp. nov. , a novel anaerobic lipolytic bacterium isolated from bovine rumen. *International Journal of Systematic and Evolutionary Microbiology* , 2006 ,56(Pt 3) : 625–628.
- [26] DN K. Rumen microbial ecosystem. *Curr Sci* 2005 ,89: 124–135.
- [27] Sankar M , Delgado O , Mattiasson B. Isolation and

characterization of solventogenic, cellulase-free xylanolytic Clostridia from cow rumen. *Water Science and Technology: a Journal of the International Association on Water Pollution Research*, 2003, 48(4): 185-188.

[28] Reilly K, Attwood GT. Detection of *Clostridium proteoclasticum* and closely related strains in the rumen by competitive PCR. *Applied and Environmental Microbiology*, 1998, 64(3): 907-913.

Gene diversity of the bacterial 48 family glycoside hydrolase(GH48) in rumen environment

Jingting Li^{1 2}, Xiaoyun Su¹, Yan Tian³, Zhiyang Dong¹, Songnian Hu³, Li Huang¹, Xin Dai^{1*}

¹State Key Laboratory of Microbial Resource, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

²University of Chinese Academy of Sciences, Beijing 100101, China

³Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

Abstract [Objective] The gene diversity of the bacterial 48 family glycoside hydrolase(GH48) in rumen environment was studied and new gene resources for efficient cellulose degradation were provided. **[Methods]** A pair of *gh48* degenerate primers was designed through sequences alignment of the *gh48* gene sequences from ruminal *Ruminococcus*. The total DNA and RNA were extracted from two rumen samples and cDNA was synthesized through reverse transcription from total RNA. Four *gh48* gene clone libraries were constructed and analyzed. **[Results]** In total 455 *gh48* gene sequences were obtained from the 4clone libraries. Sequence similarity among the 455 gene sequences varies between 58.65% and 100%. They fell into 66 species with the sequence similarity $\geq 89\%$, and divided into 5 different clusters. OTU65 in cluster C represents an abundant *gh48* gene which in both DNA and cDNA clones libraries, accounting for 36.4% and 19.5% respectively. Our study reveals rich gene diversity of the 48 family glycoside hydrolase and provided new gene resources for cellulose degradation.

Keywords: rumen bacteria, lignocellulose degradation, 48 family glycoside hydrolase, gene diversity

(本文责编: 王晋芳)

Supported by the National Basic Research Program of China(2011CB100804) and by the National Natural Science Foundation of China(31240050)

* Corresponding author. Tel: +86-10-64807418/64807430; Fax: +86-10-64807429; E-mail: daixin@im.ac.cn

Received: 27 April 2013/Revised: 7 June 2013