

硬粒小麦品种 Langdon 低分子量谷蛋白 亚基新基因的克隆

高双成, 王世华, 史国安, 孔祥生, 李友军

(河南科技大学 农学院, 河南 洛阳 471003)

摘要: 为了进一步了解小麦 *Glu-3* 位点的基因组成, 采用 PCR 方法从硬粒小麦品种 Langdon 中克隆了 2 个低分子量谷蛋白亚基(LMW-GS)新基因, 暂命名为 *LMW-LDN1*、*LMW-LDN2*, GenBank 注册号为 DQ659333 和 DQ659334。这 2 个基因都具有 LMW-GS 的典型结构特征, 编码区全长分别为 1 005 bp 和 894 bp, 编码 334 个和 297 个氨基酸。通过 *Langdon1D(1A)* 和 *Langdon1D(1B)* 代换系定位分析表明, *LMW-LDN1* 和 *LMW-LDN2* 位于 *Glu-A3* 位点。

关键词: 硬粒小麦; 低分子量谷蛋白亚基; 基因克隆; 序列分析

中图分类号: S512.1 **文献标志码:** A **文章编号:** 1004-3268(2013)08-0008-04

Cloning of Novel LMW-GS Genes from Durum Wheat Cultivar Langdon

GAO Shuang-cheng, WANG Shi-hua, SHI Guo-an, KONG Xiang-sheng, LI You-jun

(Agricultural College, Henan University of Science & Technology, Luoyang 471003, China)

Abstract: The low-molecular-weight glutenin subunits (LMW-GS) encoded at *Glu-3* loci are important for the end-use quality of durum wheat. The present paper reported two novel LMW-GS genes (*LMW-LDN1* and *LMW-LDN2*) isolated from durum wheat cultivar Langdon using PCR method. The full coding regions of *LMW-LDN1* (GenBank accession number: DQ659333) and *LMW-LDN2* (GenBank accession number: DQ659334) were composed of 1005 and 894 nucleotides, which were translated into proteins of 334 and 297 amino acids, respectively. They all had the typical structure of LMW-GS genes. The two genes were assigned to *Glu-A3* locus through the analysis of *LDN1D(1A)* and *LDN1D(1B)* substitution lines.

Key words: durum wheat; LMW-GS; gene cloning; sequence analysis

小麦种子贮藏蛋白占小麦种子总蛋白的 80% 以上, 主要存在于胚乳中。按其在胚乳中的天然构象可分为两大类: 单体蛋白(即醇溶蛋白)和依靠分子间二硫键连接起来的聚合体蛋白(麦谷蛋白)。根据十二烷基硫酸钠-聚丙烯酰胺凝胶电泳(SDS-PAGE)迁移率, 麦谷蛋白可分为分子量约为 70~140 kD 的高分子量麦谷蛋白亚基(HMW-GS)和分子量为 12~50 kD 的低分子量麦谷蛋白亚基(LMW-GS)。HMW-GS 分别由位于第一同源群染色体长臂上的 *Glu-A1*、*Glu-B1*、*Glu-D1* 基因

(统称 *Glu-1*) 编码, 而 LMW-GS 则分别由位于其短臂上的 *Glu-A3*、*Glu-B3*、*Glu-D3* 基因(统称 *Glu-3*) 编码^[1-2]。小麦低分子量谷蛋白对小麦的加工品质有重要影响, 但由于其受多基因控制, 且不易与醇溶蛋白分开, 所以对其研究的程度不如对小麦高分子量谷蛋白深入。在小麦低分子量谷蛋白基因克隆方面, 截至目前, 在 GenBank 中已注册的 LMW-GS 核苷酸序列有 586 条, 它们主要来自普通小麦、硬粒小麦、栽培一粒小麦、粗山羊草、中间偃麦草等材料。序列分析表明, 低分子量谷蛋白基因

收稿日期: 2013-01-26

基金项目: 国家自然科学基金项目(30571158); 河南科技大学人才科学研究基金项目(09001321)

作者简介: 高双成(1976-), 男, 山西夏县人, 副教授, 主要从事植物基因组学研究。E-mail: gsczml@163.com

有4个主要的结构域——20个氨基酸组成的信号肽、13个氨基酸组成的短的N末端、富含谷氨酰胺的重复序列区和C末端。小麦低分子量谷蛋白基因长度的变化主要源于富含谷氨酰胺的重复序列区重复单位次数的不同^[3]。Langdon是一个古老的硬粒小麦品种,Nieto-Taladriz等在研究硬粒小麦品种低分子量谷蛋白亚基组成时将其作为标准品种^[4]。Langdon品种的低分子量谷蛋白亚基组成为5+8+9+13+16,亚基5由Glu-A3编码,其他亚基(8+9+13+16)由Glu-B3编码^[4]。目前从Langdon品种中分离了5个低分子量谷蛋白基因序列(GenBank注册号为AJ293097、AJ293098、AJ293099、Y14104、AY146587),其中4个为Glu-A3位点基因,1个为Glu-B3位点基因。本研究根据已报道的低分子量谷蛋白基因保守序列设计PCR引物,克隆硬粒小麦Langdon品种中其他的低分子量谷蛋白基因,旨在为进一步研究低分子量谷蛋白基因在小麦加工品质中的作用奠定基础。

1 材料和方法

1.1 材料

硬粒小麦(*T. durum*)品种Langdon及其1A和1B的代换系—LDN1D(1A)、LDN1D(1B)由UDSA-ARS, Northern Crop Science Laboratory的Xu S S博士提供;携带Glu-3基因的BAC克隆(来自Langdon BAC文库)由USDA-ARS, Western Regional Research Center的Gu Y Q博士提供。

1.2 方法

1.2.1 硬粒小麦和BAC质粒DNA的提取 取硬粒小麦品种Langdon及其代换系的幼嫩叶片,液氮冷冻研磨成细粉后,采用酚-氯仿法^[5-6]进行DNA提取。BAC质粒DNA采用碱裂解法提取。

1.2.2 硬粒小麦LMW-GS基因编码区的PCR扩增 根据NCBI数据库中报道的LMW-GS基因设计了扩增低分子量谷蛋白编码区的1对兼并保守引物,Primer1为5'-ATGAAGACCTTCCTC(G/A)TCTTTG-3',Primer2为5'-TCAGTAG(G/A)CACCAACTC(C/G)G(G/A)T-3'。PCR反应总体积为25 μ L,其中含10 \times PCR Buffer 2.5 μ L,MgCl₂ 1.8 mmol/L,4种dNTP各200 μ mol/L,引物0.2 μ mol/L,Taq DNA聚合酶1 U,基因组或BAC DNA 50~100 ng。PCR反应条件为95 $^{\circ}$ C预变性5 min,然后95 $^{\circ}$ C变性45 s,60 $^{\circ}$ C退火45 s,72 $^{\circ}$ C延伸1 min,共35个循环,最后72 $^{\circ}$ C延伸10 min。扩增产物在1.5%琼脂糖凝胶上分离,恒定

电压下电泳。

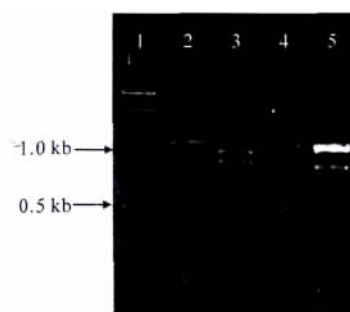
1.2.3 硬粒小麦PCR产物回收和克隆 以DNA凝胶回收试剂盒(天根生化)回收目的DNA片段,并与pCF-T载体连接。连接体系为6 μ L,其中包括目的DNA片段4 μ L、pCF-T载体50 ng、连接混合液1 μ L,在22~26 $^{\circ}$ C条件下反应5 min。取2 μ L连接产物加入到50 μ L TOP10感受态细胞中,42 $^{\circ}$ C 1.5 min热激转化,然后加入200~300 μ L SOC培养基,200 r/min、37 $^{\circ}$ C培养45 min。取200 μ L复苏菌液涂布于含氨苄青霉素、X-gal和IPTG的LB平板上,37 $^{\circ}$ C过夜培养。

1.2.4 LMW-GS基因阳性克隆的筛选及序列测定与分析 挑取白色单菌落,对菌体进行PCR扩增,筛选阳性克隆。PCR反应所用引物、条件同1.2.2。扩繁阳性克隆,用质粒小提试剂盒(天根生化)提取质粒进行测序。测序PCR反应总体积为10 μ L,其中含5 \times Buffer 1.75 μ L, BigDyeV 3.1 0.5 μ L,通用引物M13F或M13R 0.32 μ mol/L,质粒DNA 100~150 ng在96 $^{\circ}$ C预变性5 min。PCR反应条件为96 $^{\circ}$ C变性10 s,50 $^{\circ}$ C退火5 s,60 $^{\circ}$ C延伸3 min,共30个循环。反应结束后,用乙醇/醋酸钠溶液纯化样品,然后用甲酰胺溶解PCR产物,样品变性后在ABI3730XL测序仪上测序。为了保证序列的准确性,每个PCR产物测了5个阳性克隆。序列用DNASTar进行分析。

2 结果与分析

2.1 LMW-GS基因PCR扩增、目的片段的克隆与测序

利用保守引物Primer1/Primer2对硬粒小麦Langdon、LDN1D(1A)、LDN1D(1B)代换系基因组总DNA及BAC质粒DNA进行PCR扩增,用1.5%的琼脂糖凝胶分离扩增产物。在Langdon品种中检测到3条带,其大小分别为0.9、1.0、1.2 kb左右(图1),其中1.2 kb带在LDN1D(1A)代换系中没有扩增,说明该片段来源于1A染色体;0.9 kb的带在LDN1D(1A)、LDN1D(1B)代换系均有扩增,说明A、B基因组均携带有该片段的序列。但由于该引物在BAC B10上仅有1.2 kb和0.9 kb 2条扩增带,而且1.2 kb的带已被定位于1A染色体上,故0.9 kb片段应来源于1A染色体。对1.2 kb和0.9 kb这2条带分别回收与纯化后,将其连接到pCF-T载体上并转化。对白斑进行PCR扩增,确定为阳性克隆后进行测序,获得了2个LMW-GS基因,分别命名为LMW-LDN1和LMW-LDN2,GenBank的注册号为DQ659333和DQ659334。



1. 100 bp DNA ladder (Fermentas); 2. Langdon;
3. LDN1D(1A); 4. LDN1D(1B); 5. BAC B10

图 1 PCR 产物在 1.5% 琼脂糖凝胶上的电泳结果

2.2 硬粒小麦品种 Langdon LMW - GS 基因序列分析

LMW-LDN1 和 LMW-LDN2 的核苷酸序列长度分别为 1 005 bp 和 894 bp, 无内含子, 是典型的 LMW - GS 基因序列。LMW-LDN1 所测的基因序列长度较 PCR 扩增片段少了约 200 bp, 这可

能与克隆过程中重复区的缺失有关^[7]。根据通用三联体密码, 可翻译成含 334 个和 297 个氨基酸残基的蛋白质(图 2), 即先是由 20 个氨基酸残基组成的信号肽, 紧接其后的是由 13 个氨基酸残基组成的 N 端保守区, 其序列为 MENSHPGLERPS。N 末端之后是重复区和 C 末端区。重复区分别由 129 个和 92 个氨基酸残基组成, 富含脯氨酸(P)和谷氨酰胺(Q)。C 末端区由 172 个氨基酸残基组成。LMW-LDN1 的重复区比 LMW-LDN2 多 5 个重复单元(PPFSQQQQ, PVLPQQ, PPFSQQQ, LPPFSQQ, LP-PFSQQQQ)。2 个 LMW - GS 基因含有 8 个半胱氨酸残基, 其中 7 个位于 C-末端区, 1 个位于中间重复区。半胱氨酸残基与已知的 LMW - GS 基因中其位置相对一致。C-末端区的第 36 位氨基酸发生了替代, 苏氨酸(T)被甲硫氨酸(M)替代; C-末端倒数第 5 个氨基酸的变化是由于兼并引物导致。



“—”表示缺失; “.”表示终止密码; “△”表示半胱氨酸

图 2 LMW-LDN1 与 LMW-LDN2 基因推导氨基酸序列的比较

2.3 LMW-LDN1 和 LMW-LDN2 与 Langdon 中已克隆的 LMW - GS 基因的序列比对分析

与 GenBank 中已注册的从 Langdon 品种中克隆的低分子量谷蛋白的氨基酸序列比对分析表明, LMW-LDN1 和 LMW-LDN2 与 Glu - A3 位点的 Td Glu - A3 - 1(来源于 BAC 序列 AY146587)、AJ293097 和 AJ293099 的同源性为 73.1% ~ 77.1%; 与 Glu - B3 位点的 Y14104 的同源性为 80.3% ~ 83.2%。所以 LMW-LDN1 和 LMW-LDN2 是从 Langdon 中克隆的新的 Glu - A3 位点上的低分子量谷蛋白基因。目前已克隆的 Langdon 低分子量谷蛋白基因的聚类如图 3。

2.4 LMW-LDN1 和 LMW-LDN2 与其他已克隆的 LMW - GS 基因的序列比较

根据通用三联体密码将 LMW-LDN1 和 LMW-

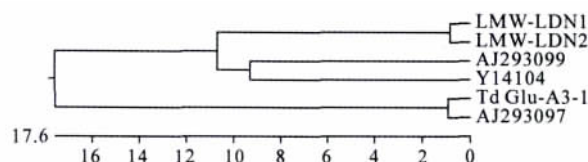


图 3 硬粒小麦 Langdon 品种中已克隆的低分子量谷蛋白基因序列的聚类分析

LDN2 的 DNA 序列翻译为氨基酸序列后, 用 DNAs-tar 软件进行 MegAlign, 结果发现, 这 2 个序列与来源于硬粒小麦 (*T. durum*)、普通小麦 (*T. aestivum*) Glu - B3 基因位点编码的 LMW - GS 相似性最高, 为 89% ~ 96% (表 1), 说明本研究所克隆的 LMW-LDN1 和 LMW-LDN2 是与 Glu - B3 位点基因成员相似的新的 Glu - A3 基因。同时也说明了小麦低分子量谷蛋白基因家族组成的复杂性。

表1 LMW-LDN1 和 LMW-LDN2 基因与其他 LMW-GS 基因氨基酸的相似性比较

来源(品种)	GenBank 登录号	编码位点	LMW-LDN1 氨基酸相似性/%	LMW-LDN2 氨基酸相似性/%	参考文献
<i>T. aestivum</i> (Norin61)	AB062858	未知	96.7	95.3	[8]
<i>T. aestivum</i> (Norin61)	AB062853	<i>Glu-B3</i>	94.9	95.3	[8]
<i>T. aestivum</i> (Norin61)	AB062856	<i>Glu-B3</i>	94.9	95.3	[8]
<i>T. durum</i> (Lira)	Y18159	未知	94.6	95.0	[3]
<i>T. aestivum</i> (Yecora)	Y17845	<i>Glu-B3</i>	93.4	93.6	[7]
<i>T. durum</i> (Line 21)	AJ007746	<i>Glu-B3</i>	89.3	88.9	[3]

3 讨论

根据 LMW-GS 的 N-末端第 1 个氨基酸序列可将其分为 LMW-s 型、LMW-m 型和 LMW-i 型,其起始氨基酸分别为丝氨酸(S)、甲硫氨酸(M)和异亮氨酸(I)。LMW-s 型亚基占主要部分,其 N-末端区氨基酸序列以 SHIPGL 开始。LMW-m 型变异类型多, METSHIPGL、METSRIPL 以及 METSCIPGL 为其主要类型^[9]。LMW-LDN1 和 LMW-LDN2 均属于 LMW-m 型, N-末端区氨基酸序列以 MENSHPGL 开始,这与从普通小麦 Norin61(AB062853、AB062862)和 Yecora Rojo(Y17845)以及硬粒小麦 Lira42(Y18159)中克隆的 LMW-GS 基因的 N-末端序列相同。Langdon 品种中 AJ293097 和 Td Glu-A3-1 为 LMW-i 型, AJ293099 为 LMW-m 型中的 MDTSCIPGL, 由此可见, Langdon 品种中 *Glu-A3* 位点的低分子量谷蛋白基因 N-末端序列类型比较丰富。

小麦低分子量谷蛋白亚基是受多基因控制的。根据 Southern 杂交的结果,估计六倍体小麦基因组中低分子量谷蛋白基因的拷贝数为 10~15 个^[10]至 35~40 个^[11-12]。Johal 等^[13]从粗山羊草(*Aegilops tauschii*) D 基因组 BAC 文库中获得 24 个携带低分子量谷蛋白基因的阳性 BAC 克隆,经 DNA 指纹图谱等分析,将 24 个 BAC 克隆划分为 7 个组,每组取一个代表性克隆进行低分子量谷蛋白基因的测序分析,结果表明,每个 BAC 含有一个低分子量谷蛋白基因,7 个基因中,4 个是有功能的,3 个为假基因,可以推测 D 基因组有 7 个低分子量谷蛋白基因。硬粒小麦品种 Langdon 是四倍体,含 AB 基因组,如果按一个基因组有 7 个低分子量谷蛋白基因计算, Langdon 品种应有 14 个低分子量谷蛋白基因。目前,研究者们已从 *Glu-A3* 位点克隆了 6 个基因(AJ293097、AJ293098、AJ293099、AY146587、DQ659333、DQ659334),其中一个为假基因(AJ293098);从 *Glu-B3* 位点克隆了一个基因(Y14104),虽然尚未获得 *Glu-3* 位点基因家族的全部成员,但这对理解 *Glu-3* 位点的复杂性奠定了很好的基础。

另外,本试验中对含有目的基因的重组质粒进行

测序后发现,基因 LMW-LDN1 比 PCR 扩增结果短了约 200 bp,以前也有过类似的报道^[7],*E. coli* 不同重组克隆重复区内常常会发生 50~200 bp 的缺失,其分子机制尚不十分清楚。为了避免碱基的缺失, Masci 等^[7]认为应采用 PCR 产物直接测序,但效果有待进一步探索。

参考文献:

- [1] Payne P I. Genetics of wheat storage protein and the effect of allelic variation on bread-making quality[J]. Annu Rev Plant Physiol, 1987, 38: 141-153.
- [2] Singh N K, Shepherd K W. Linkage mapping of the genes controlling endosperm proteins in wheat genes on the short arms of group 1 chromosomes[J]. Theor Appl Genet, 1988, 66: 628-641.
- [3] D'Ovidio R, Marchitelli C, Ercoli Cardelli L, et al. Sequence similarity between allelic Glu-B3 genes related to quality properties of durum wheat[J]. Theor Appl Genet, 1999, 98: 455-461.
- [4] Nieto-Taladriz M T, Ruiz M, Martinez M C, et al. Variation and classification of B low-molecular-weight glutenin subunit alleles in durum wheat[J]. Theor Appl Genet, 1997, 95: 1155-1160.
- [5] Sharp P J, Chao S, Desai S, et al. The isolation, characterization and application in Triticeae of a set of wheat RFLP probes identifying each homoeologous chromosome arm[J]. Theor Appl Genet, 1989, 78: 342-348.
- [6] Devos K M, Gale M D. The use of random amplified polymorphic DNA marker in wheat[J]. Theor Appl Genet, 1992, 84: 567-572.
- [7] Masci S, D'Ovidio R, Lafiandra D, et al. Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer[J]. Plant Physiol, 1998, 118: 1147-1158.
- [8] Ikeda T M, Nagamine T, Fukuoka H. Identification of new low-molecular-weight glutenin subunit genes in wheat[J]. Theor Appl Genet, 2002, 104: 680-687.
- [9] D'Ovidio R, Masci S. The low-molecular-weight glutenin subunits of wheat gluten[J]. Journal of Cereal Science, 2004, 39: 321-339.
- [10] Harberd N P, Bartels D, Thompson R D. Analysis of the gliadin multigene loci in bread wheat using nullisomic-tetrasomic lines[J]. Molecular and General Genetics, 1985, 198: 234-242.
- [11] Cassidy B G, Dvorak J, Anderson O D. The wheat low-molecular-weight glutenin genes: Characterization of six new genes and progress in understanding gene family structure[J]. Theor Appl Genet, 1998, 96: 743-750.
- [12] Sabelli P, Shewry P R. Characterization and organization of gene families at the Gli-1 loci of bread and durum wheat[J]. Theor Appl Genet, 1991, 83: 428-434.
- [13] Johal J, Gianibelli M C, Rahman S, et al. Characterization of low-molecular-weight glutenin gene in *Aegilops tauschii*[J]. Theor Appl Genet, 2004, 109: 1028-1040.