

DOI: 10.3724/SP.J.1006.2014.01340

一年生簇毛麦 α -醇溶蛋白基因的分离、原核表达与功能鉴定

杨帆¹ 陈其皎^{1,2} 高翔^{1,2,*} 赵万春^{1,2,*} 强琴琴¹ 吴丹¹ 孟敏¹

¹西北农林科技大学农学院, 陕西杨凌 712100; ²陕西省小麦新品种培育工程研究中心, 陕西杨凌 712100

摘要: 醇溶蛋白是面筋的主要成分之一, 对小麦品质具有重要影响。根据数据库中全长 α -醇溶蛋白基因设计了 1 对通用引物, 从 5 份一年生簇毛麦(*Dasypyrum villosum*)品系中共得到 52 条序列, 长度在 816~873 bp 之间(GenBank 登录号为 KJ004676~KJ004727)。核酸序列分析表明, 其中有 8 条假基因, 有 1 条(KJ004680)缺失终止密码子。推导氨基酸序列显示, KJ004677、KJ004686、KJ004714 和 KJ004696 含有 1 个额外的 Cys, 其中, 前 3 条序列由于 Tyr→Cys 所致, 而 KJ004696 则由于 Ser→Cys 突变。序列间的差异主要出现在 N-端重复区和多聚谷氨酰胺 I 区, 根据 N 端重复区多肽序列的差异将一年生簇毛麦 α -醇溶蛋白分为 5 种类型。为了分析具有额外 Cys 的 α -醇溶蛋白所具有的品质效应, 选取 KJ004708 (具有典型的 6 个 Cys)和 KJ004714 (具有 1 个额外的 Cys)分别构建表达载体, IPTG 诱导后均得到分子量约 30 kD 的蛋白, 与理论值相符; 目的条带经切胶串联质谱鉴定证明, 这 2 个 α -醇溶蛋白基因在大肠杆菌中正确表达。对表达的蛋白亚基进行纯化、复性和低温冷冻干燥, 经 4 g 粉质仪分析表明, KJ004708 和 KJ004714 均能改善面团的加工品质, 其中具有 1 个额外 Cys 的 KJ004714 亚基对面粉品质的改善更为显著。

关键词: 一年生簇毛麦; α -醇溶蛋白; 原核表达; 品质分析; 蛋白质串联质谱鉴定

Cloning, Prokaryotic Expression and *in vitro* Functional Analysis of α -Gliadin Genes from *Dasypyrum villosum*

YANG Fan¹, CHEN Qi-Jiao^{1,2}, GAO Xiang^{1,2,*}, ZHAO Wan-Chun^{1,2,*}, JIANG Qin-Qin¹, WU Dan¹, and MENG Min¹

¹ College of Agronomy, Northwest A&F University, Yangling 712100, China; ² New Varieties Cultivation of Wheat Engineering Research Centre of Shaanxi Province, Yangling 712100, China

Abstract: Gliadin, which has a great effect on wheat quality, is one of main components in gluten. According to the full lengths of α -gliadin genes deposited in NCBI database, a conserved primer pair was designed to clone α -gliadin genes in five *Dasypyrum villosum* lines. A total 52 sequences (816 to 873 bp in length) were isolated (GenBank accession numbers: KJ004676 to KJ004727) including eight pseudogenes and another sequence KJ004680 without stop codon. Deduced amino acid sequence analysis showed that KJ004677, KJ004686, and KJ004714 contain an extra Cys from the Tyr → Cys mutation, whereas, the extra Cys in KJ004696 resulted from the Ser → Cys mutation. Amino acid variation mainly occurred in N-terminal repetitive region and polyglutamine domain I. Variation in N-terminal repetitive region formed five groups in the 43 α -gliadins. To study the effects of an extra Cys on dough quality, we constructed the prokaryotic expression vectors for KJ004708 (with the typical six Cys residues) and KJ004714 (with an extra Cys) and obtained proteins of ~30 kD from *Escherichia coli* BL21(DE3) under the induction of isopropyl- β -D-thiogalactoside (IPTG) with the predicted molecular weight. These expressed proteins were verified by matrix-assisted laser desorption-ionization time-of-flight MALDI-TOF/TOF tandem mass spectrometry analysis. The result showed that these α -gliadins were expressed correctly in *E. coli*. After purification, renaturation, and freeze-drying process, the functions of the expressed proteins were tested with 4 g Farinograph. Both KJ004708 and KJ004714 had positive effects on flour quality, especially KJ004714 with an extra Cys.

本研究由“十二五”农村领域国家科技计划项目(2011AA100501), 国家现代农业产业技术体系建设专项(CARS-3-2-47), 国家自然科学基金项目(31171538, 30900896), 中央高校基本科研业务费专项资金项目(QN2009007)和西北农林科技大学唐仲英育种基金项目(A212020912)资助。

* 通讯作者(Corresponding authors): 高翔, E-mail: gx@nwsuaf.edu.cn, Tel: 13709124775; 赵万春, E-mail: zhaowc2009@hotmail.com, Tel: 13110439969

第一作者联系方式: E-mail: wenzhifan2011@126.com, Tel: 15829719716

Received(收稿日期): 2014-01-10; Accepted(接受日期): 2014-06-04; Published online(网络出版日期): 2014-06-13.

URL: <http://www.cnki.net/kcms/detail/11.1809.S.20140613.1004.001.html>

Keywords: *Dasypyrum villosum*; Alfa-gliadin; Prokaryotic expression; Functional analysis; MALDI-TOF/TOF tandem mass spectrometer

小麦是我国主要粮食作物之一, 品质改良是小麦的重要育种目标。但是在普通小麦中发现的能够编码优质蛋白的基因并不多, 并且由于现代农业导致小麦品种越来越单一化, 使得种内遗传变异逐渐丧失, 造成新选育的小麦品种遗传基础越来越狭窄, 限制了小麦品质的进一步改良。因此, 通过在小麦近缘种属中寻找新的优质基因资源来拓宽小麦的遗传基础成为小麦品质育种的一条有效途径。

一年生二倍体簇毛麦 [*Dasypyrum villosum*, $2n=14$, VV (syn. *Haynaldia villosa*)], 原产于地中海北部和高加索地区, 是小麦遗传改良的重要种质资源^[1-3]。一年生簇毛麦具有许多重要的小麦病害抗性基因资源, 如抗黑穗病(*Ustilago nuda* 和 *U. triticii*)^[4]、白粉病(*Erysiphe graminis*)^[5-6]、纹枯病(*Pseudocercospora herpotricoides*)^[7]、锈病(*Puccinia graminis* 和 *P. recondite*)^[8]和根腐病(*Rhizoctonia solani*)^[9]等多种病害, 同时也具有耐盐^[10]等抗逆特点。一年生簇毛麦的籽粒蛋白含量高, 在 1V、6VS 和 7V 染色体上, 已经发现含有编码种子储藏蛋白的基因位点 *Glu-V1*^[11]、*Glu-V3*^[12]、*Gli-V2*^[13]和 *Wsp-1*^[14]。此外, 一年生簇毛麦中还携带矮秆^[15]、产量^[16-17]和面筋强度^[18]的相关基因。

小麦胚乳储藏蛋白是决定小麦品质, 尤其是小麦加工品质的主要因素。它主要由麦谷蛋白(glutenins)和麦醇溶蛋白(gliadins)组成, 为面筋的主要成分, 是决定面团黏弹性的主要因素^[19]。麦醇溶蛋白约占小麦胚乳贮藏蛋白总量的 50%~60%^[20-21], 其中 α -醇溶蛋白占总醇溶蛋白的 25%^[22]。它包含信号肽(signal peptide)、N-端重复区(N-terminal repetitive region)、多聚谷氨酰胺 I 区(polyglutamine repeats I)和 N-端非重复区(unique non-repetitive domain I), 最后是多聚谷氨酰胺 II 区(polyglutamine repeats II)和 C-端非重复区(unique non-repetitive domain II)^[23]。目前, 已经从小麦及其近缘物种中分离了许多 α -醇溶蛋白基因, 但国内外尚未报道关于一年生簇毛麦中 α -醇溶蛋白基因对小麦品质贡献的研究。本研究利用简并引物分离一年生簇毛麦中的 α -醇溶蛋白基因, 对克隆得到的部分 α -醇溶蛋白基因进行原核表达、亲和层析及体外功能鉴定, 为研究一年生簇毛麦中 α -醇溶蛋白亚基对小麦品质改良的作用提供基础。

1 材料与方法

1.1 簇毛麦 α -醇溶蛋白基因的克隆

用微量 CTAB 法^[24]提取一年生簇毛麦 TA10220、TA10224、TA10227、TA10232 和 TA10249 (由美国堪萨斯州大学小麦遗传与基因组资源中心 Bikram S. Gill 博士实验室提供)的基因组 DNA。利用 Primer Premier 5, 根据 GenBank (<http://www.ncbi.nlm.nih.gov/>)中已注册的 α -醇溶蛋白基因序列的保守区域设计扩增基因全长编码区的引物(al-GliF: 5'-ATG AAG ACC TTC CTC ATC TTT GTC-3'; al-GliR: 5'-CCA TGT TTG AAC TAG TAT AGG TCG G-3')。PCR 体系为 $5 \times$ PrimeSTAR buffer (Mg^{2+} plus) 5 μ L, 2.5 mmol L⁻¹ dNTP Mixture 2 μ L, 0.1 mmol L⁻¹ 上下游引物各 1 μ L, 模板 DNA 100 ng, 2.5 U μ L⁻¹ PrimeSTAR HS DNA Polymerase (TaKaRa) 0.2 μ L, 加 ddH₂O 至反应体积为 25 μ L。PCR 程序为 94℃ 5 min; 94℃ 45 s、50℃ 50 s、72℃ 90 s, 30 个循环; 72℃ 8 min。PCR 扩增产物经 1.5% 琼脂糖凝胶电泳检测后, 切胶回收目的片段。用 pEASY-T1 载体连接目的片段并转化大肠杆菌 *Trans1-T1* (北京全式金生物技术有限公司), 在选择性培养基上过夜培养, 划线后从每个克隆挑选 3 个独立阳性菌落送至南京金斯瑞生物科技有限公司测序。

1.2 生物信息学分析

用 DNAMAN (6.0)推导分析 α -醇溶蛋白的序列; 采用 NCBI 在线 BLAST (<http://blast.ncbi.nlm.nih.gov/>)进行序列的同源性比对; SignalP 4.1 Server 在线完成氨基酸序列信号肽分析; 蛋白质二级结构的预测采用 PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>); 以 MEGA 5.0 构建系统演化树; 参照 van Herpen 等^[25]和李光蓉等^[26]报道的方法分析识别 α -醇溶蛋白序列的毒性多肽位点。

1.3 簇毛麦 α -醇溶蛋白原核表达和品质鉴定

设计 1 对表达引物 Ex.al-gliF (5'-GTA CTT AGA GTT CCA GTG CCG C-3')/Ex.al-gliR (5'-TTA GTT ACT ACC GGT GCT ACC AAA T-3'), 分别对含有 2 条代表性序列的阳性质粒进行 PCR 扩增(KJ004708 和 KJ004714), 回收目的片段并分别连接至表达载体 pEASY-E1, 转化表达菌株 BL21 (DE3)(北京全式金生物技术有限公司), 挑取阳性菌落接种至 LB 中(氨苄青霉素, 100 mg mL⁻¹), 37℃ 过夜培养后按照体

积比 1 : 100 将培养物接种到新鲜 LB 中培养至 $A_{600} = 0.5 \sim 0.8$, 取 1 mL 过夜培养的菌液作为阴性对照, 向剩余的菌液中加入诱导剂 IPTG (终浓度 1 mmol L^{-1}) 诱导 6 h (37°C , 200 转 min^{-1})。9300 $\times g$ 离心 10 min 收集菌体, 加入 $1 \times$ SDS 凝胶加样缓冲液重悬, 煮沸 7 min, 以 12% SDS-PAGE 电泳检测蛋白的表达, 并用消毒手术刀在 PAGE 胶上切下与预测分子量相符的正确表达的蛋白, 采用 MALDI-TOF/TOF 4800 (Applied Biosystem) 质谱仪进行串联质谱鉴定(中国科学院上海生命科学研究院蛋白质组研究分析中心), 采用一级肽指纹质量与二级肽碎片质量综合分析法 Combined (MS+MS/MS) 分析样品序列, 通过 NCBI (<http://www.ncbi.nlm.nih.gov/>) 序列同源性比对确定蛋白质的性质。

离心收集诱导后的菌体(4°C , $6800 \times g$, 10 min), 加入破碎缓冲液(50 mmol L^{-1} Tris-HCl, 100 mmol L^{-1} NaCl, 2 mmol L^{-1} EDTA, 1 mg L^{-1} 溶菌酶, pH 8.5) 在冰上放置 45 min, 超声破碎细胞并离心收集沉淀(4°C , $15000 \times g$, 10 min)。沉淀经 6 mol L^{-1} 盐酸胍的磷酸盐缓冲液预处理后, 采用 ProteinIso Ni-NTA Resin (北京全式金生物技术有限公司) 纯化和收集目的蛋白。将收集的蛋白透析 72 h, 低温冷冻干燥 48 h 后在 -20°C 冰箱中保存备用。

采用 4 g 微量粉质仪测定一年生簇毛麦 α -醇溶蛋白亚基的品质效应。在 4 g 基础面粉中加入 50 mg 纯化蛋白和适量水, 混合 4 min 后加入 0.25 mL DTT 溶液(50 mg L^{-1}) 反应 1 min, 再加入 0.25 mL KIO_3 溶液(200 mg L^{-1}) 反应 10 min, 记录粉质曲线和相应品质参数。

2 结果与分析

2.1 α -醇溶蛋白基因的克隆及相似性比对

通过分析已公布的 α -醇溶蛋白基因, 利用序列的保守区设计 1 对简并引物, 以一年生簇毛麦基因组 DNA 为模板扩增得到 52 条序列(GenBank 登录号为 KJ004676~KJ004727)。NCBI 在线 BLAST 表明, 得到的序列与已公布的多年生簇毛麦 α -醇溶蛋白基因序列相似度较高, 最高可达 94%。克隆得到的序列全长介于 816~873 bp 之间(图 1), 最常见的基因片段长度为 858 bp, 共出现 19 次。

DNAMAN 开放阅读框分析表明, 8 条序列(KJ004720~KJ004727)含有提前终止密码子, 属于假基因。其中 KJ004725 在 642 bp 处(特征区 II)缺失碱

基 C 导致阅读框移码, 其他假基因均因碱基替换所致。碱基替换在 N 端重复区出现的频率最高(3/7)。在 7 个碱基替换的类型中, 有 6 个均因密码子的第 1 位碱基 C 被 T 替换, 分别导致 $\text{CAA} \rightarrow \text{TAA}$ (3/6) 和 $\text{CAG} \rightarrow \text{TAG}$ (3/6), 而 KJ004726 是因密码子的后两位碱基同时被碱基 A 替换导致 $\text{TGG} \rightarrow \text{TAA}$ 。KJ004680 则因 3' 端倒数第 8 个密码子的第 3 位缺失碱基 T 而移码突变, 使得终止密码子缺失。

SNP 分析表明, 碱基转换占 73%, A/G 和 T/C 类型分别占 39% 和 34%; 碱基颠换只占 27%, 并以 A/T 颠换为主(58%), 其他几种碱基颠换的比例较低。

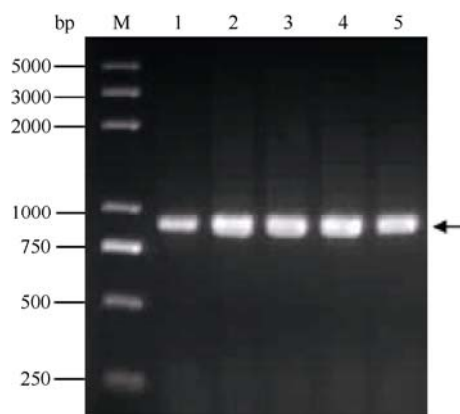


图 1 1.5% 琼脂糖凝胶扩增 α -醇溶蛋白基因序列
Fig. 1 Amplification products of α -gliadin gene sequences separated on 1.5% agarose gel
M: Trans 2K Plus DNA marker; 1: TA10220; 2: TA10224; 3: TA10227; 4: TA10232; 5: TA10249.

2.2 氨基酸序列分析

分析推导的氨基酸序列, 得到 43 条完整的 α -醇溶蛋白序列, 编码 271~290 氨基酸残基, 均具有典型的 α -醇溶蛋白结构, 分别是由 19 个氨基酸残基组成的信号肽, N-端重复区, 多聚谷氨酰胺 I 区, 特征区 I, 多聚谷氨酰胺 II 区和特征区 II。根据 N-端重复区多肽序列的差异, 将这 43 条 α -醇溶蛋白分为 5 种类型(图 2-A)。序列比对分析发现, 氨基酸变异位点大部分存在于 N-端、特征区 I 和特征区 II, 导致 Gln 和 Arg、Leu 和 Pro, 以及 Gln 和 Leu 之间的相互转换。N-端重复区和多聚谷氨酰胺 I 区多肽长度的变异主要是由于第 V 类亚基中的 4 个 α -醇溶蛋白(KJ004681、KJ004693、KJ004709 和 KJ004711)缺失了富含 Gln 的片段。本试验获得的一年生簇毛麦 α -醇溶蛋白多聚谷氨酰胺 I 区的长度介于 11~23 个氨基酸残基之间, 其中 Gln 的平均含量为 18.1 个。有 33 条序列发生了 Q \rightarrow E 突变, 其中 KJ004682 还发生了 Q \rightarrow R 突变。多聚谷氨酰胺 II 区序列都非常短, 只

含有 5 个氨基酸残基(QLLQQ)且非常保守(图 2-B), 只有 KJ004717 在这个区域的第 1 位氨基酸残基由谷氨酰胺 Q 变异为亮氨酸(LQLQQ)。

通过比较不同物种间多聚谷氨酰胺 I 和 II 区中谷氨酰胺残基数量的变异, 发现不同物种间多聚谷氨酰胺区中 Gln 的数量变异比较大。除拟斯卑尔脱山羊草和稗旱麦草外, 其他 6 个物种多聚谷氨酰胺 I 区中

Gln 的数量要多于多聚谷氨酰胺 II 区。与其他物种相比, 本研究从一年生簇毛麦中获得的 α -醇溶蛋白在多聚谷氨酰胺 II 区中 Gln 数量的变异较小(图 2-C)。

α -醇溶蛋白中 Cys 的位置和数目都非常保守, 分别在特征区 I 和特征区 II 中含有 4 个和 2 个 Cys (图 2-B)。在所得到的序列中, KJ004677、KJ004686、KJ004696 和 KJ004714 含有 7 个 Cys。除 KJ004696

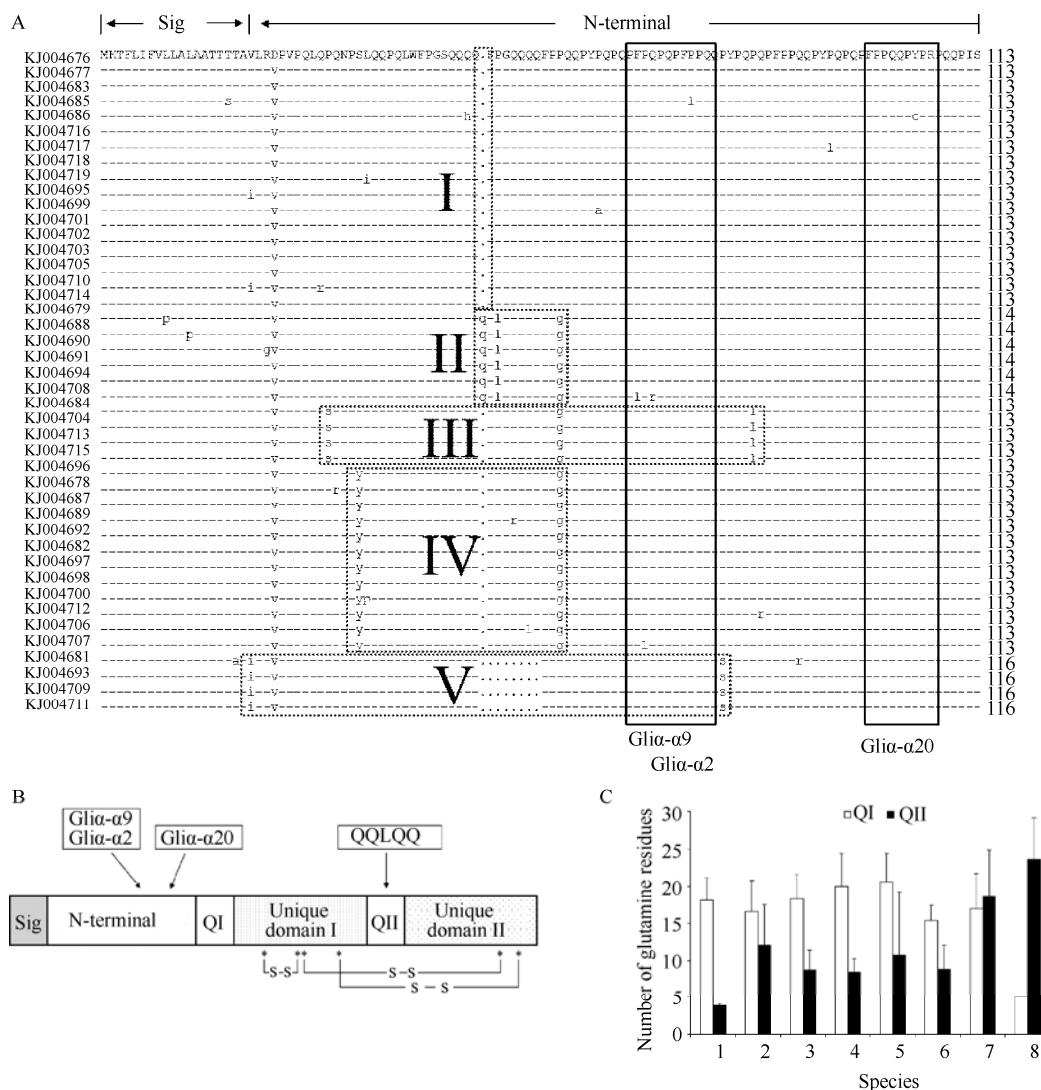


图 2 α -醇溶蛋白基因推导氨基酸序列分析

Fig. 2 Deduced amino acid sequence analysis of α -gliadin genes

A: 43 条推导氨基酸序列的毒性多肽位点和分类, 粗线框示毒性多肽位点, 虚线框示不同类型的 α -醇溶蛋白。

B: α -醇溶蛋白的结构及二硫键模式, N 端重复区(N-terminal)存在毒性多肽位点, 多聚谷氨酰胺 II 区(Q II)存在 QQLQQ 多肽序列, 星号(*)示半胱氨酸残基的位置。C: 不同物种 α -醇溶蛋白在多聚谷氨酰胺 I 区和 II 区中谷氨酰胺残基数量的变异, 误差线表示标准差。

1: 一年生簇毛麦, $n = 43$; 2: 普通小麦, $n = 26$; 3: 粗山羊草, $n = 16$; 4: 一粒小麦, $n = 15$; 5: 二粒小麦, $n = 17$; 6: 黑麦, $n = 17$;

7: 拟斯卑尔脱山羊草, $n = 7$; 8: 稗旱麦草, $n = 7$ 。

A: T-cell stimulatory epitopes and classification of 43 α -gliadin amino acid sequences, thick frames indicate T-cell stimulatory epitopes, dot line frames indicate different types of α -gliadins. B: Schematic structure of α -gliadin and intra-molecular disulfide bonds, T-cell stimulatory epitopes existed in the N-terminal, QQLQQ existed in the Q II, * indicates the positions of cysteine residue. C: Variations in mean number of glutamine residues in the polyglutamine domains (Q I and Q II) of α -gliadins among different species. Error bars show the standard deviations, 1: *Dasyphyrum villosum*, $n = 43$; 2: *Triticum aestivum*, $n = 26$; 3: *Aegilops tauschii*, $n = 16$; 4: *Triticum monococcum*, $n = 15$; 5: *Triticum dicoccoides*, $n = 17$; 6: *Secale cereale*, $n = 17$; 7: *Aegilops speltoides*, $n = 7$; 8: *Eremopyrum bonaepartis*, $n = 7$.

的额外 Cys 是由 Ser 变异而来, 其他 3 个序列的额外 Cys 均由 Tyr 变异而来(Tyr→Cys)。进一步分析额外 Cys 出现的位置, 发现其出现在 KJ004686 的 N-端重复区、KJ004696 的特征区 I、KJ004677 和 KJ004714 的特征区 II。

2.3 α-醇溶蛋白中毒性多肽位点的分析

α-醇溶蛋白中主要含有 4 个能诱发乳糜泻病的多肽因子, 分别是 Gliα-α (QGSFQPSQQ)、Gliα-α2 (PQPQLPYPQ)、Gliα-α9 (PFPQPQLPY)和 Gliα-α20 (FRPPQQPYPQ) [25-26], 人体 T 细胞与这些特定多肽结合后会引引起小肠黏膜损伤和吸收不良[27]。不同的诱发因子有特定的保守分布区域, Gliα-α 出现在特

征区 II, 而 Gliα-α2、Gliα-α9 和 Gliα-α20 均分布在 N-端重复区, 且 Gliα-α2 和 Gliα-α9 多肽序列有部分重叠(图 2-A)。将 43 条一年生簇毛麦 α-醇溶蛋白多肽序列与这 4 个毒性多肽位点对比分析表明, 所有序列中都不含有完整的 Gliα-α2、Gliα-α9 和 Gliα-α20 毒性多肽位点序列, 并缺失 Gliα-α 位点(表 1)。

比较发现, 不同物种的 α-醇溶蛋白毒性多肽位点分布差异非常大(表 2)。其中冰草属和本研究中的 5 份一年生簇毛麦品系均不含 4 个毒性多肽位点, 旱麦草属只含有 Gliα-α 位点, 黑麦属缺失 Gliα-α20 位点, 而偃麦草属、山羊草属和小麦属则含有全部 4 个毒性多肽位点(表 2)。

表 1 不同研究获得的一年生簇毛麦 α-醇溶蛋白毒性多肽位的比较

Table 1 Comparison of T cell stimulatory epitopes of α-gliadins in *Dasypyrum villosum* obtained from different studies

毒性多肽位	毒性多肽位序列	KJ004676 to KJ004719	EU325688 to EU325697 [28]
T-cell stimulatory epitopes	Sequences for T-cell stimulatory epitopes	(except KJ004680)	
Gliα-α	QGSFQPSQQ	—	<i>QGSFQPSQQ</i>
Gliα-α2	PQPQLPYPQ	<i>PQPQPFPPQ</i>	<i>LPQLTYPQ</i>
Gliα-α9	PFPQPQLPY	<i>PFPQPQPFP</i>	<i>PFLPQLTY</i>
Gliα-α20	FRPPQQPYPQ	<i>FPPQQPYPR</i>	—

粗斜体表示保守氨基酸残基, —表示缺少相应的毒性多肽位点。

Bold italic and dash indicate the conserved amino acids and the lack of T cell stimulatory epitopes, respectively.

表 2 不同物种 α-醇溶蛋白中毒性多肽位点数量的比较

Table 2 Comparison of the quantity of T cell stimulatory epitopes in α-gliadins among different species

物种	α-醇溶蛋白数	Gliα-α	Gliα-α2	Gliα-α9	Gliα-α20
Species	No. of α-gliadins				
一年生簇毛麦 <i>Dasypyrum villosum</i>	43	0	0	0	0
冰草属 <i>Agropyron</i>	15	0	0	0	0
旱麦草属 <i>Eremopyrum</i>	8	0.875	0	0	0
黑麦属 <i>Secale</i>	26	0.615	0.615	0.615	0
偃麦草属 <i>Thinopyrum</i>	120	0.419	0.015	0.015	0.015
山羊草属 <i>Aegilops</i>	130	0.759	0.616	0.536	0.812
小麦属 <i>Triticum</i>	350	0.420	0.363	0.534	0.537

冰草属、偃麦草属、山羊草属和小麦属数据来自 Chen et al. [29]; 旱麦草属和黑麦属来自 NCBI 数据库(<http://www.ncbi.nlm.nih.gov/>)。

Data of *Agropyron*, *Thinopyrum*, *Aegilops*, and *Triticum* genus originated from Chen et al. [29]; data of *Eremopyrum* and *Secale* genus were collected from NCBI (<http://www.ncbi.nlm.nih.gov/>).

2.4 蛋白质二级结构预测和分析

一年生簇毛麦 α-醇溶蛋白的二级结构主要由 α-helix 和 coil 组成。α-helix 出现的位置和数目比较保守, 在特征区 II 中除 KJ004696 含有 2 个 α-helices 外, 其他均只含 1 个 α-helix, 且在 N-端重复区不含 α-helix。成熟的一年生簇毛麦 α-醇溶蛋白中 α-helix 通常出现 5 次, 但是 KJ004677、KJ004679、KJ004682、KJ004685、KJ004694、KJ004696、KJ004700、

KJ004704、KJ004713、KJ004715 和 KJ004716 中有 6 个 α-helices, 另外 KJ004683 和 KJ004702 中只含 4 个 α-helices。在这 43 条 α-醇溶蛋白中, 除 KJ004685、KJ004691、KJ004697、KJ004707 和 KJ004713 在特征区 II 出现一次 β-strand (分别占序列的 1.1%、0.8%、1.1%、0.7%和 0.8%)外, 其他序列均不含 β-strand。

2.5 系统进化树分析

将获得的 43 条一年生簇毛麦 α-醇溶蛋白序列

表 3 原核表达的 α-醇溶蛋白串联质谱鉴定
Table 3 Identification of prokaryotically expressed α-gliadins using MALDI-TOF/TOF tandem mass spectrometer

项目 Item	KJ004708 匹配肽段 KJ004708 matched peptides (MPS)			KJ004714 匹配肽段 KJ004714 matched peptides (MPS)		
	MPS-1	MPS-2	MPS-3	MPS-4	MPS-5	MPS-6
理论分子量 Theoretical MW (Da)	858.4792	1516.8595	2751.2615	858.4792	1502.8439	2737.2458
观测分子量 Observed MW (Da)	858.4792	1516.785	2751.0918	858.4754	1502.7788	2737.1099
误差 Difference (Da)	0	-0.0745	-0.1697	-0.0038	-0.0651	-0.1359
蛋白质得分 Protein score		198			229	
蛋白得分置信度 C.I. %		100			100	
等电点 Isoelectric point		8.28			8.23	
与其他 α-醇溶蛋白序列相似度 Similarity to other α-gliadins (%)						
推导氨基酸序列 Deduced sequence	100	100	100	100	100	100
原核表达产物 Prokaryotic expression	100	100	100	100	100	100
<i>Australopyrum retrofractum</i> (ABW79851)	86	62	95	86	69	100
<i>Dasyphyrum hordeaceum</i> (ACY71755)	71	100	95	71	92	100
<i>Dasyphyrum villosus</i> (ABY48872)	71	69	86	71	62	90

匹配肽段序列分别为 SQVLQQR (MPS-1 和 MPS-4); DVLLQQHN~~I~~APIR (MPS-2)和 DVLLQQHN~~V~~APIR (MPS-5); TYQELQQCCQQLWQIPEQSR (MPS-3)和 SYQELQQCCQQLWQIPEQSR (MPS-6)。下划线表示相同多肽片段的不相同氨基酸位点。
Matched peptide sequences are SQVLQQR (MPS-1 and MPS-4); DVLLQQHN~~I~~APIR (MPS-2) and DVLLQQHN~~V~~APIR (MPS-5); and TYQELQQCCQQLWQIPEQSR (MPS-3) and SYQELQQCCQQLWQIPEQSR (MPS-6). Amino acid site variations are underlined.

表 4 添加 KJ004708 和 KJ004714 对小麦面粉质特性的影响
Table 4 Effects of KJ004708 and KJ004714 on Farinograph parameters

添加物 Additive	断裂时间 BdT (min)	形成时间 DvT (min)	稳定时间 ST (min)	带宽 WC (FU)	弱化度 DS (FU)	公差指数 MTI (FU)	粉质质量指数 FQN (mm)
DTT, KIO ₃ (control)	2.1 b	1.4 b	1.0 b	50 b	159.9 b	155 b	17.5 b
DTT, KIO ₃ , KJ004708	2.2 ab	1.6 a	1.1 b	50 b	169.9 a	170 a	20.2 a
DTT, KIO ₃ , KJ004714	2.6 a	1.6 a	1.7 a	55 a	154.9 c	150 c	21.5 a

表中数据为 3 次重复取平均值，同列中不同字母表示差异显著($P < 0.05$)。
Data are the means of three replicates. Values followed by different letters within columns are significantly different at $P < 0.05$. BdT: breakdown time; DvT: development time; ST: stability time; WC: width of curve; DS: softening of dough; MTI: mixing tolerance index; FQN: farinograph quality number.

3 讨论

一年生簇毛麦中不仅含有许多重要的抗逆位点，而且还具有能增加籽粒蛋白含量和增强面筋强度等优质性状，是小麦品种改良的重要遗传资源。目前已经培育了许多不同的小麦-簇毛麦代换系、易位系等材料，有效地拓宽了小麦的遗传资源并加快了小麦育种进度。目前已经从不同的物种中分离得到了许多 α-醇溶蛋白基因，但是假基因比例都比较高，如 van Herpen 等^[25]从小麦不同的染色体祖先中获得的 230 条 α-醇溶蛋白基因中假基因的比率就高达 87%，其他的一些研究者获得的 α-醇溶蛋白基因也存在比较高的假基因频率^[30-32]。据推测，α-醇溶蛋白

基因家族中假基因出现的频率高达 50%^[33]，而本研究从一年生簇毛麦中共获得了 52 条序列，其中仅有 8 条假基因 (15.4%)，不仅远低于前人从不同物种中得到的假基因频率，也低于 Li 等^[28]从簇毛麦中得到的假基因比例 (31.3%)，推测获得的假基因频率与不同来源的材料有密切关系。醇溶蛋白高比例假基因的发生可能是染色体非均等交换和基因大量复制过程中的突变所致^[21,34]。突变产生的类型中，碱基 C 被 T 替换是最常见的类型之一，一般认为，由于醇溶蛋白中含有大量 Gln，而编码 Gln 的密码子 CAA 和 CAG 中的第 1 位 C 一旦被 T 替换就会产生翻译的终止，产生假基因^[33]，本试验中的 6 个假基因 (75%) 属于此类。van Herpen 等^[25]曾报道此类假基因产生比

例为77.2%,而张晓霞等^[31]研究发现假基因中有95%是C-T突变类型。本研究中C-T突变频率占总突变频率的34%,仅次于A-G突变(39%),而最终假基因频率仅有15.4%,并且主要由于C-T突变造成,说明这5份一年生簇毛麦品系中较低的假基因频率并不是由于C-T突变频率较低引起的。本研究中还发现一个密码子的后两位碱基被同时替换(TGG→TAA)而产生终止密码子的例子。

α -醇溶蛋白序列结构具很高保守性,典型特征是含5个不同的区域(图2-B)。根据N端重复区多肽序列的不同,将本试验获得的 α -醇溶蛋白分为5种类型(图2-A)。研究一般认为多聚谷氨酰胺II区是 α -醇溶蛋白氨基酸序列变异比较大的一个区域^[26,31,35],而本研究得到的序列在此区域不仅非常保守而且仅含4个Gln(图2-B),Li等^[28]从一年生和多年生簇毛麦中获得的 α -醇溶蛋白在此区域分别平均含14.3个和18.6个Gln,而李光蓉等^[26]从多年生簇毛麦中得到的 α -醇溶蛋白在多聚谷氨酰胺II区也平均含9.1个Gln。贮藏蛋白亚基中Cys的含量和位置会对亚基的品质效应产生较大影响。 α -醇溶蛋白中Cys的数目和相对位置非常保守,一般含6个Cys。其中有4个位于特征区I,另外2个位于特征区II,并且第1和第2位,第3和第5位,第4和第6位的Cys可以形成3对分子内二硫键^[36]。本研究共得到4个含额外Cys的序列,额外的Cys可能参与分子间二硫键的形成,对稳定蛋白空间网络结构和改善面筋品质具有积极作用。在这4条序列中只有KJ004696是由于Ser突变为Cys,其余3条均由于Tyr突变为Cys,其中KJ004686在N-端重复区倒数第9位密码子后2位碱基同时发生转换突变(TAC→TGT),导致Tyr→Cys突变。其他的一些研究也发现在 α -醇溶蛋白中存在一个额外的Cys,但这些额外的Cys多出现在特征区II并且是由丝氨酸(Ser)变异产生的^[22,34,37-38]。Li等^[28]从一年生簇毛麦中分离了一个在特征区I含一个额外Cys的 α -醇溶蛋白,但这个Cys仍由Ser变异而来。

聚类分析发现,本研究得到的一年生簇毛麦 α -醇溶蛋白被单独聚为一类,同其他物种的亲缘关系较远。Li等^[28]通过分离一年生和多年生簇毛麦中的 α -醇溶蛋白并构建系统进化树,发现簇毛麦属和小麦属间的亲缘关系较远,同时一年生和多年生簇毛麦 α -醇溶蛋白也分别被单独聚为一类,与本研究得出的结果相似,推测簇毛麦中 α -醇溶蛋白是一个新

的 α -醇溶蛋白家族。同时本研究获得的 α -醇溶蛋白与其他一年生簇毛麦来源的 α -醇溶蛋白^[28]亲缘关系也较远,说明不同来源的一年生簇毛麦中 α -醇溶蛋白变异非常丰富。 α -醇溶蛋白中对乳糜泻病人具有毒性的多肽主要有4种,分别是Glia- α 、Glia- α 2、Glia- α 9和Glia- α 20。序列比较结果表明,本研究获得的一年生簇毛麦 α -醇溶蛋白均不含Glia- α 毒性多肽位点,且由于氨基酸残基的变异导致不含有完整的另外3个毒性多肽位点。Li等^[28]从一年生簇毛麦中得到的 α -醇溶蛋白含有完整的Glia- α 位点,但是缺失Glia- α 20毒性位点序列,另外也不含有完整的Glia- α 2和Glia- α 9位点(表1)。说明不同来源的一年生簇毛麦中 α -醇溶蛋白毒性多肽位点序列变异很大,与聚类分析的结果相符,这有助于选育不含有毒性多肽位点的新品种。

微量掺粉试验结果表明,一年生簇毛麦中两种不同类型的 α -醇溶蛋白亚基对小麦面团品质都能产生显著影响。面团的稳定时间、形成时间和断裂时间越长,带宽值愈大,同时弱化度和公差指数值越小,说明面团的面筋强度和弹性越好,而粉质质量指数则是用于综合评价面粉筋力和耐揉性一个重要依据^[34,39-40]。粉质结果显示,虽然KJ004708显著增加了弱化度和公差指数,降低了面团的面筋强度,但是对面团的形成时间和最终的粉质质量指数有显著的提升作用(表4)。含有一个额外Cys的KJ004714亚基显著作用于所有的7个主要粉质参数,添加KJ004714能够显著增加面团的稳定时间、形成时间、断裂时间、带宽和粉质质量指数,同时又显著降低面团的弱化度和公差指数,对面团面筋的强度和弹性都有显著改善(表4)。推测KJ004714中含有的额外Cys参与分子间二硫键的形成,促进蛋白质亚基间的聚合,从而显著地改善面团的结构,推测KJ004714对小麦面粉品质的改良具有很大潜力。李敏等^[34]对来源于普通小麦的一个含7个Cys的 α -醇溶蛋白品质鉴定表明,加入外源蛋白亚基后面团的断裂时间、稳定时间和粉质质量指数值显著降低,而弱化度和公差指数值却显著升高,说明此亚基对面团品质有不利影响。并且本研究和李敏等^[34]得到的 α -醇溶蛋白亚基中额外Cys均出现在特征区II,但最终对于品质的影响相差非常大,说明额外Cys对小麦品质的影响可能非常复杂,不仅与Cys的数目和位置有关, α -醇溶蛋白本身的序列也可能影响亚基的品质效应。

4 结论

利用特异引物, 从 5 份一年生簇毛麦材料中分离得到 52 条序列, 其中 8 条为假基因, 1 条由于碱基的缺失导致移码突变而缺失终止密码子。得到的 43 条完整 α -醇溶蛋白序列可被划分为 5 种不同的类型, 序列间的差异主要存在于 N 端重复区和多聚谷氨酰胺 I 区。毒性多肽位点的比较、聚类分析和质谱鉴定结果的序列比对分析发现不同来源的簇毛麦 α -醇溶蛋白序列间差异较大, 这可能对小麦品质育种产生有利的影响。本研究获得的 α -醇溶蛋白既缺失 Gli α - α 毒性多肽位点, 也不含完整的其他 3 个毒性多肽位点。KJ004708 与 KJ004714 均能改善小麦面团的品质, 说明这些一年生簇毛麦品系是小麦品质育种的优良材料, 并且 KJ004714 对品质的改善更为显著, 其所含额外 Cys 有助于小麦面筋网络结构的形成。

References

- [1] Qualset C O, Zhong G Y, De Pace C, McGuire P E. Population biology and evaluation of genetic resources of *Dasypyrum villosum*. In: Damania A B ed. Biodiversity and wheat improvement. Chichester: John Wiley & Sons, 1993. pp 227–233
- [2] De Pace C, Qualset C O. Mating system and genetic differentiation in *Dasypyrum villosum* (Poaceae) in Italy. *Plant Syst Evol*, 1995, 197: 123–147
- [3] Frederiksen S. Taxonomic studies in *Dasypyrum* (Poaceae). *Nord J Bot*, 1991, 11: 135–142
- [4] Nielsen J. Host range of the smut species *Ustilago nuda* and *Ustilago tritici* in the tribe Triticeae. *Can J Bot*, 1978, 56: 901–915
- [5] Chen X, Shi A N, Shang L M, Leath S, Murphy J P. The resistance of *H. villosa* to powdery mildew isolates and its expression in wheat background. *Acta Phytopathol Sin*, 1997, 27: 17–22
- [6] Minelli S, Ceccarelli M, Mariani M, De Pace C, Cionini P G. Cytogenetics of *Triticum* · *Dasypyrum* hybrids and derived lines. *Cytogenet Genome Res*, 2005, 109: 385–392
- [7] Yildirim A, Jones S S, Murray T D. Mapping a gene conferring resistance to *Pseudocercospora herpotrichoides* on chromosome 4V of *Dasypyrum villosum* in a wheat background. *Genome*, 1998, 41: 1–6
- [8] Jan C C, De Pace C, McGuire P E, Qualset C O. Hybrids and amphiploids of *Triticum aestivum* L. and *T. turgidum* L. with *Dasypyrum villosum* (L.) Candargy. *Z Pflanzenzücht*, 1986, 96: 97–106
- [9] Smith J G, Kidwell K K, Evans M A, Cook R J, Smiley R W. Evaluation of spring sereal grains and wild *Triticum* germplasm for resistance to *Rhizoctonia solani* AG-8. *Crop Sci*, 2003, 43: 701–709
- [10] Zhong G Y, Dvorák J. Evidence for common genetic mechanisms controlling the tolerance of sudden salt stress in the tribe Triticeae. *Plant Breed*, 1995, 114: 297–302
- [11] Montebove L, De Pace C, Jan C C, Scarascia-Mugnozza G T, Qualset C O. Chromosomal location of isozyme and seed storage protein genes in *Dasypyrum villosum* (L.) Candargy. *Theor Appl Genet*, 1987, 73: 836–845
- [12] Shewry P R, Parmar S, Pappin D J C. Characterization and genetic control of the prolamins of *Haynaldia villosa*: relationship to cultivated species of the Triticeae (rye, wheat and barley). *Biochem Genet*, 1987, 25: 309–325
- [13] Blanco A, Resta P, Simeone R, Parmar S, Shewry P R, Sabelli P, Lafiandra D. Chromosomal location of seed storage protein genes in the genome of *Dasypyrum villosum* (L.) Candargy. *Theor Appl Genet*, 1991, 82: 358–362
- [14] Liu C J, Chao S, Gale M D. *Wsp-1*, a set of genes controlling water-soluble proteins in wheat and related species. *Genet Res*, 1989, 54: 173–181
- [15] Li J M, Yang Z M, Tian H Q, Huang F, Gang P T. Somatic cell clone establishment and amphiploid synthesis in a *Triticum aestivum* · *Haynaldia villosa* intergeneric hybrid. *Hereditas* (Beijing), 1991, 13: 1–3
- [16] Mohammad P, Hossain M A, Khodarker N A, Shiraishi M. Study for morphological characteristics of species alien to wheat in Bangladesh. *Sarhad J Agric*, 1997, 13: 541–550
- [17] Okocha P I. Peculiarities of nucleo-cytoplasmic interactions in alloctoplasmic forms of wheat. *Global J Pure Appl Sci*, 1999, 5: 431–435
- [18] De Pace C, Snidaro D, Ciaffi M, Vittori D, Ciofo A, Cenci A, Tanzarella O A, Qualset C O, Scarascia Mugnozza G T. Introgression of *Dasypyrum villosum* chromatin into common wheat improves grain protein quality. *Euphytica*, 2001, 117: 67–75
- [19] 谷淑波, 于振文, 王东, 张永丽. 小麦贮藏蛋白对加工品质的影响及对环境反应. 山东农业大学学报(自然科学版), 2009, 40: 309–312
- Gu S B, Yu Z W, Wang D, Zhang Y L. Effects of wheat storage protein on processing quality and reacting to environment. *J Shandong Agric Univ* (Nat Sci Edn), 2009, 40: 309–312 (in Chinese)
- [20] Payne P I, Holt L M, Jackson E A, Law C N, Damania A B. Wheat storage proteins: their genetics and their potential for manipulation by plant breeding. *Philos Trans R Soc Lond*, 1984, 304: 359–379
- [21] Shewry P R, Halford N G. Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot*, 2002, 53: 947–958
- [22] 朱西平, 李鑫, 李雅轩, 晏月明. 普通小麦及近缘粗山羊草 α -醇溶蛋白基因的克隆、定位与进化分析. 作物学报, 2010, 36: 580–589
- Zhu X P, Li X, Li Y X, Yan Y M. Cloning, chromosomal location, and evolutionary analysis of α -gliadin genes from *Aegilops tauschii* and common wheat (*Triticum aestivum* L.). *Acta Agron Sin*, 2010, 36: 580–589 (in Chinese with English abstract)
- [23] Qi P F, Wei Y M, Yue Y W, Yan Z H, Zheng Y L. Biochemical and molecular characterization of gliadins. *Mol Biol*, 2006, 140: 713–723
- [24] Murray H G, Thompson W F. Rapid isolation of high molecular weight DNA. *Nucl Acids Res*, 1980, 8: 4321–4325
- [25] van Herpen T W, Goryunova S V, van der Schoot J, Mitreva M,

- Salentijn E, Vorst O, Schenk M F, van Veelen P A, Koning F, van Soest L J, Vosman B, Bosch D, Hamer R J, Gilissen L J, Smulders M J. Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes. *BMC Genomics*, 2006, 7: 1
- [26] 李光蓉, 任正隆, 刘成, 周建平, 杨足君. 多年生簇毛麦 α -醇溶蛋白基因的分离与序列分析. *作物学报*, 2008, 34: 1097–1103
- Li G R, Ren Z L, Liu C, Zhou J P, Yang Z J. Isolation and sequence analysis of α -gliadin genes from *Dasyphyrum brevistaratum*. *Acta Agron Sin*, 2008, 34: 1097–1103 (in Chinese with English abstract)
- [27] Molberg O, Uhlen A K, Jensen T, Flaete N S, Fleckenstein B, Arentz-Hansen H, Raki M, Lundin K E, Sollid L M. Mapping of gluten T cell epitopes in the bread wheat ancestors: implications for celiac disease. *Gastroenterology*, 2005, 128: 393–401
- [28] Li G R, Liu C, Zeng Z X, Jia J Q, Zhang T, Zhou J P, Ren Z L, Yang Z J. Identification of α -gliadin genes in *Dasyphyrum* in relation to evolution and breeding. *Euphytica*, 2009, 165: 155–163
- [29] Chen G X, Lv D W, Li W D, Subburaj S, Yu Z T, Wang Y J, Li X H, Wang K, Ye X G, Ma W, Yan Y M. The α -gliadin genes from *Brachypodium distachyon* L. provide evidence for a significant gap in the current genome assembly. *Funct Integr Genomics*, 2014, 14: 149–160
- [30] Li G R, Zhang T, Ban Y R, Yang Z J. Molecular characterization and evolutionary analysis of α -gliadin genes from *Eremopyrum bonapartii* (*Triticeae*). *J Agric Sci*, 2010, 2: 30–36
- [31] 张晓霞, 焦焱, 董振营, 李世明, 王燃, 凌宏清, 秦广雍, 王道文. 普通小麦品种小偃 54 中 α/β -醇溶蛋白编码基因的克隆与序列分析. *作物学报*, 2011, 37: 1497–1502
- Zhang X X, Jiao Z, Dong Z Y, Li S M, Wang R, Ling H Q, Qin G Y, Wang D W. Cloning and sequence analysis of α/β -gliadin genes from common wheat variety Xiaoyan 54. *Acta Agron Sin*, 2011, 37: 1497–1502 (in Chinese with English abstract)
- [32] 李玉阁, 邢冉冉, 李锁平. 栽培一粒小麦 α -醇溶蛋白新基因的克隆与序列分析. *麦类作物学报*, 2012, 32: 387–392
- Li Y G, Xing R R, Li S P. Cloning and sequence analysis of new α -gliadin genes from *Triticum monococcum*. *J Triticeae Crops*, 2012, 32: 387–392 (in Chinese with English abstract)
- [33] Anderson O D, Litts J C, Greene F C. The α -gliadin gene family: I. Characterization of ten new wheat α -gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and southern analysis of the gene family. *Theor Appl Genet*, 1997, 95: 50–58
- [34] 李敏, 高翔, 陈其皎, 董剑, 赵万春, 王明霞. 普通小麦中 α -醇溶蛋白基因(GQ891685)的克隆、表达及品质效应鉴定. *中国农业科学*, 2010, 43: 4765–4774
- Li M, Gao X, Chen Q J, Dong J, Zhao W C, Wang M X. Cloning, prokaryotic expression and *in vitro* functional analysis of α -gliadin gene from common wheat. *Sci Agric Sin*, 2010, 43: 4765–4774 (in Chinese with English abstract)
- [35] 李光蓉, 郎涛, 刘成, 周建平, 任正隆, 杨足君. 小麦新品种“成电麦 1 号” α -醇溶蛋白基因的分离与序列分析. *中国农学通报*, 2011, 27(1): 203–208
- Li G R, Lang T, Liu C, Zhou J P, Ren Z L, Yang Z J. Isolation and sequence analysis of α -gliadin genes from wheat cultivar Chengdianmai 1. *Chin Agric Sci Bull*, 2011, 27(1): 203–208 (in Chinese with English abstract)
- [36] Anderson O D, Greene F C. The α -gliadin gene family: II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes. *Theor Appl Genet*, 1997, 95: 59–65
- [37] 刘千, 龙海, 魏育明, 颜泽洪, 郑有良. 小麦品种‘川农 16’ α -醇溶蛋白基因序列分析. *中国农业科学*, 2008, 41: 2168–2173
- Liu Q, Long H, Wei Y M, Yan Z H, Zheng Y L. Sequence analysis of α -gliadin genes from wheat variety Chuannong 16. *Sci Agric Sin*, 2008, 41: 2168–2173 (in Chinese with English abstract)
- [38] Xie Z, Wang C, Wang K, Wang S, Li X, Zhang Z, Ma W, Yan Y. Molecular characterization of the celiac disease epitope domains in α -gliadin genes in *Aegilops tauschii* and hexaploid wheats (*Triticum aestivum* L.). *Theor Appl Genet*, 2010, 121: 1239–1251
- [39] 田纪春. 谷物品质测试理论与方法. 北京: 科学出版社, 2006. pp 338–340
- Tian J C. Theory and Method of Test in Grain Quality. Beijing: Science Press, 2006. pp 338–340 (in Chinese)
- [40] 姜薇莉, 孙辉, 凌家煜. 粉质质量指数(FQN)对于评价小麦粉品质的实用价值研究. *中国粮油学报*, 2004, 19(2): 42–48
- Jiang W L, Sun H, Ling J Y. Applicability of FQN in evaluation of wheat flour quality. *J Chin Cereals Oils Assoc*, 2004, 19(2): 42–48 (in Chinese with English abstract)