二穗短柄草 b 型 *Farinin* 基因的克隆与进化分析

SUBBURAJ Saminathan，朱耿锐，李晓辉，胡英考，晏月明
（首都师范大学遗传与生物工程重点实验室，北京 100048）

摘要：Farinin 是小麦和山羊草属等物种中最古老的储藏蛋白家族成员之一，对面团特性具有重要影响。本研究对二穗短柄草（*Brachypodium distachyon* L.）种子 Farinin 蛋白基因进行了分子遗传学分析。采用等位基因特异 PCR（allele-specific PCR，AS-PCR）方法从 14 个不同二穗短柄草材料中分离、克隆了 51 个 b 型 *Farinin* 基因。序列比对显示，克隆的基因与先前报道的小麦和山羊草属物种中的燕麦类似蛋白基因具有高度的相似性。氨基酸序列表明，小麦中 b 型 Farinin 分为 b1、b2、b3 和 b4 四种类型，而在二穗短柄草中只含有 b3 和 b4 两种类型。通过系统进化分析，发现短柄草 Farinin 与普通小麦中成员的关系更近，也进一步证实了 Farinin 家族存在 Group 1（b1、b2 和 b3）和 Group 2（b4）两大亚家族，而 b4 型的 Farinin 是这个家族中较为古老的成员。

关键词：植物学；二穗短柄草；Farinin；分子克隆；进化
中图分类号：Q37　　**文献标识码**：A　　**文章编号**：1674-2850(2016)09-0887-06

Molecular characterization and evolutionary analysis of b-type *Farinin* genes in *Brachypodium distachyon* L.

SUBBURAJ Saminathan, ZHU Gengrui, LI Xiaohui, HU Yingkao, YAN Yueming
(*Key Laboratory of Genetics and Biotechnology*, *Capital Normal University*, *Beijing* 100048, *China*)

Abstract: Farinin is one of the oldest members of the gluten family in wheat and *Aegilops species*, and related to dough properties. Here, we performed the detailed molecular genetic study on *Farinin* genes in *Brachypodium distachyon* L. A total of 51 b-type *Farinin* genes were cloned and characterized from 14 different *B*. *distachyon* L. accessions by allele-specific PCR (AS-PCR). All genes were highly similar to those previously reported from wheat and *Aegilops* species. The identification of deduced amino acid sequences showed that b-type Farinins across Triticeae genomes could be classified as b1-, b2-, b3-, and b4-type Farinin, while *B*. *distachyon* L. had only b3- and b4-type Farinin. Phylogenetic analysis suggested that *Brachypodium* Farinin may have closer relationships with common wheat and further confirmed four different types of b-type Farinin present in Triticeae and *Brachypodium* genomes, corresponding to b1, b2, b3 (Group 1), and b4 (Group 2). The b4-type could be considered older in the Farinin family.
Key words: botany; *Brachypodium distachyon* L.; Farinin; molecular cloning; evolution

0 引言

醇溶谷蛋白是谷物籽粒中最主要的储藏蛋白，在小麦中又分为醇溶蛋白（prolamins）和麦谷蛋白（glutenins），是面筋的主要成分，是小麦加工品质的主要决定因素[1]。除了这些谷蛋白外，还有其他类

基金项目：高等学校博士学科点专项科研基金（20111108110003）
作者简介：SUBBURAJ Saminathan（1987—），男，博士研究生，主要研究方向：植物分子遗传
通信联系人：晏月明，教授，主要研究方向：植物分子遗传. E-mail: yanym@cnu.edu.cn

型的贮藏蛋白，如大麦醇溶蛋白、黑麦醇溶蛋白以及燕麦蛋白等，它们共同的结构特征是氨基酸序列含有较高的脯氨酸和谷氨酰胺，因而被定义为谷蛋白范畴。KASARDA 等[2]曾鉴定出了一种新型的谷蛋白，它与燕麦蛋白有相似的结构因而命名为燕麦类似蛋白（avenin-likeprotein，ALP）。ALP 又进一步分为 a 型和 b 型蛋白[2]。其中，b 型蛋白通常含有 18 或 19 个半胱氨酸残基，它们影响谷蛋白的功能特性。当 b 型 ALP 融入谷蛋白中时，可以明显提升面团的黏性。而另一种新鉴定的蛋白——Farinin 与 ALP 有很高的结构相似性，它最初是由 KAN 等[3]在小麦品种 Bute 86 中鉴定出来的。“Farinin”术语适用于小麦及近缘物种中发现的 ALP，被认为是小麦及山羊草属中最古老的储藏蛋白家族成员之一，对面团特性具有重要影响。

二穗短柄草（*Brachypodium distachyon* L.）作为一种新型的小麦属模式植物，它的基因组已经测序完成[4~5]。对其蛋白质组学研究表明，它包含多种球蛋白（globulins）、醇溶谷蛋白（prolamins）以及其他高低分子量类似蛋白[6~7]。但是有关 *Farinin* 基因及其编码蛋白的信息却非常有限。本研究以二穗短柄草作为研究材料，克隆了 b3 和 b4 型的 *Farinin* 基因，并解析其编码氨基酸的结构特征，通过系统进化树分析 b1、b2、b3 和 b4 型 *Farinin* 基因的进化关系。

1 材料与方法

1.1 实验材料

本研究共使用 14 个短柄草材料：包括 7 个二倍体：Bd2（PI185133）、Bd3（PI185134）、Bd21（PI254867）、Bd21-3、Bd12、Bd14（PI239713）和 BdABR1；2 个四倍体品系：Bd347、Bd16（PI239715）；5 个六倍体：Bd4（PI208216）、Bd10（PI226452）、Bd13（PI233228）、Bd17（PI239716）和 BdISKP1. 小麦品种中国春（CS；*Triticum aestivum* L.，2*n*=6×7=42，AABBDD）作为对照。

1.2 DNA 提取与 PCR 扩增

大约取两周大的短柄草幼苗 0.5 g，使用 CTAB 法提取基因组 DNA. 首先，参考序列，利用生物学软件 Primer 5.0 设计两对特异的 AS-PCR 引物（PF1：5'-ATGAAGGTCTTCATCCTGGC-3'，PR1：5'-CTAG CACGCACCACCAGTGTA-3'和 PF2：5'-ATGAAGACAATGTTCATCCTCG-3'，PR2：5'-CTAGCAGGTA CCATCGGTATAGC-3'）用于扩增短柄草 *Farinin* 基因开放阅读框（open reading frame，ORF）。PCR 反应程序：首先 94℃预变性 5 min；然后 94℃变性 45 s，60℃退火 80 s，72℃延伸 2 min，这一过程进行 34 个循环；最后，72℃补平 10 min. PCR 反应在常规 PCR 仪中进行。

1.3 系统进化树构建

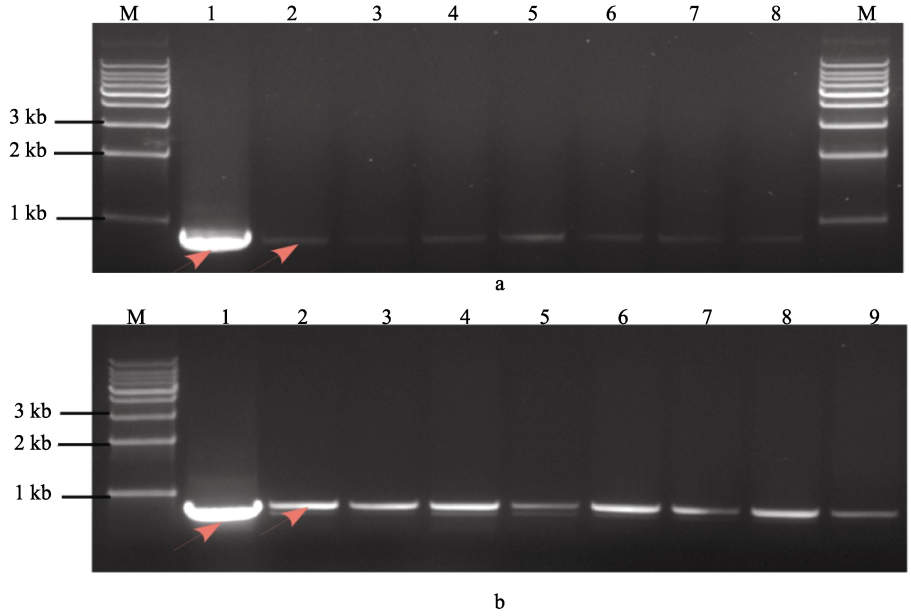
使用短柄草 Farinin 氨基酸序列及其他物种成员，如小麦、山羊草等，构建了一个 NJ（Neighbour-Joining）树。构建结果在 PhyML 3.0 平台（http://www.phylogeny.fr）上分析。具体参数包括：序列使用 ClusterW v2.0.3，比对后除去 gaps，JTT 替换模式，重复 1 000 次。

2 结果与分析

2.1 短柄草 b 型 *Farinin* 基因的鉴定

尽管有关短柄草基因组 *Farinin* 基因及其编码蛋白的信息还不清楚，但是课题组以前的工作克隆并分析了短柄草中沉默的醇溶谷蛋白基因[7~9]。据此，总共设计了两对引物（PF1+PR1；PF2+PR2）用于克隆短柄草中 b 型的 *Farinin* 基因。前者用于扩增 b1、b2 和 b3 型的 *Farinin* 基因[2, 10]，而后者用于克隆新型

的 *Farinin* 基因[3]。总之，这两对引物都是用于扩增短柄草 b 型 *Farinin* 基因的 ORFs，实验材料包括短柄草二倍体、四倍体和六倍体品系，另外中国春用作对照。如图 1 所示，两对引物扩增产物的片段长度大约在 900 bp，这与 b 型 *Farinin* 基因相一致。其中，b3 型的基因大约在 855 bp，而新型的 b 型基因在 843 bp 左右。研究总共克隆得到了 28 个基因和 23 个新型的 *Farinin* 基因，但是没有鉴定到 b1 和 b2 型基因。



注：1～9 分别代表中国春、短柄草品系 Bd21、Bd21-3、Bd2、Bd3、Bd4、Bd10、Bd347 和 Bd14；M 为 1 kb 的 DNA 梯度电泳图谱；目标条带用箭头指出

图 1 短柄草 b 型 *Farinin* 基因的鉴定

Fig. 1 Identification of b-type *Farinin* genes isolated from *Brachypodium*

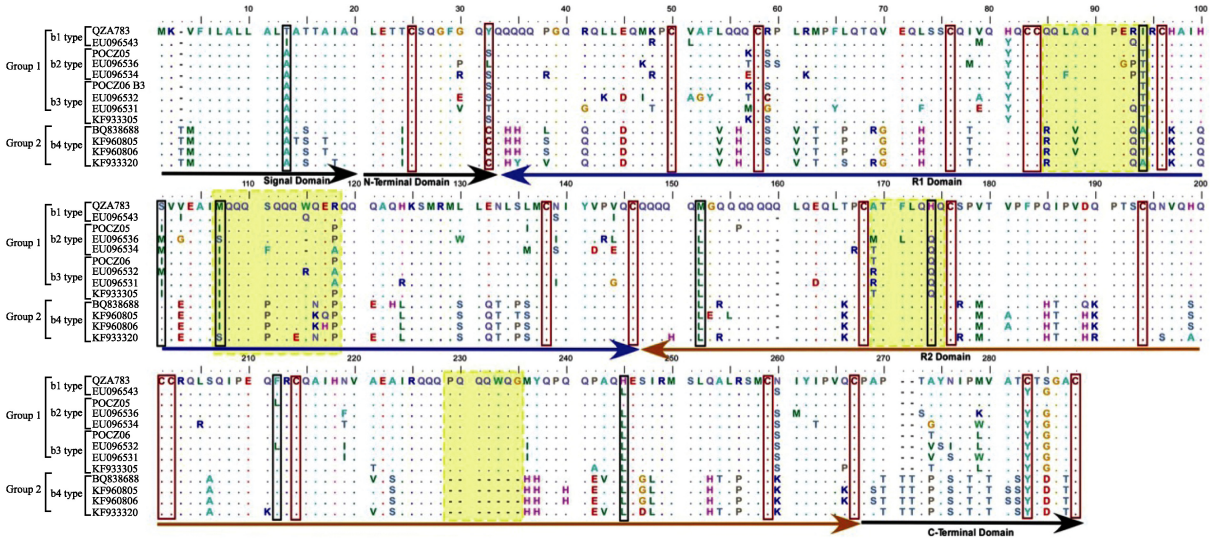
a—b3 型 *Farinin* 基因的完整 DNA ORFs；b—b4 型 *Farinin* 基因的完整 DNA ORFs

a-Complete DNA ORFs of b3-type *Farinin* genes; b-Complete DNA ORFs of b4-type *Farinin* genes

2.2 b 型 Farinin 蛋白的结构特征

所有来自短柄草品系的 *Farinin* 基因推测的氨基酸序列与小麦和山羊草属中推测的氨基酸序列有较高的相似性。因此，选择典型的 b3 型（KF933305）和新型的 b 型（KF933320）Farinin 氨基酸序列与 11 个已经鉴定到的 Farinin：小麦 Cadenza 品种 QZA783、POCZ05 和 POCZ06[2]，山羊草属中 EU096531、EU096532、EU09634、EU09636、EU09643、KF960805 和 KF960806[10]以及一个来自小麦品种 Bute 86 中新的 b 型 Farinin[3]作比较分析。如图 2 所示，这些 Farinin 都含有一个相同的结构组成：保守的信号肽、N 端区域、R1 和 R2 的重复区以及 C 端区域。除了 KF933320 在 N 端 38 号位置多一个半胱氨酸残基（Cys）外，其他 18 个 Cys 在这两个品系中的位置是保守的。氨基酸序列表明小麦中 b 型 Farinin 分为 b1、b2、b3 和 b4 四种类型，而在二穗短柄草中只含有 b3 和 b4 两种类型。依据这些序列结构，绘制了这 4 种 b 型 Farinin 蛋白的结构图（如图 3 所示）。由图 3 可知，与 b4 型的 Farinin 蛋白结构相比，b1、b2 和 b3 型的蛋白结构比较保守且序列相似度较高。不同类型的 Farinin 蛋白序列的结构差异主要在单核苷酸多态性（single nucleotide polymorphism, SNP）变异。在信号肽区，b1 型 Farinin 在 13 号位置含有一个保守的苏氨酸（threonine，T）或异亮氨酸（isoleucine，I），而 b2 和 b3 型蛋白在这个位置为丙氨酸（alanine，A）。在 N 端区域，这四种类型蛋白都分别含有几段保守序列，如 b1 型的 GFG/RQY，

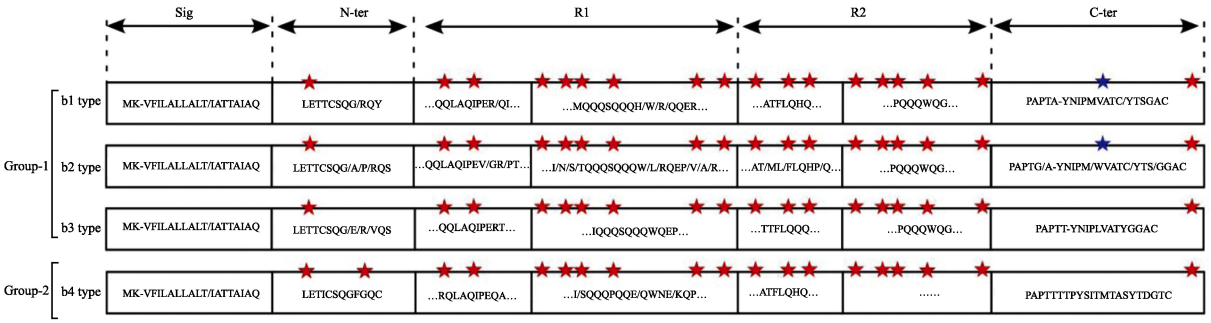
b2 型的 GFG/A/P/RQS. 而在 R1 重复区，IQQQSQQQWQE 在 b3 型的保守性要高一些，而其他类型在这部分有较大差异。相比而言，b1 和 b4 型的 Farinin 序列在 169～175 号位置有较保守的序列（ATFLQHQ），而另一个非常保守的六肽（PQQQWQG）却在 b4 型中表现缺失现象。在 C 端区，b4 型 Farinin 序列与其他三种的差异大些。



注：各蛋白区域（信号肽、N 端区、重复区：R1 和 R2，以及 C 端区）分别用黑色、蓝色和橘黄色箭头表示；保守的半胱氨酸残基和差异较大的残基分别用红色和黑色框表示；差异的 b 型 Farinin 结构用黄色阴影表示

图 2　13 个 b 型 *Farinin* 基因推测的氨基酸比较分析

Fig. 2　Deduced amino acid sequence comparison of 13 b-type *Farinin* genes



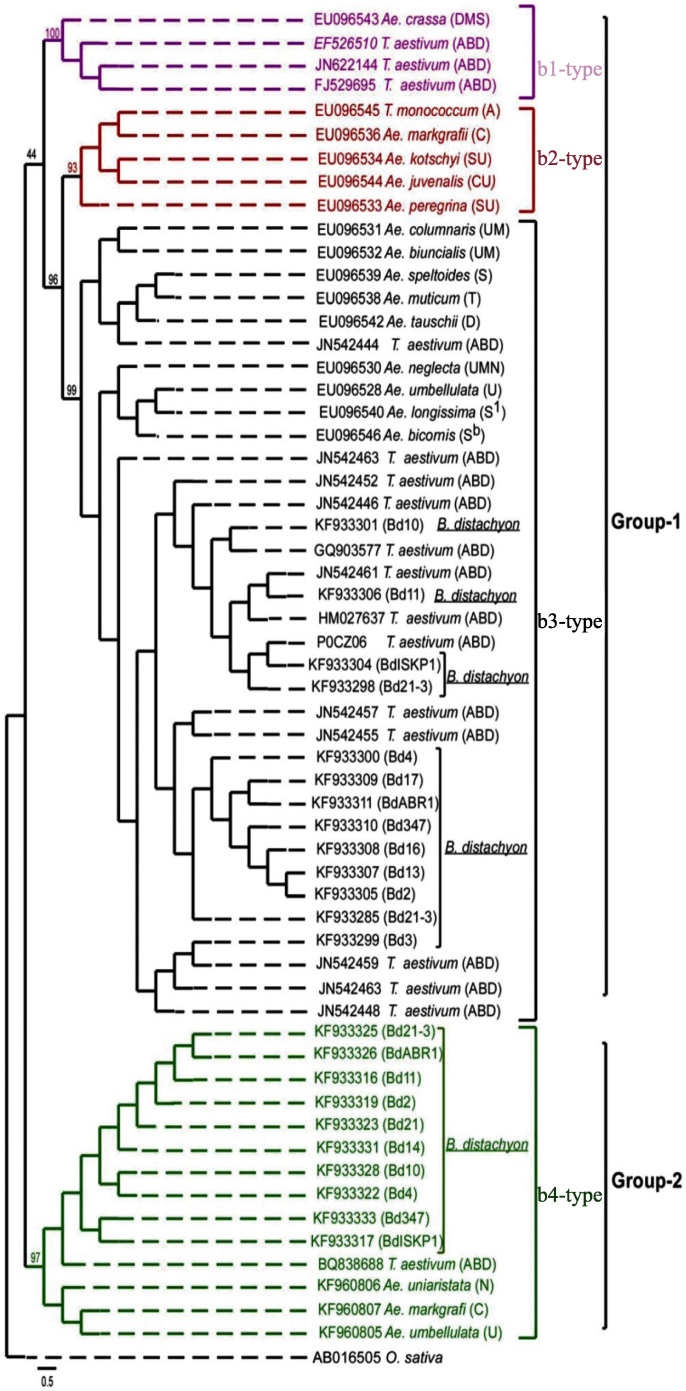
注：Sig 表示信号肽；N-ter 表示 N 端区域；R1、R2 表示重复区；C-ter 表示 C 端区域。明显的差异残基展现在各类型区域内。把手的半胱氨酸和额外的半胱氨酸残基分别用红色和蓝色星号表示。每个区域内的虚线表示 b 型的 Farinin 比较后删除保守序列部分

图 3　b 型 Farinin 通常的结构特征

Fig. 3　General structural features of b-type Farinin

2.3　系统进化树分析

通过构建系统进化树分析 b 型 Farinin 的起源及进化特征（如图 4 所示）。在这个进化树中，总共有 58 个成员，包括短柄草（23）、小麦（18）及山羊草（17），还有一个水稻成员作为外类群。与图 2 和图 3 展示的结果相同，所有的 Farinin 成员都分为两组：一组为 b1、b2、b3 型 Farinin（44）；另一组为 b4 型的成员（14）。由图 3 可知，小麦及山羊草属中的 Farinin 成员分布在所有的四个亚家族中，而短柄草中的成员却只分布在亚家族 3（b3）和亚家族 4（b4）组内。很明显的是，短柄草 Farinin 成员与来自小麦属中的成员亲缘关系更近。



注：最大似然进化树是基于氨基酸序列使用 PhyML 3.0 建成。在这个进化树中，典型的 b 型 Farinin 分为 b1、b2、b3 和 b4 型四种 Farinin，划分为两组（Group 1 和 Group 2）

图 4 短柄草以及小麦和山羊草属中 b 型 Farinin 进化树

Fig. 4 Phylogenetic tree of b-type Farinin from *Brachypodium* along with wheat and *Aegilops* species

3 讨论

Farinin 基因被认为是小麦及其近缘物种中多基因家族成员，缺少大量的重复序列[2, 10]。目前，有关短柄草 *Farinin* 基因序列的差异鲜有报道。在本研究中，从 14 个短柄草品系中鉴定了 51 个 b 型的 *Farinin* 基因，预测的氨基酸序列表明 Farinin 序列的差异主要由 SNP 引起。自从鉴定到小麦及其近缘种中含有

广泛的等位变异，它们的分类与进化分析就变得十分重要。目前的研究就是参照 *Farinin* 基因的序列变异特征进行了一个新的分类系统。Farinin 的结构分析表明，短柄草 Farinin 并不含 b1 和 b2 型的 Farinin，只含有较为保守的 b3 和 b4 型 Farinin，其与小麦属中的同源性要高于山羊草属的 Farinin.

KASARDA 等[2]曾经指出燕麦蛋白及类似蛋白形成的亚家族与醇溶蛋白和低分子量谷蛋白的关系更为紧密。这有可能是因为 *Farinin* 基因在进化的过程中，一些基因如醇溶蛋白在复制时发生分化导致[11]。随着时间的分化，*Farinin* 基因家族逐渐形成自己的进化体系。研究表明，*Farinin* 是个多基因家族，主要包括 b 型家族成员（b1～b4）。依据 Farinin 序列结构特征发现，b1、b2 和 b3 的保守性要明显高于 b4，说明 b4 型的 Farinin 是较早分化出去的 b 型成员；而从系统进化树上看，b4 型的 Farinin 作为一个整体被划分为单独的分支，可能是 Farinin 家族的祖先。结构保守性分析与系统进化分析预测的结果是相符的。

4 结论

给出了二穗短柄草 b 型 Farinin 的结构特征并分析了四种 b 型 *Farinin* 基因的系统进化关系。分子特征表明，短柄草中 b 型 Farinin 与小麦和山羊草属中的 Farinin 结构类似、保守性强。但是短柄草中仅含有 b3 和 b4 型两种 Farinin，且与小麦及其近缘物种的关系密切。从系统进化树可以得出，短柄草在进化过程中失去了 b1 和 b2 型的 *Farinin* 基因，仅保留 b3 和 b4 型的 *Farinin* 基因，并且 b4 型的 *Farinin* 是这个家族较为古老的成员。研究结果为短柄草及其近缘物种 *Farinin* 基因家族的进化和起源提供了非常有用的信息。

[参考文献] (References)

[1] DELCOUR J A, JOYE I J, PAREYT B, et al. Wheat gluten functionality as a quality determinant in cereal-based food products[J]. Annual Review of Food Science and Technology, 2012, 3(1): 469-492.

[2] KASARDA D D, ADALSTEINS E, LEW E L J, et al. Farinin: characterization of a novel wheat endosperm protein belonging to the prolamin superfamily[J]. Journal of Agricultural and Food Chemistry, 2013, 61(10): 2407-2417.

[3] KAN Y C, WAN Y F, BEAUDOIN F, et al. Transcriptome analysis reveals differentially expressed storage protein transcripts in seeds of *Aegilops* and wheat[J]. Journal of Cereal Science, 2006, 44(1): 75-85.

[4] International *Brachypodium* Initiative. Genome sequencing and analysis of the model grass *Brachypodium* distachyon[J]. Nature, 2010, 463(7282): 763-768.

[5] BRENCHLEY R, SPANNAGL M, PFEIFER M, et al. Analysis of the allohexaploid bread wheat genome (*Triticum aestivum*) using comparative whole genome shotgun sequencing[J]. Nature, 2012, 491(7426): 705-710.

[6] LARRE C, PENNINCK S, BOUCHET B, et al. *Brachypodium distachyon* grain: identification and subcellular localization of storage proteins[J]. Journal of Experimental Botany, 2010, 61(6): 1771-1783.

[7] WANG S, WANG K, CHEN G, et al. Molecular characterization of LMW-GS genes in *Brachypodium distachyon* L. reveals highly conserved Glu-3 loci in *Triticum* and related species[J]. BMC Plant Biology, 2012, 12(1): 8-12.

[8] SUBBURAJ S, CHEN G, HAN C, et al. Molecular characterisation and evolution of HMW glutenin subunit genes in *Brachypodium distachyon* L.[J]. Journal of Applied Genetics, 2014, 55(1): 27-42.

[9] CHEN G X, LÜ D W, LI W D, et al. The α-gliadin genes from *Brachypodium distachyon* L. provide evidence for a significant gap in the current genome assembly[J]. Functional & Integrative Genomics, 2014, 14(1): 149-160.

[10] CHEN P, WANG C, LI K, et al. Cloning, expression and characterization of novel avenin-like genes in wheat and related species[J]. Journal of Cereal Science, 2008, 48(3): 734-740.

[11] GAO S, GU Y Q, WU J, et al. Rapid evolution and complex structural organization in genomic regions harboring multiple prolamin genes in the polyploid wheat genome[J]. Plant Molecular Biology, 2007, 65(1): 189-203.