

李国锋, 葛 敏, 吕远大. Opaque2 转录因子对玉米  $\alpha$ -醇溶蛋白基因家族成员表达的影响[J]. 江苏农业学报, 2015, 31(6): 1224-1231.

doi: 10.3969/j.issn.1000-4440.2015.06.006

## Opaque2 转录因子对玉米 $\alpha$ -醇溶蛋白基因家族成员表达的影响

李国锋<sup>1</sup>, 葛 敏<sup>2</sup>, 吕远大<sup>2</sup>

(1. 江苏省农业科学院六合基地, 江苏 南京 210014; 2. 江苏省农业科学院农业生物技术研究所/江苏省农业生物学重点实验室, 江苏 南京 210014)

**摘要:** 醇溶蛋白是玉米主要的储存蛋白, 其中  $\alpha$ -醇溶蛋白含量最为丰富, 对玉米籽粒品质有重要影响。本研究利用全基因组数据鉴定  $\alpha$ -醇溶蛋白基因家族, 并从基因保守 motif、玉米籽粒不同发育时期基因的表达谱以及转录因子 Opaque2 对该家族成员表达的影响 3 个方面分析玉米  $\alpha$ -醇溶蛋白基因家族。结果显示, 玉米参考基因组中存在 39 个  $\alpha$ -醇溶蛋白基因, 其中 *z1A*、*z1B*、*z1C* 和 *z1D* 4 个亚族成员分别为 12 个、8 个、14 个和 5 个。基因保守 motif 分析发现该家族基因序列高度保守。表达谱分析结果显示该家族表达模式差异较大, 其中 *z1A* 亚族在籽粒发育时期表达量最高。利用 RNA-seq 数据剖析在种子发育阶段 Opaque 2(O2) 野生型和突变体编码该家族成员的表达差异, 结果显示 O2 突变体中 *z1C* 亚族成员表达显著下调, *z1A* 亚族表达略有下调, *z1B* 和 *z1D* 亚族成员表达量没有显著的变化。本研究在重新注释  $\alpha$ -醇溶蛋白家族成员信息的基础上, 对该家族的表达谱进行了分析, 为今后玉米籽粒品质育种提供了必要遗传信息。

**关键词:** 玉米;  $\alpha$ -醇溶蛋白; 基因家族表达谱; Opaque 2

**中图分类号:** S513 **文献标识码:** A **文章编号:** 1000-4440(2015)06-1224-08

## Differential expression of $\alpha$ -zein family genes regulated by Opaque2 transcription factor

LI Guo-feng<sup>1</sup>, GE Min<sup>2</sup>, LÜ Yuan-da<sup>2</sup>

(1. Luhe Research Station, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China; 2. Institute of Agro-biotechnology, Jiangsu Academy of Agricultural Sciences/Provincial Key Laboratory of Agrobiotechnology, Nanjing 210014, China)

**Abstract:** Zein proteins are main repository of nitrogen, which comprise approximately 50% of the total proteins in mature kernels. Zein proteins can be divided into four groups based on their molecular weights,  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  zeins. In contrast to the  $\beta$ ,  $\gamma$  and  $\delta$  zeins which were encoded by relatively small gene families,  $\alpha$ -zein proteins have four different gene sub-families (*z1A*, *z1B*, *z1C*, and *z1D*) and about 39 genes on maize B73 genome.  $\alpha$ -zein gene expression profiling analysis showed that their expression patterns diverged greatly and *z1A* had the highest expression level. Conserved motif analysis revealed  $\alpha$ -zein gene family was highly conserved. Opaque 2 mutant, carrying a leucine-zipper motif, downregulates  $\alpha$ -zein

收稿日期: 2015-06-07

基金项目: 江苏省农业科技自主创新基金项目[ CX(14)2009 ]

作者简介: 李国锋(1973-), 男, 江苏东台人, 硕士, 副研究员, 研究方向为栽培育种。(Tel) 025-57686218; (E-mail) ligf@jaas.ac.cn

genes expression, resulting in a so called opaque phenotype, which is widely employed in the creation of germplasm to develop quality protein maize because of the increased lysine and tryptophan accumulation. However, expression analyses of specific gene member of  $\alpha$ -zein reg-

ulated by *Opaque 2* were not well characterized. In this study, RNA-seq was employed to dissect the individual gene expression at seed developing stages and compared mRNA abundance between wild type and *Opaque 2*. The results exhibited that *Opaque 2* majorly downregulated the expression of *z1C* subfamily, slightly reduced *z1A* genes expression, and played insignificant roles in *z1B* and *z1D* genes expression. Together with the genome annotation of  $\alpha$ -zein family and their responding expression profiling, the study implied the genetic information for future improvement of maize grain quality.

**Key words:** maize;  $\alpha$ -zein; expression profiling; *Opaque 2*

玉米作为世界最大的种植作物之一,在国民经济中占据重要地位,全球对其需求量正在不断攀升。醇溶蛋白(Zein)是玉米最主要的储藏蛋白(占籽粒总蛋白的 50%以上<sup>[1-2]</sup>),对玉米籽粒品质有重要影响。醇溶蛋白根据其溶解性、氨基酸组成以及其他物理性质可以分为  $\alpha$ 、 $\beta$ 、 $\gamma$  和  $\delta$  4 类,其中  $\alpha$ -醇溶蛋白( $\alpha$ -zein)占总醇溶蛋白的 60%,含量最为丰富。玉米  $\alpha$ -醇溶蛋白富含谷氨酸、亮氨酸、脯氨酸和丙氨酸,但缺乏赖氨酸、色氨酸等几种人体和动物必需的氨基酸,且因聚集二硫键和糖基化的作用而难溶于水,在非反刍动物体内不易消化吸收,故其食用和营养价值不高<sup>[3-4]</sup>。Prasanna 等<sup>[5]</sup>通过 SDS-PAGE 电泳分析成熟籽粒中醇溶蛋白含量,发现几乎所有的优质蛋白玉米籽粒中醇溶蛋白含量较普通玉米显著较低,尤其是  $\alpha$ -醇溶蛋白,因此降低  $\alpha$ -醇溶蛋白含量一度成为提高玉米营养品质研究的重要目标之一。然而研究结果表明,过低的  $\alpha$ -醇溶蛋白虽可提高胚乳蛋白质中赖氨酸和色氨酸的含量并大大改善玉米的营养品质,但同时引起玉米胚乳质地松软和产量无法提高等问题<sup>[6]</sup>。

$\alpha$ -醇溶蛋白由一个庞大的超级基因家族编码,根据序列相似性又分为 4 个基因亚族:*z1A*、*z1B*、*z1C*、*z1D*,其中 *z1A*、*z1B* 和 *z1D* 3 个基因亚族属于 19 000  $\alpha$ -醇溶蛋白,*z1C* 为 22 000  $\alpha$ -醇溶蛋白<sup>[7-8]</sup>。Song 等<sup>[6]</sup>利用 BAC 文库测序,发现在玉米 B73 自交系中 19 000  $\alpha$ -醇溶蛋白基因家族组成了玉米基因组上最长的连续重复基因序列区域,19 000  $\alpha$ -醇溶蛋白基因家族中,*z1A* 包含 12 个基因成员,*z1B* 包含 8 个基因成员,*z1D* 包含 5 个成员,它们分别位于 3 条染色体(第 1、第 4、第 7 条染色体)的 4 个不同基因组区域内。这 25 个基因成员 DNA 序列较为保守且基因转录方向基本一致(Head-to-tail orientation),但是其中只有 12 个基因具有完整的编码区域,另外 13 个基因由于突变等未知原因无法形成完整的转录本。Song 等<sup>[8-9]</sup>利用 BAC 文库测序,基于比较基

因组学分析,在 B73 和 BSSS53 自交系中,发现 22 000  $\alpha$ -醇溶蛋白基因家族分别由 15 个和 23 个成员组成,它们位于第 4 号染色体的短臂上。*z1C* 成员大致上是呈现串联阵列,基因转录方向完全一致。而在第 4 号染色体的短臂上距离着丝粒-20 cM 另外一个区域内存在一个游离于串联阵列之外的 *z1C* 成员 *floury2*。这些 *z1C* 成员在 B73 和 BSSS53 中分别只有 7 个和 8 个具有完整的编码区域。尽管  $\alpha$ -醇溶蛋白超级基因家族的大小和复杂度在不同的自交系中有所不同,但是序列却高度保守。

Feng 等<sup>[10]</sup>结合实时定量 PCR 和随机克隆测序的研究方法,对玉米自交系 B73 胚乳发育时期  $\alpha$ -醇溶蛋白超基因家族(41 个基因成员)的表达模式进行研究,结果表明该基因家族中仅 18 个基因表达,表达水平差异很大。该家族所有亚族(*z1A*、*z1B*、*z1C*、*z1D*)在授粉后 10~34 d 均表现出上下摆动不规则的表达模式,总体而言 *z1A* 具有最高表达量,然后表达量由高到低依次为 *z1C*、*z1D* 和 *z1B*。亚族基因成员的表达模式也存在显著差异,但是,由于  $\alpha$ -醇溶蛋白基因家族成员在编码序列存在高度的相似性,实时定量 PCR 对部分基因无法精确区分。二代测序技术如 RNA-seq 具有高敏感性和高通量等特征,可精确检测到基因家族中相似基因的不同转录本的单碱基差异。利用 RNA-seq 获得的高通量数据结合序列相似性比较和覆盖深度计算,可以对玉米  $\alpha$ -醇溶蛋白基因家族成员基因表达进行高精度解析。

玉米中首个发现能调控  $\alpha$ -醇溶蛋白的反式因子为 *Opaque2*(O2),属于 bZIP 转录因子家族,O2 基因在玉米胚乳中特异表达。O2 突变体中  $\alpha$ -醇溶蛋白基因的表达下调,导致籽粒中赖氨酸和色氨酸的含量增加,但也导致籽粒胚乳粉质呈现出透明的表型<sup>[11-13]</sup>。基于上述原因该突变体广泛应用于优质蛋白玉米种质资源的创制研究,但是 *Opaque2* 对  $\alpha$ -醇溶蛋白基因家族成员的表达是否均存在影响并没有得到很好的解析。

本研究拟根据  $\alpha$ -醇溶蛋白基因家族成员的保守序列修订该基因家族在玉米 B73 染色体上的位置,并利用 RNA-seq 数据对玉米自交系 B73 胚乳中从授粉至成熟各时期编码  $\alpha$ -醇溶蛋白的各个基因的表达模式进行分析,解析该基因家族成员的表达特性,并通过随机克隆 Sanger 测序法对结果进行验证,同时比较 *Opaque2* 野生型和突变型的编码  $\alpha$ -醇溶蛋白基因家族成员的表达差异。

## 1 材料与方法

### 1.1 玉米 $\alpha$ -醇溶蛋白基因的鉴定与分类

基与最新版 (V3) 玉米自交系 B73 参考基因组数据 (<http://www.maizegdb.org/>), 利用 tBlastx 方法以  $\alpha$ -醇溶蛋白 cDNAs (NCBI) 为目标序列在玉米基因组中搜寻同源区域, 并利用生物信息学方法获得同源区域对应基因的位置及核酸序列。然后利用 ClustalX 软件对玉米  $\alpha$ -醇溶蛋白基因进行多重序列比对, 并根据序列相似性进行人工校正分类。

### 1.2 玉米 $\alpha$ -醇溶蛋白基因保守 Motif 分析

为了进一步解析  $\alpha$ -醇溶蛋白基因的结构和保守性, 我们利用 MEME 在线分析工具<sup>[14]</sup> (<http://meme-suite.org/index.html>) 对本研究所鉴定的玉米  $\alpha$ -醇溶蛋白基因的保守 Motif 进行分析。其中最大保守 motif 数目设置为 10 条, motif 长度设置为 6~50 bp。

### 1.3 玉米 $\alpha$ -醇溶蛋白基因的表达谱分析及验证

为了进一步解析玉米  $\alpha$ -醇溶蛋白基因表达模式, 我们利用玉米自交系 B73 胚乳自授粉到成熟 17 个发育时期 RNA-Seq 数据<sup>[15]</sup>, 系统分析  $\alpha$ -醇溶蛋白基因在玉米 B73 中的表达谱模式。RNA-Seq 数据的处理及  $\alpha$ -醇溶蛋白基因家族表达量 FPKM 值的获得均参照本实验室前期研究方法<sup>[16]</sup>, 最终利用 MEV4.9 软件绘制玉米  $\alpha$ -醇溶蛋白基因家族的 Heatmap 图, 并分析表达模式。为验证本研究所利用 RNA-Seq 测序结果的可靠性, 利用随机克隆 Sanger 测序法对结果进行了验证, 比较 2 种测序方法中玉米 B73 授粉 14 d 后胚乳中  $\alpha$ -醇溶蛋白基因相对表达量的一致性。

### 1.4 *Opaque2* 对玉米 $\alpha$ -醇溶蛋白基因表达的影响

利用 Li 等<sup>[17]</sup> 文章中野生型和突变体 *Opaque2* 授粉后 15 d 籽粒的 RNA-seq 数据, 并对

其进行处理。最终整理出野生型和突变体中  $\alpha$ -醇溶蛋白基因家族所有成员的相对表达量, 比较分析转录因子 *Opaque2* 对玉米  $\alpha$ -醇溶蛋白基因表达的影响。

## 2 结果与分析

### 2.1 玉米 $\alpha$ -醇溶蛋白基因的确定及分类

利用 tBlastx 方法在玉米基因组中共比对搜索到 39 条  $\alpha$ -醇溶蛋白基因序列。利用 ClustalX 软件对基因序列进行多重比对和人工校正, 玉米  $\alpha$ -醇溶蛋白基因分类结果见表 1 (参照前人研究结果对其命名), 其中 28 条序列与目标序列相似度为 100.00%, 其余 11 条序列相似度高达 97.75% 以上, 说明本研究比对结果可信度高。*z1A* 亚族包含 12 个基因成员 (*Z19A1-1*~*Z19A1-9*, *Z19A2-1*~*Z19A2-3*); *z1B* 包含 8 个成员 (*Z19B1-1*~*Z19B1-6*, *Z19B2-1*~*Z19A2-2*); *z1D* 包含 5 个成员 (*Z19D1-1*~*Z19D1-3*, *Z19D2-1*, *Z19D3-1*), 3 个亚族分别位于第 4、第 7 和第 1 染色体上; *z1C* 亚族 14 个成员位于第 4 染色体上, 其中 13 个呈现串联阵列转录方向完全一致, 仅 AZS22-16 游离于串联阵列之外, 该基因为 *floury2*。19 000  $\alpha$ -醇溶蛋白共 25 个 (*z1A*, *z1B* 和 *z1D*), 22 000  $\alpha$ -醇溶蛋白共 14 个 (*z1C*)。  $\alpha$ -醇溶蛋白超家族基因长度均值为 872 bp, 其中 *z1B* 亚族基因长度均值最长 (1 507 bp), 其余 3 个亚族 (*z1A*, *z1C* 和 *z1D*) 基因长度均值相近, 约 700 bp。39 个  $\alpha$ -醇溶蛋白基因家族中有 30 个可以对应到前人预测的基因 ID 上, 9 个基因还未预测出基因 ID 的表示为“-”。

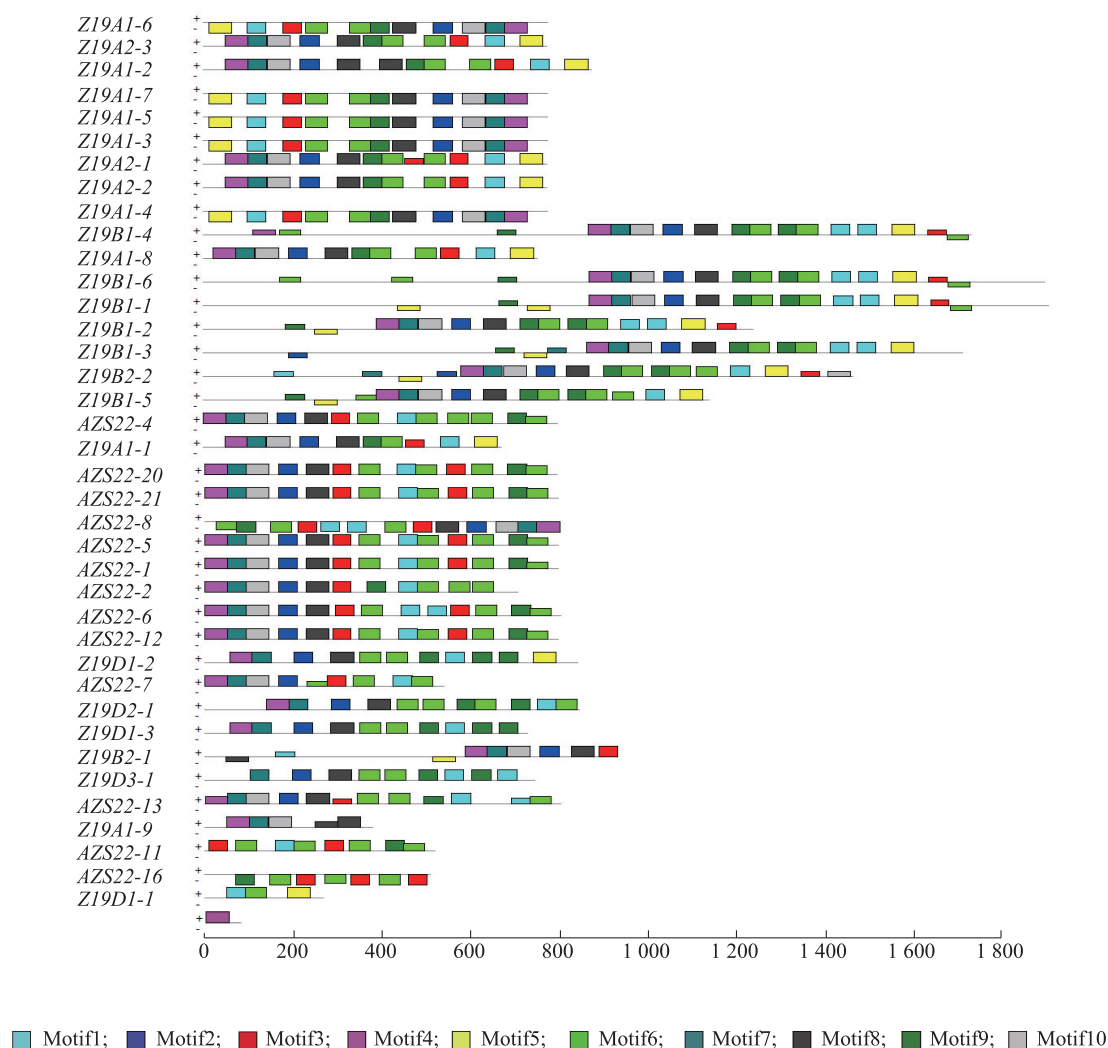
### 2.2 玉米 $\alpha$ -醇溶蛋白基因保守 Motif 分析

通过 MEME 在线分析工具对本研究所鉴定的 39 个  $\alpha$ -醇溶蛋白基因的保守 motif 进行分析, 结果如图 1 (图中  $\alpha$ -醇溶蛋白基因家族成员按照比对 *E* 值大小排列) 所示。图 1 显示,  $\alpha$ -醇溶蛋白基因家族成员含有的保守 motif 数目和排列位置总体相似度高, 尤其在亚族成员间相似性非常高, 说明该家族基因序列高度保守。10 条 motif 长度为 40~50 bp, 其中 motif 6 含量最为丰富, 在  $\alpha$ -醇溶蛋白基因中存在 2~4 个重复, 此外 motif 9、motif 5 和 motif 1 也存在重复。  $\alpha$ -醇溶蛋白基因家族中 AZS22-3、AZS22-11、AZS22-16、*Z19D1-1*、*Z19A1-9* 基因所含 motif 少 (1~4 个), 其余基因家族成员含有 motif 数目较为一致。

表 1 玉米  $\alpha$ -醇溶蛋白基因Table 1  $\alpha$ -zein gene family in maize

亚族	基因	基因位置	长度 (bp)	基因 ID	相似度 (%)	RNA-Seq 测序 相对表达量	Sanger 测序 相对表达量
<i>z1A</i>	<i>Z19A1-1</i>	染色体 4:5495175~5495849	675	—	100.00	0.000 846	0
	<i>Z19A1-2</i>	染色体 4:5522254~5523132	879	GRMZM2G404459_T01	100.00	1.318 423	0.535 278
	<i>Z19A1-3</i>	染色体 4:5560581~5561360	780	GRMZM2G353268_T01	100.00	11.674 450	7.653 391
	<i>Z19A1-4</i>	染色体 4:5548833~5549612	781	GRMZM2G353272_T01	99.74	10.202 170	8.026 811
	<i>Z19A1-5</i>	染色体 4:5544803~5545582	780	GRMZM2G053120_T01	100.00	8.688 846	7.285 333
	<i>Z19A1-6</i>	染色体 4:5540363~5541142	780	GRMZM2G059620_T02	100.00	7.307 772	6.752 203
	<i>Z19A1-7</i>	染色体 4:5535901~5536680	780	GRMZM2G059620_T03	100.00	6.442 894	8.445 043
	<i>Z19A1-8</i>	染色体 4:5594920~5595675	756	GRMZM2G018193_T01	100.00	0.005 128	0.118 546
	<i>Z19A1-9</i>	染色体 4:5599189~5599569	381	—	100.00	0.000 046	0
	<i>Z19A2-1</i>	染色体 4:29951782~29952558	777	GRMZM2G008341_T01	100.00	8.140 605	6.065 812
	<i>Z19A2-2</i>	染色体 4:29963346~29964122	777	GRMZM2G008913_T01	100.00	9.803 198	7.384 106
	<i>Z19A2-3</i>	染色体 4:30089932~30090708	777	GRMZM2G026939_T01	100.00	0.331 056	1.211 093
<i>z1B</i>	<i>Z19B1-1</i>	染色体 7:18755067~18757677	1 911	AF546188.1_FGT001	100.00	0.007 885	0.680 645
	<i>Z19B1-2</i>	染色体 7:18779184~18780427	1 244	AF546188.1_FGT002	100.00	0.096 407	1.134 669
	<i>Z19B1-3</i>	染色体 7:18789010~18790726	1 717	AF546188.1_FGT003	100.00	0.016 149	0.208 785
	<i>Z19B1-4</i>	染色体 7:18864555~18866290	1 736	AF546188.1_FGT005	100.00	9.474 392	7.008 083
	<i>Z19B1-5</i>	染色体 7:18867341~18868484	1 144	AF546188.1_FGT006	100.00	0.034 027	0.078 340
	<i>Z19B1-6</i>	染色体 7:18876589~18878491	1 903	AF546188.1_FGT007	100.00	7.697 904	5.979 487
	<i>Z19B2-1</i>	染色体 7:18923572~18924505	934	—	100.00	0.011 965	0.023 988
	<i>Z19B2-2</i>	染色体 7:18954235~18955704	1 470	GRMZM2G025763_T01	100.00	0.001 413	0.060 967
<i>z1D</i>	<i>Z19D1-1</i>	染色体 1:161788738~161789 006	269	AF546187.1_FGT001	100.00	0	0
	<i>Z19D1-2</i>	染色体 1:161790654~161791499	846	AF546187.1_FGT007	100.00	1.445 349	4.779 829
	<i>Z19D1-3</i>	染色体 1:161795434~161796165	732	AF546187.1_FGT009	100.00	0.008 367	0.091 825
	<i>Z19D2-1</i>	染色体 1:162006771~162007620	850	—	100.00	1.320 240	2.652 358
	<i>Z19D3-1</i>	染色体 1:162096303~162097050	748	—	100.00	0.000 308	0.359 442
<i>z1C</i>	<i>AZS22-1</i>	染色体 4:5094190~5094990	801	GRMZM2G088441_T01	99.88	0.720 629	0.167 829
	<i>AZS22-2</i>	染色体 4:5098300~5099009	710	GRMZM2G388461_T01	99.58	0.392 931	0.315 566
	<i>AZS22-3</i>	染色体 4:5114102~5114180	79	—	100.00	0.032 511	0.476 707
	<i>AZS22-4</i>	染色体 4:5122517~5123317	801	GRMZM2G044625_T01	99.88	2.776 458	8.762 073
	<i>AZS22-5</i>	染色体 4:5127723~5128523	801	GRMZM2G045387_T01	100.00	0.319 236	0.699 289
	<i>AZS22-6</i>	染色体 4:5141270~5142076	807	GRMZM2G346895_T01	99.88	0.149 884	1.443 707
	<i>AZS22-7</i>	染色体 4:5145381~5145923	543	—	99.00	2.717 977	0.371 357
	<i>AZS22-8</i>	染色体 4:5158097~5158897	801	GRMZM2G044152_T01	99.88	1.883 320	2.153 810
	<i>AZS22-11</i>	染色体 4:5080126~5080647	522	—	100.00	0.001 106	0.017 195
	<i>AZS22-12</i>	染色体 4:5083169~5083969	801	GRMZM2G088365_T01	99.88	0.036 762	0.279 716
	<i>AZS22-13</i>	染色体 4:5086264~5087070	807	—	99.88	0.000 029	0.015 603
	<i>AZS22-16</i>	染色体 4:21335099~21335609	511	GRMZM2G397687_T01	100.00	3.042 744	5.258 347
	<i>AZS22-21</i>	染色体 4:5172023~5172823	801	GRMZM2G160739_T01	97.75	2.094 780	1.684 603
	<i>AZS22-20</i>	染色体 4:5176140~5176937	798	GRMZM2G160739_T03	97.87	1.801 817	1.818 151



图1 玉米 $\alpha$ -醇溶蛋白基因保守 motif 分析Fig.1 Conserved motif analyses of  $\alpha$ -zein genes in maize

### 2.3 玉米籽粒发育时期 $\alpha$ -醇溶蛋白基因表达谱分析

对玉米胚乳自授粉至成熟 17 个发育时期 RNA-Seq 数据进行预处理和统计分析后,获得了玉米 $\alpha$ -醇溶蛋白基因对应表达量的 FPKM 值,根据 FPKM 值利用 MEV4.9 软件绘制出表达量的 Heatmap 图谱(图 2)。玉米 $\alpha$ -醇溶蛋白基因家族成员表达模式差异较大,Z19A1-9、Z19D1-1 和 AZS22-13 这 3 个基因表达量最低,在图 2 中 17 个时期 FPKM 均值低于 1,Z19A1-1、Z19A1-8、Z19B2-2、Z19D1-3、Z19D3-1 和 AZS22-11 这 6 个基因表达量也很低,FPKM 均值低于 50,Z19A2-2 表达量最高,FPKM 均值为 29 176。其中 $\alpha$ -醇溶蛋白超家族 4 个亚族表达量 FPKM 均值大小依次为 *z1A* (13 593)、*z1B* (4 910)、*z1C* (4 187) 和 *z1D* (2 589)。

玉米 $\alpha$ -醇溶蛋白基因在胚乳中从授粉后 10 d 开始表达,授粉 38 d 后醇溶蛋白基因几乎不表达。 $\alpha$ -醇溶蛋白 *z1A* 亚族约 58% (7/12) 的成员和 *z1B* 亚族约 25% (2/8) 的成员在授粉 12~32 d 后的胚乳中表达量持续较高,*z1D* 亚族约 40% (2/5) 的成员在授粉后 28 d 的胚乳中出现最高表达量,*z1C* 亚族约 43% (6/14) 的成员在授粉 22~28 d 后的胚乳中表现出最高的表达量。

为验证本研究所利用 RNA-seq 数据的可靠性,我们将 $\alpha$ -醇溶蛋白基因家族成员在授粉 14 d 后 Sanger 测序与二代测序结果进行了比较。如表 1 所示,2 种测序方法所得表达量数值非常一致,说明本研究分析所得 $\alpha$ -醇溶蛋白基因的表达谱数据非常可信。



a: 授粉 6 d 后胚乳;b: 授粉 8 d 后胚乳;c: 授粉 10 d 后胚乳;d: 授粉 12 d 后胚乳;e: 授粉 14 d 后胚乳;f: 授粉 16 d 后胚乳;g: 授粉 18 d 后胚乳;h: 授粉 20 d 后胚乳;i: 授粉 22 d 后胚乳;j: 授粉 24 d 后胚乳;k: 授粉 26 d 后胚乳;l: 授粉 28 d 后胚乳;m: 授粉 30 d 后胚乳;n: 授粉 32 d 后胚乳;o: 授粉 34 d 后胚乳;p: 授粉 36 d 后胚乳;q: 授粉 38 d 后胚乳。图中黑白颜色深浅表示表达量的高低,白色表示最低表达,黑色表示最高表达。

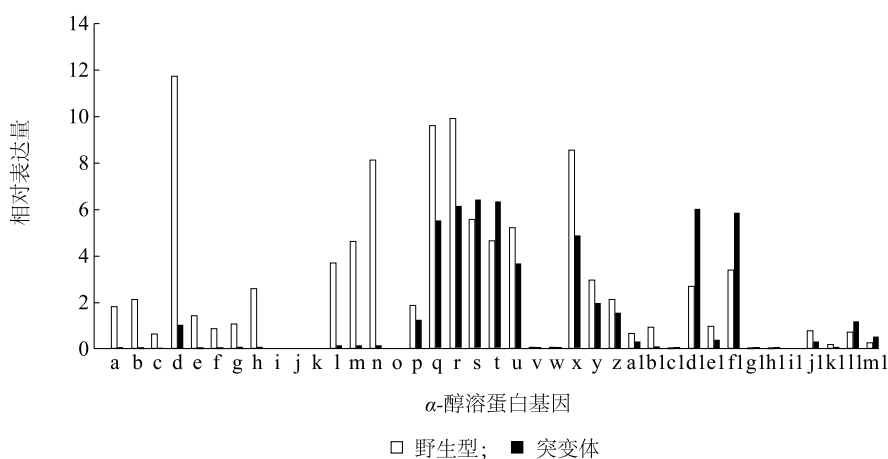
图 2 玉米  $\alpha$ -醇溶蛋白基因表达模式分析  
Fig.2 Expression patterns of  $\alpha$ -zein genes in maize

2.4 Opaque2 对玉米  $\alpha$ -醇溶蛋白基因表达的影响

根据野生型和突变体 *Opaque2* 授粉后 15 d 籽粒的 RNA-seq 数据,最终整理出野生型和突变体中  $\alpha$ -醇溶蛋白基因相对表达量。如图 3 所示,突变体中 *z1C* 亚族成员相对表达量较野生型显著下调,*z1A* 亚族成员表达量略有下调,但 *z1B* 和 *z1D* 亚族成员表达量没有显著的变化。以上结果说明 *Opaque2* 突变主要影响 *z1C* 基因亚族成员的表达,而对 *z1B* 和 *z1D* 的表达影响不显著,转录因子 *Opaque2* 为 *z1C* 亚族基因表达所必须的反式作用因子。

3 讨论

本研究基于玉米 v3 版基因组数据共鉴定  $\alpha$ -醇溶蛋白基因 39 个,其中 19 000  $\alpha$ -醇溶蛋白家族成员 25 个,22 000  $\alpha$ -醇溶蛋白家族成员 14 个。Song 等<sup>[8]</sup>在玉米 BSSS53 中共鉴定 22 000  $\alpha$ -醇溶蛋白 23 个,Song 等<sup>[7]</sup>在玉米 B73 中鉴定 19 000  $\alpha$ -醇溶蛋白 25 个,Song 等<sup>[8]</sup>分别在玉米 B73 和 BSSS53 鉴定 22 000  $\alpha$ -醇溶蛋白 15 个和 23 个,Xu 等<sup>[18]</sup>报道玉米 B73 中 19 000 和 22 000  $\alpha$ -醇溶蛋白数目分别为 26 个和 15 个,Feng 等<sup>[9]</sup>在玉米 B73 中鉴定出 19 000 和 22 000  $\alpha$ -醇溶蛋白数目分别为 25 个和 16 个。2011



a:AZS22-1;b:AZS22-2;c:AZS22-3;d:AZS22-4;e:AZS22-5;f:AZS22-6;g:AZS22-7;h:AZS22-8;i:AZS22-11;j:AZS22-12;k:AZS22-13;l:AZS22-16;m:AZS22-20;n:AZS22-21;o:Z19A1-1;p:Z19A1-2;q:Z19A1-3;r:Z19A1-4;s:Z19A1-5;t:Z19A1-6;u:Z19A1-7;v:Z19A1-8;w:Z19A1-9;x:Z19A2-1;y:Z19A2-2;z:Z19A2-3;a1:Z19B1-1;b1:Z19B1-2;c1:Z19B1-3;d1:Z19B1-4;e1:Z19B1-5;f1:Z19B1-6;g1:Z19B2-1;h1:Z19B2-2;i1:Z19D1-1;j1:Z19D1-2;k1:Z19D1-3;l1:Z19D2-1;m1:Z19D3-1。

图3 *Opaque 2* 对玉米  $\alpha$ -醇溶蛋白基因表达的影响

Fig.3 Effects of *Opaque 2* on the expression of  $\alpha$ -zein genes in maize

年 Miclaus 等<sup>[19]</sup>发现玉米 B73 中  $\alpha$ -醇溶蛋白有 41 个(19 000 和 22 000  $\alpha$ -醇溶蛋白分别为 26 个和 15 个), BSSS53 中  $\alpha$ -醇溶蛋白有 48 个(19 000 和 22 000  $\alpha$ -醇溶蛋白分别为 25 个和 23 个)。以上结果显示不同时期所鉴定的玉米  $\alpha$ -醇溶蛋白基因数目有所不同,目前仍没有一致结果。本研究是在前人研究结果的基础上,利用 NCBI 数据库中所有  $\alpha$ -醇溶蛋白基因的 cDNA 序列在玉米基因组中搜寻同源序列,并利用生物信息学方法定位到相关基因并对其位置进行修正,还得到对应基因的预测基因 ID。总体而言,本研究所鉴定的玉米  $\alpha$ -醇溶蛋白基因家族更为全面和准确,为进一步研究该基因家族提供了一定的依据。

$\alpha$ -醇溶蛋白基因家族各亚族成员在第 1、第 4 和第 7 条染色体 6 个区域内分布,与 Miclaus 等<sup>[19]</sup>研究结果相似。通过对  $\alpha$ -醇溶蛋白基因家族的保守 motif 分析发现该家族基因序列高度保守,某些 motif 存在重复现象。以上结果说明  $\alpha$ -醇溶蛋白基因家族在进化上同源关系很近,存在串联重复现象。Feng 等<sup>[10]</sup>发现  $\alpha$ -醇溶蛋白基因家族 41 个成员中仅 18 个基因表达。本研究利用 RNA-seq 获得高通量数据,结合序列相似性比较和覆盖深度计算,发现玉米  $\alpha$ -醇溶蛋白基因家族中大部分基因均表达(约 30 个),以上结果说明 RNA-seq 技术具有高敏感性,

可精确检测到基因家族中相似基因不同转录本的单碱基差异。

$\alpha$ -醇溶蛋白受顺式因子和反式因子协同调控。保守的顺式因子主要位于  $\alpha$ -醇溶蛋白基因转录起始位点上游约 300 bp 处,这一保守区被称为胚乳盒。胚乳盒含有 3 类不同的保守序列模块:第 1 类是胚乳模块(EMs);第 2 类是 GCN4 结合位点类似模块(GLMs);第 3 类是醇溶谷蛋白盒(P 盒)。P 盒序列与 EMs 序列相同。P 盒和 GLMs 可以介导  $\alpha$ -醇溶蛋白基因的胚乳特异性表达。玉米中第一个发现调控  $\alpha$ -醇溶蛋白的反式因子是一个 bZIP 类蛋白,称为 *Opaque2* (O2), O2 基因在玉米胚乳中特异表达,所编码的蛋白结合 GLMs 位点,正向调控玉米 22 000  $\alpha$ -醇溶蛋白、b-32 基因和 *cyPPDK1* 基因的表达<sup>[13]</sup>。Vicente-Carbajosa 等<sup>[20]</sup>发现另一个调节  $\alpha$ -醇溶蛋白基因表达的反式因子 PBF (Prolamin box binding factor), PBF 属于植物 Cys2-Cys2 锌指结构 DNA 结合蛋白 Dof 成员,专化性结合在  $\alpha$ -醇溶蛋白基因启动子的 P 盒区。同样, PBF 蛋白质可以特异地与 bZIP 蛋白质相互作用,促进 bZIP 蛋白质与启动子中靶序列的结合。PBF 和 O2 在玉米胚乳组织中以相同的方式表达,都是在玉米醇溶蛋白基因转录激活之前开始积累。O2 与 PBF 的相互作用对于玉米醇溶蛋白基因的转录激活是非常必要的<sup>[21]</sup>。

O2 的自然突变或 RNAi 抑制可显著降低 22 000  $\alpha$ -醇溶蛋白和其他  $\alpha$ -醇溶蛋白基因的表达,导致胚乳粉质、产量及千粒质量降低,籽粒水分含量偏高,易破碎和易感染病虫害,因而使 O2 的自然突变无法直接用于育种<sup>[22-23]</sup>。目前,未见有关 PBF 的自然突变或 RNAi 抑制后导致玉米性状表现改变的报道,究其原因因为 PBF 除了参与  $\alpha$ -醇溶蛋白的调控,还可能涉及到种子发育的其他关键途径。本研究发现  $\alpha$ -醇溶蛋白基因家族序列高度保守,但表达谱差异较大,可能与上述顺式和反式协同调控有关。本研究发现转录因子 Opaque2 主要调控 *z1C* 亚族(22 000  $\alpha$ -醇溶蛋白)的表达(与前人研究结果一致)<sup>[12]</sup>,对 *z1A* 亚族的表达也有影响,但对 *z1B* 和 *z1D* 亚族的表达影响作用不明显。

#### 参考文献:

- [1] LARKINS B A, BRACKER C E, TSAI C Y. Storage protein synthesis in maize: isolation of zein-synthesizing polyribosomes[J]. Plant Physiology, 1976, 57(5): 740-745.
- [2] 康美玲,田忠景,张倩倩. 利用醇溶蛋白电泳图谱分析不同玉米品种的遗传多样性[J]. 江苏农业科学,2013,41(10): 70-72.
- [3] HARTINGS H, LAURIA M, LAZZARONI N, et al. The zein mutants opaque-2 and opaque-7 disclose extensive changes in endosperm metabolism as revealed by protein, amino acid, and transcriptome-wide analyses[J]. BMC Genomics,2011, 12: 41.
- [4] HOLDING D R, HUNTER B G, CHUNG T, et al. Genetic analysis of opaque2 modifier loci in quality protein maize[J]. Theor Appl Genet, 2008, 117(2): 157-170.
- [5] PRASANNA B M, VASAL S K, KASSAHUN B, et al. Quality protein maize[J]. Current Science-Bangalore, 2001, 81(10): 1308-1319.
- [6] WU Y, MESSING J. Novel genetic selection system for quantitative trait loci of quality protein maize[J]. Genetics,2011, 188(4): 1019-1022.
- [7] SONG R, MESSING J. Contiguous genomic DNA sequence comprising the 19-kD zein gene family from maize[J]. Plant Physiology, 2002, 130(4): 1626-1635.
- [8] SONG R, LLACA V, LINTON E, et al. Sequence, regulation, and evolution of the maize 22-kD alpha zein gene family[J]. Genome Research, 2001, 11(11): 1817-1825.
- [9] SONG R, MESSING J. Gene expression of a gene family in maize based on noncollinear haplotypes[J]. Proc Natl Acad Sci USA, 2003, 100(15): 9055-9060.
- [10] FENG L, ZHU J, WANG G, et al. Expressional profiling study revealed unique expressional patterns and dramatic expressional divergence of maize alpha-zein super gene family[J]. Plant MolBiol, 2009, 69(6): 649-659.
- [11] SCHMIDT R J, BURR F A, BURR B. Transposon tagging and molecular analysis of the maize regulatory locus opaque-2[J]. Science, 1987, 238(4829): 960-963.
- [12] HARTINGS H, MADDALONI M, LAZZARONI N, et al. The O2 gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators[J]. EMBO J, 1989, 8(10): 2795-2801.
- [13] MULLER M, DUES G, BALCONI C, et al. Nitrogen and hormonal responsiveness of the 22 kDa alpha-zein and b-32 genes in maize endosperm is displayed in the absence of the transcriptional regulator Opaque-2[J]. Plant Journal, 1997, 12(2): 281-291.
- [14] BAILEY T L, BODEN M, BUSKE F A, et al. MEME SUITE: tools for motif discovery and searching[J]. Nucleic Acids Res, 2009, 37(Web Server issue): W202-W208.
- [15] CHEN J, ZENG B, ZHANG M, et al. Dynamic transcriptome landscape of maize embryo and endosperm development[J]. Plant Physiology, 2014, 166(1): 252-264.
- [16] 葛敏,吕远大,张体付,等. 玉米 YABBY 基因家族的全基因组鉴定与分析[J]. 江苏农业学报,2014,30(6): 1267-1272.
- [17] LI C, QIAO Z, QI W, et al. Genome-wide characterization of cis-acting DNA targets reveals the transcriptional regulatory framework of opaque2 in maize[J]. Plant Cell, 2015, 27(3): 532-545.
- [18] XU J H, MESSING J. Organization of the prolamin gene family provides insight into the evolution of the maize genome and gene duplications in grass species[J]. Proc Natl Acad Sci USA, 2008, 105(38): 14330-14335.
- [19] MICLAUS M, XU J H, MESSING J. Differential gene expression and epiregulation of alpha zein gene copies in maize haplotypes[J]. PLoS Genetics, 2011, 7(6): e1002131.
- [20] VICENTE-CARBAJOSA J, MOOSE S P, PARSONS R L, et al. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2[J]. Proc Natl Acad Sci USA,1997, 94(14): 7685-7690.
- [21] HWANG Y S, CICERI P, PARSONS R L, et al. The maize O2 and PBF proteins act additively to promote transcription from storage protein gene promoters in rice endosperm cells[J]. Plant Cell Physiology, 2004, 45(10): 1509-1518.
- [22] WU Y, MESSING J. Rescue of a dominant mutant with RNA interference[J]. Genetics, 2010, 186(4): 1493-1496.
- [23] WU Y, MESSING J. RNA interference-mediated change in protein body morphology and seed opacity through loss of different zein proteins[J]. Plant Physiology, 2010, 153(1): 337-347.

(责任编辑:陈海霞)