

黄娟, 邓娇, 石桃雄, 等. 荞麦 3 个种子蛋白基因的序列和表达分析[J]. 华南农业大学学报, 2017, 38(1): 23-30.

荞麦 3 个种子蛋白基因的序列和表达分析

黄娟, 邓娇, 石桃雄, 梁成刚, 张启迪, 陈庆富
(贵州师范大学 荞麦产业技术研究中心, 贵州 贵阳 550001)

摘要:【目的】阐明荞麦 *Fagopyrum* 中 3 个种子蛋白基因序列的特征及其在种子发育过程中的作用, 为探讨荞麦种子发育过程分子机制提供理论依据。【方法】利用本地 Perl 程序、BLASTP、MEGA6、PLANTCARE 等, 分析了荞麦 3 个种子蛋白基因(13S 球蛋白基因、11S 球蛋白基因和谷蛋白基因)序列及其编码蛋白质的进化关系, 预测启动子区域的顺式作用元件。通过荧光定量 PCR 技术分析 3 个基因在甜荞 *F. esculentum*、苦荞 *F. tataricum* 和金荞 *F. cymosum* 种子发育过程中的表达情况。【结果】13S 球蛋白基因含有 4 个外显子和 3 个内含子, 蛋白序列长 515 aa, 与苦荞中的一个 13S 球蛋白(GeneBank 登录号: ABI32184.1)仅有 6 个氨基酸不同, 推测二者为同一个基因, 在苦荞灌浆中期表达量最高; 11S 球蛋白基因含有 8 个外显子和 7 个内含子, 蛋白序列长 914 aa, 与欧洲慈姑中的一个 11S 球蛋白(GeneBank 登录号: ABB77213.1)亲缘关系最近, 在苦荞灌浆中期表达量最高; 谷蛋白基因含有 3 个外显子和 2 个内含子, 蛋白序列长 355 aa, 与甜菜中的一个 B5 型谷蛋白(GeneBank 登录号: XP_010683769.1)亲缘关系最近, 在甜荞灌浆中期表达量最高。3 个基因的启动子区域均含有大量响应逆境和激素的顺式作用元件, 以及在胚乳中表达所需的顺式作用元件。【结论】本研究获得了 1 个 13S 球蛋白基因序列、1 个 11S 球蛋白基因序列和 1 个谷蛋白基因序列, 这 3 个基因可能在胚乳中优势表达, 并且参与到逆境胁迫和激素响应的过程中; 灌浆中期是种子蛋白积累的关键时期, 3 个基因均在这一时期大量表达, 这为后续进一步研究基因功能奠定了基础。

关键词: 荞麦; 种子蛋白; 灌浆期; 基因表达
中图分类号: S517 **文献标志码:** A **文章编号:** 1001-411X(2017)01-0023-08

Sequence and expression analyses of three seed protein genes of buckwheat

HUANG Juan, DENG Jiao, SHI Taoxiong, LIANG Chenggang, ZHANG Qidi, CHEN Qingfu
(Research Center of Buckwheat Industry Technology, Guizhou Normal University, Guiyang 550001, China)

Abstract:【Objective】To clarify the sequence features and functions of three seed protein genes of buckwheat, providing a theoretical foundation for studying the molecular mechanism during seed development. 【Method】Using local Perl script, BLASTP, MEGA6 and PLANTCARE, the sequences and evolutionary relationships of three buckwheat seed protein genes, namely 13S globulin gene, 11S globulin gene and glutelin gene, were analyzed. The cis-regulatory elements in the promoter region were predicted. Using the qRT-PCR technology, the expression patterns of three genes were analyzed during seed development of *Fagopyrum esculentum*, *F. tataricum* and *F. cymosum*. 【Result】The 13S globulin gene included four exons and three introns. The protein length was 515 aa. This 13S globulin only had six different amino acid sites compared with another 13S globulin in *F. tataricum* (ABI32184.1), thus they might be the same gene. The expression level was the highest in the middle filling stage of *F. tataricum*. The 11S globulin gene included eight exons and seven introns. The protein length was 914 aa,

收稿日期: 2016-03-06 优先出版时间: 2016-12-28
优先出版网址: <http://www.cnki.net/kcms/detail/44.1110.s.20161228.0937.030.html>
作者简介: 黄娟(1988—), 女, 博士, E-mail: huang200669@163.com; 通信作者: 陈庆富(1966—), 男, 教授, 博士, E-mail: cqf1966@163.com
基金项目: 国家自然科学基金(31471562, 31171609); 国家现代农业产业技术体系专项资金(CARS-08-A4); 贵州省荞麦工程技术研究中心项目(黔科合农G字【2015】4003号); 贵州师范大学博士基金资助项目(11904/0516031, 11904/0516026)

and it was most closely related to a 11S globulin in *Sagittaria sagittifolia* (CAA70334.1). The expression level was the highest in the middle filling stage of *F. tataricum*. The glutelin gene included three exons and two introns. The protein length was 355 aa, and it was most closely related to a B5-type glutelin in *Beta vulgaris* (XP_010683769.1). The expression level was the highest in the middle filling stage of *F. esculentum*. The promoter regions of three genes contained various cis-acting elements which responded to stresses and phytohormones, as well as cis-acting elements required for endosperm expression. 【Conclusion】 The sequences of 13S globulin, 11S globulin and glutelin genes are obtained. These genes might be preferentially expressed in endosperm, and involved in stress and hormone responsive processes. The middle filling stage is the key stage for seed protein accumulation, and three genes are massively expressed at this stage, which provides a solid foundation for studying gene function in the future.

Key words: buckwheat; seed protein; filling stage; gene expression

荞麦属于蓼科荞麦属 *Fagopyrum* 1 年生双子叶草本植物,是国际粮农组织公认的药食兼用粮种。中国是荞麦种植面积最大、产量最高的国家^[1],且主要分布在西部贫困山区,成为这些地区的主要杂粮和经济作物^[2]。种子是荞麦的主要食用部分,荞麦种子中不仅富含人体所必需的蛋白质、淀粉、脂肪等大量元素,还含有丰富的膳食纤维、维生素、矿物质等微量元素,以及禾本科粮食作物所不具备的黄酮、类黄酮等重要的次生代谢物^[3],具有降三高、抗肿瘤、抗氧化、抗疲劳、增强免疫力等功效^[4]。蛋白质是荞麦成熟种子中的重要组成物质之一,荞麦籽粒富含中、低分子量蛋白质,缺乏高分子量蛋白质组分^[5]。荞麦种子蛋白质质量分数(w)平均为 12.89%,不同种间差异很大:大粒组野生荞麦为 14.92%、大粒组栽培荞麦为 12.52%、小粒组荞麦为 10.39%^[3, 6-7]。按照蛋白溶解度的不同,荞麦中种子蛋白的成分可以分为清蛋白、球蛋白、醇溶蛋白和谷蛋白 4 类,其中清蛋白的含量最高,占荞麦种子蛋白总含量的 71.4%;球蛋白和谷蛋白次之,均占 9.1%;醇溶蛋白的含量最少,占荞麦种子蛋白总含量的 7.4%^[8]。

对荞麦种子蛋白相关编码基因进行研究,有助于从分子水平上了解种子蛋白合成机制。灌浆期是种子营养物质积累的主要时期,研究该时期相关基因的表达情况对于鉴定荞麦中种子蛋白基因具有重要意义。通过对荞麦灌浆期种子全长 cDNA 文库的筛选和分析,共获得了 1 106 个有效 EST 序列,这些序列主要参与抗病、储藏蛋白等^[9]。通过筛选基因组文库,获得了 17 个 13S 球蛋白基因^[10]。本课题组前期完成了荞麦种子的转录组学分析(SRA 登录号:DRP001013),从中鉴定到了 3 个表达水平较高的种子蛋白基因[13S 球蛋白(13S Globulin)基因、11S 球蛋白(11S Globulin)基因和谷蛋白(Glutelin)基因],

并获得这 3 个基因的全长、氨基酸序列和基因组序列。本研究通过生物信息学方法对这 3 个基因的结构、进化关系以及基因上游 1 500 bp 启动子进行了分析,并分析了 3 个基因在荞麦种子灌浆期的表达情况,为探讨荞麦种子发育过程分子机制提供理论依据,也为改善荞麦种子蛋白含量和组分提供了优异基因资源。

1 材料与方法

1.1 种子蛋白基因序列的获取和生物信息学分析

从本课题组测序的荞麦基因组数据中获得 3 个基因(13S 球蛋白基因、11S 球蛋白基因和谷蛋白基因)的编码序列、氨基酸序列和基因组序列,利用 NCBI 中 BLASTP 对其进行注释,利用在线 EXPASY (http://web.expasy.org/compute_pi/) 预测蛋白的等电点和相对分子质量;利用 GSGD2.0 (<http://gsds.cbi.pku.edu.cn/>) 预测基因结构;13S 球蛋白、11S 球蛋白和谷蛋白的同源序列从 NCBI 中获得,用 MEGA6 构建系统进化树。根据基因组注释文件的位置信息,利用本地 Perl 程序从基因组中调取基因上游 1 500 bp 启动子序列,用 PLANTCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 预测启动子区域的顺式作用元件。

1.2 材料种植和取样

选择甜荞 *F. esculentum*、苦荞 *F. tataricum* 和金荞 *F. cymosum* 代表品种各 1 份,即丰甜 1 号、米苦荞 104 和红心金荞,均种植于贵州师范大学荞麦产业技术研究中心试验基地,常规大田管理,分别于授粉后第 5 天(灌浆前期)、第 10 天(灌浆中期)和第 20 天(灌浆末期)取样,取回的样品在冰上剥壳,剔除差异较大的种子,严格控制种子的大小和时期,所有的样品保存于 -80 ℃ 冰箱备用。

1.3 实时荧光定量 PCR(qRT-PCR)

RNA 提取采用 Plus 植物总 RNA 提取试剂盒(天根生化科技有限公司, 北京), 具体操作按照说明书进行。用 Nanodrop 2000 微量紫外分光光度计检测 RNA 浓度及质量, 用 $10\text{ g}\cdot\text{L}^{-1}$ 的琼脂糖凝胶电泳检测 RNA 的完整性。按照 PrimeScript™ RT Reagent Kit (Perfect Real Time)(宝生物工程有限公司, 大连)说明书反转录合成 cDNA。

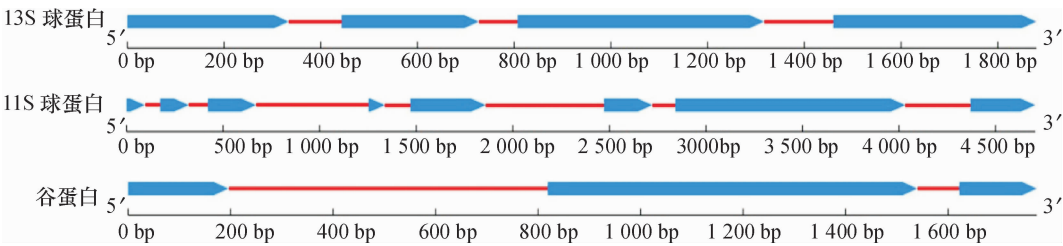
利用 Primer Premier 6.0 软件, 设计 qRT-PCR 引物。引物设计标准是: 扩增长度 50 ~ 250 bp, 温度 $(60\pm2)\text{ }^{\circ}\text{C}$, 引物长度 20 ~ 22 bp, GC 质量分数 45% ~ 55%^[11]。13S 球蛋白基因的上下游引物分别为 5'-ACGCCATAACCAGTCCGATTGC-3'、5'-AAAC-GCTTCCTCGCCTTCTCCT-3'; 11S 球蛋白基因的上下游引物分别为 5'-ACAGTGTGCTGGTGTTCCTTC-3'、5'-TGGTGTTTCATCCTGGAAGTCTGC-3'; 谷蛋白基因的上下游引物分别为 5'-GCATGGCATGGCTCTCAA-CTGT-3'、5'-GCACCAAGTCCAACCTCACCAA-3'。以荞麦 *Actin* 基因为内参(GenBank 登录号: HQ398855.1), 上下游引物分别为 5'-ACCTTGCTGGAC-GTGACCTTAC-3'、5'-CCATCAGGAAGCTCATAGTTC-3'。

qRT-PCR 的反应体系和程序按照 SYBR® Premix Ex Taq™(Tli RNaseH Plus, 宝生物工程有限公司, 大连)说明书进行, 相对表达量用 $2^{-\Delta\Delta C_t}$ 计算^[12], 数据统计和作图使用 Originlab 公司的 Origin8.0 进行(<http://www.originlab.com/>)。

2 结果与分析

2.1 种子蛋白基因的序列分析

2.1.1 13S 球蛋白基因 核酸序列长 1 548 bp。基因结构分析(图 1)表明, 该基因含有 4 个外显子(分别为 334、284、510、420 bp)和 3 个内含子(分别为 109、80、143 bp)。编码的蛋白序列全长 515 aa, 蛋白的等电点为 5.69, 相对分子质量为 58 330(表 1)。将其与 GenBank 中荞麦和其他物种中的 13S 球蛋白构建进化树, 结果(图 2)表明: 13S 球蛋白与苦荞中已报道的一个 13S 球蛋白(又名过敏蛋白, GenBank 登录号: ABI32184.1)的亲缘关系最近, 仅有 6 个氨基酸的不同(相似性 98%), 推测二者为同一个基因; 荞麦属甜荞和苦荞中 20 个 13S 球蛋白聚在一起, 它们的亲缘关系较其他物种中的 13S 球蛋白更近; 甜菜 *Beta vulgaris*、胡杨 *Populus euphratica* 和烟草 *Nicotiana sylvestris* 中的 13S 球蛋白聚为一类。



图中蓝色箭头表示外显子, 红色线条表示内含子。

图 1 3 个种子蛋白基因的结构

Fig.1 Structures of three seed protein genes

2.1.2 11S 球蛋白基因 核酸序列长 2 745 bp。基因结构分析(图 1)表明, 该基因含有 8 个外显子(分别为 95、148、249、85、390、251、1 191、336 bp)和 7 个内含子(分别为 79、100、582、132、612、119、337 bp)。编码的蛋白序列全长 914 aa, 蛋白的等电点为 5.87, 相对分子质量为 103 350(表 1)。将其与 GenBank 中其他物种中的 11S 球蛋白构建进化树, 结果(图 3)表明: 本研究中的 11S 球蛋白与欧洲慈姑 *Sagittaria sagittifolia*(GenBank 登录号: CAA70334.1)和猕猴桃 *Actinidia chinensi*(GenBank 登录号: ABB77213.1)中的 11S 球蛋白亲缘关系最近(相似性 45%)。

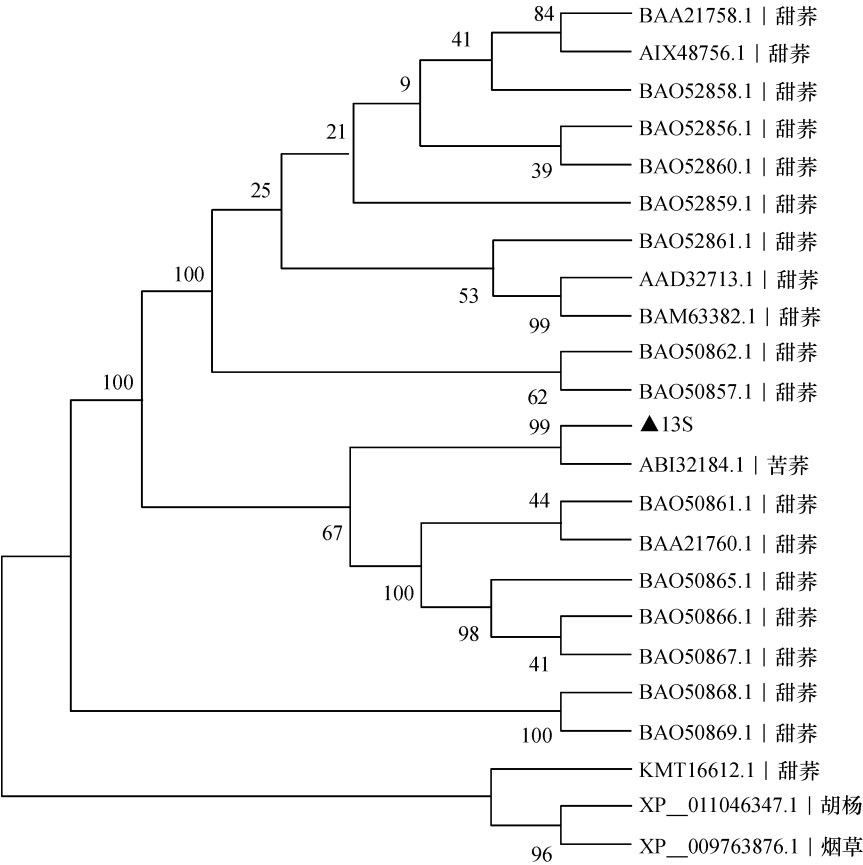
2.1.3 谷蛋白基因 核酸序列长 1 068 bp。基因结构分析(图 1)表明, 该基因含有 3 个外显子(分别为 196、721、151 bp)和 2 个内含子(分别为 623、82 bp)。编码的蛋白序列全长 355 aa, 蛋白的等电点为

6.45, 相对分子质量为 37 890(表 1)。将其与 GenBank 中其他物种中的谷蛋白构建进化树(图 4), 结果表明: 荞麦谷蛋白与甜菜 *Beta vulgaris* 中 B5 型谷蛋白亲缘关系最近(GenBank 登录号: XP_010683769.1, 相似性 81%); 荞麦、甜菜和拟南芥 *Arabidopsis lyrata* 3 个物种中的谷蛋白聚为一类, 其他物种中谷蛋白聚为一类。

表 1 3 个种子蛋白基因的基本信息统计

Tab.1 Information of three seed protein genes

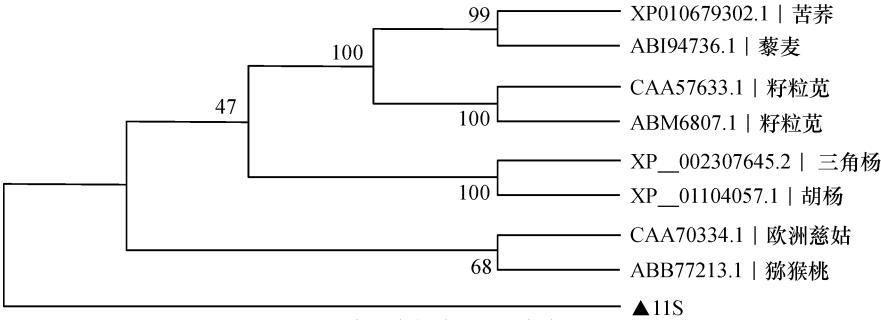
基因	核酸长度 /bp	蛋白长度 /aa	等电点	相对分子 质量
13S 球蛋白基因	1 548	515	5.69	58 330
11S 球蛋白基因	2 745	914	5.87	103 350
谷蛋白基因	1 068	3 55	6.45	37 890



黑色三角标注为本研究中的目的蛋白。

图 2 13S 球蛋白基因编码蛋白质的进化关系

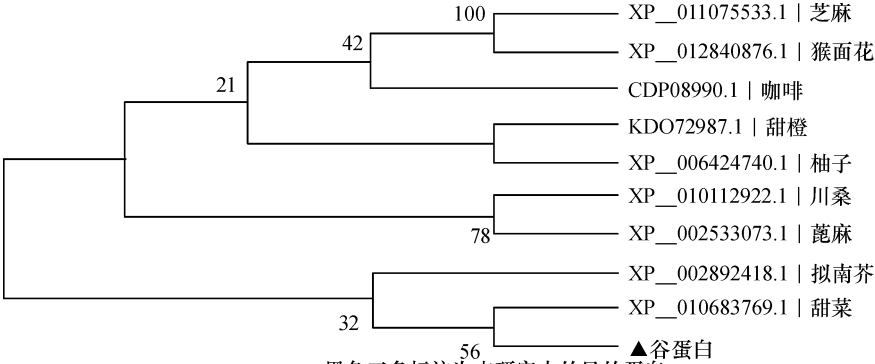
Fig. 2 Evolutionary relationship of the protein sequences encoded by 13S globulin gene



黑色三角标注为本研究中的目的蛋白。

图 3 11S 球蛋白基因编码蛋白质的进化关系

Fig. 3 Evolutionary relationship of the protein sequences encoded by 11S globulin gene



黑色三角标注为本研究中的目的蛋白。

图 4 谷蛋白基因编码蛋白质的进化关系

Fig. 4 Evolutionary relationship of the protein sequences encoded by glutelin gene

2.2 种子蛋白基因启动子序列的分析

2.2.1 13S 球蛋白基因启动子 13S 球蛋白基因启动子区域的顺式作用元件列于表 2。13S 球蛋白基因的 ATG 上游含有 21 个真核生物转录起始所必需的 TATA 框^[13]和 5 个控制转录频率的 CAAT 框^[14]; 在其 ATG 上游 1 500 bp 内含有 5 个响应生物和非生物胁迫的顺式作用元件, 即 4 个厌氧响应元件 (ARE)^[14]和 1 个真菌激发元件 (Box-W1)^[15], 表明该基因可能响应低氧环境和受真菌诱导; 13S 球蛋白基因启动子区含有 10 个响应激素的元件, 即 4 个脱

落酸响应元件 (3 个 ABRE 和 1 个 CE1)^[13, 16], 4 个茉莉酸甲酯响应元件 (2 个 CGTCA-motif 和 2 个 TGACG-motif)^[17], 2 个水杨酸响应元件 (TCA-element)^[18], 表明该基因可能对脱落酸、茉莉酸甲酯和水杨酸 3 类激素有响应; 13S 球蛋白基因启动子区含有一个 MYB 转录因子的结合位点 (MBS)^[19], 表明 MYB 转录因子可能能够结合在该基因的启动子区域, 诱导它的表达; 此外, 还含有 6 个胚乳表达所需要的顺式作用元件 (Skn-1_motif)^[20], 表明该基因可能在胚乳中优势表达。

表 2 13S 球蛋白基因和 11S 球蛋白基因的启动子分析
Tab.2 Promoter analysis of the 13S globulin gene and 11S globulin gene

基因	顺式作用元件	数量	序列	功能
13S 球蛋白基因	TATA-box	21	TATAAA	转录起始必需元件 ^[13]
	CAAT-box	5	CAAT	转录增强元件 ^[14]
	ARE	4	TGGTTT	厌氧诱导元件 ^[14]
	Box-W1	1	TTGACC	响应真菌诱导 ^[15]
	ABRE	3	TACGTG	响应脱落酸 ^[13]
	CE1	1	TGCCACCGG	响应脱落酸 ^[16]
	CGTCA-motif	2	CGTCA	响应茉莉酸甲酯 ^[17]
	TGACG-motif	2	TGACG	响应茉莉酸甲酯 ^[17]
	TCA-element	2	GAGAAGAATA	响应水杨酸 ^[18]
	MBS	1	CGGTCA	MYB 结合位点 ^[19]
	Skn-1_motif	6	GTCAT	胚乳表达调控元件 ^[20]
11S 球蛋白基因	TATA-box	2	TATAA	转录起始必需元件 ^[13]
	CAAT-box	8	CCAAT	转录增强元件 ^[14]
	WUN-motif	1	TCATTACGAA	伤害响应元件 ^[18]
	HSE	1	CNNGAANN TTCNNG	高温响应元件 ^[21]
	LTR	2	CCGAAA	低温响应元件 ^[22]
	TC-rich repeats	1	ATTTTCTTCA	响应防御和胁迫 ^[23]
	ABRE	1	TACGTG	响应脱落酸 ^[13]
	CGTCA-motif	1	CGTCA	响应茉莉酸甲酯 ^[17]
	TGACG-motif	1	TGACG	响应茉莉酸甲酯 ^[17]
	TCA-element	2	GAGAAGAATA	响应水杨酸 ^[18]
	TGA-element	2	AACGAC	响应生长素 ^[18]
	Skn-1_motif	3	GTCAT	胚乳表达调控元件 ^[20]

2.2.2 11S 球蛋白基因启动子 11S 球蛋白基因启动子区域的顺式作用元件列于表 2。11S 球蛋白基因的 ATG 上游含有 2 个真核生物转录起始所必需的 TATA 框^[13]和 8 个控制转录频率的 CAAT 框^[14]; 在其 ATG 上游 1 500 bp 内含有 5 个响应非生物胁迫的顺式作用元件, 即 1 个伤害响应元件 (WUN-motif)^[18]、1 个热胁迫响应元件 (HSE)^[21]、2 个低温响应元件 (LTR)^[22]和 1 个防御和胁迫响应元件 (TC-rich repeat)^[23], 表明该基因可能对伤害、高温、低温

等非生物胁迫有响应; 11S 球蛋白基因启动子区含有 7 个响应激素的元件, 即 1 个脱落酸响应元件 (ABRE)^[13]、2 个茉莉酸甲酯响应元件 (1 个 CGTCA-motif 和 1 个 TGACG-motif)^[17]、2 个水杨酸响应元件 (TCA-element)^[18]和 2 个生长素响应元件 (TGA-element)^[18], 表明该基因可能对脱落酸、茉莉酸甲酯、水杨酸和生长素 4 类激素有响应; 此外, 还含有 3 个胚乳表达所需要的顺式作用元件 (Skn-1 _motif)^[20], 表明该基因可能在胚乳中表达。

2.2.3 谷蛋白基因启动子 谷蛋白基因启动子区域的顺式作用元件列于表 3。谷蛋白基因的 ATG 上游含有 38 个真核生物转录起始所必需的 TATA 框^[13]和 10 个控制转录频率的 CAAT 框^[14]；在其 ATG 上游1 500 bp 内含有 3 个响应非生物胁迫的顺式作用元件,即 1 个热胁迫响应元件 HSE^[21]和 2 个防御和胁迫响应元件(TC-rich repeat)^[23],表明该基因可能对高温等非生物学胁迫有响应;11S 球蛋白基因启动子区含有 8 个响应激素的元件,即 2 个脱落酸响应元件(ABRE)^[13]、2 个茉莉酸甲酯响应元件(1 个 CGTCA-motif 和 1 个 TGACG-motif)^[17]、2 个乙

烯响应元件(ERE)^[24]和 1 个赤霉素响应元件(GARE-motif)^[18], 1 个水杨酸响应元件(TCA-element)^[18],表明该基因可能对脱落酸、茉莉酸甲酯、乙烯、赤霉素和水杨酸等 5 类激素有响应;谷蛋白基因启动子区含有 1 个参与黄酮合成的 MYB 转录因子结合位点(MBSII)^[25],表明这一类的 MYB 转录因子可能能够结合到该基因启动子区域,诱导它表达。此外,还含有 2 个胚乳表达所需的顺式作用元件(1 个 Skn-1_motif 和 1 个 GCN4_motif)^[20, 26],表明该基因可能在胚乳中表达。

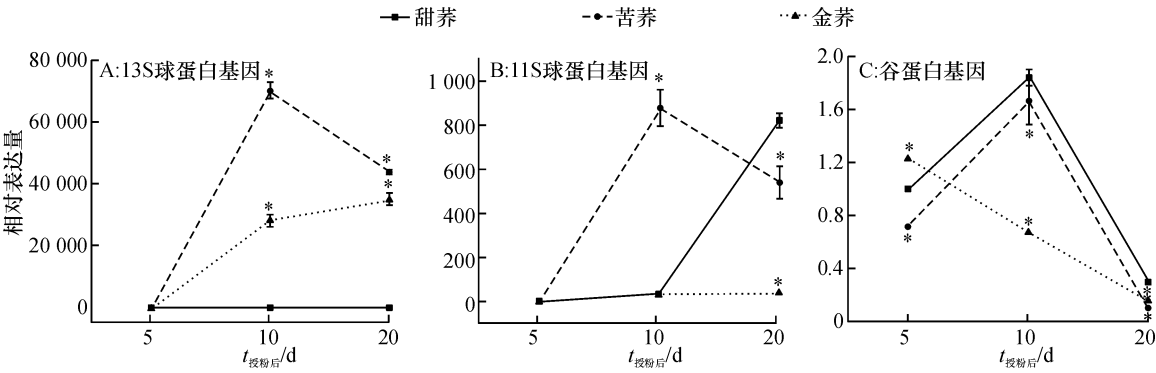
表 3 谷蛋白基因的启动子分析
Tab.3 Promoter analysis of the glutelin gene

顺式作用元件	数量	序列	功能
TATA-box	38	TATAA	转录起始必需元件 ^[13]
CAAT-box	10	CAAAT	转录增强元件 ^[14]
TC-rich repeats	2	ATTTTCTCCA	响应防御和胁迫 ^[23]
HSE	1	AAAAAATTTC	高温响应元件 ^[21]
ABRE	2	GACACGTGGC	响应脱落酸 ^[13]
CGTCA-motif	1	CGTCA	响应茉莉酸甲酯 ^[17]
TGACG-motif	1	TGACG	响应茉莉酸甲酯 ^[17]
ERE	2	ATTTCAAA	响应乙烯 ^[24]
GARE-motif	1	TCTGTTG	响应赤霉素 ^[18]
TCA-element	1	CCATCTTTTT	响应水杨酸 ^[18]
MBSII	1	AAAAGTTAGTTA	参与黄酮合成的 MYB 结合位点 ^[25]
GCN4_motif	1	TGTGTCA	胚乳表达调控元件 ^[26]
Skn-1_motif	1	GTCAT	胚乳表达调控元件 ^[20]

2.3 种子蛋白基因在荞麦种子发育过程中的表达

3 个种子蛋白基因在甜荞、苦荞和金荞种子不同发育时期的表达情况如图 5 所示: 1) 13S 球蛋白基因在苦荞中的表达量最高,在金荞中次之,在甜荞中最低;在甜荞种子发育过程中,该基因的表达没有明

显的变化,表明该基因在这一过程不起主要作用;在苦荞种子发育过程中,该基因在灌浆前期表达量较低,然后逐渐升高,到灌浆中期表达量最高,随后逐渐降低,表现出低-高-次高的规律;在金荞种子发育过程中,该基因的表达随着种子灌浆过程逐渐升



* 表示相同时间与甜荞种子相比,差异达到极显著水平(P<0.01,Bonferroni 法)。

图 5 种子蛋白基因在荞麦种子发育过程中的表达

Fig.5 Expression patterns of three seed protein genes in the process of seed development

高,表现出递增的规律(图5A)。2)11S球蛋白基因在苦荞中表达量最高,在甜荞中次之,在金荞中最低;在甜荞种子发育过程中,该基因在灌浆前期和中期无明显表达,在灌浆末期大量表达,表现出低-低-高的规律;在苦荞种子发育过程中,该基因在灌浆前期低丰度表达,然后逐渐升高,到灌浆中期表达量最高,随后逐渐降低,表现出低-高-次高的规律;在金荞种子发育过程中,该基因的表达没有明显的变化,表明该基因在这一过程中不起主要作用(图5B)。3)谷蛋白基因在甜荞中的表达量最高,在苦荞中次之,在金荞中最低;在甜荞种子发育过程中,该基因在灌浆前期有一定的表达,随灌浆进程表达量逐渐升高,在灌浆中期最高,随后快速下降,在灌浆末期表达量最低,表现出中-高-低的规律;在苦荞种子发育过程中,该基因的表达与在甜荞种子发育过程中一致,也表现出中-高-低的规律;在金荞种子发育过程中,该基因在灌浆前期的表达量最高,随后逐渐降低,表现出递减的规律(图5C)。

3 讨论与结论

蛋白质是荞麦种子中的重要营养物质之一,按照蛋白溶解度的不同,荞麦中种子蛋白的成分可以分为清蛋白、球蛋白、醇溶蛋白和谷蛋白4类。目前荞麦中球蛋白的编码基因已经鉴定了19个,均为13S球蛋白基因(18个来自甜荞、1个来自苦荞)^[10,27]。本研究获得的球蛋白基因与苦荞中已报道的一个球蛋白基因(GenBank登录号:ABI32184.1)十分相似,在氨基酸水平上仅有6个氨基酸的不同,推测二者为同一个基因;将ABI32184.1在细菌中表达,重组蛋白的大小为65 400,通过EXPASY对本研究中获得13S球蛋白进行预测,其相对分子质量为58 330,二者差异较小,说明通过EXPASY预测的蛋白相对分子质量比较可靠,具有一定的参考价值^[27]。荞麦中11S球蛋白、清蛋白、醇溶蛋白和谷蛋白的编码基因目前鲜见报道,本研究获得了1个11S球蛋白和1个谷蛋白的基因序列,这为后续进一步研究基因功能奠定了基础。

通过对3个种子蛋白基因上游启动子序列的分析,挖掘了许多响应胁迫和激素及胚乳表达所必需的顺式作用元件。植物的生长发育除了遗传调控外,还会受到环境的干扰,恶劣环境会对植物的生长发育产生不良影响,如高温、冻害、损伤、缺氧、病原菌侵染等^[28]。植物为了避免环境因素导致的不良影响,会诱导相关基因的大量表达,来抵制或消除这些不良环境造成的弊端,维持植物正常的生长发育^[29]。

3个种子蛋白基因的启动子区域都有响应胁迫的顺式作用元件,表明这3个基因可能在荞麦种子灌浆过程中参与对对各种胁迫的响应中,从而维持种子正常的生长发育过程。植物激素的作用主要有2个:一个是响应植物的逆境胁迫;另一个是植物某一生长发育过程所必需^[30]。茉莉酸甲酯主要参与植物的抗逆和衰老过程,水杨酸主要参与植物的抗逆过程,脱落酸主要参与植物的干旱胁迫和种子萌发过程,生长素主要促进植物的生长发育,乙烯促进植物的衰老和成熟,赤霉素则促进植物的生长发育及开花过程^[31]。在本研究3个种子蛋白基因的启动子中发现了这几类激素的响应元件,表明这些基因可能是这些激素作用的下游信号途径中的一员,与荞麦种子发育过程中对逆境的响应、营养物质的积累和种子成熟有关。水稻种子中的蛋白质主要积累在胚乳外周区域^[32],本研究的3个种子蛋白基因的启动子区域都含有胚乳表达所需的顺式作用元件(Skn-1_motif和GCN4_motif),这表明荞麦种子中蛋白质的积累部位可能和水稻相似,也在胚乳中。

在甜荞、苦荞和金荞种子发育过程中,这3个基因表达较高的时期均集中在灌浆中期,说明灌浆中期是种子蛋白积累的关键时期,这一发现与水稻中的结论一致^[32]。3个基因在不同材料中的表达水平不一致,13S球蛋白基因在苦荞和金荞中表达量较高,11S球蛋白基因在甜荞和苦荞中表达量较高,谷蛋白基因则在3个材料中均有较高水平的表达,这说明不同材料中相关蛋白的编码基因并不是一致的,这些基因可能存在一定的冗余,当一个编码基因不表达时,则由其同源基因来代替它完成功能,从而完成种子的发育过程。

参考文献:

- [1] 林汝法,周小理,任贵兴,等. 中国荞麦的生产与贸易、营养与食品[J]. 食品科学, 2005,26(1): 259-263.
- [2] 向达兵,彭镰心,赵钢,等. 荞麦栽培研究进展[J]. 作物杂志, 2013(3): 1-6.
- [3] GUO X, YAO H. Fractionation and characterization of tartary buckwheat flour proteins[J]. Food Chem, 2006, 98(1): 90-94.
- [4] 清源. 苦荞的价值及开发利用现状[J]. 南方农业, 2014,18(27): 131-133.
- [5] 朱新产,王宝维,魏益民. 荞麦种子蛋白组分差异研究[J]. 种子, 2000(6): 9-10.
- [6] 郭玉珍,陈庆富,杨兰英,等. 栽培及野生荞麦资源的种子蛋白质含量研究[C]//林汝法,池田清和. 苦荞产业经济国际论坛论文集. 太原:山西省食品科学学会, 2006.

- [7] 王燕,陈庆富. 甜荞、苦荞和大野荞的高分子量种子蛋白亚基研究[J]. 安徽农业科学, 2010,38(33): 18678-18680.
- [8] 刘拥海,俞乐,肖迪. 荞麦种子蛋白质组分分析[J]. 种子, 2006,25(12): 31-33.
- [9] 贾鹏. 苦荞种子灌浆期 cDNA 文库的筛选与分析[D]. 杨凌:西北农林科技大学, 2012.
- [10] SANO M, NAKAGAWA M, OISHI A, et al. Diversification of 13S globulins, allergenic seed storage proteins of common buckwheat[J]. Food Chem, 2014, 155(11): 192-198.
- [11] ARTICO S, NARDELI S M, BRILHANTE O, et al. Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data[J]. BMC Plant Biol, 2010, 10(1): 1-12.
- [12] LIVAK K J, SCHMITTGEN T D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method[J]. Methods, 2001, 25(4): 402-408.
- [13] YAMAGUCHI-SHINOZAKI K, SHINOZAKI K. *Arabidopsis* DNA encoding two desiccation-responsive *rd29* genes[J]. Plant Physiol, 1993, 101(3): 1119-1120.
- [14] VAUTERIN M, FRANKARD V, JACOBS M. The *Arabidopsis thaliana dhps* gene encoding dihydrodipicolinate synthase, key enzyme of lysine biosynthesis, is expressed in a cell-specific manner[J]. Plant Mol Biol, 1999, 39(4): 695-708.
- [15] RUSHTON P J, TORRES J T, PARNISKE M, et al. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley *PR1* genes[J]. EMBO J, 1996, 15(20): 5690-5700.
- [16] SHEN Q, UKNES S J, HO T H. Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene[J]. J Biol Chem, 1993, 268(31): 23652-23660.
- [17] ROUSTER J, LEAH R, MUNDY J, et al. Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain[J]. Plant J, 1997, 11(3): 513-523.
- [18] PASTUGLIA M, ROBY D, DUMAS C, et al. Rapid induction by wounding and bacterial infection of an *S* gene family receptor-like kinase gene in *Brassica oleracea*[J]. Plant Cell, 1997, 9(1): 49-60.
- [19] NASH J, LUEHRSEN K R, WALBOT V. *Bronze-2* gene of maize: Reconstruction of a wild-type allele and analysis of transcription and splicing[J]. Plant Cell, 1990, 2(11): 1039-1049.
- [20] ADACHI M, KANAMORI J, MASUDA T, et al. Crystal structure of soybean 11S globulin: glycinin A3B4 homo-hexamer[J]. Proc Natl Acad Sci USA, 2003, 100(12): 7395-7400.
- [21] SUN S W, CHUNG M C, LIN T Y. The structure and expression of an *hsc70* gene from *Lycopersicon esculentum*[J]. Gene, 1996, 170(2): 237-241.
- [22] WHITE A J, DUNN M A, BROWN K, et al. Comparative analysis of genomic sequence and expression of a lipid transfer protein gene family in winter barley[J]. J Exp Bot, 1994, 45(12): 1885-1892.
- [23] DIAZ-DE-LEON F, KLOTZ K L, LAGRIMINI L M. Nucleotide sequence of the tobacco (*Nicotiana tabacum*) anionic peroxidase gene[J]. Plant Physiol, 1993, 101(3): 1117-1118.
- [24] ITZHAKI H, WOODSON W R. Characterization of an ethylene-responsive glutathione S-transferase gene cluster in carnation[J]. Plant Mol Biol, 1993, 22(1): 43-58.
- [25] KOES R E, SPELT C E, VAN DEN ELZEN P J, et al. Cloning and molecular characterization of the chalcone synthase multigene family of *Petunia hybrida*[J]. Gene, 1989, 81(2): 245-257.
- [26] KIM S Y, WU R. Multiple protein factors bind to a rice glutelin promoter region[J]. Nucleic Acids Res, 1990, 18(23): 6845-6852.
- [27] SHIGEMORI S, YONEKURA S, SATO T, et al. Expression of the immunoreactive buckwheat major allergenic storage protein in *Lactococcus lactis*[J]. Appl Microbiol Biotechnol, 2013, 97(8): 3603-3611.
- [28] 徐慧妮,王秀峰. 植物生长发育逆境胁迫及其研究方法[C]//中国园艺学会. 中国园艺学会第七届青年学术讨论会论文集. 北京:中国园艺学会, 2006:887-893.
- [29] 刘辉,李德军,邓治. 植物应答低温胁迫的转录调控网络研究进展[J]. 中国农业科学, 2014, 47(18): 3523-3533.
- [30] 熊国胜,李家洋,王永红. 植物激素调控研究进展[J]. 科学通报, 2009,54(18): 2718-2733.
- [31] 王丽萍,李志刚,谭乐和,等. 植物内源激素研究进展[J]. 安徽农业科学, 2011,39(4): 1912-1914.
- [32] 兰莹. 水稻心白突变体 *flo2* 灌浆期籽粒中 miRNA 鉴定与表达分析[D]. 南京:南京农业大学, 2010.

【责任编辑 周志红】