

玉米脱水素基因家族的鉴定与分析

赵 阳, 王 玉, 蔡慧林, 程备久, 马 庆*

(安徽农业大学生命科学学院, 安徽省作物生物学重点实验室, 合肥 230036)

摘 要: 脱水素(dehydrin, DHN)属于 LEA 蛋白第二家族成员, 是一种植物中广泛存在的亲水性蛋白, 在干旱、低温和高盐等非生物胁迫的环境中起到重要作用。通过生物信息学方法对玉米全基因组 DHN 家族成员进行了鉴定, 并进一步对其系统发育、基因结构、染色体定位和基因复制以及表达模式进行了系统分析。结果显示, 在玉米 DHN 基因家族中共有 5 个家族成员, 多物种系统发育进化树分析、基因结构以及基序分析都表明 DHN 家族成员在进化上具有高度的保守性。基因复制和物种间微共线性分析表明, 在 5 个玉米 DHN 基因中存在着 1 对片段复制基因 (*ZmDHN1-ZmDHN2*), 玉米、高粱和水稻 3 个物种间存在 2 对直系同源基因 (*ZmDHN2-Sb04g032250.1*, *ZmDHN2-Os02g44870.1*)。通过转录组表达数据分析表明, 玉米 DHN 家族基因在不同发育时期具有不同的组织表达模式; 同时, 诱导表达模式分析表明 *ZmDHN* 基因的表达受到盐和干旱胁迫的显著诱导。该研究结果将为进一步鉴定玉米 DHN 家族重要的基因成员并对其开展功能分析奠定基础。

关键词: 玉米; 脱水素; 进化分析; 基因复制; 表达模式

中图分类号: S513

文献标识码: A

文章编号: 1672-352X (2015)05-0657-09

Identification and characterization of the dehydrin gene family in maize

ZHAO Yang, WANG Yu, CAI Huilin, CHENG Beijiu, MA Qing

(Key Laboratory of Crop Biology of Anhui Province, School of Life Science, Anhui Agricultural University, Hefei 230036)

Abstract: Dehydrins (DHN) is a hydrophilic protein widely distributed in plants. DHNs belong to class II in LEA protein family and have important roles in response to abiotic stresses, such as cold, drought and salt, etc. In this study, we performed a bioinformatics analysis to identify the members of the DHN family in maize. Subsequently, we carried out a systematic analysis of these genes, including phylogenetic relationship, gene structure, chromosome location, gene duplication and expression patterns. The results showed that a total of 5 DHN members were identified in maize, and the phylogenetic relationship, gene structure and motif analysis indicated that genes in the DHN family are highly conserved. Gene duplication and interspecies microsynteny analysis demonstrated that one segmental duplication pair (*ZmDHN1-ZmDHN2*) existed in maize DHN family, while two orthologous gene pairs (*ZmDHN2-Sb04g032250.1* and *ZmDHN2-Os02g44870.1*) were detected among maize, rice and sorghum. Analysis of transcriptional data indicated that the DHN genes had different expression patterns at different developmental stages. In addition, the expression of five *ZmDHN* genes was significantly induced under salt and drought stresses. This study provides a basis for the selection of important DHN genes for further functional analysis in maize.

Key words: maize; dehydrin; phylogenetic analysis; gene duplication; expression pattern

胚胎发育晚期丰富蛋白(late embryogenesis proteins abundant, LEA 蛋白)是生物体中广泛存在的一类与渗透调节有关的家族蛋白。研究表明, 该类蛋白在植物遭受低温、干旱和高盐等非生物胁迫

的过程中发挥着重要作用, 维持着植物细胞的正常代谢活动。LEA 蛋白首先发现于胚胎发育后期的棉花子叶中^[1], 随后, 在小麦^[2]、大豆^[3]、玉米^[4]等多种高等植物中均被发现。根据对这些已报道蛋白的

收稿日期: 2015-05-29

基金项目: 国家自然科学基金(31201217)和安徽省自然科学基金(1508085QC64)共同资助。

作者简介: 赵 阳, 博士, 讲师。E-mail: zhaoyang0521@163.com

* 通信作者: 马 庆, 副教授。E-mail: mqqmmq@126.com

氨基酸序列同源性以及蛋白的保守结构域等一些重要特征, 可将 LEA 蛋白进一步分为 6 个家族^[5-7]。

脱水素 (dehydrin, DHN) 属于 LEA 蛋白第二家族成员, 是一种植物中广泛存在的亲水性蛋白^[8], 一般由 82~575 个氨基酸组成, 不同成员分子量差异较大, 分布于 9~200 kDa^[9]。DHN 的一个重要结构特征是具有 3 个高度保守区域: K、S 和 Y 片段^[10]。K 片段一般位于蛋白序列的 C 端, 由 15 个氨基酸(EKKGIMDKIKEKLP)组成, 富含赖氨酸, 所形成的双亲水 α -螺旋是其亲水性的重要结构基础^[11]。S 片段是由 4 到 10 个丝氨酸残基组成, 相关证据表明 DHN 在信号肽的引导下进入细胞核可能是 S 片段的磷酸化所致^[12]。Y 片段则位于蛋白的 N 末端, 其保守氨基酸序列为 V/TDE/QYGNP, 该片段与一些细菌和植物分子伴侣的核酸结合位点有同源性^[9]。根据 K、S 和 Y 片段的分布数量可以将 DHN 分为 5 类: Kn、SKn、KnS、YnKn 和 YnSKn^[13]。除了以上 3 个保守区域外, 有的 DHN 还包含一些保守性稍差的区域, 这些区域富含极性氨基酸和甘氨酸, 称作 Ψ 片段, 它们通常分布于 K 片段之间^[7]。

大量研究表明, DHN 基因能在植物胚胎发育后期及处于干旱、低温、盐碱等非生物逆境的组织中大量表达^[10], 具有稳定细胞膜、清除自由基、结合金属离子以及分子伴侣等重要功能^[14]。例如: Puhakainen 等^[15]证明在拟南芥中过表达多个脱水素基因, 其植株抗冻能力获得显著提高。Brini 等^[16]报道小麦脱水素 DHN5 蛋白在体外条件下, 能够增强葡糖氧化酶(GOD/POD)和 β -葡糖苷酶(bglG)这 2 种酶的活性和热稳定性, 表明 DHN5 对酶在体外遭受热害时具有保护作用, 发挥了分子伴侣的功能。在抵抗重金属胁迫上, 芥菜中 2 个脱水素基因 *BjDHN2/BjDHN3* 被证明受重金属诱导, 并且在重金属胁迫下, 其转基因烟草丙二醛含量和电解质外渗量比野生型要低, 表明脱水素可以降低重金属胁迫造成的膜脂过氧化, 以此保护生物膜结构稳定^[17]。Yin 等^[18]研究发现将马铃薯的 *DHN24* 基因转入黄瓜中能显著提高黄瓜对低温的忍耐能力。Saavedra 等^[19]证明盐处理或高渗处理后, 小立碗蕨的脱水素突变体无法恢复生长, 而野生型却可以恢复生长。RoyChoudhury 等^[20]报道将水稻脱水素基因 *Rab16A* 导入烟草中, 转基因植株的抗旱和抗盐能力与对照组相比都明显获得提高。

玉米是一种重要的粮食作物, 在工农业生产中占有重要的地位。但玉米的生长和发育极易受到各

种环境胁迫的影响, 对其产量和品质造成了严重的影响。目前, DHN 家族在拟南芥和水稻等植物中都被报道^[21-22], 而对玉米 *DHN* 基因家族基因的报道, 多为单基因的研究, 在全基因组水平上对玉米 *DHN* 基因的结构、进化特征、表达模式等开展系统的研究尚未有报道。本研究利用生物信息学方法, 对玉米全基因组序列中的 DHN 家族的基因进行鉴定和分析, 为进一步鉴定和研究该家族成员的生物学功能奠定重要的基础。

1 材料与方法

1.1 玉米 *DHN* 基因家族成员的鉴定

首先通过网站 (<http://www.maizesequence.org/index.html>) 下载玉米 (*Zea mays* L.) 自交系 B73 的基因组、蛋白及 cDNA 序列数据。然后利用 DNATOOLS 软件对这些数据分别构建本地数据库。利用 Pfam 数据库 (<http://pfam.sanger.ac.uk/>)^[23] 中获得的 DHN 保守结构域 (PF00257) HMM 模型与已经构建好的本地数据库进行 BlastP 搜索, 初步筛选出含有 DHN 典型结构域的蛋白序列。将候选 DHN 蛋白基因序列通过 MEGA4.0 软件中提供的 Clustal W 工具进行多序列比对^[24-25], 手动删除重复基因并利用 Pfam (<http://pfam.sanger.ac.uk/search>) 和 SMART (<http://smart.embl-heidelberg.de/>)^[26] 网站进行辅助验证, 确定所获得序列均含有 DHN 保守结构域, 最终确定 *DHN* 基因家族成员的数量。利用 ExPASy (<http://www.expasy.org/tools/>)^[27] 在线分析工具对所获得 DHN 蛋白序列的分子量和等电点等进行预测。

1.2 系统进化树的构建

对所有玉米 *DHN* (*ZmDHN*, *Zea mays* Dehydrin) 氨基酸序列运用 Clustal W 工具进行多序列比对, 然后利用 MEGA4.0 软件对完成比对的序列进行系统进化树构建^[24-25], 构建方法采用邻接法 (Neighbor-Joining method, NJ), bootstrap 值设为 1000, 以此分析玉米全基因组中 *DHN* 基因的进化关系。同时, 为了进一步分析 *DHN* 基因在不同物种中的系统进化关系, 用上述同样的方法构建了玉米、拟南芥和水稻的 *DHN* 复合进化树。

1.3 *ZmDHN* 基因结构及染色体定位分析

将 *ZmDHN* 基因的 CDS 序列与相应基因序列分别导入 Gene structure display server (GSDS) (<http://gsds.cbi.pku.edu.cn/>)^[28] 在线分析工具进行序列比对分析, 获得玉米所有 *DHN* 基因的内含子分布信息以及 *DHN* 基因中内含子/外显子的边界。另外, 采用在线基序分析工具 MEME (Multiple expectation

maximization for motif elicitation) (http://meme.sdsc.edu/meme4_3_0/intro.html)^[29]对玉米 *DHN* 基因进行序列的保守基序分析。通过玉米基因组测序数据库获得基因在染色体上的物理位置, 然后使用 MapInspect 软件 (http://www.plantbreeding.wur.nl/uk/software_mapinspect.html)进行染色体定位分析。

1.4 基因复制和 *Ka/Ks* 分析

基因复制通常分为串联复制和片段复制。根据相关研究的定义, 串联复制基因被定义为两个旁系同源基因在染色体上被间隔的基因少于或等于 5 个, 大多以基因簇形式存在。片段复制基因是指位于同一染色体复制区域内的旁系同源基因^[30]。在复制基因鉴定后, 进一步对 *Ks* 值以及 *Ka/Ks* 比值进行计算, 分析其选择压力及其选择模式。复制基因对的 CDS 序列使用 MEGA4.0 软件进行比对, 去掉 gap 后再将比对结果导入 DnaSP v5.0 软件中进行 *Ks* 值以及 *Ka/Ks* 比值分析^[31]。滑动窗口分析 (sliding window analysis) 则采用以下参数: 窗口大小, 150 bp; 步移大小, 9 bp。对于复制事件时间估算, 采用公式 $T = Ks/2\lambda \times 10^{-6}$ Mya 进行计算, 将这些数值转换为以百万年 (Mya) 为单位的分歧时间 (T), 其中 $\lambda=6.5 \times 10^{-9}$ 代表草本植物每年每个位点的替换速率^[30,32]。

1.5 物种间的微共线性分析

使用 MCScanX 方法^[33]筛选玉米与高粱, 玉米与水稻之间的直系同源基因对。然后使用 Circos(<http://circos.ca/>)^[34]方法绘制直系同源基因之间的关系图, 其中基因位于染色体上的位置采用 Perl 软件实现。

1.6 *ZmDHN* 转录组表达谱分析

为了探究 *ZmDHN* 基因在各个组织中的表达水平, 玉米 B73 各个发育时期的转录组数据被用来进行基因的组织表达模式分析, 转录组数据共包含 18 个组织^[35]。用 *ZmDHN* 基因序号查找对应组织的表达 FPKM (fragments per kilobase of exon per million fragments mapped) 数值, 然后通过公式 $\log_2(\text{FPKM}+1)$ 将表达数值进行转换, 再用 R 语言和 Bioconductor 软件 (<http://www.bioconductor.org/>) 将表达数值转换为热图^[36]。

1.7 *ZmDHN* 基因的胁迫诱导表达模式分析

玉米 B73 自交系种植于人工气候室内, 温度设定为 $(28 \pm 2)^\circ\text{C}$, 光照条件为 14 h 光照/10 h 黑暗。长至 3 叶期时开始进行胁迫处理。盐胁迫和干旱胁迫处理分别采用 $200 \text{ mmol}\cdot\text{L}^{-1}$ NaCl 和 20% (w/v) PEG-6000 浇灌玉米根部, 处理时间为 0 h、3 h、6 h

和 12 h。将不同处理时间的叶片用液氮速冻, 然后保存于 -80°C 冰箱内, 待提取 RNA, 每一处理阶段设置 3 次重复。

采用 Trizol 法 (Invitrogen, Carlsbad, USA) 提取玉米叶片的总 RNA, 然后用 DNase I (Invitrogen) 消化去除残留的 DNA, 防止基因组污染。用 1% 的琼脂糖凝胶电泳验证提取的 RNA 的质量和含量, 再使用 NanoDrop ND-1000 (NanoDrop, Wilmington, USA) 分光光度计测定其浓度。根据反转录试剂盒 (Promega, Madison, USA) 操作说明将 $1 \mu\text{g}$ 的 RNA 反转录成第一链 cDNA。实时荧光定量 PCR 使用 ABI 7300 Real-Time (Applied Biosystems, USA) 系统完成。基因特异性引物序列设计使用 Primer Express 3.0 (Applied Biosystems) 软件。PCR 反应总体积为 $25 \mu\text{L}$, 包含 $2.0 \mu\text{L}$ 的 cDNA, $400 \text{ nmol}\cdot\text{L}^{-1}$ 的引物和 $12.5 \mu\text{L}$ 的 SYBR Green Master Mix reagent (Applied Biosystems), PCR 反应参数如下: 95°C 预变性 10 min; 95°C 15 s, 60°C 1 min, 共 40 个循环。反应结束后, 获得产物的溶解曲线。采用 $2^{-\Delta\Delta\text{CT}}$ 对获得的信号和数据进行处理^[37]。每个基因做 3 次技术重复。

2 结果与分析

2.1 玉米基因组中 *DHN* 基因的鉴定

通过 Blast 同源性搜索, 从玉米 B73 全基因组中鉴定到 5 个 *DHN* 基因家族成员 (表 1), 利用 Pfam 和 SMART 工具进行蛋白质结构分析, 这 5 个玉米 *DHN* 蛋白均含有 *DHN* 特征性结构域。根据染色体位置顺序将这些玉米 *DHN* 基因分别命名为 *ZmDHN1-5*, 其中 *ZmDHN1*、-2、-4 和 -5 基因编码的蛋白拥有 2 个 *DHN* 结构域。将 5 个玉米 *DHN* 蛋白氨基酸序列递交到 ExPasy 在线分析系统, 对其等电点、分子量等蛋白质生化属性进行统计分析 (表 1)。结果表明, 玉米 *DHN* 基因编码的蛋白的长度从 168~326 个氨基酸不等, 等电点范围从 5.51~8.78。

2.2 *ZmDHN* 家族基因结构及系统进化分析

为了进一步分析玉米 *DHN* 基因的差异和亲缘关系, 利用 MEGA4.0 软件通过邻接法进行系统进化树的构建 (图 1A), 共获得了 2 个姊妹对, 分别为 *ZmDHN1/2* 和 *ZmDHN4/5*, 其进化枝自展值高达 100%, 表明具有极高的同源性。为了进一步证明系统进化树构建的可靠性, 我们利用 GSDBS (<http://gsds.cbi.pku.edu.cn/>) 软件对玉米 5 个 *DHN* 基因的结构进行了比较和分析, 通过将玉米 *DHN* 基因的 CDS 编码序列和基因组序列进行比对, 获得了

玉米 *DHN* 基因的结构示意图(图 1B), 从图中我们可以看出 *DHN* 基因主要可以分为 3 个部分: 内含子、外显子和 UTR 区。每个 *DHN* 基因的结构较为类似, 均含有 2 个外显子和 1 个内含子。此外, 我们对内含子插入相位也进行了分析。一个内含子在基因中插入的区段存在 3 种相位: 相位 0(剪切发生在密码子之间)、相位 1(剪切发生在密码子的第 1 个核酸之后)和相位 2(剪切发生在密码子的第 2 个核酸之后)。如图 1B 所示, 所分析的 5 个玉米 *DHN* 基因内含子相位均为 0, 这些结果说明了玉米 *DHN* 家族内的成员在进化上较为保守。

在密码子之间)、相位 1(剪切发生在密码子的第 1 个核酸之后)和相位 2(剪切发生在密码子的第 2 个核酸之后)。如图 1B 所示, 所分析的 5 个玉米 *DHN* 基因内含子相位均为 0, 这些结果说明了玉米 *DHN* 家族内的成员在进化上较为保守。

表 1 玉米 5 个 *DHN* 基因的序列属性

Table 1 Sequence characteristics of the 5 *DHN* genes identified in maize

基因名称 Gene name	基因编号 Sequenced ID	氨基酸长度 Size (aa)	分子量/kDa MW	等电点 pI	染色体 Chromosome
<i>ZmDHN1</i>	GRMZM2G373522_T01	289	31.467	5.51	4
<i>ZmDHN2</i>	GRMZM2G147014_T01	290	31.441	6.05	5
<i>ZmDHN3</i>	GRMZM2G079440_T01	168	17.076	8.78	6
<i>ZmDHN4</i>	GRMZM2G052364_T01	325	31.828	8.56	8
<i>ZmDHN5</i>	GRMZM2G098750_T01	326	31.690	7.37	9

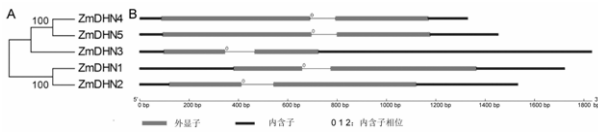


图 1 玉米 *DHN* 蛋白的进化(A)和基因结构分析(B)
Figure 1 Phylogenetic relationship (A) and gene structure (B) analysis of *DHN* proteins in maize

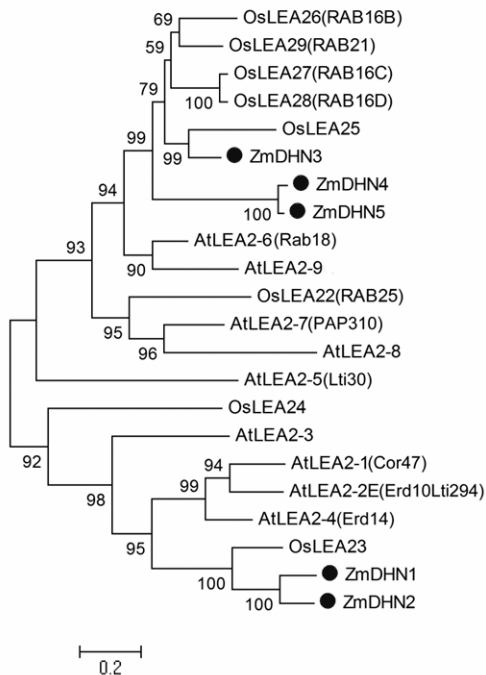


图 2 玉米、水稻和拟南芥 *DHN* 蛋白的进化关系分析
Figure 2 Phylogenetic relationships of *DHN* proteins in maize, rice and *Arabidopsis*

为进一步明晰不同物种间 *DHN* 基因的亲缘进化关系, 利用上述方法, 构建一个多物种 *DHN* 基因家族的复合进化树(图 2), 该进化树包含本研究通过生物信息学方法筛选到的 5 个玉米 *DHN* 基因家

族成员, 同时还包含了在其他物种研究过程中已经报道的 8 个水稻 *DHN* 基因和 9 个拟南芥 *DHN* 基因。结果显示, *ZmDHN1/2* 与 *OsLEA23* 聚类在一个分支中, 且与拟南芥中已报道的功能基因 *AtLEA2-1* (*Cor47*), *AtLEA2-2* (*Erd10Lti2945*), *AtLEA2-4* (*Erd14*) 进化关系较近(进化自展值达到 90%以上), 而拟南芥中已报道的这些基因与耐盐, 抗旱等功能有关, 由此可以推测玉米 *ZmDHN1/2* 基因可能也存在这些功能。另外, *ZmDHN3/4/5* 与水稻中报道的 *DHN* 基因家族成员 *OsLEA26* (*RAB16B*), *OsLEA27* (*RAB16C*), *OsLEA28* (*RAB16D*)等进化关系较近, 这些基因同样在植物的非生物胁迫反应中起到重要作用, 同样表明 *ZmDHN3/4/5* 可能与这些功能相关。

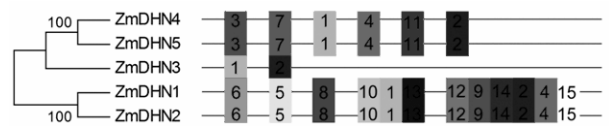


图 3 玉米 *DHN* 蛋白 15 个保守基序的分布
Figure 3 Distribution of 15 putative conserved motifs of *DHN* proteins in maize

2.3 *ZmDHN* 蛋白保守基序分析

使用 MEME 在线分析工具对玉米 *DHN* 蛋白的保守基序进行了分析, 在玉米 *DHN* 家族中共鉴定了 15 个保守基序(图 3)。每个保守的基序使用 Pfam 和 SMART 进行了注释。基序 1、2 和 4 编码 *DHN* 家族的特征结构域 *DHN*, 其中基序 1 和 2 存在于每个 *DHN* 基因成员中, 而 *ZmDHN1/2/4/5* 均包含基序 1、2 和 4。其他的 12 个基序包括 3、5、6、7、8、9、10、11、12、13、14 和 15 主要为未知功能的基

序。尽管这些基序功能目前未知,但其保守的序列结构可能暗示了其特定的生物学功能,每个保守基序的序列和相关信息列于表 2 中。

表 2 玉米 DHN 蛋白 15 个保守基序信息

Table 2 Information of the 15 conserved motifs in maize DHN proteins

基序 Motif	宽度 Width	保守基序氨基酸序列 Conserved amino acid sequences
1	29	SSSSSSSSEDDGMGCRRNGLIEKIKEK
2	17	EKKGFMDKIKEKLPQGH MCHYQQRRRQVDEYGNVCVPGDHGVQ
3	158	GQQLGECCGGYGVCGCGSYGGQQCGY GPCGCGCHDCRGYDGSHPGYGCCGCG VDDCGPGCYCCCCGRCHDCHGGVCG IQDGCCGLCGGLHGCIGMCCCCGCGCH GCCMLDCGVLGGGGHCCCGMPC
4	57	RECCCQPCCHCYPCCYVCNVCCCPNE KKGLLEKIMDKIPGYHKDYDQEQHCCCC
5	27	NDVKDRGLLDSLLGRNKHDDDDQEKNNQ
6	11	MEDERNCCQH
7	29	RPGCCGVCGCGGCFPHCCEHKCGGILRR
8	14	ELCCGMEKVCVCEP
9	18	FPCPCPPHVECHDCCVV
10	6	LLCKLH
11	9	YGGCCDCCY
12	15	VDCHCHHQEECHRPH
13	9	LKEKCKEKL
14	9	DDCKCCCPP
15	6	EHRCS

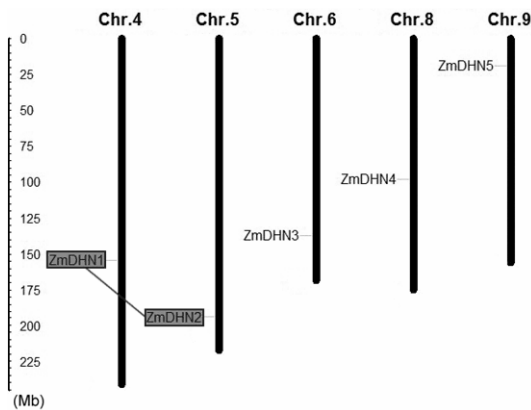


图 4 玉米 DHN 基因在染色体上的分布和基因复制

Figure 4 Chromosomal locations and gene duplication of maize DHN genes on the chromosomes

2.4 玉米 DHN 基因染色体定位和基因复制分析

玉米 5 个 DHN 基因在 10 条染色体上的分布是非随机的。由于基因家族成员较少,每条染色体上分布的基因数量差异不大(图 4)。ZmDHN1 位于第 4 号染色体, ZmDHN2 位于第 5 号染色体, ZmDHN3

位于第 6 号染色体, ZmDHN4 位于第 8 号染色体, ZmDHN5 位于第 9 号染色体, 基因在染色体上的分布未发现基因簇现象, 说明在 5 个基因中不存在串联复制的现象。进一步对玉米 DHN 基因家族的片段复制基因进行了鉴定, 根据进化关系和这些基因在染色体上同源区段的分布情况, ZmDHN1 与 ZmDHN2 具有非常相似的基因结构和紧密的进化关系。因此, 将 ZmDHN1 与 ZmDHN2 这对基因定义为片段复制基因, 参与了基因片段复制事件(图 4)。

2.5 物种间微共线性分析

为了进一步阐明玉米 DHN 基因的进化起源以及直系同源关系, 我们通过 MScanX 方法筛选了玉米与高粱, 玉米与水稻之间的直系同源基因对。通过两两循环比对, 共获得了 2 对直系同源基因, 分别为 ZmDHN2-Sb04g032250.1 和 ZmDHN2-Os02g44870.1(图 5)。我们注意到 ZmDHN2 在水稻和高粱中同时存在直系同源基因, 说明这些基因共同起源于一个祖先基因, 并推测这些基因是由基因复制产生的, 且在进化上较为保守。

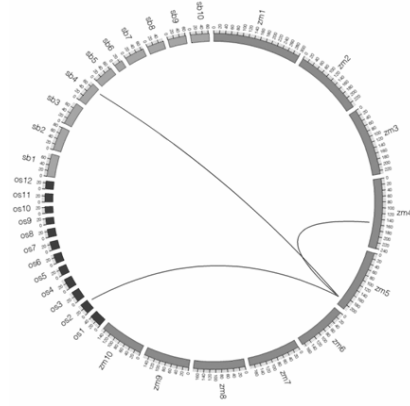


图 5 玉米、高粱和水稻间 DHN 基因的微共线性分析

Figure 5 Microsynteny of DHN genes across maize, sorghum and rice

为了更好地理解 DHN 基因家族所受到的进化约束力, 还进行了共线性区段的 Ka/Ks 分析以及复制时间的计算。一般认为推演基因进化时间的过程中通常利用非同义替换率与同义突变率的比值 (Ka/Ks) 来测验编码序列所经历的环境选择压力。从结果数据可以发现, 这 3 对复制基因的 Ka/Ks 比值都小于 1(表 3), 说明 DHN 基因家族主要经历了纯化选择的作用。同时, 还对 DHN 复制基因进行了滑动窗口分析, 从分析结果可以获取在编码序列不同位点上的 Ka/Ks 比值(图 6), 与总的 Ka/Ks 比值一样, 大多数的位点受到了中度或者强烈的纯化选择(负向选择)作用。另外, 本研究估算了与 DHN 基因

对应的片段复制事件发生的时间。依据分子钟理论，重复基因间的同义替换率(K_s)在进化过程中随时间变化的规律是相似的，因此在本研究过程中利用 K_s

值来估算包含重复基因的片段重复事件发生的时间。结果表明 3 对 DHN 基因对应的重复片段复制发生的时间大约在 1.19~4.93 百万年前(表 3)。

表 3 DHN 复制基因对的 K_a/K_s 分析以及复制时间的计算

Table 3 K_a/K_s analysis and estimated divergence time for the duplicated DHN gene pairs

DHN 复制基因对 Duplicated DHN gene pairs	K_a	K_s	K_a/K_s	复制时间/百万年 Duplication time
$ZmDHN1-ZmDHN2$	0.114	0.205	0.556	1.58
$ZmDHN2-Sb04g032250.1$	0.071	0.155	0.460	1.19
$ZmDHN2-Os02g44870.1$	0.214	0.641	0.333	4.93

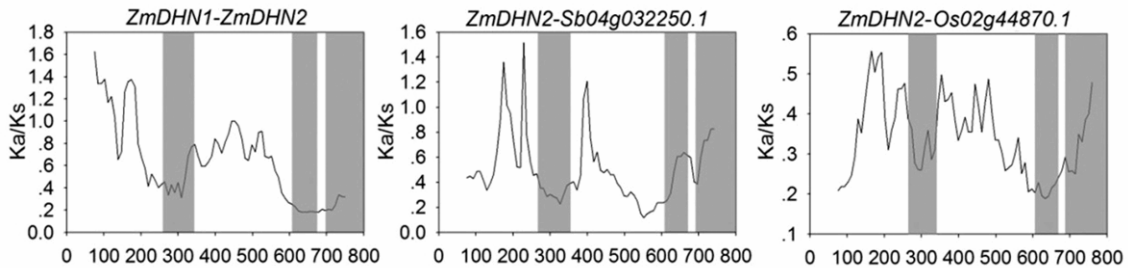


图 6 DHN 复制基因对的滑动窗口分析

Figure 6 Sliding window plots of duplicated DHN gene pairs

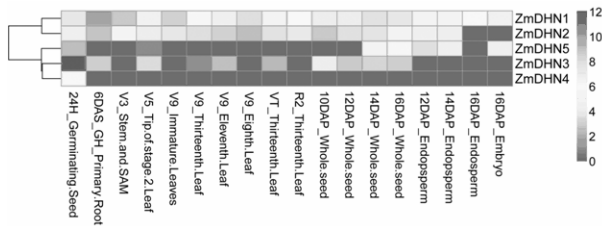


图 7 玉米 DHN 基因在不同生长发育时期表达模式分析

Figure 7 Expression profiles of $ZmDHNs$ at different development stages

表 4 玉米 DHN 基因的实时荧光定量 PCR 引物

Table 4 Gene-specific primers of maize DHN genes used for qRT-PCR analysis

基因 Gene	引物序列 Primer sequence
$ZmDHN1-F$	GACGACGACACGAAGACTCAGA
$ZmDHN1-R$	CGTCTTCCGGCTTCTTGTG
$ZmDHN2-F$	AAGCCAAAAGGCACTGAAGAAG
$ZmDHN2-R$	ACAGAACAGATCAGCAGGCTAGCTA
$ZmDHN3-F$	CGAAGACTCGTTTAGGAGAAACCA
$ZmDHN3-R$	TGCTGACCGTACTCCATGATG
$ZmDHN4-F$	TAAAGCCAGCCAAAGTTCAGAAG
$ZmDHN4-R$	TGGTAGTGCGCCATACTTCCT
$ZmDHN5-F$	AAGCCAGCCAAAGTTCAGACTT
$ZmDHN5-R$	TGGTAGTGCGCCATACTTCCT
$Actin1-F$	TGCCCCGAGGTGCTGTT
$Actin1-R$	ATGGACGCCAGGCGATT

2.6 $ZmDHN$ 转录组表达谱分析

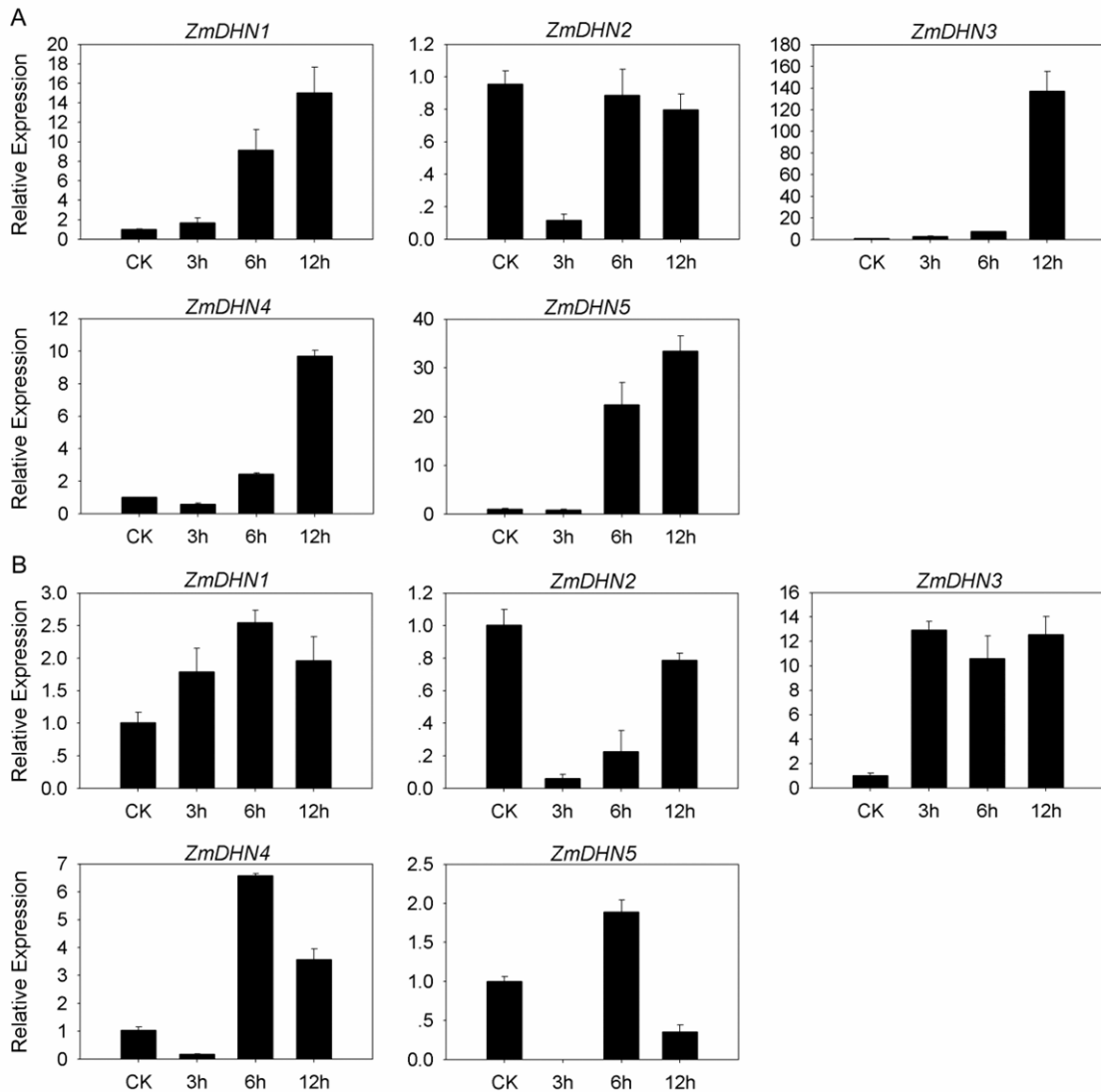
基因在整个生命过程中存在着时空表达的差异，而这些差异往往与生物体的生长和发育密切相关。为了了解玉米 DHN 基因的时空表达模式，利用 B73 自交系转录组表达数据对 5 个基因在不同发育时期的组织和器官中的表达模式进行了分析。如图 7 所示，从热图上可以发现， $ZmDHN1$ 和 $ZmDHN2$ 基因在所测定的组织中均有表达，且表达水平较高，而其余 3 个基因则主要在萌发的种子中表达水平较高，在其他组织表达相对较低。

2.7 玉米 DHN 基因诱导表达模式分析

研究表明， DHN 基因在植物的非生物胁迫应答反应中具有重要作用。为筛选玉米中参与抗逆应答反应的 DHN 基因，对 3 叶期的玉米幼苗进行了盐和干旱胁迫处理。通过 Primer Express 3.0 软件设计了基因的特异性引物(表 4)，然后通过实时荧光定量 PCR 技术对 5 个 $ZmDHN$ 基因的盐和干旱诱导表达模式进行了分析。如图 8A 所示，5 个基因在不同盐处理时间的诱导表达模式均与对照进行了比较分析， $ZmDHN1$ 、 $ZmDHN3$ 、 $ZmDHN4$ 和 $ZmDHN5$ 基因的表达量表现为显著上调， $ZmDHN2$ 基因在处理 3 h 表达为下调，在处理 6 h 后表达水平恢复与 CK 水平相一致。值得注意的是，在 4 个诱导表达上调的基因中， $ZmDHN3$ 和 $ZmDHN5$ 基因的表达模式表现较大幅度的变化(0~180 和 0~40)。另外，对 5

个 *ZmDHN* 基因的干旱诱导表达模式也进行了比较。如图 8B 所示, *ZmDHN1* 和 *ZmDHN3* 基因的表达水平表现为显著上调, *ZmDHN2* 表现为下调。另外, *ZmDHN4* 和 *ZmDHN5* 基因在处理 3 h 表达水平

下调,而在 6 h 后表达水平表现出明显的上调趋势。综上所述, *DHN* 基因均可以受到盐和干旱胁迫处理的显著诱导表达, 这些结果为进一步鉴定这些基因的功能奠定了重要基础。



CK: 对照植株; 3 h、6 h、12 h: 不同处理时间; A: NaCl 处理; B: PEG 处理

CK: normal plant; 3 h, 6 h and 12 h: different treatment time; A: salt treatment; B: PEG treatment

图 8 玉米 *DHN* 基因的盐和干旱胁迫表达模式分析

Figure 8 Expression patterns of *ZmDHNs* under salt and drought stresses

3 讨论

脱水素广泛存在于各种植物中, 属于胚胎发育晚期丰富蛋白第二家族成员。脱水素基因能够在植物胚胎发育后期及处于干旱、低温、盐碱等逆境的植株中大量表达, 具有稳定细胞膜、清除自由基、结合金属离子、保护细胞免受脱水伤害、分子伴侣的保护特性和 DNA/RNA 结合特性等功能^[9,14]。目前, 基于全基因组水平上研究 *DHN* 基因家族在拟

南芥和水稻等物种中均已报道, 而在玉米中尚未被报道。玉米基因组序列的公布使得利用生物信息学手段研究该基因家族的进化和功能成为可能^[38]。

在本研究中, 对玉米 *DHN* 基因家族做了系统全面的分析, 包括氨基酸多重序列比对、系统发育分析、保守基序、基因结构及染色体定位分析、基因复制和表达模式等。从玉米 B73 全基因组中共鉴定了 5 个非重复的玉米 *DHN* 基因, 这与其他已报道物种中的 *DHN* 基因数量较为接近, 例如拟南芥

中共有 9 个^[21], 水稻中 8 个^[22], 杨树中 10 个^[39], 葡萄中仅有 4 个^[40]。与其他类型的基因家族基因数量相比, DHN 属于数量较少的一类, 表明 DHN 家族蛋白在进化上具有高度的保守性。通过 NJ 法构建多物种系统进化树, 对玉米、水稻和拟南芥 *DHN* 基因的亲缘进化进行了比较。在系统进化树中, 发现玉米 *DHN* 基因与水稻和拟南芥的 *DHN* 基因进化关系均较为接近。同时, 通过对基因结构的分析, 发现聚类在同一分支中的基因不论是在结构上还是在保守基序上都是高度相似的, 不仅支持了系统进化树构建的可信性, 同时也为 *DHN* 基因家族的保守性进化特征提供了进一步的支撑。

基因复制事件是促进基因组和遗传系统进化的主要驱动力^[41], 研究表明玉米基因组进化历程复杂, 期间主要经历了 2 次基因组复制。第一次是较为古老的基因组复制, 大约发生在草本植物基因组分化之前(5000~7000 万年前), 最终形成了古老的加倍事件; 另一次大约发生在 500 万年前, 导致了高粱和玉米的基因组分化^[42-43]。在本研究中, 发现在玉米 *DHN* 基因家族中存在着一对片段复制基因 *ZmDHN1-ZmDHN2*, 另外在高粱和水稻中还各获得 1 对直系同源基因, 分别为 *ZmDHN2-Sb04g032250.1*, *ZmDHN2-Os02g44870.1*。注意到 *ZmDHN2* 在水稻和高粱中同时存在直系同源基因, 说明这些基因在基因组进化的早期可能起源于一个共同的祖先基因, 即起源于一个共同的复制体。通过对这 3 对 *DHN* 基因的片段复制发生的时间进行计算, 发现这些基因发生复制的时间大约在 1.19~4.93 百万年前, 大约发生在玉米基因组第 2 次复制事件之后, 这些结果在一定程度上也说明了高等植物基因组中 *DHN* 基因家族的进化相对缓慢且高度保守。

利用玉米 B73 自交系的转录组测序数据, 对玉米不同生长发育时期的组织表达模式进行了分析。从热图上可以发现, *ZmDHN1* 和 *ZmDHN2* 在各个组织中均有表达, 而 *ZmDHN3*、-4 和 -5 的表达则主要在萌发的种子中, 说明这 3 个基因在该时期可能具有重要的生物学功能。目前, 一些参与干旱, 低温, 盐胁迫诱导的 *DHN* 基因在一些物种中已经被鉴定和分离出来, 功能也得到相继证实, 这些研究结果表明通过调控 *DHN* 的表达可以显著提高植物的抗逆性^[15,17-18]。然而, 玉米中这类基因在非生物胁迫应答方面的研究还相对较少。因此, 对玉米幼苗进行了盐和干旱胁迫处理下的表达模式分析, 以此来筛选和鉴定可能在植物非生物胁迫应答反应中具有重要功能的基因。结果显示 5 个 *DHN* 基因对

盐和干旱胁迫的处理均表现出不同程度的响应。其中, *ZmDHN1*、-3、-4 和 -5 基因表现出较大幅度的诱导表达, 同时结合这些基因的多物种系统进化分析结果, 充分说明这些基因在非生物胁迫应答中具有重要的作用。这些研究结果为进一步鉴定该家族重要的 *DHN* 功能基因并开展功能分析提供了重要的参考依据。

参考文献:

- [1] Dure L III, Greenway S C, Galau G A. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis [J]. *Biochemistry*, 1981, 20(14): 4162-4168.
- [2] Ried J L, Walker-Simmons M. Group 3 late embryogenesis abundant proteins in desiccation-tolerant seedlings of wheat (*Triticum aestivum* L.) [J]. *Plant Physiol*, 1993, 102(1): 125-131.
- [3] Yue-ie C H, Chen Z Y, Hsieh J S, et al. Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds [J]. *Plant Mol Biol*, 1995, 29(4): 863-868.
- [4] Campbell S A, Crone D E, Ceccardi T L, et al. 40 kDa maize (*Zea mays* L.) embryo dehydrin is encoded by the *dhn2* locus on chromosome 9 [J]. *Plant Mol Biol*, 1998, 38(3): 417-423.
- [5] Bray E A. Molecular responses to water deficit [J]. *Plant Physiol*, 1993, 103(4): 1035-1040.
- [6] Wise M J, Tunnacliffe A. POPP the question: what do LEA proteins do? [J]. *Trends Plant Sci*, 2004, 9(1): 13-17.
- [7] 陈雷, 李磊, 李金花, 等. 植物 LEA 蛋白及其功能[J]. *中国农学通报*, 2009, 25(24): 143-146.
- [8] Dure L III, Crouch M, Harada J, et al. Common amino acid sequence domains among the LEA proteins of higher plants [J]. *Plant Mol Biol*, 1989, 12(5): 475-486.
- [9] 翟大勇, 沈黎明. 脱水蛋白研究进展[J]. *生物化学与生物物理进展*, 1998, 25(2): 119-122.
- [10] Allagulova C R, Gimalov F, Shakirova F, et al. The plant dehydrins: structure and putative functions [J]. *Biochemistry (Moscow)*, 2003, 68(9): 945-951.
- [11] Close T J. Dehydrins: a commonality in the response of plants to dehydration and low temperature [J]. *Physiol Plantarum*, 1997, 100(2): 291-296.
- [12] Jensen A B, Goday A, Figueras M, et al. Phosphorylation mediates the nuclear targeting of the maize Rab17 protein [J]. *Plant J*, 1998, 13(5): 691-697.
- [13] 马杰, 刘翠芳, 李灵之, 等. 非生物胁迫下植物脱水素的研究进展[J]. *生物学杂志*, 2012, 29(1): 71-74.
- [14] 邢鑫, 刘洋, 李德全. 植物脱水素的结构和功能[J]. *植物生理学通讯*, 2010(3): 268-276.
- [15] Puhakainen T, Hess M W, Mäkelä P, et al. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis* [J]. *Plant Mol Biol*, 2004, 54(5): 743-753.
- [16] Brini F, Saibi W, Amara I, et al. Wheat dehydrin DHN-5

- exerts a heat-protective effect on β -glucosidase and glucose oxidase activities [J]. *Biosci Biotech Bioch*, 2010, 74(5): 1050-1054.
- [17] Xu J, Zhang Y X, Wei W, et al. BjDHNs confer heavy-metal tolerance in plants [J]. *Mol Biotechnol*, 2008, 38(2): 91-98.
- [18] Yin Z, Rorat T, Szabala B M, et al. Expression of a *Solanum sogarandinum* SK 3-type dehydrin enhances cold tolerance in transgenic cucumber seedlings [J]. *Plant Sci*, 2006, 170(6): 1164-1172.
- [19] Saavedra L, Svensson J, Carballo V, et al. A dehydrin gene in *Physcomitrella patens* is required for salt and osmotic stress tolerance [J]. *Plant J*, 2006, 45(2): 237-249.
- [20] Roychoudhury A, Roy C, Sengupta D N. Transgenic tobacco plants overexpressing the heterologous lea gene Rab16A from rice during high salt and water deficit display enhanced tolerance to salinity stress [J]. *Plant Cell Rep*, 2007, 26(10): 1839-1859.
- [21] Bies-Etheve N, Gaubier-Comella P, Debures A, et al. Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana* [J]. *Plant Mol Biol*, 2008, 67(1/2): 107-124.
- [22] Wang X S, Zhu H B, Jin G L, et al. Genome-scale identification and analysis of LEA genes in rice (*Oryza sativa* L.) [J]. *Plant Sci*, 2007, 172(2): 414-420.
- [23] Finn R D, Mistry J, Schuster-Böckler B, et al. Pfam: clans, web tools and services [J]. *Nucleic Acids Res*, 2006, 34(suppl 1): D247-D251.
- [24] Tamura K, Dudley J, Nei M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0 [J]. *Mol Biol Evol*, 2007, 24(8): 1596-1599.
- [25] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Res*, 1994, 22(22): 4673-4680.
- [26] Letunic I, Doerks T, Bork P. SMART 6: recent updates and new developments [J]. *Nucleic acids Res*, 2009, 37(suppl 1): D229-D232.
- [27] Artimo P, Jonnalagedda M, Arnold K, et al. ExpASY: SIB bioinformatics resource portal [J]. *Nucleic Acids Res*, 2012, 40(Web Server Issue): W597-603.
- [28] Guo A Y, Zhu Q H, Chen X, et al. GSDS: a gene structure display server [J]. *Yi Chuan*, 2007, 29(8): 1023-1026.
- [29] Bailey T L, Elkan C. The value of prior knowledge in discovering motifs with MEME [J]. *Proc Int Conf Intell Syst Mol Biol*, 1995(3): 21-29.
- [30] Peng X, Zhao Y, Cao J, et al. CCCH-type zinc finger family in maize: genome-wide identification, classification and expression profiling under abscisic acid and drought treatments [J]. *PloS one*, 2012, 7(7): e40120.
- [31] Rozas J, Sanchez-DelBarrio J C, Messeguer X, et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods [J]. *Bioinformatics*, 2003, 19(18): 2496-2497.
- [32] Quraishi U M, Abrouk M, Murat F, et al. Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution [J]. *Plant J*, 2011, 65(5): 745-756.
- [33] Wang Y, Tang H, Debarry J D, et al. MCSanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity [J]. *Nucleic Acids Res*, 2012, 40(7): e49.
- [34] Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*, 2009, 19(9): 1639-1645.
- [35] Sekhon R S, Lin H, Childs K L, et al. Genome-wide atlas of transcription during maize development [J]. *Plant J*, 2011, 66(4): 553-563.
- [36] Zhao Y, Tong H, Cai R, et al. Identification and characterization of the RCI2 gene family in maize (*Zea mays*) [J]. *J Genet*, 2014, 93(3): 655-666.
- [37] Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method [J]. *Methods*, 2001, 25(4): 402-408.
- [38] Schnable P S, Ware D, Fulton R S, et al. The B73 maize genome: complexity, diversity, and dynamics [J]. *Science*, 2009, 326(5956): 1112-1115.
- [39] Liu C C, Li C M, Liu B G, et al. Genome-wide identification and characterization of a dehydrin gene family in poplar (*Populus trichocarpa*) [J]. *Plant Mol Biol Rep*, 2012, 30(4): 848-859.
- [40] Yang Y Z, He M Y, Zhu Z G, et al. Identification of the dehydrin gene family from grapevine species and analysis of their responsiveness to various forms of abiotic and biotic stress [J]. *BMC Plant Biol*, 2012: 12.
- [41] Moore R C, Purugganan M D. The early stages of duplicate gene evolution [J]. *Proc Natl Acad Sci U S A*, 2003, 100(26): 15682-15687.
- [42] Salse J, Bolot S, Throude M, et al. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution [J]. *Plant Cell*, 2008, 20(1): 11-24.
- [43] Gaut B S. Evolutionary dynamics of grass genomes [J]. *New Phytol*, 2002, 154(1): 15-28.