

# 大豆 *Abhydrolase\_3* 基因家族进化分析

王 娇,张培培,程 浩,喻德跃

(南京农业大学 大豆研究所/国家大豆改良中心/作物遗传与种质创新国家重点实验室,江苏 南京 210095)

**摘 要:**异黄酮是大豆的重要次生代谢物,参与植物与微生物互作。2-羟基异黄酮脱水酶(hydroxyisoflavanone dehydratase, HID)催化2-羟基异黄酮形成稳定的异黄酮。*HID*属于*Abhydrolase\_3*基因家族,该基因家族具有多种功能,但该基因家族在大豆中的进化模式尚待研究。为了研究*Abhydrolase\_3*基因家族在大豆中的进化模式,本文在大豆基因组中鉴定了62个*Abhydrolase\_3*基因,串联和片段复制是该基因家族主要扩增方式。根据系统进化关系,将大豆*Abhydrolase\_3*基因家族划分为8个亚家族,其中*HID*所在的亚家族I基因数量最多,并发生多次基因扩增事件。对大豆*Abhydrolase\_3*基因家族结构分析表明,不同亚家族具有不同的基序。多态性分析表明,亚家族I、III和V具有较高的核苷酸差异,并受到放松的自然选择。基因表达分析表明,除了亚家族II和IV外,其它亚家族的基因在大豆不同组织中有较高表达;亚家族I、III、IV、V和VI基因受病原菌诱导表达。结果说明*HID*所在的亚家族I存在基因扩增和功能分化,与病原菌互作相关的基因具有较高的遗传多样性并受病原菌诱导表达。

**关键词:**大豆;*Abhydrolase\_3*基因家族;进化;功能;基因复制

**中图分类号:**S565.1      **文献标识码:**A      **DOI:**10.11861/j.issn.1000-9841.2017.06.0842

## Evolution Analysis of *Abhydrolase\_3* Gene Family in Soybean

WANG Jiao, ZHANG Pei-pei, CHENG Hao, YU De-yue

(Soybean Research Institute /National Center for Soybean Improvement/National Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China)

**Abstract:** Isoflavonoids are important secondary metabolites of soybean and involved in the interactions with pathogens. 2-hydroxyisoflavanone dehydratase (HID) catalyzes the dehydration of 2-hydroxyisoflavanones to the formation of stable isoflavonoids. *HID* belongs to the *Abhydrolase\_3* gene family, which is involved in plenty of functions. To study the evolutionary patterns of *Abhydrolase\_3* gene family in soybean, 62 *Abhydrolase\_3* genes were identified from soybean genome. Tandem and segmental duplications are the primary duplicated patterns of this gene family. Based on the phylogenetic relationships, eight subfamilies were classified. Subfamily I contains the largest number of genes, in which several times of gene duplications occur. Genes in different subfamilies were characterized by distinct patterns of motif combinations. Genetic polymorphic analysis showed homologous genes in subfamilies I, III and V possess a higher level of nucleotide diversity, and experience a relaxed selection pressure. The expression analysis showed that genes in most subfamilies except II and IV show high levels of expression in different tissues, and the expression levels of genes in subfamilies I, III, IV, V and VI were induced by pathogens. The results showed that gene expansion and functional divergence exist in subfamily I, which contains *HID*, and that genes involved in the interaction with pathogens possess higher levels of genetic diversity and the expressions are induced by pathogens.

**Keywords:** Soybean; *Abhydrolase\_3* gene family; Evolution; Function; Gene duplication

异黄酮是一类重要的植物次生代谢物,主要分布在豆科植物中,在植物与微生物相互作用中发挥重要作用<sup>[1-2]</sup>。异黄酮由异黄酮代谢途径合成,在豆科特异的异黄酮合成酶(2-hydroxyisoflavanone synthase, IFS)的催化下,形成异黄酮类化合物2-羟基异黄酮<sup>[3]</sup>。2-羟基异黄酮不稳定,可以自发脱水或在2-羟基异黄酮脱水酶(hydroxyisofla-

vanone dehydratase, HID)的催化下形成异黄酮<sup>[4]</sup>。HID包含*Abhydrolase\_3*(PF07859)结构域,属于*Abhydrolase\_3*基因家族。*Abhydrolase\_3*基因家族属于 $\alpha/\beta$ 水解酶超家族, $\alpha/\beta$ 水解酶超家族在生物界中广泛分布<sup>[5]</sup>。 $\alpha/\beta$ 水解酶的核心是 $\alpha/\beta$ 折叠,具有保守的活性位点,包含催化三联体(catalytic triad):GX SXG结构域中的丝氨酸位点、酸性残基(植物多

为天冬氨酸)和组氨酸,还有高度保守的 HGG 盒,在水解过程中间体的稳定方面发挥重要作用<sup>[6]</sup>。植物中含有 *Abhydrolase\_3* 结构域的基因通常被分为植物羧酸酯酶(plant\_carboxylesterase, CXE)基因家族或激素敏感脂肪酶类基因家族(hormone-sensitive\_lipase\_like family, HSL)<sup>[5]</sup>。在拟南芥中鉴定了 20 个羧酸酯酶基因,属于 *Abhydrolase\_3* 基因家族<sup>[7]</sup>。

*Abhydrolase\_3* 基因家族除了参与次生代谢物异黄酮的合成外,还参与其它次生代谢物合成。如在郁金香(*Tulipa gesneriana*)中,郁金香苷 A、B 转化酶(PosA-converting enzyme, TgTCEA; PosB-converting enzyme, TgTCEB)催化次生代谢物-郁金香苷 A 和 B(Tuliposides A 和 B)转变为具有抗菌活性的郁金香内酯(tulipalins),TgTCEA 和 TgTCEB 在这里不具有水解酶活性,而具有分子内酯交换酶的活性<sup>[8]</sup>。

除了参与植物次生代谢,*Abhydrolase\_3* 基因家族在植物中还具有其它多种功能。首先,具有酯酶水解活性,如拟南芥的 CXE8 可以水解 4-硝基苯基乙酸盐和 4-硝基苯基丁酸盐<sup>[9]</sup>,CXE12 水解前除草剂 2,4-二氯苯氧乙酸甲酯(2,4-D-methyl)至植物毒性酸 2,4-二氯苯氧乙酸(2,4-dichlorophenoxyacetic acid,2,4-D)<sup>[10]</sup>,CXE18 可以水解甲基 7-羟基香豆素酯类<sup>[11]</sup>。

其次,*Abhydrolase\_3* 基因家族参与植物的抗病抗逆反应,如烟草(*Nicotiana tabacum*) *hsr203j* 基因受病原菌感染诱导,可能在超敏反应中调节细胞凋亡<sup>[12]</sup>。辣椒(*Capsicum annuum*) *PepEST* 基因受到链格孢菌(*Alternaria brassicola*) 感染的诱导,在病原菌抗性反应中发挥作用<sup>[13]</sup>。拟南芥 CXE8 与 *PepEST* 同源,响应真菌感染诱导,在抵御病原菌感染中发挥作用<sup>[9]</sup>。

另外,*Abhydrolase\_3* 基因家族参与植物激素信号通路。如赤霉素(gibberellin, GA)受体 *GID1* (*GIBBERELLIN INSENSITIVE DWARF1*) 基因属于 *Abhydrolase\_3* 基因家族。*GID1* 在 GA 信号通路中,与 GA 连接形成 *GID1*-GA 复合物,该复合物与 DELLA 蛋白(GA 作用的负调节子)相互作用,降解 DELLA 蛋白,从而激活 GA 功能<sup>[14-15]</sup>。拟南芥 *Abhydrolase\_3* 基因家族成员 *AT5G15860* 编码异戊烯半胱氨酸甲基转移酶(isoprenylcysteine methyltransferase, ICME),参与脱落酸(abscisic acid, ABA)信号通路,是 ABA 信号正调节子<sup>[16]</sup>。

虽然迄今为止有很多关于 *Abhydrolase\_3* 基因家族功能的报道,但是 *Abhydrolase\_3* 基因家族在大豆中的进化模式尚待研究。本文关注 *Abhydrolase\_3*

基因家族在大豆中进化模式,并与其它 6 个高等植物进行比较。鉴定了包含大豆在内的 7 个植物中的 *Abhydrolase\_3* 基因家族成员并研究了该基因家族的进化模式,旨在为 *Abhydrolase\_3* 基因家族进一步功能研究提供基础。

## 1 材料与方法

### 1.1 *Abhydrolase\_3* 基因家族鉴定

首先在大豆(*Glycine max*)中鉴定了 *Abhydrolase\_3* 基因家族成员。其次,与 4 个双子叶植物:菜豆(*Phaseolus vulgaris*)、苜蓿(*Medicago truncatula*)、葡萄(*Vitis vinifera*)、拟南芥(*Arabidopsis thaliana*), 2 个单子叶植物:水稻(*Oryza sativa*)、高粱(*Sorghum bicolor*) 比对,鉴定其包含的 *Abhydrolase\_3* 基因家族成员。基因组序列和相应的注释下载自 phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>)。使用如下方法鉴定 *Abhydrolase\_3* 基因家族成员:首先,以 *Abhydrolase\_3* 结构域的氨基酸一致序列(Pfam: PF07859)作为问询序列,在相应的基因组中进行 BLASTP, E 值设置为 1.0。然后,在 PFAM HMMs 中对蛋白序列打分,显著匹配 *Abhydrolase\_3* 结构域的蛋白序列被保留<sup>[17]</sup>。

### 1.2 串联和片段复制

对 *Abhydrolase\_3* 基因家族成员在基因组上的位置进行调查,串联复制定义为相邻的 *Abhydrolase\_3* 基因间隔 < 200 kb<sup>[17-18]</sup>。片段复制基因的鉴定来自于植物基因组复制数据库(Plant Genome Duplication Database PGDD; <http://chibba.agtec.uga.edu/duplication/>)<sup>[19]</sup>。

### 1.3 序列比对和进化分析

使用氨基酸序列指导核苷酸编码序列排序。在 MEGA 6.0 中使用 Clustal W 的默认选项进行多序列比对<sup>[20-21]</sup>。系统进化树基于自展检验的邻接法和 p-distance 模型由 MEGA 6.0 构建。自展分析设置为 1 000 个重复对节点的稳定性进行评估。核苷酸差异使用 Jukes 和 Cantor 校正<sup>[22]</sup>。同义(Ks)和非同义替换率(Ka)基于 Nei and Gojobori 方法计算<sup>[23]</sup>。Ka 和 Ks 的比值(Ka/Ks)被用做自然选择压力检测。

### 1.4 基因结构分析

使用 MEME (multiple expectation-maximization for motif elicitation) 鉴定分析 *Abhydrolase\_3* 基因家族的 10 个保守基序<sup>[24]</sup>。这些基序的 *P* 值 ≤ 10<sup>-6</sup> 且相互间不重叠。

### 1.5 基因表达分析

*Abhydrolase\_3* 基因家族的组织表达数据来源于

SoyBase (<http://www.soybase.org/soyseq/>) 网站上 14 个组织的 RNA-Seq 数据,包含 3 个营养组织(根、叶和根瘤)和 11 个不同发育时期的生殖生长组织(花、荚和种子)。基因在不同组织的表达量经过标准化处理,即表达量( reads )/长度( kb )/百万( million ) ( reads/kb/million, RPKM ) 经  $\log_{10}$  转化,如果  $RPKM=0$ , 则  $\log_{10}$  值设为  $-3$ 。大豆锈菌诱导表达数据从植物表达数据库( plant expression database, <http://www.plexdb.org>)下载( 试验 GM2 )。该试验利用大豆基因组表达芯片( genechip soybean genome arrays, Affymetrix ),调查含有 *Rpp2* 抗病基因的大豆抗病植株 PI230970 和不含有 *Rpp2* 抗病基因的感病植株 Embrapa-48,在锈菌( *Phakopsora pachyrhizi* ) 侵染后 7 d 的 10 个不同时间点( 6, 12, 18, 24, 36, 48, 72, 96, 120, 168 h ),侵染和未侵染大豆叶片中基因表达量的变化,  $|\log_2(\text{试验组}/\text{对照组})| \geq 1$  的基因被认为表达量发生明显改变<sup>[25-26]</sup>。在锈菌侵染后的某个或某几个时间点中表达量发生明显改变的基因为受到诱导基因;不受诱导基因为锈菌侵染后任何一个时间点基因表达量都未发生明显改变。表达数据使用 Cluster 3.0 和 Java TreeView 1.1.6 软件以热图形式呈现,其中红色和绿色分别

代表上调和下调表达。

2 结果与分析

2.1 *Abhydrolase\_3* 基因家族的鉴定

由于 *HID* 是大豆异黄酮代谢通路的重要基因,属于 *Abhydrolase\_3* 基因家族,因此首先在大豆( *Glycine max* )中进行 *Abhydrolase\_3* 基因家族鉴定。在大豆中鉴定出 62 个 *Abhydrolase\_3* 基因,其中 30 个基因( 48.4% ) 由串联复制产生, 43 个基因( 69.4% ) 由片段复制产生。为了与其它植物比较,在菜豆、苜蓿、葡萄、拟南芥、水稻和高粱中分别鉴定出 39,46,38,23,69 和 66 个 *Abhydrolase\_3* 基因,说明该基因家族在其它植物中也发生了不同程度的基因扩增( 表 1 )。进一步分析该基因家族在其它植物中的扩增模式,结果发现,在其它植物中有 30.4% ~ 71.0% 的基因由串联复制产生,有 19.6% ~ 38.5% 的基因是由片段复制产生的。可见,串联复制和片段复制是植物中 *Abhydrolase\_3* 基因家族扩增的主要方式,其中在大豆中片段复制是主要的扩增方式。大豆由片段复制产生的基因最多,这与之前的研究一致<sup>[17]</sup>,可能与大豆近期发生两次全基因组复制有关<sup>[27]</sup>。

表 1 不同物种中 *Abhydrolase\_3* 基因家族的基因数目  
Table 1 Number of *Abhydrolase\_3* genes in different species

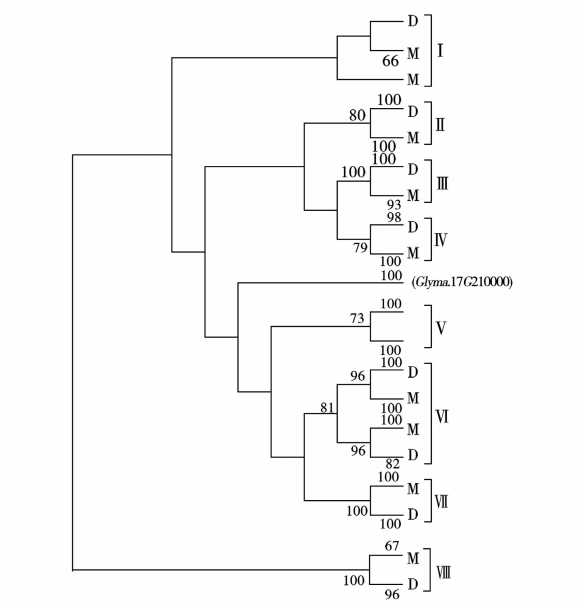
物种 Species	数目 Number of gene	串联复制数目( 比例 ) Number ( percent ) of tandem duplication	片段复制数目( 比例 ) Number ( percent ) of segmental duplication
大豆 <i>Glycine max</i>	62	30 ( 48.4% )	43 ( 69.4% )
菜豆 <i>Phaseolus vulgaris</i>	39	21 ( 53.8% )	15 ( 38.5% )
苜蓿 <i>Medicago truncatula</i>	46	29 ( 63.0% )	9 ( 19.6% )
葡萄 <i>Vitis vinifera</i>	38	21 ( 55.3% )	9 ( 23.7% )
拟南芥 <i>Arabidopsis thaliana</i>	23	7 ( 30.4% )	8 ( 34.8% )
水稻 <i>Oryza sativa</i>	69	49 ( 71.0% )	15 ( 21.7% )
高粱 <i>Sorghum bicolor</i>	66	44 ( 66.7% )	21 ( 31.8% )
总数 Total	343	201 ( 58.6% )	118 ( 34.4% )

2.2 大豆 *Abhydrolase\_3* 基因家族系统进化分析

为研究大豆 *Abhydrolase\_3* 基因间的系统发生关系,首先对大豆的 *Abhydrolase\_3* 基因进行系统进化分析( 图 1 )。按照自展值  $> 70\%$ , 将大豆 *Abhydrolase\_3* 基因家族分为 8 个亚家族( 亚家族 I 到 VIII,每个亚家族包含至少 2 个基因 )。除 *Glyma.17G210000* 不在划分的亚家族外,其余基因均位于不同的亚家族。每个亚家族含有的基因成员的数目不等,亚家族 I 包含 21 个基因成员,而亚家族 VI-

II 仅包含 2 个基因成员( 表 2 )。由于每个亚家族内的基因起源于一个共同祖先,因此亚家族内发生次数不等的基因扩增。为了解各个亚家族基因的扩增模式,分别统计每个亚家族串联复制和片段复制的基因数目。从表 2 可以看出,每个亚家族包含数目不等的串联复制和片段复制基因。其中基因数目最多的亚家族 I 包含串联复制和片段复制基因数量最多,分别有 14 和 13 个基因来自于串联复制和片段复制。



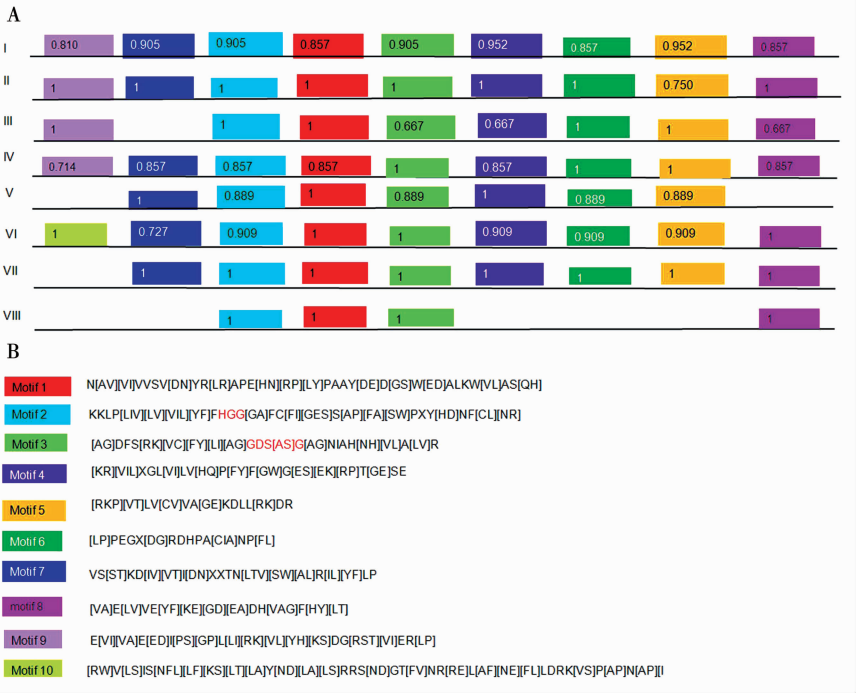


D: 双子叶植物 (大豆、菜豆、苜蓿、葡萄和拟南芥); M: 单子叶植物 (水稻和高粱)。  
D: Dicotyledon (*Glycine max*, *Phaseolus vulgaris*, *Medicago truncatula*, *Vitis vinifera* and *Arabidopsis thaliana*); M: Monocotyledon (*Oryza sativa* and *Sorghum bicolor*).

图2 7个植物 *Abhydrolase\_3* 基因家族系统进化树  
Fig.2 Phylogenetic tree of *Abhydrolase\_3* gene family in 7 plants

2.3 *Abhydrolase\_3* 基因家族结构分析

为了进一步寻找大豆中不同亚家族所包含的保守序列,使用 MEME 预测该基因家族具有的基序 (motif), 并分析每个亚家族的基序组成和出现频率,出现频率低于 50% 的基序不显示 (图 3)。基序 1、2 和 3 在每个亚家族中出现频率均高于 66.7%, 说明这些基序为不同亚家族间共同存在的保守基序。*Abhydrolase\_3* 基因家族一般包含保守的 GX-SXG 和 HGG 结构域<sup>[6]</sup>, 与此一致的是, HGG 和 GX-SXG 序列分别位于保守基序 2 和 3 中, 为不同亚家族所共有基序。其余基序 (基序 4~10) 并非存在于每个基因亚家族中, 为特异基序。不同亚家族中包含不同的基序, 亚家族 I、II 和 IV 包含基序 1~9, 亚家族 III 包含除基序 7 和 10 外的其它 8 个基序, 亚家族 V 包含基序 1~7, 亚家族 VI 包含除基序 9 外的其它 9 个基序, 亚家族 VII 包含基序 1~8 基序, 亚家族 VIII 仅具有 1、2、3 和 8 四个基序。可以看出, 亚家族 I~VII 除个别基序外, 共有其它多数基序, 而亚家族 VIII 与其它亚家族仅共有部分基序, 这之前将亚家族 I~VII 和 VIII 划分为不同分组一致<sup>[28]</sup>。另外, 虽然不同亚家族具有不同的基序, 但基序的排列顺序在不同亚家族中具有保守性。



A: 不同基序在基因上的排列顺序, 数字表示基序出现的频率; B: 基序及其氨基酸序列, 红色标注 GX-SXG 和 HGG 保守序列。  
A: The arrangement of different motifs on the genes, numbers in the box indicate the occurrence frequency of a motif;  
B: Motifs and the corresponding amino acid sequences, red color indicates conserved sequences of GX-SXG and HGG.

图3 *Abhydrolase\_3* 基因家族不同亚家族基序组成及频率

Fig.3 Motif construction and usage frequency of different subfamilies in *Abhydrolase\_3* gene family

2.4 *Abhydrolase\_3* 基因家族多态性分析

为了研究不同亚家族间的核苷酸差异及进化动力,对大豆同源基因对间的相关遗传参数进行计算(表 3)。这些同源基因位于共线性片段上,由大豆基因组复制而产生。除了亚家族 VIII 没有片段复制基因对外,其余亚家族间具有数目不等的同源基因对,亚家族 I 具有 6 对片段复制基因,亚家族 III 具有 1 对片段复制基因。亚家族间核苷酸存在很大差异,亚家族 I、III 和 V 具有较高的核苷酸差异,分

别为 0.131、0.226 和 0.098,其余亚家族核苷酸差异较低。

非同义替换率和同义替换率的比值( $Ka/Ks$ )可以用来衡量每个亚家族所受到的选择压。 $Ka/Ks$  等于 1 支持中性选择, $Ka/Ks$  大于 1 暗示存在正选择,而  $Ka/Ks$  小于 1 支持纯化选择。从表 3 看,每个亚家族的  $Ka/Ks$  值均小于 1,说明这些亚家族受到纯化选择。亚家族 I、III 和 V 的  $Ka/Ks$  值高于其它亚家族,说明这些亚家族受到放松的纯化选择压。

表 3 大豆片段复制基因核苷酸差异、同义、非同义替换率  
Table 3 Nucleotide diversity, synonymous and nonsynonymous substitution rate in segmentally duplicated genes of *Glycine max*

亚家族 Subfamily	基因对数 Pairs No.	核苷酸差异 Nucleotide diversity	同义替换率 Synonymous substitution rate/Ks	非同义替换率 Non-synonymous substitution rate/Ka	非同义与同义替换率比值 Ratios of non-synonymous to synonymous substitutions ( $Ka/Ks$ )
I	6	0.131	0.275	0.094	0.358
II	2	0.063	0.135	0.042	0.317
III	1	0.226	0.455	0.166	0.365
IV	3	0.086	0.218	0.049	0.222
V	2	0.098	0.159	0.080	0.487
VI	3	0.071	0.204	0.034	0.149
VII	2	0.059	0.153	0.031	0.203

2.5 基因表达分析

为了鉴定大豆 *Abhydrolase\_3* 基因家族的组织表达模式,根据 RNA 测序数据分析它们在 14 个组织中的表达情况。62 个 *Abhydrolase\_3* 基因中有 58 个基因可以找到表达数据(表 4,图 4A)。其中 30 个基因在大部分组织中均有较高表达量,另外 28 个基因在大部分组织中未检测到表达或表达量较

低。亚家族 I、III、V、VI 和 VIII 的基因可以检测到两种组织表达模式,即有些基因在大部分组织中有较高表达,有些基因在大部分组织中未检测到表达或表达量较低。亚家族 VII 的基因在大部分组织中有较高表达。亚家族 II、IV 的基因成员在大部分组织中未检测到表达或表达量较低。

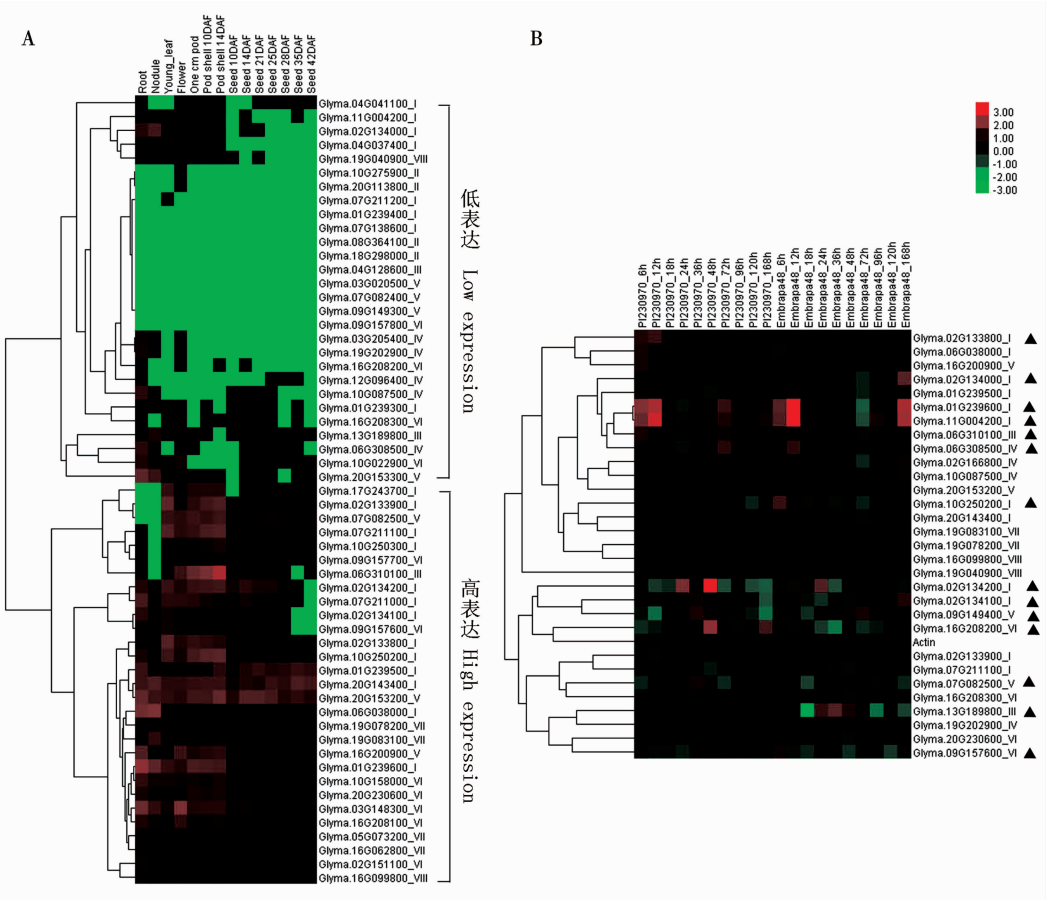
表 4 不同亚家族组织和诱导表达模式  
Table 4 Expression patterns of different subfamilies in different tissues and in response to pathogens

表达模式 Expression pattern	亚家族 Subfamily								总计 Total
	I	II	III	IV	V	VI	VII	VIII	
组织表达 Tissue expression	21	4	3	5	8	11	4	2	58
高表达 High expression	13	0	1	0	4	7	4	1	30
低表达 Low expression	8	4	2	5	4	4	0	1	28
诱导表达 Expression under pathogen infection	12	0	2	4	4	4	2	2	30
受到诱导 Induced	7	0	2	1	2	2	0	0	14
不受诱导 Non-induced	5	0	0	3	2	2	2	2	16



已知 *Abhydrolase\_3* 基因家族参与植物和微生物互作,为研究大豆 *Abhydrolase\_3* 基因对生物胁迫的响应,利用锈菌 (*Phakopsora pachyrhizi*) 侵染大豆后的表达数据调查它们在病原菌诱导后的表达模式<sup>[25]</sup>。在所调查的 62 个大豆 *Abhydrolase\_3* 基因中,30 个基因有表达数据,这些基因来自除亚家族 II 外的其它各亚家族(表 4,图 4B)。以组成型表达的 *Actin* 基因为参照,*Actin* 基因在诱导后表达量变化不明显,说明该基因不受诱导表达。亚家族 I、

IV、V 和 VI 的部分基因在锈菌侵染后表达量发生变化,说明这些基因的表达受锈菌诱导,另外一部分基因在锈菌侵染后表达量变化不明显,这些基因的表达不受锈菌诱导。亚家族 III 的两个基因的表达均受锈菌诱导,如 *Glyma. 13G189800* 在锈菌侵染感病植株 (*Embrapa-48*) 18 h 后下调表达,在 24 和 36 h 上调表达,在 96 h 后又下调表达。亚家族 VII 和 VIII 的基因成员锈菌诱导后,表达量无明显变化,说明这些基因不受锈菌诱导。



黑色三角形:受病原菌诱导表达基因。  
Black triangles: Gene expression levels were induced by pathogen.

图 4 *Abhydrolase\_3* 基因家族组织表达 (A) 和病原菌诱导表达模式 (B)  
Fig. 4 Expression levels of *Abhydrolase\_3* genes in various tissues (A) and in response to pathogens (B)

3 结论与讨论

异黄酮是大豆重要的次生代谢物,参与固氮菌共生和植物抗性反应,在抗性反应中,除本身具有抗病原菌活性外,还参与具有广谱抗性的大豆抗毒素的合成<sup>[1-2]</sup>。2 - 羟基异黄酮脱水酶 (hydroxy-isoflavanone dehydratase, HID) 可以高效催化 2 - 羟基异黄酮脱水形成异黄酮<sup>[4]</sup>。HID 属于 *Abhydrolase\_3* 基因家族,本文对 *Abhydrolase\_3* 在大豆中的进化

模式进行研究。*Abhydrolase\_3* 基因家族属于  $\alpha/\beta$  水解酶超家族,该超家族包含多种酶,如羧酸酯酶、激素敏感脂肪酶等,在动植物和微生物中广泛分布,说明其古老起源<sup>[6]</sup>。*Abhydrolase\_3* 基因家族的原始功能是具有酯酶水解酶活性<sup>[7]</sup>,在此基础上,该基因家族分化出多种功能,如 HID 参与大豆次生代谢。另外,该基因家族还参与抗病抗逆反应和激素信号通路等<sup>[13-14]</sup>。*Abhydrolase\_3* 基因家族的功能分化可能由于进化过程中发生基因复制、突变等,导

致蛋白结构和酶活的变异,从而进化出多种多样的功能<sup>[29]</sup>。

此研究将 *Abhydrolase\_3* 基因家族划分为 8 个亚家族,从基因功能来看,亚家族 I 包含与次生代谢相关的大豆 2-羟基异黄酮脱水酶编码基因 *HID* (*Glyma.01G239600*)和郁金香的郁金香苷 A 和 B 转化酶编码基因 *TgTCEA* 和 *TgTCEB*,另外还包含拟南芥具有水解活性的 *CXE12* 基因<sup>[4, 8, 10]</sup>。亚家族 III 包含来自烟草、豌豆(*Pisum sativum*, *Ps*)和番茄(*Lycopersicon esculentum*, *Le*)的 *hsr203j* 基因,参与植物的抗病抗逆反应<sup>[12]</sup>。亚家族 V 包含来自辣椒的 *CaPepEST* 基因和拟南芥的 *CXE8* 基因,它们在病原菌抗性反应中发挥作用<sup>[9, 13]</sup>。亚家族 VI 包含 3 个拟南芥赤霉素受体基因 *GID1A*(*AT3G05120*)、*GID1B* (*AT3G63010*)、*GID1C*(*AT5G27320*)和一个水稻赤霉素受体基因 *OsGID1*<sup>[30-31]</sup>。亚家族 VIII 包含拟南芥异戊烯半胱氨酸甲基转移酶基因 *ICME* (*AT5G15860*),参与脱落酸 ABA 信号反应<sup>[16]</sup>。由此可见,不同亚家族间以及亚家族内存在功能分化。

大豆中不同亚家族的基因多态性分析表明亚家族 I、III 和 V 基因多态性较高,受到放松的自然选择。这些亚家族都与病原菌互作相关,如亚家族 I 与次生代谢物异黄酮和郁金香苷等合成相关,这些次生代谢物参与病原菌-植物互作<sup>[4, 8]</sup>,亚家族 III 和 V 包含 *hsr203J*、*CaPepEST* 和 *CXE8* 等基因,这些基因家族都与植物-病原菌互作相关<sup>[9, 12-13]</sup>,诱导表达分析表明亚家族 I、III 和 V 基因受病原菌诱导表达,支持这些亚家族与病原菌互作相关。亚家族 VI 基因多态性较低,受到纯化选择。该亚家族包含赤霉素受体基因 *GID1*,为赤霉素信号反应中所必需,研究表明 *GID1* 介导的 GA 信号通路在维管束植物中普遍存在,具有保守性,因此受到严格的纯化选择作用,变异较少<sup>[32]</sup>。这些结果说明不同亚家族的进化模式不同,与病原菌互作相关的基因受到放松的纯化选择作用,变异速度快,具有较高的遗传多样性;具有保守激素受体功能的亚家族,受到严格的纯化选择作用,变异速度慢,遗传多态性低。

除了亚家族间存在功能分化外,亚家族内也存在功能分化。亚家族 I 具有最多的基因数目,主要由串联复制和片段复制产生。除了包含与植物次生代谢和病原菌互作相关的编码大豆 2-羟基异黄酮脱水酶的基因 *HID*、具有分子内酯交换酶的活性的编码郁金香苷 A、B 转化酶的基因 *TgTCEA* 和

*TgTCEB* 外,亚家族 I 还包含具有水解活性的拟南芥 *CXE12* 基因,该基因可以将外源物质前除草剂 2,4-二氯苯氧乙酸甲酯水解至有毒性的 2,4-二氯苯氧乙酸<sup>[10]</sup>。这些基因在系统进化树上聚集在同一支,说明它们起源于同一共同祖先。另外,*CXE12*、*HID*、*TgTCEA* 和 *TgTCEB* 在亚家族 I 内分别位于不同的分支,可以推断亚家族 I 内基因自共同祖先分化后,发生了基因复制和功能分化,在原始酯酶水解功能的基础上(如 *CXE12*),发生新功能化,进化出特异的功能。如大豆的 *HID* 可以特异催化 2,4'-二羟基甲氧基异黄酮和 2-羟基-4'-甲氧基异黄酮产生异黄酮<sup>[4]</sup>,郁金香的 *TgTCEA* 和 *TgTCEB* 可以特异催化郁金香苷 A 和 B 转变为郁金香内酯<sup>[8]</sup>。这些特异的催化功能可能与某些催化活性位点的变异有关,如 *HID* 将催化三联体中的丝氨酸活性位点取代为苏氨酸,失去了水解酶的作用<sup>[4]</sup>。

总之,对大豆 *Abhydrolase\_3* 基因家族进化模式的研究为其功能研究提供线索。随着越来越多 *Abhydrolase\_3* 基因家族功能的研究,*Abhydrolase\_3* 基因的进化模式与功能分化间的关系将越来越清晰。

参考文献

[1] Puepke S G. The genetic and biochemical basis for nodulation of legumes by rhizobia[J]. Critical Reviews in Biotechnology, 1996, 16(1):1-51.

[2] Dixon R A. Natural products and plant disease resistance[J]. Nature, 2001, 411(6839):843-847.

[3] Jung W, Yu O, Lau S M, et al. Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes[J]. Nature Biotechnology, 2000, 18(2):208-212.

[4] Akashi T, Aoki T, Ayabe S. Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis [J]. Plant Physiology, 2005, 137(3):882-891.

[5] Lenfant N, Hotelier T, Velluet E, et al. ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins: Tools to explore diversity of functions[J]. Nucleic Acids Research, 2013, 41(Database issue):D423-429.

[6] Ollis D L, Cheah E, Cygler M, et al. The alpha/beta hydrolase fold[J]. Protein Engineering, 1992, 5(3):197-211.

[7] Marshall S D, Putterill J J, Plummer K M, et al. The carboxylesterase gene family from *Arabidopsis thaliana*[J]. Journal of Molecular Evolution, 2003, 57(5):487-500.

[8] Nomura T, Murase T, Ogita S et al. Molecular identification of tuliposide B-converting enzyme: A lactone-forming carboxylesterase from the pollen of tulip [J]. The Plant Journal, 2015, 83(2):252-262.



- [9] Lee S, Hwang S, Seo Y W, et al. Molecular characterization of the *AtCXE8* gene, which promotes resistance to *Botrytis cinerea* infection[J]. Plant Biotechnology Reports, 2012, 7(1):109-119.
- [10] Gershater M C, Cummins I, Edwards R. Role of a carboxylesterase in herbicide bioactivation in *Arabidopsis thaliana* [J]. The Journal of Biological Chemistry, 2007, 282(29):21460-21466.
- [11] Cummins I, Landrum M, Steel P G, et al. Structure activity studies with exnbiotic substrates using carboxylesterases isolated from *Arabidopsis thaliana* [J]. Phytochemistry, 2007, 68(6):811-818.
- [12] Baudouin E, Charpentreau M, Roby D, et al. Functional expression of a tobacco gene related to the serine hydrolase family-esterase activity towards short-chain dinitrophenyl acylesters[J]. European Journal of Biochemistry, 1997, 248(3):700-706.
- [13] Ko M K, Jeon W B, Kim K S, et al. A *Colletotrichum gloeosporioides*-induced esterase gene of nonclimacteric pepper (*Capsicum annuum*) fruit during ripening plays a role in resistance against fungal infection[J]. Plant Molecular Biology, 2005, 58(4):529-541.
- [14] Hirano K, Ueguchi-Tanaka M, Matsuoka M. GID1-mediated gibberellin signaling in plants[J]. Trends in Plant Science, 2008, 13(4):192-199.
- [15] Voegele A, Linkies A, Muller K, et al. Members of the gibberellin receptor gene family GID1 (*GIBBERELLIN INSENSITIVE DWARF1*) play distinct roles during *Lepidium sativum* and *Arabidopsis thaliana* seed germination[J]. Journal of Experimental Botany, 2011, 62(14):5131-5147.
- [16] Huizinga D H, Omosogbon O, Omery B, et al. Isoprenylcysteine methylation and demethylation regulate abscisic acid signaling in *Arabidopsis* [J]. The Plant Cell, 2008, 20(10):2714-2728.
- [17] Wang J, Chu S, Zhu Y, et al. Positive selection drives neofunctionalization of the UbiA prenyltransferase gene family[J]. Plant Molecular Biology, 2015, 87:383-394.
- [18] Holub E B. The arms race is ancient history in *Arabidopsis*, the wildflower[J]. Nature Reviews Genetics, 2001, 2(7):516-527.
- [19] Lee T H, Tang H, Wang X, et al. PGDD: A database of gene and genome duplication in plants[J]. Nucleic Acids Research, 2013, 41(Database issue):D1152-1158.
- [20] Tamura K, Stecher G, Peterson D, et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0 [J]. Molecular Biology and Evolution, 2013, 30(12):2725-2729.
- [21] 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 Research, 1994, 22(22):4673-4680.
- [22] Lynch M, Crease T J. The analysis of population survey data on DNA sequence variation [J]. Molecular Biology and Evolution, 1990, 7(4):377-394.
- [23] Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions[J]. Molecular Biology and Evolution, 1986, 3(5):418-426.
- [24] Bailey T L, Boden M, Buske F A, et al. MEME SUITE: Tools for motif discovery and searching [J]. Nucleic Acids Research, 2009, 37(Web Server issue):W202-208.
- [25] van de Mortel M, Recknor J C, Graham M A, et al. Distinct bi-phasic mRNA changes in response to Asian soybean rust infection [J]. Molecular Plant-microbe Interactions, 2007, 20(8):887-899.
- [26] Wang Q, Wang J, Yang Y, et al. A genome-wide expression profile analysis reveals active genes and pathways coping with phosphate starvation in soybean [J]. BMC Genomics, 2016, 17(1):192.
- [27] Schmutz J, Cannon S B, Schlueter J, et al. Genome sequence of the palaeopolyploid soybean[J]. Nature, 2010, 463(7278):178-183.
- [28] Lan P, Li W, Wang H, et al. Characterization, sub-cellular localization and expression profiling of the isoprenylcysteine methyl-esterase gene family in *Arabidopsis thaliana* [J]. BMC Plant Biology, 2010, 10:212.
- [29] Gershater M C, Edwards R. Regulating biological activity in plants with carboxylesterases [J]. Plant Science, 2007, 173(6):579-588.
- [30] Nakajima M, Shimada A, Takashi Y, et al. Identification and characterization of *Arabidopsis* gibberellin receptors[J]. The Plant Journal, 2006, 46(5):880-889.
- [31] Ueguchi-Tanaka M, Ashikari M, Nakajima M, et al. *GIBBERELLIN INSENSITIVE DWARF1* encodes a soluble receptor for gibberellin[J]. Nature, 2005, 437(7059):693-698.
- [32] Hirano K, Aya K, Matsuoka M, et al. Molecular determinants that convert hormone sensitive lipase into gibberellin receptor[J]. Protein and Peptide Letters, 2012, 19(2):180-185.