



大豆 *BADH* 基因的生物信息学、多样性及功能分析

张永芳^{1,2,3}, 李富恒¹, 吕怡杰², 邱丽娟³

(1. 东北农业大学 生命科学学院, 黑龙江 哈尔滨 150030; 2. 山西大同大学 生命科学学院, 山西 大同 037009; 3. 中国农业科学院 作物科学研究所/国家农作物基因资源与遗传改良重大科学工程/农业农村部种质资源与生物技术重点开放实验室, 北京 100081)

摘要:为研究大豆中甜菜碱醛脱氢酶(*BADH*)在大豆中的分子机理及生物多样性,通过基因序列同源比对搜索到两个与水稻同源的大豆 *BADH* 基因,利用 Phytozome、Netphosk 3.0 Server 等网站及数据库对其进行生物信息学分析,结合已有的重测序数据分析 1 598 份大豆种质中 *BADH* 基因序列的多态性,并测定突变体的 2AP 含量。结果表明:大豆中有两个与水稻 *BADH* 基因同源性较高的基因序列 *Gm06g186300* 和 *Gm 05g0330550*, 基因全长分别为 3 959 和 6 008 bp, 编码蛋白均属于稳定蛋白质,二者同源性高达 90%。两个蛋白的氨基酸数目、分子量和等电点等非常相似;亲水性和磷酸化位点分布较一致,均无跨膜结构域,在细胞核出现的可能性最大,且二者进化关系极其相近; α 螺旋和无规则卷曲是两者的主要成分,*BADH1* 和 *BADH2* 有 1 个共同的结构域 Pfam-Aldedh, 且分布大致相同。对 1 598 份大豆种质 *BADH* 基因序列突变位点进行分析,有 29 份材料 *BADH2* 基因发生 6 个碱基序列突变导致了移码突变,这种突变可能与大豆的芳香性表型有关。对 *BADH1* 及 *BADH2* 在大豆不同部位的基因表达量分析发现各部位的表达量相差很大,均以根部表达量最高,可能与根部较高的抗旱能力有关。

关键词:大豆;*BADH1*; *BADH2*; 单倍型特性分析; 结构; 突变位点; 芳香性

Bioinformatics, Diversity and Function Analysis of *BADH* Gene in Soybean

ZHANG Yong-fang^{1,2,3}, LI Fu-heng¹, LYU Yi-jie², QIU Li-juan³

(1. School of Life Science, Northeast Agricultural University, Harbin 150030, China; 2. School of Life Science, Shanxi Datong University, Datong 037009, China; 3. Institute of Crop Sciences, Chinese Academy of Agricultural Science/The National Key Facility for Crop Gene Resources and Genetic Improvement/Key Laboratory of Germplasm & Biotechnology (Ministry of Agriculture and Rural Affairs), Beijing 100081, China)

Abstract: In order to study the molecular mechanism and biodiversity of *BADH* in soybean, two soybean *BADH* genes with the same origin as rice were identified by sequence homology analysis blast function of NCBI website. Bioinformatics software Phytozome, Netphosk 3.0 Server and other websites and databases were further utilized to analyze the protein characteristics and structural characteristics of these two genes. The polymorphism of *BADH* gene sequence in 1 598 soybean germplasm was analyzed with resequencing data and the content of 2AP was determined. The results showed that there were two gene sequences *Gm06g186300* and *Gm05G0330550* with high homology with rice *BADH* gene in soybean, the total length of which were 3 959 and 6 008 bp, respectively. It was found that the encoded proteins were stable, and the homology was as high as 90%. *BADH1* and *BADH2* were very similar in the number of amino acids, molecular weight, isoelectric point and so on. The distribution of hydrophilic and phosphorylation sites of *BADH1* and *BADH2* were consistent, and there was no transmembrane domain in the nucleus and the evolutionary relationship between them was very similar. According to the prediction scores of the secondary structures (α Helix, β angle, irregular curl, extended chain) of *BADH1* and *BADH2* amino acid sequences, it can be seen that α helix and random coil were the main components of them. *BADH1* and *BADH2* had only one common domain Pfam aldedh, and the distribution was roughly the same. The mutation sites of *BADH* Gene of 1 598 soybean germplasms were analyzed. Frame shift mutation was found in 6 base sites of *BADH2* gene in 29 soybean germplasms. The aromatic component 2AP of soybean germplasm with gene mutation was identified and the nucleotide sequence of the locus may be related to the aromatic phenotype of soybean. *BADH1* and *BADH2* gene expression analysis in different parts of soybean showed that the expression levels of each part were very different, and the highest expression level was found in the root, which may be related to the higher drought resistance ability of the root.

Keywords: Soybean; *BADH1*; *BADH2*; Haplotype characteristic analysis; Structure; Mutation site; Aromaticity

大豆(*Glycine max*)在中国已有几千年的种植历史,因富含人体必须的氨基酸和脂肪酸,可提高人体免疫力、抗衰老、促进人体消化吸收而被广泛食用。随着人们生活水平的提高,人们对大豆的品质

需求愈来愈高,香味是影响大豆食用品质的重要特征之一,较多研究表明,大豆的香味不同是由于甜菜碱醛脱氢酶基因(*betaine aldehyde dehydrogenase2*, *BADH2*)抑制了香味来源的主要物质 2AP 的

收稿日期:2020-08-08

基金项目:“十三五”重点研发计划课题(2016YFD0100201)。

第一作者:张永芳(1982—),女,博士,讲师,主要从事植物学研究。E-mail:zyf_208@163.com。

通讯作者:邱丽娟(1963—),女,博士,研究员,主要从事植物遗传育种研究。E-mail:qiuqilujuan@caas.cn;

李富恒(1963—),男,博士,教授,主要从事植物学研究。E-mail:Lifuheng1963@126.com。

生物合成^[1-2],当其变异后能够显著促进香气成分2-乙酰-1-吡咯啉(2-acetyl-1-pyrroline,2AP)的合成,2AP是一种类似爆米花香的挥发性化合物,这种香气成分存在于各种植物^[3-4]、动物^[5]、微生物^[6-7]及食品中^[8-9]。2AP含量高低直接影响消费者的偏好和接受度,含量较高的豆制品的市场价值翻倍。Arikiti等^[10]发现大豆中的*BADH2*编码的甜菜碱脱氢酶与水稻*Os2AP*同源,促进大豆2AP的生物合成,芳香大豆同时缺乏甜菜碱脱氢酶活性。Juwattanasomran等^[11]遗传分析表明鲜食豆Kaori香味是*BADH2*基因控制的,其第10外显子存在1个SNP的无义突变,使得*BADH2*蛋白功能缺失,合成芳香物质2AP,导致大豆具有稻香味^[11]。Juwattanasomran等^[12]以Chammame为材料,发现*BADH2*外显子存在2 bp缺失,导致移码突变及蛋白提前终止,致使蛋白无功能,产生香味物质2AP。研究也表明*BADH*是一种渗透调节物质,通过保护细胞膜蛋白上的一些大分子帮助植物对抗逆境胁迫来提高植物的适应能力,从而抵抗干旱。生物学家Weretilnyk等^[13]在1990年首次克隆了菠菜的*BADH*基因。Hanson等^[14]和Arakaw等^[15]应用转基因技术在植物中表达正常功能的*BADH*基因,发现甘氨酸甜菜碱含量会明显提高,说明甘氨酸甜菜碱的合成与*BADH*有关,是合成甘氨酸甜菜碱的重要酶。郭北海等^[16]和郭岩等^[17]将*BADH*基因转入小麦和山菠菜后均获得*BADH*基因高效表达的植株,并发现转基因植株高度抗旱。说明*BADH*在控制香气的形成及渗透调节中起着重要作用,研究其基因及相应的蛋白结构及生物学特性、多样性对改良大豆品质、培育抗旱大豆新品种、增加粮食产量、促进农业可持续发展具有重要意义。目前有关大豆*BADH*基因多样性的研究较少。本研究利用生物信息学同源比对方法得到大豆中的*BADH*基因,分析其基因结构及在不同组织器官中表达量、理化性质、结构差异及进化关系。分析1 598份大豆种质*BADH*基因序列的突变位点,以分析大豆中该类基因的多样性。测定突变材料中*BADH2*基因催化合成的芳香成分2AP的含量,以Juwattanasomran等^[11]报道的香味及非香大豆材料为对照,对比基因突变材料和普通材料中2AP的含量。研究旨在为深入开展*BADH*单倍型特性分析、酶学特性的分子机理、基因多样性的研究提供重要理论依据。

1 材料与方法

1.1 材料

NCBI(<https://www.ncbi.nlm.nih.gov/>)网站中的植物*BADH*基因序列。本实验室已完成重测序的1 598份国内外大豆种质。

1.2 方法

1.2.1 基因进化分析 在NCBI(<https://www.ncbi.nlm.nih.gov/>)网站使用水稻抗旱基因*BADH*编码序列同源比对得到不同植物的*BADH*氨基酸序列,采用最大似然法构建植物*BADH*基因序列的系统进化树,分析大豆与其它植物*BADH*基因的亲缘关系。

1.2.2 大豆*BADH*同源基因编码序列分析 从Phytozome网站(<https://phytozome.jgi.doe.gov>)下载大豆*BADH1*及*BADH2*基因序列,绘制两个基因编码序列的结构图,利用网站(<http://multalin.toulouse.inra.fr/multalin/>)对上述两个序列进行多序列比对。

1.2.3 二级和三级结构分析 通过PSIPRED(<http://bioinf.cs.ucl.ac.uk/psipred/>)在线网站预测和分析*BADH1*和*BADH2*蛋白质二级结构;利用CDD(<https://www.ncbi.nlm.nih.gov/cdd>)网站分析蛋白功能结构域;利用SWISS-MODEL(<https://swissmodel.expasy.org/>)构建两个蛋白质的三级结构模型。

1.2.4 蛋白质特性分析 利用ProtParam(<https://web.expasy.org/protparam/>)分析*BADH*蛋白质分子量、稳定性、等电点和亲水性等。基于HMM法,在TMHMM(<http://www.cbs.dtu.dk/services/TMHMM-2.0/>)网站分析蛋白质的跨膜区。利用PSORT II prediction(<https://psort.hgc.jp/form2.html>)软件预测*BADH*蛋白的亚细胞定位。利用NetPhos 3.1Server软件预测*BADH*蛋白的磷酸化位点。

1.2.5 大豆*BADH*的多态性分析 以Phytozome网站下载的Williams 82大豆*BADH1*和*BADH2*基因编码区序列为参考序列,根据天津诺和致源公司对1 598个大豆个体采用全基因组鸟枪法进行的全基因组重测序结果,筛选获得1 598个大豆的*BADH1*及*BADH2*基因的序列,并对其外显子区域核苷酸和编码氨基酸单倍型发生的变异和数目进行统计分析。

1.2.6 *BADH2*突变材料2AP含量测定 以Juwattanasomran等^[11]报道的香味材料KaoriHime及非香材料CM60为对照,选取这29份材料中的27份(有2份未取到)进一步测定和对比*BADH2*可能催化合成的2AP的含量。

磨制成熟期大豆豆粉,准确称取0.4 g,剪至直径2 cm大小,置于装有1.5 mL酒精的10 mL离心管中,用漩涡混合器(型号:XH-C)震荡混匀后超声萃取20 min,静置2 h后于低温离心机4 ℃ 12 000 r·min⁻¹离心10 min,取1 mL上清并用25 μm滤膜过滤于气质小瓶中(32 mm×11 mm)。

使用气相色谱质谱仪测定2AP含量,GC-MS 2010色谱柱,DB-WAX毛细管柱(30 m×0.18 mm

×0.25 μm),柱温升温程序为 60 ℃ 保持 2 min, 10 ℃·min⁻¹ 升至 100 ℃,然后以 30 ℃·min⁻¹ 升至 230 ℃,保持 5 min;进样口压力为 78.0 kPa,进样口温度为 180 ℃;载气为高纯(纯度>99.999%)氦气;恒压不分流进样,进样量为 1 μL。MS 条件:电子轰击(EI)离子源,离子源温度为 200 ℃;离子化能量为 70 eV;接口温度为 180 ℃;检测器电压为 0.1 kV,全扫描方式,扫描范围为 m/z 35~500。检测普库为 NIST 库。每份材料测定 3 次,取平均值,计算 27 份大豆种质 2AP 含量。

进一步对材料进行分级,分级标准为:先计算出参试品种 2AP 的总体平均数(\bar{x})和标准差(σ),然后按第 1 级 $[x_i \geq \bar{x} + \sigma]$ 、第 2 级 $[x_i \geq \bar{x} + 0.1\sigma]$ 、第 3 级 $[x_i < \bar{x}]$ 划分,属于 1 级的为香型材料。

1.2.7 *BADH* 同源基因的序列在各组织器官的表达分析 在 Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>)中搜索 *BADH* 基因在大豆各组织器官的表达量等并进行比较,得出其差异性。

1.3 数据分析

采用 Excel 2013 对数据进行处理分析。

2 结果与分析

2.1 大豆 *BADH* 基因序列和编码蛋白结构分析

2.1.1 基因进化分析 使用水稻抗旱基因 *BADH*

编码序列在 NCBI 网站同源比对得到 19 个物种的 *BADH* 氨基酸序列,构建的系统进化树如图 1 所示。19 种不同物种包括:水稻(*Oryza sativa Japonica Group*)、玉米(*maize*)、短花药野生稻(*Oryza brachyantha*)、短柄草(*Brachypodium distachyon*)、结缕草(*Zoysia tenuifolia*)、羊草(*Leymus chinensis*)、小麦(*Triticum aestivum*)、大麦(*Hordeum vulgare subsp*)、三羊草(*Aegilops tauschii*)、陆地棉(*Gossypium hirsutum*)、拟南芥(*Arabidopsis thaliana*)、木本棉(*Gossypium arboreum*)、欧洲油菜(*Brassica napus*)、萝卜(*Raphanus sativus*)、亚麻芥(*mitochondrial-like Camelina sativa*)、醉蝶花(*Tarenaya hassleriana*)、可可树(*Theobroma cacao*)、大豆乙醛脱氢酶(*Glycine max BADH1*)、大豆过氧化物酶乙醛脱氢酶(*Glycine max BADH2*)。按照种属关系可分为两支,分别为双子叶和单子叶植物,大豆 *BADH1* (*Glyma. 06g186300*) 和 *BADH2* (*Glyma. 05g033500*) 与双子叶植物可可树、陆地棉、木本棉、拟南芥、萝卜、亚麻芥、醉蝶花同源性较高聚为一类;单子叶植物三羊草、小麦、欧洲油菜、玉米、短花药野生稻、大麦、水稻、羊草、短柄草与大豆 *BADH* 基因亲缘关系较远,聚为一类。且随着时间变化双子叶植物 *BADH* 遗传变异度小于单子叶植物。

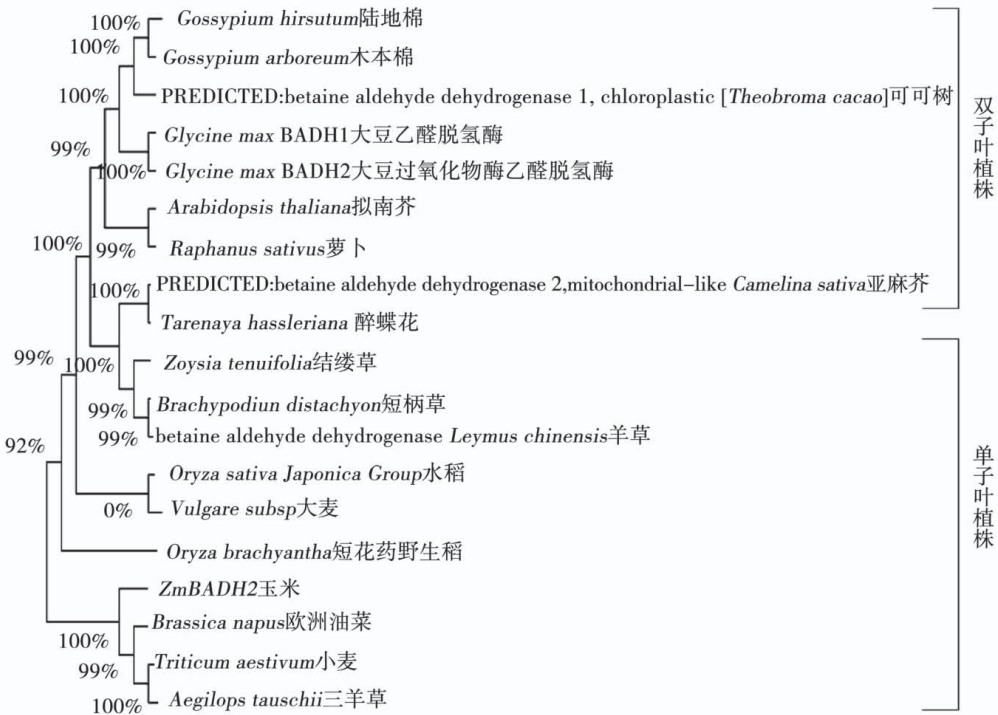


图 1 植物 *BADH* 氨基酸序列系统进化树

Fig. 1 Phylogenetic tree of plant *BADH* amino acid sequences

2.1.2 *BADH1* 与 *BADH2* 基因编码序列比较 大豆 *BADH1* 及 *BADH2* 基因结构如图 2 所示,两个基

因分别位于 5 号和 6 号染色体,基因全长分别为 6 008和 3 959 bp,均为断裂基因,编码区核苷酸序

列长度分别为1 512 和1 467 bp,编码蛋白长度分别为463 和448 个氨基酸,基因编码区分别有15 和16 个外显子,基因结构一致性较高。两个基因编码的蛋白序列同源性达90% (图3)。

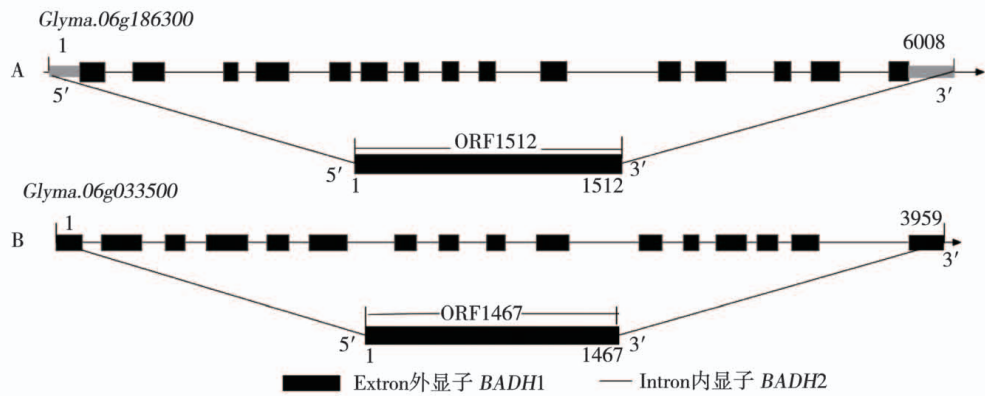
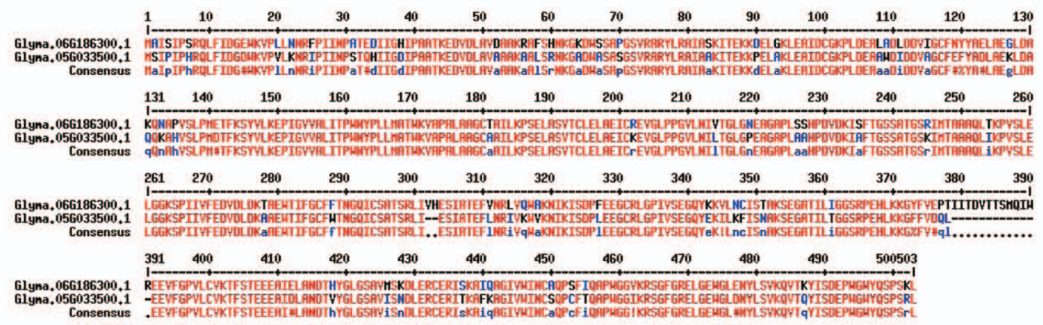


图2 *BADH1* (A) 和 *BADH2* (B) 基因结构示意图

Fig. 2 Schematic diagram of *BADH1* (A) and *BADH2* (B) gene structure

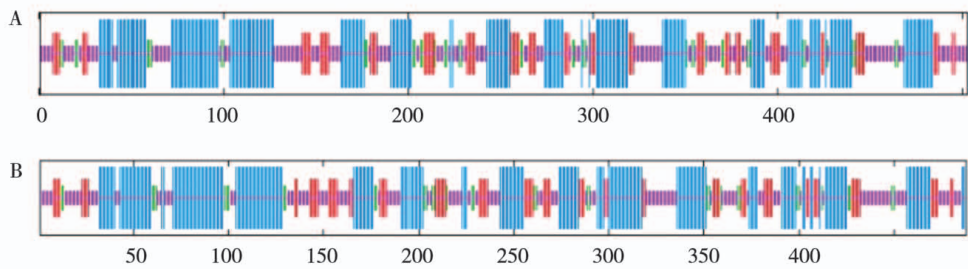


红色部分为序列相同部位。
The red part is the same part of the sequence.

图3 *BADH1* 和 *BADH2* 基因编码序列比对图

Fig. 3 Sequence alignment of *BADH1* and *BADH2* encoding sequences

2. 1. 3 *BADH1* 和 *BADH2* 蛋白序列二级结构及结构域分析 大豆 *BADH1* 和 *BADH2* 蛋白序列二级结构预测结果如图4所示,在两个多肽链中 α 螺旋分别出现44.15%和43.65%, β 转角分别出现8.15%和7.17%,无规则卷曲分别出现34.99%和32.58%,延长链分别出现15.31%和16.60%。结果说明 α 螺旋和无规则卷曲是 *BADH1* 和 *BADH2* 蛋白中的主要成分。



蓝色区域代表 α 螺旋;绿色区域代表 β 转角;橙色区域代表无规则卷曲;红色区域为延长链。
The blue area represents the α helix; The green area represents the β angle; The orange area represents the irregular curl; The red area represents the extended chain.

图4 *BADH1* (A) 和 *BADH2* (B) 氨基酸序列二级结构预测图

Fig. 4 Secondary structure prediction diagram of *BADH1* (A) and *BADH2* (B) amino acid sequence

大豆 *BADH1* 和 *BADH2* 氨基酸序列功能结构域分析结果如图5所示,两个多肽链属于ALDH-SF超家族、醛脱氢酶家族,该家族具有氧化还原酶活性,可以催化醛基氧化为羧基。*BADH1* 的3个功能结构域分别为: Feature 1, catalytic residues [active site] (258 – 293); Feature 2, catalytic residues [active

site] (160 – 293) ; Feature 3, NAD (P) binding site [chemical binding site] (155 – 293) 。*BADH2* 的3 个功能结构域分别为: Feature 1, catalytic residues [active site] (250 – 293) ; Feature 2, catalytic residues [active site] (160 – 293) ; Feature 3, NAD (P) binding site [chemical binding site] (154 – 293) 。

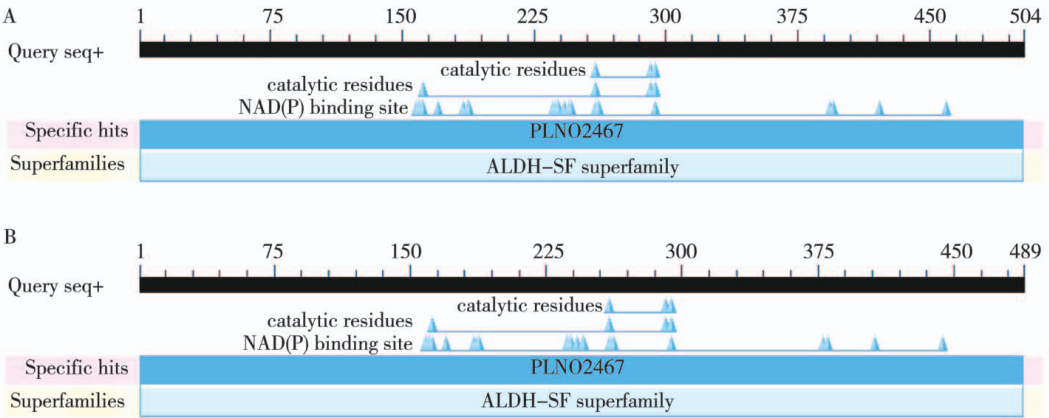


图 5 大豆 *BADH1* (A) 和 *BADH2* (B) 蛋白的结构域预测

Fig. 5 Domain prediction of soybean *BADH1* (A) and *BADH2* (B) proteins

2. 1. 4 *BADH1* 和 *BADH2* 蛋白序列三级结构预测

BADH1 和 *BADH2* 蛋白质的三级结构模型如图 6 所示,两个基因的蛋白质三级结构稍有差异,与该

蛋白质的同源模型 PDB:1jm7. 1A 对比,两者一致性达 31. 25% 。

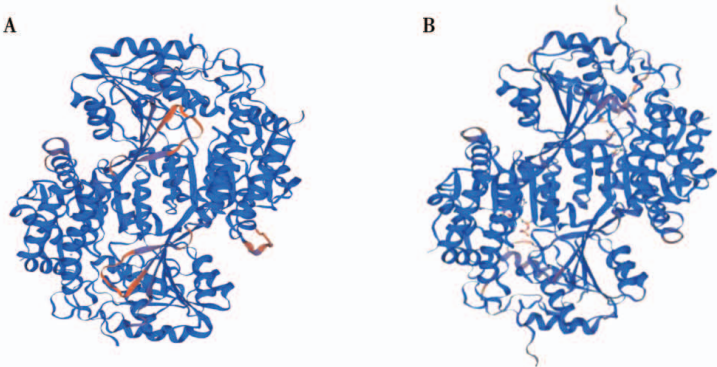


图 6 *BADH1* (A) 和 *BADH2* (B) 蛋白三级结构模型

Fig. 6 Tertiary protein structure model of *BADH1* (A) and *BADH2* (B)

2. 2 大豆 *BADH1* 和 *BADH2* 蛋白特性分析

2. 2. 1 理化性质分析 蛋白序列理化性质分析结果如表 1 所示,*BADH1* 和 *BADH2* 蛋白的氨基酸数目、分子量、pI 和不稳定指数等相似度都比较高,它们的氨基酸数目相差 15,分子量相差 1 867. 9,pI 相

差 0. 4,不稳定指数相差 6. 65,脂肪指数相差 3. 56,两者均表现出稳定性,这是由遗传的同源性决定的,在 *BADH1* 和 *BADH2* 蛋白序列中含量较高的氨基酸均为 Ala、Glu、Leu、Ile、Arg 和 Asn。

表 1 *BADH1* 和 *BADH2* 蛋白理化性质分析

Table 1 Physical and chemical properties of *BADH1* and *BADH2* protein

蛋白 Protein	编码氨基酸数目 Number of encoding amino acids	分子量 Molecular weight	pI	不稳定指数 Instability index	脂肪指数 Aliphatic index	含量最丰富的氨基酸 The most abundant amino acid						
						Ala	Glu	Gly	Leu	Ile	Arg	Asn
<i>BADH1</i>	503	54739. 84	5. 2	33. 55	94. 27	9. 3	7. 8	7. 8	8. 7	8. 2	3. 4	3. 2
<i>BADH2</i>	488	52871. 94	4. 8	26. 90	97. 83	11. 3	7. 0	7. 6	9. 2	8. 4	3. 1	2. 7

2.2.2 亲疏水性分析 蛋白质的亲水性是维持蛋白质高级结构稳定性的重要因素之一。应用 ProtScale 在线网站对大豆的 BADH1 和 BADH2 氨基酸序列进行亲疏水性分析的结果如图 7 所示,二者小于 0 的数值稍多于大于 0 的数值,其中,BADH1 蛋白的最高值为 2.256,在第 155,156 个氨基酸处,疏水性最强;最低值为 -2.322,在第 63 个氨基酸处,亲水性较强;BADH2 蛋白的最高值为 2.378,在第 180 个氨基酸处,疏水性最强;最低值为 -2.0,在第 365 和 366 个氨基酸处,亲水性较强。说明

BADH1 和 BADH2 蛋白属于亲水蛋白。

2.2.3 跨膜结构域分析 进一步利用在线工具 TMHMM 2.0 Server 对大豆跨膜结构域进行的预测分析结果表明,BADH1 和 BADH2 蛋白均没有跨膜结构域(图 8)。说明该蛋白定位于细胞质基质或细胞器基质中,是一种非分泌蛋白或膜蛋白,不能引导蛋白质的跨膜运输。

2.2.4 亚细胞定位预测 利用 PSORT II prediction 软件对 BADH 蛋白亚细胞定位的预测结果显示,二者均在细胞核出现的可能性最大(图 9)。

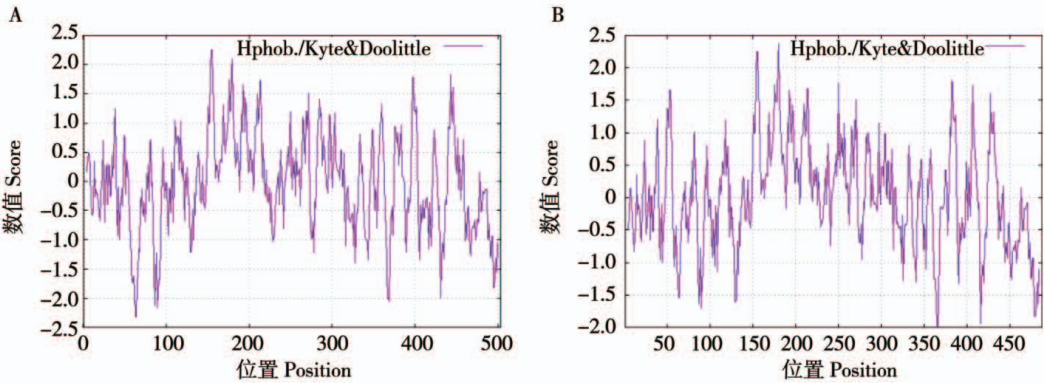


图 7 BADH1 (A) 和 BADH2 (B) 的亲疏水性分析
Fig. 7 Hydrophobic analysis of BADH1 (A) and BADH2 (B)

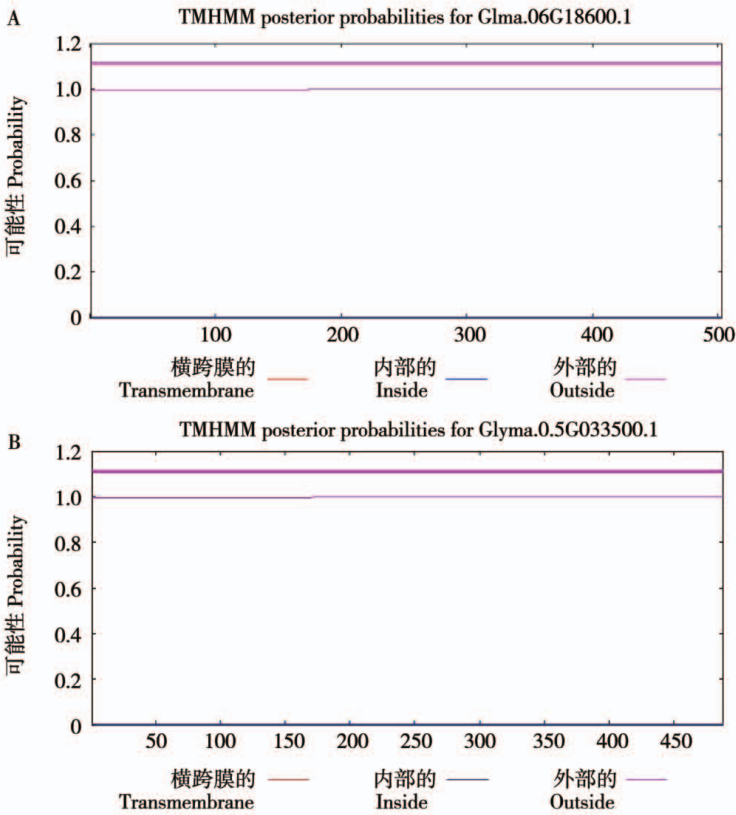


图 8 BADH1 (A) 和 BADH2 (B) 蛋白的跨膜结构域的预测
Fig. 8 Projections of the transmembrane domains of BADH1 (A) and BADH2 (B)

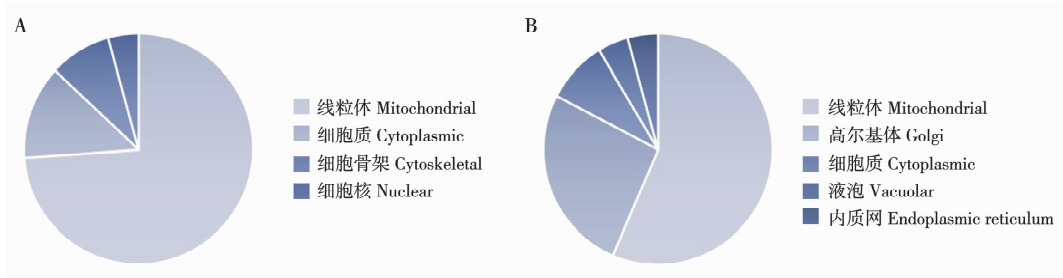


图 9 BADH1 (A) 和 BADH2 (B) 蛋白的亚细胞定位

Fig. 9 Subcellular localization of BADH1 (A) and BADH2 (B) proteins

2. 2. 5 磷酸化位点分析 蛋白质磷酸化是生物体内重要的共价修饰之一,与信号转导、细胞周期、生长发育及癌症机理有重要关系,预测其磷酸化位点有重要意义。运用 Netphosk 3.0 Server 在线工具对大豆磷酸化位点的预测结果如图 10 所示。

BADH1 的丝氨酸磷酸化位点有 38 个,分别为 S4、S7、S59、S67、S68、S72、S83、S137、S145、S187、S191、S226、S227、S235、S239、S240、S244、S258、S265、S295、S298、S305、S324、S338、S350、S354、S364、S386、S405、S423、S427、S437、S453、S465、S481、S490、S499 和 S501;苏氨酸磷酸化位点有 28 个,分别位 T32、T43、T86、T142、T159、T169、T181、T193、T215、T237、T242、T248、T254、T278、T282、T289、T297、T308、T351、T358、T378、T381、T384、T385、T403、T406、T417、T486;酪氨酸磷酸化位点有 11 个,分别为 Y77、Y120、Y121、Y146、Y163、Y342、

Y373、Y419、Y479、Y488、Y497。共 77 个磷酸化位点,其中 S88、S354、S405 和 Y342 数值偏高,达到了 0.99 以上。

BADH2 的丝氨酸磷酸化位点有 32 个,分别为 S2、S31、S59、S68、S70、S72、S137、S145、S187、S191、S239、S240、S244、S258、S265、S295、S298、S303、S322、S336、S348、S352、S362、S390、S408、S412、S435、S450、S466、S475、S484、S486;苏氨酸磷酸化位点有 23 个,分别 T32、T43、T86、T142、T159、T169、T193、T215、T237、T242、T248、T282、T289、T297、T306、T356、T359、T388、T391、T402、T442、T446、T471;酪氨酸磷酸化位点有 9 个,分别为 Y77、Y121、Y146、Y163、Y340、Y404、Y464、Y473、Y482。共 64 个磷酸化位点,其中 Y340、S352 和 S390 数值偏高,达 0.98 以上。可以看出两种蛋白磷酸化位点分布一致性比较高。

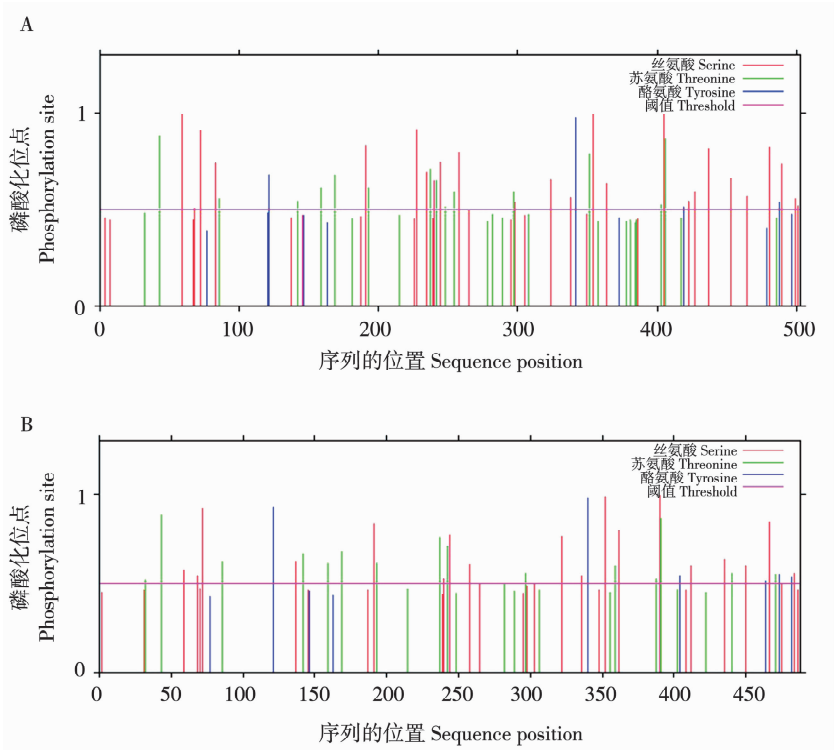


图 10 BADH1 (A) 和 BADH2 (B) 磷酸化位点预测

Fig. 10 Prediction of BADH1 (A) and BADH2 (B) phosphorylation sites

2.3 大豆 BADH 变异分析

2.3.1 大豆 BADH 多态性分析 对 1 598 份大豆重测序数据中 BADH1 及 BADH2 核苷酸和编码氨基酸序列变异情况的分析结果如表 2 所示,1 598 份大豆品种 *BADH1* 基因外显子区域产生的突变包括 4 种非同义突变和 1 种同义突变;*BADH2* 基因外显子区域产生的突变包括 8 种非同义突变、1 种同义突变、6 种移码突变。值得注意的是有 29 份材料的 *BADH2* 基因外显子区域的 6 个位点(包括 41 位、

44 位、53 位、57 位、60 位、64 位)碱基发生了移码突变,推测可能这 6 个位点与该基因对应的抗旱或香味性状有关。

2.3.2 *BADH2* 变异材料 2AP 含量分析 以香味材料 KaoriHime 及非香材料 CM60 为对照,对 27 份 *BADH2* 变异材料中 2AP 含量的测定结果如表 3 所示,*BADH2* 变异材料的 2AP 含量均高于普通材料,说明这些 *BADH2* 位点的突变可能导致了大豆芳香性的改变。

表 2 大豆 BADH 的多态性分析
Table 2 Polymorphism analysis of soybean BADH

基因 Gene	突变类型 Mutation type	碱基变化 Base change			氨基酸序列变化 Amino acid sequence change			材料数目 The number of material
		位点 Site	原碱基 The original base	突变碱基 Mutant base	位点 Site	原氨基酸 The original amino acid	突变氨基酸 Mutant amino acid	
<i>BADH1</i>	同义突变	273	A	T	91	亮氨酸	亮氨酸	45
	非同义突变	370	C	T	124	谷氨酸	赖氨酸	1
	非同义突变	388	G	C	130	丙氨酸	甘氨酸	2
	非同义突变	1334	A	T	445	异亮氨酸	赖氨酸	3
	非同义突变	1495	G	C	499	组氨酸	天冬氨酸	15
<i>BADH2</i>	非同义突变	69	C	G	23	谷氨酸	天冬氨酸	2
	非同义突变	194	A	T	65	苯丙氨酸	酪氨酸	1
	非同义突变	288	C	G	96	谷氨酰胺	谷氨酸	2
	非同义突变	452	T	C	151	谷氨酸	甘氨酸	1
	同义突变	666	T	C	222	甘氨酸	甘氨酸	1
	非同义突变	817	C	T	273	天冬氨酸	丝氨酸	2
	非同义突变	1399	G	T	467	谷氨酰胺	赖氨酸	1
	非同义突变	1430	A	T	477	亮氨酸	谷氨酰胺	3
	非同义突变	1480	G	T	494	脯氨酸	苏氨酸	1
	移码突变	41	CCGTCT	C	13	异亮氨酸、天冬氨酸	缺失	29
	移码突变	44	T	TG	14	天冬氨酸	插入谷氨酸	29
	移码突变	53	ACTTCCAG	A	17	天冬氨酸 色氨酸	缺失	29
	移码突变	57	G	GAATT	19	脯氨酸	插入天冬酰胺	29
	移码突变	60	GAC	G,AAC	20	脯氨酸	缺失	29
	移码突变	64	TG	T	21	亮氨酸	缺失	29

表 3 *BADH2* 变异大豆 2AP 含量
Table 3 2AP content of different soybean materials with mutant *BADH2*

编号 No.	基因突变材料 Gene mutation material	2AP 含量 2AP content /(mg·L ⁻¹)	香型 Fragrance type	编号 No.	基因突变材料 Gene mutation material	2AP 含量 2AP content /(mg·L ⁻¹)	香型 Fragrance type
1	KaoriHime	0.36	香	15	哈 13-2958	0.74	香
2	CM60	0.05	非香	16	克 11-1619	0.48	香
3	合丰 44	0.65	香	17	红丰 11	0.25	香
4	合丰 38	0.64	香	18	争光 1 号	0.55	香
5	衢鲜 1 号	0.46	香	19	吉林 3 号	0.17	香

续表 3

编号 No.	基因突变材料 Gene mutation material	2AP 含量 2AP content /(mg·L ⁻¹)	香型 Fragrance type	编号 No.	基因突变材料 Gene mutation material	2AP 含量 2AP content /(mg·L ⁻¹)	香型 Fragrance type
6	临河小粉青	0.40	香	20	青脸豆	0.61	香
7	石官营青豆	0.35	香	21	南通黄油果子	0.32	香
8	合农 97	0.52	香	22	南农 99-10	0.46	香
9	合农 70	0.62	香	23	18NQ0849	0.77	香
10	蒙豆 33	0.67	香	24	红面豆	0.40	香
11	合丰 30	0.49	香	25	通化平定香	0.67	香
12	吉育 701	0.48	香	26	黑生 101	0.88	香
13	南农大红豆	0.50	香	27	绥农 82	0.65	香
14	哈 14-2028	0.73	香	28	吉林 21	0.25	香

2.4 大豆 *BADH* 在各组织器官的表达分析

不同部位的基因相对表达量分析结果如图 11 所示,*BADH1* 和 *BADH2* 基因在植株各部位的表达量相差很大,*BADH1* 在植株各部位表达量均较

BADH2 高,且在根中表达量最高;*BADH2* 在植株各部位表达量相差较小,在根毛中表达量最高。结果表明 *BADH1* 和 *BADH2* 在根部的表达量都相对较高,说明两个基因可能与根部较高的抗旱能力有关。

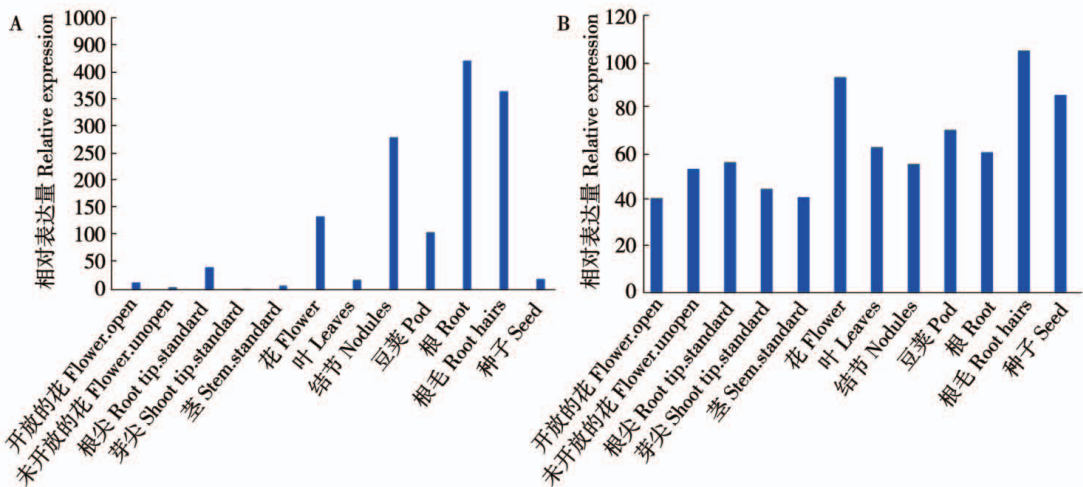


图 11 大豆不同部位 *BADH1* (A) 和 *BADH2* (B) 基因相对表达量

Fig. 11 Relative expression levels of *BADH1* (A) and *BADH2* (B) genes in different parts of soybean

3 讨论

BADH 是甜菜碱合成的关键酶之一,也是很好的胁迫抗性基因之一,可以催化有毒的甜菜碱醛转化为无毒性的甜菜碱,尤其在干旱、盐碱等逆境胁迫下 *BADH* 大量表达合成并积累甜菜碱,从而通过调节细胞内渗透压维持细胞膜稳定性,保护细胞内酶活性^[18]。在大豆中存在两种编码甜菜碱脱氢酶的同源基因,分别为 *BADH1* 和 *BADH2*,系统进化树显示 *BADH1* 和 *BADH2* 来源于同一个祖先,它们在结构和功能上具有明显的相似性,均编码甜菜碱醛脱氢酶。*BADH* 基因编码氨基酸序列的系统进化树显示该双子叶植物 *BADH* 基因聚为一类,这与 Arikiti 等^[18] 利用 *BADH* 基因的蛋白序列进行系统进化分析表明近缘物种聚在一起的结果非常相似,说

明该基因在氨基酸和蛋白质水平上进化模式一致,且在双子叶和单子叶植物中发生了较大变异。

BADH 蛋白在大豆中被推测定位于细胞核,但在甜菜中该蛋白被定位于叶绿体中,在玉米、高粱等植物中定位于过氧化物酶体中^[19],说明不同物种 *BADH* 亚细胞定位并不同,可能与其功能有关。*BADH* 蛋白不含跨膜结构域,属于稳定的亲水蛋白,这与弄庆媛等^[20] 报道的甜菜菠菜、山菠菜、辽宁蓬碱、玉米中的研究结果均一致,说明 *BADH* 蛋白理化性质在不同物种中表现相对保守。

BADH1 和 *BADH2* 是大豆中与水稻编码甜菜碱醛脱氢酶 *BADH* 基因同源的两个基因。本研究分析了二者在香味成分中的多态性,发现 *BADH2* 基因的突变导致酶活性降低,而大豆香味物质成分 2AP 的含量增加。而 *BADH1* 与香味无关,其机理可能是

BADH2 基因功能丢失导致 γ 氨基丁醛和吡咯啉的增加导致 2AP 积累^[21], 这与水稻中报道的只有 *BADH* 的功能与稻香有关相似^[22]。

4 结 论

本研究利用生物信息学方法,通过水稻 *BADH* 基因同源比对分析,得到大豆 *BADH1* 和 *BADH2* 基因,二者编码的蛋白质序列一致性较高,可以达到 90%。最大似然法构建系统进化树表明二者极其相近,可作为生物遗传分化和分子进化中的重要依据;这两个基因均为断裂基因,分子量、等电点、氨基酸数目均表现出相对一致性;二级结构分析可知 α 螺旋和无规则卷曲是两者二级结构中的主要成分,且两者组成和结构也极其相似;两个蛋白质三级结构有差异,但相似性也比较高;二者蛋白性质比较稳定,氨基酸都表现为亲水性,亚细胞定位可知两者于细胞核出现的可能性最大,无跨膜结构域;根据单倍型及表型分析,*BADH2* 可能与大豆香味性状功能有关。本研究结果为后期深入开展 *BADH* 单倍型、基因多样性分析等研究提供重要理论依据,同时也为培育大豆抗旱、香味新品种奠定了基础。

参考文献

[1] Khin M M, Brigitte C, Ange-Marie R, et al. Specific patterns of genetic diversity among aromatic rice varieties in Myanmar[J]. Rice, 2012, 5(20): 1-13.

[2] Zhang Z, Xu M, Yang Y, et al. *BADH2*, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-Acetyl-1-Pyrroline, a major component in rice fragrance[J]. The Plant Cell, 2008, 20(7): 1850-1861.

[3] Widjaja R, Craske J, Wootton M. Comparative studies on volatile components of non-fragrant and fragrant rices[J]. Journal Science of Food Agriculture, 1996, 70(2): 151-161.

[4] Yoshihashi T. Quantitative analysis on 2-acetyl-1-pyrroline of an aromatic rice by stable isotope dilution method and model studies on its formation during cooking[J]. Journal Food Science, 2002, 67(2): 619-622.

[5] Brahmachary R L, Sarkar P, Dutta J. The aroma of rice ... and tiger[J]. Nature, 1990, 26:344.

[6] Adams A, De K N. Formation of pyrazines and 2-acetyl-1-pyrroline by *Bacillus cereus*[J]. Food Chemistry, 2007, 101(3): 1230-1238.

[7] Snowdon E M, Bowyer M C, Grbin P R, et al. Mousy off-flavor: A review[J]. Journal of Agricultural and Food Chemistry, 2006, 54(18): 6465-6474.

[8] Schieberle P, Wener G. Potent odorants of the wheat bread crumb differences to the crust and effect of a longer dough fermentation [J]. Zeitschrift für Lebensmittel-Untersuchung und Forschung, 1991, 192(2): 130-135.

[9] Buttery R G, Ling L C, Juliano O B, et al. Cooked rice aroma

and 2-acetyl-1-pyrroline [J]. Journal of Agricultural and Food Chemistry, 1983, 31: 823-826.

[10] Arikiti S, Yoshihashi T, Wanchana S, et al. Deficiency in the amino aldehyde dehydrogenase encoded by *GmAMADH2*, the homologue of rice Os2AP, enhances 2-acetyl-1-pyrroline biosynthesis in soybeans (*Glycine max* L.) [J]. Plant Biotechnol Journal, 2010, 122(2): 311-316.

[11] Juwattanasomran R, Somta P, Chankaew S. A SNP in *GmBADH2* gene associates with fragrance in vegetable soybean variety ‘Kaori’ and SNAP marker development for the fragrance[J]. Theoretical and Applied Genetics, 2011, 122: 533-541.

[12] Juwattanasomran R, Somta P, Kaga A. Identification of a new fragrance allele in soybean and development of its functional marker[J]. Molecular Breeding, 2012, 29: 13-21.

[13] Weretilnyk E A, Hanson A D. Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought[J]. Proceedings of the National Academy of Sciences of the United States of America, 1990, 87(7): 2745-2749.

[14] Hanson A D, May A M, Grumet R, et al. Betaine synthesis in chenopods: Localization in chloroplasts [J]. Proceedings of the National Academy of Sciences of the USA, 1985, 82: 3678-3682.

[15] Arakawa K, Takabe T, Sugiyama T, et al. Purification of betaine aldehyde dehydrogenase from spinach leaves and preparation of its antibody [J]. Japanese Biochemical Society, 1987, 101: 1485-1488.

[16] 郭北海, 张艳敏, 李洪节, 等. 甜菜碱醛脱氢酶(*BADH*)基因转化小麦及表达[J]. 植物学报, 2000, 42(3): 279-283. (Guo B H, Zhang Y M, Li H J, et al. Transformation of wheat with a gene encoding for the betaine aldehyde dehydrogenase (*BADH*) [J]. Acta Botanica Sinica, 2000, 42(3): 279-283.)

[17] 郭岩, 张莉, 肖岗, 等. 甜菜碱醛脱氢酶基因在水稻中的表达及转基因植株的耐盐性研究[J]. 中国科学, 1997, 27(2): 151-155. (Guo Y, Zhang L, Xiao G, et al. Expression of Betaine-aldehyde dehydrogenase gene in rice and salt tolerance of transgenic plants [J]. Science in China, 1997, 27(2): 151-155.)

[18] Arikiti S, Yoshihashi T, Wanchana S, et al. Deficiency in the amino aldehyde dehydrogenase encoded by *GmAMADH2*, the homologue of rice Os2AP, enhances 2-acetyl-1-pyrroline biosynthesis in soybeans (*Glycine max* L.) [J]. Plant Biotechnology Journal, 2011, 9(1): 75-87.

[19] Wood A J, Saneoka H, Rhodes D, et al. Betaine aldehyde dehydrogenase in sorghum [J]. Plant Physiology, 1996, 110: 1301-1308.

[20] 弄庆媛, 麦秀英, 周虹, 等. 植物甜菜碱醛脱氢酶的生物信息学分析[J]. 北方园艺, 2012(8): 120-124. (Nong Q Y, Mai X Y, Zhou H, et al. Bioinformatics analysis of betaine aldehyde dehydrogenase in plants [J]. Northern Horticulture, 2012(8): 120-124.

[21] Bradbury L M T, Henry R J, Jin Q S, et al. A perfect marker for fragrance genotyping in rice [J]. Molecular Breed, 2005, 16: 279-283.

[22] He Q, Yu J, Kim T S, et al. Resequencing reveals different domestication rate for *BADH1* and *BADH2* in rice (*Oryza sativa*) [J]. Plos One, 2015, 10(8): e0134801.