Isolation and Molecular Marker Detection of Badh2 Gene from Aromatic Rice Germplasm Resources in Southern Henan

Bo Peng*
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Yu Zhu
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Zi-Yu Wang
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Juan Peng
Xinyang Station of Plant Protection and Inspection, Xinyang, China

Lu-Lu He
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Xia-Yu Tian
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Zi-Yue Liu
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Ling Sun
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Ying Wang
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Xiao-Hua Song
Xinyang Academy of Agricultural Science, Xinyang, China

Hui-Long Li
Xinyang Academy of Agricultural Science, Xinyang, China

Yan-Yang Sun
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Rui-Hua Pang
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Jin-Tao Li
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Quan-Xiu Wang
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

*Corresponding Author
Wei Zhou
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Hong-Yu Yuan
College of Life Sciences and Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains, Xinyang Normal University, Xinyang, China

Abstract

The production of aroma in aromatic rice is due to the increase of 2-acetyl-1-pyrroline (2-AP) precursor substances caused by the functional deletion of Badh2 gene on chromosome 8, and the accumulation of 2AP makes rice produce aroma. In this study, Badh2 gene was isolated and cloned from 18 representative aromatic rice cultivars in Southern Henan, and the bioinformatics analysis of Badh2 gene was carried out. Meanwhile, seven functional molecular markers developed by Badh2 gene were used to detect and analyze Badh2 gene in 18 aromatic rice varieties from Southern Henan. The results showed that the coding region of Badh2 gene was 1509 bp in length. It contained 15 exons and 14 introns, and encoded 503 amino acids. There are many types of variation of the Badh2 gene in the 18 aromatic rice varieties. According to the variation of Badh2 gene, the tested aromatic rice varieties could be divided into three groups, among which Xinxianggeng 1, Xiangnuo 25, Heixiangdao 193 and Xiangbao 2 were concentrated in group I, while the other 14 kinds of aromatic rice were concentrated in group II. Seven functional molecular markers of Badh2 gene were used to detect different varieties mutation types in exon 2, exon 4–5, exon 7 and exon 13 of Badh2 gene. No aromatic rice varieties with different mutation types were found in promoter region, exon 12 and exon 14 of Badh2 gene. Therefore, our results provide important information for understanding the genetic basis of fragrant genes in aromatic rice germplasm resources in Southern Henan and breeding new varieties of high-quality aromatic rice using molecular marker-assisted selection.

Keywords: Aromatic rice; Badh2 gene; Gene cloning; Bioinformatics analysis; Molecular marker detection.

1. Introduction

Rice (Oryza sativa L.) is the staple food for more than 3 billion people in the world and provides about 25% of their energy [1, 2] which is one of the most important food crops. For most of the population in Southeast Asia, rice provides more than 35% of its energy [3, 4]. The world population is expected to grow at a rate of 25% over the next 30 years and reach 10 billion [5]. With the improvement of people's living standard, the demand for high-quality rice is increasing [6]. Rice quality traits are complex, including nutritional quality, appearance quality, cooking and eating quality, and consumers tend to pay more attention to cooking and eating quality of rice [7-9]. As one of the cultivated rice types, aromatic rice is favored by consumers at home and abroad because of its unique aroma [10]. Over the past decade, the market share of aromatic rice has gradually increased, and the price of aromatic rice has been higher [11, 12]. Therefore, it has important economic value and broad application prospects to conduct in-depth research on aromatic rice germplasm resources, cultivate new varieties with high quality and yield, and apply them in production practice.

More than 200 volatile substances have been isolated and identified in rice [10, 13]. 2-Acetyl-1-pyrroline (2-AP) is considered as one of the main volatile substances in aromatic rice [13, 14], and 2-AP can be detected at low concentrations. Previous studies have found that the aroma of gene is located on chromosome 8 of rice, and the aroma of rice is caused by the mutation of exons 2 and 7 of Badh2 gene. In exon 7 of betaine aldehyde dehydrogenase 2 gene, there are eight deletions and three polymorphisms of deoxynucleotide mutations, which result in the inability of the protein enzymes transcribed and translated from betaine aldehyde dehydrogenase 2 gene to perform their normal functions and make common rice emit aroma that it should not have [11, 15-18]. There are different scholars who argue that the loss of 7 bases at the second exon of Badh2 gene may be the main cause of rice fragrance [19-22]. In different aromatic rice varieties, in addition to one 7 bp deletion in exon 2, there may also be an 803 bp deletion in exons 4 and 5 of Badh2 gene [18, 23, 24], and a mutation site exists in exons 1, 10, 13 and 14 [23-26]. Further studies revealed that there were insertion, deletion or single nucleotide mutation sites in exon 1 and intron 1 of Badh2 gene, promoter region and its 5'-UTR region [18, 27, 28]. Aroma differences in rice are largely determined by allelic variations in the Badh2 gene. Aroma-related rice varieties often include 8 bp deletion in exon 7 and 3 single nucleotide polymorphisms (SNPs), or 7 bp deletion in exon 2 [18, 23]. Therefore, when the Badh2 gene mutates in the coding region or regulatory region, it can produce non-biologically active betaine dehydrogenase, resulting in the aroma of rice, which provides an important theoretical basis for the cultivation of new aromatic rice varieties.

The genetic basis of rice fragrant genes is complex. There is less than one pair of genes controlling rice aroma, or more than 2–4 pairs or more. This may be due to: (1) The genetic basis of different aroma types (such as popcorn, jasmine, violet and pecan) is different; (2) There are interactions between fragrant genes and various environmental factors (illumination, temperature, soil fertility, etc.); (3) The diversity of aroma components in rice, and the methods and techniques for identifying different types of aroma are not perfect at present [29, 30]. There are abundant aromatic rice germplasm resources in the South of Henan, among which the aroma is different. It is speculated that there may be some variation in the fragrant Badh2 gene in aromatic rice. The Badh2 gene was isolated and cloned from 18 representative fragrant rice varieties in Southern Henan. Sequence analysis of the Badh2 gene was carried out. Seven functional molecular markers of the Badh2 gene were used for detection and analysis. Therefore, the
results of this study will further reveal the genetic basis of fragrant genes in Southern Henan aromatic rice and improve important information for the cultivation of new varieties of high-quality aromatic rice.

2. Materials and Methods

2.1. Materials for Testing

All the materials in this study are *japonica* rice varieties. Among them, 18 aromatic rice resources from Southern Henan are widely representative, 14 conventional aromatic rice and 4 *waxy* aromatic rice (Table 1). All aromatic rice varieties were sown in the same experimental field of Xinyang Academy of Agricultural Sciences in 2016. Each variety was planted in 2 rows, 12 plants per row, and the row spacing was 16.5 × 26.4 cm. Routine cultivation and management of common field was carried out from sowing to seed maturity.

| Number | Variety     | Source                                      | Type                      |
|--------|-------------|---------------------------------------------|---------------------------|
| 1      | Xianggeng 805 | Xinyang Academy of Agricultural Sciences   | *Japonica* conventional   |
| 2      | Changxianggeng 101 | Xinyang Academy of Agricultural Sciences | *Japonica* conventional   |
| 3      | Xiangnuo 1862  | Henan Academy of Agricultural Sciences      | *Waxy* aromatic rice      |
| 4      | Zhengxianggeng 11 | Henan Academy of Agricultural Sciences    | *Japonica* conventional   |
| 5      | Nongxianggeng 4 | Agricultural University of Heunodefinednan | *Japonica* conventional   |
| 6      | Xiangnuo 25  | Institute of Agricultural Sciences, Xixian County, Henan Province | *Waxy* aromatic rice      |
| 7      | Xiangbao 2  | Xinyang Institute of Agriculture and Forestry | *Japonica* conventional   |
| 8      | Exiang 1  | Xinyang Institute of Agriculture and Forestry | *Japonica* conventional   |
| 9      | Xiangfeng 916 | Xinyang Academy of Agricultural Sciences | *Japonica* conventional   |
| 10     | Huaibinxiangxiangdao | Institute of Agricultural Sciences, Huaibin County, Henan Province | *Japonica* conventional   |
| 11     | Nongxianggeng | Agricultural University of Heunodefinednan | *Japonica* conventional   |
| 12     | Xinxianggeng 1 | Xinyang Academy of Agricultural Sciences | *Japonica* conventional   |
| 13     | Xiangbao 1  | Xinyang Institute of Agriculture and Forestry | *Japonica* conventional   |
| 14     | Xinxiangruo 933 | Xinyang Academy of Agricultural Sciences | *Waxy* aromatic rice      |
| 15     | Heixianggruo 1926 | Xinyang Institute of Agriculture and Forestry | *Waxy* aromatic rice      |
| 16     | Xiangxiang 1 | Institute of Agricultural Sciences, Huaibin County, Henan Province | *Japonica* conventional   |
| 17     | Heixiangdao 193 | Xinyang Institute of Agriculture and Forestry | *Japonica* conventional   |
| 18     | Exiang 2  | Xinyang Institute of Agriculture and Forestry | *Japonica* conventional   |

2.2. Isolation and Cloning of *Badh2* Gene

Genome-wide DNA extraction from all aromatic rice leaves by CTAB method [4, 31]. PCR primers such as Table 2. The PCR reaction system includes DNA template 2 μl, 10 × Buffer 2.5 μl, MgCl$_2$ 1.5 μl of 25 mmol·L$^{-1}$, dNTP 2.0 μl of 2.5 mmol·L$^{-1}$, forward and reverse primers (12.5 mmol·L$^{-1}$) of 1.0μl, Taq enzyme of 5 U·μl$^{-1}$ of 0.4 μl, supplemented with sterile water to 20 μl. The reagents used were purchased from Shanghai Biotechnology Co., Ltd. The conditions of PCR reaction were as follows: pre-denaturation at 94℃ for 5 min, denaturation at 94℃ for 30 s, annealing at 55–61℃ for 30–90 s, extension at 72℃ for 10 min after 32 cycles, and preservation at 4℃. After the target fragment was recovered by Axygen gel recovery kit, the target fragment was connected to the carrier of pMD19-T (purchased from Takara Company). Systems: Solution I 5 μL, pMD19-T 1 μL, gel recovery product X μL, DDW 4-X pMD19-T 1 μL, gel recovery product 4 μL. Conditions: 16 ℃, overnight. The above products were
transformed into DH5α (purchased from Solarbio Company) and grown overnight at 37 °C. A part of the monoclones grown on the plate was selected to grow in liquid LB medium containing ampicillin. After cultured in shaking bed at 37 °C and 180 rpm for 5~6 h, the positive bacteria were detected by PCR, and then sent to Anhui General Company for sequencing. Finally, the sequence was spliced completely.

Table-2. PCR amplification and sequencing primers of Badh2 gene in rice

| Number | Primer name | Primer sequence  |
|--------|-------------|------------------|
| 1      | Badh2 P1-F  | GTCTCTCCTCCAACATGCTC |
| 2      | Badh2 P3-R  | CAAGCGTCTCTAGTCTAGCC |
| 3      | Badh2 P3-F  | GTGCGTATCTCTGTCTGG |
| 4      | Badh2 P4-R  | CCATCAAGAGAAGATGTTC |
| 5      | Badh2 P5-F  | CUTCCTGGTAATGCAGCTC |
| 6      | Badh2 P10-R | CCGTGTCATCACTAATCCTCC |

2.3. Bioinformatics Analysis Software and Related Web Sites

The nucleotides and structures of Badh2 gene were analyzed, the amino acid sequences encoded by Badh2 were speculated, and the properties and functional domains of the proteins were analyzed and predicted by bioinformatics. The relevant tool software and web site for Badh2 gene sequence analysis are shown in Table 3.

Table-3. The main tools used in bioinformatics analysis of the structure and function of Badh2 gene in rice

| Commonly used resources | Web sites and software |
|------------------------|-----------------------|
| Analysis of Physicochemical Properties of Prot | http://www.expasy.org/tools/protparam.Html |
| Hydrophilic and hydrophobic analysis of proteins | http://web.expasy.org/protscale |
| Protein signal peptide prediction | http://www.cbs.dtu.dk/services/SignalP |
| Prediction of transmembrane domain | http://www.cbs.dtu.dk/services/TMHMM |
| Prediction of protein secondary structure | http://pbil.ibcp.fr/htm/index.php |
| Spatial conformation modeling of proteins | http://swissmodel.expasy.org |
| Analysis of Protein Functional Domains | SMART website |
| Construction of phylogenetic tree | MEGA6.0 Software |
| Basic BLAST retrieval | http://www.ncbi.nlm.nih.gov/BLAST |

2.4. Detection of Gene Functional Markers

Badh2 gene has many allele forms in natural population, and different alleles have different aroma. Seven functional markers of Badh2 gene have been developed for insertion, deletion or single nucleotide mutation sites in exons 2, 4, 5, 7, 12, 13 and 14 of 5'-UTR region of Badh2 gene [11, 28, 32]. Among them, functional marker FME14 is a sequence marker of digestion amplification polymorphism (The cleaved amplified polymorphic sequence marker, CAPS), and the rest are common PCR molecular markers. Detailed information about the amplified products, primer sequence and annealing temperature of each Badh2 gene functional marker is shown in Table 4. All the primers were synthesized by Nanjing Kingsley Biotechnology Co., Ltd. and the preliminary tests showed that the seven functional markers of Badh2 gene were stable in the reaction system and the results were clear and reliable.

Table-4. Functional markers information of Badh2 gene in this experiment

| Alleles     | Functional markers | Primer sequence (5'-3') | Annealing temperature | Enzyme | PCR products size (bp) |
|-------------|-------------------|------------------------|-----------------------|--------|----------------------|
| Badh2-UTR   | FMU1-2            | F:TCCCACCACCACCTCCACA | 61                    | –      | Frgrant:160          |
|             |                   | R:ACGAAGAGCTGCGCTGC   |                       |        | Non-frgrant:163      |
| Badh2-E2    | FME2              | F:GGGAGGCGCTGAAGAGGA  | 60                    | –      | Frgrant:100          |
|             |                   | R:GGGTAGTCCACACCTACCTT |                      |        | Non-frgrant:1107    |
| Badh2-E4-5  | FME4-5            | F:TGCTGGTATGCTTTGAGTA | 55                    | –      | Frgrant:321          |
|             |                   | R:GTCTAGCACACTGAAAGAGCCA |                    |        | Non-frgrant:1123    |
| Badh2-E7    | FME7              | F1:CCGCTGGTCTCTTTCATC | 55                    | –      | Frgrant:583,169     |
|             |                   | R1:TGGAATCGGTTAAAAAGATTATGCC |                  |        | Non-frgrant:591,449 |
|             |                   | F2:GAGCAGCTGAAATATATACC |                     |        |                     |
|             |                   | R2:TTGCAATCTCTGCTTGG  |                       |        |                     |
3. Results and Analysis

3.1. Isolation and Cloning of Badh2 Gene from Aromatic Rice in South Henan

Using the leaf DNA of aromatic rice in Southern Henan as template, the expected DNA fragment could be amplified by using the primers in Table 2 to amplify the expected DNA fragment (Fig. 1). The target fragments were recovered by gel recovery kit and linked to cloning vector pMD19-T. The fragments were transformed into DH5α and grew overnight at 37°C. After monoclonal detection, the positive bacterial liquid was sent to Anhui General Company for sequencing. Finally, the sequence of Badh2 gene of all 18 varieties of aromatic rice germplasm resources in Southern Henan was obtained. Badh2 gene contains 15 exons and 14 introns (Fig. 2), which encodes about 503 amino acids. There are many types of variation in Badh2 gene among 18 aromatic rice varieties in Southern Henan, which may be the reason for the difference of aroma among different aromatic rice varieties in Southern Henan.

| Badh2-E12 | FME12 | F:TTGGTCCAGTGCTCTGTGTG | 58 | Frgrant:189 |
|-----------|-------|-------------------------|----|-------------|
|           |       | R:GCACCAGCCAGACCATAAC   |    | Non-frgrant:192 |
| Badh2-E13 | FME13 | F:TTGGTCCAGTGCTCTGTGTG | 58 | Frgrant:195 |
|           |       | R:GCACCAGCCAGACCATAAC   |    | Non-frgrant:192 |
| Badh2-E14 | FME14 | F:TCGATGCCGGAATTATCTGGGTG | 61 | Bsl I |
|           |       | R:TCCCCACGGCTCATCGGAGG |    | Frgrant:60,205 |

Fig-1. PCR amplification of Badh2 gene in some representative aromatic rice varieties

A, using Badh2 P5-F and Badh2 P10-R primers amplification; B, using Badh2 P1-F and Badh2 P3-R primers amplification; C, using Badh2 P3-F and Badh2 P4-R primers amplification; M, Marker; 1~17 respectively represent aromatic rice varieties are shown in Table 1.

Fig-2. Sequence structure of Badh2 gene coding region

3.2. Amino Acid Sequence Coded by Badh2 Gene and Analysis of its Physicochemical Properties

The Badh2 allele in Xiangfeng 916 aromatic rice widely used in production was analyzed by bioinformatics. The full length of Badh2 allele in Xiangfeng 916 aromatic rice was 7111 bp, encoding 503 amino acids. The isoelectric point of the protein was 5.36, the size of the protein was 54.68 KDa, and its instability coefficient was 72.32. The amino acid sequence analysis of Badh2 gene showed that the content of alanine in Badh2 protein was the
highest, 11.1%. Secondly, the contents of glycine and glutamate were 8.5%. Among the Badh2 protein sequences, 66 negative-charged amino acids (such as aspartic acid, Asp) and glutamic acid accounted for 13.1% of the total amino acid residues, while 57 positive-charged amino acids (such as arginine, Arg) and lysine (Lys) accounted for 11.3% of the total amino acid residues. The fat index and instability index of Badh2 protein were 86.92 and 35.87, which indicated that Badh2 protein was relatively stable.

The hydrophobicity of rice Badh2 protein was predicted by Prot Scale online bioinformatics analysis tool. The results showed that the lower the score, the stronger the hydrophilicity, and the higher the score, the stronger the hydrophobicity. At the same time, the results showed that the most hydrophilic amino acid of Badh2 protein was arginine (Arg), with a score of -2.722, and the most hydrophobic amino acid was glycine (Gly) with a score of 2.067. Overall, the hydrophilic amino acids of Badh2 protein were significantly more than those of hydrophobic amino acids (Fig. 3). Therefore, Badh2 protein of aromatic rice is a relatively hydrophilic protein.

3.3. Prediction of Transmembrane Structure and Signal Peptide Analysis of Badh2 Gene Encoding Protein

The existence of transmembrane domain of Badh2 protein in aromatic rice (www.cbs.dtu.dk/services/TMHMM) was analyzed by TMHMM. The results showed that there was no transmembrane domain of Badh2 protein (Fig. 4).

SignalP 4.0 online analysis software was used to predict the protein encoded by Badh2 gene (http://www.cbs.dtu.dk/services/SignalP). As shown in Fig. 5, the 35th alanine residue of Badh2 protein may be the original splicing site of signal peptide, but the highest predictive score and the highest signal peptide score are 0.110 and 0.153, respectively. Both predictive values are low. Comprehensive analysis results show that Badh2 protein is likely to have no signal peptide sequence.
3.4. Structural Analysis of Badh2 Protein in Aromatic Rice

The secondary structure of Badh2 protein in rice was analyzed by online protein analysis website (https://npsa-prabi.ibcp.fr/cgi-bin). The results showed that there were 190 amino acids in Badh2 protein, accounting for 37.77% of the total sequence; 74 amino acid stretching fragments, accounting for 14.71% of the total sequence; 239 amino acid compositions were irregularly curled, accounting for 47.51% of the total sequence (Fig. 6). Using SWISS-MODEL software on-line (http://swissmodel.expasy.org/repository), the tertiary structure of Badh2 protein was predicted and homologous modeling was carried out. The results showed that the random curl and α-helix structure occupied most of the (Fig 7), which was consistent with the secondary structure of Badh2 protein.
3.5. Cluster Analysis of Badh2 Gene in Southern Henan Aromatic Rice

According to the DNA sequence of Badh2 gene of all the 18 aromatic rice germplasm resources in Southern Henan, the DNA sequence of Badh2 gene of aromatic rice in Southern Henan was aligned by MEGA7 software, and the cluster map of Badh2 gene of 18 aromatic rice germplasm resources in Southern Henan was drawn by MEGA7 software. The results showed as shown Fig 8: Xinxianggeng 1, Xiangnuo 25, Heixiangdao 193 and Xiangbao 2 were clustered in group I, while the remaining 14 kinds of aromatic rice were clustered in group II. In group II, Nongxianggeng, Huai bin Xiangxiang rice, Xiangnuo 1862, Hei Xiangnuo 1926 and Xiangxiang 1 gathered in group A, ZhengXianggeng, Xinxianxianggeng 933, Exiang 2, Xianggeng 805, Changxianggeng 101, Xiangfeng 916, Exiang 1 and Xiangbao 1 gathered in group B, and Nongxianggeng 4 was different from group A or group B in Badh2 gene sequence (Fig. 8).

![Cluster analysis of Badh2 gene in aromatic rice from South Henan](image)

3.6. Molecular Marker Detection and Analysis of Badh2 Gene in Aromatic Rice in South Henan

The genomic DNA of 18 aromatic rice varieties from South Henan was used as template. According to some of the variation loci, Badh2 gene functional markers (Table 4) were used to amplify PCR and detected by non-denaturing polyacrylamide gel electrophoresis. Functional molecular marker FME2 was used to detect the fragments. The results showed that seven aromatic rice varieties, including Xianggeng 805, Changxianggeng 101, Nongxianggeng, Xiangbao 1, Exiang 1, Nongxianggeng 4 and Xinxianxianggeng 1, could amplify 100 bp fragments, indicating that there was a 7 bp deficiency in exon 2 of Badh2 gene in all seven aromatic rice varieties. Lost fragments belong to the same variation type of aroma gene. A 107 bp fragment was amplified from the other 11 fragrant rice varieties in Southern Henan, indicating that the 11 aromatic rice varieties in Southern Henan did not belong to the mutation type of exon 2 of Badh2 gene (Table 5). Badh2 gene functional marker FME4-5 was used for detection and analysis respectively. The results showed that Xianggeng 805 and Nongxianggeng 4 could amplify a fragment with the size of 321 bp. The remaining 16 aromatic rice varieties in Southern Henan province all amplified a fragment with the size of 1123 bp. It indicates that only Xianggeng 805 and Nongxianggeng 4 of the two aromatic rice materials in Southern Henan belong to exon 4~5 mutation types of Badh2 gene. Using the Badh2 gene functional marker FME7, it was found that Zhengxianggeng 11, Xinxianxianggeng 933, Xiangfeng 916, Exiang 2, Huai bin xiangxianggeng, Xuangnuo 1862, Heixianggeng 1926 and Xiangxiang 1 could amplify two fragments with sizes of 583 and 169 bp. The results indicated that there was an 8 bp deletion fragment and 3 bp mutation in the 7th exon of Badh2 gene in these 5 aromatic rice materials, which belonged to the same mutation type of fragrant gene. Badh2 gene functional marker FME13 was used to detect and analyze the fragments. The results showed that Xiangbao 2, Xianggeng 25 and Heixianggengdao 193 could amplify 195 bp fragments, while the other 15 varieties could only amplify 192 bp fragments. Badh2 gene was detected by functional markers FMU1-2, FME4-5, FME2 2 and FME14 of 18 fragrant rice varieties in Southern Henan, respectively. No polymorphic loci were found (Table 5). Therefore, the Badh2 gene of 18 aromatic rice varieties in Southern Henan may not be mutated at 5'-UTR region, exon 12 and exon 14.

![Detection results of functional markers of Badh2 gene in 18 aromatic rice materials](image)

Table 5. Detection results of functional markers of Badh2 gene in 18 aromatic rice materials

| Alleles          | Mutation type                        | Fragrant rice varieties                   |
|------------------|--------------------------------------|------------------------------------------|
| Badh2-UTR        | 3 bp deletion in 5'UTR region         |                                          |
| Badh2-E2         | 7 bp deletion in exon 2               | Xianggeng 805, Changxianggeng 101,       |
|                  |                                      | Nongxianggeng, Xiangbao 1, Exiang 1,     |
|                  |                                      | Nongxianggeng 4, Xinxianxianggeng 1      |
| Badh2-E4-5       | 803 bp deletion of exon 4 and 5       | Xianggeng 805, Nongxianggeng 4           |
| Badh2-E7         | 8 bp deletion and 3 bp                | Zhengxianggeng 11, Xinxianxianggeng 933,|
4. Discussion

4.1. Biological Function of Badh2 Gene

Previous studies have shown that a mutation in the Badh2 gene on chromosome 8 causes the rice to produce fragrance, while the Badh2 protein encoded by Badh2 gene in non-aromatic rice has the activity of betaine dehydrogenase. The loss of betaine dehydrogenase activity in the mutant type and the accumulation of 2-AP lead to the production of rice flavor [11, 22, 24]. Through sequencing and analysis of rice fragrant genes, most researchers believe that the deletion of exon 2 or 7 of Badh2 gene results in the accumulation of 2-AP [24] and aroma production. But the fact that some rice varieties have high levels of 2-AP but no Badh2 alleles suggests that there may be other Non-badh2 alleles that contribute to fragrance [26]. Most researchers only focus on the differences between exons of aromatic rice and non-aromatic rice, but for eukaryotes, introns play an important role in gene expression and regulation besides exons of genes [33, 34]. The results showed that the Badh2 gene of 18 representative aromatic rice varieties in Southern Henan contained 15 exons and 14 introns (Fig. 2), encoding about 503 amino acids, and there were many types of mutations in the Badh2 gene of 18 aromatic rice varieties in Southern Henan. This may be the main reason for the difference of aroma among different aromatic rice varieties in the aromatic rice germplasm resources of Southern Henan. Further analysis of Badh2 gene sequence showed that aromatic rice varieties in Southern Henan could be divided into two groups. Xinxianggeng 1, Xiangnuo 25, Heixiangdao 193 and Xiangbao 2 were clustered in group I, while the other 14 aromatic rice varieties were clustered in group II. These results provide important information for the future development and utilization of Badh2 gene in Southern Henan.

4.2. Regulation Mechanism of Badh2 Gene

So far, although it has been confirmed that the loss of Badh2 gene function can lead to the production of rice aroma substance 2-AP, how Badh2 protein regulates the biosynthesis of 2-AP remains unclear. Betaine aldehyde dehydrogenase plays an important role in the biosynthesis of 2-acetyl-1-pyrroline. However, how betaine aldehyde dehydrogenase plays a role in the synthesis of 2-acetyl-1-pyrroline and whether other substances are involved in the regulation are still unclear and need to be further studied. Meanwhile, gamma-aminobutyraldehyde, as the precursor of 2-acetyl-1-pyrroline synthesis, may also play a key regulatory role in the biosynthesis of 2-acetyl-1-pyrroline. The loss of Badh2 protein function in aromatic rice may lead to the accumulation of gamma-aminobutyraldehyde in the body, and then convert gamma-aminobutyraldehyde to 1-pyrrolineid, and eventually synthesize a large amount of 2-acetyl-1-pyrroline (2-AP [22]). In non-aromatic rice, Badh2 protein has the catalytic activity of betainaldehyde dehydrogenase, which may convert gamma-aminobutyric acid from gamma-aminobutyric acid in rice, but inhibit the synthesis of 2-AP precursor, 1-pyrrolineid, and ultimately make it impossible to synthesize 2-AP. Therefore, the biosynthetic pathway of 2-AP and related regulatory networks and mechanisms are still unclear. In this study, the Badh2 allele of xiangfeng 916, which is widely used in production, was taken as an example to conduct bioinformatics analysis of its Badh2 allele. The results showed that Badh2 protein was a relatively hydrophilic protein with good stability, and signal peptide sequence and transmembrane structure were not found in Badh2 (Fig. 3–5). Further analyzing the senior Badh2 protein structure prediction, the result shows that in the Badh2 protein structure, random coil and alpha helix occupy a large (Fig 7). The structure of protein plays a decisive role in its function. This study conducted structural prediction analysis of Badh2 protein, providing important clues for in-depth analysis of 2-AP biosynthesis pathway and related regulatory networks and mechanisms.

4.3. Application of Badh2 Gene in Molecular Breeding of Aromatic Rice

Since Badh2 gene was isolated and cloned in japonica rice, multiple variation sites have been found in the gene [11, 18, 24, 35], and a series of molecular markers have been designed for the identification of aroma genes, the screening of different aromatic rice varieties and the cultivation of new varieties of aromatic rice. At present, there are at least 17 mutation loci in Badh2 gene, which are distributed in the 5'-UTR region of Badh2 gene, the 1st exon, the junction between the 1st exon and the 1st intron, the 2nd exon, between the 4th exon and the 5th exon, and the 7th, 8th, 10th, 12th, 13th and 14th exons [11]. For these mutation sites, multiple pairs of molecular markers have been designed at the junction of the 1st exon and the 1st intron of Badh2 gene, the 2nd exon, the 4th exon and the 5th exon, the 7th, 12th, 13th and 14th exons to identify the mutation type of Badh2 gene [11, 19, 16, 24, 27, 25, 28, 36]. On the basis of these molecular markers, a total of 7 functional markers of rice flavor genes were further developed (Table 4) and verified by using isolated populations [11, 24, 27, 36]. Among them, functional marker FMU1-2 can not only detect the 8 bp deletion mutation of Badh2 gene in the 7th exon, but also detect whether there is an 8 bp deletion in the 5'-UTR region of Badh2 gene. Functional markers FME2-7, FME7, FME12-3, FME13 and FME14 were used to identify the variation types of aroma genes in exons 2, 7, 12, 13 and 14 of Badh2, respectively. Functional markers FME14 belongs to CAPS (the cleaved amplified polymorphic sequence) tag, so the PCR product detection, need the help of restriction enzymes Bst.

| Mutation | Description | Genotype |
|----------|-------------|----------|
| Badh2-E12 | 3 bp deletion of exon 12 | – |
| Badh2-E13 | 3 bp insertion of exon 13 | Xiangbao 2, Xiangnuo 25, Heixiangdao 193 |
| Badh2-E14 | 1 bp insertion of exon 14 | – |

### Table 4

| Genotype | Mutation in exon 7 | Genotype |
|----------|-------------------|----------|
| –         | 916, Exiang 2, Huai binxiongxiangdao, Xiangnuo 1862, Heixiangdao 1926, Xiangxiang 1 |

...
In this study, seven functional molecular markers were used to detect the mutation types of Badh2 gene in southern Henan aromatic rice. The results showed that there were mutations in exon 2, exon 4-5, exon 7 and exon 13 of the Badh2 gene in 18 Southern Henan aromatic rice varieties. Interestingly, there may be no variation in the 5'-UTR region, the 12th exon, and the 14th exon of the Badh2 gene (Table 5). The results of these molecular markers were in good agreement with the results of Badh2 gene cluster analysis in Southern Henan (Fig. 8). Therefore, the detection and analysis of functional molecular markers of these aroma genes provide an accurate, rapid and effective means for molecular breeding of aromatic rice.

5. Conclusion

According to the variation of Badh2 gene in 18 aromatic rice varieties in Southern Henan, the tested varieties can be divided into three groups. Xinxianggeng 1, Xiangnuo 25, Heixiangdao 193 and Xiangbao 2 are clustered in group I, while the remaining 14 fragrant rice varieties are clustered in group II. Seven functional molecular markers of Badh2 gene were used to detect different mutation types in exon 2, exon 4-5, exon 7 and exon 13 of Badh2 gene. No aromatic rice varieties with different mutation types were found in other locations of Badh2 gene. Therefore, the results of this study provide an important basis for using molecular marker-assisted selection to breed new high-quality aromatic rice varieties.

Acknowledgements

This work was financially supported by National Natural Science Foundation of China (U1604110, 31600992, 31801332), Key Scientific and Technological Research Projects in Henan Province (182102110442, 192102110119), Special Research Project of Teacher Education Linkage Development Community in Southern Henan (2019-GTYYB-01), Research Project of Teacher Education Curriculum Revolution of XYNU (2019-JSJJYYJ-10), Nanhu Scholars Program for Young Scholars of XYNU (20160504), Scientific Research Innovation Project for Postgraduate of XYNU (2018KYJJ47), Key Scientific Research Projects of Universities in Henan Province (19A180030), Institute for Conservation and Utilization of Agro-bioresources in Dabie Mountains.

References

[1] Peng, B., Kong, H. L., Li, Y. B., Wang, L. Q., Zhong, M., Sun, L., Gao, G. J., Zhang, Q. L., Luo, L. J., et al., 2014a. "OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice." Nat Commun., vol. 5, p. 4847.
[2] Tian, Z. X., Qian, Q., Liu, Q. Q., Yan, M. X., Liu, X. F., Yan, C. J., Liu, G. F., Gao, Z. Y., Tang, S. Z., et al., 2009. "Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities." Proc. Natl. Acad. Sci. USA., vol. 106, pp. 21760–21765.
[3] Kusano, M., Yang, Z. G., Okazaki, Y., Nakabayashi, R., Fukushima, A., and Saito, K., 2015. "Using metabolomic approaches to explore chemical diversity in rice." Molecular Plant, vol. 8, pp. 58–67.
[4] Peng, B., Wang, L. Q., Fan, C. C., Jiang, G. H., Luo, L. J., Li, Y. B., and He, Y. Q., 2014b. "Comparative mapping of chalkiness components in rice using five populations across two environments." BMC Genet., vol. 15, p. 49.
[5] Hickey, L. T., Hafeez, A. N., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C. M., Tester, M., Gao, C. X., Godwin, I. D., Hayes, B. J., et al., 2019. "Breeding crops to feed 10 billion." Available: https://doi.org/10.1038/s41587-019-0152-9
[6] Fitzgerald, M. A., Mccouch, S. R., and Hall, R. D., 2009. "Not just a grain of rice: the quest for quality." Trends in Plant Science, vol. 14, pp. 133–139.
[7] Guo, T., Liu, X. L., and Wan, X. Y., 2011. "Identification of a stable quantitative trait locus for percentage grains with white chalkiness in rice." Journal of Integrative Plant Biology, vol. 53, pp. 598–607.
[8] Takayuki, K. and Jun, M., 2018. "Identification and characteristics of quantitative trait locus for grain protein content, TGP12, in rice Oryza sativa L." Euphytica, vol. 214, p. 165.
[9] Daygon, V. D., Prakash, S., Calingacion, M., Riedel, A., Owenden, B., Snell, P., Mitchell, J., and Fitzgerald, M., 2016. "Understanding the jasmine phenotype of rice through metabolite profiling and sensory evaluation." Metabolomics, vol. 12, p. 3.
[10] Myint, K. M., Arikit, S., and Wanchana, S., 2012. "A PCR-based marker for a locus conferring the aroma in Myanmar rice Oryza sativa L." Theoretical and Applied Genetics, vol. 125, pp. 887–896.
[11] He, Q. and Park, Y. J., 2015a. "Discovery of a novel fragrant allele and development of functional markers for fragrant rice." Molecular Breeding, vol. 35, p. 217.
[12] Bradbury, L. M., Fitzgerald, T. L. and, Henry, R. J., 2005a. "The gene for fragrance in rice." Plant Biotechnology Journal, vol. 3, pp. 363–370.
[13] Daygon, V. D., Prakash, S., and Calingacion, M., 2016. "Understanding the jasmine phenotype of rice through metabolite profiling and sensory evaluation." Metabolomics, vol. 12, pp. 1–15.
[14] Plows, J. F., Yu, X. Y., and Broadhurst, R., 2017. "Absence of a gestational diabetes phenotype in the LepRdb/+ mouse is independent of control strain, diet, misty allele, or parity." Scientific Reports, vol. 7, p. 45130.
[15] Bradbury, L. M., Henry, R. J., and Jin, Q., 2005b. "A perfect marker for fragrance genotyping in rice." Molecular Breeding, vol. 16, pp. 279–283.
ally important traits of Basmati—ant allele domesticated in India. "Food Chemistry," vol. 142, pp. 383–391.

[18] Shao, G. N., Tang, S. Q., and Chen, M. L., 2013. "Haplotype variation at Badh2, the gene determining fragrance in rice." Genomics, vol. 101, pp. 157–162.

[19] Chen, S., Wu, J., and Yang, Y., 2006. "The fgr gene responsible for rice fragrance was restricted within 69 kb." Plant Science, vol. 171, pp. 505–514.

[20] Chen, S., Yang, Y., and Shi, W., 2008. "Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance." Plant Cell, vol. 20, pp. 1850–1861.

[21] Niu, X., Tang, W., and Huang, W., 2008. "RNAl-directed downregulation of OsBADH2 results in aroma, 2-acetyl-1-pyrroline, production in rice Oryza sativa L." BMC Plant Biology, vol. 8, pp. 300–311.

[22] Shan, Q. W., Zhang, Y., and Chen, K. L., 2015. "Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology." Plant Biotechnology Journal, vol. 13, pp. 791–800.

[23] Shi, Zhao, G. C., Xu, X. L., and Li, J. Y., 2014. "Discovery of a new fragrance allele and development of functional markers for identifying diverse fragrant genotypes in rice." Molecular Breeding, vol. 33, pp. 701–708.

[24] Michael, J. K., Mariafe, N. C., Melissa, A. F., and Susan, R. M., 2009. "The origin and evolution of fragrance in rice Oryza sativa L." Proc. Natl. Acad. Sci. USA, vol. 106, pp. 14444–14449.

[25] 2015. "Discovery and mapping of genomic regions governing economically important traits of Basmati rice." BMC Plant Biology, vol. 15, p. 207.

[26] Amarawathi, Y., Singh, R., Singh, A. K., Singh, V. P., Mohapatra, T., Sharma, T. R., and Singh, N. K., 2008. "Mapping of quantitative trait loci for OsBADH2 expression in rice Oryza sativa L." Molecular Breeding, vol. 21, pp. 49–65.

[27] Ootsuka, K., Takahashi, I., Tanaka, K., Itani, T., Tabuchi, H., Yoshihashi, T., Tonouchi, A., and Ishikawa, R., 2014. "Genetic polymorphisms in Japanese fragrant landraces and novel fragrant allele domesticated in northern Japan." Breeding Sci, vol. 64, pp. 115–124.

[28] Shao, G. N., Tang, A., and Tang, S. Q., 2011. "A new deletion mutation of fragrance gene and the development of three molecular markers for fragrance in rice." Plant Breeding, vol. 130, pp. 172–176.

[29] Luo, W., Guo, T., and Yang, Q., 2014. "Stacking of five favorable alleles for amylase content, fragrance and disease resistance into elite lines in rice Oryza sativa L. By using four HRM-based markers and a linked gel-based marker." Molecular Breeding, vol. 34, pp. 805–815.

[30] Peng, B., Sun, Y. F., Chen, B. Y., Sun, R. M., Kong, D. Y., Li, X. W., Song, X. H., Li, H. L., et al., 2017. "Research progress of fragrance gene and its application in rice breeding." Chin Bull Bot, vol. 52, pp. 797–807.

[31] Jiang, G. H., He, Y. Q., and Xu, C. G., 2004. "The genetic basis of stay-green in rice analyzed in a population of doubled haploid lines from an indica by japonica cross." Theoretical and Applied Genetics, vol. 108, pp. 688–698.

[32] Xu, X. L., Zhao, G. C., and Li, J. Y., 2011. "Development of molecular markers used to identify two types of fragrant rice and analysis of mutation sites of BADH2 gene in 24 varieties of fragrant rice." Plant Diversity and Resources, vol. 33, pp. 667–673.

[33] Fu, H., Kim, S. Y., and Park, W. D., 1995. "High-level tubule expression and sucrose inducibility of a potato SucS4 sucrose synthase gene require 5' and 3' flanking sequences and the leader intron." Plant Cell, vol. 7, pp. 1387–1394.

[34] Lee, B. H., Won, S. H., Lee, H. S., Miyao, M., Chung, W. I., Kim, I. J., and Jo, J., 2000. "Expression of the chloroplast-located small heat shock protein by oxidative stress in rice." Gene, vol. 245, pp. 283–290.

[35] He, Q., Yu, J., Kim, T. S., Cho, Y. H., Lee, Y. S., and Park, Y. J., 2015b. "Resequencing reveals different domestication rate for BADH1 and BADH2 in rice Oryza sativa L." PLoS One, vol. 10, p. e0134801.

[36] Shi, Yang, Y., Chen, S. H., and Xu, M. L., 2008. "Discovery of a new fragrance allele and the development of functional markers for the breeding of fragrant rice varieties." Molecular Breeding, vol. 22, pp. 185–192.