Badh2, Encoding Betaine Aldehyde Dehydrogenase, Inhibits the Biosynthesis of 2-Acetyl-1-Pyrroline, a Major Component in Rice Fragrance

Saihua Chen, Yi Yang, Weiwei Shi, Qing Ji, Fei He, Ziding Zhang, Zhukuan Cheng, Xiangnong Liu, and Mingliang Xu

INTRODUCTION

Fragrant rice (Oryza sativa) is gaining widespread popularity among consumers worldwide (Bhattacharjee et al., 2002); thus, its market price is much higher than that of nonfragrant rice (Qiu and Zhang, 2003). A mixture of 114 different volatile compounds was detected in the flavor of cooked rice (Yajima et al., 1978). One of them, 2-acetyl-1-pyrroline (2AP), is a potent flavor component in rice fragrance. By contrast, its two recessive alleles, badh2-E2 and badh2-E7, induce 2AP formation. Badh2 was found to be transcribed in all tissues tested except for roots, and the transcript was detected at higher abundance in young, healthy leaves than in other tissues. Multiple Badh2 transcript lengths were detected, and the complete, full-length Badh2 transcript was much less abundant than partial Badh2 transcripts. 2AP levels were significantly reduced in cauliflower mosaic virus 35S-driven transgenic lines expressing the complete, but not the partial, Badh2 coding sequences. In accordance, the intact, full-length BADH2 protein (903 residues) appeared exclusively in nonfragrant transgenic lines and rice varieties. These results indicate that the full-length BADH2 protein encoded by Badh2 renders rice nonfragrant by inhibiting 2AP biosynthesis. The BADH2 enzyme was predicted to contain three domains: NAD binding, substrate binding, and oligomerization domains. BADH2 was distributed throughout the cytoplasm, where it is predicted to catalyze the oxidation of betaine aldehyde, 4-aminobutyraldehyde (AB-ald), and 3-aminopropionaldehyde. The presence of null badh2 alleles resulted in AB-ald accumulation and enhanced 2AP biosynthesis. In summary, these data support the hypothesis that BADH2 inhibits 2AP biosynthesis by exhausting AB-ald, a presumed 2AP precursor.

In rice (Oryza sativa), the presence of a dominant Badh2 allele encoding betaine aldehyde dehydrogenase (BADH2) inhibits the synthesis of 2-acetyl-1-pyrroline (2AP), a potent flavor component in rice fragrance. By contrast, its two recessive alleles, badh2-E2 and badh2-E7, induce 2AP formation. Badh2 was found to be transcribed in all tissues tested except for roots, and the transcript was detected at higher abundance in young, healthy leaves than in other tissues. Multiple Badh2 transcript lengths were detected, and the complete, full-length Badh2 transcript was much less abundant than partial Badh2 transcripts. 2AP levels were significantly reduced in cauliflower mosaic virus 35S-driven transgenic lines expressing the complete, but not the partial, Badh2 coding sequences. In accordance, the intact, full-length BADH2 protein (903 residues) appeared exclusively in nonfragrant transgenic lines and rice varieties. These results indicate that the full-length BADH2 protein encoded by Badh2 renders rice nonfragrant by inhibiting 2AP biosynthesis. The BADH2 enzyme was predicted to contain three domains: NAD binding, substrate binding, and oligomerization domains. BADH2 was distributed throughout the cytoplasm, where it is predicted to catalyze the oxidation of betaine aldehyde, 4-aminobutyraldehyde (AB-ald), and 3-aminopropionaldehyde. The presence of null badh2 alleles resulted in AB-ald accumulation and enhanced 2AP biosynthesis. In summary, these data support the hypothesis that BADH2 inhibits 2AP biosynthesis by exhausting AB-ald, a presumed 2AP precursor.
located in this fgr region: Cah and Mccc2, encoding eukaryotic-type carbonic anhydrase and 3-methylcrotonyl-CoA carboxylase β-chain, respectively (Chen et al., 2006).

In cocoa fermentation, 2AP is synthesized from either L-Pro or L-Orn by Bacillus cereus (Romanczyk et al., 1995). In fragrant rice, L-Pro is a precursor to 2AP, providing the nitrogen of the pyrroline ring but not the carbon of the acetyl group (Yoshihashi et al., 2002). 2AP formation positively correlates with an accumulation of Pro, resulting in the strong aroma of the Thailand fragrant rice variety Khao Dawk Mali 105, which is grown in the drought region Tung Kula Rong Hai (Yoshihashi et al., 1999). Interestingly, Pro has been shown to be involved in osmoregulation and to accumulate in rice plants under diverse types of stress, including drought (Yang and Kao, 1999). It has been suggested that 2AP is synthesized via the polyamine pathway, in which Orn, via 1-pyrroline, is the major source of the nitrogen in 2AP (A. Vanavichit, T. Yoshihashi, S. Wanchana, S. Areekit, D. Saengsraku, W. Kamolsukyunyong, J. Lanceras, T. Toojinda, and S. Tragoonrnung, unpublished data). To date, however, the biosynthetic pathway of 2AP has not been demonstrated clearly. Therefore, the identification of the fragrant-related gene(s) would provide valuable insight regarding the mechanism of 2AP biosynthesis.

RESULTS

Identification and Cloning of Fgr Candidates

We previously mapped fgr to a 69-kb region flanked by the markers L02 and L06 (Chen et al., 2006). Here, fgr was further localized to an interval between L04 and L06 (Figure 1). Sequence alignment between the Fgr-containing nonfragrant rice varieties indica cv 93-11 and japonica cv Nipponbare revealed considerable variation within this L04/L06 interval, which had a physical distance of 62 kb in 93-11 but only 35 kb in Nipponbare. In Nipponbare, a 30-kb DNA segment rich in transposition-related genes was inserted inside an interval of L02/L03. In 93-11, a large DNA insert was present between L04 and L05. Three putative genes in the L04/L06 interval, shared by both 93-11 and Nipponbare, were considered to be candidates for the Fgr gene in these nonfragrant varieties. These genes are Cah, Mccc2, and Badh2, encoding putative eukaryotic-type carbonic anhydrase, 3-methylcrotonyl-CoA carboxylase β-chain, and betaine aldehyde dehydrogenase, respectively (Figure 1).

After screening the indica Nanjing11 and japonica Suyunuo BAC libraries, we obtained two and four positive BAC clones in the Fgr (Nanjing11; nonfragrant) and fgr (Suyunuo; fragrant) regions, respectively. Two BAC clones from each library were sufficient to cover the regions of interest and are shown in Figure 1. Three Fgr candidates and three fgr candidates from these BAC clones (corresponding to Cah, Mccc2, and Badh2 in each case) were subcloned into the expression vector pCambia1302 (pCAM). Coding regions for Cah, Mccc2, and Badh2 were determined to be 2499, 6506, and 5846 bp, respectively. Maps of the resulting six clones are show in Supplemental Figure 1 online. The inserts of these six Fgr/fgr constructs were sequenced and compared with their counterparts in the indica cv 93-11 (http://rise.genomics.org.cn) and the japonica cv Nipponbare (http://www.gramene.org). Sequence alterations specific in the fragrant variety Suyunuo were observed only in the badh2 gene, including an 8-bp deletion (5’-GATTATGG-3’) and three single nucleotide polymorphisms in exon 7 (see Supplemental Figure 2A and Supplemental Data Set 1 online). Furthermore, we sequenced another 23 fragrant rice varieties at their badh2 loci. Apart from the above badh2 allele (designated badh2-E7), a new null badh2 allele (named badh2-E2) with a 7-bp deletion (5’-CGGGCGC-3’) in exon 2 was identified (Shi et al., 2008). In addition, other nucleotide sequence divergences were also

Figure 1. Positive BAC Clones, Predicted Genes, Developed Markers Used for Mapping, and Genomic Sequences in the Fgr Regions for the indica cv 93-11 and the japonica cv Nipponbare.

Similar genomic regions are shaded in gray, while major insertions are shown in white (regions A and B). Dark arrows show three putative genes, Cah, Mccc2, and Badh2, within the L04/L06 region that are common to both the indica and japonica rice subspecies and, therefore, were considered to be candidates for the Fgr gene. BAC clones used for subcloning are shown above and below the genomic region.
observed within introns in two fragrant rice varieties, cv Wuxiangjing9 and Suyunuo, including a TT deletion in intron 2 and a repeated (AT)n insert in intron 4.

An alignment of amino acid sequences of Badh2, badh2-E2, and badh2-E7 alleles, together with a highly homologous rice Badh1 gene as reference, is provided in Supplemental Figure 2B online. The intact BADH2 shows 75.6% identity with the BADH1 amino acid sequence. The truncated BADH2 from badh2-E7 consisted of a presumed peptide with 251 residues. Another truncated BADH2 from badh2-E2 consisted of a shorter peptide with 82 residues.

**Confirmation of Badh2 Function in Rice Fragrance**

Because rice fragrance is a recessive trait, three Fgr constructs (pCAM-Badh2, pCAM-Cah, and pCAM-Mccc2) were transformed into the fragrant rice recipient Wuxiangjing (japonica, badh2-E7) and transgenic lines were identified (see Supplemental Figure 3 online). Details and photographs showing the steps of the Agrobacterium tumefaciens–mediated rice transformation procedure are available from the authors upon request. All 30 plantlets regenerated from the calli transformed with pCAM-Badh2 were confirmed to be positive. Only two and three transgenic plantlets were obtained from the calli transformed with pCAM-Cah and pCAM-Mccc2, respectively (see Supplemental Figures 3A and 3B online).

The 2AP contents were determined for all regenerated plantlets by gas chromatography–mass spectrometry (GC-MS) (see Supplemental Figure 4 online). Transgenic lines carrying either the exogenous Cah or Mccc2 gene were fragrant, and their 2AP levels were as high as those of nontransgenic lines. By contrast, all Badh2 transgenic lines were nonfragrant with no or very low 2AP content (Table 1; see Supplemental Table 1 online). The Badh2 transgenic lines were further self-pollinated to generate T1 progeny, which were analyzed by PCR using Badh2 genespecific primers (see Supplemental Figure 5 online). Segregation of the exogenous Badh2 gene in T1 progeny was consistent with the 3:1 ratio expected for monogenic inheritance. All T1 progeny with the exogenous Badh2 genes were nonfragrant with reduced 2AP levels, while individuals without the exogenous Badh2 gene were fragrant.

The construct pCAM-Badh2 was further transformed into another four fragrant rice varieties (Zhenxiangjing5, Guanglingxiangjing, Wuxiangjing9, and Xiangjing111) that harbor the other null badh2-E2 alleles. As expected, transgenic lines showed significantly reduced 2AP levels compared with nontransgenic lines (Table 2; see Supplemental Table 2 online). The data from the above analyses indicate that the Badh2 gene can compensate for the defective functions of both badh2-E2 and badh2-E7 to render transgenic lines nonfragrant.

**Transcription Profiling of the Badh2/badh2-E7 Alleles**

In the nonfragrant rice cv Nanjing11, Badh2 transcripts were detected in all tissues tested except for the roots. The Badh2 transcript was more abundant in the initial fully expanded leaf than in other tissues (Figure 2). Interestingly, a similar transcription pattern with a lower transcription level was observed for the badh2-E7 allele in the fragrant rice cv Wuxiangjing (Figure 2). This indicates that sequence alterations specific to the badh2-E7 allele significantly reduce its transcription levels but cause only a slight change in its transcription pattern.

**Multiple Transcription Start Sites Are Present in the Badh2 Gene**

To identify the transcription start point, rapid amplification of cDNA ends (RACE) analysis was performed on mRNA from leaves of nonfragrant rice cv Nanjing11. As expected, the 3′-RACE product was 1013 bp, the size predicted from the Badh2

---

Table 1. 2AP Levels and Their Significant Differences among Three Kinds of Transgenic Lines and Nontransgenic Lines

|                 | Nontransgenic and Transgenic Lines | 2AP Levels (ng/g; mean ± SD) | Significance
|-----------------|-----------------------------------|------------------------------|-------------
|                 | No. of Plantlets                   |                              | At 5%       | At 1%       |
| Nontransgenic lines | 11                                 | 21.33 ± 7.37 a               | A           |
| Cah transgenic lines | 2                                 | 18.94 ± 1.32 a               | A           |
| Mccc2 transgenic lines | 4                                 | 20.52 ± 3.75 a               | A           |
| Badh2 transgenic lines | 11                                | 1.99 ± 1.03 b                | B           |

*a2AP levels with different letters are significantly different at P < 0.05 (lowercase letter) or at P < 0.01 (uppercase letter).*

---

![Figure 2. Examination of the Badh2/badh2 Transcription Levels in Various Plant Tissues Using Real-Time RT-PCR.](image)
The Badh2 Gene Inhibits 2AP Synthesis

Table 2. 2AP Levels in Nontransgenic and Transgenic Plantlets Using Four Fragrant Rice Varieties (badh2-E2) as Recipients

<table>
<thead>
<tr>
<th>Fragrant Rice Recipients (badh2-E2)</th>
<th>Transgenic Lines (Badh2·badh2-E2/badh2-E2)</th>
<th>Nontransgenic Lines (badh2-E2/badh2-E2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Plantlets</td>
<td>2AP Levels (ng/g; means ± SD)</td>
</tr>
<tr>
<td>Xiangjing111</td>
<td>3</td>
<td>5.75 ± 0.46</td>
</tr>
<tr>
<td>Guanglingxiangjing2</td>
<td>2</td>
<td>8.21 ± 1.83</td>
</tr>
<tr>
<td>Wuxiangjing9</td>
<td>3</td>
<td>9.54 ± 1.15</td>
</tr>
<tr>
<td>Zhenxiangjing5</td>
<td>3</td>
<td>7.92 ± 2.99</td>
</tr>
</tbody>
</table>

*No data available.

gene sequence (http://www.gramene.org). However, to our surprise, the majority of the 5′-RACE products were only ~500 bp, much shorter than the 800 bp expected from the Badh2 gene (Figure 3). Sequence analysis of the 5′-RACE products showed that the major transcription start point moved downstream of the predicted Badh2 start codon, resulting in partial Badh2 transcripts. Moreover, the Badh2 coding sequence (CDS) appeared to have variable transcription start points, since different leader sequences were observed that corresponded to exon 2, intron 2, or exon 3. A search for a possible Badh2 cDNA sequence in the database (www.ncbi.nlm.nih.gov) identified a Badh2 cDNA clone (National Center for Biotechnology Information accession number AK060461) with an even shorter coding sequence (1095-bp CDS) compared with the partial Badh2 cDNA (1182-bp CDS) identified in this study.

To confirm the presence of multiple Badh2 transcripts, we designed three forward primers (FP1, FP2, and FP3) and one common reverse primer (RP) based on the Badh2 predicted sequence (see Supplemental Figure 6 and Supplemental Table 4 online). Using real-time RT-PCR, the primer pair FP1/RP generated very small amounts of PCR products, while the other two primer pairs, FP2/RP and FP3/RP, generated much larger amounts (Figure 4A). Likewise, no PCR product was observed in RT-PCR with the primer pair FP1/RP, although strong PCR bands were amplified using the other two primer pairs, FP2/RP and FP3/RP (Figure 4B).

The relative abundance of different Badh2 transcripts was also estimated by RNA gel blot hybridization. Two probes were designed based on the predicted Badh2 cDNA sequence (see Supplemental Figure 6 online). Probe A covers the region from ~74 to 244 bp, and this probe hybridizes specifically to the complete Badh2 transcript. Probe B corresponds to the Badh2 cDNA segment from 367 to 1325 bp, which hybridizes to all possible Badh2 transcripts. When hybridized to probe A, almost no signal was detected for all RNA samples (Figure 5). When hybridized to probe B, however, strong signals were observed in two fragrant rice varieties (Nangjing11 and Nipponbare) but not in two fragrant rice varieties (Wuxiangjing and Wuxiangjing9) (Figure 5). The major signal band was estimated to be ~1.5 kb, corresponding to the size of the partial Badh2 transcript (~100-bp 5′ untranslated region [UTR], 1182-bp CDS, and 210-bp 3′ UTR) detected in the RACE reaction.

The data from the above RACE, real-time RT-PCR, RT-PCR, and RNA gel blot analysis indicated the following: (1) the major transcription start point was downstream of the predicted Badh2 start codon, resulting in very few complete Badh2 transcripts but abundant partial Badh2 transcripts; and (2) transcription of the badh2 gene was severely suppressed in fragrant rice varieties.

**The Intact BADH2 Protein Reduces the 2AP Levels**

As multiple Badh2 transcripts were present in nonfragrant rice varieties, we were anxious to know which one in fact is associated with rice fragrance. Therefore, we overexpressed both the complete and partial Badh2 coding sequences in the fragrant rice recipient Wuxiangjing to examine the effect on fragrance. When driven by the CaMV35S promoter, overexpression of the predicted complete Badh2 CDS (pCAM35S:GL) rendered transgenic lines nonfragrant and caused a reduction in 2AP levels (Table 3; see Supplemental Table 3 online). CaMV35S-driven overexpression of the partial Badh2 CDS (pCAM35S:GM, pCAM35S:GS, pCAM35S:CM, and pCAM35S:CS) had no influence on rice fragrance (Table 3; see Supplemental Table 3 online).

Additional evidence that BADH2 inhibits the production of rice fragrance was obtained from our protein gel blot analysis. The intact 55-kD BADH2 protein with 503 amino acids was observed in the nonfragrant rice varieties (Nanjing11 and 3037) as well as in nonfragrant transgenic lines harboring either the native Badh2 gene or the CaMV35S-driven complete Badh2 CDS (Figure 6). However, no BADH2 protein appeared in nontransgenic lines (Table 3).

![Figure 3. The 5′- and 3′-RACE Products of the Badh2 Coding Sequence.](image-url)
in the fragrant rice varieties (Wuxiangjing9 and Suyunuo), non-transgenic lines, or fragrant transgenic lines with the CaMV35S-driven partial Badh2 CDS (Figure 6). We concluded that the intact 503-amino acid BADH2 protein encoded by the complete Badh2 gene inhibits 2AP synthesis and thus renders rice non-fragrant.

The Predicted Three-Dimensional Structure of BADH2

Screening of the protein database allowed us to identify a human mitochondrial aldehyde dehydrogenase (ALDH2; Protein Data Bank code 1o04; x-ray resolution, 1.42 Å) as the structural template for BADH2 (Perez-Miller and Hurley, 2003). The template ALDH2 shared 42% sequence identity with BADH2 (see Supplemental Figure 7 online). The predicted three-dimensional model of BADH2 could be divided into three domains: a NAD binding domain (residues 9 to 124 and 152 to 262), an oligomerization domain (residues 129 to 151 and 480 to 486), and a substrate binding domain (residues 263 to 464) (Figure 7A). Based on the active sites of ALDH2 (Johansson et al., 1998), eight residues conserved between ALDH2 and BADH2 were annotated as the potential active sites in BADH2. Two residues, Asn-162 and Cys-294, acting in a catalytic role, are responsible for interacting with the substrate oxygen. Six residues, Tyr-163, Leu-166, Trp-170, Glu-260, Cys-453, and Trp-459, are predicted to form the substrate binding pocket (Figure 7B).

Figure 4. Relative Abundance of different Badh2 Transcripts. (A) Real-time RT-PCR. Tissues A, B, and C represented leaf tissues of Nanjing11 (A), immature seeds 14 d after the flowering of Nanjing11 (B), and leaf tissues of Wuxiangjing (C). Total RNA from these three tissues was extracted and used for real-time PCR. Error bars represent the SD of the transcript levels determined from the three independent real-time PCRs. (B) RT-PCR. 5’-RACE products from the leaf tissue of Wuxiangjing were subjected to RT-PCR. Lanes 1, 2, and 3 corresponded to use of the primer pairs FP1/RP, FP2/RP, and FP3/RP, respectively.

Figure 5. Relative Abundance of Different Badh2 Transcripts Estimated by RNA Gel Blot Hybridization. (A) Probe A (lanes A) and probe B (lanes B) were synthesized and digoxigenin-labeled with an asymmetry PCR labeling system. Probes A and B were analyzed by electrophoresis on a 1% agarose gel to check their quantities and qualities. (B) Total RNA extracted from four rice varieties was analyzed by electrophoresis on a 1.2% formaldehyde agarose gel. (C) Images of RNA gel blot hybridization using probes A and B. Very weak signals were present in two fragrant rice varieties (lanes 1 and 3). M, RNA marker. Lanes 1 to 4 in (B) and (C) corresponded to the rice cv Wuxiangjing (japonica, badh2-E7), Nanjing11 (indica, Badh2), Wuxiangjing9 (japonica, badh2-E2), and Nipponbare (japonica, Badh2), respectively.
The badh2-E7 allele, due to its 8-bp deletion in exon 7, encodes a presumably truncated BADH2 that lacks 252 C-terminal residues. The missing C-terminal amino acids cover the entire substrate binding domain and partial oligomerization domains, thus rendering the truncated BADH2 nonfunctional. Likewise, the badh2-E2 allele encodes an even shorter truncated BADH2 (82 residues) that is not predicted to have any function as an aldehyde dehydrogenase.

**Subcellular Localization of BADH2**

Because BADH2 is highly expressed during cell division, young panicles of rice plants were used to determine its subcellular localization. Immunodetection using an anti-BADH2 antibody and a fluorescein-conjugated secondary antibody showed strong fluorescent signals for BADH2 in the inflorescence meristem during cell division. A significant amount of BADH2 signal was seen in the cytoplasm. By contrast, none of the BADH2 signal was observed to be localized in the nucleus (Figure 8).

**In Vitro Expression of the Badh2/badh2 Alleles**

We predicted that in vitro expression of the Badh2/badh2-E7 cDNAs would result in the intact 503–amino acid polypeptide (encoded by the complete Badh2 cDNA), the partial 393–amino acid polypeptide (encoded by the partial Badh2 cDNA), the truncated 251–amino acid polypeptide (encoded by the complete badh2-E7 cDNA), and the truncated 141–amino acid polypeptide (encoded by the partial badh2-E7 cDNA) BADH2 proteins, respectively. Since each construct expressed a fusion protein in which the tagged portion (Nus, His, and S tags) was ~61 kD, the fusion proteins were estimated to be 116 kD (PET43.1a-NFPCDNA), 104 kD (PET43.1a-NFPCDNA), 88 kD (PET43.1a-FcCDNA), and 76 kD (PET43.1a-FpCDNA), respectively. We observed the expected fusion proteins to be produced in _Escherichia coli_ in an isopropylthio-β-D-galactoside–dependent manner (see Supplemental Figure 8 online), confirming that these cDNAs could produce the appropriate proteins.

**BADH2 Exhibits Aldehyde Dehydrogenase Activity**

The predicted three-dimensional structure of BADH2 suggested that it belongs to the NAD-dependent aldehyde dehydrogenase family, characterized by the typical Aldedh substrate binding domain (Steinmetz et al., 1997; http://pfam.sanger.ac.uk/search/sequence). The BADH enzymes in sugar beet (Beta vulgaris), spinach (_Spinacia oleracea_), and oat (_Avena sativa_) all showed broad substrate specificities, catalyzing the oxidation not only of betaine aldehyde (Bet-ald) but also of other ω-aminaldehyde (Trossat et al., 1997; Incharoensakdi et al., 2000; Livingstone et al., 2003). To elucidate the role of BADH2 in 2AP biosynthesis, we characterized its aldehyde dehydrogenase activity as well as its substrate specificity.

The intact BADH2 (503 residues) and partial BADH2 (393 residues) were expressed in _E. coli_ and purified using nickel-nitritoltriacetic acid agarose (see Supplemental Figure 9 online). The intact BADH2 showed high Bet-ald dehydrogenase activity, with a rapid increase in A~340 within the first 10 min after the enzymatic reaction (Figure 9). In addition, the intact BADH2 also showed strong 4-aminobutyraldehyde (AB-ald) and 3-amino-propionaldehyde (AP-ald) dehydrogenase activities (Figure 9). By contrast, the partial BADH2 polypeptide showed only background activity for all three aldehyde substrates (Figure 9). These data indicate that the intact BADH2 of rice, like BADH of sugar beet and spinach, showed high aldehyde dehydrogenase activity and wide substrate specificities.
DISCUSSION

We have shown that the badh2 locus of rice constitutes the fgr gene that determines fragrance. Of three Fgr candidates in the rice genomic region to which Fgr (corresponding to the lack of fragrance) was mapped previously, only Badh2 significantly reduced 2AP content and thus rendered fragrant rice nonfragrant. A nonfunctional allele, either badh2-E7 or badh2-E2, is required for rice fragrance. The presence of the dominant (Fgr) or recessive (fgr) allele, determining the nonfragrant versus the fragrant trait, respectively, corresponds completely with the presence of the functional Badh2 or null badh2 allele, respectively.

Compared with the functional Badh2 allele, both nonfunctional badh2-E2 and badh2-E7 alleles had very low transcription levels, as revealed by real-time RT-PCR and RNA gel blot analysis. This indicates that loss-of-function mutations in Badh2 severely suppress mRNA transcription. Theoretically, deletions in the Badh2 gene sequence, for example, the observed 7-bp deletion in exon 2 and an 8-bp deletion in exon 7, would cause frame shifts and result in truncated BADH2 proteins. This was confirmed by in vitro expression in which the Badh2 cDNA resulted in an intact 503–amino acid BADH and the badh2-E7 cDNA resulted in a truncated 251–amino acid BADH (see Supplemental Figure 8 online). However, no such truncated BADH2 proteins were detected in vivo in the fragrant rice cv Wuxiangjing9 and Suyunuo by protein gel blot hybridization (Figure 6). This suggests that the major reason that the badh2 alleles are nonfunctional is the suppression of both transcription and translation, rather than being a result of the truncation of the BADH2 protein.

We conducted a RACE analysis to determine the transcription start point and found that the 5'-RACE products for both the Badh2 and badh2 alleles were much shorter than expected. Subsequent real-time RT-PCR, RT-PCR, and RNA gel blot analyses all showed that the complete Badh2 transcript was much less abundant than partial Badh2 transcripts. However, protein gel blots confirmed that the BADH2 protein had a mass

Figure 7. The Three-Dimensional Structure of BADH2 and the Annotated Active Sites.

(A) Representation of BADH2. Domain coloring is as follows: the oligomerization domain is shown in red, the substrate binding domain is shown in orange, and the NAD binding domain is shown in green. To position the NAD cofactor in the predicted model, the NAD cofactor from the crystal structure of ALDH2 is superimposed into the predicted BADH2 model and is indicated by gray spheres.

(B) The annotated active sites of BADH2. Asn-162 and Cys-294 are catalytic residues that are predicted to interact with the substrate oxygen. Tyr-163, Leu-166, Trp-170, Glu-260, Cys-453, and Trp-459 are predicted to form the substrate binding pocket. The side chain atoms in these highlighted residues are represented as sticks and are colored as follows: green, carbon; blue, nitrogen; red, oxygen; orange, sulfur.

Images in both panels were prepared using PyMol (http://pymol.sourceforge.net).

Figure 8. Indirect Subcellular Immunodetection of BADH2.

(A) Green signal represents BADH2 detected with fluorescein-conjugated goat anti-rabbit antibody.

(B) Blue signal indicates nuclei stained with 4,6-diamidino-phenylindole (DAPI).

(C) The merged image of BADH2 and DAPI-stained nuclei.
Analysis of these data indicated that (1) full-length Badh2 gene resulted in low levels of BADH2 protein (Figure 6).

Badh2 be maintained in an equimolar ratio with foniopropionaldehyde, AP-ald, and AB-ald. AB-ald is known to substrates structurally similar to Bet-ald, such as 3-dimethylsulfoxoniopropionaldehyde, AP-ald, and AB-ald. AB-ald is known to substrates structurally similar to Bet-ald, such as 3-dimethylsulfoxide (SD from three independent experiments with 1 mM Bet-ald and 50 μM AB-ald). Consumption of AB-ald by converting it into GABA inhibits 2AP synthesis, whereas the accumulation of AB-ald results in increased 2AP synthesis.

In this study, the intact BADH2 protein encoded by the complete Badh2 gene sequence was shown to influence this critical switch, possibly due to its strong AB-ald dehydrogenase activity. From its activity in vitro, we conclude that, in nonfragrant rice, the BADH2 enzyme converts AB-ald into GABA, inhibiting 2AP biosynthesis. In fragrant rice lacking intact BADH2, failure to convert AB-ald into GABA due to the absence of BADH2 enzymatic activity results in AB-ald accumulation, which activates 2AP biosynthesis. Interestingly, the low level of BADH2 detected in some transgenic lines overexpressing Badh2 might not completely inhibit the consumption of AB-ald, resulting in a small quantity of 2AP (see Supplemental Table 3 online). Apart from BADH2, another homologous protein, BADH1, is encoded by the rice genome. Although we have not yet investigated the enzymatic activity of BADH1, we predict that BADH1 will show a low affinity for AB-ald but high affinities for other aldehyde substrates. It is unlikely that BADH1 has the same aldehyde dehydrogenase activity and substrate specificities as BADH2 in rice, since BADH2 clearly confers the nonfragrant trait and loss of 2AP accumulation.

In higher plants, betaine is well known as a nontoxic or protective cytoplasmic osmolyte, allowing normal growth of plants in a saline or arid environment. Betaine is synthesized via a two-step oxidation of choline, and the second step (from Bet-ald to betaine) is catalyzed by BADH. In barley, two BADH isozymes (BBD1 and BBD2) were reported, and both of them are induced to higher levels by salt, drought, and abscisic acid treatments (Nakamura et al., 2001). The rice BADH2 protein in this study showed high betaine aldehyde dehydrogenase activity, suggesting that Badh2 also may play a key role in osmoregulation in nonfragrant rice. However, the absence of BADH2 in fragrant rice did not negatively affect normal growth. For example, the Thailand fragrant rice variety Khao Dawk Mali 105.

Because the absence of BADH2 protein results in fragrance, this suggests that Badh2 is not directly involved in 2AP biosynthesis. Alternative possibilities to explain the effect of BADH2 are that the BADH2 enzyme is involved in a competing pathway in which one of the 2AP precursors serves as a BADH2 substrate (Bradbury et al., 2005) or that BADH2 participates in 2AP catabolism. Trossat et al. (1997) showed that the sugar beet BADH catalyzes the oxidation not only of Bet-ald but also of other substrates structurally similar to Bet-ald, such as 3-dimethylsulfoxoniopropionaldehyde, AP-ald, and AB-ald. AB-ald is known to be maintained in an equimolar ratio with 3-1-pyrroline, an intermediate 2AP precursor. On the other hand, AB-ald can be converted into 4-aminobutyric acid (GABA) (A. Vanavichit, T. Yoshihashi, S. Wanchana, S. Areekit, D. Saengsraku, W. Kamolsukuyunyong, J. Lanceras, T. Toojinda, and S. Tragoonrung, unpublished data). It was reported that GABA was found in lower amounts in leaves of the aromatic isogenic line than the nonaromatic counterpart (A. Vanavichit, T. Yoshihashi, S. Wanchana, S. Areekit, D. Saengsraku, W. Kamolsukuyunyong, J. Lanceras, T. Toojinda, and S. Tragoonrung, unpublished data). Therefore, AB-ald levels appear to be an important factor regulating the rate of 2AP biosynthesis.

Because the absence of BADH2 protein results in fragrance, this suggests that Badh2 is not directly involved in 2AP biosynthesis. Alternative possibilities to explain the effect of BADH2 are that the BADH2 enzyme is involved in a competing pathway in which one of the 2AP precursors serves as a BADH2 substrate (Bradbury et al., 2005) or that BADH2 participates in 2AP catabolism. Trossat et al. (1997) showed that the sugar beet BADH catalyzes the oxidation not only of Bet-ald but also of other substrates structurally similar to Bet-ald, such as 3-dimethylsulfoxoniopropionaldehyde, AP-ald, and AB-ald. AB-ald is known to be maintained in an equimolar ratio with 3-1-pyrroline, an intermediate 2AP precursor. On the other hand, AB-ald can be converted into 4-aminobutyric acid (GABA) (A. Vanavichit, T. Yoshihashi, S. Wanchana, S. Areekit, D. Saengsraku, W. Kamolsukuyunyong, J. Lanceras, T. Toojinda, and S. Tragoonrung, unpublished data). It was reported that GABA was found in lower amounts in leaves of the aromatic isogenic line than the nonaromatic counterpart (A. Vanavichit, T. Yoshihashi, S. Wanchana, S. Areekit, D. Saengsraku, W. Kamolsukuyunyong, J. Lanceras, T. Toojinda, and S. Tragoonrung, unpublished data). Therefore, AB-ald levels appear to be an important factor regulating the rate of 2AP biosynthesis. Consumption of AB-ald by converting it into GABA inhibits 2AP synthesis, whereas the accumulation of AB-ald results in increased 2AP synthesis.

In this study, the intact BADH2 protein encoded by the complete Badh2 gene sequence was shown to influence this critical switch, possibly due to its strong AB-ald dehydrogenase activity. From its activity in vitro, we conclude that, in nonfragrant rice, the BADH2 enzyme converts AB-ald into GABA, inhibiting 2AP biosynthesis. In fragrant rice lacking intact BADH2, failure to convert AB-ald into GABA due to the absence of BADH2 enzymatic activity results in AB-ald accumulation, which activates 2AP biosynthesis. Interestingly, the low level of BADH2 detected in some transgenic lines overexpressing Badh2 might not completely inhibit the consumption of AB-ald, resulting in a small quantity of 2AP (see Supplemental Table 3 online). Apart from BADH2, another homologous protein, BADH1, is encoded by the rice genome. Although we have not yet investigated the enzymatic activity of BADH1, we predict that BADH1 will show a low affinity for AB-ald but high affinities for other aldehyde substrates. It is unlikely that BADH1 has the same aldehyde dehydrogenase activity and substrate specificities as BADH2 in rice, since BADH2 clearly confers the nonfragrant trait and loss of 2AP accumulation.

In higher plants, betaine is well known as a nontoxic or protective cytoplasmic osmolyte, allowing normal growth of plants in a saline or arid environment. Betaine is synthesized via a two-step oxidation of choline, and the second step (from Bet-ald to betaine) is catalyzed by BADH. In barley, two BADH isozymes (BBD1 and BBD2) were reported, and both of them are induced to higher levels by salt, drought, and abscisic acid treatments (Nakamura et al., 2001). The rice BADH2 protein in this study showed high betaine aldehyde dehydrogenase activity, suggesting that Badh2 also may play a key role in osmoregulation in nonfragrant rice. However, the absence of BADH2 in fragrant rice did not negatively affect normal growth. For example, the Thailand fragrant rice variety Khao Dawk Mali 105.
(badh2-E7/badh2-E7) grows well in the arid region Tung Kula Rong Hai (Yoshishashi et al., 1999; Bradbury et al., 2005). This suggests that other Badh genes in the rice genome can compensate for the defective null badh2 alleles and allow for tolerance to salinity and drought stresses. Apart from the known Badh1 gene on chromosome 4, we found at least two other genes encoding putative BADHs on chromosome 7. The rice Badh1 gene on chromosome 4 showed high homology with the Badh1 genes in barley and sorghum (Sorghum bicolor) (Bradbury et al., 2005). More research is required to elucidate the functions of these Badh genes in both tolerance to salinity/drought stresses and the regulation of 2AP biosynthesis.

In summary, the intact 503-amino acid BADH2 encoded by the complete Badh2 gene inhibits 2AP biosynthesis by converting AB-ald (a presumed 2AP precursor) to GABA, whereas the absence of BADH2 due to nonfunctional badh2 alleles results in AB-ald accumulation and thus turns on the pathway toward 2AP biosynthesis.

METHODS

Construction of Rice BAC Libraries and Identification of Positive BAC Clones

Both a local Chinese fragrant rice cv, Suyunuo (Oryza sativa japonica), and a nonfragrant rice cv, Nanjing11 (Oryza sativa indica), were used to construct BAC libraries. The leaf tissues were collected from etiolated shoots after a 2-week culture at 25°C. The construction and screening of the BAC library were performed as described by Xu et al. (2001) with some modifications. Two restriction enzymes, HindIII and BamHI, were used to partially digest rice genomic DNA. The markers in the fgr region, including L02, L03, L04, L05, and L06 (see Supplemental Table 4 online), were used to screen both the Suyunuo and Nanjing11 BAC libraries.

Cloning of the Fgr/fgf Candidates from the Positive BAC Clones

The complete genomic sequences of the Fgr region from both Nipponbare and 93-11 were obtained from sequences deposited in the databases (http://rise.genomics.org.cn and http://www.gramene.org). The restriction sites flanking the three candidate genes (Cah, Mcc2, and Badh2) were identified and used to subclone the intact Fgr/fgf candidates from the corresponding BAC clones. For our studies, an intact candidate gene must contain at least the 1.3-kb promoter, the full coding region, and the 3′ UTR region. For each of the three genes, a DNA fragment with the intact candidate gene was then subcloned into the expression vector pCAM (http://www.cambia.org.au/daisyl/cambia/home.htm).

Rice Transformation and Complementation Tests

The constructs containing Fgr candidates were introduced into Agrobacterium tumefaciens strain EHA105 (Hood et al., 1993) and then used to infect the fragrant Wuxiangjing callus. Plantlets regenerated from hygromycin-resistant calli were transplanted into large culture pots and grown in a culture room under a 14-h photoperiod with a room temperature of 26 ± 2°C. Leaf tissue of each independently regenerated plantlet was harvested to examine both genotype and fragrance. The marker tagged to each Fgr candidate was used to identify positive transgenic lines (primers listed in Supplemental Table 4 online). The fragrance trait for each regenerant plantlet was first investigated using the smelling method (Chen et al., 2006). Thereafter, an aliquot of 0.3 g fresh leaf was subjected to 2AP extraction at 65°C for 1 h in a solution consisting of ethanol and 0.459 ng/mL 2,4,6-trimethylpyridine as an internal standard (Bergman et al., 2000). After centrifugation at 12,000 rpm (relative centrifugal field of 13,400g) for 30 min, the supernatant was directly used to measure 2AP concentrations with the GC-MS method (Itani et al., 2004). The smelling method is unreliable due to its inherent subjectivity; while the GC-MS method can directly estimate 2AP level; therefore, it is more accurate than the smelling method in evaluating rice fragrance.

Real-Time RT-PCR

Real-time RT-PCR was conducted on both the nonfragrant rice cv Nanjing11 and the fragrant rice cv Wuxiangjing. Rice tissues, including the initial fully expanded leaf, the last fully expanded leaf, the stem, young panicles at the booting stage, panicles at the heading stage, panicles 15 d after the heading stage, and roots, were all collected for total RNA extraction. After digestion with RNase-free DNase I, an aliquot of 500 ng of total RNA was used for one-step real-time RT-PCR using a Quantiti-SYBR Green RT-PCR kit (Qiagen). Each RNA sample was assayed in triplicate. For a negative control, total RNA without reverse transcription was directly used for real-time PCR. With a reference Actin gene, the relative amount of the Badh2 transcript is presented as 2^\(-\Delta CT\) according to the ΔCT method described in the Real-Time PCR Applications Guide (Bio-Rad Laboratories; http://www.bio-rad.com). A fluorescence threshold was set up at a value of <0.05 to ensure the amplification in the logarithmic phase. The ΔCT value was obtained by subtracting the Ct (threshold cycles) number of the Actin gene from that of the Badh2 gene (ΔCT = Ct_Badh2 - Ct_Actin). The ΔCT value was converted to the linear form in terms of 2^\(-\Delta CT\) for statistical analysis. Analysis of variance was performed using the SAS system (SAS Institute), and the expression level of each sample was represented as mean 2^\(-\Delta CT\) ± SD.

RACE Reactions and Cloning of cDNAs

Leaf tissues were collected from both the nonfragrant rice cv Nanjing11 and the fragrant rice cv Wuxiangjing. Total RNA was isolated from leaf tissue with the LiCl method (Sambrook and Russell, 2001), followed by mRNA isolation using the PolyATtract System III kit (Promega). The synthesis of first-strand cDNA (RACE-ready cDNA) was catalyzed by PowerScript reverse transcriptase (BD SMART RACE cDNA amplification kit; BD Biosciences). The Badh2-specific primers were designed for the Fgr/fgf alleles based on the predicted coding sequences and were used for the amplification of 3′- and 5′-RACE-ready cDNAs (see Supplemental Table 4 online). The amplicons (RACE products) were then cloned into the vector pGEM-T (Promega) for sequencing. The sequences of 3′- and 5′-RACE products were merged into full-length cDNAs based on the condition that both RACE products shared the identical overlapping region.

RNA Gel Blot Analysis

Leaf tissues were collected from four rice varieties, Wuxiangjing (japonica, badh2-E7), Nanjing11 (indica, Badh2), Wuxiangjing9 (japonica, badh2-E2), and Nipponbare (japonica, Badh2), and used for total RNA extraction with the LiCl method (Sambrook and Russell, 2001). For RNA gel blot analysis, all RNA samples (~20 μg each) were separated on 1.2% formaldehyde agarose gels and then blotted onto a nylon membrane (Amersham Pharmacia Biotech.). The probes were synthesized and digoxigenin-labeled with an asymmetric PCR labeling system (Zhang et al., 2005). RNA gel blot hybridization was repeated twice to ensure the reproducibility of hybridization signals.

Functional Analysis of Badh2 Transcripts

We obtained cDNAs corresponding to the complete Badh2 transcript (1512-bp CDS) and two partial Badh2 transcripts (1182-bp CDS and 1095-bp CDS) via RT-PCR. In parallel, we amplified genomic DNA segments
corresponding to these three Badh2 transcripts from the pCAM-Badh2 construct. All Badh2 fragments were cloned into the expression vector pCAM under the control of the CaMV3SS promoter. In total, three genomic Badh2 overexpression constructs (pCAM3SS:GL, pCAM3SS:GM, and pCAM3SS:GS) and two Badh2 cDNA overexpression constructs (pCAM3SS:CM and pCAM3SS:CS) were successfully made. Here, L, M, and S represent the long (1512 bp, complete CDS), medium (1182 bp, partial CDS), and short (1095 bp, partial CDS) Badh2 transcripts, respectively. These five constructs were separately transformed into the fragrant rice recipient Wuxiangjing via Agrobacterium. The positive transgenic plantlets were subjected to 2AP assay with the GC-MS method.

**Protein Gel Blot Analysis**

Leaf tissues were harvested for total protein extraction using the trichloroacetic acid–acetone precipitation method (Darnel et al., 1988). The quantity of each sample was estimated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). All protein samples (quantity of each sample was estimated using a NanoDrop ND-1000 spectrophotometer). The fusion protein containing the Nus flag was expressed by induction with 0.5 mM isopropylthio-β-D-galactoside at 18°C overnight (see Supplemental Figure 8 online). The induced E. coli cells were collected and lysed by sonication on ice. The fusion proteins present in the supernatant were subjected to electrophoresis on 12% SDS-polyacrylamide gels.

The fusion proteins from PET43.1a-NFccDNA and PET43.1a-NFPCcDNA were purified using nickel-nitritetriacetic acid agarose affinity chromatography columns (Qiagen) and assayed for their aldehyde dehydrogenase activities. The Bet-ald chloride was purchased from Sigma-Aldrich. The diethyldiacetals of BA-ald and AP-ald were obtained from Sigma-Aldrich and TCI American, respectively. The preparation of three aldehyde substrates (Bet-ald, AP-ald, and AN-ald) and standard enzymatic reactions were conducted as described by Trossat et al. (1997). The enzymatic activity of BADH2 was measured at 0, 10, 20, 30, and 60 min following the initiation of reactions by monitoring the increase in A₅₀₀ with a spectrophotometer. Three replicates for each reaction were performed.

**Subcellular Immunodetection of BADH2**

The young panicles (~3 cm in length) were harvested and fixed for 1 h with 4% (w/v) paraformaldehyde in 20 mL of 1× PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 2 mM KH₂PO₄, pH 8.0) followed by three washes in 1× PBS (5 min each). The spikelet was squashed in 1× PBS solution on a microscope slide covered with a cover slip. After soaking in liquid nitrogen and removing the cover slip, the slide was dehydrated through an ethanol series (70, 90, and 100%) prior to being soaked in liquid nitrogen and removing the cover slip, the slide was dehydrated through an ethanol series (70, 90, and 100%) prior to being embedded in paraffin wax. Sections (7 μm) were taken and deparaffinized in xylene. The slides were rehydrated through a series of ethanol washes (100, 90, and 70%) and then washed in Tris-HCl, pH 7.5, 0.15 M NaCl, and 0.5% blocking reagent. After three rounds of washing in PBS, fluorescein-conjugated goat anti-rabbit antibody was added to the slides. The slides were first incubated with a humid chamber at 37°C for 3 h with BADH2 antibody diluted 1:500 in TNB buffer (0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl, and 0.5% blocking reagent). After three rounds of washing in PBS, fluorescein-conjugated goat anti-rabbit antibody was added to the slides. The slides were counterstained with 4,6-diamidino-phenylindole in an antifade solution (Vector Laboratories). The signals were observed using a Leica TCS SP5 confocal system. Captured images were enhanced and pseudocolored by Adobe Photoshop CS2 software.

**Supplemental Data**

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Maps of Three Fgr Constructs Derived from the Nonfragrant cv Nanjing11 and Three fgr Constructs Derived from the Fragrant cv Suyuno.

**Supplemental Figure 2.** Characterization of Nucleotide Acid Sequences and Amino Acid Sequences of the Badh1, Badh2, badh2-E2, and badh2-E7 Alleles.

**Supplemental Figure 3.** Identification of Transgenic Lines Using the Gene-Tagged Markers.

**Supplemental Figure 4.** Chromatogram of Rice Leaf Tissue Extracts.

**Supplemental Figure 5.** Genotypes at the Exogenous Badh2 Gene in T1 Progeny.

**Supplemental Figure 6.** Maps of the Three Badh2 Transcripts and Locations of the Primers and Probes Used for Investigation of Badh2 Transcription.

**Supplemental Figure 7.** Sequence Alignment between BADH2 and ALDH2.

**Supplemental Figure 8.** In Vitro Expression of the Badh2/badh2 cDNAs in E. coli Cells.

**Supplemental Figure 9.** Purification of the Fusion Proteins Containing the Intact or Partial BADH2.

**Supplemental Table 1.** 2AP Levels in Each Transgenic and T1 Progeny. 

**Supplemental Table 2.** 2AP Levels in Transgenic and Nontransgenic Lines Using the Four Fragrant Rice Recipients Containing the badh2-2 Alleles.

**Supplemental Table 3.** 2AP Levels in Each Nontransgenic Line and T1 Progeny.

**Supplemental Table 4.** 2AP Levels in Each Transgenic Line and T1 Progeny.

**Supplemental Table 5.** 2AP Levels in Each Transgenic Line and T1 Progeny.
Supplemental Table 4. List of All Primers Used in This Research.

Supplemental Data Set 1. Sequence Alignment among the \textit{Badh2} cDNA and Its Genomic Genes from Various Rice Varieties.

ACKNOWLEDGMENTS

We thank Jiming Jiang and Jason Walling for their critical review of the manuscript. We also are grateful to Jufei Lu for his assistance in the field testing. HongMei Wang and Qing Cao for their guidance in tissue culture, Xin Li and ShuZhu Tang for their help in collection of rice varieties, and Lu Jiang, ChunXia Liu, and Yan Zhang for their technical assistance. This research was financially supported by the National High-Tech 863 Program of the People’s Republic of China (Grant 2002AA224041) and the National Natural Science Foundation of China (Grant 30771318).

Received February 20, 2008; revised June 4, 2008; accepted June 17, 2008; published July 3, 2008.

REFERENCES


Trossat, C., Rathinasabapathi, B., and Hanson, A.D. (1997).


Badh2, Encoding Betaine Aldehyde Dehydrogenase, Inhibits the Biosynthesis of 2-Acetyl-1-Pyrroline, a Major Component in Rice Fragrance

Saihua Chen, Yi Yang, Weiwei Shi, Qing Ji, Fei He, Ziding Zhang, Zhukuan Cheng, Xiangnong Liu and Mingliang Xu

Plant Cell 2008;20:1850-1861; originally published online July 3, 2008;
DOI 10.1105/tpc.108.058917

This information is current as of June 29, 2017