Proceedings of 12th Agricultural Research Symposium (2013) 11-15

Detection of Novel Allele and Protein Structure of Mutated Fragrant Gene in Sri Lankan Aromatic Rice

D.M.S. DISSANAYAKA[1], N.S. KOTTEARACHCHI[1],
O.V.D.S.J. WEERASENA[2] and W.A.M. PEIRIS[2]

[1]*Department of Biotechnology, Faculty of Agriculture and Plantation Management, Wayamba University of Sri Lanka, Makandura, Gonawila (NWP)*

[2]*Institute of Biochemistry, Molecular Biology and Biotechnology, University of Colombo, Colombo 03*

ABSTRACT

Fragrance in rice is caused by the mutations occurred in *badh2* (betaine aldehyde dehydrogenase) gene. Other than the predominant allele, *badh2.1* additional mutations have been identified in *badh2* gene. As the fragrance originated in Sri Lankan fragrance rice, cannot be explained by *badh2.1* allele, this study was carried out to detect the mutation in the 14th exon of *badh2* gene and if mutated, to predict the aberrant protein structure due to novel allele. Rice DNA was extracted and amplified the 14th exon region by sequence specific primers. Purified PCR fragments were sequenced. The sequence analysis revealed that fragrant accessions in Sri Lanka possessed 'G' insertion in the 14th exon creating *badh2.7* allele and it produced only 476 amino acids in comparison to the wildtype. Three dimensional structure for BADH2 and its mutant was constructed using PsAMADH2 (amino aldehyde dehydrogenase 2 from *Pisum sativum*) as template. Predicted three dimensional model of BADH2 could be divided into three domains: a coenzyme binding domain, an oligomerization domain, and a substrate binding domain. Predicted three dimensional protein structure for mutant showed loss of part of oligomerization and coenzyme binding domain which would putatively abolish protein function and result in fragrance. Therefore, it can be assumed that abolished protein structure created by novel mutation, 'G' insertion, might be the causal factor for the fragrance in most of Sri Lankan aromatic rice.

KEYWORDS: Fragrant gene, Mutation, Protein structure prediction, Rice

INTRODUCTION

One of the highly valued grain quality traits in rice is fragrance. It determines the premium price in both domestic and international markets. The gene responsible for fragrance was a single recessive gene (*fgr*) located on chromosome eight (Sood and Siddiq, 1978; Huang *et al.*, 1994; Jin *et al.*, 2003). Bradbury *et al.* (2005) suggested that *fgr* gene encodes betaine aldehyde dehydrogenase (*badh2*) and reported an eight base pairs (bp) deletion and three single nucleotide polymorphisms (SNPs) in seventh exon that created the recessive *badh2.1* allele. This functional mutation created a premature stop codon leading to loss of BADH2 protein function and account for the accumulation 2-acetyl-1-pyrroline (2AP); the major compound responsible for the characteristic aroma (Buttery *et al.*, 1982; Paule and Powers, 1989; Petrov *et al.*, 1996). 2AP is found in all parts of plants of fragrant rice varieties except in roots (Buttery *et al.*, 1983). Hence, biochemical pathway that leads to 2AP synthesis has not been fully understood, two pathways of 2AP biosynthesis in rice were proposed: BADH2 dependent 2AP synthesis (Bradbury *et al.*, 2008; Chen *et al.*, 2008) and BADH2 independent 2AP synthesis (Huang *et al.*, 2008).

The gene model for *badh2* contains 15 exons and 14 introns (Bradbury *et al.*, 2005). Kovach *et al.* (2009) have identified nine mutations scattered in the exons 1, 2, 10, 13 and 14 of *badh2* gene that are responsible for the fragrant phenotype in addition to the mutation exist in the 7th exon and they were named starting from *badh2.2* to *badh2.10* consecutively. Hence, one base pair insertion in the 14th exon named as *badh2.7*, has been identified as the causal factor for the fragrance of most *Oriza sativa* subpopulation, *aus* habituation in India, Sri Lanka and Nepal. Proving this fact it was identified that most of aromatic rice germplasms in Sri Lanka do not carry *badh2.1* allele but they had shown elevated levels of 2AP (Kottearchchi *et al.*, 2010).

The objective of this study was to detect the presence of *badh2.7* allele in fragrant rice varieties originated in Sri Lanka and to study the molecular conformational changes due to *badh2.7* allele by computational modelling approach.

MATERIALS AND METHODS

Plant Material

Seeds of *Oryza sativa* L., representing aromatic accessions of Sri Lankan, Suwanda

Al (Acc.No. 04366), Suwadal (Acc.No. 10646) and Kuruluwee (Acc.No.04903) were obtained from the gene bank of Plant Genetic Resources Centre, Sri Lanka.

DNA Extraction and PCR Assay

Five seeds from each accession were planted on wetted Petri dishes. Leaves of three week old rice seedlings were used for the extraction of genomic DNA. DNA was extracted from rice leaves using a method previously reported by (Anushka *et al.*, 2008). Extracted DNA samples were quantified by 0.8% agarose gel containing 0.5 µg/ml ethidium bromide. PCR was performed in a Bio-Rad (My cycler ™) Thermal cycler. PCR reaction consisted of 5 µl of diluted template (50 ng/µl), 1.5 µl of 10x PCR buffer, 1.2 µl of 2 mM dNTPs, 1.12 µl of 20 pmol/µl primers and 0.18 µl of 5 u/µl *Taq* DNA polymerase (Dream Taq, Fermentas) in total volume of 15 µl. Amplification of 14th exon was conducted with two specific primers designed based on the sequence of *Oriza sativa indica* (EU770319.1), using National Centre for Biotechnology Information (NCBI) Primer-BLAST tool (Anon, 2012a). The primers sequences were; Forward: 5'CAA GTG AAG GGG ATT G 3' Reverse: 5'ACC AAA GGC ATG ATG TCA GGT CG 3'. Following amplification conditions were maintained: an initial denaturation at 95^0C for 1 min, 35 cycles of 95^0C for 30 sec, 56.9^0C for 30 sec, 72^0C for 1 min and 72^0C for 7 min. Amplified PCR product was electrophoresed (Electrophoresis unit, MUPIDexu, England) on 1.3% agarose gel containing 0.5 µg/ml ethidium bromide. The gel was run at 5 V/cm in 0.5x TBE buffer. Gel was visualized using ultraviolet transilluminator.

Purification of Amplified Product and Sequencing

The separated gel bands were excised under the ultraviolet light. The bands were collected into four eppendorf tubes. Each eppendorf tube was filled with 300 µl of binding buffer (Fermentas, GeneJET™ PCR Purification Kit) and kept at 65^0C for 10 min in a dry bath. Solutions were transferred into four separate DNA binding columns. Columns were centrifuged at 13,000 rpm for 1 min. The column then was placed into fresh tube and 30 µl of distilled water was added. Tubes were heated at 65^0C for 5 min. Finally the solution was centrifuged at 13,000 rpm for 1 min. Sequencing was performed with the Big Dye terminators V2.0 cycle sequencing reaction kit and ABI Prism automated DNA sequencer.

Multiple Sequence Alignment

A multiple sequence alignment was conducted for the DNA sequences obtained from sequencing, using MEGA 4 software. Open reading frames for fragrant rice varieties were constructed using Open Reading Frame (ORF) Finder of National Centre for Biotechnology Information (NCBI) (Anon, 2012b). Amino acid sequences of fragrant rice varieties were aligned with wild type using MEGA 4 software.

Prediction of Three Dimensional Structures of BADH2 Protein and Its Mutant

Homology models for BADH2 protein and its mutant were built using Swiss-Model Automatic Modelling Mode (Anon, 2011a). The crystal structure of plant amino aldehyde dehydrogenase (Protein Data Base code 3IWJB) with resolution of 2.15^0 was used as template. Template identification was carried out with Swiss Model Workspace Template identification tool (Anon 2011b).

RESULTS AND DISCUSSION

Multiple Sequence Alignment

Elevated levels of 2AP of aromatic rice germplasm in Sri Lanka that do not carry *badh2.1* allele raise the question on what causal factors affect the fragrance of such varieties. As the possibility of existence of novel fragrant allele name as *badh2.7* in South Asian region representing *aus* subpopulation (Kovach *et al.*, 2009), 14th exon region of *badh2* gene was sequenced in traditional fragrant accessions. The DNA multiple sequence alignment (Figure 1) of three accession revealed that Suwanda Al (Acc.No. 04366), Suwadal (Acc.No. 10646), Kuruluwee (Acc.No.04903) contained 'G' nucleotide insertion in the 14th exon region of *badh2* gene in comparison to the wild type, (*Oriza sativa indica* group (EU770319.1) confirming the existence of *badh2.7* allele. The Amino acid sequence alignment (Figure 2) revealed that, the gene codes only 476 amino acids and introduces stop codon creating a putatively truncated protein.

Kuruluwee (04903)	TGCTTCTGCCAAGCTCCATGGGGGCGGGAACAAGCGCAGCGGCTTTGGACGCGAGCTCGGAGAAGGGT
Suwanda Al (04366)	TGCTTCTGCCAAGCTCCATGGGGGCGGGAACAAGCGCAGCGGCTTTGGACGCGAGCTCGGAGAAGGGT
Suwadal (10646)	TGCTTCTGCCAAGCTCCATGGGGGCGGGAACAAGCGCAGCGGCTTTGGACGCGAGCTCGGAGAAGGGT
Wild type (EU 770319.1)	TGCTTCTGCCAAGCTCCATGGGG-CGGGAACAAGCGCAGCGGCTTTGGACGCGAGCTCGGAGAAGGGT

Figure 1. Multiple sequence alignment of 14th exon region ('G' insertion indicated by arrow mark)

Mutated BADH2	GIIWVNCSQPCFCQAPWG--------REQ------------------AQRLWTRARRRGH-
BADH2	GIIWVNCSQPCFCQAPWGGNKRSGFGRELGEGGIDNYLSVKQVTEYASDEPWGWYKSPSKL

Figure 2. The amino acid sequence alignment showing the altered sequence coded by *badh2.7*

Prediction of Three Dimensional Structures of BADH2 Protein and Its Mutant

In order to detect the structural changes due to *badh2.7* allele, protein structure was predicted and compared with the wild type BADH2 protein. Three dimensional structure for BADH2 and its mutant was constructed using amino aldehyde dehydrogenase 2 from *Pisum sativum* (PsAMADH2) as a template which was closely related to BADH2 protein with 76.75% sequence identity. The predicted three dimensional model of BADH2 (Figure 3a) could be divided into three domains: a coenzyme binding domain, an oligomerization domain, and a substrate binding domain. The coenzyme binding domain in BADH2 is formed by residues 1-131,152-261 and 453-479. The substrate binding domain spread from 262 to 452 and oligomerization domain formed by residues 132-151 and 480-503. Based on the coenzyme binding and substrate binding sites of PsAMADH2 (Tylichova *et al.*, 2010) nine residues conserved between BADH2 and PsAMADH2. Six residues Glu-188, Thr-159, Lys-185, Thr-242, Ser-239 and Trp-161 are responsible for interacting with coenzyme. Three residues Asn-162, Cys-294, Glu-260 are responsible for interacting with substrate.

'G' nucleotide insertion in the 14th exon encoded truncated BADH2 protein (Figure 3b) that lack part of C-terminal. This C- terminal contains part of oligomerization and coenzyme binding domains. But this truncation does not affect the conserved coenzyme binding sites or catalytic sites. Therefore functional mutation creating *badh2.7* allele critically affects the oligomerization domain in the C-terminal. BADH2 belongs to Aldehyde Dehydrogenase (ALDH) super family (Kotchoni *et al.*, 2010). Oligomerization domain of ALDHs forms intersubunit contacts between monomers and determine the binary and quaternary structure of the protein (Rodriguez-Zavala and Weiner, 2001). Due to dimeric state of BADH2 protein (Wongpanya *et al.*, 2011), loss of oligomerization domain function might influence the maintenance of stable dimmers. This unstable dimmer of BADH2 protein is also disturbs the catalytic function (Munoz-Clares *et al.*, 2010). Ultimately this functional mutation produces non functional protein, which might enhance the 2AP synthesis.

Predicted models accuracy highly dependent on when the sequence identity is more than 50%. Chen *et al.* (2008) used three dimensional model of BADH2 using a human mitochondrial aldehyde dehydrogenase (Protein Data Bank code 1o04) as template which showed only 42% sequence identity. In this study we used PsAMADH2 as template with 76.75% sequence identity as the accuracy of the predicted model is highly dependent on the sequence identity between target and template.

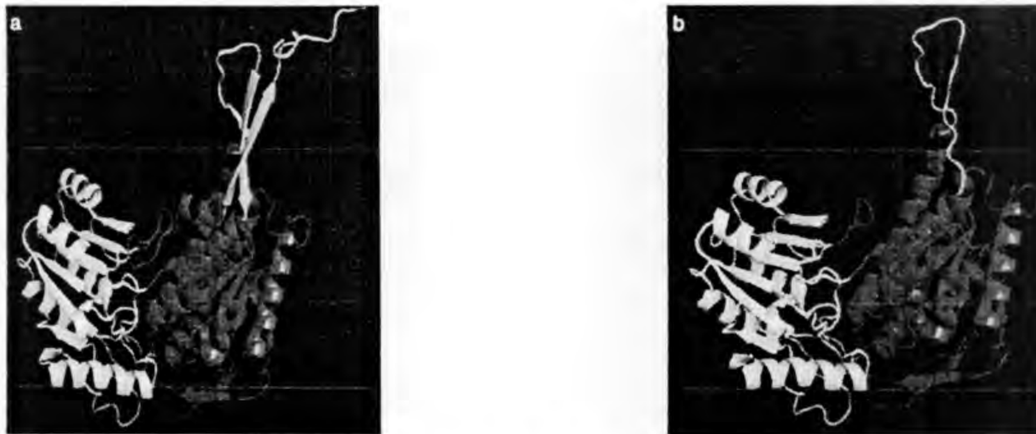

Figure 3. The three dimensional structures of BADH2 and its mutant. (a) Predicted three dimensional structure of BADH2 protein. (b) Predicted three dimensional structure of mutated BADH2 protein.

CONCLUSIONS

The results of the multiple sequence alignment revealed that 'G' insertion in the 14th exon suggesting that there is a possible mutation leading to loss of function in *badh2* gene in Suwanda Al (Acc.No. 04366), Suwadal (Acc.No. 10646) and Kuruluwee (Acc.No.04903). The predicted three dimensional model of BADH2 could be divided into three domains: a coenzyme binding domain, an oligomerization domain, and a substrate binding domain. Protein structure for mutant showed loss of part of oligomerization and coenzyme binding domain which would putatively abolish protein function and result in fragrance. Further studies are necessary to design marker/s associated with *badh2.7* allele and validate using Sri Lankan aromatic rice varieties that could not be detected by the markers developed by Bradbury *et al.* (2005).

ACKNOWLEDGEMENTS

The authors would like to offer their gratitude to the National Research Council (Grant No. 09-11) who financially supported this research work.

REFERENCES

Anon, (2011a). Swiss-Model Automatic Modelling Mode. Available from: http://swissmodel.expasy.org/workspace/. (Accessed on 20th February 2013).

Anon, (2011b). Swiss Model Workspace Template identification tool. Available from:http://swissmodel.expasy.org/workspace/. (Accessed on 19th February 2013).

Anon, (2012a). National Centre for Biotechnology Information (NCBI) Primer-BLAST tool. Available from: http://www.ncbi.nlm.nih.gov/tools/primer-blast/. (Accessed on 13th January 2013).

Anon, (2012b). NCBI, Open Reading Frame (ORF) Finder. Available from: http://www.ncbi.nlm.nih.gov/gorf/gorf.html. (Accessed on 14th January 2013).

Anushka, K., Kottearachchi, N.S. and Attanayaka, D.P.S.T.G. (2008). Identification of fragrance gene (*fgr*) in Sri Lankan rice varieties using Polymerase Chain Reaction based molecular markers. In Proceedings of the 8th Agricultural Symposium, 13-14 August, 2008. Wayamba University of Sri Lanka. 182-187.

Bradbury, L.M.T., Fitzgerald, T.L., Henry, R.J., Jin, Q.S. and Waters, D.L.E. (2005). The gene for fragrance in rice. *Plant Biotechnology Journal*, **3**, 363–370.

Bradbury, L.M.T., Gillies, S.A., Brushett, D.J., Waters, D.L.E. and Henry, R.J. (2008). Inactivation of an aminoaldehyde dehydrogenase is responsible for fragrance in rice. *Plant Molecular Biology*, **68**, 439–49.

Buttery, R.G., Lilng, L.C. and Juliano, B.O. (1982). 2-Acety-1-pyrroline: An important aroma component of cooked rice. *Chemistry & Industry*, **12**, 958–959.

Buttery, R.G., Ling, L.C., Juliano, B.O. and Turnbaugh, J.G. (1983). Cooked rice aroma and 2-acetyl-1-pyrroline. *Journal of Agriculture and Food Chemistry*, **31**, 823–826.

Chen, S., Yang, Y., Shi, W., Ji, Q., He, F., Zhang, Z., Cheng, Z., Liu, X. and Xu, M. (2008). Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. *Plant Cell*, **20**, 1850–61.

Huang, N., McCouch, S.R., Mew, T., Parco, A. and Guiderdoni, E. (1994). Development of an RFLP map from a doubled haploid population in rice. *Rice Genetic Newsletter*, **11**, 134–137.

Huang, T.C., Teng, C.S., Chang, J.L., Chuang, H.S., Ho, C.T. and Wu, M.L. (2008). Biosynthetic mechanism of 2- acetyl-1-pyrroline and its relationship with Δ1-pyrroline-5-carboxylic acid and methylglyoxal in aromatic rice (*Oryza sativa* L.) callus. *Journal of Agriculture and Food Chemistry*, **56**, 7399–404.

Jin, Q.S., Waters, D., Cordeiro, G.M., Henry, R.J. and Reinke, R.F. (2003). A single nucleotide polymorphism (SNP) marker linked to the fragrance gene in rice (*Oryza sativa* L.). *Plant Science*, **165**, 359–364.

Kotchoni, S.O., Jimenez-Lopez, J.C., Gao, D., Edwards, V., Gachomo, E.W., Margam, V.M. and Seufferheld, M.J. (2010). Modelling-Dependent Protein Characterization of the Rice Aldehyde Dehydrogenase (ALDH) Superfamily Reveals Distinct Functional and Structural Features. *PLoS ONE*, **5** (7), e11516.

Kottearachchi, N.S., Priyangani, E.G.D. and Attanayaka, D.P.S.T.G. (2010). Identification of fragrant gene, *fgr*, in traditional rice varieties of Sri Lanka. *Journal of National Foundation Sri Lanka*, **38**(2), 137-141.

Kovach, M.J., Calingacion, M.N., Fitzgerald, M.A. and McCouch, S.R. (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.).*Proceedings of the National Academy of Sciences of the USA*, **106**, 14444–14449.

Munoz-Clares, R.A., Diaz-Sanchez, A.G., Gonzalez-Segura, L. and Montiel, C. (2010). Kinetic and structural features of betaine aldehyde dehydrogenases: Mechanistic and regulatory implications. *Archives of Biochemistry and Biophysics*, **493**, 71–81.

Paule, C.M. and Powers, J.J. (1989). Sensory and chemical examination of aromatic and nonaromatic rices. *Journal of Food Science*, **54**, 343–346.

Petrov, M., Danzart, M., Giampaoli, P., Faure, J. and Richard, H. (1996). Rice aroma analysis: discrimination between a scented and a non-scented rice. *Sciences des Aliments*, **16**, 347–60.

Rodriguez-Zavala, J. and Weiner, H. (2001). Role of the C-terminal tail on the quaternary structure of aldehyde dehydrogenases. *Chemico-Biological Interactions*, 130–132, 151–160

Sood, B.C. and Sidiq, E.A. (1978). A rapid technique for scent determination in rice. *Indian Journal of Genetics* and *Plant Breeding*, **38**, 268–271.

Tylichova, M., Kopecny, D., Morere, S., Briozzo, P., Lenobel, R., Snegaroff, J. and Sebela, M. (2010). Structural and Functional Characterization of Plant Aminoaldehyde Dehydrogenase from *Pisum sativum* with a Broad Specificity for Natural and Synthetic Aminoaldehydes. *Journal of Molecular Biology*, **396**, 870-882.

Wongpanya, R., Boonyalai, N., Thamm achuchourat, N., Horata, N., Arikit, S., Myint, K.M., Vanavichit, A. and Choowongkomon, K. (2011). Biochemical and enzymatic study of rice BADH wild-type and mutants: an insight into fragrance in rice. *Protein Journal*, **30**, 529–538.