Jurnal
Teknologi

Full Paper

MOLECULAR PHYLOGENY AND STRUCTURE PREDICTION OF RICE RFT1 PROTEIN

Article history
Received
8 June 2015
Received in revised form
17 September 2015
Accepted
15 January 2016

Shahkila Mohd Arif[a], Abdulrahman Mahmoud Dogara[a], Nyuk Ling Ma[b], Mohd Shahir Shamsir[c], Sepideh Parvizpour[c], Zaidah Rahmat[a,d]*

[a]Department of Biotechnology and Medical Engineering, [c]Department of Biosciences and Health Sciences, Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]School of Fundamental Science, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia
[d]Innovation Center in Agrotechnology for Advanced Bioprocessing (ICA), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

*Corresponding author
zaidahrahmat@utm.my

Abstract

Rice is one of the most important species in the family of Poaceae. As one of the major crop that is consumed by world population, it is cultivated commercially in many parts of the world. Hence, the phylogeny study of this crop is crucial as a step for improvement of its breeding programs. Phylogenetic relationship among 12 rice cultivars that originated from two common sub-species; Indica and Japonica were inferred by comparing protein sequence data sets derived from its flowering time gene, namely *RFT1* and analyzed using maximum parsimony (MP) method. The predicted structure of RFT1 protein was generated by I-TASSER server and analyzed using YASARA software. The result showed that the cultivars were classified into two major groups, where the first group (Japonica) evolved first followed by the second group (Indica). The findings suggested that some cultivars had a close relationship with each other even it is originates from different varieties. The relationships among these cultivars provide useful information for better understanding of molecular evolution process and designing good breeding program in order to generate new cultivar.

Keywords: Rice, *RFT1*, phylogenetic analysis, predicted structure

Abstrak

Padi merupakan salah satu spesis yang terpenting dalam keluarga Poaceae. Sebagai salah satu tanaman utama yang dimakan oleh populasi dunia, ia ditanam secara komersial di merata dunia. Oleh itu, kajian filogeni tanaman ini adalah penting sebagai langkah untuk penambahbaikan program pembiakannya. Hubungan filogenetik antara 12 kultivar padi yang berasal dari dua sub-spesies; Indica dan Japonica telah dirumus dengan membandingkan set data jujukan protein yang diperolehi dari gen berbunga, iRFT1 dan dianalisis menggunakan kaedah maksimum parsimoni (MP). Struktur ramalan protein RFT1 telah dijana oleh perisian I-TASSER dan dianalisis menggunakan perisian YASARA. Hasilnya menunjukkan kultivar-kultivar ini dapat diklasifikasikan kepada dua kumpulan utama, di mana kumpulan pertama (Japonica) berevolusi terlebih dahulu diikuti kumpulan kedua (Indica). Hasil kajian mencadangkan sesetengah kultivar mempunyai hubungan yang rapat antara satu sama lain walaupun berasal dari varieti yang berbeza. Hubungan antara kultivar ini memberi informasi yang berguna bagi pemahaman yang lebih baik berhubung proses evolusi molekul dan mereka bentuk program pembiakan yang baik untuk penghasilan kultivar baru.

Kata kunci: Padi, RFT1, analisis filogeni, struktur ramalan

1.0 INTRODUCTION

Rice or its scientific name *Oryza sativa* L. is one of the most prominent members of the family Poaceae that is cultivated around the world and also function as model plant for genomic and breeding research [1]. It is believed to be domesticated from a common wild rice, *Oryza Rufipogon* Griff. in Asia thousands of years ago and consist of two major sub-species, which is Indica (*Xia*) and Japonica (*Keng*) [2-3]. Rice is the second major cereal crop being cultivated throughout the world and feeding more than half of the world population [4]. It has to be increased up to 50% in order to cater the world population by 2025 [5]. Nowadays, there are many new cultivars being produced in order to improve the yield and quality of the rice. For instance, in Malaysia, two new local varieties, MR220CL1 and MR220CL2 were developed by cross-breeding two local varieties namely MR200 and MR219 with American variety, IMI-TR No. 1770 [6]. Literally, one of the factors that determine the effectiveness of hybrid production is the well-defined phylogenetic relationship [7] and good information of genetic diversity that leads to a well-designed breeding program. However, the phylogenetic relationship of rice cultivars is poorly studied and remains unclear as there are large numbers of rice cultivars being cultivated all over the world. Nevertheless, with the advancement of molecular techniques, phylogenetic study based on amino acid sequence is possible and more reliable to compare with conventional techniques that are based on breeding behavior, cytological markers, ecological distribution and morphological markers in order to assess genetic diversity and relationships among cultivars [8].

In this study, phylogenetic relationship of 13 selected cultivars from two major varietal, Indica and Japonica including one partial sequence from our own study was inferred by using amino acid sequences of rice mobile florigen namely RICE FLOWERING LOCUS T 1 (RFT1). Rice Flowering Locus T 1 (RFT1) is one of the florigen in rice. The florigens; RFT1 and Hd3a (Heading date 3a) regulate flowering time via a dynamic mechanism. In general these florigens regulate flower development in rice through translocation from the leaf to the shoot apical meristem to activate downstream flowering genes. However, they both act under different photoperiod. Both are orthologs of *Arabidopsis* FT (FLOWERING LOCUS T) where Hd3a activates flowering under short day (SD) and RFT1 under long day (LD) [9-11].Earlier studies on *RFT1* showed correlation between earlier flowering time and higher *RFT1* mRNA level [12] leading to suggestion that at least among japonica cultivars, the *RFT1* is the main activator to induce flowering under LD. The finding indicates the importance of *RFT1* under LD as in the same condition, late flowering had been observed in *RFT1* RNAi lines [11]. Additional support based on the phylogenetic analyses of *FT*-like genes showed rapid divergence of *RFT1* from rice displaying high variation [13]. Hence, adaptation of rice to temperate climate regulated by *RFT1* was documented [14, 12].

Nevertheless, in the photoperiod pathway of rice flowering, RFT1 gene and protein could be activated both under SD and LD. Our focus on RFT1 arises due to the environment of the rice grown in Malaysia which could be categorized under LD. Furthermore, our focus of study was on upland rice which requires minimum amount of water for its growth and can sustain drought condition. The drought-tolerant variety is able to produce rice yield at approximately similar time as the wetland version and for some of the varieties, the upland version can produce earlier yield by almost a month. According to [15], flowering time is one of the important factors considered in determining regional and seasonal adaptation in rice and it is recognized as major trait to be selected in breeding program, which is the reason RFT1 was selected in the present study. Additionally, the predicted structure of RFT1 protein was also generated in this study. To date there is no study reported for RFT1 protein structure either by computational or experimental work. The predicted structure is important as it serves as the starting point to study the protein and provide fundamental understanding of the protein [16].

2.0 EXPERIMENTAL

2.1 Isolation of *RFT1* Gene from Malaysian Upland Rice Variety Hitam

RNA was extracted from the leaf of rice at 11 weeks old using EasyPure Plant RNA kit (Transgen Biotech) as per manufacturer recommendation. cDNA was later synthesized from the RNA using MMLV reverse transcriptase enzyme (Promega) according to manufacturer recommendation (manuscript submitted for publication). Amplification of RFT1 gene was carried out on the cDNA using primers EX1F (5'TGGCTAGCTTAACCTTCCTG3') and EX1R (5'GTCTACCATCACCTGTAGGT3') at 94°C for 5 min followed by 35 cycles of 94°C for 1 min, 56°C for 30 sec, 72°C for 2 min and final extension at 72°C for 5 min. The amplicon was run on 1% (w/v) agarose gel, band of interest excised, purified and sent for sequencing. Sequences were analyzed via BLAST with 99% identity to *RFT1* from *Oryza sativa* variety Chin Galay.

2.2 Phylogeny Analysis

CAP3 sequence assembly program was used to contig *RFT1* gene sequence from sequencing result. The sequence was translated into protein sequence with a total length of 120 residues. Meanwhile, the protein sequences of RFT1 from 12 rice cultivars, Basmati 370 (BAH30236), Pokkali (BAO03221), Bleiyo (BAJ53916), Muha (BAJ53912), Deng Pao Zhai (BAJ53911) and Kasalath (BAH30234) of Indica and Khao Nam Jen (BAJ53917), T65 (BAF92712), Dianyu 1

(BAO03202), Nipponbare (BAB78480), Khau Mac Kho (BAJ53915) and Modi (BAH30237) of Japonica and the outgroup, *Oryza glumipatula* (BAH56284) were obtained from GenBank (http://www.ncbi.nlm.nih.gov/). These 12 cultivars were chosen based on the quality and origin [17-22] while *O. glumipatula*, an American wild rice [23] was selected as an outgroup as it share the same genus, which is Oryza. The sequences were saved in FASTA format and multiple sequence alignment was performed by using Clustal X and the phylogeny reconstruction analysis based on Maximum Parsimony method with Tree-Bisection-Reconnection (TBR) algorithm was conducted using PAUP version 5.10 (MEGA 5.1 Beta 3) software.

2.3 Structure Prediction

The predicted structure of RFT1 protein on three selected cultivar was generated by I-TASSER server and analyzed using YASARA software.

3.0 RESULTS AND DISCUSSION

3.1 Molecular Phylogeny of RFT1 Protein from Rice

In evolutionary study, the construction of cladogram is of paramount important as it function in inferring and elucidating evolutionary relationship among species being studied. As for this study 14 amino acid sequences in the data sets were analyzed with all positions that contain gaps and missing data eliminated. However, the RFT1 sequence obtained from Malaysian upland rice was only a partial sequence amounting to 120 amino acid positions. The multiple sequence alignment (Figure 1) revealed that all 14 sequences are highly conserved from the region beginning at the start of the RFT1 protein until 69 amino acid position with 99% degree of similarity. The highly conserve region among the varieties indicates slow evolutionary rate [24] which implies low substitution and mutation rate. As we only obtained a partial sequence of our variety (Hitam), we expect the overall sequence similarity value to be lower. As shown in Figure 1, the 6 Japonica cultivars were highly conserved in their amino acid sequences except for Nipponbare which lacked the last amino acid Proline (P). It was noted that one of the Indica cultivar Basmati370 shared similar conserved amino acid as the other Japonica cultivars. Although our study produced partial sequence of the overall protein, it was of interest that our variety which is of Indica cultivar also showed high similarity to the Japonica cultivars. Nevertheless, among the other 5 of the Indica cultivars, there were 6 amino acid changes observed at different loci as indicated by regions highlighted either in white or yellow (Figure 1). This seemed to be lesser than the one reported by [25] that revealed 16 amino acid changes in *RFT1*. However, their work was conducted based on the nucleotide sequence translated into amino acid via genetic code. Additionally, they shared a finding that higher diversity was observed in Indica as compared to Japonica. As the Indica cultivars with the exception of Basmati 370 were not 100% conserved, substitution in its amino acid sequences was observed that could contribute to the diversity of this cultivar.

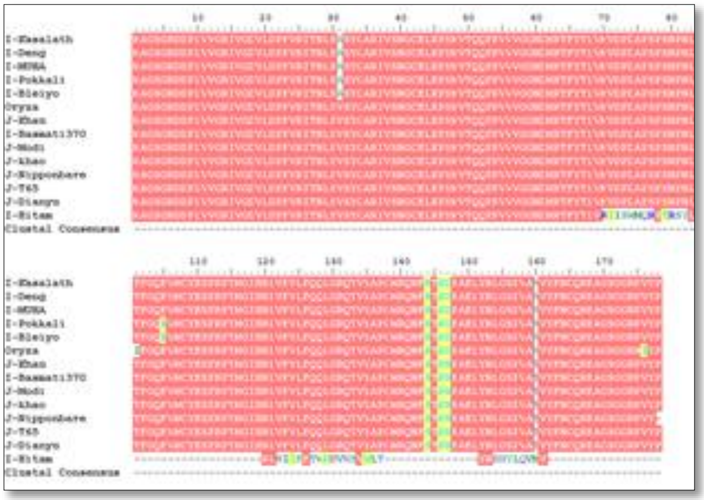

Figure 1 Multiple sequence alignment of 13 rice cultivars and the outgroup used is *Oryza glumipatula* (conserved region highlighted in pink)

The relationship of all rice cultivars for the analysis was presented in a cladogram. The cladogram (Figure 2) shows the consensus phylogenetic tree of 13 rice cultivars with a consistency index (CI) of 1 and retention index (RI) of 1 which indicates that there is no homoplasy and the character is totally steady with phylogeny [26].

As shown in the consensus tree (Figure 2), the 13 cultivars of rice were classified into two major groups.

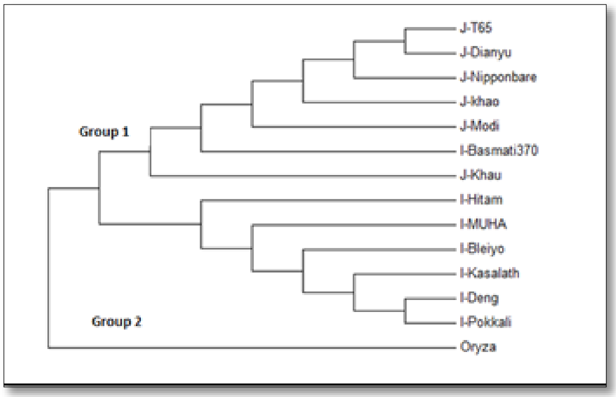

Figure 2 Strict consensus tree derived from the parsimony analysis of RFT1 protein

Group 1 is further divided into two sub clades and comprised of all Japonica cultivars along with Basmati 370 (Indica). This finding is in line with the previous study done by [27] that demonstrated a close evolutionary relationship between Japonica and

Basmati varieties that is based on the fragrance characteristic.

On the other hand, group 2 comprised of the other 5 Indica cultivars along with our variety (Hitam) containing further sub clades. At general glance of the sequence alignment (Figure 1), Hitam was observed to be highly conserved for the first 69 amino acids to the 6 Japonica cultivars and Basmati 370 (Indica) with 100% similarity. However, upon construction of cladogram, Hitam was shown to progress similarly to the other 5 Indica cultivars and forming a sub-clade of its own. Since we have yet to obtain full RFT1 sequence of Hitam, it would be interesting to know if it would later be grouped into the first clade of all Japonica cultivars and Basmati 370 or remain in the second clade and retaining a sub-clade of its own. Early implication of the result signifies that Hitam which is an upland rice variety may have its own evolutionary trait as drought-tolerant rice varieties. Additionally, the findings were supported by distinct phenotypic characteristic of both cultivars. For instance, in terms of size and appearance, Indica cultivars possess long and slender grains while Japonica cultivars possess short and stumpy grains [28, 29]. Furthermore, most of Japonica cultivars originated from oriental countries like Japan, Taiwan, China, Laos and Vietnam while Indica cultivars originated from many broad countries, like India, Indonesia, and Thailand [19]. Hence, there is a degree of similarity and diversity of RFT1 protein among these two cultivars where the suggestion that RFT1 have high variability in Asia was recommended based on *RFT1* gene phylogenetic analysis [13]. Moreover, from the phylogenetic tree obtained, the evolutionary pattern of the cultivars can be predicted. The cladogram suggested that the first group that is predominant by Japonica type of rice evolve earlier as compared to the second group that consisted of all Indica type of rice based on the length of the nodes that branch out from the outgroup.

3.2 RFT1 Protein Structure Prediction

In order to further determine the similarity and relatedness between these two varieties, predicted model of RFT1 protein (Figure 3) was constructed using I-TASSER server [30-32] for Nipponbare (Japonica), Kasalath and Basmati 370 (Indica). We did not proceed with the structure prediction of Hitam from our study due to its partial sequence. Nevertheless, since RFT1 protein structure was never published, we predicted the protein structure of this florigen from established sequence. The quality of the predicted structure was determined by the value of the C-score. Higher C-score value implies higher confidence of the predicted structure [30-32]. In this study, all three predicted structures have C-score value close to 1 which indicates high confidence level.

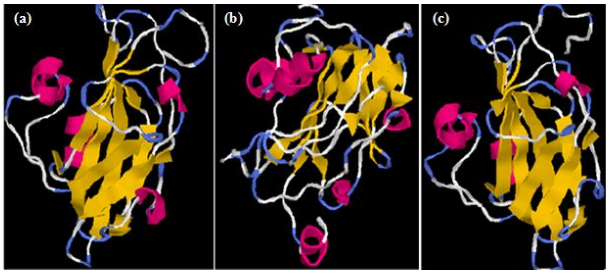

Figure 3 Predicted model of RFT1 protein using I-TASSER showing Nipponbare (left), Basmati 370 (middle) and Kasalath (right)

Nipponbare was chosen to represent Japonica cultivar while Kasalath for Indica as both cultivars were widely studied and reported. An additional Basmati 370 was selected as it was from Indica but in the cladogram (Figure 2) it was grouped among the Japonica based on its sequence. The predicted structure was further analyzed by using YASARA software based on few parameters namely secondary structure, B-factor and the mass of the object (Table 1).

Table 1 Comparison of predicted structure of RFT1 protein

Parameter	Nipponbare	Basmati 370	Kasalath
2˚ structure (%)	11.9 helix, 26.6 sheet, 7.9 turn, 51.4 coil, 2.3 3-10 helix	15.2 helix, 23 sheet, 6.7 turn, 52.2 coil, 2.8 3-10 helix	11.8 helix, 28.7 sheet, 11.8 turn, 47.8 coil, 0.0 3-10 helix
B-factor	1.0	2.1	1.0
Mass (g/mol)	18247.605	18329.644	18285.631

As shown in Table 1, there is not much of difference in the mass of all three cultivars. Similar result was observed on the secondary structure where small difference in terms of degree of α-helix, β-sheet, turn, coil (or loop) and 3_{10}-helices (parahelices) which are the main angles that define the secondary structure of a protein. The main structural motif in protein secondary structures are α-helices and β-sheets although in α-helix rich protein α to β transition are sometimes observed as universal deformation mechanism [33]. The only obvious difference among the three cultivars in their secondary protein structure parameter was the absence of 3_{10}-helices value (0%) in Kasalath. The 3_{10}-helix is almost similar to α-helix however in identical amino acid restudies, the former turns and coils tightly to provide longer and thinner structure [34] which has been said to be less stable than the latter. Most 3_{10}-helices are very short contributing only up to 4% of total amino acid residues. Occasionally there are longer 3_{10}-helices which could be found in several proteins with possible role in protein folding [35]. We did not expect any stark difference in the function of RFT1 of the three cultivars even with Kasalath having none of the 3_{10}-helix. The only clear difference among the analysis of predicted structure of the three cultivars is the B-factor. Basmati 370 had the highest number (2.1) as compared to Nipponbare and Kasalath (1.0). B-factor or atomic displacement parameter is one of the crucial parameter in examining protein structure as it gives indication of the degree of mobility for each of the atoms in the protein structure [36] thus give dynamics information of the protein. High B-factor implies high mobility while low B-factor indicates low mobility (structural rigidity) [37]. It is suggested that RFT1 protein of Basmati 370 is highly dynamic and more flexible. This could be seen from the cladogram in Figure 2 as it evolved earlier compared to the other 5 Indica cultivars. This could be due to the fact that protein size and mobility is decreasing during evolution [37]. As RFT1 is a specific long day (LD) florigen in rice that regulates flowering, the predicted structure of this protein in selected cultivars along with its phylogenetic analysis could assist in understanding the molecular function of flowering.

4.0 CONCLUSION

In conclusion, findings from this study showed that the phylogenetic analysis inferred classification of Japonica and Indica into two distinct groups, where Japonica evolved earlier than Indica. It thus implies that RFT1 could function as potential biomarker in distinguishing rice cultivars. The findings also suggested that some cultivars had a close relationship with each other although it originates from different areas which give better insight with regards to the relationship among these cultivars. Useful information could be gained for better understanding of molecular evolution process and developing good breeding program in order to generate new cultivar with excellent quality. Data mining study of RFT1 is important as long day flowering in rice in temperate countries requires adaptation and depended on the activation of florigen.

Acknowledgement

We thank the Universiti Teknologi Malaysia (UTM) and Ministry of Higher Education (MOHE) for funding this project (Flagship Research University Grant vote: 00G65), Prof. Dr. Roji Sarmidi (project leader) and Zamalah for sponsoring the graduate student scholarship.

References

[1] Upadhyay, P., Neeraja, C. N., Kole, C., Singh, V. K. 2012. Population Structure and Genetic Diversity in Popular Rice Varieties of India as Evidenced from SSR Analysis. *Biochemical Genetics*. 50: 770-783.

[2] Oka, H. I. 1988. *Origin of Cultivated Rice*. Elsevier, Tokyo.

[3] Zhimin, A. 1999. Origin of Chinese Rice Cultivation and its Spread East *Beijing Cultural Relics*. 2: 63-70.

[4] Sohrabi, M., Rafii, M. Y., Hanafi, M. M., Siti Nor Akmar, A., Latif, M. A. 2012. Genetic Diversity of Upland Rice Germplasm in Malaysia Based on Quantitative Traits. *The Scientific World Journal*. 1-9.

[5] Datta, S. K. 2004. Rice Biotechnology: A Need for Developing Countries. *AgBioForum*. 7(1&2): 31-35.

[6] Hershey, C. H. 2010. *Malaysia Launches Two New Varieties of Rice. An Electronic Newsletter of Applied Plant Breeding*. Cornell University.

[7] Nishiyama, K., Choi, Y. A., Honsho, C., Eiadthong, W., Yonemori, K. 2006. Application of Genomic *in situ* Hybridization for Phylogenetic Study between *Mangifera indica* L. and Eight Wild Species of *Mangifera*. *Scientia Horticulturae*. 110: 114-117.

[8] Oka H. I. 1958. Intervarietal Variation and Classification of Cultivated Rice. *Indian Journal of Genetics and Plant Breeding*. 18: 79-89.

[9] Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S., Simamoto, K. 2007. Hd3a Protein is a Mobile Flowering Signal in Rice. *Science*. 316: 1033-1036.

[10] Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S., Shimamoto, K. 2008. *Hd3a* and *RFT1* are Essential for Flowering in Rice. *Development*.135: 767-774.

[11] Komiya, R., Yokoi, S., Shimamoto, K. 2009. A Gene Network for Long-day Flowering Activates *RFT1* Encoding a Mobile Flowering Signal in Rice. *Development*. 136: 3443-3450.

[12] Naranjo, L., Talon, M., Domingo, C. 2014. Diversity of Floral Regulatory Genes of Japonica Rice Cultivated at Northern Latitudes. *Genomics*. 15: 101.

[13] Hagiwara, W.E., Uwatoko, N., Sasaki, A., Matsubara, K., Nagano, H., Onishi, K., Sano, Y. 2009. Diversification in Flowering Time Due to Tandem *FT-Like* Gene Duplication, Generating Novel Mendelian Factors in Wild and Cultivated Rice. *Molecular Ecology*. 18: 1537-1549

[14] Itoh, H. and Izawa, T. 2013. The Coincidence of Critical Day Length Recognition for Florigen Gene Expression and Floral transition under Long-Day Conditions in Rice. *Molecular Plant*. 6(3): 635-649.

[15] Yano, M., Ebana, K., Matsubara, K., Hori, K., Yamanouchi, U., Mizubayashi, T., Kono, I., Shomura, A., Ito, S., Ando, T. 2011. Uncovering of Major Genetic Factors Generating Naturally Occurring Variation in Heading date among Asian Rice Cultivars. *Theoretical and Applied Genetics*. 122(6): 1199-1210.

[16] Khor, B. Y., Tye, G. J., Lim, T. S., Choong, Y. S. 2015. General Overview on Structure Prediction of Twilight-zone Proteins. *Theoretical Biology and Medical Modelling*. 12(15): 1-11.

[17] Tuteja, N., Gill, S. S. 2013. *Plant Acclimation to Environmental Stress*. New York, USA: Springer.

[18] Takayuki, O., Mihoko, Y., Hiromoto, Y., Tetsu K. 2011. The Biotron Breeding System: A Rapid and Reliable Procedure for Genetic Studies and Breeding in Rice. *Plant Cell Physiology*. 52(7): 1249-1257.

[19] Kazuhiro, S., Seishi, I., Shima, E., Hisayuki, M., Eiko, H., Chiharu, K., Yuri, K., Atsuhiko, K., Takuro, S., Kiwamu, M., Tadashi, S. 2010. Impact of Plant Genotype and Nitrogen Level on Rice Growth Response to Inoculation with *Azospirillum sp.*Strain B510 under Paddy Field Conditions. *Soil Science and Plant Nutrition*. 56: 636-644.

[20] Takehisa, H., Yasuda, M., Fukuta, Y., Kobayashi, N., Hayashi,N., Nakashita,H., Abe, T., Sato,T. 2009. Genetic Analysis of Resistance Gene in an Indica-type Rice (*Oryza Sativa* L.), Kasalath, Using DNA Markers. *Breeding Science*. 59: 253-260.

[21] Chandi, G. K., and Sogi. D. S. 2008. Characterization of Traditional (Basmati 370) and Developed (Pusa Basmati 1) Basmati Rice. *International Journal of Food Properties*. 11: 910-918.

[22] Jain, S., Jain, R. K., McCouch, S. R. 2004. Genetic Analysis of Indian Aromatic and Quality Rice (*Oryza sativa* L.) Germplasm Using Panels of Fluorescently-labeled Microsatellite Markers. *Theoretical and Applied Genetics* 109: 965-977.

[23] Veasey, E. A., Bressan, E. D. A., Zucchi, M. I., Vencovsky, R., Cardim, D. C., da Silva, R. M. 2011. Genetic Diversity of American Wild Rice Species. *Scientia Agricola*. 68(4).

[24] Clegg, M. T. 1993. Review: Chloroplast Gene Sequences and the Study of Plant Evolution. *Proceedings of the National Academy of Sciences (PNAS) USA*. 90: 363-367.

[25] Ogiso-Tanaka, E., Matsubara, K., Yamamoto, S., Nonoue, Y., Wu, J., Fujisawa, H., Ishikubo, H., Tanaka, T., Ando, T., Matsumoto, T., Yano, M. 2013. Natural variation of *RICE FLOWERING LOCUS T 1* Contributes to Flowering Time Divergence in Rice. *PLOS ONE*. 8(10): e75959.

[26] Nei, M., and Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. New York: Oxford University Press, Oxford.

[27] Kovach, M. J., Calingacion, M. N., Fitzgerald, M. A., McCouch, S. R. 2009. The Origin and Evolution of Fragrance in Rice (*Oryza sativa* L.). *Proceedings of the National Academy of Sciences (PNAS)*. 106(34): 14444–14449.

[28] Khush, G. S. 2000. Taxonomy and Origin of Rice. In Singh, R. K., Singh, U. S., Khush, G. S. (ed.). Aromatic Rices. Oxford & IBH, New Delhi. 5-13.

[29] IRRI, International Rice Research Institute. 1996. *International Rice Research Notes* (Vol. 2): International Rice Research Institute (IRRI).

[30] Yang, Z. 2008. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinformatics*. 9(40).

[31] Roy, A., Kucukural, A., Yang, Z. 2010. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nature Protocols*. 5: 725-738

[32] Roy, A., Jianyi, Y., Yang, Z. 2012. COFACTOR: An Accurate Comparative Algorithm for Structure-based Protein Function Annotation. *Nucleic Acids Research*. 40: W471-W477.

[33] Qin, Z., Buehler, M.J. 2010. Molecular Dynamics Simulation of the Alpha–Helix to Beta–Sheet Transition in Coiled Protein Filaments: Evidence for a Critical Filament Length Scale. *Physical Review Letters*. 104: 198304.

[34] Vieira-Pires, R.S., Morais-Cabral, J.H. 2010. 3_{10} Helices in Channels and Other Membrane Proteins. *Journal of General Physiology*. 136(6): 585-592.

[35] Armen, R., Alonso, D.O., Daggert, V. 2003. The Role of Alpha-, 3_{10}- and pi-helix in Helix → Coil Transitions. *Protein Science*. 12: 1145-1157

[36] Smith, D. K., Radivoiac, P., Obradovic, Z., Dunker, A. K., Zhu, G. 2003. Improved Amino Acid Flexibility Parameters. *Protein Science*. 12(5): 1060-1072.

[37] Navare, C., and Banerji, A. 2012. Residue Mobility has decreased during Protein Evolution. *An Electronic Newsletter* Cornell University. Cornell University Library.