TROPICAL GENETICS

Volume 3, No 2, November 2023
https://ojs.genetikawan-muda.com/index.php/tg

Original Research

Analysis of genetic variations in bHLH protein (Rc) gene sequences in rice (Oryza) NCBI PopSet 2496581476 using in-silico RFLP

Nurul Hasanah, Violita, Afifatul Achyar

Department of Biology, Faculty of Science and Mathematics, Universitas Negeri Padang, 25131, Kota Padang

Corresponding author(s): afifatul.achyar@fmipa.unp.ac.id

Article Info

Article history:
Received 26 Juni 2023
Accepted 6 Agustus 2023
Available online 30 November 2023

Keywords:
Oryza, Rc-bHLH, RFLP, BstuI enzymes, in-silico

How to cite:
Hasanah, et.al. 2023. Analysis of genetic variations in bHLH protein (Rc) gene sequences in rice (Oryza) NCBI PopSet 2496581476 using in-silico RFLP. *Tropical Genetics,* 3(2): 60-65

Abstract

Rice is a plant from the Graminae Family that is spread almost throughout the region. In rice, there is an Rc gene that codes for the formation of the basic loop-helix-loop (bHLH) protein, the Rc gene (Rc-bHLH) functions as a transcription of colour pigments in brown rice. This research was conducted computationally from several related sites. The study used restriction enzyme BstUI with a side recognition of 5'-CG'CG-3'. The purpose of this study was to analyze the genetic variation of the Rc gene sequence encoding the bHLH protein in NCBI PopSet rice 2496581476 using RFLP in silico. The results showed that there were genetic variations of the Rc-bHLH gene, 2 allele variations were present in 21 rice sequences using the restriction enzyme BstUI.

Introduction

Population genetics is a field of biology that studies the genetic composition of biological populations and changes in genetic composition. Genetic diversity and variation can occur due to changes in the nucleotides that make up DNA that may affect the phenotype of an organism. Mutation, recombination, migration, and natural selection determine the degree of genetic diversity and variation (Ellegren 2009). Populations that have a high degree of genetic variation indicate that they respond adaptively to environmental changes (Arimurti 2017).

Rice (genus Oryza) is a most important food grain because its consumed by more than half of the world's population for their sustainable livelihood (World Rice Production 2019). Rice is a plant from the Graminae or Poaceae Family, a diploid plant with a relatively small genome size when compared to other cereal plants, which is ~400 Mb (Chen et al. 2013). In this study, genetic variation analysis of the Rc gene coding for the basic loop-helix-loop (bHLH) protein in rice was carried out. The Rc gene encoding the basic helix-loop-helix (bHLH) protein found on chromosome 7 in the rice

genome is marked with the marker RC12 at the exon position 7 (Utami et al. 2010). Total of 167 Rc-bHLH genes have been identified in the rice genome (Li et al. 2006).

The Rc gene is a transcription factor for the color pigment protein prothocyanidin in red rice seeds. To date, only two loci for the red rice trait have been reported; Rc, which encodes a bHLH transcription factor, and Rd, which encodes dihydroflavonol-4-reductase (DFR) (Sweeney et al. 2006; Furukawa et al. 2007). Mbanjo et al. (2019) have cinducted screening for the 14-bp deletion within the Rice Rc locus known to result in loss of red pigmentation in Rice grains. The results are in accordance with Sweeney (2006) research, this deletion is known to result in the loss of red pericarp (Sweeney et al. 2006). This screen revealed that the functional allele of the Rc gene (red pigmented pericarp) was preponderant among the accessions screened (Mbanjo et al. 2019). The Rc gene encodes the formation of the bHLH protein which serves as a transcription factor that binds to specific DNA target sites, having a central role in metabolic, physiological, and plant development processes. Simultaneously bHLH protein activates the genes associated to anthocyanin biosynthesis in pericarp and developed black grain rice (Roy & Shil 2020). In addition, bHLH protein also plays a role in plant responses to abiotic stresses such as drought, salinity, low temperatures, and metals (Qian et al. 2021).

One molecular marker that is often used for genetic variation analysis is RFLP. Restriction Fragment Length Polymorphism (RFLP) is a DNA marker that uses DNA-DNA hybridization (Southern blot) in the detection process. The DNA detected is a DNA fragment resulting from the cutting of the DNA genome by restriction enzymes (Dwiyani 2016). Before performing RFLP in the laboratory, RFLP can be done by in silico method.

In silico tests are tests carried out by computational methods or virtually. The use of in silico tests is to predict, hypothesize, give new discoveries or new advances in medicine and therapy (Hardjono 2013; A. Achyar et al. 2021). In silico can be used as a method to approach real conditions into computer-based simulations using specific application or software programs.

Based on the explanation above, this study aims to analyze the genetic variation of the Rc gene encoding the basic loop-helix-loop (bHLH) protein in rice using molecular markers Restriction Fragment Length Polymorphism (RFLP) in silico.

Materials and Methods

1. Materials

The gene sequence in rice used in this study is the Rc gene that produces the basic helix-loop-helix (bHLH) protein. The sequence is downloaded in fast form from the NCBI website (https://www.ncbi.nih.gov/popset) with the PopSet identity number 2496581476. The sequence was submitted by Mudhale, A., Sar, P., Kumar, J., Bhowmick, P.K., Basak, N., Patra, B.C., Bisht, D.S., Iquebal, M.A., Vinod, K.K., Goapala Krishnan, S., Banerjee, A., Mandal, N.P. and Roy, S. in their study of Assessment of Genetic Diversity and Polymorphisms Associated with Key Domestication Genes in Rice Germplasm Collected from Riverine Ecology of Majuli and Adjoining Areas in Assam, Northeastern India. The PopSet has 21 gene sequences with Accession number (Table 1.) (https://www.ncbi.nlm.nih.gov/popset/?term=2496581476).

Table 1. Sample of bHLH protein (Rc) gene sequence of rice in NCBI

No.	Acc number of rice isolate	No.	Acc number of rice isolate	No.	Acc number of rice isolate
1.	OP542207.1	8.	OP542214.1	15.	OP542221.1
2.	OP542208.1	9.	OP542215.1	16	OP542222.1
3.	OP542209.1	10.	OP542216.1	17.	OP542223.1
4.	OP542210.1	11.	OP542217.1	18.	OP542224.1
5.	OP542211.1	12.	OP542218.1	19.	OP542225.1
6.	OP542212.1	13.	OP542219.1	20.	OP542226.1
7.	OP542213.1	14.	OP542220.1	21.	OP542227.1

Methods

a. Restriction Enzyme Candidate Screening Method
Screening of restriction enzyme candidates was carried out on the http://insilico.ehu.es/restriction site by selecting the tool “compare restriction patterns of many sequences”. This tool will compare the restriction patterns of many DNA sequences tested as well as the restriction enzymes that cut them (Achyar et al. 2021). Upload the fasta file from the downloaded sequence then click "go to next step" and the sequence alignment results appear. The same sequence will be removed for easy analysis. The option "only restriction enzymes with known base (No, N, R, Y, ...)" is checked because it only uses obvious restriction enzymes (A,G,C, and T) so that the recognition side of restriction enzymes is certain. Then click "get the list restriction enzymes" to get the restriction enzyme candidate to be used in the next stage.

b. RFLP in-Silico
In silico RFLP or computational restriction is performed using https://www.benchling.com/sign site. First, we need to register using email, this site is free so that after registration can sign up and be used indefinitely. Next, import the fasta sequences that have been downloaded in the project folder by clicking (+) symbol and selecting the option "import DNA/RNA sequences". Click the scissors symbol in the right corner and type the name of the restriction enzyme in the "find enzymes" column then select "run digest" to perform the restriction. The restriction electrophoregram can be seen on the "virtual digest" menu.

Results and Discussion

1. Screening of restriction enzyme candidates
Based on the results of restriction enzyme screening, there are 28 sides of restriction recognition for the bHLH protein (Rc) gene sequence in rice (Oryza). One side of such recognition is the 5'-CG'CG-3' recognized by the enzyme BstUI. This restriction enzyme was chosen because it has a variation in the bHLH protein (Rc) gene sequence in the PopSet 2496581476.
The restriction enzyme BstUI was isolated from the *E.coli* strain that carries the cloned BstUI gene from *Bacillus stearothermophilus* U458 (Z. Chen). BstUI is a type IIP restriction enzyme

that has isoschimers AccII, Bsh1236I, BspFNI, BstFNI, and MvnI, all of which have blunt ends. This enzyme is associated with restriction endonucleases B, time-saver qualified restriction enzymes products and can be used in Fast Cloning applications: accelerate your cloning workflows with reagents from NEB, restriction enzyme digestion (New England Biolabs 2023).

2. RFLP in silico

   RFLP method as one method to determine polymorphism in studying the evolutionary history of human populations (lineage/genealogy) and to determine the presence of mutations (Siti et al. 2013). With the development of bioinformatics software technology, restriction and visualization of restricted fragments can be done in silico which aims to predict genotyping results before conducting RFLP in real time in the laboratory. In addition to prediction, this technology can also be used to further explore DNA sequences available in the GenBank NCBI database (Achyar et al. 2021).

Figure 1. Electrophorogram of rice Rc gene sequence restriction results using BstUI restriction enzyme in silico. Ket: Kiri (Ladder Life 1 kb Plus); (1) OP542211.1; (2) OP542215.1; (3) OP542216.1; (4) OP542217.1; (5) OP542219.1; (6) OP542220.1; (7) OP542222.1; (8) OP542225.1; (9) OP542226.1; (10) OP542227.1; (11) OP542207.1; (12) OP542208.1; (13) OP542209.1; (14) OP542210.1; (15) OP542212.1; (16) OP542213.1; (17) OP542214.1; (18) OP542218.1; (19) OP542221.1; (20) OP542223.1; (21) OP542224.1

Restriction with the BstUI enzyme on 21 DNA sequences of the Rc-bHLH gene resulted in two allele variations. The first allele produces one DNA band with a length of 400 bp for the *Oryza sativa* sequence and 370 bp for the *Oryza rufipogon* sequence while the second allele produces two DNA bands with a length of 320 bp and 110 bp (Figure 1).

Table 2. Frequency of bHLH Protein (Rc) Gene Allele of NCBI Rice Based on In Silico RFLP Results

Restriction Enzyme	Recognition Site	Fragment Size (bp)	Allele	Number of Fragment Presence (N = 21)	Fragment Presence Percentage (%)	Allele Frequency
BstUI	CG'CG	400 (*O. sativa*) 370 (*O. rufipogon*)	1	10	47,62	0,48
		320 110	2	11	52,38	0,52

In Table 2 it can be seen that the percentage of presence of the DNA band size of the first allele is 47.62% and the percentage of the presence of the second allele is 52.38%. The first allele appears with a frequency of 0.48 while the second allele appears with a frequency of 0.52. The presence of genetic variations in the Rc gene coding for the bHLH protein is caused by gene mutations, where some alleles are not passed on to later plants. Two loci have been identified using classical genetic analysis, Rc (brown pericarp and seed coat) and Rd (red pericarp and seed coat). Rc in the absence of Rd produces brown seeds, whereas Rd alone has no phenotype (white rice). Although Rc is referred to as a mutant allele because its phenotype differs from that of common rice cultivars, the action of Rc is dominant over white pericarp (rc) (Sweeney et al. 2006).

Conclusion

Based on the results of RFLP in silico, it can be concluded that there are genetic variations, there are 2 allele variations in 21 sequences of the Rc gene coding for bHLH protein found in NCBI rice with PopSet 2496581476.

References

Achyar, Afifatul et al., 2021. Analysis of Genetic Variations in Poly Gene Sequences in Dengue Virus 2 Using In-Silico RFLP. *Bioscience*, 5(1), pp.80–86.

Achyar, A., Atifah, Y. & Putri, D.H., 2021. In Silico Study of Developing a Method for Detecting Pathogenic Bacteria in Refillable Drinking Water Samples. *Journal of Physics: Conference Series*, 1940(1).

Arimurti, A.R.R., 2017. Keanekaragaman Genetik Nyamuk Vektor Filariasis Culex quinquefasciatus Say, 1823 (Diptera: Culicidae) di Kota dan Kabupaten Pekalongan Dengan Metode PCR-RAPD. *Lembaga Penelitian dan Pengabdian Masyarakat Universitas Muhammadiyah Surabaya*, 87(1,2), pp.149–200.

Chen, H. et al., 2013. Development of Genomics-based Genotyping Platforms and Their Applications in Rice Breeding. *Current Opinion in Plant Biology*, 16(2), pp.247–254. Available at: http://dx.doi.org/10.1016/j.pbi.2013.04.002.

Dwiyani, R., 2016. Penggunaan Penanda Molekuler untuk Pemuliaan Tanaman. In *Bahan Ajar Mata Kuliah Ilmu Benih dan Pemuliaan Tanaman*. Universitas Udayana Press.

Ellegren, H., 2009. Is Genetic Diversity Really Higher in Large Populations? *Journal of Biology*, 8(4).

Furukawa, T. et al., 2007. The Rc and Rd Genes are Involved in Proanthocyanidin Synthesis in Rice Pericarp. *Plant Journal*, 49(1), pp.91–102.

Hardjono, 2013. *Kimia Organik Dasar*, Yogyakarta: Universitas Gadjah Mada Press.

Li, X. et al., 2006. Genome-wide Analysis of basic/Helix-Loop-Helix Transcription Factor Family in Rice and Arabidopsis. *Plant Physiology*, 141(4), pp.1167–1184.

Mbanjo, E.G.N. et al., 2019. Exploring the Genetic Diversity within Traditional Philippine Pigmented Rice. *Rice*, 12(1).

NEB - New England Biolabs, 2023. *BstUI*. Available at: https://international.neb.com/products/r0518-bstui#Product Information [Accessed June 23, 2023].

Production, W.R., 2019. World Rice Production 2019/2020.

Qian, Y. et al., 2021. Regulatory Mechanisms of bHLH Transcription Factors in Plant Adaptive Responses to Various Abiotic Stresses. *Frontiers in Plant Science*, 12(June).

Roy, S.C. & Shil, P., 2020. Black Rice Developed Through Interspecific Hybridization (O. sativa x O. rufipogon): Origin of Black Rice Gene from Indian Wild Rice. *bioRxiv*, pp.1–35.

Siti, Heli H.M., G. Gun G., Desy N., A.S.N., 2013. Variasi Urutan Nukleotida Daerah D-Loop DNA Mitokondria Manusia pada Dua Populasi Asli Indonesia Tenggara. , pp.440–446.

Sweeney, M.T. et al., 2006. Caught Red-handed: Rc Encodes a basic Helix-Loop-Helix Protein Conditioning Red Pericarp in Rice. *Plant Cell*, 18(2), pp.283–294.

Utami D., W., Ilhami A., Hanarida, I., 2010. DNA Fingerprinting of Local Rice Germplasm Using The Specific Markers for Red Rice. *Berita Biologi, Jurnal Ilmu-Ilmu Hayati*, 10(2), pp.143–150.