Advances in Tropical Biodiversity and Environmental Sciences
8(1): 20-25, February, 2024
e-ISSN:2622-0628
DOI: 10.24843/ATBES.2024.v08.i01.p04
Available online at: https://ejournal1.unud.ac.id/index.php/atbes/article/view/261

Primer Design and in Silico PCR for Detection Microsatellite Locus on Cassava (*Manihot esculenta*) as an Early Study of Genetic Diversity of Gluten Free Food Crops

Ni Putu Senshi Septiasari[1]* and Ni Putu Yuni Astriani Dewi[2]

[1] Diploma IV Department of Medical Laboratory Technology, Faculty of Health Sciences, Bali International University, Jl. Seroja Gg. Jeruk, Tonja, Kecamatan Denpasar Timur, Denpasar, Bali, Indonesia
[2] Department of Biology Education, Faculty of Education, University of Timor, Jl. El Tari Km 09, Kelurahan Sasi Kota Kefamenanu, NTT, Indonesia
*Corresponding author: senshiseptia@iikmpbali.ac.id

Abstract. Food allergy is a hypersensitive reaction of the body to certain substances that should not be harmful. One of the basic ingredients of foods that are reported to contain allergens is wheat. The type of allergen in wheat is found in gluten, which produces the protein gliadin. The protein gliadin is thought to be the cause of allergic reactions, especially in children and people who have celiac disease. Local food ingredients such as cassava can be used as alternative raw materials for gluten-free food products. The genetic diversity test of cassava is needed to determine the genetic variation of the cassava population and the ability to adapt to natural surroundings. Study of genetic diversity of a plant population can be analyzed by DNA fingerprinting technique using microsatellite molecular markers in the PCR method. Primer selection is a crucial stage because the position of the primer attachment will determine the success of the amplification process. Primer design and in silico PCR assay were carried out as a preliminary study to design the right primer to attach to a specific template. The method used digitally with the help of the NCBI Pick Primer page and FastPCR software for PCR stimulation, knowing the position of the primer attachment, precisely determining the primer and the length of the resulting amplicon. The results obtained in the form of 5 pairs of primer sequences attached to the microsatellite polymorphic locus in nucleus. The length of the amplicon produced in each primer was 167 bp, 195 bp, 155 bp, 161 bp and 112 bp. The basis of primer selection consists of the length of the primary base, the TM value (melting temperature), the percentage of GC amount and the possibility of the formation of dimers between primers.

Keywords: cassava; in silico; microsatellite locus; primary design

I. INTRODUCTION

Food allergy is the body's hypersensitive response to certain substances in food. Food allergies can occur at any age, starting in the first 2 years of life. The percentage of food allergy incidence in children ranges from 6-8%, much higher than the proportion of allergies in adults, which is 2.5% [1],[2]. Symptoms caused by allergies range from mild symptoms (swelling, itching, nasal congestion, and eye disorders) to severe symptoms (the occurrence of anaphylaxis which leads to death) [3].

Some food sources that cause allergies in the body are food ingredients in the form of nuts such as peanuts, nuts from trees, soybeans, animal milk (especially from cows), eggs, food sources from the sea such as fish and shellfish, and wheat [3],[4]. Wheat allergy is one of the most common allergies in children and people with celiac disease. The causative agent of wheat allergy is the protein gliadin, which is produced by wheat gluten. Wheat allergy mediated by Ig-E antibodies has been reported to affect 0.4-1% of American and British children. Wheat allergy in childhood and about 80% resolved by age 5 years. However, if a child has a combination of atopic dermatitis, the tolerance for healing can be lower, which is around 29% for 4 year olds, 56% for 8 year olds and 70% for 14 year old children [5].

Gluten is found in wheat flour which is made from wheat. Gluten-free foods have been widely circulated in the market, one of the local food ingredients that can be used as a substitute for wheat flour is cassava-based mocaf flour. Mocaf flour is fermented flour which is rich in beta

carotene. Cassava is a food source that contains lots of carbohydrates and can be used as a basic ingredient in agriculture [6]. The increasing demand for cassava in Indonesia is in line with the development of the cassava-based industry in the future. According to that around 2025 it is estimated that the demand for cassava will increase by around 30 million tons. With the increased demand for cassava, production will need to increase by up to 27% [7]. To increase cassava production, efforts are needed to improve varieties through conventional and modern biotechnology [8].

Genetic diversity can determine various biological phenomena in plants and their correlation with phenotypes which can be used as a guide to determine the ability to adapt to natural surroundings [9]. Genetic diversity can be analyzed by DNA fingerprinting techniques using molecular markers. Molecular markers can identify plants at the species, variety, clone, individual and population levels. The PCR method is a molecular analysis carried out by enzymatically duplicating certain DNA fragments [10]. The PCR method uses DNA markers that have high polymorphism, one of which is the Simple Sequence Repeat (SSR) marker. SSR is a polymorphic short tandem copying of nucleotides present at microsatellite loci [8].

The main components of PCR are primers, DNA templates, nucleotides (dNTP or deoxynucleotide triphosphate), and DNA polymerase enzymes. Primers are single-stranded sections of DNA that are segmented with the DNA template. Selection of primers is a crucial step in PCR because a specific primer design determines the success of the amplification process. Primary design in silico can be done with a computer (software). The in silico test serves to save processing time allocation prior to the in vitro PCR stage and ensures that the primer used in PCR sticks to the right location [11]. One of the common software used in designing primers is Pick Primer on the NCBI (National Center for Biotechnology Information) page. Meanwhile, PCR stimulation was carried out with the help of FastPCR software.

Cassava plants are known to have microsatellite markers spread across each chromosome. The characterization of cassava plants was determined by genotype using agro-morphology and SNP detection by sequencing analysis [6]. This method is difficult to carry out due to the limited availability of tools. Research is needed that can provide a reference for the use of conventional PCR tools to support the study of genetic diversity, especially in Indonesia.

This study aims as a preliminary study to obtain specific primer sequences at the cassava microsatellite locus and to stimulate these primers at the in silico PCR stage. The designed primer is expected to be used as an initial study of the genetic diversity of cassava (*Manihot esculenta*), a gluten-free food plant.

II. METHOD

Cassava gene sequences (*Manihot esculenta*) that was used to design the primer was the complete sequence of Cassava (*Manihot esculenta*) with accession number NC_035162.2. The gene was downloaded from Genbank NCBI (http://www.ncbi.nlm.nih.gov). The primary design was also done on the same page as the Primary Pick mode. The primer designed on the page was selected to be the primary candidate according to the specific target region. The primary criteria that must be met are primers that have a nucleotide length of between 18-30 bp, base composition of G and C which has a percentage of between 40-60%, the difference in temperature Tm between primers does not exceed 5^0C, and the absence of primary secondary structures such as hairpins or self dimers.

The specificity of the primers that have been designed is carried out by PCR stimulation in FastPCR software to determine the position of the primer attachment, the accuracy of the determination of the primers and the length of the resulting amplicon.

III. RESULTS AND DISCUSSION

In silico primer design on the NCBI page with the cassava (*Manihot esculenta*) genome reference sequence with accession number NC_035162.2. The determination of accession numbers can be seen in Figure 1. The primary design begins with tracing the microsatellite locus using FastPCR software. The cassava (*Manihot esculenta*) genome obtained from the NCBI website was traced for its microsatellite loci on each chromosome. Chromosomes have an intron region that does not encode for amino acids and has a high mutation rate. It is in this intron region that there is usually a locus with repeated base sequences called a microsatellite locus. Microsatellite loci or commonly called SSR (Simple Sequence Repeat) markers are easy to detect so they can be applied in research. SSR markers have been identified in cassava and can be used to identify genetic diversity in cassava populations [8].

Microsatellite DNA detection from FastPCR software search results, found 5 kinds of microsatellite loci from 5 different chromosomes. The five microsatellite loci sequences obtained were then designed as primers using the NCBI Pick primer page. Pick primers produce 10 pairs of primers on each chromosome, then the best pair of primers were choosen. The selected primer pairs are presented in Table 1.

The repeat type at the microsatellite locus was obtained from the FastPCR software, while the primer sequence was obtained from the primer design via the NCBI Pickprimer page. The following are the criteria for determining the primary selection which are used as a reference for selecting candidates for the primary pair (Table 2).

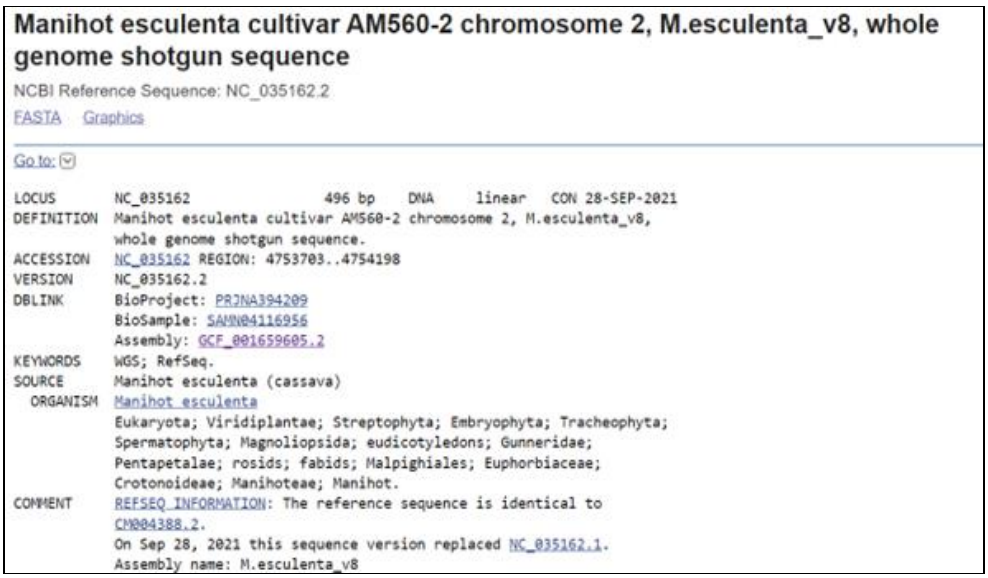

Fig.1. Genome accession number determination of *Manihot esculent*

TABLE 1.
CASSAVA PRIMARY DESIGN RESULTS (*Manihot esculenta*)

No.	Name of Primer	Repetition type (microsatellite loci)	Sekuen (5`-3`)
1.	M.esc2	(GT)8, (GA)10,GC, (GA)12	F : ACCATGTAAATGAAGACTGGAAAG R : AGGTTGGATGCTTGAAGGAATA
2.	M.esc15	(GA)16	F : TGGCAGCCAAACATGAATTA R : AGAGGAGAAGCACCAAGCA
3.	M.esc12	(CT)19	F : GGAAGCTGTCCCCTATTCACA R : ACAGGCCTTGATCATTTTTAGCC
4.	M.esc3	(CA)16	F : AGGGCGAGGGGTTTATGTAG R : AACATGCACAAGCAGAAAACA
5.	M.esc1	CA(8)	F : TACGGCAACTGCCAAAACAG R : CCGTTATTGTTCCTGGTCCT

TABLE 2.
PRIMARY SELECTION CRITERIA

No.	Name of Primer	Primary length	Value of Tm	GC%	Self 3` Complementarity
1.	M.esc2 F	24	57,18	40,50	0,00
	M.esc2 R	22	57,41	40,91	2,00
2.	M.esc15 F	20	55,90	40.00	2,00
	M.esc15 R	20	58,36	50,00	0,00
3.	M.esc12 F	21	59,00	52,38	0,00
	M.esc12 R	23	59,55	43,48	1,00
4.	M.esc3 F	20	58,87	55,00	0,00
	M.esc3 R	21	57,48	40,10	0,00
5.	M.esc1 F	20	59,33	50,00	1,00
	M.esc1 R	20	57,22	50,00	1,00

The specific primer design will affect the success of template DNA amplification. Factors that influence the selection of a specific primer pair are: the length of the primary base, the value of Tm (melting temperature), the percentage of the number of GCs and the possibility of forming dimers between primers [12].

The optimal primer length for the amplification process is at 18-30 oligonucleotides. Primers that have less than 18

bases can cause the primers to stick to non-specific places. While the bases in the primers number more than 30 bases, a hybridization process will occur between primers and will inhibit DNA amplification [13].

Melting temperature (Tm) is the temperature at which double strands of DNA separated by 50% become single DNA. The optimal value of Tm ranges from 50-65ºC and the difference between primers does not exceed 5ºC. If the Tm temperature is too high, it will affect the DNA amplification process, but if it is too low, it will produce non-specific DNA bands due to the attachment of primers from unwanted sequences [14]. The Tm value can also be calculated manually using the calculation Tm = 2(A+T) + 4(G+C) [15].

The percentage of G-C also affects the primary specifications. The base percentages of Guanine and Cytosine in primers range optimally from 40 to 60%. The percentage of G-C can affect the Tm value of the primer and the double splitting of the DNA template [13]. A low primary percentage will lead to a decrease in efficiency in the PCR process [15, 16]. The last factor that influences the selection of primers is the possibility of the emergence of dimers between primers. The dimer at the 3` end of the primer should not be more than 3 bases because it will affect the specificity of the primer [14].

In silico PCR stimulation was performed to pair pairs of primers with cassava (*Manihot esculenta*) genome sequences using FastPCR software (Figure 2). In silico PCR functions in predicting and demonstrating the process of gluing the primer sequences that occur on the template DNA sequence, which will minimize the occurrence of errors when PCR is carried out.

PCR results performed in silico showed that 5 pairs of designed primers could attach to the cassava (*Manihot esculenta*) genome sequence. The results of the in silico PCR test included the position of the primer on the nucleotide template and the estimated size of the amplicon (PCR product) (Table 3).

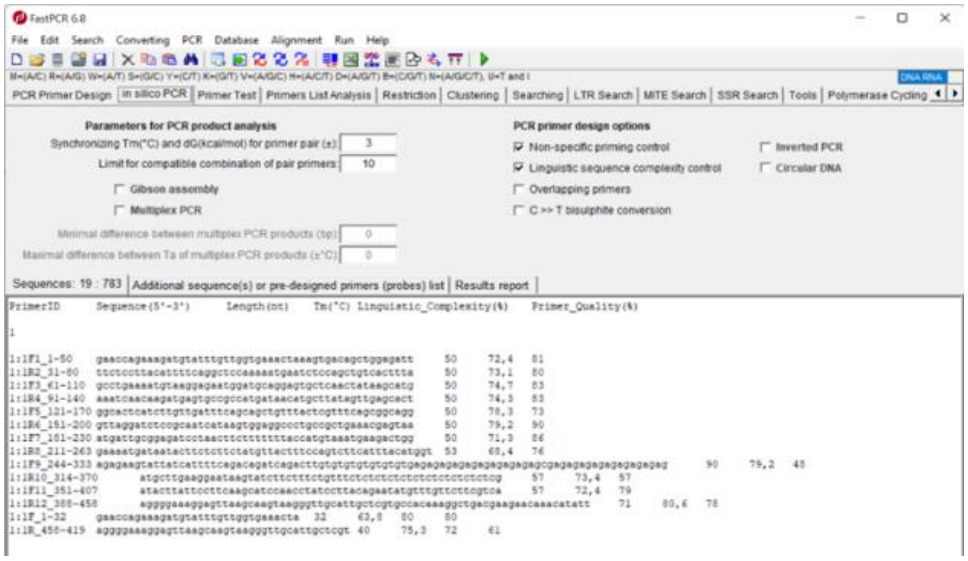

Fig.2. Display FastPCR Software

TABLE 3.
IN SILICO PCR RESULTS ON FASTPCR SOFTWARE

No	Name ofPrimer	Primary Positions on Chromosomes	Primary attachment position in nucleotides	PCR Product Size (bp)
1.	M.esc2 F	Chromosome 2	4.753.932 – 4.753.955	167
	M.esc2 R		4.754.077 – 4.754.098	
2.	M.esc15 F	Chromosome 15	2.911.092 – 2.911.111	195
	M.esc15 R		2.911.267 – 2.911.285	
3.	M.esc12 F	Chromosome 12	5.761.975 – 5.761.995	155
	M.esc12 R		5.762.107 – 5.762.129	
4.	M.esc3 F	Chromosome 3	22.473.468 – 22.473.487	161
	M.esc3 R		22.473.608 – 22.473.628	
5.	M.esc1 F	Chromosome 1	37.485.267 - 37.485.286	112
	M.esc1 R		37.485.359 - 37.485.378	

The five primer pairs amplify the microsatellite loci located on each chromosome. On the NCBI website, cassava is reported to have 18 chromosomes, so the possibility of a microsatellite locus in a different position is still very high. Microsatellite markers identified in cassava have a total of around 2146 markers [8]. Meanwhile, another report stated that the total number of SSR markers on cassava was up to 3,376 markers [17]. Another study reported that the cassava genetic map had a SNP (Single Nucleotide Polymorphism) value of 57.4% are located in coding sequence [18]. So this shows that many primers are needed to be able to detect mutations in base sequences in cassava genetics. From this research it is known that the study of microsatellite loci in cassava is still needed to be used as a reference to determine the genetic diversity of cassava in a population.

Genetic diversity is needed for plant breeding because diversity occurs due to mutations in individual plants that cause genetic diversity in one population. This can be used to determine the adaptability of plants to the environment and can improve quality in terms of production and develop seeds from superior plants [19].

Another advantage of studying genetic diversity is knowing the genetic traits of plants. Genetic bad traits in the same plant as the parent will affect the quality and quantity of the plant. By studying genetic diversity in certain plants, it is hoped that it can improve the characteristics of these plants so that it is expected to increase quality, such as production quality and freedom from disease [20].

IV. CONCLUSIONS

The in silico primer design produced 5 pairs of primers which successfully developed microsatellite loci in cassava (*Manihot esculenta*). The amplicon lengths produced for each primer were 167 bp, 195 bp, 155 bp, 161 bp and 112 bp. The basis for choosing a primer consists of the length of the primary base, the value of TM (melting temperature), the percentage of the number of GC and the possibility of forming dimers between primers.

ACKNOWLEDGMENT

From the research that has been done, we as the authors would like to thank various parties who provided support and motivation so that this research could be carried out smoothly.

REFERENCES

[1] Salsabila, H. Y., Yasmin, A.M. Putera, A. Baskoro. 2021. Correlation between nutritional status and children's activity with food allergy: A cross-sectional study. *Annals of Medicine & Surgery*. 68(1): 1-4

[2] Duhita KAN. 2018. Prevalensi Penyakit Alergi pada Anak Usia 6-7 Tahun dan 13-14 Tahun di Daerah Istimewa Yogyakarta. Universitas Gadjah Mada.

[3] Suseno, R., Palupi, N. S., & Prangdimurti, E. (2017). Alergenisitas Sistem Glikasi Isolat Protein Kedelai-Fruktooligosakarida (Allergenicity Properties of Soy Protein Isolate-Fructooligosaccaride Glycation Systems). *Agritech Journal* 36(4): 450-461.

[4] Gupta, R. S., Dyer, A. A., Jain, N., & Greenhawt, M. J. (2013). Childhood food allergies: current diagnosis, treatment, and management strategies. *Mayo Clinic Proceedings* 88(5): 512–526.

[5] Iweala, O. I., Choudhary, S. K., & Commins, S. P. 2018. Food Allergy. In Current Gastroenterology Reports. *Curr Gastroenterol Rep.* 20(5): 17:28

[6] Karim, K. Y., Ifie, B., Dzidzienyo, D., Danquah, E. Y., Blay, E. T., Whyte, J. B. A., Kulakow, P., Rabbi, I., Parkes, E., Omoigui, L., Norman, P. E., & Iluebbey, P. 2020. Genetik characterization of cassava (*Manihot esculenta* Crantz) genotypes using agro-morphological and simple nucleotide polymorphism markers. *Physiology and Molecular Biology of Plants* 26(2): 317–330.

[7] Suryana, A. 2016. Kebijakan Penelitian dan Pengembangan Ubi kayu untuk Agroindustri dan Ketahanan Pangan . Lokakarya Pengembangan Ubi Kayu Di Balitkabi. Malang.

[8] Ramadanti, Nur Ayu, Hartati, Putri, Dwi Hilda, Hartati, N.Sri. 2020. Molecular Characteristics of Cassava Carvita 25 Somaclonal Variant Using SSR Marker. *Jurnal Ilmu Dasar*. 21(2):87-96

[9] Agarwal, M., Shrivastava, N., & Padh, H. 2008. Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep*. 27(4), 617–631.

[10] Karihaloo, J. L. 2015. DNA Fingerprinting Techniques for Plant Identification. Plant Biology and Biotechnology. Vol II: Plant Genomics and Biotechnology pp. 205–221. Springer.

[11] Deratih, B. P., & Achyar, A. 2021. Primer Design and in Silico PCR for Detection Shigella Sp. on Refilled Water Samples Desain Primer dan PCR In Silico untuk Deteksi *Shigella* Sp. pada Sampel Air Minum Isi Ulang. *Serambi Biologi*. 6(1):1-6

[12] Saraswati, H., Dwi Wahyuni, F., Bioteknologi, P., Ilmu-ilmu Kesehatan, F., & Esa Unggul, U. 2019. Desain Primer Secara In Silico untuk Amplifikasi Gen cryIII dari *Bacillus thuringiensis* Isolat Lokal. *Indonesian Journal of Biotechnology and Biodiversity*. 3(1):33-38

[13] Sasmitha, V.L., Sanna Yustiantara, P., dan S. C. Yowani. 2018. Desain DNA Primer secara In Silico sebagai Pendeteksi Mutasi Gen gyrA Mycrobacterium

tuberculosis untuk Metode Polymerase Chain Reaction. *Cakra Kimia (Indonesian E-Journal of Applied Chemistry*. 6(1): 63-69

[14] Pradnyaniti, D.G., S.C. Yowani. 2013. Desain Primer secara in silico untuk Amplifikasi Fragmen Gen rpoB Mycobacterium tuberculosis dengan Polymerase Chain Reaction (PCR). *Jurnal Farmasi Udayana*. 2(3):124-130

[15] Borah, P. 2011. Primer Designing For PCR. *Science Vision*, 11(3): 134–136.

[16] Sasmito, D.E. K., Kurniawan R., Muhimmah, I. 2014. Karakteristik Primer pada Polymerase Chain Reaction (PCR) untuk Sekuensing DNA: Mini Review. Prosiding Seminar Nasional Informatika Medis (SNIMED).

[17] Ferguson, M., Rabbi, I., Kim, D.-J., Gedil, M., Lopez-Lavalle, L. A. B., & Okogbenin, E. 2012. Molecular Markers and Their Application to Cassava Breeding: Past, Present and Future. *Tropical Plant Biology*, 5(1): 95–109.

[18] Soto, J.C., Ortiz J.F., Perlaza-Jiménez L., Vásquez A.X., Lopez-Lavalle L.A.B., Mathew B., 2015. A genetic map of cassava (Manihot esculenta Crantz) with integrated physical mapping of immunity-related genes. *BMC Genomics*.16(1):190.

[19] Yunita R. 2009. Pemanfaatan Variasi Somaklonal dan Seleksi in Vitro dalam Perakitan Tanaman Toleran Cekaman Abiotik. *Jurnal Litbang Pertanian*. 18(4):142–8.

[20] Basuki, TMD. 2018. Analisis Keragaman Genetik Sepuluh Aksesi Cabai Merah (*Capsicum annuum* L.) Indonesia Menggunakan Marka Random Amplified Polymorphic DNA. Universitas Jendral Soedirman.