J. Trop. Agric. and Fd. Sc. 46(1)(2018): 17 – 23

Characterisation of 13 polymorphic simple sequence repeat (SSR) markers for cassava (*Manihot esculenta* Crantz)

[Pencirian 13 penanda polimorfik *simple sequence repeat* (SSR) bagi ubi kayu (*Manihot esculenta* Crantz)]

I. Siti Norhayati[1], N.S. Nasarudin[2], A.R. Shahril Firdaus[1], M.Y. Muhammad Fairuz[1] and J. Mohd Azwan[1]

[1] Genebank and Seed Bank Research Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor, Malaysia
[2] Horticulture Research Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor, Malaysia

Abstract

Cassava is widely cultivated as an annual crop for carbohydrates source in tropical and subtropical regions. It provides food security against famine due to its inherent tolerance to stressful environments. The present study aims to provide preliminary information on 13 polymorphic SSR markers to both researchers and breeders, as it is essentially useful to manage their breeding program. These 13 SSR markers displayed a wide range of polymorphism whereby the number of alleles ranged from 2 (SSRY106) to 10 (SSRY324) with an average of 5.5 alleles per locus. The highest polymorphism information content (PIC) is shown in marker SSRY324 with the value of 0.9999, while SSRY106 showed the lowest value of 0.3893. The observed and expected heterozygosity varied from 0.2549 to 0.9391, and 0.3902 to 0.8192 with an average of 0.6309 and 0.6067, respectively. These markers indicated sufficient levels of polymorphism that could further be employed in studies on population and genetic variation, as well as cultivar identification of cassava in Malaysia.

Keywords: heterozygosity, allele, molecular markers, genetic diversity, breeding programme

Introduction

Due to its inherent tolerance to stressful environments, the starchy root crop cassava (*Manihot esculenta* Crantz) is considered as a food security source against famine (El-Sharkawy 2004 and Fregene et al. 2003) and the edible roots of cassava can be kept under ground for a long period of time, i.e. between 6 to 24 months after planting (Alves 2002). Cassava (2n = 36) is native in Latin America and grown mostly in tropical countries. It contributes to approximately 233 million tonnes of global production. Half of the total world production comes from Africa, 33% from Asia and 15% from America (Rabbi et al. 2012). Despite its importance, the cultivation of cassava in Malaysia has declined to only 2,700 hectares with a production of 37,200 tonnes of fresh root in 2010 (MOA 2010). Besides being consumed fresh, cassava starch is also extracted to produce monosodium glutamate (MSG) by local food industries as flavour enhancer (Mohamed Noor et al. 2012).

Article history
Received: 21.8.17
Accepted: 3.11.17

Authors' full names: Siti Norhayati Ismail, Noor Syahira Nasarudin, Mohd Shahril Firdaus Ab Razak, Muhammad Fairuz Mohd Yusof and Mohd Azwan Jaafar
E-mail: yati@mardi.gov.my

Cassava is commonly propagated by stem cutting and with regard to breeding programs, cassava is propagated by seed pollination. However, it is worth highlighting that seeds take longer time to develop, are smaller in size and are less vigorous compared to stem cuttings (Alves 2002). Although cassava has monoecious flowers, self-pollination is prevented due to protogyny, a phenomenon where the female organ matures before the male (Rabbi et al. 2012). Consequently, seedlings will genetically be segregated through cross-pollinations under natural conditions. Farmers use these seedlings as planting materials occasionally, which contributed to the genetic diversity of cassava (Lokko et al. 2006).

Studies on genetic diversity provide important information with regards to predicting heterosis or hybrid vigorousness. Heterosis is a phenomenon whereby progenies, which are obtained from crossing diverse parental lines within or between species yield significant increase in F_1 value over mid-parent value, in which the progenies obtained superior phenotypes compared to both parents (Birchler et al. 2010 and Virmani 2012). The uses of molecular markers are of utmost importance in evaluating the degree of variation within an observed population and hence, enable breeders to select the parental lines to be used in a breeding program.

Simple sequence repeat (SSR) is one of the most widely employed molecular markers in genetic diversity studies (Fregene et al. 2003; Moyib et al. 2007; Asare et al. 2011 and Turyagyenda et al. 2012). Studies of genetic diversity of pro-vitamin A (Esuma et al. 2012) and cassava mosaic virus (Lokko et al. 2006) in cassava germplasm by means of using SSR markers were previously carried out. Notably, SSRs are co-dominant markers, multi-allelic, display high reproducibility rate and polymorphism level and amenable for high throughput automation platform. Given such a background, the objective of this study is to assess the polymorphism of cassava germplasm collection in Malaysian Agricultural Research and Development Institute (MARDI).

MARDI has released several starch and table varieties namely Black Twig, Medan, Perintis, MM92, Sri Kanji 1, Sri Kanji 2 and Sri Pontian from its cassava breeding and selection program (Tan 2001). These clones were chosen from a pool of cassava genetic resources (germplasm) maintained at MARDI resulting from a slightly modified breeding and selection scheme from those applied at Centro International de Agriculture Tropical (CIAT) and International Institute of Tropical Agriculture (IITA) (Ceballos et al. 2012).

Besides genetic studies, polymorphic markers can also be applied in breeding programmes in order to facilitate the selection of important traits in cassava such as high yield, disease resistance, high carotene level and low cyanide content in cassava varieties. In future, the development of high yielding and high quality cassava varieties can be enhanced through the use of marker-assisted selection (MAS) or marker-assisted backcrossing (MAB).

Materials and methods

Molecular markers and cassava germplasm in MARDI

Table 1 shows the 13 selected SSR markers from previous studies (see Raghu et al. 2007 and Moyib et al. 2007). These 13 SSRs were genotyped on 231 individual plants from a total of 77 accessions obtained from MARDI's cassava germplasm collection that were maintained in field. These cassava accessions were acquired from local farmers and from international research centres such as CIAT. Four months after planting, young leaf samples from individual plants were collected prior to DNA extraction.

Table 1. Information on the thirteen polymorphic SSR markers used in the study

Primer ID	Locus	Primer sequences (5'-3')	Repeat motive	Repeat type	No. of alleles	Ta (°C)	Product size	Null Allele	H°	H^e	PIC
Ca_1	SSRY8	F:AGTGGTTTGAGAAGACTGGTGA R: TTTCCAAAATGGAACTTCAAA	$(CA)_{14}$ CT$(CA)_2$	Di	8	57.7	287 – 314	Yes	0.6089	0.6714	0.9942
Ca_3	SSRY28	F:TTGACATGAGTGATATTTTCTTGAG R: GCTGCGTGCAAAACTAAAAT	$(CA)_3(AT)_3AC(AT)_2$	Di	7	53.9	174 – 200	No	0.7263	0.6946	0.9971
Ca_4	SSRY32	F: CAAATTTGCAACAATAGAGAACA R: TCCACAAAGTCGTCCATTACA	$(CA)_{11}$	Di	5	56.0	303 – 325	No	0.8458	0.6618	0.9798
Ca_6	SSRY39	F: TCAATGCATAGGATTTTGAAAGTA R: AATGAAATGTCAGCTCATGCT	$(CT)_{24}AT(CT)_3(AT)_3$	Di	8	53.9	294 – 327	No	0.7778	0.7086	0.9980
Ca_13	SSRY235	F: CAGCTTTGCCATCCAATTTT R:CAGCAAAATGACATGAGTGTATCTC	NI	NI	3	56.0	231 – 263	No	0.6383	0.6397	0.8620
Ca_18	SSRY45	F: CAAGTGGATGAGCTACGCAA R: CCGCTTAACTCCTTGCTGTC	$(CT)_{27}$	Di	3	53.9	256 – 286	No	0.5255	0.4512	0.6705
Ca_21	SSRY50	F: CAAGAACGCCAATATGCTGA R: GGCTGCTTTACCTTCTACTCAGA	$(CA)_6(N)_6(GA)_{31}$	Di	4	56.4	237 – 250	No	0.4394	0.4335	0.7192
Ca_24	SSRY78	F: ACAATTCATCATGAGTCATCAACT R: CCGTTATTGTTCCTGGTCCT	$(CT)_{22}$	Di	5	58.8	264 – 291	No	0.8350	0.6535	0.9755
Ca_29	SSRY106	F: CTGATCAGCAGGATGCATGT R: GCAGTAAAACCATTCCTCCAA	$(CT)_{24}$	Di	2	59.7	258 – 301	Yes	0.2549	0.3902	0.3893
Ca_14	SSRY324	F: CGCTTACAACACCACCTTCA R: GCTTGATCTCAGCCATGTCA	NI	NI	10	61.0	192 – 290	Yes	0.6349	0.8192	0.9999
Ca_16	SSRY3	F: AACGTAGGCCCTAACTAACCC R: ACAGCTCTAAAAACTGCAGCC	$(CA)_{17}$	Di	4	56.0	107 – 127	No	0.4365	0.3616	0.6451
Ca_20	SSRY51	F: GCTGCGTGCAAAACTAAAAT R: TTGACATGAGTGATATTTTCTTGAG	$(CT)_{11}CG(CT)_{11}(CA)_{18}$	Di	6	62.7	174 – 196	Yes	0.5389	0.6289	0.9850
Ca_28	SSRY111	F: CACTCCGTTGCAGGCATTA R: CGATCTCAGTCGATACCCAAG	$(GA)_{29}$	Di	6	47.5	226 – 261	No	0.9391	0.7725	0.9990
		Average			5.5		56.4231		0.6309	0.6067	0.8627

Di = Dinucleotide; NI = No information; T*a* = Annealing temperature; H° = Observed heterozygosity; H^e = Expected heterozygosity; PIC = Polymorphic information content

DNA extraction, quality analysis and normalisation

Total genomic DNA of individual plants were isolated by means of using a standard high throughput plant DNA extraction method established at the Centre for Marker Discovery and Validation (CMDV) laboratory, MARDI Headquarters, Serdang. Approximately 1 g of cassava leaf was ground using TissueLyser (Qiagen, Netherlands). Subsequently, 600 μl of DNA extraction buffer (2% PVP, 4 mM DIECA, 5 mM ascorbic acid, 1.4 M $NaCl_2$, 100 mM Tris-HCL pH 8.0 and 20 mM EDTA) was added to the ground leaf sample before incubated in water bath (Memmert, German) at 65 °C for 1 hour. An equal volume of cold isopropanol was employed to precipitate the DNA and was further washed with 70% ethanol. The DNA pellet was air dried and re-suspended in 50 μl TE (Tris-EDTA) buffer. The quality of isolated DNA was observed on 0.8% pre-stained (Ethidium Bromide) agarose gel and visualised under ultraviolet light (UV), Quantum ST4 3000 (Thermo Scientific, USA). The DNA concentration was measured using Thermo Labsystems Floroskan Ascent^{TM} (Thermo Scientific, USA) and was subsequently normalized to 30 ng/μl using automated workstation, Janus (Perkin Elmer, USA).

PCR amplification and fragment analysis

Polymerase chain reaction (PCR) amplification was performed using thermal cycler (Applied Biosystems, USA) in a total reaction volume of 10 μl. The reaction mixture consisted of 10X PCR buffer, 50 mM $MgCl_2$, 2 mM dNTPs, 10 μM of forward primer anchored with M13 tail (Schuelke 2000), 10 μM of unlabelled reverse primer and 5 μM of M13 fluorescent dyes (FAM, VIC, PET or NED), 5U Taq Polymerase (Invitrogen, USA) and approximately 30 ng/μl of template DNA. The PCR thermal conditions were: 5 min at 95 °C, followed by 30 cycles for 1 min at the annealing temperature of each primer pair (see *Table 1*) and 1 min at 72 °C, and the final step at 72 °C for 10 min. Next, fragment analysis was conducted using DNA Analyser (model: ABI3730XL) (Applied Biosystems, USA). GeneScan^{TM} 500 LIZ was employed as size standard, which can identify DNA fragments ranging from 35 to 500 base pair (bp). DNA fragment produced by DNA analyser was scored by means of using GeneMapper Software Version 4.0 (Applied Biosystems, USA).

Statistical analysis

Initially, the genotypic data was formatted in Microsoft Excel^{TM} and was converted into several different formats using CONVERT software (Glaubitz 2004). Micro-Checker (version 2.2) was utilised to analysed the presence of null alleles (Van et al. 2004), while the number of alleles (N_A) per locus was calculated using POPGENE (version 1.31) (Yeh et al. 1999). GENEPOP (version 3.4) was employed to calculate the observed and expected heterozygosity (Raymond and Rausset 2004) and the polymorphic information content (PIC) was calculated using Power Marker software (version 3.5) (Liu and Muse 2005).

Results and discussion

A total of 13 SSR markers were successfully genotyped to 77 cassava accessions held in MARDI. Detailed information for each SSR marker is demonstrated in *Table 1*. The number of alleles detected across loci ranged from 2 alleles for primer locus SSRY106 to 10 for SSRY324 with an average of 5.5 alleles per locus. This finding suggested that the SSR markers employed in the present study possess a wide degree of polymorphism, which can thus be used in further genetic studies.

It is noteworthy that the polymorphism information content (PIC) value represents the strength of markers to evaluate polymorphism in a population. The PIC values of 13 SSR markers were calculated and the most informative marker was determined as the one carrying the highest PIC value. As regards the present study,

the PIC value ranged from 0.3893 for primer locus SSRY106 to 0.9999 for SSRY324 with an average of 0.8627. According to Liu and Cordes (2004), PIC value is highly dependent on the number of detected alleles and their frequency of distribution. Hence, the researchers found that SSRY324 with the highest PIC value had the highest number of alleles, while SSRY106 with the lowest PIC value had the lowest number of alleles.

In addition, a high number of alleles would produce a high value of expected heterozygosity, which indicates high value of genetic diversity. The observed and expected heterozygosity in this study varied from 0.2549 to 0.9391, and 0.3616 to 0.8192 with an average of 0.6309 and 0.6067 respectively (*Table 1*). Apart from that, SSRY106 consistently exhibited the lowest observed and expected heterozygosity with the lowest number of alleles, while SSRY324 showed the highest expected heterozygosity with the highest number of alleles.

The occurrence of null alleles was detected at four primer loci namely SSRY8, SSRY106, SSRY324, and SSRY51 (*Table 1*). There are several potential causes of null alleles including: 1) poor primer annealing due to point mutations or indels, and 2) preferential amplification of short alleles, which are also known as 'partial nulls' due to inconsistent quality of the DNA template (Dakin and Avise 2004). Furthermore, it can also be the case of false impression of null alleles presence that is caused by biological factors such as Wahlund's effects or inbreeding (Dakin and Avise 2004 and Chakraborty et al. 1992).

Conclusion

It was revealed in the present study that the selected SSR markers showed sufficient levels of polymorphism that can be utilised in genetic variability studies on our local cassava germplasm collection. It is noteworthy that these polymorphic SSR markers would be valuable for breeders to identify or resolve genetic diversity of cassava germplasm in order to estimate higher heterosis in a breeding programme. Also, it is important for genetic improvement on commercial varieties or the development of new cassava cultivars.

Acknowledgement

This project was funded by the Rancangan Malaysia Ke-10 Projek Mega and the National Key Economic Area (NKEA) Entry Point Project 14 (EPP 14) grant. We would like to acknowledge our deepest thanks to our project members from CMDV and members from the former Rice and Industrial Crops Research Centre.

References

Alves, A.A.C. (2002). Cassava botany and physiology. In *Cassava: Biology, production and utilization*. (R.J. Hillocks, J.M. Thresh and A.C. Bellotti ed.), p. 67 – 90. New York: CABI Publishing

Asare, P.A., L.K.A. Galyuon, J.K. Sarfo, and J.P. Tetteh. (2011). Morphological and molecular based diversity studies of some cassava (Manihot esculenta crantz) germplasm in Ghana. *African Journal of Biotechnology* 10(63): 13900 – 13908

Birchler, J.A., Yao, Chudalayandi, H.S., Reiner, A. and Veitia, R.A. (2010). Heterosis. *The Plant Cell* 22(7): 2105 – 2112

Ceballos, H., Kulakow, P. and Hershey, C. (2012). Cassava breeding: Current status, bottlenecks and the potential of biotechnology tools. *Tropical Plant Biol*. 5: 73 – 87

Chakraborty, R., De Andrade, M., Daiger, S.P. and Budowle, B. (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Ann. Hum. Genet*. 56: 45 – 57

Dakin, E.E., and J.C. Avise. Microsatellite null alleles in parentage analysis. (2004). *Heredity* 93: 504 – 509

El-Sharkawy, M.A. (2004). Cassava biology and physiology. *Plant Molecular Biology* 56: 481 – 501

Esuma, W., Rubaihayo, P., Pariyo, A., Kawuki, R., Wanjala B., Nzuki I., Jagger J.W. Harvey and Baguma Y. (2012). Genetic diversity of provitamin A cassava in Uganda. *Journal of Plant Studies* 1 (1): 60 – 71

Fregene, M.A., Suarez, M., Mkumbira, J., Kulembeka, H.E., Ndedya, Kulaya, A., Mitchel, S. Gullberg, U., Rosling, H., Dixon, A.G.O., Dean, R. and Kresovich, S. (2003). Simple sequence repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. *Theor. Appl. Genet*. 107: 1083 – 1093

Glaubitz, J. (2004). Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes* 4(2): 309 – 310

Liu, K. and Muse, S.V. (2005). PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics Application Note* 21(9): 2128 – 2129

Liu, Z.J. and Cordes, J.F. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238: 1 – 37

Lokko, Y., Dixon, A., Offei, S., Danquah, E. and Fregene, M. (2006). Assessment of genetic diversity among African cassava Manihot esculenta Grantz accessions resistant to the cassava mosaic virus disease using SSR markers. *Genetic Resources and Crop Evolution* 53: 1441 – 1453

Mohamed Noor, N., Shariff, A. and Abdullah, N. (2012): Slow pyrolysis of cassava wastes for biochar production and characterization. *Iranica Journal of Energy and Environment* 3: 60 – 65

Moyib, O.K., Odunola, O.A. and Dixon, A.G.O. (2007). SSR markers reveal genetic variation between improved cassava cultivars and landraces within a collection of Nigerian cassava germplasm. *African Journal of Biotechnology* 6: 23 2666 – 2674

Rabbi, I.Y., Kulembeka, H.P., Masumba, E., Marri, P.R. and Ferguson, M. (2012). An EST-derived SNP and SSR genetic linkage map of cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics* 125(2): 329 – 342

Raghu, D., Tairo, F.M. and Rweyemamu, C.L. (2007). Morphological and simple sequence repeats (SSR) based fingerprinting of South Indian cassava germplasm. *International Journal of Integrative Biology* 1(2): 141 – 148

Raymond , M. and Rousset, F. *Genepop on the web*. Retrieved on 28 May 2017 from http://genepop.curtin.edu.au/accessed

Schuelke, M. (2000): An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18(2): 233 – 234

Tan, S.L. (2001). Cassava breeding and agronomy research in Malaysia during the past 15 years. Paper presented at the Present Situation and Future Research and Development Needs in Proc. 6th Regional Workshop, Ho Chi Minh

Turyagyenda, L.F., Kizito, E.B., Ferguson, M.E., Baguma, Y., Harvey, J.W., Gibson, P., Wanjala, B.W., Osiru, D.S.O. (2012). Genetic diversity among farmer-preferred cassava landraces in Uganda. *African Crop Science Journal* 20: 15 – 30

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. and Shipley, P. (2004). Micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535 – 538.

Virmani, S.S. (2012). *Heterosis and hybrid rice breeding*. Springer Science and Business Media

Yeh, F.C., Yang, R.C. and Boyle. T. (1999). *Popgene, version 1.31. Microsoft window-based freeware for population genetic analysis. Quick User Guide*. Alberta: University of Alberta

Abstrak

Ubi kayu ditanam secara meluas sebagai tanaman tahunan yang merupakan sumber karbohidrat dan di kawasan tropikal dan subtropikal. Ia dianggap sebagai satu sumber keselamatan makanan terhadap kebuluran kerana toleransi sedia ada dengan persekitaran yang tertekan. Kajian ini bertujuan untuk menyediakan informasi kajian awal mengenai 13 penanda molekul SSR kepada penyelidik dan pembaik baka yang mana ia boleh digunakan dalam pengurusan program pembiakbakaan mereka. Sebanyak 13 penanda molekul SSR ini menunjukkan tahap polimorfisme yang tinggi yang mana bilangan alel adalah antara 2 (SSRY106) hingga 10 (SSRY324) dengan purata 5.5 alel per lokus. Penanda SSRY324 mempunyai nilai kandungan maklumat polimorfik (PIC) tertinggi iaitu 0.9999, manakala SSRY106 mempunyai nilai yang terendah iaitu 0.3893. Keheterozigotan yang diperhatikan dan dijangka adalah berbeza antara 0.2549 hingga 0.9391, dan 0.3902 hingga 0.8192 dengan purata masing-masing antara 0.6309 dan 0.6067. Penanda-penanda ini menunjukkan tahap polimorfisme yang mencukupi untuk digunakan dalam kajian populasi dan variasi genetik serta pengenalpastian genotip varieti ubi kayu di Malaysia.