Full Length Research Paper

Isolation and characterization of microsatellite repeats in Marama bean (*Tylosema esculentum*)

M. Takundwa¹, P. M. Chimwamurombe^{1*}, K. Kunert² and C. A. Cullis³

¹Department of Biological Sciences, University of Namibia, P. Bag 13301, Windhoek, Namibia. ²Department of Plant Sciences, University of Pretoria, 0001 South Africa. ³Department of Biology, Case Western Reserve University, Cleveland, Ohio, USA.

Accepted 30 November, 2009

Marama bean [*Tylosema esculentum* (Burchell) Schreiber] occurs naturally in arid and dry parts of Southern Africa, including Botswana, Namibia and South Africa. Due to the high nutrient value of the seeds and tubers, rich in protein, oil and starch, it is a potential crop for arid areas where few conventional crops can survive. The objectives of this study were to design microsatellite primers based on microsatellite repeats identified in Marama bean and to determine their ability to detect polymorphisms for use in diversity characterization. DNA fragments of *T. esculentum* containing microsatellite loci were isolated by enrichment of genomic DNA by a modified FIASCO technique and sequenced. Nine Marama bean microsatellite libraries enriched for (AAG)₇, (GTT)₇, (AGG)₇, (GAG)₇, (CA)₁₀, (CT)₁₀, (TCC)₇, (CA)₁₅ and (CAC)₇ were created and thirty microsatellite primers based on the microsatellite regions of marama bean were designed and screened for polymorphism. Polymorphic microsatellites that were identified can be used as markers in future breeding programs. The screening for the polymorphic SSRs is being carried out to reach a target of 50 or more polymorphic SSR markers. It appears that the marama bean germplasm has abundant polymorphic microsatellites as 77% of the 30 loci screened initially were polymorphic.

Key words: Marama bean, microsatellite markers, primers, polymorphism, maker assisted selection (MAS).

INTRODUCTION

Marama bean, a future crop

According to the most recent accounts, the genus Tylosema (Schweinf.), Torre and Hillc. (Leguminosae, Caesalpinioidea) comprises four currently accepted species [T.esculentum (Burch.) A. Schreib., T. fasssoglense (Schweinf.) Torre and Hillc. *T*. argenteum (Chiov.) Brenan and *T*. humifusum (Pic.Serm and Roti Mich.) Brenan] occurring in eastern and central tropical Africa (Coetzer and Ross, 1977; Thulin, 1993). Marama remote and arid regions (Keegan and van Staden, 1981). Due to the high nutrient value of the seeds and tubers, rich in protein, oil and starch, marama bean has the potential to become a productive crop in arid areas where few conventional crops can survive (Keetshajwang et al., 1998).

Marama bean is a wild tuber producing and nonnodulating legume, the seeds and tubers are edible after roasting and cooking respectively. Unfortunately, marama bean is not vet domesticated but remains a wild plant with so much potential. Generally, the plant is low yielding, produces few seeds (one or two seeds per pod) so collecting seed in the wild is not a sustainable way of reducing the malnutrition problems in this Southern African region. The plant needs to be developed into a crop and further developed into desirable cultivars with high vielding early character and maturing (Chimwamurombe, 2008; Nepolo et al., 2009).

The need for molecular breeding tools for marama bean

Information on genetic variation is a prerequisite for the improvement of any plant species by breeding programs. The natural populations of Marama bean are under pressure from both overgrazing and human exploitation

^{*}Corresponding author. E-mail: pchimwa@unam.na. Tel: +264612063358, +264812744678. Fax: +264612063791.

of the seeds: therefore a detailed knowledge of the genetic structure of these populations is required for developing the remaining wild germplasm (Naomab, 2004). Markers based on polymorphic DNA fragments that are independent of the growing environment of a plant and can be unambiguously scored are one class of molecular markers that are emerging and being extensively used in plant diversity analysis and as a tool in plant breeding through Marker Assisted Selection (MAS) (Karp et al., 1996; Karp et al., 1997).

Microsatellite markers

Microsatellites (also known as simple sequence repeats, SSRs, or hypervariable sequences) are arrays of short tandem repeat motifs of 1 - 5 base pairs in length. In recent years, microsatellites have become a popular tool for genetic mapping and genome analysis (Chen et al., 1997; Li et al., 2000) genotype identification and variety protection (Senior et al., 1998) seed purity evaluation and germplasm conservation (Brown et al., 1996), diversity studies (Xiao et al., 1996) and marker assisted breeding (Weising et al., 1998) among other uses. This popularity stems from the high allelic diversity or hypervariability found at the SSR loci, the co-dominance of alleles and the relatively simple polymerase chain reaction (PCR)based screening methods that are reproducible. In addition, microsatellites occur frequently and randomly in eukaryotic genomes and are highly informative markers (Wang et al., 1994). For measuring genetic diversity, assigning lines to heterotic groups and genetic fingerprinting, microsatellites provide power of discrimination equal to or greater than that of restriction fragment length polymorphisms (RFLP) in a more cost effective manner (Smith et al., 1997; Senior et al., 1998).

Plant microsatellites have been developed for germplasm conservation, cultivar identification and for assessing genetic diversity not only in crops such as tomato (He et al., 2003), sweet orange (Novelli et al., 2006), peanut (Fergurson et al., 2004) and common bean (Gaitán-Solis et al., 2002) but also in perennial plants such as the Melaleuca (the tea tree) (Rossetto et al., 1999), Pinus (Kutil and Williams, 2001), *Cryptomeria* (Moriguchi et al., 2003) and Litchi (Viruel and Hormaza, 2004). In the past SSR's were expensive to develop and thus often limited to applications to the major commercial crops (Scott et al., 2000).

Until recently, developing SSR markers for new species was a laborious and costly exercise thus limiting their potential applications. Characterisation relied on searches through published sequences, only possible for a minority of species, or on the production of highly involved genomic libraries. Microsatellites can be isolated directly from total genomic DNA libraries, cDNA libraries, libraries enriched for specific microsatellites (Maguire et al., 2000) or from sequences deposited at the GenBank as previously reported in *Phaseolus vulgaris* and Vigna

for example (Yu et al., 1999).

Recently, the development of SSR enrichment techniques, in which selective genomic libraries containing pre-screened inserts are prepared, has increased the efficiency of SSR characterisation in new species (Zane et al., 2002). The availability of such technology opens new opportunities for large scale SSR characterization of species with no previous knowledge of genome sequence such as marama bean. The Fast Isolation by AFLP of Sequences Containing repeats (FIASCO) technique is one such method and it was used here. The method is fast and simple and many unnecessary steps found in other protocols are eliminated (Zane et al., 2002). The protocol relies on the extremely efficient digestion-ligation reaction of the amplified fragment length polymorphism procedure (AFLP, Vos et al., 1995).

Objectives

The objectives of this study were to design microsatellite primers based on conserved flanking regions of microsatellite repeats isolated by a modified FIASCO technique in Marama bean and to determine their ability to detect polymorphisms for use in genetic diversity characterization of Marama bean.

MATERIALS AND METHODS

DNA extraction and enrichment for SSR's using the FIASCO technique

DNA was extracted from each of the 361 plant samples collected from 12 sampling sites. The 12 sites were composed of 11 different geographic locations in Namibia (Figure 1) and 1 location in South Africa- Pretoria (not shown in the map). The DNeasy mini protocol for purification of total DNA from plant tissue was used. Genomic DNA was isolated from the plant tissue according to the manufacturer's protocol. The DNA was stored in clearly labelled microcentrifuge tubes at -20°C. DNA with a concentration of 25 -250 µg/µl was collected and the concentration was determined on a 1% vagaries gel stained with ethidium bromide using known molecular weight standards. The DNA samples from Omipanda (Figure 1) were used for the purposes the FIASCO experiment. The Fast Isolation by AFLP of Sequences Containing microsatellite repeats (FIASCO) method has been used to isolate microsatellites in other plants successfully and a modified method was applied to marama bean (Zane et al., 2002). Nine Marama bean microsatellite libraries enriched for (AAG) 7, (GTT) 7, (AGG) 7, (GAG) 7, (CA) 10, (CT) 10, (TCC) 7, (CA) 15 and (CAC) 7 were created using the modified FIASCO technique. The modifications included use of Msp I, Csp 6I and Sau 3A restriction enzymes in the restriction ligation reaction instead of Msel as well as the use of more biotinylated probes. The enriched libraries were sent for 454 sequencing at Ingaba biotechnologies laboratory in South Africa. Individual DNA samples from each of the individuals, which were collected from the 12 sampling locations, were pooled into one sample for each of the sites. Equal amounts from individual samples were taken into the pooled sample for each location. The 12 pooled DNA samples were then used for screening the microsatellites.

The concentration of the pooled samples was determined both by



Figure 1. The 11 sites in Namibia, which were used to collect samples for DNA extraction.

spectroscopy and electrophoresis to ensure the template concentration used in each PCR reaction was equal. All the samples were diluted to 10 $ng/\mu l$.

polymorphic or monomorphic based on the separation of amplification products and banding patterns generated on the agarose gels in the different DNA templates.

Primer design and PCR amplification

The Simple Sequence Repeat Identification Tool (SSRIT) which identifies perfect SSRs in any given sequence was used to identify the microsatellites in the contig files obtained from 454 sequencing at Inqaba biotechnologies laboratory (Temnykh et al., 2001). Sequences containing perfect microsatellites with sufficient sequence information of the flanking region were used to design PCR primers complementary to the flanking region of the microsatellites. Thirty microsatellite primers pairs were designed from the sequence data of the contig files using Primer 3 software available online (http://fokker.wi.mit.edu/primer3/). The primers were synthesized by Inqaba biotechnologies laboratory in South Africa and then used for amplification of pooled marama DNA from the 12 different localities. Each microsatellite primer marker was given a name consisting of the prefix "MARA" followed by a number (001 - 030). The initial 30 microsatellite primer pairs were used in PCR amplification of marama DNA from the 12 ecotypes. The reaction was performed in a total volume of 20 µL for each of the 30 primer pairs. The PCR programme for microsatellite amplification consisted of initial denaturation step of 94 °C for 4 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 60 s and extension at 72°C for 2 min. A final extension at 72°C for 5 min was performed. The same pooled DNA template samples were used in all 30 reactions to be able to distinguish the SSR primers as monomorphic or polymorphic with regard to the panel of 12 marama pooled DNA samples used. Agarose gel (2%) visualization of PCR products was then used to determine if a primer pair was

RESULTS

The first 30 microsatellite primer pairs screened showed polymorphisms in 23 out of the 30 primers pairs screened (Figure 2). The other 7 microsatellite primers screened showed clearly monomorphic banding patterns on amplification of the different marama DNA samples (Figure 3).

The primers designed were assessed for their ability to detect polymorphism using electrophoresis gels. The microsatellites were described as polymorphic (Figure 2), giving more than one form, or monomorphic (Figure 3), giving one form. Table 1 below provides sequence information of the 30 primers used, the microsatellite repeat targeted as well as a description of the locus as monomorphic or polymorphic. The microsatellites ranged from 8 base pairs to 24 base pairs in length. The primer MARA 003 (Figure 2) amplified 5 alleles in marama bean whereas primer MARA 004 (Figure 3) amplified only a single allele. The polymorphic primers such as MARA 003 can thus be used in genetic variation studies while the monomorphic primers will not be applicable for such studies.



Figure 2. Gel electrophoresis photograph showing PCR amplification products with microsatellite primer MARA 003 for DNA samples from 12 different localities Namibia and South Africa. The gel was run at 90 V for 65 min on a 2% agarose gel matrix. An example of a polymorphic microsatellite.



Figure 3. Gel electrophoresis photograph showing PCR amplification products with microsatellite primer MARA 004 for DNA samples from 12 different localities in Namibia and South Africa. The gel was run at 90 V for 65 min on a 2% agarose gel matrix. An example of a monomorphic microsatellite.

DISCUSSION AND CONCLUSION

Microsatellite primers based on conserved flanking regions of microsatellite repeats isolated by a modified FIASCO technique in marama bean were designed and used to determine their ability to detect polymorphisms for future use in genetic variation studies. This is the first report of microsatellite markers for the species *Tylosema esculentum*. It appears that the marama bean germplasm is rich with polymorphic microsatellites as 77% of the loci screened initially were polymorphic. This is in line with

previous investigations on the microsatellite regions of plants which are known to be commonly occurring (Tautz, 1993; Brown et al., 1996). Furthermore, it was noted that the microsatellites above 15 base pairs in length were able to detect more polymorphisms in the marama bean genome. Dinucleotide and trinucleotide repeat motifs were the most common type of microsatellite identified in the marama bean genome. The primers designed for the amplification of longer SSR's were mostly polymorphic in comparison to the primers that targeted mainly 8 - 15 base pair SSR's (Table 1). All the monomorphic loci were Table 1. The 30 microsatellite primers designed for Marama bean and the classification of each primer pair as polymorphic or monomorphic.

Marama SSR primer pairs (first 30)					
Primer	Left primer (L)	Right primer (R)	SSR	Repeats	Polymorphic/ monomorphic
MARA001	GCACAACCAATTTCCTGCTT	TCCCTCACTGGCCTATATCC	GAG	5	monomorphic
MARA002	CTCCCTCCTCCTCGTAG	CGGGAGCAAATAGACCCTTT	ACC	8	polymorphic
MARA003	TCTCACCGACCGGGTCTC	CCTCTATCCCGCTCCCTATC	CTC	5	polymorphic
MARA004	TGCAGGCTTCACCAGAGTAA	TCTAAGACTGCGCACACAGC	GA	5	monomorphic
MARA005	GCTATCCGAGGGAGGATCA	GTGTCTATGTGTGCGCGTGT	AG	6	polymorphic
MARA006	GCTATCCGAGGGAGGATCA	TCCCATTCAGCCATTTTAGG	TG	7	polymorphic
MARA007	TATCCGAGGGAGGATCATGT	TCACATCCTAAGACTCGAACTTCA	AC	6	polymorphic
MARA008	GCTGGTCCATGGCTTCAT	TTTGTAATCGGTTGACACTTTGA	TG	5	polymorphic
MARA009	GGGAGGATCAACCTCAACAA	TGTACAAAAAGCAGGCTCCA	GAA	5	polymorphic
MARA010	TGTGCTATCCGAGGGAGGAT	ACGTCGCGATTAAACAAACC	AAG	7	polymorphic
MARA011	TGTCAACGCTTACGTTGGTC	TCATTTGAAACCCTTGTACTGC	TC	8	polymorphic
MARA012	ATATGGTGGCTCGTCGATGT	GCACATAATTCGAACAGAACACA	AG	5	Polymorphic
MARA013	GCTTCTCGTACATGGGCTTT	GCATTTATCGCGAATACAGCA	TC	5	Polymorphic
MARA014	GGTGGTGGTGTAGGAGGAGA	GACTTGAGTGCATGCCATTT	AGG	5	Polymorphic
MARA015	ACTCATCCCGCTCCTAAGGT	AAACAGGCTCGTATTTTATCTTCG	TG	5	Polymorphic
MARA016	TTCAATTTTCTTCACCACAAACTC	ACAGGAAGGTCTTCCACAGC	CA	7	Polymorphic
MARA017	ACCCTTGAATTGTGGTAGGG	ACTGTGCTATCCGAGGGAGA	CT	6	Polymorphic
MARA018	ATTTTGGCTTTACCGCACAC	AGCACTCTCCAGCCTCTCAC	CTTGA	3	Polymorphic
MARA019	CCGGAACAGGAGAAGCTATG	TCAACTTTTGCAATGAACGAA	CTT	4	Polymorphic
MARA020	TGTCTTCCCCTCCTCTTCCT	TTGACACTTTGGGACTGCTG	CAG	4	Polymorphic
MARA021	GAGGGAGGATCACCACTCAG	TGGCCATCAATCATGTTACG	TGT	4	Polymorphic
MARA022	CCCCTGTACCCAAGACTCTG	TCCATGAAGTCAGGAGAAGGA	TAGC	3	Polymorphic
MARA023	ATGGGGATACTCCCGAAACT	AATGGGAGCAAGAATTTCCA	AAAG	3	monomorphic
MARA024	CCAAGAGTGGGGATGAAAGA	TTGGAATAGTTCCCCCTTCC	AGA	4	Polymorphic
MARA025	CACGTGGGTTGTACTTATCTGC	TAATGTGTTGAGCGCCGTAG	TCTTC	3	Polymorphic
MARA026	GCTGTTGGGAACCGTAGAAA	CCTATTGTCAGTGTAAGCAACCA	TC	4	Monomorphic
MARA027	TTGTTCCAAACCACAGGTCA	TGGCCATCTCCCAATTTTAC	CA	4	Monomorphic
MARA028	CTCCGCATCTGACTTCAAAA	CCTCCTCTCCCTGATTTTCC	AGA	3	Monomorphic
MARA029	CCGAGGGAGTAGTGCTTCAT	CGCCACTTAGCATTTGCTTT	TG	4	Polymorphic
MARA030	GAGCCAAAGCCATGATCCTA	CCCATGTTGTATATTCGTGGAA	CAA	3	Monomorphic

from microsatellites that were less than 15 base pairs in length. The 23 loci that gave polymorphic bands will thus be used in the marama breeding programme to link traits to microsatellite markers designed here. The traits of interest identified so far for marama bean include the number of seeds per pod produced by the plants, internode length, early flowering and time to germination to mention a few.

The loci that gave monomorphic bands: MARA 001, MARA 004, MARA 023, MARA 026, MARA 027, MARA 028 and MARA 030 are of particular interest as well as they were conserved across such a geographically diverse collection of germplasm in Namibia and a South African population (Figures 1 and 3).

These could be genes of interest linked to particular traits of marama bean and this could be established by sequencing such amplification products and carrying out BLAST searches to see if they are linked to any genes of interest. The study confirms the usefulness of SSR's in

providing a quick method of investigating genetic variation in plant populations. The FIASCO technique enrichment technique applied to marama bean here can also now be applied to other plant species being investigated in Namibia if SSR markers are sought.

The high percentage of polymorphic SSR loci (77%) in marama bean is one indication of a relatively high level of genetic diversity in marama bean. A study being carried out in line with this one will now seek to establish the extent of this variation within and between marama bean populations using the SSR markers identified here. The data obtained here also suggests that variation lies within populations rather than between populations (Figures 2 and 3) and this is in agreement with previous studies done using RAPD and AFLP markers (Naomab, 2004; Monaghan and Halloran, 1996). The microsatellite markers will be used in germplasm characterisation of interesting traits for marker assisted selection in the marama bean breeding program.

ACKNOWLEDGEMENTS

The work described in this study was supported by funding from the Kirk house Trust, United Kingdom (to PC) and by a World-Wide Learning Experience grant, a McGregor Fund Initiative in the College of Arts and Sciences to (CAC). Mr. E Nepolo is acknowledged for assistance with creating some of the microsatellite enriched libraries created on a visit to Case Western Reserve University on a travel and learning grant supported by the Kirk house Trust.

REFERENCES

- Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F, Kresovich S (1996). Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [Sorghum bicolor (L.) Moench]. Theor. Appl. Genet. 93: 190-198.
- Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR (1997). Development of a microsatellite framework map providing genome wide coverage in rice (*Oryza sativa* L.). Theor. Appl. Genet. 95: 553-567.
- Chimwamurombe P (2008). ABS and creation of an enabling environment for innovation, is it an issue for SADC countries? Marama bean domestication: an ABS case. Building Bridges 3: 5-7. characterization and cross-species amplification in *Phaseolus* sp. Crop Science 42:2128-2136.
- He C, Poysa V, Yu K (2003). Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among Lycopersicon esculentum cultivars. Theor. Appl. Genet. 106: 363-373.
- Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkins T (1997). Molecular tools in plant genetic resources conservation: A guide to technologies. IPGRI Technical Bulletin 47, IPGR, Rome.
- Keegan AB, van Staden J (1981). Marama Bean, Tylosema esculentum, a plant worthy of cultivation. S. Afr. J. Sci. 77:387-387.
- Karp A, Ketshajwang KK, Seberg O, Buiatti M (1996). Molecular techniques in the assessment of botanical diversity. Ann. Bot. 78: 143-149.
- Kutil BL, Williams CG (2001). Triplet-repeat microsatellites shared among hard and soft pines. J. Hered. 92:327-332.
- Coetzer LA, Ross JH (1977). Tylosema. Trees in South Africa 28: 77-80.
- Fergurson ME, Burow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, Mitchel S (2004). Microsatellite identification and characterization in peanut (*A. Hypogea* L.). Theor. Appl. Genet. 108:1064-1070.
- Gaitán-Solis E, Duquqe MC, Edwards KJ, Tohme J (2002). Microsatellite repeats in Common Bean (Phaseolus vulgaris): Isolation, among hard and soft pines. J. Hered. 92:327-332.

- Li CD, Rossnagel BG, Scoles GJ (2000, in press). The development of oat microsatellite markers and their use in identifying *Avena* species and oat cultivars. Theor. Appl. Genet.
- Moriguchi Y, Iwata H, Ujino-Ihara T, Yoshimura K, Taira H, Tsumura Y (2003). Development and characterization of microsatellite markers for Cryptomeria japonica D. Don. Theor. Appl. Genet. 106:751-758.
- Naomab E (2004). Assessment of genetic variation in natural populations of Marama bean (*Tylosema esculentum*) using molecular markers. A thesis submitted to the University of Namibia for the degree of Master of Science.
- Nepolo E, Takundwa M, Chimwamurombe PM, Cullis CA, Kunert K (2009). A review of geographical distribution of marama bean [*Tylosema esculentum* (Burchell) Schreiber] and genetic diversity in the Namibian germplasm. Afr. J. Biotech. 8(10):2088-2093.
- Novelli VM, Cristofani M, Souza AA, Machado MA (2006). Development and characterization of polymorphic microsatellite markers for the sweet orange (*Citrus sinensis* L. Osbeck). Genetics and Molecular Biology 29:90-96.
- Rossetto M, Melauchan A, Harris FCL, Henry RJ, Baverstock PR, Lee LS, Maguire TL, Edwards KJ (1999). Abundance and polymorphism of microsatellite markers in the tea tree (*Melaleuca alternifolia*, Myrtaceae). Theoret. Appl. Genet. 98: 1091-1098.
- Senior ML, Murphy JP, Goodman MM, Stuber CW (1998). Utility of SSRs for determining genetic similarities and relationships in maize using an agarose system. Crop Science 38:1088-1098.
- Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997). An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPs and pedigree. Theoret. Appl. Genet. 95:163-173.
- Tautz D (1993). Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ (eds) DNA fingerprinting: State of the Science. Birkhäuser Verlag, Basel pp. 21-28.
- Thulin M (1993). Fabacea (Leguminosae). In: Thulin M, ed. Flora of Somalia, vol 1. Pteridophyta, Gymnospermae, Angiospermae (Annonacea- Fabacea). Kew: Royal Botanic gardens pp. 360-361.
- Viruel MA, Hormaza JI (2004). Development, characterization and variability analysis of microsatellites in lychee (Litchi chinensis Sonn., Sapindaceae). Theor. Appl. Genet. 108: 896-902.
- Vos P, Hogers R, Bleeker M (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414.
- Wang Z, Weber JL, Zhang G, Tansksley SD (1994). Survey of plant short tandem DNA repeats. Theor. Appl. Genet. 88:1-6.
- Weising K, Winter P, Huttel B, Kahl G (1998). Microsatellite markers for molecular breeding. J. Crop Production 1:113-143.
- Xiao J, Li J, Yuan L, McCouch SR, Tanksley SD (1996). Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. Theor. Appl. Genet. 92:637-643.
- Zane L, Bargelloni L, Patarnello T (2002). Strategies for microsatellite Isolation: A review. Mol. Ecol. 11: 1-16.