ORIGINAL ARTICLE

Development of microsatellite markers from an enriched genomic library of pumpkin (*Cucurbita moschata* L.)

Nuchjaree Watcharawongpaiboon[1] and Julapark Chunwongse[2]

Abstract

Watcharawongpaiboon, N. and Chunwongse, J.

Development of microsatellite markers from an enriched genomic library of pumpkin (*Cucurbita moschata* L.)

Songklanakarin J. Sci. Technol., 2007, 29(5) : 1217-1223

Development of microsatellite markers in Pumpkin (*Cucurbita moschata* L.) was performed using the biotin-streptavidin enrichment procedure. One hundred and thirty three clones were randomly selected. After sequence analysis of 31 randomly picked positive colonies, 100% of the colonies were found to contain microsatellite sequences, and 9 primer pairs were designed. Five of the primers tested could amplify pumpkins DNA and can be used for genetic purity testing of the commercial hybrids. This paper reports the first isolation and utilization of microsatellite markers in pumpkin.

Key words : pumpkin, simple sequence repeats (SSR), hybridity testing

[1]Ph.D. Stydent in Agricultural Biotechnology, Center for Agricultural Biotechnology, [2]Ph.D. (Plant Breeding), Assist. Prof. Department of Horticulture, Faculty of Agriculture, Kasetsart University, Kamphaeng Saen, Nakhon Pathom, 73140 Thailand.

Corresponding E-mail: julapark.c@ku.ac.th

Received, 17 July 2006 Accepted, 1 November 2006

บทคัดย่อ

นุชจรี วัชรวงษ์ไพบูลย์[1] และ จุลภาค คุ้นวงศ์[2]

การพัฒนาดีเอ็นเอเครื่องหมายชนิดไมโครแซทเทลไลท์ในฟักทอง (*Cucurbita moschata* L.)

ว. สงขลานครินทร์ วทท. 2550 29(5) : 1217-1223

การพัฒนาดีเอ็นเอเครื่องหมายชนิดไมโครแซทเทลไลท์ในฟักทอง ได้ใช้วิธีการเพิ่มปริมาณชิ้นส่วนดีเอ็นเอด้วย biotin-streptavidin โดยสามารถคัดเลือกโคลนแบบสุ่มได้จำนวน 133 โคลน การหาลำดับนิวคลีโอไทด์จำนวน 31 โคลน พบไมโครแซทเทลไลท์ดีเอ็นเอ 100 % และสามารถออกแบบไพรเมอร์ได้จำนวน 9 คู่ และไพรเมอร์จำนวน 5 คู่ สามารถใช้สำหรับการตรวจสอบความบริสุทธิ์ทางพันธุกรรมในลูกผสมพันธุ์การค้าได้งานวิจัยครั้งนี้เป็นครั้งแรกที่มีรายงาน คือ การใช้ประโยชน์ของเครื่องหมายโมเลกุลแบบไมโครแซทเทลไลท์ในฟักทอง

[1]ศูนย์เทคโนโลยีชีวภาพเกษตร [2]ภาควิชาพืชสวน คณะเกษตร มหาวิทยาลัยเกษตรศาสตร์ อำเภอกำแพงแสน จังหวัดนครปฐม 73140

It is necessary to examine the genetic purity of commercial hybrid pumpkins before exporting or selling of the seeds. The test of seed genetic purity is to assure the uniformity and stability of the field performance and yield. Grow out test (GOT) for genetic purity of hybrid pumpkin has been traditionally conducted by the seed industry in Thailand and is based on morphological criteria that can be influenced by environmental conditions. Different environmental variation may cause variation in morphology. Some traits need to be identified during the reproductive stage, which delays the test. GOT is limited by the accuracy, time and labor cost. Additionally, the GOT requires a high degree of expertise for phenotypic identification

The uses of protein markers such as isozyme analysis are an established method that can be used for determination of genetic purity. However this method is limited due to many factors affecting protein expression including development of plant tissue and the environment. Low number of markers restricted polymorphism may also affect the utility of these markers (Walter *et al.*, 1989).

Random Amplified Polymorphic DNA (RAPD) is a useful method for hybridity testing although the limitations of this method are the reliability and reproducibility of the technique. If the reaction conditions change of such as the thermal cycle, *Taq* polymerase or DNA and primer concentrations the results could be varied. RAPD markers cannot be used for identification of heterozygous individuals from the homozygous ones (Buscher *et al.*, 1993).

Microsatellites or Simple Sequence Repeats (SSRs) are a class of repetitive sequences, which are abundant and distributed throughout all eukaryotic genomes (Hamanda and Kakunaga, 1982). They consist of a variable number of tandem repeats form 1-6 nucleotides. Short DNA repeats in tandem tend to be inaccurately replicated during the DNA synthesis, and generate new alleles with different number of repeating units. A variable number of repeats results from slipped strand mispairing of the DNA polymerase during DNA replication (Tautz *et al.*, 1986, Levinsen and Gutman, 1987). This length variation is the source of polymorphisms in SSR loci and can be identified by the polymerase chain reaction (PCR) (Saiki *et al.*, 1985) using primer pairs specific to the flanking SSR regions (Litt and Luty 1989, Weber and May 1989). These simple sequence length polymorphism (SSLP) markers are codominant, multiallelic and have a higher level of DNA polymorphism than other marker systems (Powell *et al.*, 1996). The SSRs are distributed throughout a genome, and have been proven to be useful as genetic markers and for cultivar identification (Cantini *et al.*, 2001, Wu and Tanksley, 1993).

In the present paper, we describe the development and application of five useful

microsatellite markers from a pumpkin species (*Cucurbita moschata* L.) for hybridity testing

Materials and Methods

Plant materials

Pumpkin (*Cucurbita moschata* L.) accession number CM147 from Tropical Vegetable Research Center (TVRC), Kasetsart University, Kamphaeng Saen, Thailand, was used for DNA extraction and the construction of enriched genomic libraries. Four pumpkin parental lines and the corresponding hybrids (PK1, PK2, PK3, and PK4) were kindly provided by Chia Tai Co.,Ltd, Seed Company, Thailand, and were used for hybridity testing.

Construction of the enriched genomic libraries

Plant genomic DNA extraction

Pumpkin genomic DNA accession CM147 was extracted from young fresh leaves of 7-days-old seedling using the micro preparation procedure described by Fulton *et al.* (1995) and DNA concentration was quantified by comparison to a DNA standard using the ethidium bromide stained agarose gel. The concentration was adjusted to be 100 ng/μl and 5 μl of the DNA was used for further digestion.

Digestion and ligation of linker to DNA fragments

Pumpkin genomic DNA was cut with 10 unit of *Tru* 9I (Promaga) for 1 hour at 65°C then, 100 pmol *Tru* 9I adapter (GAC GAT GAG TCC TGAG), 1 unit T4 DNA ligase (Promega), 10 mM ATP were added and incubated at 37°C for further 3 hours.

Amplification of DNA template

The ligation mixture was diluted 10 fold with dH_2O.and 5 ng of DNA and used for PCR amplification with 10 μM of the *Tru* 9I primer corresponding to the sequence of the *Tru* 9I adapter, in the 50 μl of PCR reaction containing the following: 1X PCR buffer, 10mM dNTP, 1.5 mM $MgCl_2$, 10 μM *Tru* 9I primer, and 1 unit of *Taq* polymerase. The profile of PCR is 94°C for 30 sec, 56°C for 1 minute and 72°C for 1 minute for a total of 20 cycles.

Capture of specific repetitive sequences with biotinylated oligonucleotide

Biotinylated (5') simple sequence repeat (SSR) oligonucleotides with sequence of B-$(CT)_{15}$ (0.5 μg) were used to capture DNA fragments containing repetitive sequences (50 μl). Streptavidin coated Dynal bead M280 (Dynal, Norway) (10 μg/μl) was added to capture the biotinylated oligonucleotide probes, which bind the SSR containing DNA fragments complementarily. The bound complex was drawn to the Magnetic Separation Stands (Promaga inc. USA), where DNA fragments were washed seven times and were cleaned through the Nucleo Spin® Extract (Macherey-Nagel) spin column. The eluate was used for PCR amplification with 10 μM of the *Tru* 9I primer. In the 50 μl of PCR reaction contained the solution mix as following: 1X PCR buffer, 10 mM dNTP, 1.5 mM $MgCl_2$, 10 μM *Tru* 9I primer and 1 unit of *Taq* polymerase. The profile of PCR was 94°C for 30 sec, 56°C for 1 minute, and 72°C for 1 minute for total of a 35 cycles. Finally, the PCR amplified products were used for the construction of the DNA library.

DNA Library construction

The enriched DNA fragments were ligated to the pGEM-T Easy Vector (Promega; USA) and transformed into the competent *E. coli* DH10B cells by electroporation. The transformants were selected on LB agar plates containing 0.1 g/ml amplicillin with 0.2 M X-Gal and 0.1 M IPTG. The resulting DNAs library were selected by blue-white colony screening. Single white colonies were randomly picked for culture.

DNA Library screening by dot blot hybridization

The individual clones were cultured in 3 ml LB liquid medium over night for plasmid DNA extraction using alkaline lysis (Sambrook *et al.*, 1989). DNA of each clone was dot blotted to the

Hybond-N+ nylon membrane (Amersham, UK) and hybridized with $(CT)_{10}$ repeat probe and detected on filter using BCIP/NBT Phosphate Substrate System (KPL; Kirkegard Perry Laboratories) for the positive clones containing the SSR sequences.

DNA sequencing of the positive clones

DNA of the SSR containing clones was submitted to the sequencing facility, Macrogen Sequencing Service, Republic of Korea, Which conducted the sequencing reaction based on Big Dye chemistry using M13F universal primer. The results were sent back and used for further analysis.

Primer design

DNA sequence regions flanking the SSR sequences were used for primer design using web service portals Gene Fisher <http://bibiserv.toofak./uni-biclefeld.de/genefisher/> or Primer 3 <http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi/> and also by manual selection with the criteria of 20-28 nucleotide primer length, Tm of 55-65°C, G + C content of 45-65%. Oligonucleotides were synthesized by the Bioservice Unit (BIOTEC, Thailand).

Primer testing and optimization

SSR primers were tested and optimized on pumpkin accession CM147 and the commercial hybrid (Sri Muang 016) by PCR amplification. PCR reaction was performed in a total volume of 20 µl using the reaction mixture containing 50 ng of DNA, 1X PCR buffer, 10 mM dNTP, 1.5-3 mM $MgCl_2$, 10 µM each of forward and reverse primers, and 0.4 unit of *Taq* polymerase. The profile for the amplification was 94°C for 1 minute, 45-55°C for 30 second, and 72°C for 1 minute for total of 35 cycles. The reactions were stopped by adding of 10 µl of sequencing stop dye (98% formamide, 0.025% Bromphenolblue, 0.025% Xylene Cyanol, 10 mM 0.5 EDTA pH8.0). The amplified products were separated in 4.5% denaturing polyacrylamide gel in 1 X TBE buffer. After a migration at 50 v. for 2 h., the DNA fragments were stained using silver nitrate protocol (Promega, USA). Sizes of the amplified fragments were estimated using *Phi*X174 /*Hinf* I marker. Allele scoring and sizing were done using the software Photo-Capt software version 99.01 (Vilber Laumat, France).

Hybridity testing

PCR of the four pumpkin parental lines and the corresponding hybrids using the optimized primer pairs were performed to identify the hybridity of each hybrid.

Results

Isolation of SSR markers in pumpkin

The microsatellite-enriched library was constructed using the enrichment procedure. One hundred and thirty three colonies were randomly picked from the primary transformation plates and 31 positive clones (23.31%) were sequence identified. All 31 clones were found to contain SSR sequences (GA repeats). Among these microsatellite-containing fragments, we could design 9 (29.03%) primer pairs from the 31 unique SSR fragments. The remaining 22 sequences showed the SSR sequence locating near the cloning site and were not be able to design the primers. Five out of nine primer pairs can clearly amplify pumpkin DNA (Figure 1). Repeat motif, primer sequences, expected allele size and annealing temperatures of the 5 microsatellite loci are given in Table 1.

Hybridity testing

Five pumpkin SSR primer pairs (PKCT 47, 62, 111, 122, and 133) were used for the polymorphism survey on four pumpkin hybrid sets (PK1, 2, 3, and 4). Polymorphisms could be detected in all sets with the PKCT122 primer and 3 sets of PK2, PK3 and PK4 by the primer PKCT111 (Figure 2). The PKCT62 can be used to detect the polymorphism of the set PK4 (data not shown).While the monomorphic could be detected in all sets with PKCT133 and PKCT47 (data not shown).

Table 1. Characteristic of pumpkin microsatellite markers; repeat motif, primer sequences, and expected size of allele. All same the optimum annealing temperature of 50°C.

Locus	Repeat motif	Primer sequence (5'-3')	Expected size (bp)
PKCT47	$(GA)_{21}$	F: GGTCCCAATAATAGCAACCAA R: GTGGGACACATCTTGAGCA	273
PKCT62	$(CT)_{23}$	F: GAAGTTCGTGGTCTGTGCAAGTC R: CCTGAGTAACCTCCGTGCTTCC	300
PKCT111	$(GA)_{23}$	F: GTTGCAGCGACCGTTCTTCTTC R: GCATCTGAAGACGATGCGTCGT	189
PKCT122	$(CT)_{20}$	F: CTAAACAGGATGCCTCTGACAC R: CGGGATTTCCGAAACAACGT	211
PKCT133	$(GA)_{12}$	F: TCGGAATCGTCTTCAGCAATAGTC R: TCCTCTTCCATTCCACTTTCTCCT	238

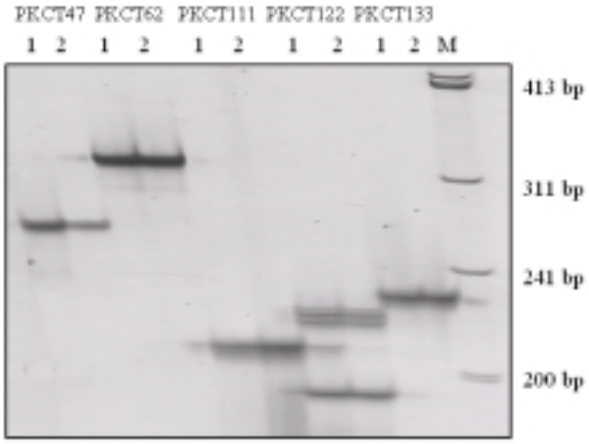

Figure 1. The DNA pattern of PCR amplified products using primer PKCT47, PKCT62, PKCT111, PKCT122, and PKCT133 respectively. Lane 1 is CM147 DNA, lane 2 is Srimuang 016 DNA and lane M is a molecular weight marker DNA.

Discussion

Microsatellite marker has become one of the most widely used molecular markers for genetic analysis in recent years. The enrichment of DNA fragments through the binding of the SSR probe is simple and efficient approach for SSR isolation and has been successfully applied to a number of plant genomes (Edwards *et al.*, 1996; Patricia *et al.*, 2004). By employing these procedures from other plant species, we have successfully applied the procedures to isolate pumpkin microsatellite markers.

Traditionally, microsatellite loci have been isolated from partially digested genomic libraries of small insert size. Several thousand clones need to be screened through colony/plaque hybridization using repetitive DNA probes (Panaud *et al.*, 1995; Broun and Tanskley 1996). Nevertheless, this can be extremely tedious and inefficient for plant species with low microsatellite frequencies (0.059-5.8%) (Zane *et al.*, 2002). In contrast, the enrichment procedure is faster and more cost effective than the library screening approach.

The enrichment efficiency of this experiment was found to be extremely high (100%),

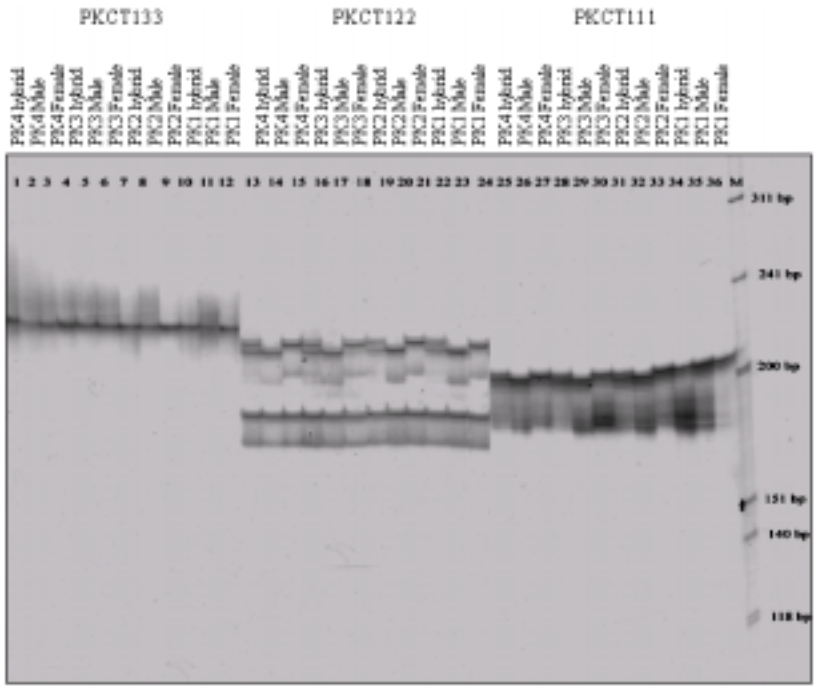

Figure 2. The DNA pattern of PCR amplified products using primer PKCT133 (lane 1-12), PKCT122 (lane 13-24), and PKCT111 (lane 25-36) in four hybrid pumpkin sets, PK4, PK 3, PK 2, and PK1 respectively (hybrid, male, and female for each genotype respectively). Lane M is a molecular weight marker DNA.

comparable to the enriched DNA libraries of the other plant species where the efficiency were between 50% and 90% (Butcher *et al.*, 2000), thus considerably reduced the cost of the sequence identification. However, the majority of the SSR containing clones (70.97%) were found to be at the proximal end, adjacent to the cloning site, hence it was not possible to design flanking primers. Changing the enzymes used in digestion of the genomic DNA might provide the better position of SSR in the clones than the *Tru* 9I used in this experiment.

From five primer pairs that detected the polymorphism, two highly polymorphic pairs were identified and can be used in hybridity testing of commercial lines. These primer pairs will be useful in not only hybridity testing but can also in verification of genetic stock. With slight modification of the protocol, several more markers can be generated and used in extensive genetic studies. This paper is the first report on isolation and utilization of microsatellite markers in pumpkin.

Acknowledgements

The authors gratefully acknowledge the Center for Agricultural Biotechnology, Kasetsart University, for a Ph.D. grant and Kasetsart University Research and Development Institute for a research grant. The Tropical Vegetable Research Center (TVRC), Kasetsart University, and Chia Tai Co., LTD in Thailand provided plant materials.

References

Broun, P. and. Tanksley, S.D. 1996. Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol. Gen. Genet., 250 : 39-49.

Buscher, N., Zyprian, E. and Blaich, R. 1993. Identification of grapevine cultivars by DNA analysis

Pitfalls of random amplified polymorphic DNA techniques using 10 mer primers. Vitis, 32 : 187-188.

Butcher, P.A., Decroocq, S., Gray, Y. and Moran, G.F. 2000. Development, inheritance and cross-species amplification of microsatellite markers from *Acacia mangium*. Theor. Appl. Genet., 101 : 282-1290.

Cantini, C., Lezzoni, A.F., Boritzki, M. and Struss, D. 2001. DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. J. Am. Soc. Hortic. Sci., 126 : 205-209.

Edwards, K.J., Barker, J.H.A., Daly, A., Jones, C. and Karp, A. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques, 20 : 758-760.

Fulton, T.M., Chunwongse, J. and Tanksley, S.D. 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol. Biol. Rep., 13 : 207-209.

Hamanda, H. and Kakunaga T. 1982. Potential Z-DNA forming sequences are highly dispersed in the human genome. Nature, 298 : 396-398.

Levinsen, G. and Gutman, G.A. 1987. Slipped-strand mispairing: a major mechanism for DNA evolution. Mol. Biol. Evol., 4 : 203-221.

Litt, M. and Luty, J.A. 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle acting gene. Am. J. Hum. Genet., 44 : 397-401.

Panaud, O., Chen, X. and McCouch, S. 1995. Frequency of microsatellite sequences in rice *Oryza sativa* L. Genome, 38 : 1170-1176.

Patricia S.R., Tuli., C.L.L., Rodrigo, L.T., Glaucia, S.C.B., Jose, A.B. and Marcio, E. 2004. Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). BMC Plant Biology: http://www.biomedcentral.com/1471-2229/4/9 [access date July 14, 2004]

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed, 2 : 225-238.

Saiki, R.K., Scharf, S., Faloona, F.A., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. 1985. Enzymatic amplification of f1-globin sequences and restriction site analysis for diagnosis of sickle cell anemia. Science, 230 : 1350-1354.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press, New York.

Tautz, D., Trick, M. and Dover, G. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature, 322 : 652-653.

Walter, T.W., Posluszny, U. and Kevan, P.G. 1989. Isoenzyme analysis of the grape (Vitis). I.A practical solution. Can. J. Bot., 67 : 2894-2899.

Weber, J. and May, P.E. 1989. Abundant class of human DNA polymorphism, which can be typed using the polymerase chain reaction. Am. J. Genet., 44 : 388-396.

Wu, K.S., and Tanksley, S.D. 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol. Gen. Genet., 241 : 225-235.

Zane, L., Bargelloni, L. and Patarnello, T. 2002. Strategies for microsatellite isolation: A review. Mol. Ecol., 11 : 1-16.