January 2017. Horticultural Plant Journal, 3 (1): 13–16.

Horticultural Plant Journal

Available online at www.sciencedirect.com
The journal's homepage: http://www.journals.elsevier.com/horticultural-plant-journal

Development of Insertion and Deletion Markers for Bottle Gourd Based on Restriction Site-associated DNA Sequencing Data

WU Xinyi, XU Pei, WU Xiaohua, WANG Baogen, LU Zhongfu, and LI Guojing *

Institute of Vegetables, State Key Lab Breeding Base for Sustainable Control of Plant Pest and Disease, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China

Received 28 October 2016; Received in revised form 20 December 2016; Accepted 4 January 2017

Available online 12 June 2017

A B S T R A C T

Bottle gourd is an important cucurbit crop worldwide. To provide more available molecular markers for this crop, a bioinformatic approach was employed to develop insertion–deletions (InDels) markers in bottle gourd based on restriction site-associated DNA sequencing (RAD-Seq) data. A total of 892 Indels were predicted, with the length varying from 1 bp to 167 bp. Single-nucleotide InDels were the predominant types of InDels. To validate these InDels, PCR primers were designed from 162 loci where InDels longer than 2 bp were predicated. A total of 112 InDels were found to be polymorphic among 9 bottle gourd accessions under investigation. The rate of prediction accuracy was thus at a high level of 72.7%. DNA fingerprinting for 4 cultivars were performed using 8 selected Indels markers, demonstrating the usefulness of these markers.

Keywords: bottle gourd; InDels markers; RAD-Seq; DNA fingerprinting

1. Introduction

Bottle gourd or calabash [*Lagenaria siceraria* (Mol.) Standl.] (2n = 2x = 22), also known as opo squash or long melon, is diploid belonging to the genus *Lagenaria* of the Cucurbitaceae family with a genome size of ~334 Mb (Beevy and Kuriachan, 1996; Achigan-Dako et al., 2008). Bottle gourd is believed to originate from Africa and independently domesticated in Africa and Asia 10 000 years ago (Erickson et al., 2005). Today, bottle gourd has a wide range of distribution and application all over the tropics (Heiser, 1979, 1989; Erickson et al., 2005). Fresh fruit of bottle gourd is a common vegetable in many regions of Asia and Africa (Morimoto and Mvere, 2004); dry fruits can be used as containers, pipes, floats, music instruments, for medicine, artistic endeavors, etc (Heiser, 1979). Bottle gourd is also widely used as the rootstock for watermelon to defend soil-borne disease and low soil temperature (Lee, 1994; Yetisir and Sari, 2003).

Molecular markers are indispensible for genetic research (McCouch et al., 1997; Joshi et al., 2001; Nagaraju et al., 2002; Ni et al., 2002). Popular classical DNA markers include restriction fragment length polymorphism markers (RFLP; Botstein et al., 1980), amplified fragment length polymorphism markers (AFLP; Vos et al., 1995) and simple sequence repeat markers (SSR; Tautz, 1989). With the advances of high-throughput sequencing technology, new generation of DNA markers has been emerging. Among them, single-nucleotide polymorphisms (SNPs) and insertions–deletions polymorphisms (InDels) markers are the most promising in that they are reliable, genome-distributed and highly abundant. Unlike SNPs, for which genotyping still needs special equipment and the cost is high, InDels marker analysis can be performed with ordinary PCR thermocyclers and agarose gel electrophoresis systems. By far InDels marker has been developed and successfully applied in wheat, rice, *Arabidopsis*, *Brassica rapa* and many other agricultural crops (Hayashi et al., 2006; Raman et al., 2006; Hou et al., 2010; Liu et al., 2013).

DNA marker resources for bottle gourd are still limited, until now most of markers published in limited studies are random amplified polymorphic DNA (RAPD) markers and inter simple

* Corresponding author. Tel.: +86 571 86419189
E-mail address: ligj@mail.zaas.ac.cn

Peer review under responsibility of Chinese Society for Horticultural Science (CSHS) and Institute of Vegetables and Flowers (IVF), Chinese Academy of Agricultural Sciences (CAAS)

http://dx.doi.org/10.1016/j.hpj.2017.01.011

sequence repeats (ISSR) markers (Decker-Walters et al., 2001; Morimoto et al., 2006; Bhawna et al., 2014; Srivastava et al., 2014). In our laboratory, Xu et al. (2011) developed 400 SSR markers through partial sequencing of the bottle gourd genome. Recently, RAD-Seq technology was applied to a bottle gourd F_2 population for developing SNP markers (Xu et al., 2014). RAD-Seq is one of the next-generation sequencing (NGS) method that sequencing the DNA only flanking specific restriction enzyme sites to produce a 'complexity reduction' genome resources, which ligated an adapter containing multiplex identifiers (MIDs) in the reduced-representation libraries (RRLs) (Miller et al., 2007). This set of RAD-Seq data also provides an opportunity for developing InDels markers by mining sequence variations between the parental lines. The objective of this study was to identify InDels polymorphisms between the female parent 'Hangzhou Gourd' and the male parent 'J129' to develop InDels markers. To our knowledge, this is the first report of InDels marker development in bottle gourd. These markers will be valuable for genetic research and breeding of bottle gourd in future works.

2. Materials and methods

2.1. *RAD-Seq data source*

The RAD-Seq data of two bottle gourd cultivars 'Hangzhou Gourd' and 'J129' previously generated in the authors' lab (Xu et al., 2014) were used in this study. 'Hangzhou Gourd' is a landrace collected from Southern China, and 'J129' is also a landrace collected from Northeast China. These two cultivars were sequenced using Illumina HiSeq 2000 (Majorbio Pharm Technology Co., Ltd, Shanghai) in 2013. The sequencing volume is 2 105 562 reads for 'Hangzhou Gourd' and 3 038 816 reads for 'J129'.

2.2. *Bioinformatic methods*

Raw sequence reads of 'Hangzhou Gourd' and 'J129' were trimmed to 85 nucleotides from the 3′ end to ensure more than 90% of the nucleotides have a quality value above Q30 (equals 0.1% sequencing error) and more than 99% above Q20 (equals 1% sequencing error). Reads of low quality, including those shorter than 85 bp after trimming or with ambiguous barcodes, were discarded (Xu et al., 2014). To detect InDels polymorphisms, a four-step procedure was used: (1) The sequences were clustered based on *Eco*R I recognition sequence at one side of each read using *Ustacks* under default parameters (Catchen et al., 2011, Stacks package, http://catchenlab.life.illinois.edu/stacks/). Sequences from different samples were discriminated by barcodes; (2) The stacks of 'Hangzhou Gourd' and 'J129' were compared in a pairwise way allowing no gap using *Cstacks*. Only 1 bp mismatch was allowed in this step; (3) The other end (containing no *Eco*R I recognition sequence) of reads were clustered based on the above stacking results, then the reads from each cluster of each sample were assembled by *Phred* under default parameters (Ewing and Green, 1998); (4) The contigs generated from each sample were aligned by blat (Kent, 2002; https://genome.ucsc.edu/FAQ/FAQblat.html), and the putative InDels loci were identified based on the blat results using an in-house PERL script. The criteria for blat results filtering were: gap number = 1; maximum mismatch = 1 bp.

Table 1 The primer sequences of polymorphic InDels markers developed to construct DNA fingerprints

Marker name	InDels size	Forward primer	Reverse primer	T_m/°C
BID039	4	ttaacttttccacaagcatttt	ggaaggagtacctcatccataa	57
BID046	4	caactccacagatagttgaaca	tctttgtgggtctttctcaat	56
BID052	4	ccataccaaaatacaaccaaga	tggttagttgatggtgtgaat	55
BID089	9	cgcttctggtttataggtttac	gctaaaccaatcaaaacctaag	55
BID096	10	ccaactcgacattttgattc	ggccacacactttcttttatg	57
BID105	12	gaagctcaagaaaatgaaaatc	tgaaacgcgaagaataagaa	56
BID122	17	aagttcaataaccgaaagaaaa	cgaaaaccaaatggttacaaat	58
BID137	28	atgtaagtccgcccgtg	ttcaattttagtgacattgggt	56

2.3. *InDels marker development*

Primer 3 (Untergasser et al., 2012) was used to design PCR primers based on the flanking sequences of predicted InDels loci. The length of primers was set to 17–22 bp. The annealing temperature (T_m) was set to between 55 °C and 60 °C.

2.4. *Marker validation*

InDels marker validation was performed on 9 bottle gourd accessions, which were 'Anji Gourd', 'Zhepu 2', 'Zhepu 8', 'J129', 'Hangzhou Gourd', 'Puxian Gourd', 'Ningbo Yekaihua', 'Nanxiu Gourd', and 'Niutuipu'. Genomic DNA was extracted from young leaves of two-week old seedlings using the DNA extraction kit (TIANGEN Co. Ltd, Beijing), following the manufacturer's instructions. Polymerase chain reaction (PCR) was performed in a volume of 12.5 μL containing 10–20 ng of template, 2.5 pmol of each of the primers, 2.5 nmol of each of the dNTPs, 18.6 nmol $MgCl_2$, 0.1 U *Taq* DNA polymerase and 1 × PCR buffer. The PCR profile included 1 cycle of 94 °C 3 min, 35 cycles of 94 °C 30 s, 55–60 °C (depending on the primers) 40 s and 72 °C 40 s, and a final extension at 72 °C for 5 min. The PCR products were separated in 8% non-denaturing polyacrylamide gels (Acr:Bis = 19:1) at room temperature with 1 × TBE buffer and visualized by silver staining (Bassam et al., 1991).

2.5. *DNA fingerprinting with InDels markers*

Eight InDels markers amplifying clear bands (BID039, BID046, BID052, BID089, BID096, BID105, BID122, BID137; Table 1) were selected to construct DNA fingerprints for 4 bottle gourd cultivars, which are 'PS79-1', 'Peel Round Gourd', 'Nanxiu Gourd' and 'Shaoxing Long Melon'. Each sample was genotyped with 8 InDels markers and the PCR products were resolved on polyacrylamide gels.

3. Results and discussion

3.1. *InDel detection*

Using the current bioinformatic approach, 892 putative InDels were identified between the genomes of 'Hangzhou Gourd' and 'J129'. The length of the insertion–deletion of these InDels ranged from 1 bp to 167 bp and showed a remarkable skewed distribution. Overall, these InDels fell into 45 types according to InDel length, with 68.3% being single-nucleotide, 13.6% two-nucleotide, 3.5% three-nucleotide and 2.5% four-nucleotide InDels (Fig. 1). Thirty-four types of InDels had less than 5 members, and 19 types with large insertion–deletion

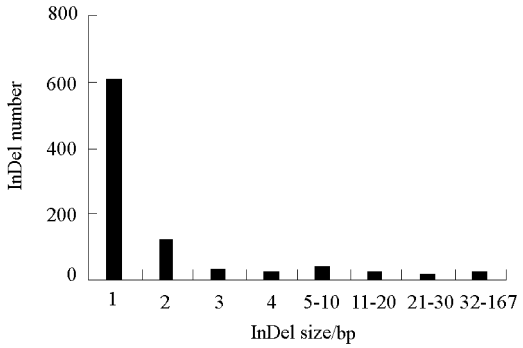

Fig. 1 The number and distribution frequency of InDels identified between 'Hangzhou Gourd' and 'J129'

(>30 bp) have only 1 individual. Compared to the case in *B. rapa* and maize where InDels detection is based on genome re-sequencing data (Liu et al., 2013, 2015), the InDels distribution pattern here is similar, suggesting that small insertion–deletion variations are more common than larger ones in plant genomes.

3.2. *Experimental validation of InDels polymorphisms*

To validate the fidelity of the predicted InDels, we designed PCR primer pairs for all the 162 putative InDels with the insertion–deletion size longer than 2 bp. PCR was performed for 9 bottle gourd accessions including 'Hangzhou Gourd' and 'J129', the InDels discovery parents. Eight primer pairs failed to amplify in each of the nine accessions, and the rest all amplified clear bands in all 9 cultivars, giving a technical success rate of 95%. Of the 154 primer pairs with successful amplification, 42 primer pairs (27.3%) gave monomorphic results and 112 primer pairs (72.7%) revealed identifiable polymorphisms among the 9 accessions (Fig. 2). The primer pairs with monomorphic bands between 'Hangzhou Gourd' and 'J129' demonstrated that they were not truly derived from InDels regions. We also noted that all the 20 predicted InDels with the insertion–deletion size > 33 bp fell into this group, indicating that prediction of larger InDels is less accurate. Among the primer pairs showing polymorphisms, the total allele number in the 9 accessions detected per primer pair ranged from 2 to 5 (Fig. 2). More than half (62 markers) of the primer pairs amplified bi-allelic products ('J129' and 'Hangzhou Gourd' allele, respectively) and about one-third (40 markers) generated tri-allelic products (plus either the heterozygous allele of 'J129' and 'Hangzhou Gourd' or a novel allele). The 112 polymorphic InDels markers therefore will be useful for genetic research and molecular marker assisted breeding in bottle gourd.

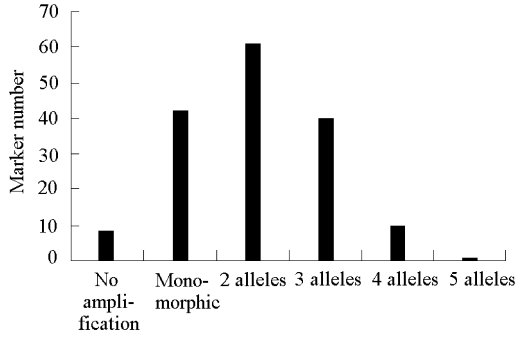

Fig. 2 The amplification results of the 162 primer pairs in the 9 accessions

3.3. *Fingerprinting of 4 cultivars using the InDels markers*

As an example showing the application of the InDels markers, DNA fingerprints of 4 cultivars were constructed using 8 selected InDels markers. These 8 InDels markers were selected according to their amplification performances in the nine accessions, each being able to reveal at least 3 allelic polymorphisms. As shown in Fig. 3, each cultivar was readily distinguishable from the others based on at least one diagnostic marker. For example, 'PS79-1' and 'Peel Round Gourd' were clearly distinguished by the polymorphic bands amplified by BID052; 'Nanxiu Gourd' had a unique low-molecular weight band amplified by BID096, and 'Shaoxing Long Melon' had two unique bands amplified by BID046 and BID137, respectively. In addition, compared with SSR markers in bottle gourd (Lu et al., 2012), these InDels markers amplified less and more specific bands, rendering them more reliable for

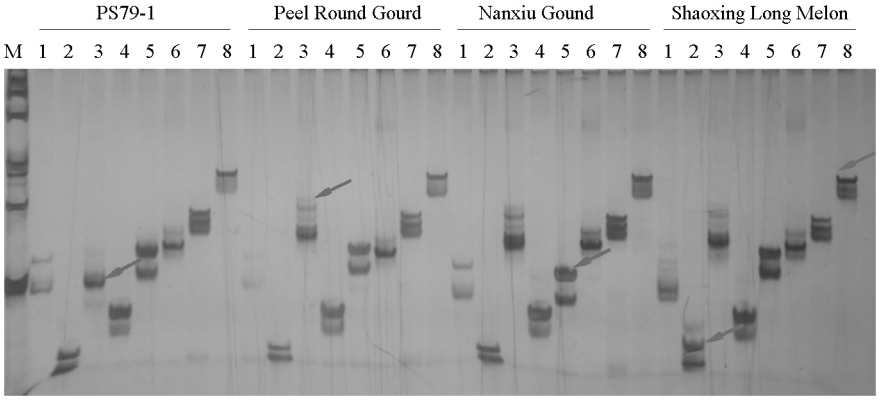

Fig. 3 DNA fingerprinting of 4 cultivars

Lanes 1–8 indicate the bands of BID039, BID046, BID052, BID089, BID096, BID105, BID122 and BID137. The first lane on the left is marker DL 500 (M). The arrows denote some unique bands.

germplasm identification. Indeed, the marker BID046 has been successfully used for hybrid seed purity inspection in the author's lab (data not shown).

Acknowledgments

This study was supported by the National Natural Science Foundation of China (31401880), China Postdoctoral Science Foundation Funded Project (2015M571900), the Natural Science Foundation of Zhejiang Province (LY14C150004), Public Project of Zhejiang Province (2015C32042) and grants from Zhejiang Academy of Agricultural Sciences (2016R23R08E04).

REFERENCES

Achigan-Dako, E.G., Fuchs, J., Ahanchede, A., Blattner, F.R., 2008. Flow cytometric analysis in *Lagenaria siceraria* (Cucurbitaceae) indicates correlation of genome size with usage types and growing elevation. Plant Syst Evol, 276: 9–19.

Bassam, B.J., Gaetano-Anollé, G., Gresshoff, P.M., 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem, 196: 80–83.

Beevy, S.S., Kuriachan, P., 1996. Chromosome numbers of south Indian Cucurbitaceae and a note on the cytological evolution in the family. J Cytol Genet, 31: 65–71.

Bhawna Abdin, M.Z., Arya, L., Saha, D., Sureja, A.K., Pandey, C., Verma, M., 2014. Population structure and genetic diversity in bottle gourd [*Lagenaria siceraria* (Mol.) Standl.] germplasm from India assessed by ISSR markers. Plant Syst Evol, 300: 767–773.

Botstein, D., White, R., Skolnick, M.H., Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet, 32: 314–331.

Catchen, J.M., Amores, A., Hohenlohe, P.A., Cresko, W.A., Postlethwait, J.H., 2011. Stacks: Building and genotyping loci de novo from short-read sequences. G3 (Bethesda), 1: 171–182.

Decker-Walters, D.S., Staub, J.E., Lopezsese, A.I., Nakata, E., 2001. Diversity in landraces and cultivars of bottle gourd (*Lagenaria siceraria*; Cucurbitaceae) as assessed by random amplified polymorphic DNA. Genet Resour Crop Evol, 48: 369–380.

Erickson, L.D., Smith, B.D., Clarke, A.C., Sandweiss, D.H., Tuross, N., 2005. An Asian origin for a 10,000-year-old domesticated plant in the Americas. Proc Natl Acad Sci USA, 102: 18315–18320.

Ewing, B., Green, P., 1998. Basecalling of automated sequencer traces using phred. II. Error probabilities. Genome Res, 8: 186–194.

Hayashi, K., Yoshida, H., Ashikawa, I., 2006. Development of PCR-based allele-specific and InDel marker sets for nine rice blast resistance genes. Theor Appl Genet, 113: 251–260.

Heiser, C.B., 1979. The Gourd Book. University of Oklahoma Press, Norman.

Heiser Jr., C.B., 1989. The evolution of plant exploitation, in: Harris, D.R., Hillman, G.C. (Eds.), Foraging and Farming. Unwin Hyman, London, pp. 471–480.

Hou, X.H., Li, L.C., Peng, Z.Y., Wei, B.Y., Tang, S.Q., Ding, M.Y., Liu, J.J., Zhang, F.X., Zhao, Y.D., Gu, H.Y., Qu, L.Y., 2010. A platform of high-density INDEL/CAPS markers for map-based cloning in *Arabidopsis*. Plant J, 63: 880–888.

Joshi, S.P., Bhave, S.G., Chowdari, K.V., Apte, G.S., Dhonukshe, B.L., Lalitha, K., Ranjekar, P.K., Gupta, V.S., 2001. Use of DNA markers in prediction of hybrid performance and heterosis for a three-line hybrid system in rice. Biochem Genet, 39: 179–200.

Kent, W.J., 2002. BLAT — The BLAST-like alignment tool. Genome Res, 12: 656–664.

Lee, J.M., 1994. Cultivation of grafted vegetables I. current status, grafting methods and benefits. Hortscience, 29: 235–239.

Liu, B., Wang, Y., Zhai, W., Deng, J., Wang, H., Cui, Y., Cheng, F., Wang, X.W., Wu, J., 2013. Development of InDel markers for *Brassica rapa* based on whole-genome re-sequencing. Theor Appl Genet, 126: 231–239.

Liu, J., Qu, J.T., Yang, C., Tang, D.G., Li, J.W., Lan, H., Rong, T.Z., 2015. Development of genome-wide insertion and deletion markers for maize, based on next-generation sequencing data. BMC Genomics, 16: 601.

Lu, Z.F., Xu, P., Wu, X.H., Wang, B.G., Wang, S., Liu, Y.H., Li, G.J., 2012. A rapid and precise SSR-based procedure for bottle gourd seed quality survey. Acta Agric Zhejiangensis, 24: 578–581. (in Chinese)

McCouch, S.R., Chen, X.L., Panaud, O., Temnykh, S., Xu, Y.B., Cho, Y.G., Huang, N., Ishii, T., Blair, M.W., 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol, 35: 89–99.

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res, 17: 240–248.

Morimoto, Y., Maundu, P., Kawase, M., Fujimaki, H., Morishima, H., 2006. RAPD polymorphism of the white-flowered gourd [*Lagenaria siceraria* (Molina) Standl.] landraces and its wild relatives in Kenya. Genet Resour Crop Evol, 53: 963–974.

Morimoto, Y., Mvere, B., 2004. *Lagenaria siceraria*, in: Grubben, G.J.H., Denton, O.A. (Eds.), Vegetables. Plant Resources of Tropical Africa 2. Backhuys Publishers/CTA, Wageningen/Leiden, pp. 353–358.

Nagaraju, J., Kathirvel, M., Kumar, R., Siddiq, E.A., Hasnain, S.E., 2002. Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc Natl Acad Sci USA, 99: 5836–5841.

Ni, J.J., Colowit, M.P., Mackill, D.J., 2002. Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci, 42: 601–607.

Raman, H., Raman, R., Wood, R., Martin, P., 2006. Repetitive indel markers within the *ALMT1* gene conditioning aluminium tolerance in wheat (*Triticum aestivum* L.). Mol Breed, 18: 171–183.

Srivastava, D., Khan, N.A., Shamim, M., Yadav, P., Pandey, P., Singh, K.N., 2014. Assessment of the genetic diversity in bottle gourd [*Lagenaria siceraria* (Molina) Standl.] genotypes using SDS-PAGE and RAPD markers. Natl Acad Sci Lett, 37: 155–161.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res, 17: 6463–6471.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer 3 — New capabilities and interfaces. Nucleic Acids Res, 40: e115.

Vos, P., Hogers, R.C.J., Bleeker, M., Reijans, M., Lee, T.V.D., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M., Zabeau, M., 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res, 23: 4407–4414.

Xu, P., Wu, X.H., Luo, J., Wang, B.G., Liu, Y.H., Ehlers, J.D., Wang, S., Lu, Z.F., Li, G.J., 2011. Partial sequencing of the bottle gourd genome reveals markers useful for phylogenetic analysis and breeding. BMC Genomics, 12: 467.

Xu, P., Xu, S.Z., Wu, X.H., Tao, Y., Wang, B.G., Wang, S., Qin, D.H., Lu, Z.F., Li, G.J., 2014. Population genomic analyses from low-coverage RAD-Seq data: A case study on the non-model cucurbit bottle gourd. Plant J, 77: 430–442.

Yetisir, H., Sari, N., 2003. Effect of different rootstock on plant growth, yield and quality of watermelon. Aust J Exp Agric, 43: 1269–1274.