Short Note: Internal Size Standard for Microsatellite Genotyping

By C. H. NG*), S. L. LEE and K. K. S. NG

Forest Research Institute Malaysia, 52109 Kepong, Selangor Darul Ehsan, Malaysia

(Received 17th March 2008)

Abstract

The in-house production of fluorescently labelled internal size standard offers the advantage of cost saving over the commercial size standard in microsatellite genotyping. Based on the reported in-house internal size standard protocol, we have improved the method by generating 21 DNA fragments (in a standard named as HM-400) with each size similar to that of the commercial size standard. The consistent amplification of the correct fragment size was optimised via primer modulation for non-templated nucleotide addition by *Taq* DNA polymerase. A total of six microsatellite loci were used to assess the accuracy of HM-400 and the mean standard deviation of the size data was 0.19. The differences between the fragment size means for samples sized using HM-400 and commercial size standard were small with an average of 0.29 bp. The production cost of HM-400 was only 10% of the cost of commercial size standard.

Key words: size standard; microsatellites; genotyping; population genetics; dipterocarps.

Microsatellites have become popular DNA markers in population genetics and are being used extensively to generate genetic information for the conservation of dipterocarps in tropical forests (NG et al., 2004, 2006). However, the commercial size standard constitutes a substantial portion of the cost in microsatellite genotyping conducted in our laboratory. In order to reduce the cost in automated genotyping, various reports have described novel ways to synthesize desirable internal size standards using a DNA template of known sequence (BRONDANI and GRATTAPAGLIA, 2001; DEWOODY et al., 2004; SYMONDS and LLOYD, 2004). In this short note, we improved the density of the fragment size below 400 bp by generating a 21-fragment size standard comparable with the commercial size standard. The amplification of the correct fragment size was tested by using T4 DNA polymerase, *Pfu* DNA polymerase and "PIGtailing" (BROWNSTEIN et al., 1996). The consistency of amplified fragments was determined by sizing six microsatellite loci amplified from *Neobalanocarpus heimii* and *Koompassia malaccensis*.

The plasmid pGEM®-3Zf(+) (hereafter called pGEM) was used to design a collective forward primer (fluorescently labelled with ROX) and 21 corresponding reverse primers to produce 21 amplification fragments within the size range of 50–400 bp. Ten nanograms of pGEM were used in 10-µL PCR reactions which consisted of 0.4 µM of each forward and reverse primers, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 unit of *Taq* DNA polymerase (Promega Corporation) and 1 X buffer. The reaction mixture was subjected to amplification using a GeneAmp PCR System 9700 (Applied Biosystems), for an initial denaturation step of 5 min at 94 °C, 35 cycles of 94 °C for 30 s, 56 °C for 30 s and 72 °C for 1 min. A final cycle of 60 min at 72 °C was used to complete the extension of any remaining products before holding the samples at 4 °C until analysed. All the PCR products were amplified individually and subsequently combined and electrophoresed on ABI PRISM 377 DNA Sequencer. Fragment sizes were determined against the commercial size standard GeneScan™ 400HD (GS-400) using GeneScan and Genotyper software (Applied Biosystems). The "home-made" peak can be differentiated from the commercial size standard by its higher intensity.

The amplified fragments consisted a mixture of true allele (T) and products sized one nucleotide greater than the true allele (T+1). The (T+1) is commonly the result of the non-templated addition of a single nucleotide to the 3'-end of the PCR product by *Taq* DNA polymerase (CLARK, 1988). The degree to which a fragment is subjected to (T+1) is relatively marker specific. For example, fragments 100 bp, 120 bp, 160 bp, 200 bp, 220 bp and 260 bp were always sized as (T+1) (data not shown). The (T+1) fragments were tried to be trimmed by using T4 DNA polymerase. The removal of 3'-overhangs in (T+1) fragment treated with T4 DNA polymerase was not efficient as many fragments with two or three nucleotide less were produced. We have also tried to produce true fragments by using *Pfu* DNA polymerase. However, there was a mixture of unspecific amplification fragments which might due to fragment degradation by 3'→5' exonuclease activity associated with proofreading function in *Pfu* DNA polymerase. Finally, the correct sizing of fragments was generated by placing the "PIG-tailing" sequence GTTTCTT on the 5'-end of reverse primers (*Table 1*). The primers were designed in a way that the products were one base pair less than the expected fragment because the PIG-tailing" sequence strongly favours adenylation (BROWNSTEIN et al., 1996). After optimizing the correct sizes of the 21 fragments, these size standards were mixed in proportional ratios according to relative fluorescence intensity in order to obtain approximate even intensity in most of the fragments (*Fig. 1*). The combined volume of the 21 fragments was ethanol precipitated and re-suspended in 10-µL TE buffer and named as home-made internal size standard (HM-400). A total of 0.25-µL HM-400 was used in each subsequent microsatellite genotyping.

*) Communicating author: CHIN HONG NG. Telephone: 603-62797622. Fax: 603-62804614. E-mail address: chinhong@frim.gov.my

DOI:10.1515/sg-2009-0012
edited by Thünen Institute of Forest Genetics

Table 1. – Primer sequences for HM-400 internal size standard.

Primer name	Primer sequence	Product size (bp)
GEMF1	5'-ROX-GGT GCC TAA TGA GTG AGC-3'	
GEM50	5'-GTTTCTT GCG CAA CGC AAT TAA TGT-3'	50
GEM60	5'-GTTTCTT CGG GCA GTG AGC GCA ACG-3'	60
GEM90	5'-GTTTCTT GCT GGC ACG ACA GGT TTC-3'	90
GEM100	5'-GTTTCTT CAT TAA TGC AGC TGG CAC-3'	100
GEM120	5'-GTTTCTT TCC CCG CGC GTT GGC CGA-3'	120
GEM150	5'-GTTTCTT CGG AAG AGC GCC CAA TAC-3'	150
GEM160	5'-GTTTCTT AGC GAG GAA GCG GAA GAG-3'	160
GEM180	5'-GTTTCTT CCG AGC GCA GCG AGT CAG-3'	180
GEM190	5'-GTTTCTT AGC CGA ACG ACC GAG CGC-3'	190
GEM200	5'-GTTTCTT CGC TCG CCG CAG CCG AAC-3'	200
GEM220	5'-GTTTCTT GCC TTT GAG TGA GCT GAT-3'	220
GEM240	5'-GTTTCTT CTG TGG ATA ACC GTA TTA-3'	240
GEM260	5'-GTTTCTT TCC TGC GTT ATC CCC TGA-3'	260
GEM280	5'-GTTTCTT CCT TTT GCT CAC ATG TTC-3'	280
GEM290	5'-GTTTCTT CTT TTG CTG GCC TTT TGC-3'	290
GEM300	5'-GTTTCTT GGT TCC TGG CCT TTT GCT-3'	300
GEM320	5'-GTTTCTT CAG CAA CGC GGC CTT TTT-3'	320
GEM340	5'-GTTTCTT CGG AGC CTA TGG AAA AAC-3'	340
GEM360	5'-GTTTCTT TGT GAT GCT CGT CAG GGG-3'	360
GEM380	5'-GTTTCTT CTG ACT TGA GCG TCG ATT-3'	380
GEM400	5'-GTTTCTT CCT GTC GGG TTT CGC CAC-3'	400

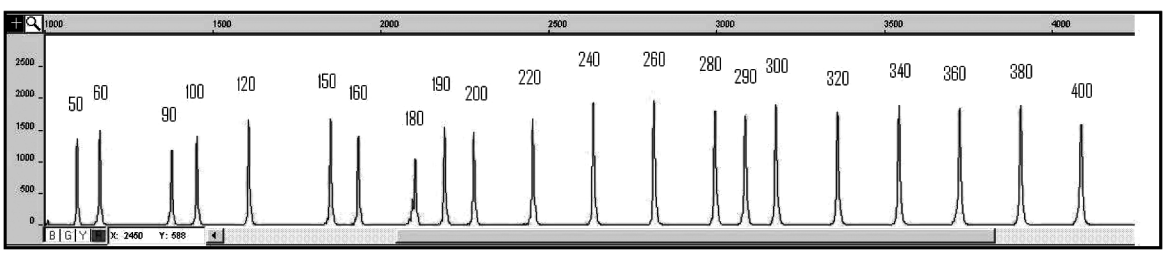

Figure 1. – Electropherogram of HM-400 internal size standard. The fragment sizes (bp) are indicated above each peak.

The consistency of HM-400 was verified using six microsatellite loci, i.e., *Hbi*161 (LEE et al., 2004a), *Sle*111a and *Sle*392 (LEE et al., 2004b), *Kma*180, *Kma*141 and *Kma*147 (LEE et al., 2006), in comparison with GS-400. *Hbi*161, *Sle*111a, and *Sle*392 were amplified on three samples of *N. heimii* and produced fragment sizes ranging from 99 to 200 bp while *Kma*180, *Kma*141, and *Kma*147 were amplified on three individuals of *K. malaccensis* and produced fragment sizes ranging from 249 to 350 bp (*Table 2*). The amplified fragments were sized for five replications (five batches for each size standard). Samples sized with HM-400 had standard deviations (SD) ranging from 0.024 to 0.489 (mean = 0.19), while samples sized with GS-400 achieved lower SD ranging from 0.000 to 0.283 (mean = 0.07). The SD of 0.489 at *Sle*111a was caused by one batch of HM-400 and this was not observed again in the subsequent batched produced. These comparisons show that all fragments sized by HM-400, on average, will be assigned a value within plus or minus 0.2 bp of one another. This is important because any fragment with SD of more than 0.5 bp will be assigned as an allele that differs by a single nucleotide (GILL et al., 1996). The differences between the fragment size means for samples sized using HM-400 and GS-400 were small, i.e., with an average of 0.29 bp. This shows that HM-400 is both highly consistent between runs and with the equivalent accuracy in fragment sizing as the commercially available standard.

Table 2. – Comparisons of fragment sizes using HM-400 and GS-400 internal size standards (bp).

Locus	HM-400		GS-400		Diff.
	Mean	SD*	Mean	SD*	means**
*Hbi*161	99.43	0.268	99.63	0.283	0.20
*Sle*111a	141.33	0.489	141.43	0.062	0.10
*Sle*392	189.59	0.043	189.90	0.000	0.31
*Kma*180	249.62	0.194	250.21	0.042	0.59
*Kma*141	309.12	0.024	309.31	0.021	0.19
*Kma*147	345.34	0.092	345.01	0.037	0.33

* SD = standard deviation.

** Diff. means = difference between the means of replicates sized with HM-400 and GS-400.

DOI:10.1515/sg-2009-0012
edited by Thünen Institute of Forest Genetics

The main reagents for the production of HM-400 are plasmid pGEM, labelled forward primer, unlabelled reverse primers, dNTPs and *Taq* DNA polymerase. The *Taq* DNA polymerase could be purchased from many vendors such as Promega Corporation (GoTaq® Flexi DNA Polymerase #M8295), Fermentas Life Sciences (*Taq* DNA Polymerase #EP0402) and New England Biolabs (*Taq* DNA Polymerase #M0273G). All of them produced similar results. Our cost analysis showed that the production cost of HM-400 was only 10% of the cost of commercial size standard. The cost for genotyping one sample using commercial size standard can be used to genotype ten samples using HM-400. In summary, the production of HM-400 is simple and cost effective, and it can become an impetus for high-throughput microsatellite genotyping without sacrificing the accuracy.

Acknowledgements

This study was supported by the IRPA research grant (09-04-01-0013-EA001) from the Ministry of Science, Technology and Innovation, Malaysia. We thank LEE, CHAI TING and TNAH LEE HONG for providing us the DNA samples and genotype data to assess the accuracy of HM-400 in comparison with GS-400 internal size standard.

Literature

BRONDANI, R. P. V. and D. GRATTAPAGLIA (2001): Cost-effective method to synthesize a fluorescent internal DNA standard for automated fragment sizing. Biotechniques **31(4)**: 793–800.

BROWNSTEIN, M. J., J. D. CARPTEN and J. R. SMITH (1996): Modulation of nontemplated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. Biotechniques **20(6)**: 1004–1010.

CLARK, J. M. (1988): Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. Nucleic Acids Research **16**: 9677–9686.

DEWOODY, J. A., J. SCHUPP, L. KENEFIC, J. BUSCH, L. MURFITT and P. KEIM (2004): Universal method for producing ROX-labeled size standards suitable for automated genotyping. Biotechniques **37(3)**: 348–352.

GILL, P., A. URQUHART, E. S. MILLICAN, N. J. OLDROYD, S. WATSON, R. PARKER and C. P. KIMPTON (1996): A new method of STR interpretation using inferential logic-development of a criminal intelligence database. International Journal of Legal Medicine **109**: 14–22.

LEE, C. T., S. L. LEE, Q. Z. FARIDAH, S. S. SIRAJ, K. K. S. NG, B. NORLIA and M. N. MAT-ISA (2006): Isolation and characterization of microsatellite markers in *Koompassia malaccensis* (Leguminosae), an important tropical timber species. Molecular Ecology Notes **6**: 1198–1201.

LEE, S. L., N. TANI, K. K. S. NG and Y. TSUMURA (2004a): Characterization of 15 polymorphic microsatellite loci in an endangered tropical tree *Hopea bilitonensis* (Dipterocarpacea) in Peninsular Malaysia. Molecular Ecology Notes **4**: 147–149.

LEE, S. L., N. TANI, K. K. S. NG and Y. TSUMURA (2004b): Isolation and characterization of 20 microsatellite loci for an important tropical tree *Shorea leprosula* (Dipterocarpaceae) and their applicability to *S. parvifolia*. Molecular Ecology Notes **4**: 222–225.

NG, K. K. S., S. L. LEE and C. L. KOH (2004): Spatial structure and genetic diversity of two tropical tree species with contrasting breeding systems and different ploidy levels. Molecular Ecology **13**: 657–669.

NG, K. K. S., S. L. LEE, L. G. SAW, J. B. PLOTKIN and C. L. KOH (2006): Spatial structure and genetic diversity of tree species with different habitat preferences within a natural forest. Tree Genetics and Genomes **3**: 121–131.

SYMONDS, V. V. and A. M. LLOYD (2004): A simple and inexpensive method for producing fluorescently labeled size standard. Molecular Ecology Notes **4**: 768–771.

Herausgeber: Johann Heinrich von Thünen-Institut. Bundesforschungsinstitut für Ländliche Räume, Wald und Fischerei.
Schriftleitung: Institut für Forstgenetik, Sieker Landstraße 2, D-22927 Großhansdorf
Verlag: J. D. Sauerländer's Verlag, Finkenhofstraße 21, D-60322 Frankfurt a. M.
Anzeigenverwaltung: J. D. Sauerländer's Verlag, Frankfurt am Main.
Satz und Druck: ADN Offsetdruck, Battenberg — Printed in Germany.

© J. D. Sauerländer's Verlag, Frankfurt a. M., 2009
ISSN 0037-5349

DOI:10.1515/sg-2009-0012
edited by Thünen Institute of Forest Genetics