Journal of Tropical Plant Physiology 13(1) (2021):15-25
https://doi.org/10.56999/jtpp.2021.13.1.11

Development of Genic SSR Markers from Transcriptome Sequences of *Baeckea frutescens* (Cucur Atap) for Genetic Diversity Assessment

Nur Nabilah Alias*, Chai Ting Lee, Nurul Farhanah Zakaria, Soon Leong Lee, Kevin Kit Siong Ng, Chin Hong Ng, Lee Hong Tnah and Norlia Basherudin
Genetics Laboratory, Forest Research Institute Malaysia (FRIM), 52109 Kepong, Selangor, Malaysia
**Email: nabilah@frim.gov.my*

Received: 20 January 2021; Revised: 2 April 2021; Accepted: 27 April 2021; Published: 6 June 2021

ABSTRACT

Baeckea frutescens or Cucur Atap, is a medicinal plant that has been traditionally used as herbal medicine and has great potential for pharmaceutical application due to its biological properties. The increasing demand for *B. frutescens* herbal medicinal products has led to its over-exploitation. Therefore, it is crucial to generate baseline genetic information in order to formulate conservation and breeding programmes for this species. In this study, a total of 26 polymorphic simple sequence repeat (SSR) markers have been developed in *B. frutescens* based on the leaf transcriptome sequence data obtained via next generation sequencing (NGS) technology. These markers were evaluated using 32 individuals from Forest Research Institute Malaysia Research Station in Setiu, Terengganu. The mean number of alleles was two and the average expected (H_E) and observed heterozygosity (H_O) were 0.484 and 0.900, respectively. Cluster analysis revealed high genetic similarity among the individuals, whereby many of them were of the same clones. It is essential to assess more individuals from other populations for better understanding on the genetic diversity of this useful medicinal plant.

Keywords: Cluster analysis; genetic diversity; medicinal plant; next generation sequencing (NGS); simple sequence repeats (SSRs).

INTRODUCTION

Baeckea frutescens, locally known as Cucur Atap, is a medicinal plant of the family Myrtaceae and subfamily Myrtoideae. It is a shrub or small evergreen tree that is found in Peninsular Malaysia, Sumatra and the coastal areas of southern China to Australia. In Peninsular Malaysia, *B. frutescens* is found on mountain tops, quartz ridge and sandy coasts of the eastern parts (Wong et al., 2010). Throughout South-East Asia, the flowers and leaves of *B. frutescens* are made into herbal tea and are used in treating dyspepsia and menstrual problems (Navanesan et al., 2015). In Malaysia and Indonesia, it is traditionally associated with health drinks and powders given to mothers after childbirth. Furthermore, the essential oil from this plant is also used in the treatment of rheumatism. Although *B. frutescens* has been widely used as traditional medicine and has shown great potential for use in pharmaceuticals because of its anti-inflammatory, antimicrobial, antimalarial and antioxidant properties (Navanesan et al., 2015), very little is known about the genetics of this species. In view of the importance and rising demand of *B. frutescens* for pharmaceutical industries, it is important to gather baseline genetic data for *B. frutescens* to facilitate conservation and breeding programmes for this species.

Simple sequence repeats (SSRs), also known as microsatellites, are short tandem repeat sequences of DNA that are useful for the detection of genetic variations because they are abundant, highly polymorphic, highly reproducible and multiallelic (Wang et al., 2018). Additionally, the co-dominance nature of SSRs allows the direct measurement of heterozygosity, which only requires small amount of DNA for data collection (Mason, 2015). As such, SSR markers have also been widely used as a powerful tool in

studies of population genetic diversity and genetic structure analysis (Yoichi et al., 2016). The emergence of next-generation sequencing (NGS) in recent years has increased the sequencing capability far beyond the traditional Sanger method. Recent publications have shown that NGS plays a low-cost and time-efficient role in polymorphic SSR discovery (Chen et al., 2014). In fact, transcriptome method has facilitated the effort of SSR discovery especially in non-model plants with no reference genome, thus it has become the method of choice nowadays (Zalapa et al., 2012; Yan et al., 2015).

In this study, we developed polymorphic SSR markers based on the transcriptome sequence of *B. frutescens* obtained via Illumina paired-end sequencing. To the best of our knowledge, this was the first study on the development of SSR markers in *B. frutescens*. The developed SSR markers can be used for genetic studies to improve the utilisation and conservation of *B. frutescens* genetic resources.

MATERIALS AND METHODS

Sample collection and DNA extraction

Leaf samples of 32 *B. frutescens* individuals were collected from Forest Research Institute Malaysia (FRIM) Research Station in Setiu, Terengganu (SPF Setiu). Total DNA was extracted from approximately 5 g of fresh leaf tissues for each sample using a modified CTAB method (Murray and Thompson, 1980) and further purified using High Pure PCR Template Preparation Kit (Roche Diagnostics, Germany). The integrity and quality of the DNA were evaluated by electrophoresis on 1% agarose gels and the concentration of the DNA samples was determined using NanoDrop 2000 spectrometer (Thermo Fisher Scientific, USA).

RNA extraction and transcriptome sequencing

Total RNA from fresh leaves of *B. frutescens* was extracted using RNeasy Plant Mini Kit (Qiagen, USA) following the manufacturer's instructions. The quality of the extracted RNA was checked using NanoDrop 2000 spectrometer (Thermo Fisher Scientific, USA) and 1% agarose gel electrophoresis, and quantified using Qubit 2.0 (Thermo Fisher Scientific, USA). The RNA integrity and quantification were verified using Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA) with the minimum RNA integrated number (RIN) value of 8. The cDNA library was constructed following the NEBNext Ultra II RNA Library Prep Kit protocol along with NEBNext Poly(A) mRNA Magnetic Isolation Module (New England BioLabs, USA). The transcriptome sequencing was carried out at the Genetics Lab, FRIM, Kepong, Selangor using paired-end Illumina MiSeq sequencing on an Illumina MiSeq (Illumina Inc., USA) with 600 cycles according to the manufacturer's instructions.

SSR identification and marker development

The quality of generated paired-end reads was checked using FastQC (Andrews, 2010). The sequences had undergone pre-processing analysis using Trimmomatic v0.32 (Bolger et al., 2014) to discard low quality sequences. *De novo* transcriptome assembly of the cleaned paired-end reads was then performed using Trinity RNA-Seq v2.4.0 with default parameters (Grabherr et al., 2011). The assembled data were deposited in the NCBI Sequence Read Archive (SRA) under the accession number PRJNA725027. SSR-containing sequences were identified using MIcroSAtellite identification tool (MISA) (Thiel et al., 2003), with the minimum number of repeats set at 10 for mono-nucleotide repeats, six for di-nucleotide motifs, and five for tri-, tetra-, penta- and hexa-nucleotide repeats. Primers were designed using Primer 3 (Rozen and Skaletsky, 2000) using an in-house script. Selection of primers was based on the criteria which were primer length of 18 to 24 bases, GC content of 50 to 60%, melting temperature of 50 to 63°C and the predicted PCR product sizes of 100 to 380 bp.

Screening of SSR primer pairs and validation of SSR loci

Initially, 70 primer pairs from the Primer 3 output that meet the criteria mentioned above were randomly selected and screened for four *B. frutescens* individuals by PCR amplification and electrophoresis on 2% agarose gels. The sizes of the PCR products were calculated by comparison with 1 kb DNA ladder. The primer pairs that produced specific-single band were then selected for 5' end fluorescent labelling using either 6-FAM or HEX, at the forward primers. The labelled primers were further screened for polymorphisms using 32 individuals from FRIM Research Station in Setiu. The multiplex PCR amplifications were performed in 8 µL reaction mixture, consisting of approximately 5 ng of DNA template, 1 × Type-it Multiplex PCR Master Mix (Qiagen, Germany) and 0.2 µM of primer mix for an initial denaturing step of 5 min at 95°C, 35 cycles of 95°C for 30 s, 55°C annealing temperature for 1 min 30 s, and 72°C for 30 s, followed by a final extension of 30 min at 60°C. The PCR products were subjected to fragment analysis using an ABI 3130xl Genetic Analyzer (Applied Biosystems, USA) with ROX 400 as the internal size standard (Applied Biosystems, USA). Genotyping was performed using GeneMarker v2.6.4 (SoftGenetics, 2010).

Evaluation of SSR markers

The 32 *B. frutescens* samples were used for SSR markers evaluation. Microsatellite Toolkit (Park, 2008) was applied to analyse the genotypic data. Genetic diversity parameters such as number of alleles (A), observed (H_O) and expected heterozygosities (H_E) as well as polymorphic information content (PIC) were estimated. MICROCHECKER v.2.2.3 (van Oosterhout et al., 2004) was used to check for scoring errors caused by stutters or large-allele dropouts and to estimate null-allele frequencies. For analysing the genetic relationship among the samples, cluster analysis was carried out based on unrooted tree construction and Unweighted Pair wise Methods with Arithmetic averages (UPGMA) (Nei and Li, 1979) to generate a dendogram via PowerMarker version 3.25 (Liu and Muse, 2005).

RESULTS AND DISCUSSION

SSR distribution in transcriptome sequences of *B. frutescens*

Transcriptome sequencing of *B. frutescens* produced 3,649,253 raw reads. After trimming, a total of 3,597,597 clean reads were obtained. The clean reads were *de novo* assembled into 96,581 contigs and were used for mining potential SSR motifs. From the total number of sequences examined, there were 9,213 identified SSRs with di-nucleotide motifs being the most enriched (8,500; 92.26%), followed by tri-nucleotides (682; 7.4%) and tetra-nucleotides (31; 0.34%) (Figure 1). *B. frutescens* is classified as dicotyledonous plants. According to Biet et al. (1999), di-nucleotide repeats are generally the most frequent SSR motif type in dicotyledonous species whereas tri-nucleotide repeats are most common in graminaceous plants such as rice, wheat and barley. The finding was comparable to the study by Peng et al. (2008) which demonstrated that di-nucleotide repeat type had the highest occurrence rate (36.03%) in cassava, followed by tri-nucleotide (31.84%) and mono-nucleotide repeat type (30.10%). The result was also consistent with a study by Liu et al. (2018), showing that di-nucleotide SSRs were the most common in *Eucalyptus*, followed by tri-nucleotide SSRs. Furthermore, similar results by Yan et al. (2015), Chen et al. (2015) and Xing et al. (2017) have shown that di-nucleotide repeats were the most common SSR repeat type in dicotyledonous plants. In contrast, Li et al. (2018) reported that tri-nucleotide repeats were the most common in centipedegrass, followed by di-, tetra-, hex- and penta-nucleotides repeats. The results were also in accordance with the findings in other plant species, for example peanut (Zhou et al., 2016) and chrysanthemum (Wang et al. 2013).

In addition, a total of 24 different types of motifs were identified which belonged to four types of di-nucleotides repeats, nine types of tri-nucleotides and 11 types of tetra-nucleotides. As shown in Figure

2, the AG/CT motif was the most predominant di-repeat (8,139; 88.34%) followed by AC/GT (264; 2.87%), AT/AT (92; 1.00%) and CG/CG (5; 0.05%). For the tri-nucleotide motifs, the AAG/CTT motif (319; 3.46%) was the most abundant tri-repeats followed by AGG/CCT (141; 1.53%) and AGC/CTG (77; 0.84%). Among tetra-repeats, ACAT/ATGT showed the largest proportion (6; 0.07%). The findings were consistent with the result of centipedegrass (Li et al., 2018) and *Vernicia fordii* (Xu et al., 2012), which reported that AG/CT showed the largest proportion in di-nucleotide repeat whereas AAG/CTT showed the largest proportion in tri-nucleotide repeat. Previous studies by Chen et al. (2015) also indicated that AG/CT and AAG/CTT were the most abundant di- and tri-nucleotide repeat motifs, respectively in dicotyledonous plants. In contrast, CCG/CGG (21; 0.23%) was a rare motif type in the transcriptome of *B. frutescens*. This was similar to other dicotyledonous plants, for instance *Dysosma versipellis* (Guo et al., 2014), *Epimedium sagittatum* (Zeng et al., 2010) and *Raphanus sativus* L. (Jiang et al., 2012).

Screening of SSR primers

Initially, 70 primer pairs with 30, 34 and 6 primer pairs for di-, tri- and tetra-nucleotide SSRs, respectively, were randomly selected for screening. Fifty-five primer pairs showed specific-single bands on 2% agarose gel electrophoresis (Figure 3) and subsequently the forward primers were fluorescently labelled to be used for fragment analysis. In this study, 15 out of 70 primer pairs were unable to amplify effectively and have been excluded from further analyses.

Furthermore, fragment analysis showed that 38 out of 55 SSR markers yielded consistent genotypes and clean peaks without any noise. Among these 38 SSR loci, 26 loci were polymorphic while the remaining 12 SSR markers were identified as monomorphic with only one allele observed per locus. The examples of the genotype profiles of selected polymorphic loci were shown in Figure 4. Out of the 26 markers developed, tri-nucleotide SSRs have the highest abundance with 14 markers, followed by eight and four markers for di- and tetra-nucleotide SSRs, respectively. The overall proportion of polymorphic primers was 47.3%, which was comparable to the polymorphic ratio of EST-SSRs in castor bean (41.1%) (Qiu et al. 2010).

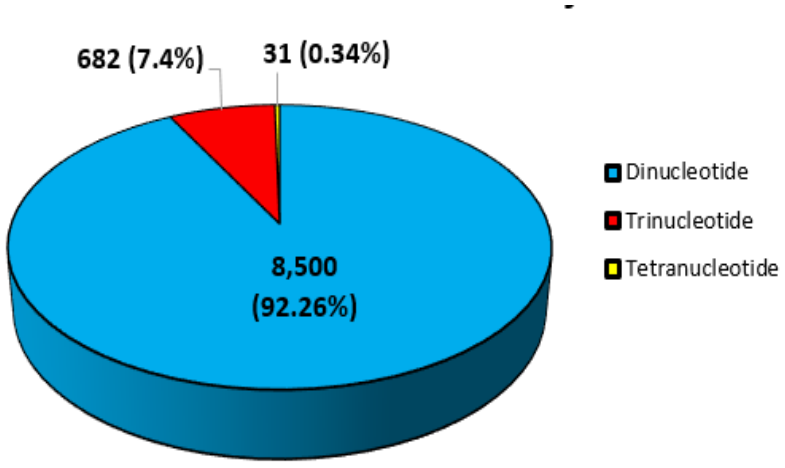

Figure 1. Di-, tri- and tetra-nucleotides SSR distribution in *B. frutescens*

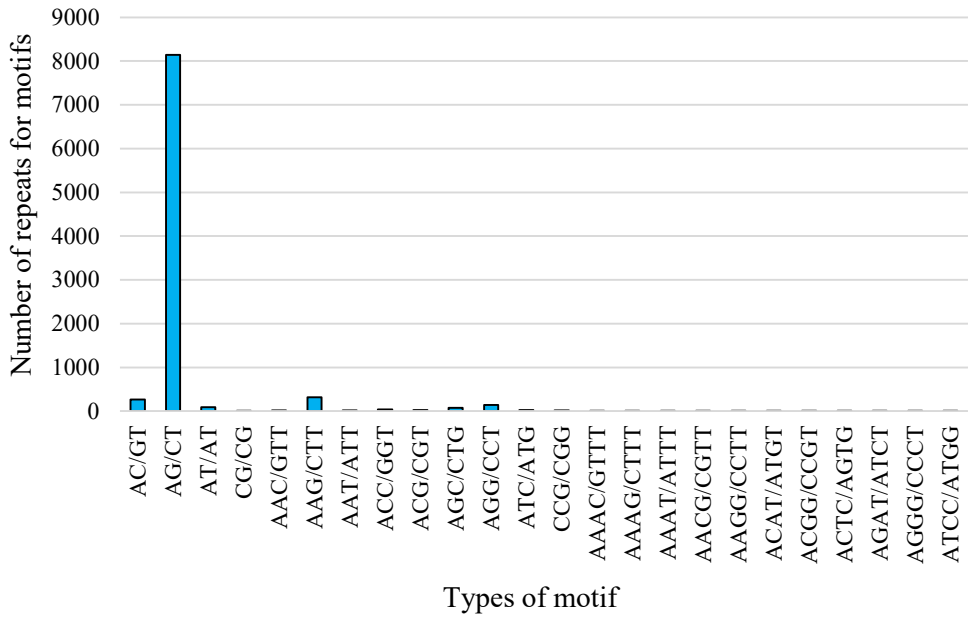

Figure 2. Distribution of SSR repeat motifs in *B. frutescens* showing details of motifs comprising di-, tri- and tetranucleotides with sequence complementary

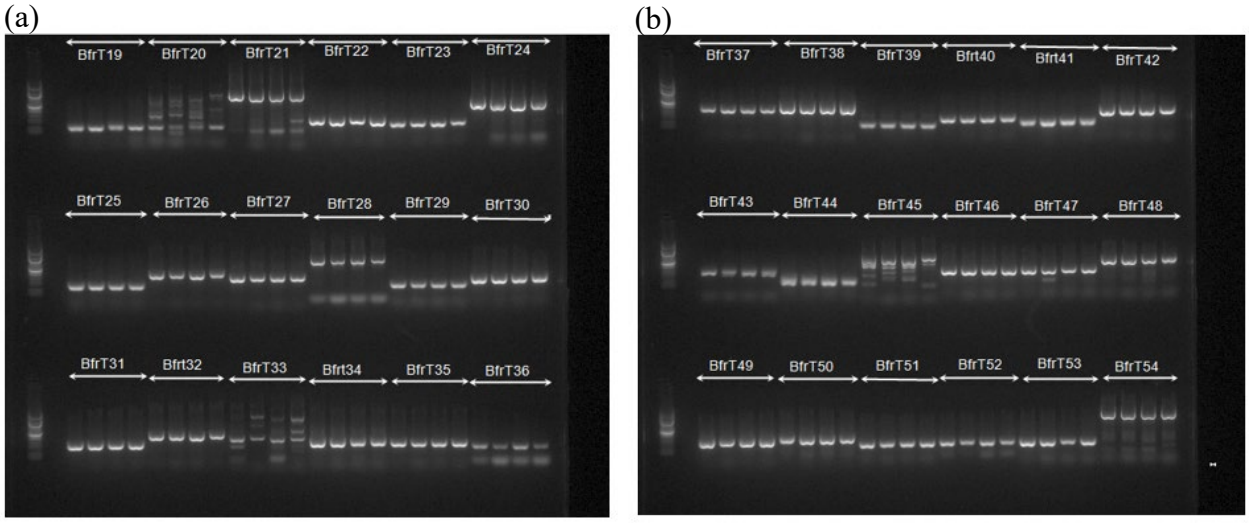

Figure 3. Part of the results for primer screening viewed in 2% agarose gels. Gel images for PCR products of (a) *Bfr*T19-*Bfr*T36 (b) *Bfr*T37-*Bfr*T54 primer pairs with 100 base pair ladders on the leftmost well of each lane.

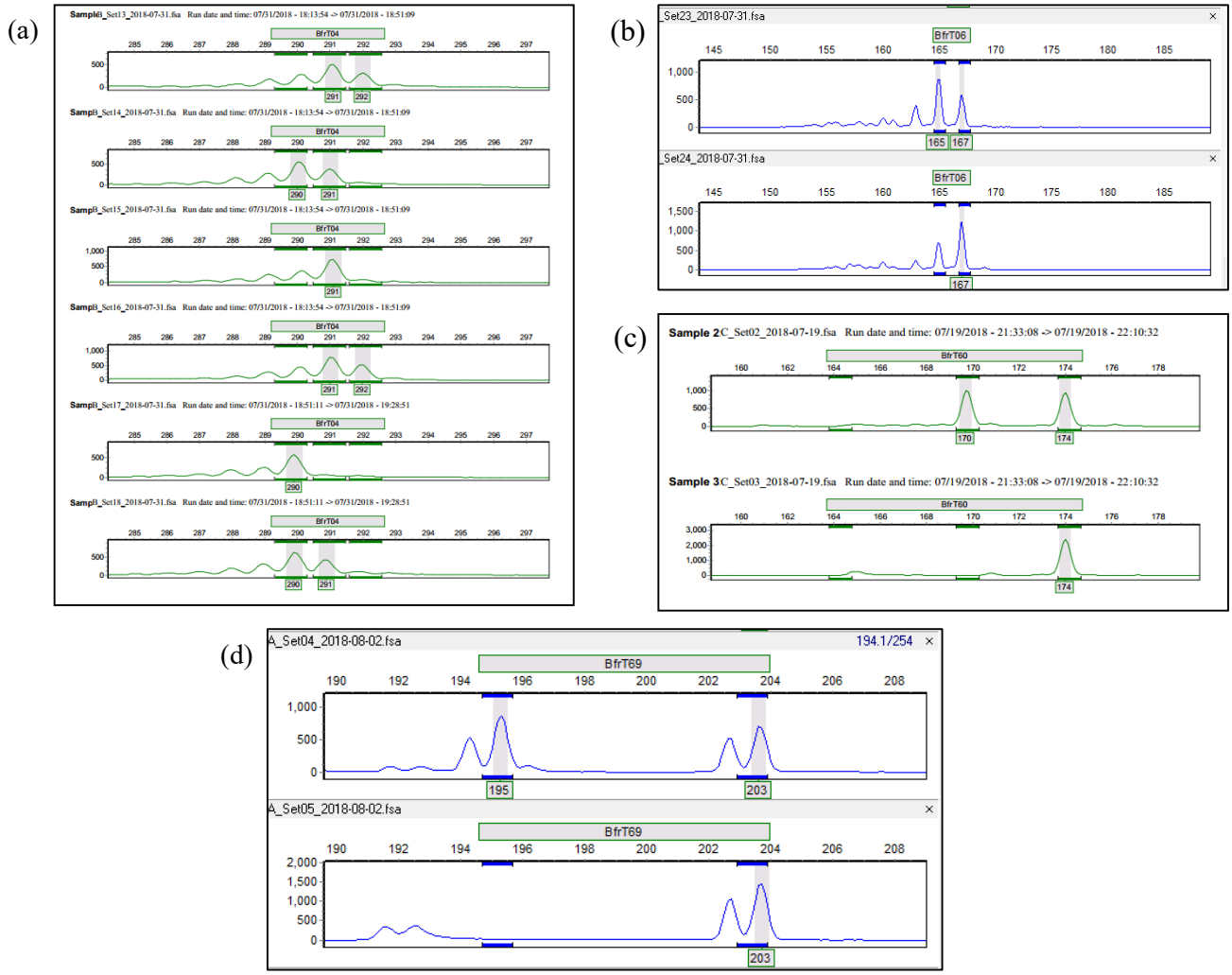

Figure 4. Genotype profiles of selected polymorphic loci. (a) *Bfr*T04 (b) *Bfr*T06 (c) *Bfr*T60 and (d) *Bfr*T69.

Evaluation of SSR markers

Genetic diversity is defined as the variation of heritable characteristics present in a population of the same species (Swingland et al., 2001). The 26 polymorphic genic SSR markers developed in this study were evaluated based on the 32 *B. frutescens* individuals. As presented in Table 1, the average expected (H_E) and observed heterozygosity (H_O) were 0.484 and 0.900, respectively, with a mean number of alleles of two. Mean number of alleles (MNA) per locus is a good indicator of the genetic polymorphism within the population (Hassen et al., 2012). However, it is largely dependent on the sample size (Hipparagi et al., 2017). In this study, the MNA was low despite substantial sample size because the samples studied were closely related.

The number of detected alleles may increase when more populations were investigated, assuming that more genetic diversity will be captured from different populations. Salunkhe et al. (2013) analysed the genetic diversity of 48 Indian emmer wheat cultivars using 47 SSR markers and identified 2 to 9 alleles per locus with a mean of 3.87. In other study on 34 Ethiopian emmer cultivars using 29 SSR markers, an average of 6.95 alleles per locus was detected (Teklu et al., 2006). This proved that larger sample size with

relatively higher diverse individuals led to higher number of alleles per locus (Rafeipour et al., 2016; Hipparagi et al., 2017).

Polymorphic information content (PIC) is an index that indicates the information content of molecular markers and is widely used to evaluate molecular marker polymorphism. The PIC value of each SSR locus is determined by both the number of alleles and their frequency distribution within the population (Aljumaah et al., 2012). In this study, the PIC ranged from 0.110 to 0.470 with an average of 0.362. The obtained average PIC is comparable to the value of 0.35 reported for 48 Indian emmer wheat individuals (Salunkhe et al., 2013). Similarly, Hipparagi et al. (2017) also reported 2.61 alleles with an average PIC value of 0.36 among 75 soybean genotypes assayed by 21 SSR markers. Previous studies by Bisen et al. (2015) detected 2.22 alleles per locus with an average PIC value of 0.199 using 16 SSR markers on 38 soybean varieties.

Table 1. Genetic diversity parameters for the 26 genic SSR markers in *B. frutescens.* Number of alleles (A), observed (H_O) and expected heterozygosity (H_E) and polymorphic information content (PIC) were estimated based on 32 individuals.

Locus	Repeat motif	Allele size range (bp)	A	H_O	H_E	PIC
*Bfr*T03	$(AG)_8$	114 - 122	2	1.000	0.508	0.375
*Bfr*T04	$(AG)_{15}$	290 - 292	3	0.344	0.572	0.470
*Bfr*T05	$(GT)_9$	123 - 145	2	1.000	0.508	0.375
*Bfr*T06	$(AG)_{20}$	165 -167	2	0.250	0.268	0.229
*Bfr*T22	$(CT)_8$	238 - 250	2	1.000	0.508	0.375
*Bfr*T25	$(TC)_8$	185 - 195	2	1.000	0.508	0.375
*Bfr*T27	$(CT)_{12}$	309 - 317	2	1.000	0.508	0.375
*Bfr*T30	$(GA)_{17}$	268 - 290	2	1.000	0.508	0.375
*Bfr*T31	$(TCG)_8$	240 - 250	2	1.000	0.508	0.375
*Bfr*T34	$(AAG)_8$	263 - 294	2	1.000	0.508	0.375
*Bfr*T40	$(GAG)_8$	192 - 200	2	1.000	0.508	0.375
*Bfr*T41	$(TCT)_8$	135 - 154	2	1.000	0.508	0.375
*Bfr*T42	$(ATG)_8$	314 - 330	2	1.000	0.508	0.375
*Bfr*T43	$(GGA)_8$	253 - 256	2	1.000	0.508	0.375
*Bfr*T46	$(CTT)_{10}$	333 - 345	2	1.000	0.508	0.375
*Bfr*T47	$(TCC)_8$	349 - 353	2	1.000	0.508	0.375
*Bfr*T49	$(AAG)_{10}$	256 - 259	2	1.000	0.508	0.375
*Bfr*T52	$(GAA)_8$	181 - 190	2	1.000	0.508	0.375
*Bfr*T56	$(GGC)_8$	212 - 216	2	1.000	0.508	0.375
*Bfr*T60	$(TTC)_8$	170 - 174	3	0.688	0.458	0.349
*Bfr*T62	$(GGA)_8$	142 - 148	2	1.000	0.508	0.375
*Bfr*T63	$(AGG)_8$	196 - 203	2	1.000	0.508	0.375
*Bfr*T65	$(AGGG)_6$	232 - 236	2	1.000	0.508	0.375
*Bfr*T66	$(TCCT)_6$	160 - 172	2	1.000	0.508	0.375
*Bfr*T67	$(GAGT)_6$	248 - 251	2	1.000	0.508	0.375
*Bfr*T69	$(TATG)_6$	195 - 203	2	0.125	0.119	0.110
		Mean	2	0.900	0.484	0.362

A marker with PIC of more than 0.5 can be considered as highly informative and polymorphic whereas PIC value in the range of 0.25 to 0.5 indicates moderate polymorphism and reasonably informative,

and PIC value of less than 0.25 is measured as low polymorphism (Marshall et al., 1998). Less informative markers lead to lower PIC values. In this study, 24 of the 26 pairs (92.3%) of primers showed moderate polymorphism while the remaining two markers (*Bfr*T06 and *Bfr*T69) showed low polymorphism (7.7%). The highest PIC value was 0.470, for locus *Bfr*T04. Null alleles usually occur due to mutations at the priming sites. In this study, null allele was detected only in one locus (*Bfr*T04) and repetition of amplification yielded the same result. In general, these 26 SSR markers were reasonably informative and could be used in future genetic analysis of *B. frutescens*.

Genetic relatedness among individuals

The close genetic relationship among the 32 individuals was confirmed by cluster analysis. Based on the 26 loci, the 32 individuals of *B. frutescens* were separated into three main clusters, corresponding to Cluster A, Cluster B and Cluster C (Figure 5). Cluster A consisted of 15 individuals with six sub-clusters. Cluster B comprised nine individuals with six sub-clusters. Cluster C consisted of eight individuals with three sub-clusters. The dendrogram generated revealed high genetic similarity among the individuals. Overall, there were 15 sub-clusters of individuals which shared similar multilocus genotype within sub-cluster (Figure 5), implying the presence of clonal plants.

CONCLUSIONS

A total of 26 polymorphic genic SSR markers with consistent and scorable genotypes based on the fragment analysis which excluded the monomorphic loci have been successfully developed for *B. frutescens.* Our study demonstrated that SSR technology is a powerful tool for evaluating genetic diversity and relationship among the *B. frutescens* individuals. However, since the gene pool of *B. frutescens* at our study site was very limited, it is essential to assess more individuals from other populations across Peninsular Malaysia to reveal information on the genetic diversity distribution and population genetic structure of *B. frutescens*. The SSR markers developed can also be utilised to evaluate *B. frutescens* germplasm collections for conservation and breeding purposes. Conservation programme guided by genetic information is crucial to prevent irreversible genetic erosion of this useful medicinal plant, while effective breeding programme will support planting of *B. frutescens* to ensure sustainable utilisation

AUTHORS CONTRIBUTION

NNA and CTL conceived and designed the work. NNA, CTL, NFZ, SLL, KKSN and CHN performed the analysis. NNA and CTL wrote the paper. LNT and NB checked and approved the submission.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

FUNDING

This study was funded by the Ministry of Energy and Natural Resources under the 11th Malaysia Plan.

ACKNOWLEDGEMENTS

We thank the supporting staff from the FRIM Genetics Laboratory (Ramli, P., Ghazali, J., Yahya, M., Yasri, B., Sharifah, T. and the late Suryani, C.S.) for field and technical assistance.

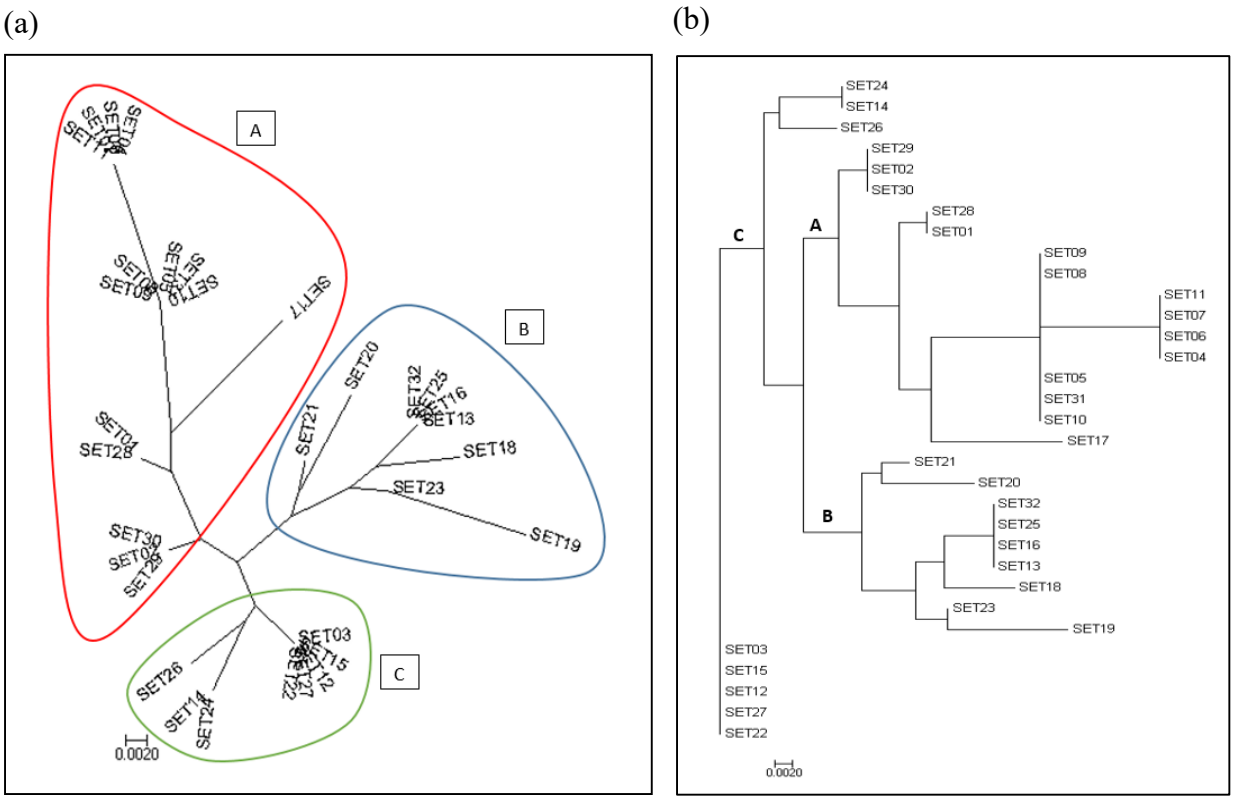

Figure 5. Cluster analysis among the *B. frutescens* samples. (a) An unrooted tree based on the proportion-of-shared-allele distances. (b) A dendrogram of genetic relationships based on UPGMA.

REFERENCES

Aljumaah, R. S., Musthafa, M. M., Al-Shaikh, M. A., Badri, O. M. and Hussein, M. F. (2012). Genetic diversity of Ardi goats based on microsatellite analysis. *African Journal of Biotechnology*, 11(100), 16539-16545.

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc [Accessed on 26 April 2019].

Biet, E., Sun, J. and Dutreix, M. (1999). Conserved sequence preference in DNA binding among recombination proteins: an effect of ssDNA secondary structure. *Nucleic Acids Research*, 27, 596-600.

Bisen, A., Khare, D., Nair, P. and Tripathi, N. 2015. SSR analysis of 38 genotypes of soybean (*Glycine max* (L.) Merr.) genetic diversity in India. Physiology and Molecular Biology of Plants 21(1): 109-115.

Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.

Chen, C. M., Sio, C. P., Lu, Y. L., Chang, H. T., Hu, C. H. and Pai, T. W. (2014). Identification of conserved and polymorphic STRs for personal genomes. *BMC Genomics*, 15 (suppl 10), S3.

Chen, L. Y., Cao, Y. N., Yuan, N., Nakamura, K., Wang, G. M. and Qiu, Y. X. (2015). Characterization of transcription and development of novel EST-SSR makers based on next-generation sequencing technology in *Neolitsea sericea* (Lauraceae) endemic to Asian land-bridge islands. *Molecular Breeding*, 35, 187.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I. and Regev, A. (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology* 29(7), 644-652.

Guo, R., Mao, Y. R., Cai, J. R., Wang, J. Y., Wu, J. and Qiu, Y. X. (2014). Characterization and cross-species transferability of EST-SSR markers developed from the transcriptome of *Dysosma versipellis* (Berberidaceae) and their application to population genetic studies. *Molecular Breeding*, 34, 1733-1746.

Hassen, H., Lababidi, S., Rischkowsky, B., Baum, M. and Tibbo, M. (2012). Molecular characterization of Ethiopian indigenous goat populations. *Tropical Animal Health Production*, 44(6), 1239-1246.

Hipparagi, Y., Singh, R., Choudhury, D. R. and Gupta, V. (2017). Genetic diversity and population structure analysis of Kala bhat (*Glycine max* (L.) Merrill) genotypes using SSR markers. *Hereditas*, 154, 9.

Jiang, L., Wang, L., Liu, L., Zhu, X., Zhai, L. and Gong, Y. (2012). Development and characterization of cDNA library based novel EST-SSR marker in radish (*Raphanus sativus* L.). *Scientia Horticulturae*, 140, 164-172.

Li, J., Guo, H., Wang, Y., Zong, J., Chen, J., Li, D., Li, L., Wang, J. and Liu, J. (2018). High-throughput SSR marker development and its application in a centipedegrass (*Eremochloa ophiuroides* (Munro) Hack.) genetic diversity analysis. *PLoS ONE*, 13(8), e0202605.

Liu, G., Xie, Y., Zhang, D. and Chen, H. (2018). Analysis of SSR loci and development of SSR primers in *Eucalyptus. Journal of Forestry Research*, 29(2), 273-282.

Liu, K. and Muse, S. V. 2005. POWERMARKER: Integrated analysis environment for genetic marker data. *Bioinformatics*, 21, 2128-2129.

Marshall, T. C., Slate, J., Kruuk, L. E. B. and Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural population. *Molecular Ecology*, 7(5), 639-655.

Mason, A. S. (2015). SSR gentyping. *Methods in Molecular Biology*, 1245, 77-89.

Murray, M. G. and Thompson, W. F. (1980). Rapid isolation of high molecular weight DNA. *Nucleic Acids Research*, 8, 4321-4325.

Navanesan, S., Wahab, N. A., Manickam, S. and Sim, K. S. (2015). Evaluation of selected biological capacities of *Baeckea frutescens*. *BMC Complementary and Alternative Medicine*, 15, 186.

Nei, M. and Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleasis. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), 5269-5273.

Ojango, J. M., Mpofu, N., Marshall, K. and Andersson-Eklund, L. (2011). Quantitative methods to improve the understanding and utilization of animal genetic resources. In: Ojango, J.M., Malmfors, B. and Okeyo, A.M. (Eds.) *Animal genetics training resource, version 3.* Nairobi, Kenya, International Livestock Research Institute and Uppsala, Sweden: Swedish University of Agricultural Sciences, pp. 38.

Park, S. D. E. (2008). *Excel Microsatellite Toolkit. Computer program and documentation distributed by the author*. Available at: http://animalgenomics.ucd.ie/sdepark/ms-toolkit/ [Accessed on 16 June 2019].

Peng, D. W., Zheng, L. C. and Zhu, H. B. (2008). Analysis of SSR information in EST resource of cassava. *Chinese Agricultural Science Bulletin*, 24(2), 433-436.

Qiu, L., Yang, C., Tan, B., Yang, J. and Liu, A. Z. (2010). Exploiting EST databases for the development and characterization of EST-SSR markers in castor (*Ricinus communis* L.). *BMC Plant Biology*, 10, 278.

Rafeipour, M., Mirzaghaderi, G., Shaaf, S. and Badakhshan, H. (2016 0. SSR assessment of the genetic diversity of emmer wheat with emphasis on Iranian landraces (*Triticum dicoccon* Schrank). *Genetic Resources and Crop Evolution*, 63, 595-600.

Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Misener, S. and Krawetz, S. A. (Eds.) *Bioinformatics Methods and Protocols. Methods in Molecular Biology™, vol 132*. New Jersey, USA, Humana Press, pp. 386.

Salunkhe, A., Tamhankar, S., Tetali, S., Zaharieva, M., Bonnett, D., Trethowan, R. and Misra, S. (2013). Molecular genetic diversity analysis in emmer wheat (*Triticum dicoccon* Schrank) from India. *Genetic Resources and Crop Evolution*, 60, 165-174.

Swingland, I. R. (2001). Definition of Biodiversity. *Encyclopedia of Biodiversity*, 1, 377-390.

Teklu, Y., Hammer, K., Huang, X. Q. and Ro¨der, M. S. (2006). Analysis of microsatellite diversity in Ethiopian tetraploid wheat landraces. *Genetic Resources and Crop Evolution*, 53, 1115-1126.

Thiel, T., Michalek, W., Varshney, R. K. and Graner, A. (2003). Exploiting EST databases for the development and characterisation of gene-derived SSR markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, 106, 411-422.

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.

Wang, H., Jiang, J., Chen, S., Qi, X., Peng, H., Li, P., Song, A., Guan, Z., Fang, W., Liao, Y. and Chen, F. (2013). Next-generation sequencing of the *Chrysanthemum nankingense* (Asteraceae) transcriptome permits large-scale unigene assembly and SSR marker discovery. *PLoS ONE*, 8, e62293.

Wang, J., Ge, Q., Li, J., Gao, H., Li, J. and Zhao, F. (2018). Identification of novel EST-SSR markers by transcriptome sequencing in ridgetail white prawn *Exopalaemon carinicauda*. *Genes & Genomics*, 40, 207-215.

Wong, K., Sugumaran, M., Lee, D. and Zahid, M. (2010). Ecological aspects of endemic plant populations on Klang Gates quartz ridge, a habitat island in Peninsular Malaysia. *Biodiversity Conservation*, 19(2), 435-447.

Xing, W., Liao, J. Y., Cai, M. Y., Xia, Q. F., Liu, Y. and Zeng, W. (2017). *De novo* assembly of transcriptome from *Rhododendron latoucheae* Franch. using Illumina sequencing and development of new EST-SSR markers for genetic diversity analysis in Rhododendron. *Tree Genetics & Genomes*, 13, 53.

Xu, W., Yang, Q., Huai, H. and Liu, A. (2012). Development of EST-SSR markers and investigation of genetic relatedness in tung tree. *Tree Genetics & Genomes*, 8(4), 933-940.

Yan, X., Zhang, X., Lu, M., He, Y. and An, H. (2015). *De novo* sequencing analysis of the *Rosa roxburghii* fruit transcriptome reveals putative ascorbate biosynthetic genes and EST-SSR markers. *Gene*, 561, 54-62.

Yoichi, W., Sakaguchi, S., Ueno, S., Tomaru, N. and Uehara, K. (2016). Development and characterization of EST-SSR markers for the genus *Rhododendron* section Brachycalyx (Ericaceae). *Plant Species Biology*, 32(4), 455-459.

Zalapa, J.E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R. and Simon, P. (2012). Using next-generation sequencing approaches to isolate simple sequence repeat (SSRs) loci in the plant sciences. *American Journal of Botany*, 99, 193-208.

Zeng, S. H., Xiao, G., Guo, J., Fei, Z. J., Xu, Y. Q., Roe, B. A. and Wang, Y. (2010). Development of a EST dataset and characterization of EST-SSRs in a traditional Chinese medicinal plant, *Epimedium sagittatum* (Sieb. Et Zucc.) Maxim. *BMC Genomics* 11, 94.

Zhou, X., Dong, Y., Zhao, J., Huang, L., Ren, X., Chen, Y., Huang, S., Liao, B., Lei, Y., Yan, L. and Jiang, H. (2016). Genomic survey sequencing for development and validation of single-locus SSR markers in peanut (*Arachis hypogaea L.*). *BMC Genomics*, 17, 420.