Usage of microsatellite markers for characterization of polyploids: a case study in reference to hexaploid bamboo species

Rajendra K. Meena*, Maneesh S. Bhandari and Harish S. Ginwal

Division of Genetics & Tree Improvement, Forest Research Institute, Dehradun-248 195, Uttarakhand, India.

* Corresponding author: Rajendra K. Meena, E-mail: rajnrscp@gmail.com

Abstract

Microsatellite markers are most valuable tools for characterization of plant genetic resources or population genetic analysis. Since they are codominant and allelic markers, utilizing them in polyploid species remained doubtful. In such cases, microsatellite markers are usually analyzed by treating them as dominant marker. In the current study, it has been showed that despite of losing the advantage of co-dominance, microsatellite markers are still powerful tool for genotyping of polyploid species because of availability of large number of reproducible alleles per locus. It has been studied by genotyping of nineteen sub populations of Dendrocalamus hamiltonii (hexaploid bamboo species) with seventeen polymorphic SSR primer pairs. Among these, ten primers gave typical banding pattern of microsatellite marker as expected in diploid species but rest seven gave unusual pattern i.e. more than two bands per locus per genotype. In such case genotyping data are generally analyzed by considering as dominant markers. Given these facts, data were analyzed in both ways as dominant and codominant. All the seventeen primer were first scored as non-allelic data and analyzed; later ten primer pairs giving standard banding pattern of microsatellite marker as expected in diploid species were analyzed as allelic data and the results were compared. The UPGMA clustering and genetic structure showed that results obtained with both the data sets were very similar, and therefore the SSR marker could be utilized to characterize polyploid species by considering them as dominant marker. The study is highly useful to widen the scope of SSR markers applications and beneficial to the researchers dealing with polyploid species.

Keywords: Microsatellite markers, dominant and codominant, Dendrocalamus hamiltonii

Research Note

Microsatellite based simple sequence repeat (SSR) markers are one of the most valuable tools for characterization of plant genetic resources because of their abundance, wide genome coverage, higher reproducibility and codominant nature. The SSRs usually generate allelic data and reveal the information of heterozygosity in the individuals of population. In a diploid species the allelic relationship could easily be established by considering single band as homozygous and double band as heterozygous locus; but in polyploid species, where genome present in multiple copies, more than two bands may be obtained, and assigning correct allelic dosage is difficult (Sampson and Byrne, 2012; Garcia-Verdugo et al., 2013; Bhandawat et al., 2019). Although, allele dosage has been successfully estimated in tetraploids (Obbard et al., 2006) and in allopolyploids (Van Dijk et al., 2012), it is often unreliable in high-order polyploids. The genetic analysis with allelic marker is highly challenging in the autopolyploid species because of random bivalent pairing during meiosis. In contrast, allopolyploid species consist of distinct set of genomes derived from their ancestral progenitors during evolutionary hybridization events. These differentiated genomes behave as meiotically independent subgenomes (homoeologues) and follow disomic inheritance but genetic studies in allopolyploids are still complicated because some of the loci are shared between homoeologues and resulted in multiple alleles during PCR amplification (Van Dijk et al., 2012).

The problem of multiple alleles was encountered in our study performed for population genetic analysis in an Indian bamboo species, Dendrocalamus hamiltonii (Meena et al., 2019). This species belongs to the lineage palaeotropical woody bamboos whose genome is allohexaploid (2n = 6x = 72) containing genomes of different ancestral lineages (Thakur et
The study was conducted to characterize the metapopulation of *D. hamiltonii* using microsatellite markers. In the study a total of 335 individual clumps belonging to 19 subpopulations were subjected to genotyping with 17 polymorphic SSR primer sets. The populations were sampled from natural range of distribution i.e. northeastern states of India (5 populations from Mizoram, 5 from Assam, 2 from Nagaland, 1 from Meghalaya and 6 from Arunachal Pradesh). The detailed geographic information of sampling location is given in supplementary table 1. The SSR primers used in the current study were validated through cross-transferability; a total of 62 primer pairs of *D. latiflorus* (Bhandawat et al., 2014) were screened in *D. hamiltonii*, and 17 of these demonstrated polymorphic banding pattern (supplementary table 2). Among these, 10 primers gave typical banding pattern of microsatellite markers i.e. 1 or 2 alleles per individual (Figure 1a) but rest 7 exhibited more than 2 bands in single individual when the PCR products resolved in high resolution agarose gel (Figure 1b). This may be attributed to the polyploid nature of the species. The multiple alleles generated with these primer sets were also reported in the parent literature where the primers were developed for *D. latiflorus* and tested their cross-transferability in 36 other bamboo species of various ploidy level (Bhandawat et al., 2014). The problem remained persistent with the different set of SSR primers developed in bamboos (Bhandawat et al., 2015; Bhandawat et al., 2019) as well as other polyploid plant species viz. *Mercurialis perennis* (Pfeiffer et al., 2011) and *Juniperus thurifera* (Teixeira et al., 2014).

![Figure 1](image.png)

**Figure 1**

Representative gel photograph of marker genotyping with SSR primer pairs in *Dendrocalamus hamiltonii*; a and b represent the different band profile generated by two primer pair DLUGMS 56 and 57 in population DH19.

In order to understand the informativeness of the SSR markers in polyploid species by treating them as dominant marker, the band profile generated by each primer was scored in two ways; first, the banding pattern obtained with all 17 primers were scored as non-allelic data in binary matrix considering each band as individual loci; secondly, ten of these primers giving diploid like banding pattern were scored as allelic data. The raw data set of all studied SSR primer loci in three representative populations is provided as supplementary data sheet 1 and 2. The allelic and non-allelic data generated from same set of markers in sampled populations were analyzed separately for estimating the measures of population genetics, phylogenetic analysis and deciphering genetic structure using different software such as GenAIEx ver 6.5 (Peakall and Smouse, 2012) and POPGENE ver 1.32 (Yeh et al., 1999) and STRUCTURE ver 2.2 (Pritchard et al., 2000). The results obtained with both the data sets were then compared. However, the diversity measures could not be directly compared because different parameters are considered to be the key indicators in analysis with dominant and codominant marker. For instance the Shannon’s information indices (I) is the key measure for dominant data, while expected heterozygosity (He) is considered as appropriate diversity measure in codominant marker analysis. But the UPGMA clustering, Principle Coordinate Analysis and structure analysis could be performed with both the data sets and gives non-quantitative results in the form of dendrogram and plots; and thus, they could be compared.

Both the data sets generated uniform clustering pattern in UPGMA dendrogram i.e. individual populations clustered in group 1, group 2 and group 3 remained same in both the dendrogram with varying level of similarity and dissimilarity between the populations (Figure 2a and b). The clustering pattern demonstrated that two major groups viz. group 1 and 3 were obtained with the clustering of populations as per their geographical distribution; group 1 consisted of four populations of Mizoram except popDH03 which remained un-clustered with any group, and group 2 was collectively formed by the other populations of Assam, Nagaland and Arunachal Pradesh which were later separated into distinct subgroups in accordance to their geographical distribution. While in group 2, a population of Assam was separated from other Assamese populations and unexpectedly clustered with the population of Meghalaya. The group wise clustering behavior of each population has been explained in our published article (Meena et al., 2019). Structure analysis was carried out to identify the ancestral gene pools and determination of inferred ancestry of individual populations in each defined cluster. Based upon the plot of K value with delta K (manually calculated ad hoc quantity of estimated Ln probability of data) (Evanno et al., 2005), the optimum K value was determined that indicates the number of ancestral gene pool exist in sampled populations. The analysis was carried out with both the data sets to estimate the optimum K value and generated the bar plots of inferred ancestry among sub populations. In the allelic data set, K=3 and K=4 showed highest peak in the plot with delta K and hence both the K values could be used for interpretation of genetic structure among subpopulations (Figure 3a). Here in, K=4 gives more precise clustering that is also in accordance with the clustering pattern obtained in UPGMA dendrogram. The majority of populations belonging to Group 1 and 2 of UPGMA dendrogram were also defined by distinct clusters in structure bar plot but the populations of group 3 were further separated in two subclusters and here also clustering was obtained in accordance to the geographical distribution with exception of one population in each subcluster (Figure 3c). The populations of Assam and Nagaland were clustered together and separated...
from that of Arunachal Pradesh. The sub clustering of group 3 populations may be attributed to the natural barrier of river Bramputra falls between the regions. Non-allelic data also showed that K=4 is the optimum number of subpopulations and revealed bar pot with similar pattern of clustering with slight variation detected for few populations that could have been arisen due to different number of loci taken in both the data sets (Figure 3b and d). Based upon the comparative study, it can be concluded that SSR markers could be used for characterizing the polyploid species by treating it as dominant marker, and it is more advantageous than other dominant markers due to presence of large number of reproducible alleles per locus (Pfeiffer et al., 2011). Definitely the advantage of co-dominance could not be realized and the information such as heterozygosity, allelic richness etc. will be lost but the results generated here will be more reliable than the dominant markers. Therefore, in majority of studies data generated with SSR markers were analyzed by transforming allelic data into binary matrix (Teixeira et al., 2014; Bhandawat et al., 2014, 2015 and 2019). Also, the quantitative as well as non-quantitative measure calculated with non-allelic data set could be used to characterize and describe the populations.

References

Bhandawat A, Sharma V, Sharma H, Sood A, Sharma RK (2014) Development and crosstransferability of functionally relevant micro-satellite markers in Dendrocalamus latiflorus and related bamboo species. J. Genet. 93: e48-55. https://doi.org/10.1007/s12041-014-0377-9

Bhandawat A, Sharma V, Singh P, Seth R, Nag A, Kaur J, Sharma RK (2019) Discovery and Utilization of EST-SSR Marker Resource for Genetic Diversity and Population Structure Analyses of a Subtropical Bamboo, Dendrocalamus hamiltonii. Biochem Genet. 57(5):652-672. https://doi.org/10.1007/s10528-019-09914-4

Bhandawat A, Singh G, Raina AS, Kaur J, Sharma RK (2015) Development of genic SSR marker resource from RNASeq data in Dendrocalamus latiflorus. J. Plant Biochem. Biotechnol 25:179-190. https://doi.org/10.1007/s13562-015-0323-9

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14(8): 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Garcia-Verdugo C, Calleja JA, Vargas P, Silva L, Moreira O, Pulido F (2013) Polyploidy and microsatellite variation in the relict tree Prunus lusitanica L.: how effective are refugia in preserving genotypic diversity of clonal taxa? Mol. Ecol. 22(6):1546-1557. https://doi.org/10.1111/mec.12194

Guo ZH, Ma PF, Yang GQ, et al. (2019) Genome Sequences Provide Insights into the Reticulate Origin and Unique Traits of Woody Bamboos. Mol. Plant. 12(10): 1353-1365. https://doi.org/10.1016/j.molp.2019.05.009

Meena RK, Bhandhari MS, Barthwal S, Ginwal HS (2019) Genetic diversity and structure of Dendrocalamus hamiltonii natural metapopulation: a commercially important bamboo species of northeast Himalayas. 3 Biotech 9: 60. https://doi.org/10.1007/s13205-019-1591-1.
Obbard DJ, Harris SA, Pannell JR (2006) Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. Heredity (Edinb) 97:296-303. https://doi.org/10.1038/sj.hdy.6800862.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetics software for teaching and research-an update. Bioinformatics 28(19): 2537-2539. https://doi.org/10.1093/bioinformatics/bts460.

Pfeiffer T, Roschanski AM, Pannell JR, Korbecka G, Schnittler M (2011) Characterization of microsatellite loci and reliable genotyping in a polyploid plant, Mercurialis perennis (Euphorbiaceae). J. Hered. 102: 479-88. https://doi.org/10.1186/s13104-017-2700-z.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2): 945–955.

Sampson JF, Byrne M (2012) Genetic diversity and multiple origins of polyploidy Atriplex nummularia Lindl. (Chenopodiaceae). Biol. J. Linn. Soc. 105:218-230. https://doi.org/10.1111/j.1095-8312.2011.01787.x.

Teixeira H, Rodriguez-Echeverria S, Nabais C (2014) Genetic Diversity and Differentiation of Juniperus thurifera in Spain and Morocco as determined by SSR. PLoS ONE 9(2): e88996. https://doi.org/10.1371/journal.pone.0088996.

Thakur A, Barthwal S, Ginwal HS (2015) Genetic diversity in bamboos: conservation and improvement for productivity. In Kaushik S, Singh YP, kumar D, Thapaliyal M and Barthwal S (eds) Bamboos in India (ed.), ENVIS Centre on Forestry, Dehradun, pp. 131-146.

Van Dijk T, Noordijk Y, Dubos T, Bink MCAM, Meulenbroek BJ, Visser RGF, Van de Weg E (2012) Microsatellite allele dose and configuration establishment (MADCE): an integrated approach for genetic studies in allopolyploids. BMC Plant Biol. 12: 25. https://doi.org/10.1186/1471-2229-12-25.

Yeh FC, Yang RC, Boyles TBJ, Ye ZH, Mao JX (1999) POPGENE Version 1.32: Microsoft Window-Based Freeware for Population Genetics Analysis, Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton.