Indian Journal of Ecology (2023) 50(5): 1315-1320
DOI: https://doi.org/10.55362/IJE/2023/4053

Manuscript Number: 4053
NAAS Rating: 5.79

Detection of Plastid-based SNPs to Resolve *Bambusa tulda-longispiculata-nutans-teres* Complex in Bamboo Taxonomy

Aditya Kumar, Sanjay Singh[1], Nuzhat Bano and Hareram Sahoo

Genetics and Tree Improvement Division, ICFRE-Institute of Forest Productivity, Ranchi-835 303, India
[1]ICFRE-Eco-Rehabilitation Centre, Prayagraj-211 002, India
E-mail: aditya@icfre.org, aditya9678@gmail.com

***Abstract*:** The SNP barcodes are a viable option for identifying cultivars or species, hence to find single nucleotide polymorphisms (SNPs) that can distinguish the *Bambusa tulda-longispiculata-nutans-teres* complex, the plastid genes *rbcL* (ribulose bi-phosphate carboxylase), *matK* (maturase K), and *rpoC1* (RNA polymerase C1) were targeted. Through amplification of these plastid genes and sequencing of amplicons, alignment, and mining of the sequence, SNPs were discovered using NovoSNP software. Over 2045 bp sequenced plastid DNA, seven SNPs were identified in targeted three barcode genes. Three SNPs were found in the *matK* gene, which was of 868 bp length, whereas two SNPs were found in each of the *rbcL* gene's 730 bp and *rpoC1*'s 447 bp lengths. Based on that, discovered that *B. nutans* differed from all other species in terms of SNPs present in the *rbcL* and *rpoC1* genes, but *B. longispiculata* was determined to be distinct from all three species with respect to SNPs present in the *matK* gene. Out of seven identified SNPs, the unique heterozygous nucleotide A/G at locus 520G/A and the nucleotide 'C' at locus 713 bp were observed in *Bambusa longispiculata,* which were absent in all the other three species. These identified SNPs can be utilized to differentiate these taxonomically closely related bamboo species.

***Keywords*:** Amplification, Bamboo complex, Plastid genes, Sequencing, SNP discovery

Bamboo is a woody perennial grass that belongs to the subfamily Bambusoideae within the family Poaceae. It is one of the fastest growing plants (Singh et al 2022). It exhibits 70 genera and more than 1200 species worldwide (Dajun and Shao-Jin 1987). Based on flowering, 67 genera of woody bamboos have been placed in the nine sub-tribes (Dransfield and Widjaja 1995 and Li 1997). India is the second-largest reservoir of bamboos with a total of 136 bamboo species belonging to 23 genera (Hauchhum and Singson 2019). Most bamboo species flower once in a lifetime and then die. Its flowering time is uncertain and varies from 1 year to 120 years, depending on the species. The uncertainty in flowering of bamboo species makes it very difficult to identify closely related bamboo species taxonomically based on morphological characters. The taxonomic identification of bamboo species based on their morphological characteristics is limited as these characters are highly influenced by environmental factors. Gao et al (2016) suggested that SNP barcodes are highly reproducible, high-throughput, and a good choice for cultivar or species identification. Therefore, it is needed to develop species-specific molecular markers to distinguish bamboo species at the genomic level. Lots of work related to genetic diversity, genetic structure, and phylogenetic relationships among bamboo species has been carried out using various molecular marker systems, v*iz.,* RFLP, ISSR, SSR, AFLP, and SCAR (Nayak and Rout 2005, Barkley et al 2005, Bhattacharya et al 2006, Sun et al 2006, Das et al 2007, Lin et al 2009). However, reports on SNP discoveries that can assist in the taxonomic classification of bamboo species are scant. Therefore, the present study was carried out to discover the SNPs in the plastid genome of the *Bambusa tulda-longispiculata-nutans-teres* complex through the amplification and sequencing of the plastid genes viz. *rbcL, matK,* and *rpoC1*. The obtained SNPs can be utilized as molecular tools for taxonomic identification of these bamboo species.

MATERIAL AND METHODS

Sample collection: To carry out the study, seven samples of *Bambusa tulda* (BT), nine samples of *Bambusa nutans* (BN), and one sample each of *Bambusa teres* (BTE) and *Bambusa longispiculata* (BL) were utilized. These samples were collected from different geographical locations in the country (Table 1).

DNA extraction and quantification: Total genomic DNA was extracted and purified from the young leaves of bamboo as per Vinod (2004). The quality and quantity of extracted genomic DNA were assessed using a bio photometer and gel electrophoresis, respectively. The UV-based bio photometer plus was used to measure the optical density (OD) of DNA samples at 260 nm and 280 nm, and the A_{260}:A_{280} ratio was

worked out for the purity index. Simultaneously, DNA samples were electrophoresed in 0.8% agarose gel (XcelGen®, #XGA) at a constant voltage of 80 V with 0.05 LmL-1 (stock-10 mg mL^{-1}) ethidium bromide (Himedia®, #MB071) in 0.5X TBE (Tris, boric acid, and EDTA) buffer for 1 h.

Amplification of plastid-based candidate genes: The pre-designed universal primers (Table 2) for *rbcL* (ribulose bi-phosphate carboxylase), *matK* (maturase K), and *rpoC1* (RNA polymerase C1) genes were synthesized through Xcelris Labs Ltd. and subjected to amplify the selected candidate genes in a panel of four bamboo species using gradient PCR and a basic PCR. A 20 µL reaction mixture was prepared by adding 50 ng of template DNA, 0.2 µM of each primer (Xcelris), 0.2 mM of dNTPs (Fermentas®, #R0192), 2.5 mM of $MgCl_2$(ThermoFisher #AB0359), and 2U Taq DNA polymerase (ThermoFisher #EP0712) with 1X green buffer (ThermoFisher), and the mixture was subjected to amplify in the gradient PCR at 94 °C for 5 minutes for initial denaturation, followed by denaturation, annealing, and extension for 35 cycles at 94 °C for 45 seconds, 49-60°C for 30 seconds at a gradient of 1°C for 45 seconds, 49-60°C for 30 seconds at a gradient of 1°C for 1 minute, and the final extension was performed at 72°C for 5 minutes. After optimization, a 50 µL reaction mixture was prepared by adding the chemicals at the above concentration, and the mixture was subjected to PCR amplification at 94 °C for 5 minutes, followed by a three-step cycle of denaturing at 94 °C for 4 seconds, annealing at 48 °C for *rbcL* and 53 °C for *matK* and *rpoC1* for 30 seconds, and extension at 72 °C for 1 minute. The cycle was repeated 35 times before a final

Table 2. PCR primers used for amplification of bamboo species

Primer	Sequence	T_m (°C)	T_a (°C)
rbcL 1F	ATGTCACCACAAACAGAAAC	56	48
rbcL724 R	TCGCATGTACCTGCAGTAGC	66	48
matK X F	TAATTTACGATCAATTCATTC	52	53
matK 5 R	GTTCTAGCACAAGAAAGTCG	58	53
rpoC1-2 F	GGCAAAGAGGGAAGATTTCG	68	53
rpoC1-3 R	TGAGAAAACATAAGTAAACGGGC	64	53

Table 1. Collection of Bamboo samples from wide geographical locations in Jharkhand and West Bengal for DNA extraction and SNP discovery

Plant code	Location	latitude	longitude	Elevation	Humidity/ Rain fall
Bambusa tulda					
BT-1	Jalpaiguri, WB	26.5215° N	88.7196° E	89 m	63% humidity, 3242 mm
BT-13	Lalgutwa, Ranchi, JH	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
BT-3	Lalgutwa, Ranchi, JH	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
BT-5	Lalgutwa, Ranchi, JH	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
BT-6	Jalpaiguri, WB	26.5215° N	88.7196° E	89 m	63% humidity, 3242 mm
BT-7	Lalgutwa, Ranchi, JH	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
BT-9	Lalgutwa, Ranchi, JH	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
Bambusa nutans					
BN-9	Sukna, Siliguri, Darjeling	26.6841 ° N	88.3506 °E	5120 m	68% humidity, 3047 mm
BN-4	Jagarnathpur, Mahulia, Ghatsila	22.6338° N	86.4342°E	103 m	85% humidity, 1241 mm
BN-10	Jamuney, Daltonganj	24.0465° N	84.0768°E	215 m	1174 mm,, 79% humidity
BN-11	Sukna, Siliguri, Darjeling	26.6841 ° N	88.3506 °E	5120 m	68% humidity, 3047 mm
BN-13	Sukna, Siliguri, Darjeling	26.6841 ° N	88.3506 °E	5120 m	68% humidity, 3047 mm
BN-15	Sukna, Siliguri, Darjeling	26.6841 ° N	88.3506 °E	5120 m	68% humidity, 3047 mm
BN-3	Kakrisol, Chakulia	22.4737° N	86.7465° E	115 m	83% humidity, 1294 mm
BN-5	Lalgutwa, Ranchi	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
BN-7	Lalgutwa, Ranchi	23.3441° N	85.3096° E	651 m	56% humidity, 1430 mm
Bambusa teres					
BTE	BSI, Kolkata	22.5726° N	88.3639° E	9.14 m	90% humidity 1582 mm
Bambusa longispiculata					
BL	Assam, Jorhat	26.7509	94.2037	116 m	95% humidity, 2244 mm

BT = Bambusa tulda, BN = Bambusa nutans, BTE = Bambusa teres, BL = Bambusa longispiculata

extension at 72 °C for 5 minutes. The amplified PCR products were electrophoresed for 90 minutes on 1.5% agarose (XcelGen®, #XGA), containing 0.5 μgmL^{-1} ethidium bromide (Himedia®, #MB071), at 80 V.

DNA sequencing and SNP discovery: The single prominent band obtained through *rbcL*, *matK,* and *rpoC1* primers in a set of bamboo germplasm was subjected to sequence by the di-deoxy chain termination method in an automated DNA sequencer (3500 Genetic Analyzer of AB applied biosystem) through Sanger's method. The obtained sequenced data was blasted against the NR (non-redundant) data base in NCBI (https://www.ncbi.nlm.nih.gov/) to find the similarity and identity of the amplified product. The sequenced data of the *rbcL, matK, and rpoC1* genes were aligned separately, and SNPs were searched through NovoSNP software (Weckx et al. 2005). The species-wise SNP table was generated to find the species-specific SNPs to resolve the bamboo taxonomy of four closely related bamboo species, *viz., Bambusa tulda-longispiculata-nutans-teres* complexes. The images of gel electrophoresis amplicons were processed through the programme Gene System. The size or weight of the amplicons was determined using the programme Alpha Ease 4.0. The sequenced data of each gene was validated manually on Finch TV, and the noisy peak at both ends of the nucleotide was trimmed and saved. The forward and reverse sequence reads of a gene were combined in the Bioedit programme and a gene contig was prepared. Sequenced data of a gene from all accessions were individually aligned in Cluster W and the Bioedit software (Fig. 1), and SNPs were discovered at a species level through NovoSNP software. The discovered SNPs were used to generate a phylogenetic tree by applying a maximum likelihood model with 600 bootstrap values in MEGA 7 software.

RESULTS AND DISCUSSION

In the present study, SNPs were searched under plastid barcode genes using NovoSNP software to find species-specific SNPs that could distinguish four bamboo species in *Bambusa tulda-longispiculata-nutans-teres* complexes. The 900 bp, 750 bp, and 500 bp long plastid-based *matK*, *rbcl,* and *rpoC1* genes were amplified in collected accessions of these bamboo species. The amplified amplicons were sequenced and blasted against the data base in NCBI. The eighteen accessions of four bamboo species showed 100% sequencing success, with a robust sequencing peak (Fig. 2). The BLAST result revealed that the *matK* gene of size 901 bp showed 100% query cover, a 0.0 E-value, and 99.89% identity with the *matK* gene of other bamboo species. The *rbcL* gene of size 730 bp showed 100% query cover, a 0.0 E-value, and 99.45% identity with the *rbcL* gene of *Bambusa variostriata* and other species. The *rpoC1* gene of size 445 bp showed 100% query cover, a 0.0 E-value, and 99.10% identity with the *rpoC1* gene of other bamboo species. The sequence results of *all three target* genes were aligned separately, and SNPs were searched. *A total* of seven SNPs were detected in the three genes covering 2045 bp of plastid DNA. The *matK* gene of size 868 bp showed three SNPs (Table 3), whereas the *rbcL* gene of size 730 bp and the *rpoC1* gene of size 447 bp showed two SNPs each (Table 4, 5). The *matK* gene of size 868 bp showed three SNPs where, at locus 520A/G, the base G (guanine) was predominantly present in BTE, BN, and BT, and heterozygous base A/G was present in BL. Similarly, at locus 700T/C, the base T

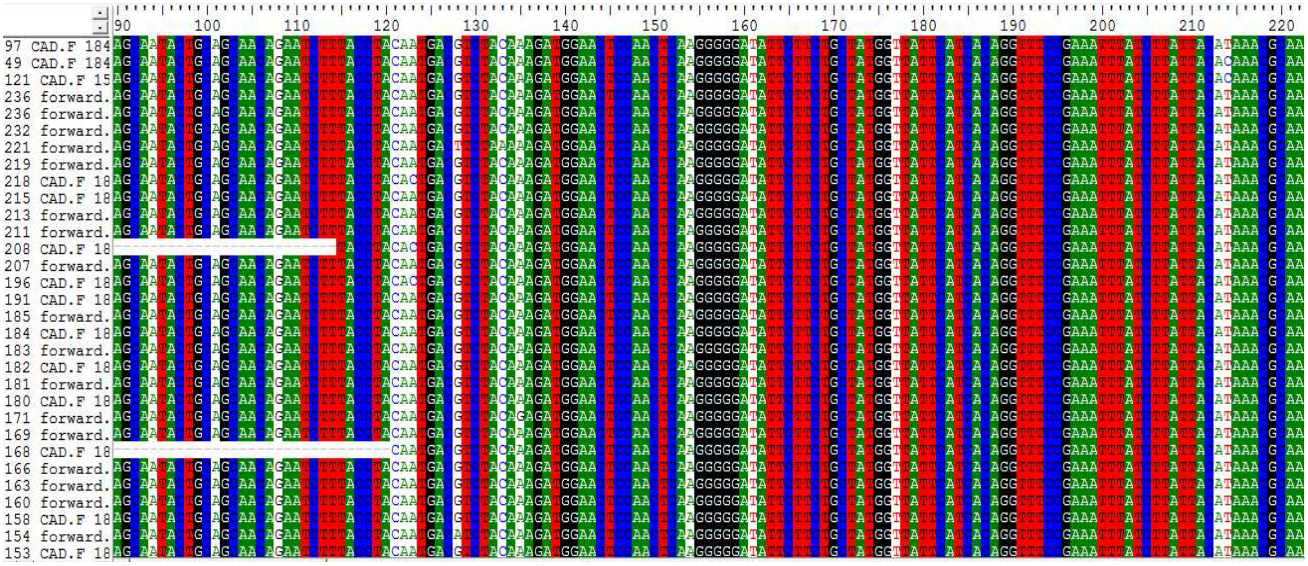

Fig. 1. Alignment of sequenced data in Bioedit software

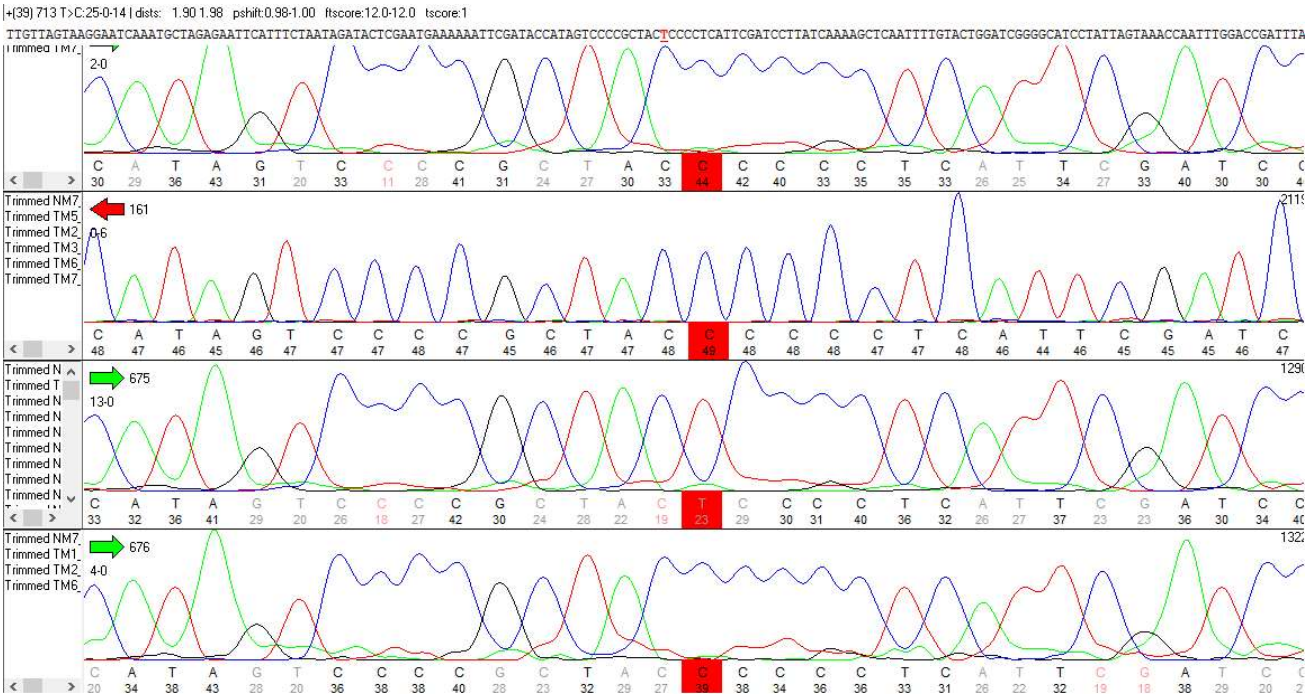

Fig. 2. Aligned peak of sequenced data in NovoSNP software. The nucleotide with red mark showed single nucleotide polymorphism

Table 3. SNPs present in 868 bp long *matK* gene in studied bamboo species

Bamboo species	No. of accessions	Read no. (Replication)	Identified SNPs		
			Loci 520	Loci 700	Loci 713
B. teres	1	4	G	T	T
B. longispiculata	1	4	A/G	C	C
B. nutans	9	25	G	T	T
B. tulda	7	14	G	T	C (1T)

Table 4. SNPs present in 730 bp long *rbcL* gene in studied bamboo species

Bamboo species	No. of accessions	Read no. (replication)	Identified SNPs	
			Loci 418A/C	Loci 667C/A
B. teres	1	4	A	C
B. longispiculata	1	4	A	C
B. nutans	9	38	5C, 4A	5A, 4C
B. tulda	7	36	7A	7C

Table 5. SNPs present in 447 bp long *rpoC1* gene in the studied bamboo species

Bamboo species	No. of accessions	Read no. (replication)	Identified SNPs	
			Loci 32A/G	Loci 77 A/G
B. teres	1	4	A	C
B. longispiculata	1	4	A	C
B. nutans	9	14	4A, 5A/G	5A, 4A/G
B. tulda	7	18	7A	7A

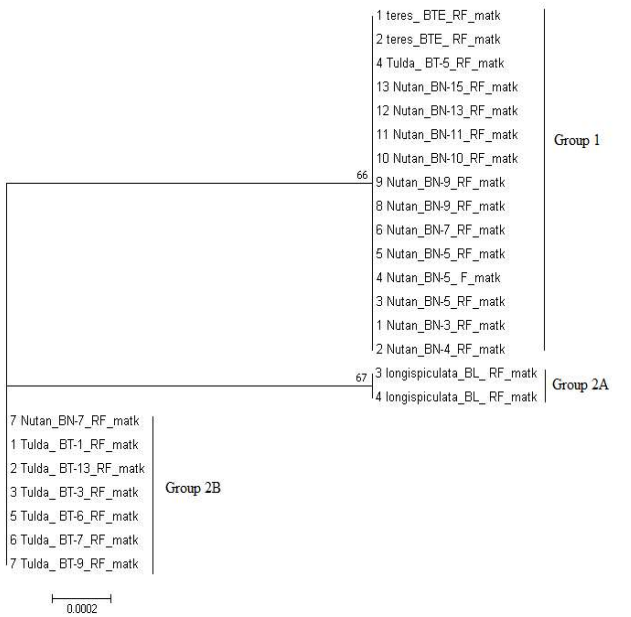

Fig. 3. Phylogenetic tree generated using SNPs variation identified under *matK* gene in four bamboo species

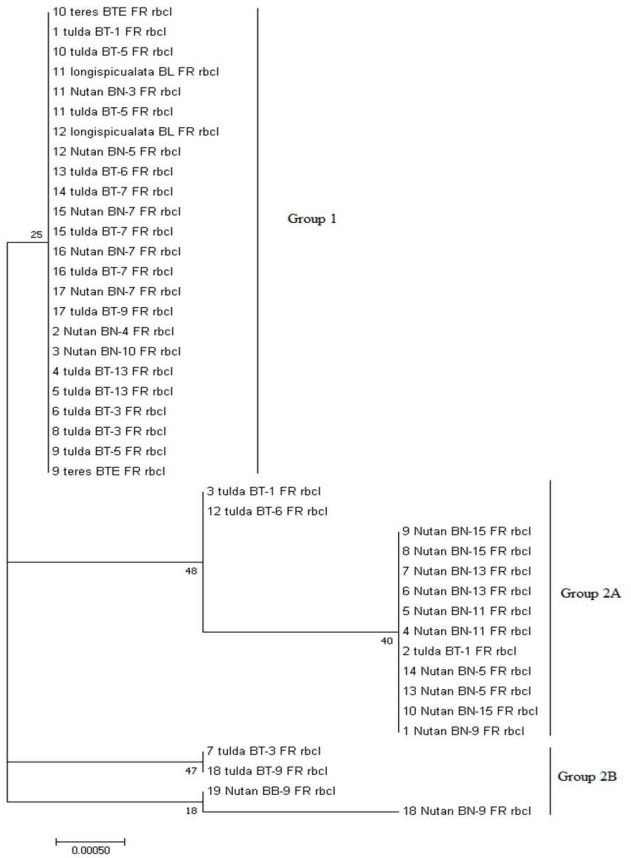

Fig. 4. Phylogenetic tree generated using SNPs variation identified under *rbcL* gene in four bamboo species

(thymine) was predominantly present in BTE, BN, and BT, and the base C was present in BL. In the case of locus 713T/C, the base T was present in BTE and BN, whereas the base C (cytosine) was predominantly present in BT and BL, respectively. The *rbcL* gene of size 730 bp showed two SNPs at locus 418A/C: base A was present in BTE, BL, BT, and BN, whereas base 'C' was only present in a few accessions of BN. Similarly, at locus 667C/A, the base C was predominantly present in BTE, BL, and BT. In BN, base A was present in five studied accessions and base C in four accessions at locus 667C/A. The *rpoC1* gene of size 447 bp showed two SNPs: at locus 32A/G, base 'A' was present in BTE, BL, BT, and BN, whereas a heterozygous base A/G was predominantly present in BN. Similarly, at locus 77A/G, the base 'A' was present in BTE, BL, BT, and BN, and a heterozygous base A/G was present in BN alone.

Based on detected seven SNPs, *B. longispiculata* was found to be different from all three species with respect to SNPs present in the matK gene (Fig. 3), whereas *B. nutans* was found to be different from all other species with respect to SNPs present in the *rbcL* and *rpoC1* genes (Fig. 4, 5). Out of seven identified SNPs, the unique heterozygous nucleotide A/G at locus 520G/A and the nucleotide 'C' at locus 713 bp were observed in *Bambusa longispiculata,* which were absent in all the other three species. In wheat, Gao et al. (2016) selected 43 SNPs from an array of 9000 SNPs that help in discriminating different varieties of wheat. Similarly, in capsicum, SNPs derived from nuclear and cytoplasmic DNA were used for taxonomic classification of distinct species of capsicum (Jeong et al. 2010). In bamboo, this is perhaps the first effort to identify SNPs to differentiate these closely related four species.

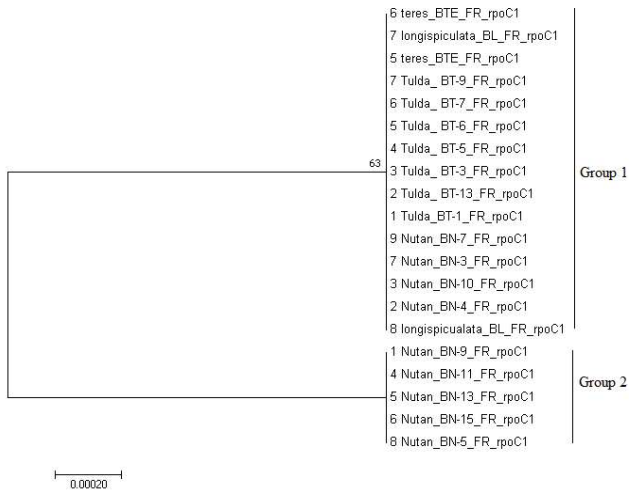

Fig. 5. Phylogenetic tree generated using SNPs variation identified under *rpoC1* gene in four bamboo species

CONCLUSION

The study demonstrates that the taxonomically closed *Bambusa tulda-longispiculata-nutans-teres* complex can be distinguished through single nucleotide polymorphism (SNP) discovery using plastid genes, i.e., *rbcL* (ribulose bi-phosphate carboxylase), *matK* (maturase K), and *rpoC1* (RNA polymerase C1). The SNPs detected in the *matK* gene can differentiate *B. longispiculata* from the other three species, whereas the SNPs detected in the *rbcL* and *rpoC1* genes can differentiate *B. nutans* from the other three species. These species-specific SNPs need to be validated in a large set of germplasm.

Acknowledgements

Authors are highly thankful to the ICFRE, Dehradun for the financial assistance to conduct the present study. The authors are also thankful to the Director, ICFRE-Institute of Forest Productivity, Ranchi for allowing to utilize institutional facilities to carry out the study.

REFERENCES

Barkley NA, Newman ML, Wang ML, Hotchkiss MW and Pederson GA 2005. Assessment of the genetic diversity and phylogenetic relationships of a temperate bamboo collection by using transferred EST-SSR markers. *Genome* **48**(4): 731-737.

Bhattacharya S, Das M, Bar R and Pal A 2006. Morphological and molecular characterization of Bambusa tulda with a note on flowering. *Annals of Botany* **98**(3): 529-535

Dajun W and Shao-Jin S 1987. *Bamboos of China*. Timber Press, Portland, Oregon, USA.

Das M, Bhattacharya S, Basak J and Pal A 2007. Phylogenetic relationships among the bamboo species as revealed by morphological characters and polymorphism analyses. *Biologia Plantarum* **51**: 667-672.

Dransfield S and Widjaja EA 1995. *Plant resources of south-east Asia No. 7: Bamboos*. Leiden, Backhuys.

Gao L, Jia J and Kong X 2016. A SNP-based molecular barcode for characterization of common wheat. *PLoS One* **11**(3): e0150947.

Hauchhum R and Singson MZ 2019. Assessment of above ground biomass and carbon storage in Bamboo species in sub-tropical bamboo forests of North-East India. *Indian Journal of Ecology* **46**(2): 358-362.

Jeong HJ, Jo YD, Park SW and Kang BC 2010. Identification of Capsicum species using SNP markers based on high resolution melting analysis. *Genome* **53**(12): 1029-1040.

Li D 1997. The flora of China Bambusoideae project-problems and current understanding of bamboo taxonomy in China. In: Chapman GP (ed) *The bamboos.* Academic Press, London, pp 61-81.

Lin XC, Ruan XS, Lou YF, Guo XQ and Fang W 2009. Genetic similarity among cultivars of *Phyllostachys pubescens*. *Plant Systematics and Evolution* **277**: 67-73.

Nayak S and Rout GR 2005. Isolation and characterization of microsatellites in *Bambusa arundinacea* and cross species amplification in other bamboos. *African Journal of Biotechnology* **4**(2): 151-156.

Singh NS, Upadhyay KK and Tripathi SK 2022. Leaf litter decomposition of *Melocanna baccifera* (Roxb.) Kurz under field and Laboratory Microcosm in Northeast India. *Indian Journal of Ecology* **49** (1): 114-118.

Sun Y, Xia N and Stapleton CM 2006. Relationships between Bambusa species (*Poaceae, Bambusoideae*) revealed by random amplified polymorphic DNA. *Biochemical Systematics and Ecology* **34**(5): 417-423.

Vinod KK 2004. Total genomic DNA extraction, quality check and quantitation. In: Proceedings of the *training programme on Classical and modern plant breeding techniques: A hands on training,* Tamil Nadu Agricultural University, Coimbatore, pp 109-121.

Weckx S, Del-Favero J, Rademakers R, Claes L, Cruts M, De Jonghe P, Van Broeckhoven C and De Rijk P 2005. NovoSNP, a novel computational tool for sequence variation discovery. *Genome Research* **15**(3): 436-442.

Received 14 April, 2023; Accepted 22 August, 2023