

Bioinformatic Testing of Genetic Variation in Plastid Gene matK for Assessing its usage in Phylogenetic Studies of Bambusa Species

Aishwarya Singh¹, Srishti Shekhar², Supreeti Das³ and Jasmeet Kaur Abat⁴

^{1,2,4}Department of Botany, Gargi College, University of Delhi

³Department of Physics, Gargi College, University of Delhi

E-mail: ¹aishwarya.singh129@gmail.com, ²srishti.shekhar05@gmail.com,

³supreeti25@gmail.com, ⁴jasmeet29@yahoo.com

Abstract—Bamboos are large perennial grasses of temperate and tropical forests worldwide, consisting of two major growth forms, the woody bamboos (tribes Arundinarieae and Bambuseae), and herbaceous bamboos (tribe Olyreae). Bamboos have great economic significance having varied uses ranging from food, to raw material for furniture and for medicinal purposes. Bambusoideae is one of the largest subfamilies in the grass family Poaceae containing around 115 genera with more than 1400 species. Deciphering evolutionary relationships among the wide range of species has always been difficult. Many coding as well as non-coding regions of chloroplast DNA (cpDNA) are often used to resolve plant phylogenies. The matK gene of chloroplast is 1500 bp long, located within the intron of the trnK and codes for maturase like protein, which is involved in Group II intron splicing. MatK gene contains high substitution rates within the species and is a well-established candidate to study plant systematics and evolution. Keeping this in view, the objective of the present investigation was to assess variation and phylogenetic relationships of Bambusa species by using the matK gene sequences. For the present study, sequences of matK from 11 species of Bambusa were obtained from GenBank, NCBI and were analyzed for variations using Multiple Sequence Alignment (MSA) tool (CLUSTAL OMEGA). Length of the sequences retrieved ranged from 1527 to 1754 bp. MSA revealed very less variations and that to only at some sites while some species like Bambusa vulgaris and B. bambos showed no sequence variation. The study underscored that matK is a conserved gene with slow rate of substitution in some regions of the gene. Thus suggesting that matK could be used at above species level rather than at species level for assessing phylogenetic variations.

1. INTRODUCTION

Bambusoideae is one of the major lineages within the grass family Poaceae containing around 115 genera with more than 1400 species. Bamboos are large perennial grasses of temperate and tropical forests worldwide, consisting of two major growth forms, the woody bamboos (tribes Arundinarieae and Bambuseae), and herbaceous bamboos (tribe Olyreae) [1]. Bamboos are concentrated in Asia, Latin America and Africa. Most of them grow in tropical and subtropical regions with good monsoon climate and heat

conditions while some are also found in temperate and even sub-arctic regions. There is a wide range in size of bamboos, varying from small annuals to giant timber bamboo. *Dendrocalamus sinicus* is the largest bamboo in the world with around 12 inches in diameter and reaching a towering 100m in height while *Raddiella vanessae*, is only a few short centimeters tall. India, the second largest producer of bamboo in the world has rich diversities of bamboos with about 130 species spreading over 18 genera.

Bamboos are of significance economic importance having varied uses ranging from food, to raw material for furniture and for medicinal purposes. Soft bamboo shoots, stems, and leaves are used as food source. All parts of bamboo have medicinal value. Among the diseases treated with extracts from the species are dysentery, diarrhea in babies, pile and epilepsy [2]. Bamboo is used in forests to maximize the production of PhytOC which has the potential to result in considerable quantities of securely bio-sequestered carbon [3]. Giant bamboo plantations are potential lignocellulosic feedstock for the production of second generation bioethanol [4]. Bamboo culms were found to be environmentally and financially better in comparison to other building materials like steel, concrete and timber. Bamboo is also used as the source of raw material for papermaking, textiles, basketry, matting, rope, house construction, furniture, bridges and fishing equipment.

The estimation of phylogenetic relationships plays a key role in understanding evolution and is an essential component of evolutionary biology. Chloroplast DNA diversity studies are routinely used for deciphering evolutionary relationships among the wide range of species. Both non-coding and coding regions are applied to the study of phylogenetic relationships among closely related taxa [5]. Maturase K (matK) and ribulose 1,5-biphosphate carboxylase/oxygenase (rbcL) genes of chloroplast genome are routinely used to assess the phylogenetic relationships amongst different plants, recently a

combination of these was used for inferring phylogenetic relationships of Indian Citron (*Citrus medica* L.) [6]. The matK gene of chloroplast is 1500 bp long, located within the intron of the trnK and codes for maturase like protein, which is involved in Group II intron splicing. MatK gene contains high substitution rates within the species and is a well-established candidate to study plant systematics and evolution [7]. Keeping this in view, the objective of the present investigation was to assess variation and phylogenetic relationships of *Bambusa* species by using the matK gene sequences.

2. MATERIALS AND METHODS

2.1 Sequence Retrieval and Bioinformatic analysis

Sequences of matK for available *Bambusa sp* were retrieved from GenBank, NCBI (<http://www.ncbi.nlm.nih.gov>). A total of 11 matK gene sequence of *Bambusa* species were retrieved (Table1). Length of the sequences retrieved ranged from 1527 to 1754 bp. For phylogenetic analysis, the sequences were subjected to Multiple Sequence Alignment (MSA) using Clustal omega available at EBI (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). For assessing interspecific relationships, phylogenetic tree was constructed using the MSA.

Table 1: List of matK sequences of *Bambusa sp.* with Accession Numbers and Size of the gene

S.no.	Name of the species	Accession no.	Size
1	<i>Bambusa vulgaris</i>	JX428387.1	1539bp
2	<i>Bambusa spp.</i>	KC734453.1	1527bp
3	<i>Bambusa bambos</i>	EU434243.1	1754bp
4	<i>Bambusa spp.</i>	KC734447.1	1527bp
5	<i>Bambusa chungii</i>	KC734450.1	1527bp
6	<i>Bambusa dolichomerithalla</i>	HM448935.1	1754bp
7	<i>Bambusa ralignensis</i>	EU434245.1	1754bp
8	<i>Bambuseae spp.</i>	EU434262.1	1754bp
9	<i>Bambusa oliveriana</i>	EU434247.1	1754bp
10	<i>Bambusa beecheyana</i>	EU434244.1	1754bp
11	<i>Bambusa multiplex</i>	EF125166.1	1600bp

3. RESULTS AND DISCUSSION

Comparative and Bioinformatic analysis of matK region was done for assessing variation and phylogenetic relationship of *Bambusa*. MatK sequences of 11 *Bambusa* species were retrieved from GenBank, NCBI. Length of the sequences retrieved ranged from 1527 to 1754 bp with *B. chungii* and *B. spp* having the smallest sequence (1527bp) while *B. dolichomerithalla*, *B. ralignensis*, *B. oliveriana*, *B. beecheyana* and *B. bambos* having the longest sequence (1754bp). These retrieved matK sequences were subjected to Multiple Sequence Alignment (MSA) (Fig. 1). MSA revealed that the

sequences were highly conserved and some species like *B. vulgaris* and *B. bambos* showed no sequence variation.



Fig. 1: Region of sequence variations (in boxe) in Multiple Sequence Alignment of matK gene sequence from 11 *Bambusa sp.* *denotes the conserved region. nucleotides

Phylogenetic analysis done using the above obtained MSA, revealed that the sequences of matK region were not able to discriminate different species of *Bambusa* (Fig. 2). This could be attributed to very less sequence variations within matK regions of selected species of *Bambusa*. Present study underscored that matK is a conserved gene with slow rate of substitution in some regions of the gene. Thus suggesting that matK could be used at above species level rather than at species level for assessing phylogenetic variations. Similar results were obtained in a study by Das et al., 2013 on Indian bamboos where they showed that matK sequence based plant DNA barcoding failed to identify *Bambusa* species from Northeast India. It was suggested that matK failed to provide species specific marker in *Bambusa* due to interspecies hybridization, introgression and polyploidization [8].

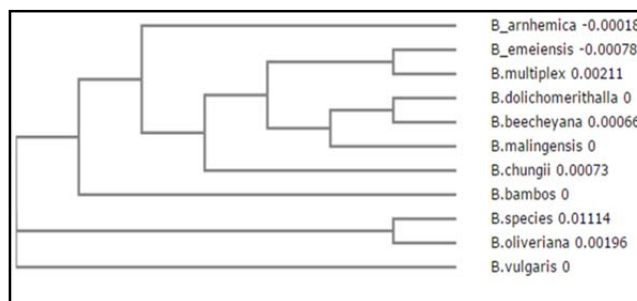


Fig. 2: Phylogenetic tree constructed using matK gene sequences of 11 *Bambusa* species

4. ACKNOWLEDGEMENTS

The research work was supported under Star College Scheme to Gargi College by Department of Biotechnology and Delhi University Innovation Project GC-305.

REFERENCES

- Li D.Z., Wang Z.P., Zhu Z.D., Xia N.H., Jia L.Z., Guo Z.H., Yang G.Y., Stapleton C., Bambuseae in Flora of China. Wu Z.Y., Raven P.H., editors. Vol. 22. Beijing & St. Louis: Science Press & Missouri Botanical Garden Press; 2006.

- [2] Ogunjinmi, A. A., H. M. Ijeomah, A. A. Aiyeloja., "Socio-economic importance of bamboo (*Bambusa vulgaris*) in Borgu local government area of Niger State, Nigeria", *Journal of Sustainable Development in Africa*, 10, 2009, 284-289.
- [3] Parr, J., Sullivan, L., Chen, B., Ye, G., Zheng, W., "Carbon bio-sequestration within the phytoliths of economic bamboo species", *Global Change Biology*, 16, 2010, 2661-2667.
- [4] Garcia-Aparicio, M., Parawira, W., Van-Rensburg, E., Diedericks, D., Galbe, M., Rosslander, C., Zacchi, G., Görgens, J., "Evaluation of steam-treated giant bamboo for production of fermentable sugars", *Biotechnology progress*, 27, 2011, 641-649.
- [5] Widmer, A., Baltisberger, M., "Extensive intraspecific chloroplast DNA (cpDNA) variation in the alpine *Draba aizoides* L. (Brassicaceae): haplotype relationships and population structure", *Molecular Ecology*, 8(9), 1999, 1405-15.
- [6] Uchoi, A., Malik, S.K., Choudhary, R., Kumar, S., Rohini, M.R., Pal, D., Ercisli, S., Chaudhury, R., "Inferring Phylogenetic Relationships of Indian Citron (*Citrus medica* L.) based on rbcL and matK Sequences of Chloroplast DNA", *Biochemical Genetics*, 54, 2016, 249-69.
- [7] Hilu, K., Liang, H., "The matK gene: sequence variation and application in plant systematics", *American Journal of Botany*, 84(6), 1997, 830.
- [8] Das, M.M., Mahadani, P., Singh, R., Karmakar, K., Ghosh, S.K., "MATK Sequence Based Plant DNA Barcoding Failed to Identify *Bambusa* (Family:Poaceae) Species from Northeast India", *Journal of Environment and Sociobiology*, 10, 2013, 49-54.