Journal of Applied Biological Sciences
E-ISSN: 2146-0108 17(3): 510-519, 2023
DOI: https://doi.org/10.71336/jabs.1246
Research Article

GENETIC DIVERSITY ANALYSIS OF CULTIVATED AND WILD MANGOES FROM BANGLADESH BASED ON PCR AMPLIFICATION PATTERN OF TANDEMLY ORGANIZED REPEAT SEQUENCES

Nusrat Sultana[1,*], Samsunnahar Oishi[1], Tanzena Tanny[2], Mohammad Abdur Rahim[3]

[1]*Jagannath University, Faculty of Life and Earth Sciences, Dhaka, Bangladesh*
[2] *National Institute of Biotechnology, Plant Biotechnology Division, Savar, Dhaka, Bangladesh*
[3] *Bangladesh Agricultural University, Department of Horticulture, Mymensingh, Bangladesh*

**Corresponding Author:*
E-mail: nusrat.bot@gmail.com

(Received 19th June 2023; accepted 15th October 2023)

ABSTRACT. *Mangifera* is an economically important genus mainly because of the delicious fruit produced by many species of this genus. Wild species available in Bangladesh of this genus are often misidentified and have ambiguous taxonomic leveling, due to the difficulty of getting the appropriate sample for plant identification. Tandem repeat diversity was analyzed from the *Mangifera* genome, to understand their impact on the genomic diversity of different species and cultivars available in Bangladesh. We use tandem repeats amplification pattern for mango diversity analysis because of the fact that satellite repeats (longer arrayed tandem repeats) are usually rapidly evolving parts of the genome, found to even diversify in closely related species, hence is useful to study genome diversity analysis. RepeatExplorer2-based bioinformatics analysis was performed to identify tandemly organized repeat sequences. Face to face PCR primers were developed from each of the consensus sequences of the identified tandem repeats sequences. Our results showed that the studied samples showed a tandem repeat-specific PCR amplification pattern that corresponds to their monomer size variation, confirming the presence of genomic organization of each repeat type. Three primer pairs targeting three tandem repeat specific clusters CL18, CL165 and CL335 showed genotype specific patterns, while CL201 and CL283 produced similar banding patterns in all the studied genotype. However, CL19 failed to follow any consistent banding pattern in all three species and cultivars except smear. Our results suggest that although cultivated *M. indica* genotypes are very similar in terms of tandem repeat diversity but the wild species collected from different locations of Bangladesh are indeed different. Therefore, in depth taxonomic investigation is necessary to solve the taxonomic ambiguity of wild *Mangifera* germplasm available in Bangladesh.

Keywords: *Mangifera indica, RepeatExplorer2, repetitive DNA, satellite DNA.*

INTRODUCTION

Genus *Mangifera* belongs to the family Anacardiaceae (also called cashew family) and consists of ca.58 species. Most of these species are distributed in tropical Asia [1]. Mango (*Mangifera indica* L.) is one of the most popular fruit crops from this genus, also known as the "king of fruits" because of its excellent aroma, color, sweetness, softness and nutritional value [1]. The center of origin and diversification of mango is South East Asia

which also includes Bangladesh. Therefore, the geographical location of Bangladesh is strategically important for mango germplasm. Around 100 different mango accessions are available in Bangladesh including wild *Mangifera* species [2]. The wild *Mangifera* species found in this country are *M. longipes* Griff., *M. sylvatica* Roxb.as well as wild *M. indica* [3-5]. Fruits and leaves of wild and cultivated mango species are rich in many nutritional and medicinal properties. Besides, wood or timber add another source of natural product from these species which is being used for plywood production [6]. However, these wild crops are often threatened due to overexploitation, environmental destruction, climate change, and land use changes. Therefore, it is very important to take the necessary steps to save these natural treasures of Bangladesh.

A significant amount of research work has been already carried out particularly on the genomics of mango such as PCR-based assessment of genetic diversity [7] genomic in-situ hybridization [8]; phylogenetic relationships of *Mangifera* species [9]; the mango SNP database [10], construction of a high definition genetic map [11]; single nucleotide polymorphism and transcriptome analysis [12]; a underway draft genome assembly [13]. The mango genome is approximately 439 Mb with 2n = 40 chromosomes [14]. In Bangladesh, mango research is mainly conducted by BARI and BFRI focused on breeding for new improved varieties, new cultivar release and other horticulture related practices [15]. Although a significant advancement has been done so far focusing on only mango genome in the current decade, the basic information on the repeat sequence evolution is still very limited considering different mango germplasm. Therefore, more study is required for crop improvement, germplasm conservation, and other systematic horticultural practice.

Repeat sequences are a major constituent of plant genome (about 30-80 %), dynamics of which in each species are directly correlated with the evolutionary time scale that species has undergone [16, 17]. Repeat sequences are normally categorized by their mode of origin, structure, distribution, size and function. The major types of repeat sequence found in plant genomes are actually transposable elements, also called selfish DNA, that have a scattered distribution throughout the genome [18]. Other major groups of repeats sequence are also known as tandem repeats based on the fact that monomer sequences are organized head to tails tandem organization. Satellite repeats and rDNA are mainly found tandemly organized in the genome of all eukaryotes [19, 20].

Repeat profiles or genomic proportions of different repeat sequences varied among different species. Therefore, species can be characterized based on their repeat profiles. For instance, depending on the plant species or family genomic proportion of satellite repeats can vary from less than 1% to more than 35% [19-23]. Generally, satellite DNA has a specific monomer unit size range from 150-180 bp to 320-360 bp, however, monomer units less or more than this range are also recorded [24, 25]. Mostly, satellite DNAs are noncoding monomer arrays, repeated several hundreds to thousands times [16]. Often satellite repeats show significant sequence similarity with other repeat sequences such as the long terminal repeats (LTR) of LTR-retrotransposons [26], MITEs [27], intergenic spacers of rDNA sequences [28], or SSR sequences or even with genic sequence. Therefore, associated sequences are often thought to be responsible for the origin and amplification of each specific type of satellite DNA. Different mechanisms are found to be responsible for these types of repeat sequence proliferation and amplification like rolling circle replication, strand slippage, unequal crossing over, or sequence duplication [16, 29].

Advancement in plant genome analysis tools reveals that tandem repeats modulate genome architecture, chromosome stability and gene regulation and thereby influence gene expression patterns [30-32]. Therefore, changes in satellite DNA directly correlated with species adaptation and environmental challenges that species have gone through [16, 30].

In this study, a research initiative will be carried out to reveal the bioinformatics analysis of repeat sequences, especially satellite repeats from the publicly available whole genome sequence data. Moreover, genomic distribution of each different satellite repeats types will be analyzed through bioinformatics and molecular analysis.

MATERIALS AND METHODS

Sample collection and preservation

A local *M. indica* L. and a wild *M. sylvatica* tree has been conserved in Botanical Garden, Dhaka University Campus. Another wild *Mangifera* species, *M. longipes* seedlings were collected from Dulahazra Safari Park, Chittagong, Chakaria, Cox's Bazar. Young leaf tissue of the tree plant has been collected and used as a study material.

Bioinformatics identification of satellite repeats monomer

Bioinformatics analysis of tandem repeats has been performed using the publicly available whole genome sequence reads. Whole genome sequence data with the sequence Id Sra SRR8281982 were uploaded in RepeatExplorer2 Galaxy instances [33,34]. Sequence reads were quality filtered with the sequence pre-processing tools available in the Repeat Explorer2 galaxy instance. Total 19999932 paired-end reads were selected for Repeat Explorer run with default parameters available in RepeatExplorer2 galaxy [33,34]. The software automatically selects 3921172 paired-end reads for repeat sequence analysis with the proportion of reads in analyzed clusters is about 33 %. The resulting output of the RepeatExplorer Galaxy instance was further inspected for satellite clusters with round graph shape and reported as satellite repeats according to the TAREAN report. The consensus sequence was extracted and used for primer designing.

Satellite specific PCR primer designing and polymerase chain reaction

For primer designing the consensus sequence of each satellite repeats and read sequences of their respective clusters were uploaded in Geneious Prime 2020 bioinformatics platform (http://www.geneious.com, [35]). Each satellite specific consensus sequence was mapped against the read sequences of their respective clusters. The most homogeneous region of the mapped reads were taken for face to face forward and reverse primer designing. Optimum Tm temperature and GC percentage were calculated for each primer pair. Length of each primer was tried to keep in the range between 16-25 bp and GC percentage as high as possible to increase the specificity of the primer. In addition, the 3' end region of each primer was tried to be constituted with either G or C to increase the specificity of the primer binding region.

Genomic DNA extraction

Genomic DNA was isolated from the young leaf tissues of each plant sample using the mini preparation CTAB method (Doyle and Doyle, 1987) with some modification

according to the protocol given by Uddin et al. 2014. For this, stock solutions were prepared beforehand such as 10% CTAB, 5 M NaCl, 0.5 M EDTA (pH 8.0), 1M Tris-HCl (pH 8.0), diluted with sterile MQ water to the working solution. The extract genomic DNA pellets were dried and dissolved in 50 microliters of TE Buffer and treated with RNaseA for 30 min at 37℃ and stored at -20 ℃.

PCR Amplification of selected mango germplasm using tandem repeat specific primer pairs

Tandem repeat specific PCR amplifications were carried out for both the genomic DNA extracted from the wild and cultivated mango genotypes. Annealing temperature for each specific primer pair was determined from the results of bioinformatics software as well as gradient PCR (Supplementary Data 1). Finally polymerase chain reactions (PCR) were performed using dream Taq polymerase and green Taq buffer (Thermo scientific). Standard PCR conditions were 94 °C for 3 min followed by 35 cycles of 94 ℃ for 50 second specific annealing temperature for 50 second, extension at 72 ℃ and final extension at 72 ℃ for 5 min. Amplified PCR products were separated by a 1.2% agarose gel, 45-50 V and 1X TAE buffer for about an hour. Well separated bands of the amplified product were visualized under UV rays after ethidium bromide staining.

RESULTS AND DISCUSSION

Bioinformatics identification of repeat sequences using RepeatExplorer pipeline

Whole genome sequence data were used for bioinformatics analysis of repeat sequences of *Mangifera indica* ssp Alphonso genome. The software automatically selected 3921172 paired-end reads for repeat sequence analysis with proportion of reads in analyzed clusters of 344 which is about 33% of total reads. Based on the graph shape of each cluster, TAREAN analysis, six clusters were recorded as tandem repeat clusters. The naming of the satellite clusters were performed based on the serial of cluster number (Table 1). After analyzing each probable satellite or tandem repeat cluster, it was found that the consensus monomer length (bp) of satellite repeats in *M. indica* ssp Alphonso genome were ranged between 119 bp - 3370 bp, genome proportion 0.012-0.26%, average GC content 31.4% - 58.8% (Table 1). Consensus monomer sequences of six tandem repeats are provided as Supplementary data 2.

Table 1. *Characteristics features of six satellite repeats identified from RepeatExplorer based whole genome sequence analysis*

Satellite name	Satellite cluster number	Consensus monomer length (bp)	Genome proportion	Average GC content (% of monomer)	Satellite probability
SatMan1	CL18	148	0.26	58.8%	0.121
SatMan2	CL19	119	0.26	33.6%	0.0376
SatMan3	CL165	337	0.082	51.3%	0.0277
SatMan4	CL201	342	0.055	33.0%	0.979
SatMan5	CL283	3370	0.026	31.4%	0.0667
SatMan6	CL335	188	0.012	51.6%	0.0284

Amplification and analysis of satellite repeats using PCR reactions

Six primer pairs were specifically designed for the amplification of the six tandem repeats family from three different *Mangifera* species (Table 2). Banding pattern differences were analyzed both among six satellite repeats as well as among three different species. It was found that in *M. indica* genome CL165 and CL201 produced expected ladder like pattern with expected band size of 400 bp and 800 bp within smear background, while monomer size of the respective satellite repeats CL165 and CL201 were 337 bp and 342 bp respectively. In contrast, CL283 produced ladder-like patterns of about 250 bp, 700 bp and 1200 bp, although the analyzed monomer size of the respective satellite repeats were 3370 bp. The remaining three primer pairs produced only smear in the PCR results in *Mangifera indica* genome, no specific satellite typical banding pattern was observed with the satellite specific primer pairs (Figure 1). The banding patterns were similar within different cultivars of mango (*M. indica*) (Supplementary Figure1).

Table 2. *Satellite primer pairs for PCR amplification of each satellite repeats*

Satellite name	Primer name	PCR Product size (bp)		Primers (5'-3')	Length	GC%	Tm
SatMan1	Man_CL18	148	Forward	CGTGTAAAACGGACGGTG	18	56	56.51
			Reverse	CGCGTGCGCAATATAGG	17	59	56.38
SatMan2	Man_CL19	119	Forward	TTTCAACTTTTAGAATTCG	21	23.81%	48.66
			Reverse	CGAAAAAACGGAACTG	16	43.75%	47.91
SatMan3	Man_CL165	337	Forward	GTGTTACAAATGGTATCAGAGC	22	40.91	55.19
			Reverse	CAAGAGGACACGCGGAG	17	64.71	57.35
SatMan4	Man_CL201	342	Forward	GTTCATAAACACTGTCGGTAC	21	41	54.58
			Reverse	GTCGTTTTGGGTTTACTAGC	20	45	54.61
SatMan5	Man_CL283	3370	Forward	CAAGTCATTTCTCACAAACCAG	22	41	56
			Reverse	GAAGTATGTGACCATCACTAGTATC	25	40	56
SatMan6	Man_CL335	188	Forward	GACGCGTTTTAGGAACAAAACC	22	45	59
			Reverse	GGCTTCAGGCCCGCAC	16	75	60.8

Comparative analysis of satellite repeats amplification among three *Mangifera* species revealed that each species have quite distinct satellite repeats banding pattern. For instance, primer CL18, only produced satellite typical dimeric bands in *M. sylvatica*, while the monomer size is aligned with expected monomer size 148 bp and 350 bp. However, the other two species (*M. indica* and *M. longipes*) produced higher bands on smear background (Figure 2).

All the three studied species produced a clear satellite typical banding pattern with the CL201 according to the monomer size 337 bp. However, in case of *M. indica* and *M. sylvatica* only dimeric bands were found but in case of *M. longipes* the highest band produced was tetramer in smeared background (Figure 2).

Primer CL283 produced similar banding patterns of about 250 bp, 700 bp and 1200 bp in all three studied species which do not match with the expected monomer size 3370 bp. Banding intensity is the highest with primer CL283 in all three studied species compared to other five primers. However, for *M sylvatica*, banding intensity was quite less compared to other two species *M. indica* and *M. longipes* (Figure 2).

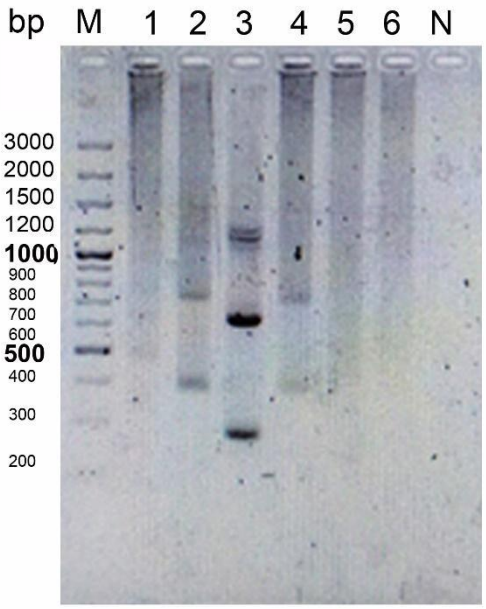

Fig. 1. *Amplification of satellite repeats from DNA of Mangifera indica using satellite specific primer pairs. Lane M = DNA Marker 200 bp., Lane 1 = Man_CL18, Lane 2 = Man_CL201, Lane 3 = Man_CL283, Lane 4 = Man_CL165, Lane 5 = Man_CL335, Lane 6 = Man_CL19, Lane N = Negative control.*

Primer CL165 produced expected satellite typical banding patterns within smeared background in all three species with highest dimeric band, while aligned with the monomer size of 337 bp (Figure 2).

Man_CL335 only produced dimeric bands in *M. sylvatica* within smeared background according to the monomer size 188 bp. In other two species, *M. indica* and *M. sylvatica* very faint bands with smeared background were observed (Figure 2).

Finally, CL19, does not produce any specific banding pattern in all three species studied, only smear was found (Figure 2).

Our results provide a first look at the mango satellite's repeated diversity and the subsequent effect on the genome diversity among different germplasm (wild and cultivated) of mango. Satellite specific markers distinguish different wild species at the molecular level and therefore has the potential to be used as inexpensive tools for taxonomic identification of different wild *Mangifer*a germplasm, or finding their distant relationship [37-39]. Especially the wild species from Bangladesh, are still not well characterized and identification is often ambiguous [4-5].

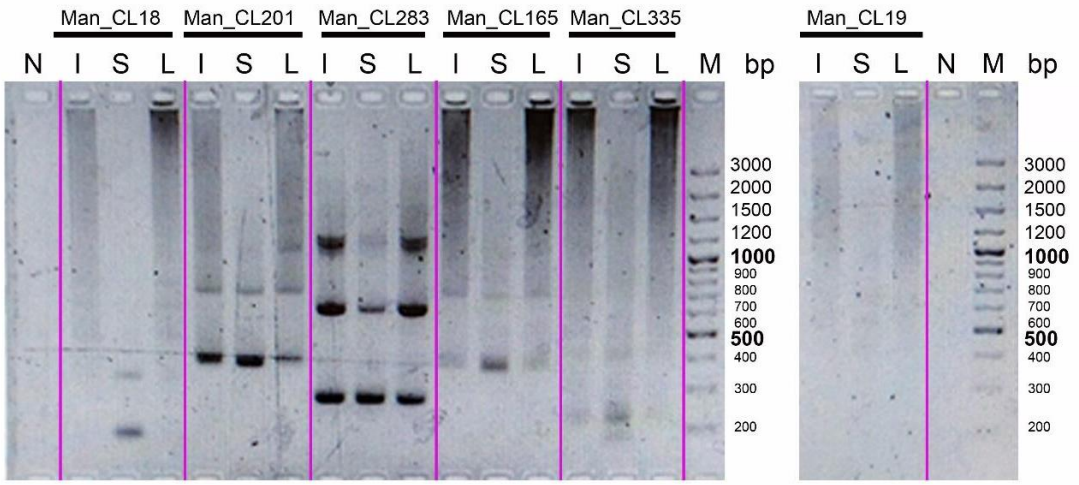

Fig. 2. *Amplification pattern of six satellite repeats from three Mangifera species. Here, Lane N = Negative control, I = Mangifera indica, S = Mangifera sylvatica, L = Mangifera longipes, M = DNA Marker 200 bp.*

CONCLUSION

Mango originated from the Indo-Burma region and therefore Bangladesh is strategically important for mango genetic resources. Several closely related mango wild species are growing here, however are not characterized properly and their taxonomic identification is often ambiguous. Since mango is a cross pollinating plant, often hybridized in natural condition, and has a long juvenile stage (seedling to flowering), it is not always easy to properly identify them or characterize them based on leaf morphology. The tandem repeat specific molecular marker has revealed that the collected wild species are indeed different from each other genomically. Therefore, it is very important to conserve the genetic diversity that exists in the natural environment.

Acknowledgement. S.O. is grateful to the National Science and Technology provided NST research fellowship for MSc. We are also grateful to the National Institute of Biotechnology for providing infrastructure support to carry out the research, especially polymerase chain reaction. We also acknowledge the research support provided by the Germplasm center of Bangladesh Agricultural University, by providing seeds of the plant material.

Conflict of Interest. The authors declared that there is no conflict of interest.

Authorship Contributions. Concept: N.S., Design: N.S., Data Collection or Processing: N.S., S.O., T.T., Analysis or Interpretation: N.S., S.O., T.T., Literature Search: N.S., Writing: N.S., S.O., Plant material: M.A.R.

Financial Disclosure. This research was funded by Jagannath University provided research project.

REFERENCES

[1] Dinesh, M.R., Hemanth, K.V., Ravishankar, K.V., Thangadurai, D., Narayanaswamy, P., Ali, Q., Kambiranda, D. and Basha, S.M. (2011): *Mangifera*. *In Wild Crop Relatives:* Genomic and Breeding Resources (pp. 61-74). Springer, Berlin, Heidelberg.

[2] Majumder, D.A.N., Hassan, L., Rahim, M.A. and Kabir, M.M. (2012): Genotypic and phenotypic variability in mango (*Mangifera indica* L.). Bangladesh Journal of Agricultural Research 37(4): 683-690.

[3] Rahman, M.A. (2017). Plant diversity in Hazarikhil wildlife sanctuary of chittagong and its conservation management. Journal of Biodiversity Conservation and Bioresource Management 3(2): 43-56.

[4] Uddin, M.Z. and Hassan, M.A. (2016): Plant diversity of Dhaka university campus, Bangladesh. Journal of the Asiatic Society of Bangladesh, Science 42(1): 49-68.

[5] Ashrafuzzaman, M., Khatun, M.M., Tunazzina, N.A. and Sarwar, A.K.M.G. (2021): Conservation of minor fruit genetic resources at the Botanical Garden, Bangladesh Agricultural University. Internafional Journal of Minor Fruits Medicinal & Aromafic Plants 7:1-18.

[6] Sattar, M.A. (2006): Plywood. Available at www.banglapedia.org/httpdocs/HT/P_0199.HTM

[7] Srivastava, A.P., Chandra, R., Saxena, S., Rajan, S., Ranade, S.A. and Prasad, V. (2007): A PCR-based assessment of genetic diversity, and parentage analysis among commercial mango cultivars and hybrids. The Journal of Horticultural Science and Biotechnology 82(6): 951-959.

[8] Nishiyama, K., Choi, Y.A., Honsho, C., Eiadthong, W. and Yonemori, K. (2006): Application of genomic in situ hybridization for phylogenetic study between *Mangifera indica* L. and eight wild species of *Mangifera*. Scientia horticulturae 110(1): 114-117.

[9] Yonemori, K., Honsho, C., Kanzaki, S., Eiadthong, W. and Sugiura, A. (2002): Phylogenetic relationships of *Mangifera* species revealed by ITS sequences of nuclear ribosomal DNA and a possibility of their hybrid origin. Plant Systematics and Evolution 231(1-4): 59-75.

[10] Iquebal, M.A., Jaiswal, S., Mahato, A.K., Jayaswal, P.K., Angadi, U.B., Kumar, N., Sharma, N., Singh, A.K., Srivastav, M., Prakash, J. and Singh, S.K. (2017): MiSNPDb: a web-based genomic resources of tropical ecology fruit mango (*Mangifera indica* L.) for phylogeography and varietal differentiation. Scientific reports, 7(1): 14968.

[11] Kuhn, D.N., Bally, I.S., Dillon, N.L., Innes, D., Groh, A.M., Rahaman, J., Ophir, R., Cohen, Y. and Sherman, A. (2017): Genetic map of mango: a tool for mango breeding. Frontiers in plant science 8: 577.

[12] Sherman, A., Rubinstein, M., Eshed, R., Benita, M., Ish-Shalom, M., Sharabi-Schwager, M., Rozen, A., Saada, D., Cohen, Y. and Ophir, R. (2015): Mango (*Mangifera indica* L.) germplasm diversity based on single nucleotide polymorphisms derived from the transcriptome. BMC Plant Biology 15(1): 277

[13] Singh, N.K., Mahato, A.K., Jayaswal, P.K., Singh, A., Singh, S., Singh, N. and Rai, V., (2016): Origin, diversity and genome sequence of mango (*Mangifera indica* L.). Indian Journal of History of Science 51(2):355-368.

[14] Arumuganathan, K. and Earle, E.D. (1991): Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9(3):208-218.

[15] Sarker, B.C. and Rahim, M.A. (2012): Vegetative growth, harvesting time, yield and quality of mango (*Mangifera indica* L.) as influenced by soil drench application of paclobutrazol. Bangladesh Journal of Agricultural Research 37(2):335-348.

[16] Garrido-Ramos, M.A. (2017): Satellite DNA: an evolving topic. Genes 8(9):230.

[17] Macas, J., Novak, P., Pellicer, J., Čížková, J., Koblížková, A., Neumann, P., Fukova, I., Doležel, J., Kelly, L.J. and Leitch, I.J. (2015): In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe Fabeae. PLoS One 10(11): e0143424.

[18] Mehrotra, S. and Goyal, V. (2014): Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. Genomics, proteomics & bioinformatics, 12(4): 164-171.

[19] Schmidt, T. and Heslop-Harrison, J.S. (1998): Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends in Plant Science *3*(5):195-199.
[20] Heslop-Harrison, J.S. and Schwarzacher, T., (2011): Organisation of the plant genome in chromosomes. The Plant Journal 66(1): 18-33.
[21] Kelly, L.J., Renny-Byfield, S., Pellicer, J., Macas, J., Novák, P., Neumann, P., Lysak, M.A., Day, P.D., Berger, M., Fay, M.F. and Nichols, R.A. (2015).:Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. New Phytologist 208(2): 596-607.
[22] Heitkam, T., Petrasch, S., Zakrzewski, F., Kögler, A., Wenke, T., Wanke, S. and Schmidt, T. (2015): Next-generation sequencing reveals differentially amplified tandem repeats as a major genome component of Northern Europe's oldest *Camellia japonica*. Chromosome Research 23: 791-806.
[23] Kirov, I.V., Kiseleva, A.V., Van Laere, K., Van Roy, N. and Khrustaleva, L.I. (2017): Tandem repeats of *Allium fistulosum* associated with major chromosomal landmarks. Molecular Genetics and Genomics 292: 453-464.
[24] Garrido-Ramos, M.A. (2015): Satellite DNA in plants: more than just rubbish. Cytogenetic and Genome Research, 146(2): 153-170.
[25] Plohl, M., Meštrovic, N., Mravinac, B. (2012): Satellite DNA evolution. In Repetitive DNA Genome Dynamic vol. 7, Garrido-Ramos, M.,A., Ed., Karger, Basel, pp.126–152.
[26] Meštrović, N., Mravinac, B., Pavlek, M., Vojvoda-Zeljko, T., Šatović, E. and Plohl, M., (2015): Structural and functional liaisons between transposable elements and satellite DNAs. Chromosome Research 23:583-596.
[27] Macas, J., Koblížková, A., Navrátilová, A. and Neumann, P. (2009): Hypervariable 3′ UTR region of plant LTR-retrotransposons as a source of novel satellite repeats. Gene 448(2): 198-206.
[28] Stupar, R.M., Song, J., Tek, A.L., Cheng, Z., Dong, F. and Jiang, J., (2002): Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. Genetics 162(3): 1435-1444.
[29] del Bosque, M.E.Q., López-Flores, I., Suárez-Santiago, V.N. and Garrido-Ramos, M.A., (2014): Satellite-DNA diversification and the evolution of major lineages in Cardueae (Carduoideae Asteraceae). Journal of Plant Research 127: 575-583.
[30] Louzada, S., Lopes, M., Ferreira, D., Adega, F., Escudeiro, A., Gama-Carvalho, M. and Chaves, R. (2020): Decoding the role of satellite DNA in genome architecture and plasticity—An evolutionary and clinical affair. Genes, 11(1): 72.
[31] Wendel, J.F., Jackson, S.A., Meyers, B.C. and Wing, R.A. (2016): Evolution of plant genome architecture. Genome Biology 17:1-14.
[32] Shapiro, J.A. and von Sternberg, R., (2005). Why repetitive DNA is essential to genome function. Biological Reviews 80(2): 227-250.
[33] Novák, P., Neumann, P., Pech, J., Steinhaisl, J. and Macas, J. (2013): RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics 29(6): 792-793.
[34] Novák, P., Neumann, P. and Macas, J. (2020): Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. Nature Protocols 15(11): 3745-3776.
[35] Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C. and Thierer, T. (2012): Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12): 1647-1649.
[36] Doyle, J.J. and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. Focus 12(13): 39-40.
[37] Nagoshi, R.N. and Meagher, R.L. (2003): FR tandem-repeat sequence in fall armyworm (Lepidoptera: Noctuidae) host strains. Annals of the Entomological Society of America, 96(3): 329-335.

[38] Grant, J.R., Pilotte, N. and Williams, S.A., (2019). A case for using genomics and a bioinformatics pipeline to develop sensitive and species-specific PCR-based diagnostics for soil-transmitted helminths. Frontiers in Genetics, 10:883.
[39] de Boer, H., Rydmark, M.O., Verstraete, B. and Gravendeel, B., (2022). Molecular identification of plants: from sequence to species. Advanced Books, 1:e98875.