Plant Tissue Cult. & Biotech. **33**(1): 57-70, 2023 (June) **ISSN 1817-3721, E-ISSN 1818-8745**
DOI: https://doi.org/10.3329/ptcb.v33i1.66904

PTC&B

Phylogenetic Relationship among Grasses, Wheat and Pearl Millet using *matK* and *rbcR* Genes

M.M.H. Molla*, M.C. Saha[1], D. Serba[2], S. K. Talukdar[1], M. A. Salam[3], M. Shahabuddin Ahmed and M. S. Islam

Molecular Biology Laboratory, Tuber Crops Research Centre, Bangladesh Agricultural Research Institute, Gazipur-1701, Bangladesh

Key words: Grasses, Wheat, Pearl millet phylogenetic analysis, *matK* and *rbcR* gene

Abstract

Fourteen genotypes of tall fescue, *Brachypodium*, wheat and pearl millet were used to find out phylogenetic relationships at their different ploidy level. Isolated chloroplast DNA of the 14 genotypes was sequenced using some selected *matK* and *rbcR* genetic markers. The *matK* and *rbcR* chloroplast gene sequences length were 1539bp and 1431bp, respectively. The consensus sequence length of *matK* gene was 1265bp in *Brachypodium* and wheat, 1266bp in both tall fescue Continental and Mditerranean species, 1269bp in pearl millet and for *rbcR* gene was only 1156bp in all studied four species. Except in *Brachypodium,* the variations of SNP for both *matK* and *rbcR* genes were observed in different ploidy levels of pearl millet and between the tall fescue Continental and Mediterranean species. Wheat with their different ploidy level gave the SNP variations only for *matK* gene. In respect of all ploidy levels, pearl millet had the highest number of SNP for both *matK* (247) and *rbcR* (130) genes. The tall fescue Continental (129) and Mediterranean (123) were in second position to show their genetic variation for *matK* gene. The *rbcR* gene had higher SNP at 5′ regions of all four species, but *matK* gene only in wheat and tall fescue and 3′ in millet and *Brachypodium* species. There was no in-del for *rbcR* gene, but the insertion of nucleotide for the *matK* gene was found highest in diploid wheat (11) and the deletion was in pearl millet (6). The phylogenetic analysis stated that their relationships were tall fescue Continental and tall fescue Mediterranean was much more related to wheat than pearl millet and *Brachypodium*.

*Author for correspondence: <mhmolla.bari@gmail.com>. [1]Molecular Lab, The Samuel Robert Noble Research Institute, Inc. 2510 Sam Noble Parkway, Ardmore, OK, USA. [2]Research Geneticist, USDA-ARS, Maricopa, Arizona, USA. [3]Bangladesh Agricultural Research Council (BARC), Farmgate, Dhaka-1215, Bangladesh.

Introduction

There are about 10,000 species in 600 to 700 genera of flowering plants in the world (Watson and Dallwitz 1992). A large number of peoples in the world depend on grass family Poaceae, which includs rice (*Oryza sativa*), wheat (*Triticum aestivum*) and maize (*Zea mays*) for a major portion of their diet. Grass covers 27% of the world's total land area and 71% of the world agricultural area. Several minor crops, e.g., pearl millet (*Pennisetum glaucum*), belonging to the same family is staple food in the different countries in the world (Zeid et al. 2010). Tall fescue (*Festuca arundinaceae*) is a major forage species belongs to this family. Genetic studies of these minor crops are slowed down due to the limitation of genetic resources. Recently molecular advancement such as gene sequence and their comparative study are important tools for the study of plant evolution by the evolutionary biologists.

Chloroplasts are organelles that are essential for transducing energy. They contain the entire enzymic process and the components necessary for photosynthesis as well as DNA, RNA and all the machineries involved in DNA replication. Both chloroplasts and the mitochondria contribute to plant development through genetic control outside the nuclear chromosomes. In higher plants this extra chromosomal control came primarily from the female parent. The maturase K (*matK*) gene has been adopted as an international standard for DNA barcoding (Group et al. 2009) and provides evidence for maternal progenitor identity, is being used in phylogenetic studies of different plant species at their taxonomic levels (Johnson and Soltis 1995, Hilu and Liang 1997, Hilu et al. 1999, Ge et al. 2002). The phylogenetic signal from *matk* gene has been used to resolve relationship from the species level to across the broad group of land plants. The mode of action of *matK* gene evolution is distinct from other chloroplast genes along with nucleotide substitution. Nevertheless, coding sequences of the *matK* gene that is the most highly evolving in sequence can provide enough information for the development of phylogenetic relationships among the intragenus group of plant (Yang et al. 2004). Moreover, the *rbcL* gene, which encodes the large subunit of ribulose-1, 5-biphosphate carboxylase/oxygenase has been widely sequenced from numerous taxa of plants and the resulting database has greatly assisted the research of plant phylogeny (Clegg and Zurawski 1992, Chase et al. 1993). Such phylogeneis, based on *rbcL* sequence were successfully obtained at the family level (Soltis et al. 1990, Bousquet et al. 1992, Morgan and Soltis 1993) and also at higher levels (Bousquet et al. 1992, Gaut et al. 1992, Chase et al. 1993). Grass family (Poaceae) had been identified using the plant DNA barcodes of *rbcL* and *matK* gene (Saadullah et al. 2016). In addition, sequences of internal transcribed spacers are the most often used in nuclear DNA sequence for discussing the relationships of the plants species (Group et al. 2001).

The coding sequences of the plastid were successfully used in phylogenetic studies at the higher taxonomic level (Lewis and Doyle 2001, Mason-Gamer et al. 2002, Nishikawa et al. 2002, Kato et al. 2003, Neel and Cummings 2004). The nuclear genomes are

generally passed to offspring from parents by the equal contribution of male and female parents, whereas chloroplast and mitochondrial genomes are generally transferred to their offspring`s following maternal inheritance (Korpelainen 2004). Based on maternal inheritance, the study of genetic relationship in chloroplast DNA between plant species is easier than biparental inheritance and in polyploid species (Matsuoka et al. 2002). Recombination frequency is usually observed in maternal inheritance, which is used to simplify the genome evolution and the theory of chloroplast DNA in most of the plant taxa. Chloroplast genome sequence provides basic information to develop a theory of the evolution to get that relationship study. There was no comprehensive phylogenetic study of representative species of tall fescue, brachypodium, wheat and pearl millet. The results of phylogenetic studies of chloroplast *matK* and *rbcR* gene sequences in members of the Poaceae family have been reported here. Such results would provide a better understanding of evolutionary history of these species, which will give the way to required improvement in the study of plant species and to use in the plant breeding and biotechnology. Our objectives were to know the phylogenetic relationship among the wheat, tall fescue, brachypodium, and pearl millet at different ploidy, progenitors and probable progenitor's level.

Materials and Methods

Fourteen genotypes from four important Poaceae species, tall fescue (*Festuca arundinaceae*), Brachypodium (*Brachypodium distachyon*), wheat (*Triticum aestivum*) and pearl millet (*Pennisetum glaucum*) were included to demonstrate relationship among four tribes of Poaceae at their different ploidy levels (Table 1). Tall fescue is grown extensively for pasture, silage and hay. It has three distinct morphotypes e.g., Continental, Mediterranean and Rizomatous. The Continental fescue grows round the year, Mediterranean fescue displays summer dormancy and is less winter hardy. However, it is an allopolyploid and found their differences in ploidy from diploid (2n=6x=14) to dodecaploid (2n=12x=84) with genome size between 5.27-5.83x 10^6 kb (Loureiro et al. 2007, Šmarda et al. 2008). Large genome size and polyploid nature of plants make difficulties to improve genetic traits for the beneficial use. *Brachypodium distachyon* L., a new model grass, has diverse use and treated as an economically important group of energy crops among the temperate grasses (Draper et al. 2001). That species is ranged from annuals to perennials. It has high inbreeding depression and self-incompatibility noted in plant improvement program through conventional breeding (Khan and Stace 1999). It also has small haploid genome (~355 Mbp) and its basic chromosome number is 5 (2n=2x=10). *B. distachyon* is being used as a model for the larger genome size polyploid crops such as bread wheat (16979 Mbp, 2n=2x=42), durum wheat (12030 Mbp, 2n=2x=28) and barley (5439 Mbp 2n=2x=14). Barley, wheat and rye are also economically important crops and they have close relationship with brachypodium (Bortiri et al. 2008, Ozdemir et al. 2008). Pearl millet (*Pennisetum glaucum* L.) is widely cultivated in dry land, rain fed

and irrigated conditions in drought-prone regions of the tropics and sub-tropics. It has two wild relatives of *P. violaceum* and *P. mollissium* and are closely related based on their reproductive system and the shared gene from the same gene pool (Brunken 1977). Its basic chromosome number varies from 5 to 9 (x=5, 7, 8 and 9) and it is heterogeneous with their reproductive behaviour (sexual or apomictic) and life cycle (annual, biennial or perennial).

The genus *Triticum*, includes world most important food crops, composed of four basic A, B, D, and G genomes which help to understand the genome constitution of the plant kingdom (Kihara 1924, Brunken 1977). Emmer and Timopheevii are the two evolutionary polyploid wheat. The evolution of cultivated wheat was from the diploid species of four wheat and goat grass through amphidiploidization. Diploid wheat and goat grass (*Aegiolops squarrosa*) donated the A and D genomes, respectively to develop the cultivated wheat, whereas B genome was a recombined genome derived from two more diploid *Aegiolops* species (Kihara 1924, Kihara 1930, McFadden and Sears 1946). It was noted that there were two different a genomes in the wheat species, A^{11} and A^{b} genome of *T. urartu* Thum. ex Gandil and (Kindt et al. 1970) of *T. boeoticum* Boiss, respectively (Kihara 1930, Kindt et al. 1970). To conduct the study, brachypodium and pearl millet's seeds were collected from Germplasms Resources Information Network (GRIN), USDA. The diploid, tetraploid and hexaploid wheat seeds were received from Kansas State University, Manhattan, USA. The Noble Research Institute *Inc*, Ardmore, Oklahoma kindly provided the Continental and Mediterranean diploid, tetraploid and hexaploid tall fescue seeds. After collection of seeds from different sources, they were sown, germinated and grown in the greenhouse of the Noble Research Institute. Seed sources and ploidy levels of entire collection are presented in Table 1.

Fresh young needle leaves (approx. 100 mg) of tall fescue, brachypodium, wheat and pearl millet were collected and cut into small spices in 2 ml collection tubes and kept in liquid Nitrogen. Leaf samples were ground into fine powder immediately using Mixer Mill Type MM 300 (Retsch Tissue Lyser, Hannover, Germany) and stored at -80^{0}C. Genomic DNA from leaves was extracted using Promega Genomic DNA Isolation Kit, USA. All steps for the DNA extraction were followed according to the manufacturer's instructions. Extracted DNA was quantified using a micro spectrophotometer (NanoDrop ND -1000, USA) at wavelengths of 260/280 and 260/230 nm, respectively. *matK and rbcR* chloroplast gene primers were used to detect the *matK and rbcR* gene. Primers were selected from the conserved regions of tall fescue, brachypodium, wheat, rye and maize that amplify specific *matK and rbcR* gene sequences accurately. Details of the primers are shown in Table 2. Primers were designed using the primer3 software (http://www.genome.wi.mit.edu/genome_software/other/ primer3.html). The major parameters for designing primers were 16-22 bases long, 20bp was optimum having 385 to 498bp amplified product size, annealing temperature 45 to 50°C and GC contents from 40-70% with 50% optimum. Primers were synthesized by Sigma-Aldrich Co. LLC, USA. A sample of 15 ng of template DNA was used for PCR amplification in a GeneAmp PCR

Systems 9700, Applied Biosystem, USA with an initial denaturation at 94°C for 2 minutes followed by 35 cycles of 1 minute denaturation at 94°C, 1 minute annealing at 45°C and elongation or extension at 72°C for 2 minutes.

After the last cycle, a final step of 5 minutes at 72°C was added to allow complete extension of all amplified fragments. Double annealing temperature like 45°C and 50°C were employed accommodate all primers tested in this study. Reactions were carried out in 20 µl volumes containing 5× PCR buffer 2.5 mM $MgCl_2$, 2.0 mM of each dNTPs, 5 Units Taq DNA Polymerase (Promega), 10 µM of each primer pair. After completing of cycling program, reaction was held at 4°C. PCR amplifications were examined on 1.5% agarose gels with ethidium bromide staining. After examination of the PCR product on 1.5% agarose gel, amplified DNA was run again in 1.25% agarose gel for extracting purified DNA for sequencing. Purification was done following the instruction of Wizard® SV Gel and PCR Clean-Up System (Promega). All steps for the DNA purification were done as per manufacture's manual. Purified DNA was quantified again using NanoDrop ND-1000 Spectrophotometer, USA at wavelengths of 260/280 and 260/230 nm, respectively. All the samples were diluted with 20 ng/µl for sequencing. PCR and sequencing were executed using the same primers. Sequencing was carried out at the Noble Research Institute Genomics core facility.

Phylogenic trees were constructed using Geneious tree builder, global alignment with free end gapes alignment type using cost matrix 65% similarities and Tamura – Nei genetic distance model UPGMA method from 14 genotypes from consensus alignment for *matK* and *rbcR* genes.

Results and Discussion

Chloroplast genes (*matK* and *rbcR*) from purified PCR product of fourteen genotypes of tall fescue, brachypodium, wheat and pearl millet were sequenced and phylogenetic analysis was done to see their relationships (Figs. 1, 2, 3 and 4). The primers were designed from the most common conserved regions to isolate the complete gene sequence. Genomic DNA amplification produced two fragments with different length from most of plant materials. The *matK* and *rbcR* gene sequence length was 1539 bp and 1431bp, respectively found from data base. The consensus sequence length of *matK* gene in brachypodium, tall fescue Continental, tall fescue Mediterranean, wheat and pearl millet was 1265, 1266, 1266, 1265, 1269bp, respectively and it was 1156bp for *rbcR* gene (data not shown). Length variations in *matK* gene may be due to mutations of plastid genomes that has been identified in higher plants (Lu et al. 2016).

Sequenced data were divided into 3 groups like 5´ (1-400bp), middle (400-800bp) and 3' (800-1200bp) for each genotype. In case of *matK* gene sequence, a total of 247 SNPs were found in pearl millet diploid, tetraploid and hexaploid species. Among the SNPs, 84 were found both in pearl millet diploid and tetraploid and 79 was in the hexaploid species. Only 6 deletions were seen in diploid species at middle region (Fig. 1, Table 3).

Diploid, tetraploid and hexaploid species of pearl millet produced 31, 28, 28; 16, 19, 16 and 37, 37, and 35 SNPs at 5', middle and 3' regions, respectively. Maximum number of SNPs was found at 3' regions followed by 5' and middle regions. There was no addition of nucleotides in pearl millet *matK* gene sequence. In case of *rbcR* gene, a total of 45, 41, 44 SNPs were recorded in pearl millet diploid, tetraploid and hexaploid species, respectively where maximum number of SNPs 22, 20, 22 have been seen at 5' regions; 14, 13, 13 SNPs at middle regions and 9, 8, 9 SNPs at 3' regions of diploid, tetraploid and hexaploid pearl millet, respectively. There was no insertion or deletion seen in any species of pearl millet. Maximum number of SNPs was found at 5' regions followed by middle and 3' regions in diploid, tetraploid and hexaploid pearl millet.

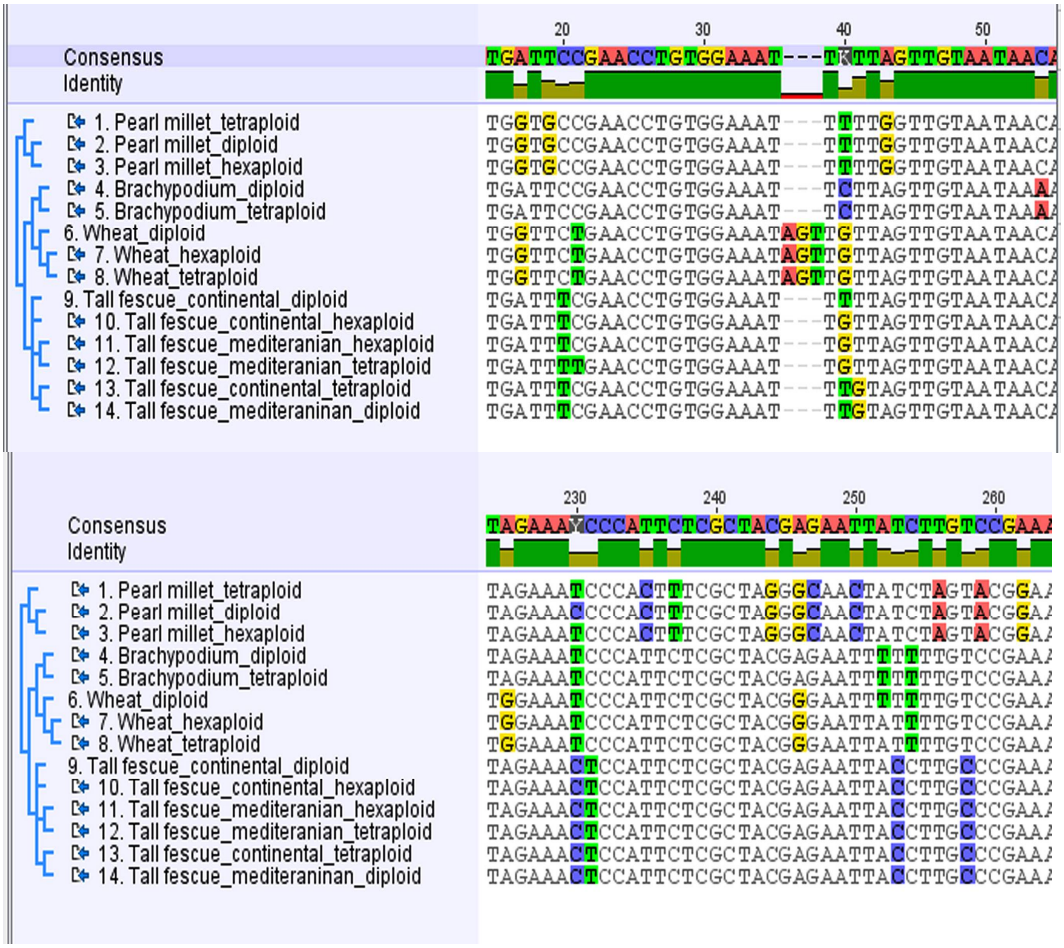

Fig. 1. Consensus alignment for *matK* gene identified SNP variations within and among four species and an insertion of 3-bp in wheat species.

Table 1. List of genotypes, their ploidy level and source of collection used in this study

Cultivar/ genotype	Tribe	Species	Ploidy	Source	Subfamily
Wheat	Triticeae	*T. tauschii*	Diploid (2n=2X=14)	Kansas State University, USA	Panicoideae
		Triticum turgidum	Tetraploid (2n=4X=28)		
		Triticum aestivum	Hexaploid (2n=6X=42)		
Tall Fescue Continental	Poeae	*Festuca arundinacea*	Diploid (2n=2X= 14)	Noble Research Institute, Oklahoma	Pooideae
		Festuca arundinacea Var. *glucescens*	Tetraploid (2n=4X=28)		
		Festuca arundinacea Var. *genuina*	Hexaploid (2N=6X=42)		
Tall Fescue Mediterranean	Poeae	*Festuca arundinacea*	Diploid (2n=2X= 14)		
		Festuca arundinacea Var. *glucescens*	Tetraploid (2n=4X=28)		
		Festuca arundinacea Var. *genuina*	Hexaploid (2N=6X=42)		
Millet	Paniceae	*Pennisetum glaucum*	Diploid (2n=2X= 18)	GRIN, USDA	Pooideae
		P. purpureum	Tetraploid (2n=4X= 36)		
		P. squamulatum	Hexaploid (2n=6X= 54)		
Brachypodium	Poeae	*Brachypodium distachyon*	Diploid (2n=2X=10)	GRIN, USDA	Pooideae
		B. hybridum	Tetraploid (2n=4x=30)		

Table 2. Primers for the full length amplification of the *matK* and *rbcR* genes.

	Forward sequence(5′- 3′)	Fragment length	Reverse sequence(5′- 3′)	Fragment length	Amplified fragment
matK1	GATTATGGATTAAATG	16	CCTATCCTATCCATTT	16	469
matK2	TCTATTCATTCAATATTT	18	GATATCAAGGAAAGGC	16	498
matK3	TTTATGGATCCTCTTATGCA	20	GGTACGAACTTTTATGCAACG	21	477
rbcR1	GATACTGATATCTTGGCAGC	20	GGTAGAGCGTGTTATGAGTG	20	385
rbcR2	GGTAACGTATTTGGTTTCAA	20	GGATTCACCGCAAATACTA	19	469
rbcR3	GTTCCTATTGTAATGCATGA	20	GCACCTGGTGCAGCAGCTAATC	22	556

A total of 48 SNPs (*matK*) were recorded in each of brachypodium diploid and tetraploid species where 14 were at 5' regions in the species, 17 and 16 were at middle regions and 17 and 18 at 3' regions of diploid and tetraploid species. Maximum number

of SNPs (17 and 18) was seen at 3' regions in brachypodium diploid and brachypodium tetraploid species, respectively followed by middle (17 and 16) and 5' regions (14) (Fig. 1, Table 3). No insertion or deletion was found in the *matK* sequence of brachypodium diploid and tetraploid species. In case of *rbcR* gene, maximum SNPs (10) were found at 5' followed by middle (9) and 3' regions (3) of brachypodium diploid and brachypodium tetraploid sequence, respectively. A total of 22 SNPs were found in each of brachypodium diploid and tetraploid species. However, 17, 11 and 10 SNPs were found in 5', middle and 3' regions of wheat tetraploid and hexaploid species, respectively where 18, 13 and 12 were observed in the 5', middle and 3' regions of wheat diploid. Maximum number of SNPs has been seen at 5' region followed by middle and 3' regions in diploid, tetraploid and hexaploid specieses of wheat (Fig. 2, Table 3). Moreover, *matK* gene sequence in wheat diploid showed 11 insertion at 5' regions and 3 insertions at 5' regions of wheat tetraploid and hexaploid. There was no nucleotide deletion at any regions of sequence in wheat diploid, tetraploid and hexaploid species. From the *rbcR* gene sequence, a total of 11 SNPs was found in each of wheat diploid, tetraploid and hexaploid species where maximum (6) SNPs were seen in 5' regions followed by middle (4) and 3' region (1) in each species.

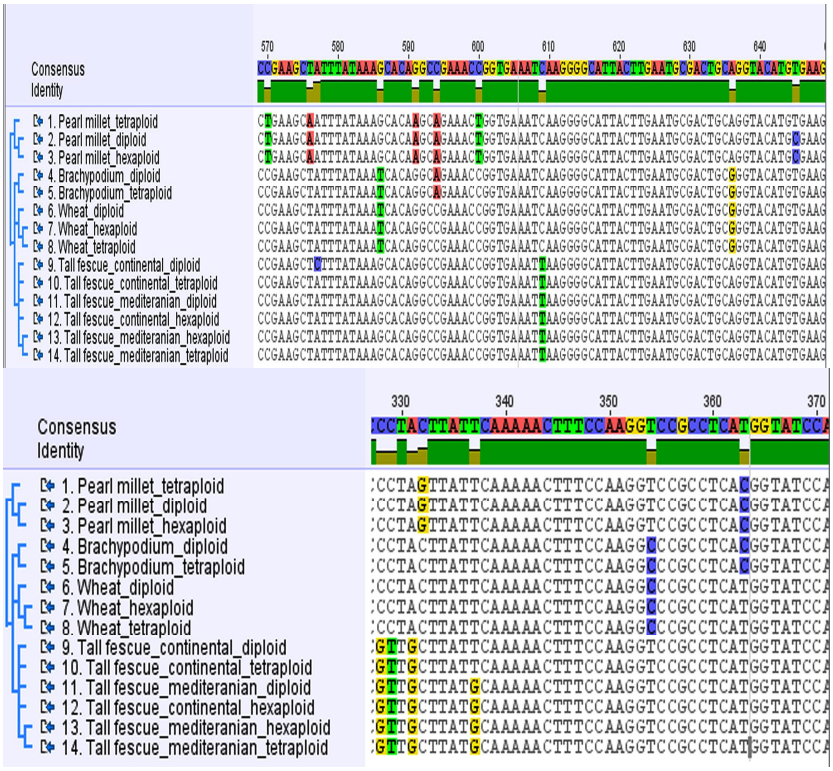

Fig. 2. Consensus alignment for *rbcR* gene within identified species- and genotype-specific SNPs in four species of the Poaceae family.

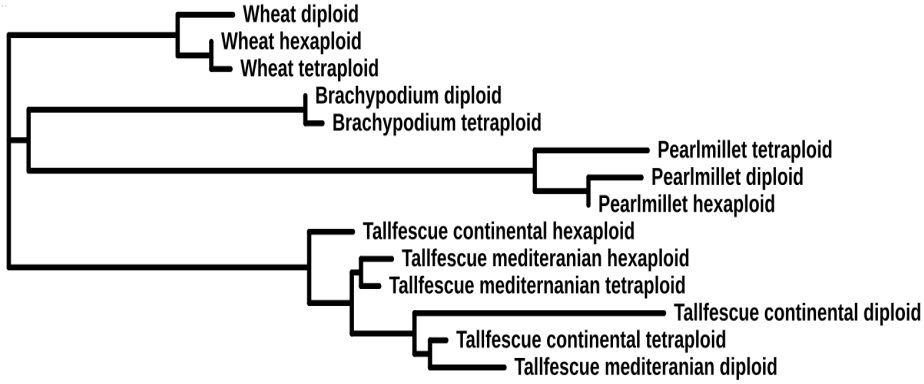

Fig. 3. Phylogenetic relationship among few members of the Poaceae family as revealed by the consensus sequences of the *matK* gene.

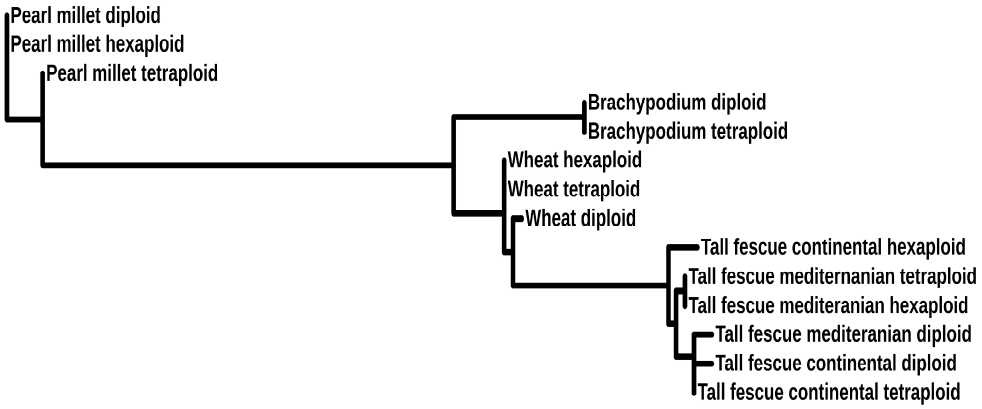

Fig. 4. Phylogenetic relationship among few members of the Poaceae family as revealed by the consensus sequences of the *rbcR* gene.

There was no insertion or deletion found in any of the species of wheat in this case (Fig. 2). These results indicated a relatively more conserved 3' region in comparison to the more variable 5' region, which also earlier reported by Hilu and Liang (1997).

Maximum number of SNPs (47) were found in tall fescue Continental diploid followed by tall fescue Continental tetraploid (46), tall fescue Mediterranean diploid (42), tall fescue Mediterranean hexaploid (41), tall fescue Mediterranean tetraploid (40) and tall fescue Continental hexaploid (36) from the *matK* sequence. In all cases, maximum number of SNPs has been found at 5' regions followed by 3' and middle regions. There was only one insertion was seen in tall fescue Continental tetraploid at middle region. In case of *rbcR* gene, maximum number of SNPs were recorded in tall fescue Continental diploid (12) followed by tall fescue Continental hexaploid (11), tall fescue Continental tetraploid (10), tall fescue Mediterranean diploid (10), tall fescue Mediterranean

tetraploid (9), tall fescue Mediterranean hexaploid (9) (Fig. 2, Table 3). The highest number of SNPs (6) was seen at 5' regions in every species but equal number of SNPs was found at middle and 3' end of tall fescue species in most of the cases. Only one SNP was found in tall fescue Mediterranean tetraploid and tall fescue Mediterranean hexaploid species at 3' region of *rbcR* gene sequence where 3 SNPs were in tall fescue Continental diploid and tall fescue Continental hexaploid and 2 SNPs were found in tall fescue Continental tetraploid and tall fescue Mediterranean diploid (Fig. 2). There was no insertion or deletion has been seen in any of the tall fescue species for *rbcR* gene sequences (Table 3).

To perform phylogenic analysis, the consensus sequences of *matK* and *rbcR* genes for each species were combined. All trees were bootstrapped 100 times. The bootstrap values were 98 for pearl millet diploid and pearl millet hexaploid and 88 times in tall fescue Mediterranean hexaploid and tall fescue Mediterranean tetraploid in *matK* and *rbcR* gene, respectively (Figs. 3-4). Members of the Tritieae and Poeae tribes grouped together in one cluster (Fig. 3). Tribe Paniceae separated cluster from all other groups. Moreover, the genera of the grasses and cereals fall into two main clades. The first clade includes three genera namely *Festuca*, *Brachypodium* and *Triticum*. The second clade consists of with *Pennisetum.* This clustering is in agreement with previously established taxonomic relationships among grass species (Kellogg 2001, Zeid et al. 2010).

Based on the sequenced data both of the *matK* and *rbcR* gene indicated that tall fescue Continental and Mediterranean are much more related with wheat than pearl millet and brachypodium. It is interesting to note that in both cases, tall fescue Mediterranean tetraploid and tall fescue Mediterranean hexaploid produced individual clades in both of *matK* and *rbcR* gene with 100 and 88 times bootstrapped, respectively (Figs. 3 and 4). But only *matK* gene sequenced data of tall fescue Continental tetraploid and tall fescue Mediterranean diploid produced similar clade with 100 times bootstrapped (Fig. 3). In case of brachypodium, wheat and pearl millet, both of *matK* and *rbcR* gene sequenced data produced separate clade with 100 times boot-strapped where only tetraploid and hexaploid wheat species goes together in a separate clade with 100 times bootstrapped (Fig. 3). But pearl millet diploid and hexaploid species also produced similar clades both in *matK* and *rbcR* gene with 98 and 100 times bootstrapped, respectively (Fig. 3 and 4). However, 99.9% similarities has been found between brachypodium diploid and brachypodium tetraploid species but 7-8% genetic variation was observed among brachypodium and tall fescue species (Table 4). Within the different species of tall fescue, there were genetic similarities which ranged from 97.5 to 99.6% (Table 4). Among the tall fescue species, there were 98.8 % similarities between Continental diploid and Continental tetraploid. Also Continental diploid and Continental hexaploid species showed 97.5% similarities.

In case of tall fescue Mediterranean ecotypes, diploid has 98.9% similarities with tetraploid and 93% similarities with Mediterranean hexaploid but tetraploid has 99.6% similarities with haxaploid. However, tall fescue Mediterranean diploid has 99.7%

similarities with Continental tetraploid (Table 4). Among the species of brachypodium and wheat, brachypodium diploid has 94% similarities with wheat diploid and 94.6% similarities with both wheat tetraploid and hexaploid (Table 4). However, within the species of wheat, tetraploid wheat has been found 100% similar with hexaploid where diploid wheat showed 5% variation with tetraploid and hexaploid (Table 4). In case of pearl millet, 99% genetic similarities were found between pearl millet diploid and hexaploid where 98.7% was recorded between tetraploid and hexaploid (Table 4). Within the millet species, *matK* gene sequence was very much conserved. The variation between millet diploid and brachypodium diploid was about 9% and brachypodium tetraploid 8.9%. Pearl millet diploid had 3.3% variation with wheat tetraploid and hexaploid. Distinct genetic variation was observed among the species of tall fescues and pearl millet which ranged from 1.8 to 2.8%. But in case of *rbcR* gene sequence, no genetic variation found between brachypodium diploid and brachypodium tetraploid species and only 2.9 to 2.75% genetic variation was observed among brachypodium and tall fescue species (Table 5). 100% similarities were found among tall fescue Mediteranian tetraploid and hexaploid species. Among all other species of tall fescue, genetic variation ranged was 99.7 to 100% (Table 5). However, within the species of wheat, tetraploid has been found 100% similarities with hexaploid where diploid wheat showed 98.1% similarities with tetraploid and hexaploid. In case of pearl millet, 100% genetic similarities were found between pearl millet diploid and hexaploid where 0.4% variation was seen among other ecotypes. Within the millet species, *rbcR* gene sequence was very much conserved. The variation among brachypodium diploid, tall fescue Continental diploid, tall fescue Mediterranean diploid, wheat diploid and pearl millet diploid genotypes were found about 2.9, 0.3, 0.2 and 4.25%, respectively (Table 5).

Chloroplast genome has become an interesting molecular tool to study the phylogenetic relationship and genetic variabilities among species due to its small size, conserved DNA sequence and maternal inheritance (15-17). In this study, we sequenced the *matK* and *rbcR* chloroplast genes from 14 genotypes of four important members of the Poaceae family to understand their DNA polymorphisms and phylogenetic relationships. The *matK* is an interesting and invaluable gene in plant systematic and evolutionary studies (Barthet and Hilu 2007). In general, *matK* was found more variable (total 714 SNPs) compared to the *rbcR* (total 268 SNPs). However, majority (51%) of the variations in *rbcR* gene was observed in the 5' region compared to the 3' region (18%). Similar trend was also observed in *matK* gene but the magnitude was less pronounced (38% in 5' vs. 36% in 3'). In general, 3' region is more conserved than the 5' region (Hilu and Liang, 1997). Both the gene sequences revealed similar phylogenetic relationship among the studies species. Diploid, tetraploid and hexaploid species of a genus followed the expected patterns. However, few mismatch was observed in tall fescue. The genomic constitution of Continental tall fescue is well characterized (Hand et al. 2010). But, the genomic constitution of Mediterranean tall fescue still remails unknown (Dierking et al.

2015). Though the progenitors of Continental tall fescue are well defined but for Mediterranean it is yet to be decided.

We sequenced *matK* and *rbcR* genes from four important members of the Poaceae family. In general, *rbcR* was found much more conserved than the *matK*. Majority of the variations were observed in the 5' compared to the 3' regions. However, variation in 5' and 3' regions was much more pronounced in *rbcR* compared to *matK*. Both the gene sequences revealed very similar phylogenetic relationship among the members of the Poaceae family. Our findings will help to identify species-specific DNA markers and to understand the evolutionary similarities/dissimilarities among the studied species.

Acknowledgements

Authors are grateful to World Bank funded National Agricultural Technology Support Project (NATP)-BARC Phase-II to provide the financial support for this research work. Principal author is also grateful to NATP-BARC to provide him Fulbright scholarship for this study in USA for six months. We are very thankful to the members of the Noble Research Institute greenhouse and genomics core facilities, USA.

References

Barthet MM and Hilu KW (2007) Expression of *matK*: Functional and evolutionary implications. Am J Bot. **94**(8): 1402-1412.

Bortiri E, Coleman-Derr D, Lazo DG, Anderson OD and **Gu YQ** (2008) The complete chloroplast genome sequence of Brachypodium distachyon: sequence comparison and phylogenetic analysis of eight grass plastomes. BMC Research Notes. **1**(1): 1-8.

Bousquet J, Strauss SH, Doerksen AH and **Price RA** (1992) Extensive variation in evolutionary rate of rbcL gene sequences among seed plants. Proceedings of the National Academy of Sciences. **89**(16): 7844-7848.

Brunken JN (1977) A systematic study of Pennisetum sect. Pennisetum (Gramineae). American Journal of Botany. **64**(2): 161-176.

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG and **Qiu YL** (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. Annals of the Missouri Botanical Garden: 528-580.

Clegg MT and **Zurawski G** (1992) Chloroplast DNA and the study of plant phylogeny: present status and future prospects. Molecular systematics of plants, Springer**:** 1-13.

Dierking R, Azhaguvel P, Kallenbach R, Saha M, Bouton J, Chekhovskiy K, Kopecký D and **Hopkins A** (2015) Linkage maps of a Mediterranean × Continental tall fescue population and their comparative analysis with other Poaceae species. The Plant Genome. **8**(1): https://doi.org/10.3835/plantgenome2014.07.0032

Draper J, Mur LA, Jenkins G, Ghosh-Biswas GC, Bablak P, Hasterok R and **Routledge AP** (2001) Brachypodium distachyon. A new model system for functional genomics in grasses. Plant physiology. **127**(4): 1539-1555.

Gaut BS, Muse SV, Clark WD and **Clegg MT** (1992) Relative rates of nucleotide substitution at the rbcL locus of monocotyledonous plants. Journal of molecular evolution. **35**(4): 292-303.

Ge S, Li A, Lu BR, Zhang SZ and **Hong DY** (2002) A phylogeny of the rice tribe Oryzeae (Poaceae) based on matK sequence data. American Journal of Botany. **89**(12): 1967-1972.

Group CPW Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS and **Erickson DL** (2009) A DNA barcode for land plants." Proceedings of the National Academy of Sciences. **106**(31): 12794-12797.

Group GPW, Barker NP, Clark LG, Davis JI, Duvall MR, Guala GF, Hsiao C, Kellogg EA, Linder HP and **Mason-Gamer RJ** (2001) Phylogeny and subfamilial classification of the grasses (Poaceae). Annals of the Missouri Botanical Garden: 373-457.

Hand ML, Cogan NOI, Stewart AV and **Forster JW** (2010) Evolutionary history of tall fescue morphotypes inferred from molecular phylogenetics of the *Lolium-Festuca* species complex. BMC Evol Biol. **10**:303-303.

Hilu KW and Liang H (1997) The *matK* gene: sequence variation and application in plant systematics. Am J Bot. **84**(6): 830-839.

Hilu KW, Alice LA and **Liang H** (1999) Phylogeny of Poaceae inferred from matK sequences. Annals of the Missouri Botanical Garden: 835-851.

Hilu KW and **Liang G** (1997) The matK gene: sequence variation and application in plant systematics. American journal of botany. **84**(6): 830-839.

Johnson LA and **Soltis DE** (1995) Phylogenetic inference in Saxifragaceae sensu stricto and Gilia (Polemoniaceae) using matK sequences. Annals of the Missouri Botanical Garden: 149-175.

Kato Y, Aioi K, Omori Y, Takahata N and **Satta Y** (2003) Phylogenetic analyses of Zostera species based on rbcL and matK nucleotide sequences: implications for the origin and diversification of seagrasses in Japanese waters. Genes & genetic systems. **78**(5): 329-342.

Kellogg EA (2001) Evolutionary history of the grasses. Plant physiology. **125**(3): 1198-1205.

Khan MA and **Stace CA** (1999) Breeding relationships in the genus Brachypodium (Poaceae: Pooideae). Nordic Journal of Botany. **19**(3): 257-269.

Kihara H (1924) Cytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rucksicht auf das Verhalten der Chromosomen und die Sterilitat in den Bastarden. Mem. Coll. Sci., Kyoto Imp. Univ. **1**: 1-200.

Kihara H (1930) Genome analyse bei Triticum und Aegilops. Cytologia. **1**(3): 263-284.

Kindt TJ, Mandy W and **Todd CW** (1970) Association of allotypic specificities of group a with allotypic specificities A11 and A12 in rabbit immunoglobulin. Biochemistry. **9**(9): 2028-2032.

Korpelainen H (2004) The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes. Naturwissenschaften. **91**(11): 505-518.

Lewis CE and **Doyle JJ** (2001) Phylogenetic utility of the nuclear gene malate synthase in the palm family (Arecaceae). Molecular phylogenetics and evolution. **19**(3): 409-420.

Loureiro J, Kopecký D, Castro S, Santos C and **Silveira P** (2007) Flow cytometric and cytogenetic analyses of Iberian Peninsula Festuca spp. Plant Systematics and Evolution. **269**(1): 89-105.

Lu RS, Li P and **Qiu YX** (2016) The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: comparative genomic and phylogenetic analyses. Front Plant Sci. **7**:2054.

Mason-Gamer RJ, Orme NL and **Anderson CM** (2002) Phylogenetic analysis of North American Elymus and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. Genome. **45**(6): 991-1002.

Matsuoka Y, Yamazaki Y, Ogihara Y and **Tsunewaki K** (2002) Whole chloroplast genome comparison of rice, maize, and wheat: implications for chloroplast gene diversification and phylogeny of cereals. Molecular Biology and Evolution. **19**(12): 2084-2091.

McFadden ES and **Sears ER** (1946) The origin of Triticum spelta and its free-threshing hexaploid relatives. Journal of heredity. **37**(4): 107-116.

Morgan DR and **Soltis DE** (1993) Phylogenetic relationships among members of Saxifragaceae sensu lato based on rbcL sequence data. Annals of the Missouri botanical Garden. 631-660.

Neel MC and **Cummings MP** (2004) Section-level relationships of North American Agalinis (Orobanchaceae) based on DNA sequence analysis of three chloroplast gene regions. BMC Evolutionary Biology. **4**(1): 1-12.

Nishikawa T, Salomon B, Komatsuda T, von Bothmer R and **Kadowaki KI** (2002) Molecular phylogeny of the genus Hordeum using three chloroplast DNA sequences. Genome. **45**(6): 1157-1166.

Ozdemir BS, Hernandez P, Filiz E and **Budak H** (2008) Brachypodium genomics. International Journal of Plant Genomics. **2008**.

Saadullah Z and **Khan M Zaib-u-Nisa A** (2016) Identification of the grass family (Poaceae) by using the plant DNA barcodes *rbcl* and *matK*. Journal Bio. Env. Sci. 2016. **8**(5): 175-186 Online available at http://www.innspub.net

Šmarda P, Bureš P, Horová L, Foggi B and **Rossi G** (2008) Genome size and GC content evolution of Festuca: ancestral expansion and subsequent reduction. Annals of botany. **101**(3): 421-433.

Soltis DE, Soltis PS, Clegg MT and **Durbin M** (1990) *rbcL* sequence divergence and phylogenetic relationships in Saxifragaceae sensu lato. Proceedings of the National Academy of Sciences. **87**(12): 4640-4644.

Watson L and **Dallwitz MZ** (1992) The grass genera of the world. PP. 1024. ISBN: 0-85198-802-4

Yang DY, Fushimi H, Cai SQ and **Komatsu K** (2004) Molecular analysis of Rheum species used as Rhei Rhizoma based on the chloroplast matK gene sequence and its application for identification. Biological and Pharmaceutical Bulletin. **27**(3): 375-425.

Zeid M, Yu J, Goldowitz I, Denton M, Costich DE, Jayasuriya C, Saha M, Elshire MR, Benscher D and **Breseghello F** (2010) Cross-amplification of EST-derived markers among 16 grass species. Field Crops Research. **118**(1): 28-35.

(Manuscript received on 10 May, 2023; revised on 10 June, 2023)