The European Journal of Plant Science and Biotechnology ©2012 Global Science Books

The Hungarian Gene Bank Collections of Common Millet (*Panicum miliaceum*) and the Application to Conservation Genetics

Gábor Gyulai[1,2*] • Richárd Lágler[1] • László Holly[2] • Lajos Horváth[2]

[1] St. István University, Institute of Genetics and Plant Breeding, Gödöllő 2103, Hungary
[2] Research Center of Agrobiodiversity, Tápiószele 2766, Hungary

Corresponding author: * **gyulai.gabor@mkk.szie.hu**

ABSTRACT

Over the past decades, the gene bank of RCAT (Research Center of Agrobiodiversity, Tápiószele, Hungary), specialized for field and vegetable crops, has collected 250 accessions of heirlooms, old varieties, landraces and breeding materials of common millet (*Panicum milaceum*). The accessions have been characterized by descriptor list and compiled according to the international gene bank standards (IPGRI). In this study, a short history of the RCAT, and a case study of common millet accessions is presented. Molecular analysis of DNA samples of twenty current millets and two archaeological seed samples excavated from the 4th and 15th centuries are compared at loci of *gln*4 (nuSSR), ISSRs and 18S mtDNA. The *in silico* analyses at loci of ITS1-5.8S-ITS2 (nuDNA), *rbc*L (cpDNA) and protein RuBisCO are also presented to trace molecular events that occurred during the evolution and domestication of common millet.

Keywords: DNA sequence analysis, domestication of common millet, gene bank
Abbreviations: aDNA, ancient DNA; **cpDNA**, chloroplast DNA; **nuDNA**, nuclear DNA; **PCR**, Polymerase Chain Reaction; **RCAT**, Research Centre of Agrobiodiversity (Tápiószele, Hungary); **SSR**, Simple Sequence Repeat DNA

INTRODUCTION

The Research Centre of Agrobiodiversity (RCAT) at Tápiószele (Hungary), founded in 1958, has been responsible for the development and maintenance of genetic resource collections of field and vegetable crops. In order to fulfill these requirements, RCAT performs all components of gene bank activity, including the collection, multiplication, characterization, medium and/or long term storage, documentation and distribution of germplasm samples. In total, 47,000 unique accessions of more than 900 crops and their wild relatives are now maintained by RCAT.

The first *Panicum* varieties (1-12) were introduced to Tápiószele in the 1950's, and a further 121 accessions were collected by 1970, and now, it consists of 259 accessions. There are 58 accessions of Hungarian origin, and among them 28 are classified as landrace and local ecotype.

The remaining larger part of the *Panicum* collection was obtained from about twenty different countries, and characterized by using a descriptor list, compiled in accordance with the international gene bank standards including 29 morphological and phenological characters (IPGRI 1985).

Common millet (*Panicum miliaceum*) is one of the most ancient grain crops of humanity with the first historical reports from 5000–3200 bc (Ho 1977). *Panicum* became a typical foods of Sumer and Northern India, together with barley (*Hordeum vulgare*), in about 2500 bc. For the nations of steppic Scythia, such as the Celtics, and Hungarians, in 2000 bc, the common millet was the first crop that produced two harvests in one year (Gyulai *et al.* 2011b). In the ancient Chinese 'Book of Poetry' (*Shih Ching*), written about 1000–500 bc, nine poems mention common millet (Keng 1974). *Panicum* spread from the Steppes through Europe via tribes of the Celts, Sarmatians, Huns, Avars and Hungarians, and also through the region of the 'Fertile Crescent' and Africa (Harlan 1971). It was the *milium* of Romans (Smith 1977). Millet was introduced to North America in the 17th century (Colosi and Schaal 1997) and several new cultivars were registered recently. Before the dominating cultivation of wheat and maize, common millet had been an important staple food in the Hungarian agriculture, but by the 20th century, its significance declined and now it is used mainly as a supplementary and second crop.

Morphological (Gyulai *et al.* 2011a; Janick *et al.* 2011) and molecular analyses (Gyulai *et al.* 2011b) have been used for improving the utilization and conservation strategies of plant genetic resources for many decades. Ancient plant varieties have provided unique data for tracing crop domestication and evolution. aDNA sequence analyses of microsatellites (nuSSRs, and ISSRs), as highly species-specific probes (Tóth *et al.* 2000; Lágler *et al.* 2005), and organelle DNAs (cpDNA and mtDNA) also provided a large amount of data by excluding laboratory contaminations (Gugerli *et al.* 2004; Lágler *et al.* 2005; Szabó *et al.* 2005; Gyulai *et al.* 2006). Microsatellite analysis was applied to herbarium samples of about 100-yr old common reed (*Phargmites australis*) to track plant invasion in North America (Saltonstall 2003). Allelic diversity of SSRs were also reliably detected in aDNAs of 4000-year old seagrass samples (*Posidonia oceanica*) (Raniello and Procaccini 2002). Melon (*Cucumis melo*) specific SSRs were used to identify an *inodorus* type melon recovered from aDNA of the 15th cent seed remains (Szabó *et al.* 2005). Flesh type of medieval watermelon (*Citrullus lanatus*) was also identified from aDNA samples (Tóth *et al.* 2007), and organellum (mtDNA) specific fragments were isolated from the 15th cent *Panicum* (Gyulai *et al.* 2006).

Multiple sequence alignment programs for *in silico* analysis of BioEdit Sequence Alignment Editor (North Carolina State University, USA; Hall 1999), MULTALIN (Combet *et al.* 2000), CLUSTAL W (Thompson *et al.* 1994), FastPCR (Kalendar *et al.* 2009) and BLAST (*Basic Local Alignment Search Tool*) of NCBI (*National Center for Bio-*

Received: 31 May, 2011. Accepted: 28 December, 2011.

Original Research Paper

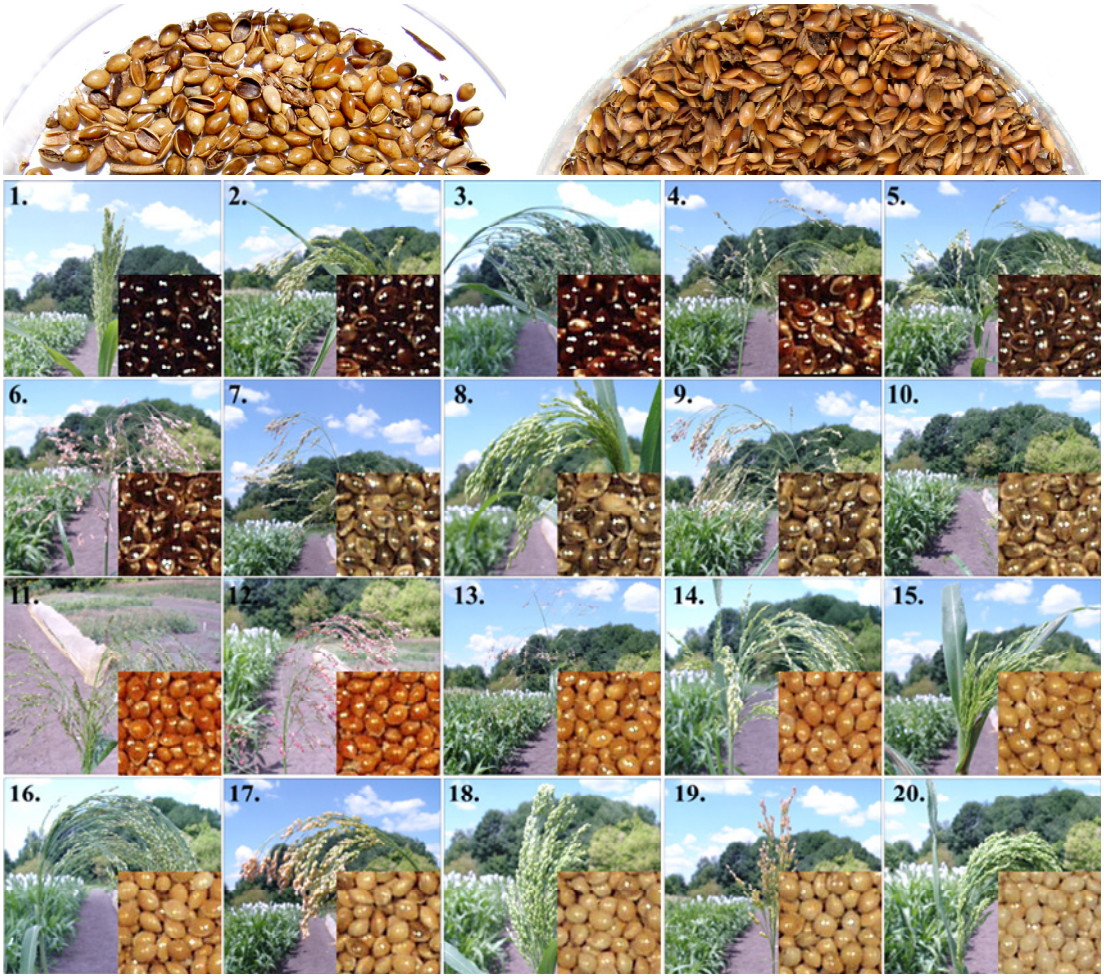

Fig. 1 Seeds and panicle types of common millet (*Panicum miliaceum*) studied. Wet-sieved and selected seeds excavated from the 4th cent (Mongolia) (*left up*) and 15th cent (Árpád King's Palace, Budapest, Hungary) (*top right*), and twenty (1-20) current cultivars (*bottom*) of the RCAT collection (Tápiószele, Hungary) used for comparative study. Cultivar names, origin (HUNG – Hungary, ITA – Italy, POL – Poland, RUS – Russia) and RCAT # are: (1) 'Tápió-A' (POL) (RCAT073416). (2) 'Tápió-C' (ITA) (RCAT073585). (3) 'Tápiói barna' (HUN) (RCAT017521). (4) 'Debreceni barnamagvú' (HUN) (RCAT017513). (5) 'Tápiószentmártoni' (lr., landrace) (HUN) (00185/01). (6) 'Tápió-B' (HUN) (RCAT017280). (7) 'Fertődi' (HUN) (RCAT017272). (8) 'Püski' (lr.) (HUN) (RCAT017296). (9) 'Rábaszentandrási' (lr.) (HUN) (RCAT017297). (10) 'Bolgár' (HUN) (RCAT017267). (11) 'Fertődi piros' (HUN) (RCAT017291). (12) 'Kecskeméti' (lr.) (HUN) (RCAT017527). (13) 'Omszkoje' (RUS) (02546/00). (14) 'Jászberényi' (lr.) (HUN) (RCAT017555). (15) 'Császárréti-2' (HUN) (RCAT017277). (16) 'Nyíregyházi' (lr.) (HUN) (RCAT017526). (17) 'Tápiószele-A' (RUS) (RCAT017509). (18) 'Fertődi fehér' (HUN) (RCAT017290). (19) 'Martonvásári-3' (HUN) (RCAT017285). (20) 'Mesterházi' (lr.) (HUN) (RCAT017494).

technology and Information) server (Altschul *et al.* 1997) provide powerful tools for analyzing sequences of DNAs, aDNAs, RNAs and proteins available in the data banks such as NCBI (Altschul *et al.* 1997) as used in the study presented here. Molecular dendrograms based on the sequence alignments were found to be the most comprehensive forms of sorting data with the ability to revealing dis/similarities and evolutionary relatedness of species compared (Kajita *et al.* 2001; Aliscioni *et al.* 2003; Wojciechowski *et al.* 2004; Lágler *et al.* 2005; Szabó *et al.* 2005; Mansour *et al.* 2009).

MATERIALS AND METHODS

Seed samples

Seed remains of common millet (*P. miliaceum*) from the 4th cent site (3rd grave, Darhan, Mongolia, 1969) (50 seeds); and 15th cent site (150 seeds) (8th well, Mansion Teleki, King's Palace of Árpád Dynasty, Budapest, Hungary, 2003) were excavated (Nyékhelyi 2003) and analyzed according to Lágler *et al.* (2005) and Gyulai *et al.* (2006, 2011b). Wet-sieved sediment seeds were processed by floatation followed by seed sorting and identification in the laboratory. Intact seeds were separated from damaged grains under microscope. For comparative analysis, 20 current *Panicum* cultivars were selected from the RCAT gene bank (**Fig. 1**).

aDNA and DNA extraction

Seeds were surface sterilized through a pre wash with detergent under tap water, soaked and shaken in commercial NaOCl 5% for 10 min, and a final wash with sterilized H_2O in a laminar hood. Aseptic seeds were ground in an aseptic mortar with liquid nitrogen. For DNA and aDNA (ancient DNA) extraction, CTAB (cethyltrimethylammonium bromide) method was used (Murray and Thompson 1980) in pools according to Biss *et al.* (2003) and Gyulai *et al.* (2006). DNA of 20 current cultivars (**Fig. 1**) was also extracted from aseptic seeds (0.1 g) in CTAB buffer, followed by an RNase-A treatment (Sigma, R-4875) for 30 min at 37°C in each case. The quality and quantity of extracted DNA samples was measured (2 µl) by a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies, Delaware, USA – BioScience, Budapest, Hungary), adjusted to concentration of 30 ng/µl with ddH_2O and subjected to PCR (Polymerase Chain Reaction) amplification.

PCR analysis

A minimum of two independent DNA preparations from each sample was used for PCR according to Lágler *et al.* (2005) and Gyulai *et al.* (2006). Fragments were scored for the presence versus absence of band profiles. Each successful reaction with scorable bands was repeated at least twice. Negative control samples of

DNA-free, primer-free, *Taq*-free and water were applied in each PCR runs.

Hot Start PCR was combined with Touchdown PCR using PCR Master Mix (Promega, #M7502) in a total volume of 25 µl, including dNTPs (dATP: 2'-deoxyadenosine 5'-triphosphate, sodium salt; dCTP: 2'-deoxycytidine 5'-triphosphate, sodium salt; dGTP: 2'-deoxyguanosine 5'-triphosphate, sodium salt; and dTTP: 2'-deoxythymidine 5'-triphosphate, sodium salt), buffer and heat stable *Taq* DNA polymerase. Touchdown PCR was performed by decreasing the annealing temperature from 66 to 56°C by 0.7°C/30 s decrements per cycle with each of the initial 12 cycles (PE 9700, Applied Biosystems), followed by a '*touchdown*' annealing temperature for the remaining 25 cycles at 56°C for 30 s with a final cycle of 72°C for 10 min and hold at 4°C.

PCR primers

PCR primer (oligonucleotide) pairs (pairs of short, single stranded DNA strands) complementary to target sequences of DNA samples were used for PCR amplifications. For SSR analysis at the *gln*4 locus primer pair agc aga acg gca agg gct act // ttt ggc aca cca cga cga) (NCBI# D14577) was used (Chin *et al.* 1996). For ISSR (inter simple sequence repeat of DNA) analysis (Zietkiewitz *et al.* 1994) primers of Fv808 - $(ag)_8$c; Fv821 - $(gt)_8$t; and Fv835 - $(ag)_8$(t/c)c (Cekic *et al.* 2001) were used (Lágler *et al.* 2005). Mitochondrial-DNA (mtDNA) at the 18S-5S-rDNA locus was amplified by primer pair gtg ttg ctg aga cat gcg cc // ata tgg cgc aag acg att cc (Petit *et al.* 1998).

ALF fragment analysis

Microsatellite fragments were forwarded for ALF (Automated Laser Fluorometer) analysis using ALF Express II (Pharmacia-Amersham, AP-Hungary, Budapest, Hungary). After heat denaturation, PCR products (2 µl) were loaded and analyzed using short thermoplate (#181123-60) with 'ReproGel High Resolution' PAGE gel (24%) (Amersham) of 40 samples capacity. Before loading, PAGE gels were UV-linkage for 15 min under ReproSet (#18-1125-64). The separation was carried out by ALF Express II DNA Analyser (#18-1125-07) at 850 V, 50 mA, 50 W, 50°C for 120 min followed by computer analysis with ALFwin Fragment Analyser 1.03 program (#18-1125-92) according to Röder *et al.* (1998), Huang *et al.* (2002) and Gyulai *et al.* (2006). The 5′-end of one oligonucleotides of each primer pair was fluorescently labeled by Cy5-CE phosphoramidite dye, indodicarbocyanine 5, 1′-*O*-(4-monomethoxytrityl)-1-*O*-(2-cyanoethyl)-(*N*,*N*-diisopropyl)-phosphoramidite (in short Cy5), which posses maximal absorbance at 643 nm, and emission at 667 nm. Cy5 stained external and internal molecular weight DNA standards were applied in each run.

Sequencing

Amplified fragments were isolated from the agarose gel with a spin column (Sigma, 56501) and subjected to automated fluorescent DNA sequencer (ABI PRISM 3100 Genetic Analyzer, Applied Biosystems, Hungary) using 16-capillary electrophoresis system. The DNA chain terminating dideoxy nucleotides (ddNTPs; 2',3'-dideoxynucleotide triphospates of ddATP, ddCTP, sddGTP, and ddTTP) were fluorescently labeled (at 5′ ends) with the four different fluorescent dyes of (1) FL (FAM) blue (5-carboxyfluoroscein); (2) NED (TAMRA) yellow (N,N,N′,N′-tetramethyl-6-carboxyrhodamine); (3) CRX (ROX) red (6-carboxy-X-rhodamine), and (4) JOE, green (2′,7′-dimethoxy-4′,5′-dichloro-6-carboxyfluorescein), and applied in 1% of the total dATP according to Sanger *et al.* (1977) following the ABI PRISM 310 manufacturer's protocol using BigDye terminator sets. For sequencing, AmpliTaq Gold™ Polymerase (0.5 U) was used in a total volume of 10 µl, containing genomic DNA (50 ng), 1X PCR buffer including 2.5 mM $MgCl_2$, dNTPs (200 µM each), and 20 pmol of each primer. Touchdown PCR for DNA amplification during sequencing were the same as used for fragment amplification.

Sequence analysis *in silico* and *in situ*

DNA sequences were analyzed by programs of BioEdit Sequence Alignment Editor (North Carolina State University, USA) (Hall *et al.* 1999), MULTALIN (Combet *et al.* 2000), CLUSTAL W (Thompson *et al.* 1994) and FastPCR (Kalendar *et al.* 2009). *In silico* analyses and the authenticity of aDNA sequences were aligned by BLAST program (Basic Local Alignment Search Tool) of NCBI server (Altschul *et al.* 1997). Molecular dendrograms were edited by computer programs MEGA4 (Tamura *et al.* 2007).

RESULTS AND DISCUSSION

Panicum is a large genus of about 450 species with diverse genetic background (**Fig. 2**). It includes common millet (*syn.*: proso millet) (*P. miliaceum*; 2n = 4x = 36), native to Asia (Aliscioni *et al.* 2003; Gyulai *et al.* 2011b). The perennial warm season bunchgrass *Panicum virgatum* (switchgrass) with polyploidy series (2n = 18, 36, 54, 72, 90, and 108) is native to the central North American tallgrass prairie. Its complete cpDNA sequence (139.619 bp) is available (NCB NC_015990*). Panicum dichotomiflorum* (fall panic grass) is a serious grass weed in the USA. *Panicum maximum* (Guineagrass) is a highly productive forage grass (Aliscioni *et al.* 2003). Based on the highly conserved DNA sequences at the nuclear ITS1-5.8S-ITS2 sequences (*Iternal Tanscribed Spacer1 - 5.8S ribosomal RNA - ITS2*) (Allan and Porter 2000), *Panicum* species clustered in three separate clades (**Fig. 2**).

Nuclear SSR, ISSR and mtDNA analyses of 15th cent *Panicum* and compared to current varieties

The aDNA samples of the 1,600-year old common millet (4th cent), and the 600-year old sample (15th cent) showed amplifiable DNA compared to current cultivars (Gyulai *et al.* 2006). After sequencing (218 bp), SSR fragment (Messier *et al.* 1996; Taylor *et al.* 1999) amplified at the *gln*4 locus of the 600-year old aDNA and the comparative current *Panicum* cv. 'Topáz' showed identical millet specific DNA sequences without SNPs (*single nucleotide polymorphism*) supposed to occur during the past centuries (Gyulai *et al.* 2006). Due to the high rate of aDNA damage, no microsatellites were amplified from the 1,600-year old millet.

BLAST and sequence alignments revealed sequence homology of *gln*4 (nuDNA) sequences of medieval and current *P. miliaceum* to *gln*4 loci of the current monocot species with the closest similarity to *Zea mays* (**Fig. 3A**). The *gln*4-linked *Panicum* SSR aligned to all (24) *Panicum* SSRs available (most of them with a consensus 3′ flanking site sequence of tagtccagatctaagcaagagaatcactag) (NCBI) showed the closest sequence similarity to EF117725 and EF117726 (**Fig. 3B**).

Based on the former (Lágler *et al.* 2005) and current ISSR-PCRs, a total of 242 ISSR fragments of 15 alleles at three loci of Fv808 (6 alleles), Fv821 (5 alleles), ISSR-Fv835 (4 alleles) were amplified, and the dendrogram analysis of fragments patterns revealed (**Fig. 3C**) that 15th cent millet showed the closest molecular similarity to current cv. 'Tápiószentmártoni' (landrace, #5) and 'Omszkoje' (#13), which indicates a phenotype reconstructing image of the medieval *Panicum* (**Fig. 1**).

The 600-year old *Panicum* organelle DNA amplified at the 18S mtDNA locus (191 bp) was identical to current *Panicum miliaceum* sequence. After alignment it showed the closest similarity to *Triticum aestivum* (**Fig. 4A**). All of the other monocots (*Oryza*, *Zea*, *Secale*, *Asparagus*) grouped in a separate clade (**Fig. 4A**). This pattern of related species was different from the phylogram based on the complete mtDNA sequences (**Fig 4B**), *Panicum virgatum* available clustered in single clade.

In the other organelle, the chloroplast phylogeny analyzed *in silico* at the *rbc*L gene locus, *Panicum* species grouped in three different clades including *P. miliaceum*

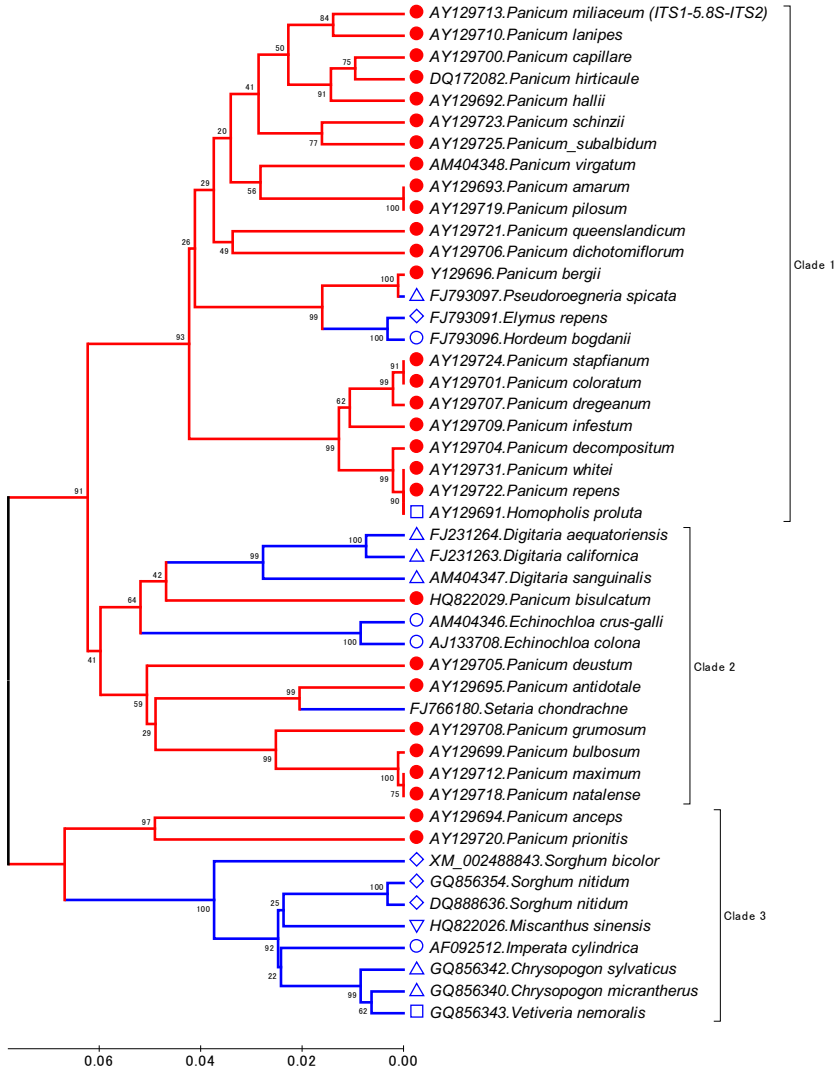

Fig. 2 ***In silico*** **ITS phylogram of** ***Panicum*** **and related species**. ITS1-5.8S-ITS2 (*Internal Transcribed Spacer1 - 5.8S ribosomal RNA - ITS2*) sequences of NCBI (Altschul *et al.* 1997) were collected, aligned by BioEdit (Hall 1999) and the consensus sequences (752 bp) were analyzed by phylogram (UPGMA, MEGA4, Tamura *et al.* 2007). Relative genetic distance (*scale*), bootstrap (x1000) values, and NCBI accession numbers are indicated.

closest to *P. repens* (**Fig. 5**) similar to the three subclades of ITS phylogram (**Fig. 2**). However, ITS phylogram indicated *P. lanipes* to be closest to *P. miliaceum* (**Fig. 2**). These results might indicate the individual origin and lives of the three DNA bearing organelles of nucleus, mitochondria and chloroplasts, even in the case of the continuous double way gene transfer between mitochondria and nucleus; however, chloroplasts transfer cpDNA to both nucleus and mitochondria but never receive DNA from them (Huang *et al.* 2003). The different speeds of changes in the numbers of synonymous (silent) and non-synonymous (amino acid changes) nucleotides (Chamary *et al.* 2009) occurring during speciation (evolution) (**Fig. 6**), might also account for the different phylograms of nuDNA (**Fig. 2**), mtDNA (**Fig 4A, 4B**) and cpDNA (**Fig. 5**). However, *Panicum* was reported to be monophyletic based on phylogeny of *ndhF* genes of cpDNA (Aliscioni *et al.* 2003).

In conclusion, common millet (*Panicum miliaceum*) gene bank accessions of RCAT (Tápiószele, Hungary) have provided plant material not only for breeders, but also for the molecular and morphological reconstructions of archaeological samples, the results of which might be useful to trace the speciation, evolution, and domestication of *Panicum* species.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Renée Malone (DIT, Dublin, Ireland) and Dr. Luther waters Jr. (Auburn University AL. USA) for the comments on the manuscript. Research was supported in part by the Project TÁMOP-4.2.2.B-10/1 “Development of a complex educational assistance and support system for talented students and prospective researchers at the Szent István University”.

REFERENCES

Aliscioni SS, Giussani LM, Zuloaga FO, Kellogg EA (2003) A molecular phylogeny of *Panicum* (Poaceae: *Paniceae*): Tests of monophyly and phylogenetic placement within the Panicoideae. *American Journal of Botany* **90**, 796-821

Allan GJ, Porter JM (2000) Tribal delimitation and phylogenetic relationships of *Loteae* and *Coronilleae* (Faboideae: *Fabaceae*) with special reference to *Lotus*: Evidence from nuclear ribosomal ITS sequences. *American Journal of Botany* **87**, 1871-1881

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipmand DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402

Biss P, Freeland J, Silvertown J, McConway K, Lutman P (2003) Successful amplification of rice chloroplast microsatellites from century-old grass samples from the park grass experiment. *Plant Molecular Biology Reporter* **21**, 249-257

Cekic C, Battey NH, Wilkinson MJ (2001) The potential of ISSR-PCR primer-pair combinations for genetic linkage analysis using the seasonal flowering locus in *Fragaria* as a model. *Theoretical and Applied Genetics* **103**, 540-546

Chamary JV, Hurst LD (2009) The price of silent mutations. *Scientific American* **300**, 46-53

Chin E, Senior L, Shu H, Smith JSC (1996) Maize simple repetitive DNA sequences: Abundance and allele variation. *Genome* **39**, 866-873

Colosi JC, Schaal BA (1997) Wild proso millet (*Panicum miliaceum*) is

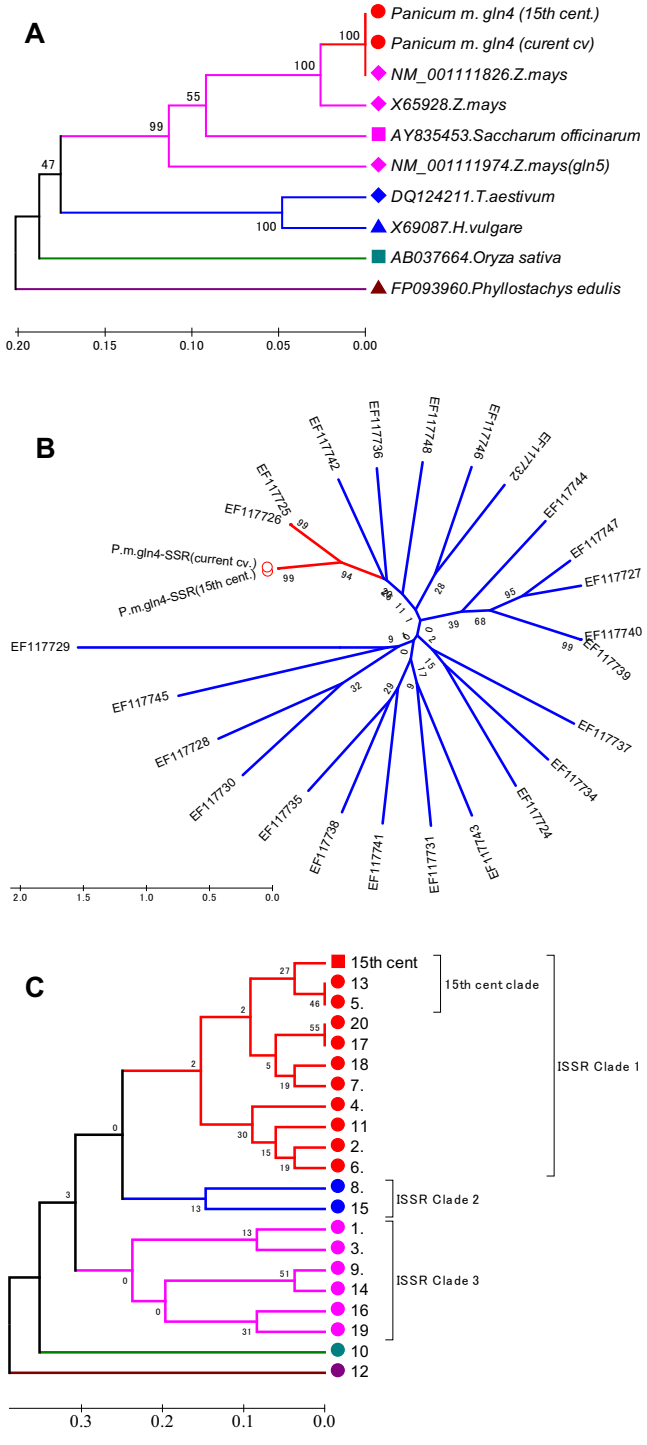

Fig. 3 (**A**) The *gln4* phylogran of *Panicum* and related species. The nuDNA SSR fragments (218 bp) linked to *gln*4 gene (*glutamine synthetase* 4) of the 15th cent millet and current *Panicum* cv. 'Topáz' was amplified, sequenced and aligned (NCBI, Altschul *et al.* 1997) followed by phylogram analysis (MEGA4, Tamura *et al.* 2007). (**B**) SSR cladogram at the *gln4*-linked loci of *Panicum miliaceum*. The *gln*4 linked SSR fragment, after sequencing, were aligned to all (24) *Panicum* SSR sequences available in NCBI (Altschul *et al.* 1997) followed by phylogram (UPGMA, radial) analysis (MEGA4, Tamura *et al.* 2007). Relative genetic distance (*scale*), bootstrap (x1000) values, and NCBI accession numbers are indicated. (**C**) The ISSR cladogram of current common millet (*Panicum miliaceum)* cultivars (1 to 20) and the 15th cent millet. A total of 242 ISSR fragments of 15 alleles amplified at three loci of Fv808 (6 alleles), Fv821 (5 alleles) and Fv835 (4 alleles) were compared. Relative genetic distance (*scale*), bootstrap (x1000) values (small numbers) are indicated. Cultivar names are listed in **Fig. 1**.

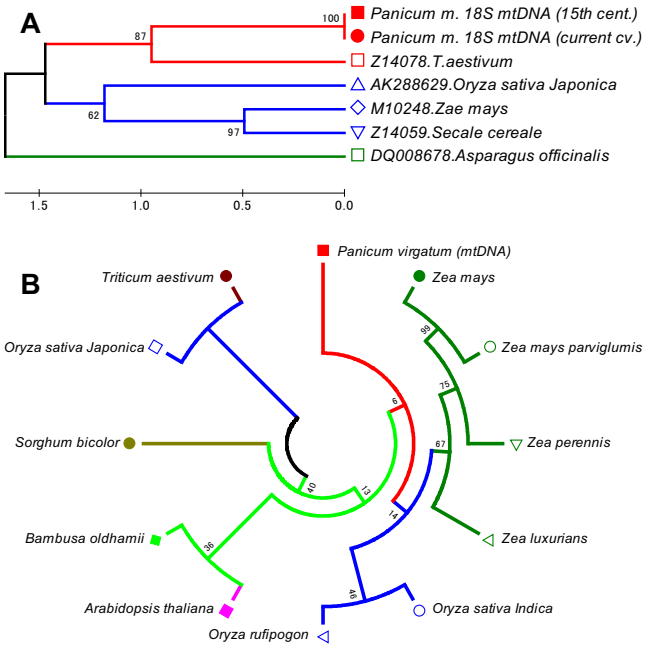

Fig. 4 (**A**) The 18S mtDNA cladogram. The 600-year-old aDNA fragment of *Panicum miliaceum* amplified at the 18S mtDNA locus, was sequenced (191 bp) and blasted to NCBI database (Altschul *et al.* 1997) followed by phylogram analysis (MEGA4, Tamura *et al.* 2007). Relative genetic distance (*scale*), bootstrap (x1000) values, and NCBI accession numbers are indicated. (**B**) Total mtDNA phylogram *in silico*. The small complete mtDNA of *Panicum virgatum* (139,619 bp) was blasted (NCBI, Altschul *et al.* 1997) to total mtDNAs of eleven related species followed by dendrogram analysis (MEGA4, Tamura *et al.* 2007). Relative genetic distance (*scale*), bootstrap (x1000) values, NCBI accession numbers, and the dicot outgroup of *Arabidopsis thaliana* are indicated. The mitom (mtDNA) sizes are: *Arabidopsis t.* (Y08051): 366,924 bp; *Bambusa oldhamii* (EU365401): 509,941 bp; *Oryza rufipogon* (AP011076): 559,045 bp; *Oryza sativa* Indica (AP011077): 434.735 bp; *Oryza sativa* Japonica (BA000029): 490,520 bp; *Panicum virgatum* (HQ822121): 139,619 bp; *Sorghum bicolor* (DQ984518): 468,628 bp; *Triticum aestivum* (AP008982): 452-528 bp; *Zea luxurians* (DQ645537): 538,368 bp; *Zea mays* (AY506529): 569,630 bp; *Zea mays parviglumis* (DQ645539): 680,603 bp; *Zea perennis* (DQ645538): 470,354 bp.

genetically variable and distinct from crop varieties of proso millet. *Weed Science* **45**, 509-518

Combet C, Blanchet C, Geourjon C, Deléage G (2000) NPS@: Network Protein Sequence Analysis. *Trends in Biochemical Sciences* **25**, 147-150

Gugerli F, Parducci L, Petit RJ (2005) Ancient plant DNA: Review and prospects. *New Phytologist* **166**, 409-418

Gyulai G, Humphreys M, Lágler R, Szabó Z, Tóth Z, Bittsánszky A, Gyulai F, Heszky L (2006) Seed remains of common millet from the 4th (Mongolia) and 15th (Hungary) centuries: AFLP, SSR and mtDNA sequence recoveries. *Seed Science Research* **16**, 179-191

Gyulai G, Malone RP, Heszky L, Waters L Jr., Kiss E (2011a) Morphogenetics of ancient *Vitis* seeds from antiquity (3rd and 11th - 15th Cents Hungary) - A genotype reconstruction. In: Gyulai G (Ed) *Plant Archaeogenetics*, Nova Science Publishers Inc., New York, pp 41-48

Gyulai G, Humphreys MO, Lágler R (2011b) aDNA extraction and molecular analysis from seed remains of ancient common millets (*Panicum miliaceum*) (4th and 15th CENT). In: Gyulai G (Ed) *Plant Archaeogenetics*, Nova Science Publishers Inc., pp 51-86

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98

Harlan JR (1971) Agricultural origins: Centers and noncenters. *Science* **174**, 468-473

Ho P-T (1977) The indigenous origins of Chinese agriculture. In: Reed CA (Ed) *Origins of Agriculture*, Mouton Publishers, Paris, pp 413-418

Huang XQ, Börner A, Röder MS, Ganal MW (2002) Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics* **105**, 699-707

Huang CY, Ayliffe MA, Timmis JN (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **422**, 72-76

IBPGR (1985) Descriptors for *Panicum miliaceum* and *P. sumatrense*. International Board for Plant Genetic Resources, Rome. IBPGR internal document

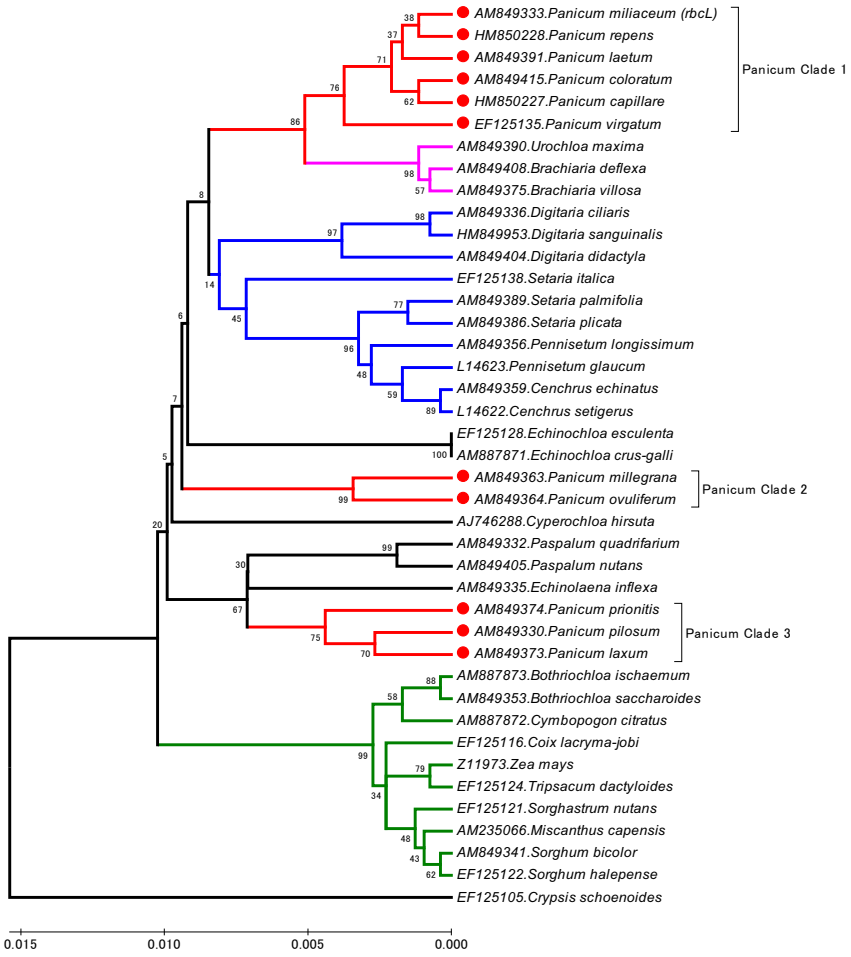

Fig. 5 The cpDNA phylogram of *rbc*L genes *in silico*. After data mining and BLAST analysis (NCBI, Altschul *et al.* 1997) the *rbc*L gene sequences, including *Panicum*, were aligned followed by dendrogram analysis (MEGA4, Tamura *et al.* 2007). Relative genetic distance (*scale*), bootstrap (x1000) values, and NCBI accession numbers are indicated.

code: AGPG:IBPGR 85/52. http://www2.bioversityinternational.org/publications/Web_version/149/

Janick J, Daunay M-C, Paris HS (2011) Plant iconography - a source of information for archaeogenetics. In: Gyulai G (Ed) *Plant Archaeogenetics*, Nova Science Publishers Inc., pp 143-159

Kajita T, Ohashi H, Tateishi Y, Bailey CD, Doyle JJ (2001) *rbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. *Systematic Botany* **26**, 515-536

Kalendar R, Lee D, Schulman AH (2009) FastPCR software for PCR primer and probe design and repeat search. *Genes, Genomes and Genomics* **3**, 1-14 and also available online: http://primerdigital.com/fastpcr.html]

Keng H (1974) Economic plants of ancient north China as mentioned in *Shih Ching* (Book of Poetry). *Economic Botany* **28**, 391-410

Lágler R, Gyulai G, Humphreys M, Szabó Z, Horváth L, Bittsánszky A, Holly L, Heszky L (2005) Morphological and molecular analysis of common millet (*P. miliaceum*) cultivars compared to an aDNA sample from the 15th century (Hungary). *Euphytica* **146**, 77-85

Mansour A, Teixeira da Silva JA, Gyulai G (2009) Assessment of molecular (dis)similarity: The role of multiple sequence alignment (MSA) programs in biological research. *Genes, Genomes and Genomics* **3**, 23-30

Messier W, Li S-H, Stewart C-B (1996) The birth of microsatellites. *Nature* **381**, 483

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**, 4321-4325

Nyékhelyi BD (2003) Monumenta Historica Budapestinensia XII. Historical Museum of Budapest, Hungary, pp 1-102

Petit RJ, Demesure B, Dumolin S (1998) cpDNA and mtDNA primers in plants. In: Karp A, Isaac PG, Ingram DS (Eds) *Molecular Tools for Screening Biodiversity*, Chapman and Hall, London, pp 256-261

Raniello R, Procaccini G (2002) Ancient DNA in the seagrass *Posidonia oceanica*. *Marine Ecology – Progress Series* **227**, 269-273

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. *Genetics* **149**, 2007-2023

Saltonstall K (2003) Microsatellite variation within and among North American lineages of *Phragmites australis*. *Molecular Ecology* **12**, 1689-1702

Szabó Z, Gyulai G, Humphreys M, Horváth L, Bittsánszky A, Lágler R, Heszky L (2005) Genetic variation of melon (*C. melo*) compared to an extinct landrace from the Middle Ages (Hungary). *Euphytica* **146**, 87-94

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* **74**, 5463-5467

Smith PM (1977) Minor crops. In: Simmonds NW (Ed) *Evolution of Crop Plants*, Longman, London, pp 301-324

Taylor JS, Durkin JMH, Breden F (1999) The death of microsatellite: A phylogenetic perspective on microsatellite interruptions. *Molecular Biology and Evolution* **16**, 567-572

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680

Tóth G, Gáspári Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research* **10**, 967-981

Tóth Z, Gyulai G, Horváth L, Szabó Z, Heszky L (2007) Watermelon (*Citrullus l. lanatus*) production in Hungary from the Middle Ages. *Hungarian Agricultural Reseach* **2007/4**, 14-19

Wojciechowski MF, Lavin M, Sanderson MJ (2004) A phylogeny of legumes (*Leguminosae*) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. *American Journal of Botany* **91**, 1846-1862

Zietkiewicz E, Rafalski A, Labuda A (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* **20**, 176-183

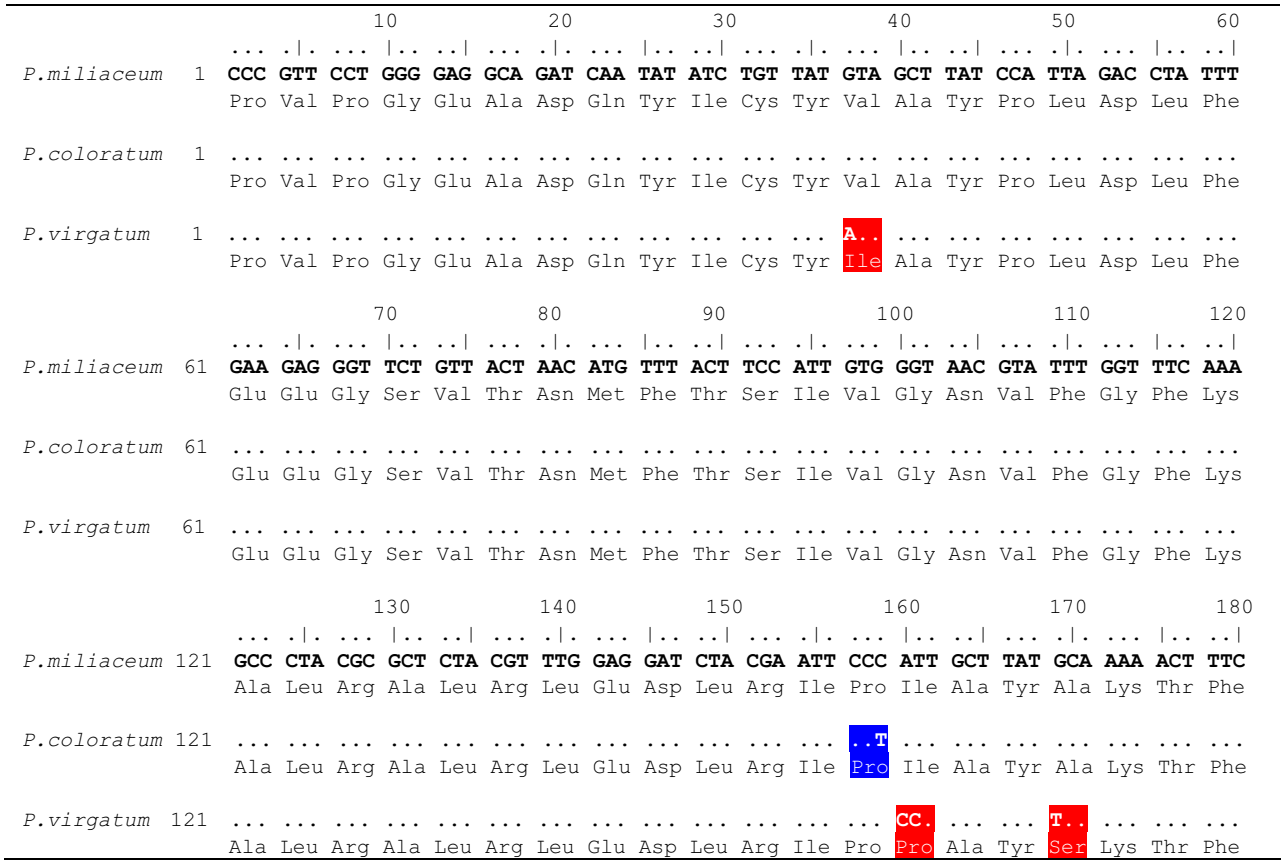

Fig. 6 Partial sequence alignments and protein translate of gene *RuBisCO* analyzed *in silico*. The genes of *RuBisCO* of three *Panicum* species of *P. miliaceum* (FR667683), *P. virgatum* (EF125135) and *P. coloratum* (AM849415) were aligned and translated (BioEdit, Hall 1999). The non-synonymous (*red*) and a synonymous (*blue*) nucleotide substitution were detected in the 180 bp stretches (*dots* indicate consensus nucleotides).