Theor Appl Genet (2001) 103:668–677

A. Benabdelmouna · M. Abirached-Darmency
H. Darmency

Phylogenetic and genomic relationships in *Setaria italica* and its close relatives based on the molecular diversity and chromosomal organization of 5S and 18S-5.8S-25S rDNA genes

Received: 28 August 2000 / Accepted: 27 January 2001

Abstract We have analyzed the phylogenetic and genomic relationships in the genus *Setaria* Beauv. including diploid and tetraploid species, by means of the molecular diversity of the 5S rDNA spacer and chromosomal organization of the 5S and 18S-5.8S-25S rDNA genes. PCR amplification of the 5S rDNA sequences gave specific patterns. All the species studied here share a common band of about 340 bp. An additional band of an approximately 300-bp repeat unit was found for *Setaria verticillata* and the Chinese accessions of *Setaria italica* and *Setaria viridis*. An additional band of 450 bp was found in the sole species *Setaria faberii*. Fluorescent in situ hybridization was used for physical mapping of the 5S and 18S-5.8S-25S rDNA genes and showed that they are localized at two separate loci with no polymorphism of chromosome location among species. Two chromosome pairs carrying the 5S and 18S-5.8S-25S rDNA clusters can now be unambiguously identified using FISH. Phylogenetic trees based on the variation of the amplified 5S rDNA sequences showed a clear separation into four groups. The clustering was dependent on the genomic composition (genome A versus genome B) and confirmed the closest relationship of *S. italica* and *S. viridis* accessions from the same geographical region. Our results confirm previous hypotheses on the domestication centers of *S. italica*. They also show the wide difference between the A and B genomes, and even clarify the taxonomic position of *S. verticillata*.

Keywords *Setaria* species · rDNA genes · Fluorescent in situ hybridization · 5S sequence analysis · Phylogeny

Communicated by H.C. Becker

A. Benabdelmouna · H. Darmency (✉)
Unité de Malherbologie et Agronomie, INRA,
BP 86510, 21065 Dijon, France
e-mail: darmency@dijon.inra.fr

M. Abirached-Darmency
Laboratoire de cytologie, INRA-SGAP,
BP 86510, 21065 Dijon, France

Introduction

Setaria Beauv. is a genus of about 125 species widely distributed in warm and temperate parts of the world. This genus is a member of the subfamily Panicoideae and the tribe Paniceae. It contains crop, wild and weed species with different breeding systems, life cycles and ploidy levels. *Setaria italica* (L.) P. Beauv., the foxtail millet, is an important grain crop used as a staple food in China, India and Japan, and is grown for silage and hay in North and South America, Australia and North Africa (Wanous 1990). Native species from Eurasia include *Setaria faberii* Herrm. (giant foxtail), *Setaria verticillata* (L.) P. Beauv. (bristly foxtail), *Setaria adhaerans* (Forsskal) Chiov. (bristly grass), *Setaria viridis* (L.) P. Beauv. (green foxtail) and *Setaria pumila* (Poiret) Roemer & Schultes (yellow foxtail). They are all aggressive colonizers in their native habitats and represent the most-widely distributed weedy foxtail species (Rominger 1962). The evolution and domestication history of *S. italica* have been studied based on morphological data and biochemical analysis. Three domestication centers have been suggested: China, Europe and a geographical area ranging from Afghanistan to Lebanon (Kawase and Sakamoto 1982, 1984, 1987; Jusuf and Pernès 1985; Darmency and Pernès 1987; Darmency et al. 1987; Li et al. 1995; Schontz and Rether 1999). Observations drawn from interspecific hybridization and hybrid pollen fertility suggest that the genus *Setaria* is organized into three gene pools. The primary gene pool is composed of the diploid species (2n = 2*x* = 18) *S. italica* and its putative wild ancestor *S. viridis* (Harlan and De Wet 1971). A secondary gene pool contains *S. adhaerans* (2n = 2*x* = 18) and the two allotetraploids *S. verticillata* and *S. faberii* (2n = 4*x* = 36; Li et al. 1942; Benabdelmouna et al. 2001). A tertiary gene pool contains *Setaria glauca* (or *S. pumila*, 4*x* to 8*x*) in addition to many other wild species (Zangré et al. 1992). However, genetic relationships among the *Setaria* species are still poorly understood. As mentioned by Li et al. (1998), understanding the genetic relationships between foxtail millet and other species of

the *Setaria* complex can provide a basis to secure successful parent selection and hybridization, as well as to organize germplasm, identify cultivars, and ensure sampling from a broad range of genetic variability.

Cytogenetic studies on foxtail millet and other *Setaria* species is difficult because of their flowering characteristics, their chromosome morphology and size. However, techniques such as fluorescent in situ hybridization (FISH) to ribosomal genes, that are very useful in studies of systematics and evolution (Mukai et al. 1991; Maluszynska and Heslop-Harrison 1993; Benabdelmouna and Darmency 1997; Zoldos et al. 1999), have not been performed on *Setaria* species so far. During the past decade, analysis of the molecular diversity of the rDNA genes (5S and 18S-5.8S-25S rDNA) and their chromosomal organization have already been shown to be suitable for the karyological characterization of species with small and similarly sized chromosomes as well as for establishing taxonomic and phylogenetic analysis (Appels and Baum 1991; Maluszynska and Heslop-Harrison 1991; Baum and Appels 1992; Baum and Johnson 1994; Schmidt et al. 1994; Kolipara et al. 1997; Scoles et al. 1988; Zoldos et al. 1999). As in all higher eukaryotes, the 5S and 18S-5.8S-25S rDNA genes are organized into separate loci. They are located at one or more pairs of sites that can be visualized on chromosomes by FISH (Castilho and Heslop-Harrison 1995; Benabdelmouna and Darmency 1997; D'Hont et al. 1998). The genes that code for 5S rRNA are organized into clusters of tandem repeats with up to thousands of copies of repeated units (Appels and Honeycut 1986). Every repeat unit consists of a transcribed region of approximately 120 bp and a non-transcribed spacer region varying in size and sequence. This variation within the non-transcribed spacer region was found to be useful for the phylogenetic reconstruction of species and even for cultivar identification (Kolchinsky et al. 1991; Baum and Bailey 1997; Baum and Johnson 1998). The 18S-5.8S-25S ribosomal RNA genes are present in several hundreds of tandemly repeated units of the three genes with intergenic spacers, organized in one or more clusters within the genome (Appels et al. 1980; Saghai-Maroof et al. 1984; Flavel 1986; Rogers and Bendich 1987). We report here the study of the genomic and phylogenetic relationships of five Eurasian species of the genus *Setaria* using the 5S rDNA molecular diversity and the intrachromosomal organization of the 5S and 18S-5.8S-25S rDNA genes by FISH on metaphase chromosomes.

Materials and methods

Plant materials

The *Setaria* wild species and the *S. italica* lines used in this study, as well as their characteristics, are listed in Table 1. Their taxonomical status has been checked according to the International Nomenclature Code (Kerguélen 1999) and their ploidy level and genome structure confirmed in a recent study (Benabdelmouna et al. 2001). Plants were grown in a greenhouse at 20°C for a 16-h day length and 15°C during the night.

PCR amplification, Southern hybridization, and sequence analysis of the 5S rDNA genes

Amplification of 5S rRNA was made by PCR amplification of total genomic DNA in the presence of the 5S-specific primers 5′-GTGACCTCCTGCGAAGTCCT-3′ and 5′-CCCATCCGTGTACTACTCTC-3′ (D'Hont et al. 1998). These primers were designed on the basis of conserved regions of the 5S wheat and barley sequences. The standard amplification mixture (50 µl) contained 5 ng of genomic template DNA, 0.4 M of each primer, 250 µM each of dATP, dCTP, dGTP and dTTP (Pharmacia), 5 µl of reaction buffer [170 mM $(NH_4)_2 SO_4$, 15 mM $MgCl_2$, 670 mM Tris-HCl, pH 8.8] and 2 units of *Taq* polymerase (BRL). Thirty amplification cycles were performed on a PTC 200 thermal-cycler (MJ Research). Each cycle consisted of denaturation at 93°C for 55 s, annealing at 60°C for 25S and primer extension at 72°C for 35 s. Amplification products were separated by electrophoresis on 2% agarose gels. They were then stained with ethidium bromide and visualized using an ultraviolet transilluminator. Verification of the 5S rDNA nature of the amplified sequences and Southern analysis were both performed using a heterologous 5S rDNA probe from *Petunia hybrida* (Benabdelmouna and Darmency 1997). The PCR gels and the BamHI-digested total genomic DNA from the chosen *Setaria* species and *S. italica* were transferred to nylon membranes (Amersham, N^+) according to Southern (1975). Membranes were then hybridized with a heterologous 5S rDNA probe labelled with ^{32}P-dCTP (3000 mCi/ml) using a random primed DNA labelling kit (Boehringer). Autoradiograms were obtained after overnight exposure at –80°C.

Each amplification product was cloned, without further digestion, in pT7 Blue T-Vector (Novagen) following the manufacturer's

Table 1 Origin of the plant material used, its chromosome number and its putative genome attribution. The binomials correspond to those agreed at the present time under the International Nomenclature Code. The accession numbers are those of the collection held at our laboratory. The superscripts indicate that the material has been collected by: [1] Prof. Ph. Jauzein, Institut National Agronomique; [2] M. Rougeron, Limagrain; [3] Prof. I.N. Morrison, University of Manitoba; [4] Prof. J. Pernès, Université Paris-XI

Species	Accession	Origin	Code	Ploidy level	Genome
S. adhaerans	97.009	Spain (Catalonia)[1]	ADH	2n = 2*x* = 18	BB
S. verticillata	97.007	France (Normandy)[1]	VERTI	2n = 4*x* = 36	AABB
S. faberii	97.008	Austria [1]	FAB	2n = 4*x* = 36	AABB
S. viridis var. *viridis*	94.185	France (Charente) [2]	VIR1	2n = 2*x* = 18	AA
S. viridis var. *pycnonoma*	94.186	France (Auvergne) [2]	VIR2	2n = 2*x* = 18	AA
S. viridis var. *viridis*	96.174	Canada (Manitoba) [3]	VIR3	2n = 2*x* = 18	AA
S. viridis var. *viridis*	92.014	Canada (Manitoba) [3]	VIR4	2n = 2*x* = 18	AA
S. viridis var. *viridis*	90.059	China (Heilongjiang) [4]	VIR5	2n = 2*x* = 18	AA
S. italica cv Amende 4	93.004	China (Heilongjiang) [4]	AM	2n = 2*x* = 18	AA
S. italica cv Jigu 11	93.036	China (Hebei) [4]	JIG	2n = 2*x* = 18	AA

instructions and transformed into *Escherichia coli* XL1-blue. Recombinant clones were plated onto selective media and screened by the blue/white method. Minipreparations of positive clones were prepared according to Maniatis et al. (1989) and sequenced using fluorescent sequencing by Genome Express (Grenoble, France) using the M13 forward and reverse universal primers. The resulting sequences were analyzed by the GCG package (Infobiogen). The *Setaria* 5S rDNA sequences were aligned against 5S rDNA from *Triticum aestivum* (TA5SDNAR, Baum and Appels 1992) using CLUSTALW (Thompson et al. 1994) then were further improved by visual examination and editing using GENEDOC (Nicholas and Nicholas 1996). Sequences were subjected to similarity searches in the GenBank and EMBL (European Molecular Biology Laboratory) databases using the National Center of Biotechnology Information (NCBI). Phylogenetic analysis and bootstrapping (100–1,000 replications) were performed using the PHYLIP package (version 3.5, Felsenstein 1995) and fastDNAML program (version 1.0.6; Olsen et al. 1994). Trees were constructed by maximum-likelihood (ML), parsimony and distance methods. ML analyses were done using the fastDNAML program. The transition-transversion ratios of 2 (default) were used to estimate a maximum-likelihood tree with Jumble (random sequence addition order) and global branch swapping options. Parsimony analysis used DNAPARS and the strict consensus tree obtained after CONSENS. For distance-based analysis, a similarity matrix was obtained by DNADIST, using Kimura's two-parameter model of distance estimation (Kimura 1980) to infer the phylogeny of 5S rDNA genes. The similarity matrix was analyzed by the FITCH algorithm and a strict consensus tree obtained using CONSENS. The degree of support for each branch on the strict consensus tree was assessed by bootstrap analysis (100–1,000 replicates). The strict consensus trees were edited by tree view (Page 1996). The GenBank accession number. of all the nucleic acid sequences used are: ADH (AF227004), AM (AF227005), AMS (AF227006), FAB (AF227007), FABL (AF227008), JIG (AF227009), JIGS (AF227010), VERTI (AF227011), VERTIS (AF227012), VIR1 (AF227013), VIR2 (AF227014), VIR3 (AF227015), VIR4 (AF227016), VIR5 (AF227017), VIR5S (AF227018); see abbreviation of plant materials in Table 1.

FISH analysis

Chromosome preparation

Chromosome preparation was done according to Benabdelmouna and Darmency (1997) with minor modifications. Briefly, root-tips of about 2 cm in length were treated with saturated alpha bromonaphthalene for 45 min to accumulate metaphases. All treatments were performed at 27°C in the dark. After rinsing with distilled water, the root-tips were fixed in 3:1 (v/v) 100% ethanol: acetic acid and stored at –20°C until use. The fixed root-tips were rinsed with distilled water and then incubated at 37°C in an enzyme mixture (0.25% pectolyase Y-23, 3% cellulase onozuka R-10, 0.4 M sorbitol, pH 5.5) for 90 min. After incubation the root-tips were transferred to fresh cold 45% acetic acid for 15 min. A single root-tip was transferred in a drop of 45% acetic acid onto a cleaned slide before gentle squashing. The cover slips were removed by immersing slides in liquid nitrogen and the preparations were dehydrated for 10 min each in a graded series of 70%, 95%, and absolute ethanol at room temperature. The air-dried slides were stored at room temperature, up to 1 month, in sealed containers until use.

DNA probes

For the 18S-5.8S-25S rDNA probe we used the 6.1-kb *Eco*RI fragment from sunflower kindly provided by P. Heizmann (Choumane and Heizman 1988). This 6.1-kb *Eco*RI fragment carries an almost complete 18S fragment, the IGS and a small part of the 25S fragment. This probe was labelled by means of nick translation (Boehringer) following the supplier's instructions. When used alone, labelling of the 6.1-kb *Eco*RI fragment was done with digoxigenin-11-dUTP or by direct incorporation of tetramethylrhodamine-6-dUTP, when used in co-localisation experiments.

The 5S rDNA probe was a mix of the cloned PCR product obtained from *S. adhaerans* (B genome) and *S. viridis* (A genome). This probe was labelled by digoxigenin-11-dUTP either by means of nick translation or by PCR amplification using the described 5S-specific primers according to the supplier's instructions.

In situ hybridization

The in situ hybridization reaction was done as described in Benabdelmouna and Darmency (1997). The hybridization mixture was prepared to a final concentration of 10 ng/µl of the digoxigenin-labelled probe, 600 ng/µl of autoclaved herring sperm DNA, 2 × SSC, 10% dextran sulfate, 0.1% SDS, 1 mM EDTA, 1 × Denhardts and 50% de-ionized formamide. Sites of probe hybridization were detected with the immunodetection procedure using fluorescein-conjugated anti-Dig antibody (green fluorescence) and chromosomes were counter-stained with propidium iodide (red fluorescence).

Co-localization of the 5S and 18S-5.8S-25S genes was done by using the same 5S rDNA probe labelled with digoxigenin-11-dUTP, detected with FITC (green fluorescence), and the same 18S-5.8S-25S rDNA probe labelled with tetramethylrhodamine-6-dUTP (red fluorescence). In this case, chromosomes were counterstained with DAPI (4′,6-diamidino-2-phenylindole, blue fluorescence). For visualization, chromosome preparations were analyzed using a Leitz epifluorescent microscope with appropriate filters, and photographs were taken on Fujicolor 400 color slide film.

Results

PCR amplification of 5S rDNA genes and Southern hybridization analysis

PCR amplification of the 5S rDNA produced the banding pattern shown in Fig. 1. The same banding pattern was observed when the PCR amplification products were hybrid-

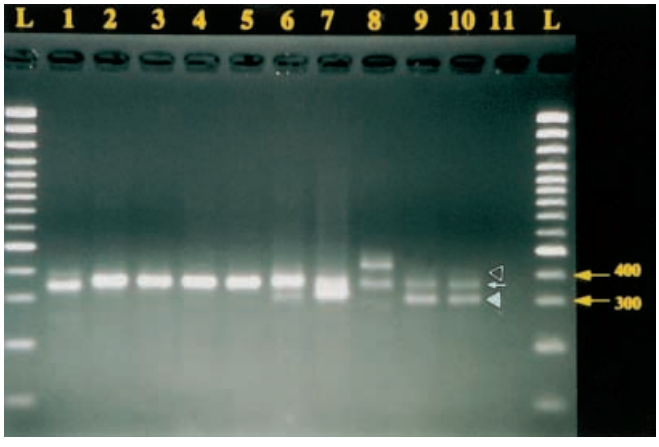


Fig. 1 PCR amplification of 5S rDNA sequences from different *Setaria* species. PCR amplification products were loaded onto 2% gel agarose and stained with ethidium bromide. *L* 100-bp size marker; lane 1 *S. adhaerans*; *lanes 2–6* five different accessions of *S. viridis*; *lane 7 S. verticillata*; *lane 8 S. faberii*; *lanes 9–10* two different accessions of *S. italica*. The *dark arrowhead* indicates the small 5S rDNA unit of about 300 bp produced in *S. verticillata*, the two accessions of *S. italica* and in only one accession of *S. viridis*. The *arrow* indicates the ''normal'' 340-bp unit amplified in all the tested species. The *open arrowhead* indicates the large 5S rDNA unit of about 450 bp observed in the sole species *S. faberii*

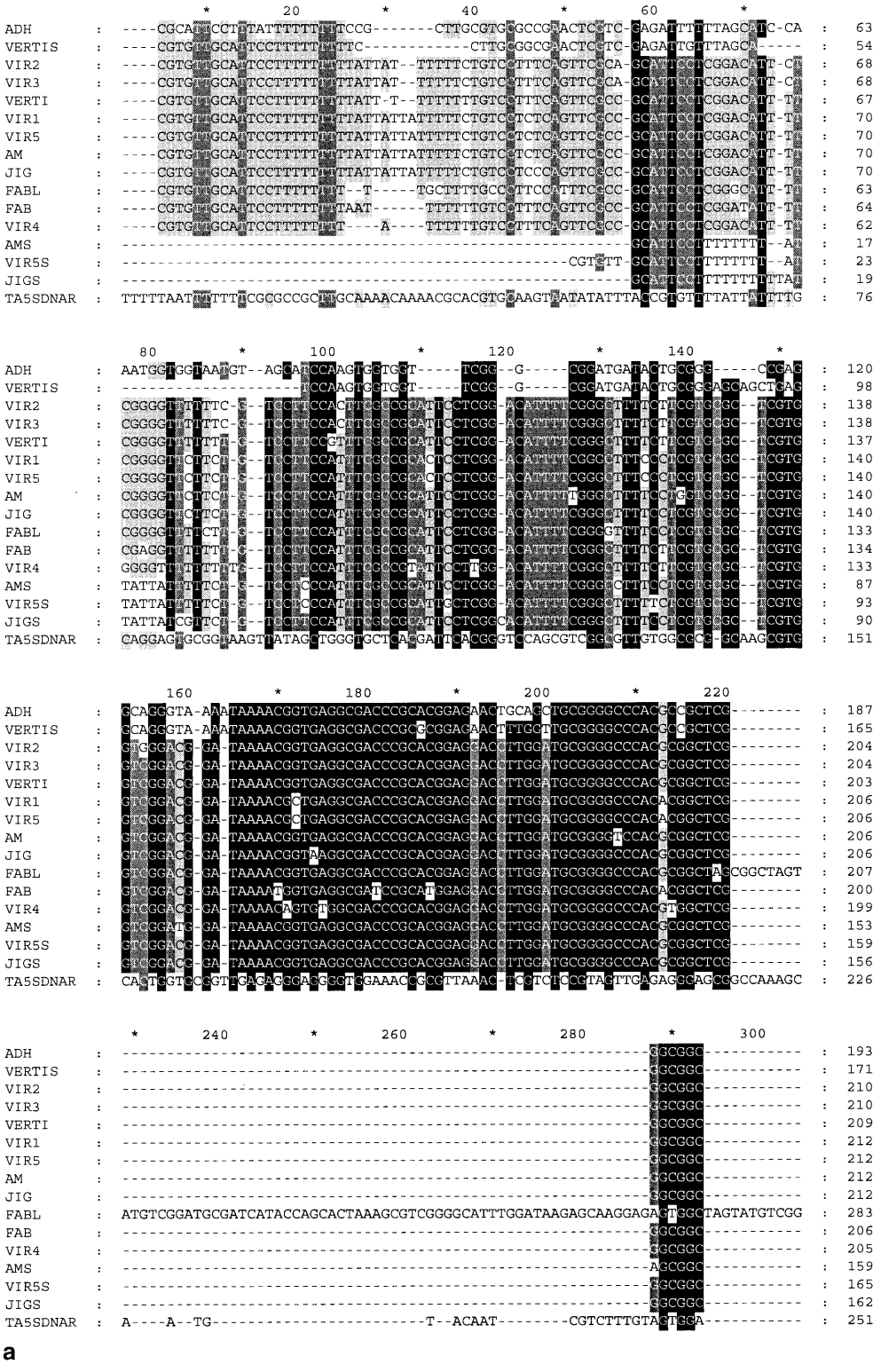
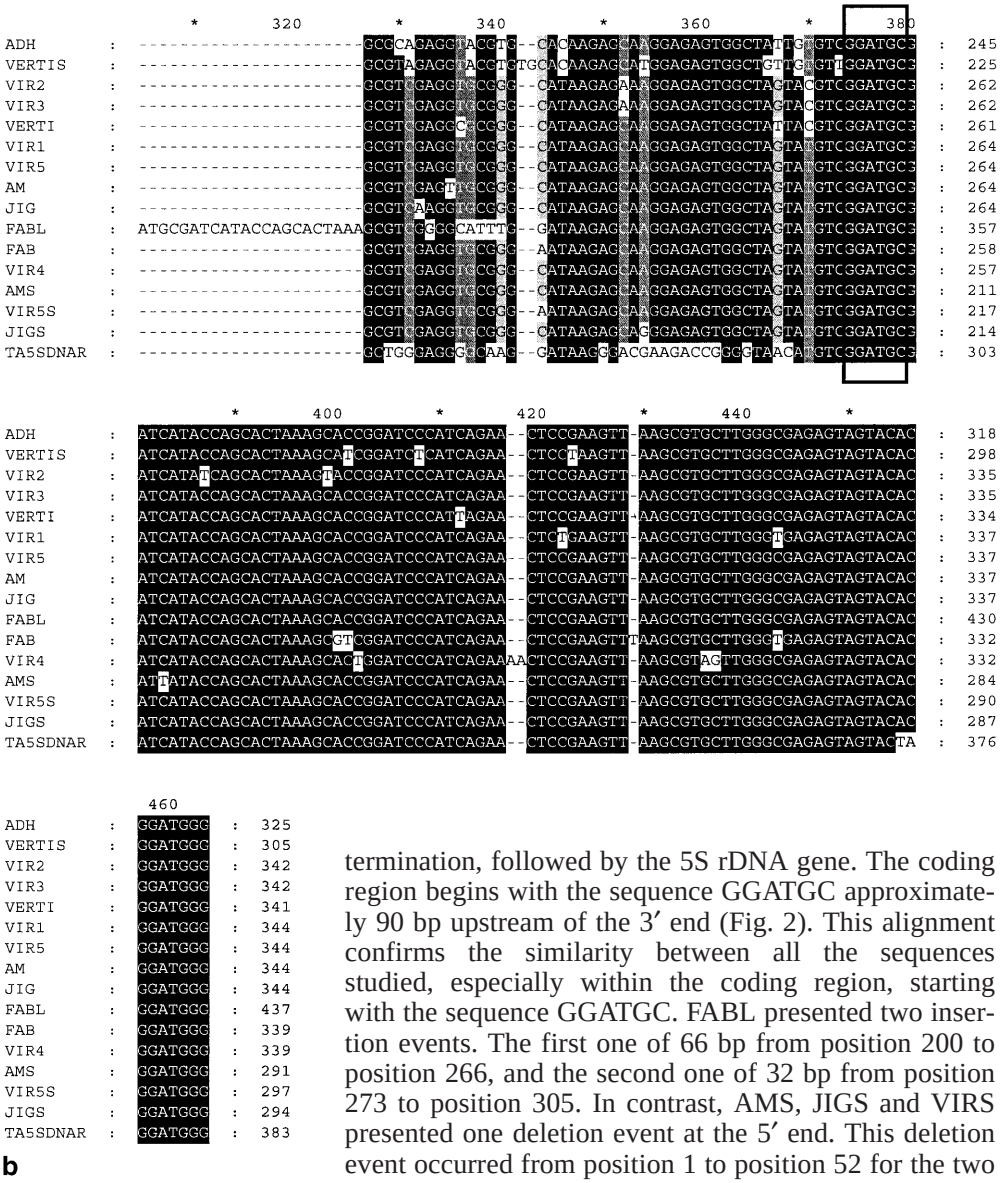


Fig. 2 CLUSTALW multiple alignment of the 5S DNA sequences amplified from the different *Setaria* species against the 5S rDNA sequence from wheat (TA5SDNAR, Baum and Appels 1992). Alignments begin with the spacer sequences followed by the coding sequences. Spacer domains start with the block of T-rich sequences at the 5′ end followed by coding regions that start with the sequence GGATGC indicated by a *retangular frame*. Three size classes were obtained, ''normal'' units of about 340 bp, small units of about 300 bp, and a large unit of 450 bp. Sequence abbreviations correspond to cloned 5S rDNA sequences obtained from: *S. adhaeran*s (ADH), five accessions of *S. viridis* (VIR1, VIR2, VIR3, VIR4, VIR5, for the normal unit, and VIR5S, for the small unit), *S. verticillata* (VERTI, for the normal unit, and VERTIS, for the small unit), *S. faberii* (FAB, for the normal unit, and FABL, for the long unit), and two accessions of *S. italica* (AM and JIG, for the normal unit, AMS and JIGS for the small unit). *Numbers* on the right indicate sequence lengths and (–) denote gaps. Primary shading (*black*) corresponds to 90% sequence similarity, secondary shading (*dark gray*) correspond to 80% sequence similarity, and tertiary shading (*light gray*) corresponds to 50% sequence similarity

ized with the heterologous 5S rDNA probe from *Petunia* (data not shown), which confirmed the 5S rDNA origin of the PCR products. Different sizes of rDNA repeats were observed. A common band of about 340 bp was produced for all species. Additional bands were observed, of about 300 bp in *S. italica*, *S. verticillata*, and in one accession only of *S. viridis*, and of about 450 bp in *S. faberi*.

Southern blots obtained after *Bam*HI digestion of total genomic DNAs and probing with the *Petunia* 5S rDNA probe showed characteristic restriction ladders in all digests (data not shown). Each restriction ladder was characterized with a specific periodicity which can be clearly related to the corresponding 5S amplification patterns shown in Fig. 1. A basic periodicity of approxi-

Fig. 2 (continued)



mately 340 bp was revealed in all the species used. An additional periodicity of 300 bp was revealed in the two lines of *S. italica*, one of the five accessions of *S. viridis*, and *S. verticillata*. In *S. faberii*, an additional periodicity of about 450 bp was observed.

Sequence analysis of 5S rRNA genes and phylogenetic analysis

The 15 PCR-amplified bands were cloned without further *Bam*HI digestion and sequenced. FABL refers to the largest band of about 450 bp, and JIGS, AMS, VIR5S and VERTIS to the smallest bands of about 300 bp. These 15 sequenced clones ranged in size from 291 to 431 bp. The sequence alignments were arranged to start with the block of T-rich sequences at the 5′ end of the non transcribed region implicated in transcription termination, followed by the 5S rDNA gene. The coding region begins with the sequence GGATGC approximately 90 bp upstream of the 3′ end (Fig. 2). This alignment confirms the similarity between all the sequences studied, especially within the coding region, starting with the sequence GGATGC. FABL presented two insertion events. The first one of 66 bp from position 200 to position 266, and the second one of 32 bp from position 273 to position 305. In contrast, AMS, JIGS and VIRS presented one deletion event at the 5′ end. This deletion event occurred from position 1 to position 52 for the two *S. italica* 5S rDNA sequences (AMS and JIGS), while it occurred from position 1 to position 46 for VIR5 S. Two smaller deletions were found in VERTIS from position 23 to position 35, and from position 68 to position 87. The overall sequence similarity between the analyzed sequences ranged from 64%, between FAB and ADH, to 99%, between VIR2 and VIR3 or between VIR1 and VIR5 (Table 2). The strict consensus trees obtained using the distance, parsimony and ML methods were essentially similar in their topology. The 5S rDNA sequences analyzed formed four major branches (Fig. 3). Branch I contained ADH and VERTIS, branch II contained VIR5 S, AMS and JIGS, branch III contained VIR1, VIR5, AM and JIG, and branch IV contained VIR4, FAB, VIR2, VIR3, VERTI and FABL. These four branches were strongly supported with bootstrap values ranging from 89% to 100% for branch I, 87% to 88% for branch II, 80% to 98% for branch III, and 60% to 80% for branch IV. The level of 5S rDNA sequence similarity within each branch ranged from 92% to 99% (Table 2).

Table 2 Sequence percent similarity for the cloned 5S rDNA sequences obtained from the different *Setaria* species. Values of percent similarity were deduced from distances between sequences calculated using the Kimura two-parameter method. Sequence abbreviations correspond to cloned 5S rDNA sequences obtained from: *S. adhaeran*s (ADH), five accessions of *S. viridis* (VIR1, VIR2, VIR3, VIR4, VIR5, for the normal unit, and VIR5S, for the small unit), *S. verticillata* (VERTI, for the normal unit, and VERTIS, for the small unit), *S. faberii* (FAB, for the normal unit, and FABL, for the long unit), and two accessions of *S. italica* (AM and JIG, for the normal unit, AMS and JIGS, for the small unit)

Sequence	ADH	VERTIS	VIR2	VIR3	VERTI	VIR1	VIR5	AM	JIG	FABL	FAB	VIR4	AMS	VIR5 S	JIGS
ADH	100														
VERTIS	92	100													
VIR2	66	73	100												
VIR3	67	74	99	100											
VERTI	67	74	96	97	100										
VIR1	67	73	94	95	95	100									
VIR5	68	74	95	96	96	99	100								
AM	67	73	95	96	96	97	98	100							
JIG	68	74	95	96	96	97	98	98	100						
FABL	67	72	91	92	93	91	92	92	93	100					
FAB	64	71	94	94	95	94	94	94	94	91	100				
VIR4	66	73	94	95	95	93	94	94	94	91	94	100			
AMS	72	75	90	92	91	91	92	92	93	90	90	90	100		
VIR5S	71	74	89	90	90	90	91	91	91	88	90	89	98	100	
JIGS	73	76	91	92	92	92	93	93	94	90	91	91	97	98	100

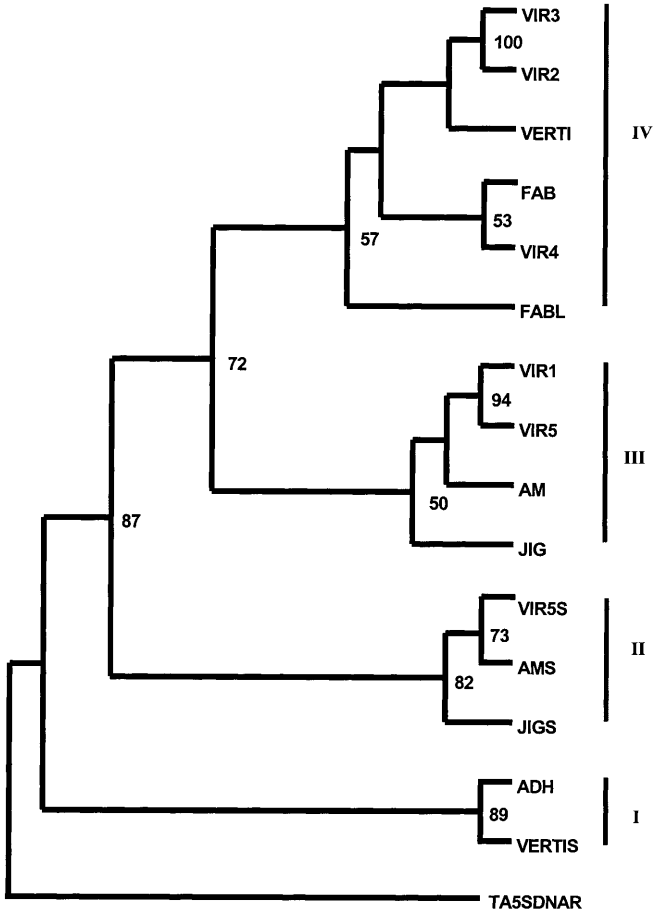


Fig 3 Phylogenetic analysis of 5S rDNA sequences amplified in *Setaria* species. The tree was rooted with TA5SDNAR from wheat. The strict consensus tree presented here was constructed using the maximum-likelihood method (fastDNAML program) with 100 bootstrap replicates. *Numbers* adjacent to branches indicate the percentage of bootstrap replicates supporting that branch. The 15 sequences could be divided into four groups (I to IV). Trees constructed using distance or parsimony methods were essentially similar in their topology to the maximum likelihood-generated tree

Chromosomal localization of the rDNA genes by FISH

Fluorescent in situ hybridization was carried out on all *Setaria* species using as a probe the cloned 5S rDNA PCR product obtained from *S. adhaerans* and *S. viridis*. Two sites of the 5S rDNA genes were revealed in the diploid species at a pericentromeric position on a pair of metacentric to sub-metacentric chromosomes. In the tetraploid species, four sites were revealed at the same pericentromeric position (Fig. 4 f–i). When the heterologous 18S-5.8S-25S rDNA probe from sunflower was used, FISH also showed two hybridization sites in the diploid species, *S. adhaerans*, *S. viridis* and *S. italica*. These two hybridization sites were located at the end of one arm of a pair of metacentric to sub-metacentric chromosomes normally associated with a secondary constriction (Fig. 4 a–c). The same probe revealed four hybridization sites in the tetraploid species *S. faberii* and *S. verticillata*. These sites were located at the end of two metacentric to sub-metacentric chromosome pairs (Fig. 4 d–e). Simultaneous co-localization of the 5S and 18S-5.8S-25S rDNA sites showed that these two gene clusters were carried by two distinct chromosome pairs (Fig. 4 j).

Discussion

Polymorphism of rDNA has often been used to investigate phylogenetic relationships among species and to organize groups of strains within a species. Recently, RFLP analysis of 18S-5.8S-25S rDNA provided interesting data on the geographical distribution and relationships among foxtail millet accessions (Fukunaga et al. 1997). Our study focuses on 5S rDNA, using PCR amplification, sequence analysis, and chromosomal localization in five *Setaria* species of the gene pool of foxtail millet. The physical mapping of the 18S-5.8S-25S and 5S rDNA genes showed that they are organized at

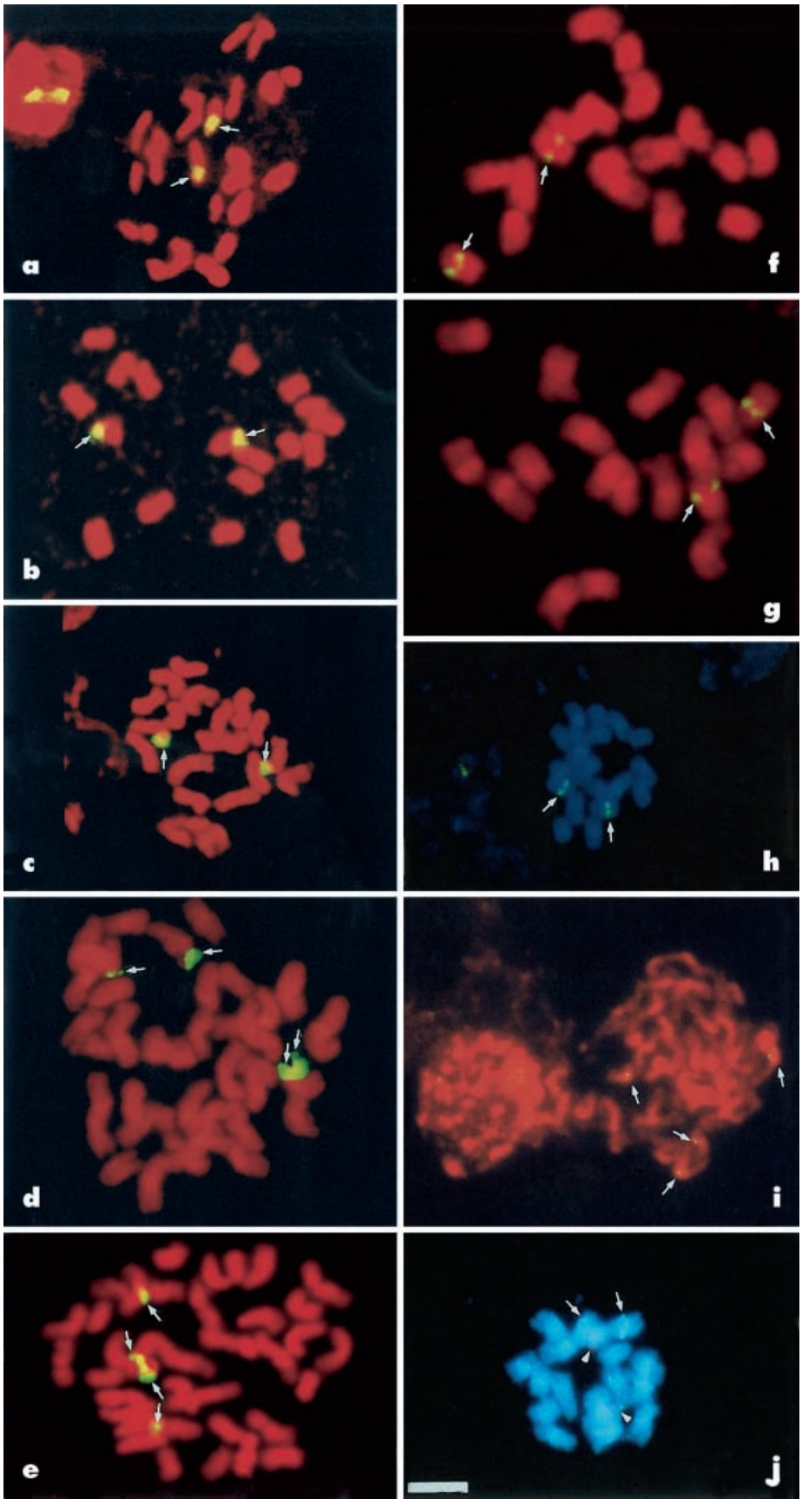


Fig. 4a–j Fluorescence in situ hybridization with the 18S-5.8S-25S rDNA probe (**a–e**), the 5S rDNA probe (**f–i**), or a combination of the two rDNA probes (**j**) to root-tip metaphases from different *Setaria* species. Chromosomes were counter-stained either with propidium iodide (red fluorescence, **a–g, i**) or with DAPI (blue fluorescence, **h** and **j**). The heterologous 18S-5.8S-25S rDNA probe from sunflower revealed two large yellow sites in the diploid species *S. adhaerans* (**a**), *S. viridis* (**b**) and *S. italica* (**c**), while four large hybridization sites were revealed in the tetraploid species *S. faberii* (**d**) and *S. verticillata* (**e**). All these 18S-5.8S-25S rDNA sites were located at the end of metacentric to sub-metacentric chromosome pairs (*arrows*). The homologous 5S rDNA probe from the cloned PCR products obtained from *S. adhaerans* (genome B) and *S. viridis* (genome A) revealed also two hybridization sites in the diploid species *S. adhaerans* (**f**), *S. viridis* (**g**) and *S. italica* (**h**), while it revealed four hybridization sites in the tetraploid species *S. faberii* and *S. verticillata* (**i**). These sites were located at a pericentromeric position (*arrows*). Co-localization of the two rDNA probes (18S-5.8S-25S rDNA labelled with rhodamine and 5S rDNA labelled with digoxigenin and revealed with fluorescein) showed that these two gene clusters were carried by two distinct chromosome pairs (**j**). *Arrows* indicate the 5S rDNA sites and *arrowheads* indicate the 18S-5.8S-25S rDNA sites. Photographs **a–g** were taken with confocal microscope while photographs **h–j** were taken with an epifluorescence microscope. Bar equals 8 µm

two separate loci. No polymorphism of chromosome location among species and genomes was observed. Two chromosome pairs can now be unambiguously identified using FISH: one pair carries the 5S rDNA cluster, while another chromosome pair has the 18S-5.8S-25S rDNA cluster. This will help to clarify the genetic map of foxtail millet that has been recently constructed (Wang et al. 1998) and to assign linkage groups to physically marked chromosomes.

The PCR-amplified products produced in this study were clearly shown to have a 5S rDNA origin. This was indicated by the PCR amplifications combined with the hybridization results with the heterologous 5S rDNA probe from *P. hybrida* and also by sequencing. This amplification pattern was confirmed by Southern analysis which produced different restriction ladders reproducing the specific periodicity revealed in the PCR amplification. One obvious cause for the ladder differences is variation in the restriction sites. This is a likely possibility for sites that are in the spacer, but unlikely for a site that is within the coding region. Variation in the restriction site might be due to C-methylation, especially when using *Bam*HI which is known to be sensitive to C-methylation. As found in our previous work on *Petunia* (Benabdelmouna and Darmency 1997) and in other work on *Triticum* (Dvorak et al. 1989), maize and different dicots (Mascia et al. 1981; Rafalski et al. 1982; Goldsbrough et al. 1983) we conclude that the ladders produced on the different *Setaria* species were caused by C-methylation of *Bam*HI restriction sites.

All the species studied here share a common band of about 340 bp. *S. adhaerans* and four of the five accessions of *S. viridis* have this unique band. An additional band of an approximately 300-bp repeat unit was found for *S. italica*, one of the five accessions of *S. viridis*, and *S. verticillata*. An additional band of 450 bp was also found in the sole species *S. faberii*. Comparison of these amplified repeat units and the published *T. aestivum* 5S rDNA sequences (TA5SDNAr, Baum and Appels 1992) revealed a high similarity between them, especially in the coding region which begins with the sequence GGATGC approximately 90 bp upstream from the 3′ end. Assuming that the size of the coding sequences of the 5S rRNA genes is 120 bp, the length of the non transcribed spacer sequence between the 5S rRNA genes in the genus *Setaria* would vary in length from 180 to 330 bp.

The overall sequence similarity between the analyzed sequences varied from 64% to 99% which indicates a high degree of interspecific nucleotide diversity (up to 36%). This result seems to be supported by the work of Kellog and Appels (1995) who also found a high degree of interspecific variation in the spacer, and even in the gene sequence, after comparison of the 5S rDNA gene sequences across 34 Triticeae species. Phylogenetic trees based on the 5S rDNA sequences amplified in *Setaria* species were rooted with TA5SDNAr from *T. aestivum.* Whatever the method, the clustering allowed us to separate these 5S rDNA sequences into four groups. The first one included 5S rDNA units amplified in the species that are thought to have genome B: *S. adhaerans* and *S. verticillata* (VERTIS). Although VERTIS shows two deletions compared to the common 340-bp band, it is clearly more similar to ADH, indicating that these deletions probably occurred after the origin of the allotetraploid *S. verticillata*. In contrast, the second band of *S. verticillata*, VERTI of 340 bp, is more similar to bands of *S. viridis* (group IV) and is likely to belong to the A genome. The other allotetraploid, *S. faberii*, that has also one set of chromosomes of the B genome, has two units included in group IV, of which one, FABL, is well-separated from the others and could belong to the B ancestor. Indeed, it is likely that the two insertions observed in FABL, adding up to 100 bp, could greatly modify its similarity with the original unit belonging to the B genome. The second group included the small unit (300 bp) amplified in the two Chinese cultivars of *S. italica* and in the Chinese accession of *S. viridis* (VIR5 S) which are known to share the same A genome. This small rDNA unit is not found in the other accessions *of S. viridis.* This similarity among Chinese material is in agreement with the theory of a domestication center for the foxtail millet in China from local Chinese *S. viridis* (Kawase and Sakamoto 1984; Jusuf and Pernès 1985; Li et al. 1998; Wang et al. 1998). More accessions of the wild species from various regions should be analyzed in order to definitely correlate the presence of the small unit to its sole Chinese origin. If so, it could provide a useful marker to distinguish historical seed-crop exchanges between different regions versus the independent regional domestication processes of foxtail millet. The third group included the ''normal'' 5S rDNA unit (340 bp) amplified in the cultivated and wild Chinese species in addition to one French accession of *S. viridis* (VIR1). The ''normal'' units of the three other accessions of *S. viridis* are located in the fourth group, together with those of the tetraploid species. This unit in the allotetraploids certainly belongs to the A genome because it is grouped with that of *S. viridis*. Separating this ''normal'' unit into two clusters denotes the huge variability of the 5S rDNA gene spacer and confirms its value for investigating phylogeny and gene flow within the gene pool of a cultivated species. Thus, the clustering supports the view of the closest relationship of the *S. italica* and *S. viridis* accessions from the same region, and it provides a new tool to investigate previous hypotheses on the domestication centers of *S. italica*. It also confirms the wide difference between the A and B genomes, and even the taxonomic position of *S. verticillata* that has been recently shown to be allotetraploid (Benabdelmouna et al. 2001), because one rDNA unit is close to that of species having the A genome while the other is similar to that of species having the B genome.

Acknowledgements This research was supported in part by the Burgundy Region Council (France) and the INCO-DC European contract IC18-CT98–0391. The authors thank Claude Humbert from the CMAB, University of Burgundy, Dijon, for his expertise in confocal microscopy observations, as well as Jacky Delbut and Alain Fleury for their technical help.

References

Appels R, Gerlach ES, Dennis H, Swift H, Peacock WI (1980) Molecular and chromosomal organization of DNA sequences coding for the ribosomal RNAs in cereals. Chromosoma 78: 293–311

Appels R, Honeycutt (1986) rDNA: evolution over a billion years. In: Dutta SK (ed) DNA systematics. II. Plant DNA. CRC Press, Boca Raton, Florida, USA, pp 81–125

Appels R, Baum BR (1991) Evolution of the NOR and 5S DNA loci in the *Triticeae*. In: Doyle JJ, Soltis PS (eds) Plant molecular systematics. Chapman and Hall, New York, pp 92–116

Baum BR, Appels R (1992) Evolutionary change at the 5S DNA loci of species in the *Triticeae*. Plant Syst Evol 183: 195–208

Baum RB, Bailey LG (1997) The molecular diversity of the 5S rRNA gene in *Kengyilia altavica* (Drobov) JL Yang, Yen & Baum (Poaceae: Triticeae): potential genomic assignment of different rDNA units. Genome 40: 215–228

Baum BR, Johnson DA (1994) The molecular diversity of the 5S rRNA gene in barley (*Hordeum vulgare*). Genome 37: 992–998

Baum BR, Johnson DA (1998) The 5S rRNA gene in sea barley (*Hordeum marinum* Hudson sensu lato): sequence variation among repeat units and relationship to the X haplome in barley (*Hordeum*). Genome 41: 652–661

Benabdelmouna A, Darmency AM (1997) Distribution and chromosomal organization of 18S-5.8S-25S and 5S rDNA in *Petunia* species. Agronomie (Plant Genet Breed) 17: 349–360

Benabdelmouna A, Shi Y, Abirached-Darmency M, Darmency H (2001) Genomic in situ hybridization (GISH) discriminates between the A and the B genomes in diploid and tetraploid *Setaria* species. Genome 44: (in press)

Castilho A, Heslop-Harrison JS (1995) Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. Genome 38: 91–96

Choumane W, Heizman P (1988) Structure and variability of nuclear ribosomal genes in the genus *Helianthus*. Theor Appl Genet 76: 481–489

Darmency H, Pernés J (1987) An inheritance study of domestication in foxtail millet using an interspecific cross. Plant Breed 99: 30–33

Darmency H, Zangré GR, Pernès J (1987) The wild weed crop complex in *Setaria italica:* a hybridization study. Genetica 75: 103–107

D'Hont A, Ison D, Alix K, Roux C, Glaszmann JG (1998) Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. Genome 41: 221–225

Dvorak J, Zhang HB, Kota RS, Lassner M (1989) Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. Genome 32: 1003–1016

Felsenstein J (1995) PHYLIP (Phylogeny Inference Package) version 3.57c. Department of Genetics, University of Washington, Seattle, USA

Flavell RB (1986) The structure and control of expression of ribosomal RNA genes. Oxford Surveys Plant Mol Cell Biol 3: 251–274

Fukunaga K, Domon E, Kawase M (1997) Ribosomal DNA variation in foxtail millet, *Setaria italica* (L.) Beauv., and a survey of variation from Europe and Asia. Theor Appl Genet 95: 751–756

Goldsbrough PB, Ellis THN, Lomonossof GP (1983) Sequence variation and methylation of the 5S RNA genes. Nucleic Acids Res 10: 4501–4514

Harlan JR, De Wet JMJ (1971) Towards a rational taxonomy of cultivated plants. Taxon 20: 509–517

Jusuf M, Pernès J (1985) Genetic variability of foxtail millet (*Setaria italica* P. Beauv.). Electrophoretic study of five isoenzyme systems. Theor Appl Genet 71: 385–391

Kawase M, SakamotoS (1982) Geographical distribution and genetic analysis of phenol color reaction in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet 63: 117–119

Kawase M, SakamotoS (1984) Variation, geographical distribution and genetic analysis of esterase isozymes in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet 67: 529–533

Kawase M, SakamotoS (1987) Geographical distribution of landrace groups classified by intraspecific hybrid pollen sterility in foxtail millet, *Setaria italica* (L.) P. Beauv. Jpn J Breed 37: 1–9

Kellogg E, Appels R (1995) Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics 140: 325–343

Kerguélen M (1999) Index synomymique de la flore de France. (http://www.inra.fr/Internet/Centres/Dijon/malherbo/fdf/accueil1.htm)

Kimura M (1980) A simplified method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120

Kolipara KP, Singh RJ, Hymowitz T (1997) Phylogenetic and genomic relationships in the genus *Glycine* Willd. based on sequences from the ITS region of nuclear rDNA. Genome 40: 57–67

Kolshinsky A, Kolesnikova M, Ananiev E (1991) ''Portraying'' of plant genomes using polymerase chain reaction amplification of ribosomal 5S genes. Genome 34: 1028–1031

Li CH, Pao WK, Li HW (1942) Interspecific crosses in *Setaria*. II. Cytological studies of interspecific hybrids involving 1, *S. faberii* and *S. italica*, and 2, a three way cross, F_2 of *S. italica* × *S. viridis* and *S. faberii*. J Hered 33: 351–355

Li Y, Wu SZ, Cao YS (1995) Cluster analysis of an international collection of foxtail millet [*Setaria italica* (L.) P. Beauv.]. Euphytica 83: 79–85

Li Y, Jia J, Wang Y, WuS (1998) Intraspecific and interspecific variation in *Setaria* revealed by RAPD analysis. Genet Res Crop Evol 45: 279–285

Maluszynska I, Heslop-Harrison JS (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. Plant J 1: 159–161

Maluszynska I, Heslop-Harrison JS (1993) Physical mapping of rDNA loci in *Brassica* species. Genome 36: 774–781

Maniatis T, Fritsch EF, Sambrook J (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Mascia PN, Rubenstein I, Phillips RL, Wang AS, Xiang LZ (1981) Localization of the 5S rDNA genes and evidence for diversity in the 5S rDNA region of maize. Gene 15: 7–20

Mukai Y, Endo TR, Gill BS (1991) Physical mapping of the 18 S.26S rRNA multigene family in common wheat: identification of a new locus. Chromosoma 100: 71–78

Nicholas KB, Nicholas JB Jr (1996) GeneDoc©: a tool for editing and annotating multiple sequence alignments. V. 1.1.004. www.cri.com/~ketchung

Olsen GJ, Matsuda H, Hagstrom R, Overbeek R (1994) FastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput Appl Biosci 10: 41–48

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358

Rafalski JA, Wiewiorowski M, Soll D (1982) Characterization and nucleotide sequence of nuclear 5S RNA genes in yellow lupin (*Lupinus luteus*). Nucleic Acids Res 10: 7635–7642

Rogers SO, Bendich AJ (1987) Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. Plant Mol Biol 9: 509–520

Rominger JM (1962) Taxonomy of *Setaria* (Gramineae) in north America. Illinois Biological Monographs No. 29. University of Illinois Press, Urbana, Illinois, USA

Saghai-Maroof MA, Soliman KM, Jorgensen RN, Allard RW (1984) Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81: 8014–8018

Schmidt T, Swharzacher T, Heslop-Harrison JS (1994) Physical mapping of rRNA genes by fluorescent in situ hybridization and structural analysis of 5S rRNA and intergenic spacer sequences in sugar beet (*Beta vulgaris*). Theor Appl Genet 88: 629–636

Schontz D, Rether B (1999) Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv.: identification and classification of lines with RAPD markers. Plant Breed 118: 190–192

Scoles GJ, Gill BS, Xin ZY, Clarke B, McIntyre CL, Chapman C, Appels R (1988) Frequent duplication and deletion events in the 5D RNA genes and the associated spacer region of the *Triticeae*. Plant Syst Evol 160: 105–122

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–509

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

Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD (1998) Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet 96: 31–36

Wanous MK (1990) Origin, taxonomy and ploidy level of foxtail millet and minor cereals. Plant Var Seeds 3: 99–112

Zangré R, Nguyen-Van E, Rherissi B, Till-Bottraud I (1992) Organisation du pool génetique de *Setaria italica* (L.) P. Beauv. et exploitation des ressources génétiques d'espèces spontanées. In: Lavoisier (ed) Complexes d'espèces, flux de gènes et ressorces génétiques des plantes. BRG, Paris, pp 87–97

Zoldos V, Papes D, Cerbah M, Panaud O, Besendorfer V, Siljak-Yakovlev S (1999) Molecular-cytogenetic studies of ribosomal genes and heterochromatin reveal conserved genome organization among 11 *Quercus* species. Theor Appl Genet 99: 969–977