Theor Appl Genet (1994) 89: 26–32

Y. Sun · D. Z. Skinner · G. H. Liang · S. H. Hulbert

Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA

Received: 24 November 1993 / Accepted: 2 February 1994

Abstract The phylogenetic relationships of the genus *Sorghum* and related genera were studied by sequencing the nuclear ribosomal DNA (rDNA) internal transcribed spacer region (ITS). DNA was extracted from 15 *Sorghum* accessions, including one accession from each of the sections *Chaetosorghum* and *Heterosorghum*, four accessions from *Parasorghum*, two accessions from *Stiposorghum*, and seven representatives from three species of the section *Sorghum* (one accession from each of *S. propinquum* and *S. halepense*, and five races of *S. bicolor*). The maize (*Zea mays*) line, H95, and an accession from *Cleistachne sorghoides* were also included in the study. Variable nucleotides were used to construct a strict consensus phylogenetic tree. The analyses indicate that *S. propinquum*, *S. halepense* and *S. bicolor* subsp. *arundinaceum* race aethiopicum may be the closest wild relatives of cultivated sorghum; *Sorghum nitidum* may be the closest 2n=10 relative to *S. bicolor*; the sections *Chaetosorghum* and *Heterosorghum* appear closely related to each other and more closely related to the section *Sorghum* than *Parasorghum*; and the section *Parasorghum* is not monophyletic. The results also indicate that the genus *Sorghum* is a very ancient and diverse group.

This research was partially supported by a Third Country Scholarship from Pioneer Hi-Bred International Incorporated Contribution 94-182-J from Kansas Agricultural Experiment Station

Communicated by Y. Gleba

Y. Sun · D. Z. Skinner · G. H. Liang · S. H. Hulbert (✉)
Genetics Program, Kansas State University, Manhattan, KS 66506, USA

D. Z. Skinner
USDA-ARS, Kansas State University, Manhattan, KS 66506, USA

G. H. Liang
Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA

S. H. Hulbert
Department of Plant Pathology, Kansas State University,

Key words *Sorghum* · *Zea mays* · Phylogeny
rDNA sequence

Introduction

The current classification of the genus *Sorghum* includes five sections, *Sorghum* (or *Eusorghum*), *Parasorghum*, *Chaetosorghum*, *Heterosorghum*, and *Stiposorghum* (Garber 1950). The section *Sorghum* includes cultivated sorghum (*S. bicolor*) and its close relatives ranging through Africa and Asia. *Parasorghum* includes about ten species found in the eastern hemisphere and central America. *Chaetosorghum* and *Heterosorghum* are monotypic and are found in the Australo-Pacific region. *Stiposorghum* consists of five species found in northern Australia. Genomic chromosome numbers of the *Sorghum* species vary from n=5 to n=20

The genus *Cleistachne* (2n=36), distributed in the tropical eastern African and western Indian regions, has been considered a sister genus to *Sorghum*, both belonging to the supergenus *Sorghastrae* (Garber 1950). Maize (*Zea mays*, 2n=20) is an important cereal crop which has the same chromosome number as *S. bicolor*. Most maize RFLP probes cross hybridize to sorghum and genetic maps of the two species share large regions of colinearity (Hulbert et al. 1990; Whitkus et al. 1992; Melake Berhan et al. 1993).

Recently-developed molecular genetic techniques have provided a powerful means to reveal phylogenetic relationships among species. By studying restriction-site variation in the chloroplast of *Sorghum* species, Duvall and Doebley (1990) raised some doubt on the monophyly of the section *Parasorghum* and suggested that the genus *Sorghum* may be either paraphyletic or polyphyletic. Their data also suggested that *Cleistachne sorghoides* is closely related to some *Sorghum* species.

Ribosomal DNA (rDNA) is well-suited for a broad range of phylogenetic studies (Hamby and Zimmer 1992) because the different components of rDNA vary in their degree of conservation. The three ribosomal gene subunits

are very conservative throughout organisms, both in their length and their nucleotide sequences, and are useful in phylogenetic analyses at broad levels. The internal transcribed spacer regions (ITS) are more divergent in their nucleotide sequences. The intergenic spacer (IGS) is most variable, and polymorphisms in nucleotide sequence as well as in length, can often be found within species (Appels and Dvorak 1982; Saghai-Maroof et al. 1984). Hamby and Zimmer (1992) used sequences of the 17*s* gene in phylogenetic studies on the 60 taxa of seed plants and Baldwin (1992) used sequences of ITS regions to study evolution in the Compositae. Springer et al. (1989) reported some within-species variation in the IGS regions of sorghum, but found the lengths and restriction maps of the rDNA arrays of sorghum, maize, and sugarcane to be very similar. Our objectives were to use sequences of internal transcribed spacer regions of rDNA to study phylogenetic relationships among *Sorghum* species along with the related genera *Cleistachne* and *Zea*.

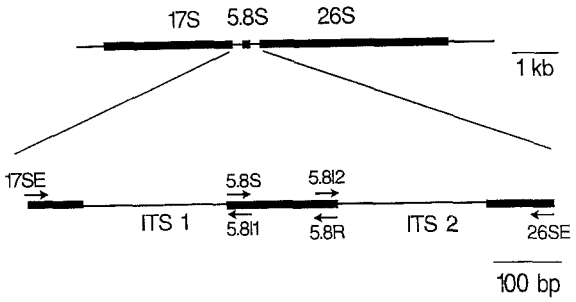

Fig. 1 The composition of the rDNA gene in *Sorghum* and its relatives. *Thick lines* represent the coding sequences for the ribosomal subunits; *thin lines* represent IGS or ITS spacer regions. *Arrows* represent the position of primer sequences used for DNA amplification and sequencing (17SE and 26SE), or only DNA sequencing (5.8S, 5.8R, 5.8I1 and 5.8I2). Primer sequences are as follow: 17SE ACGAATTCATGGTCCGGTGAAGTGTTCG; 26SE TAGAATTCCCCGGTTCGCTCGCCGTTAC; 5.8S ACGACTCTCGGCAAC; 5.8R GCGTGACGCCCAGGC; 5.8I1 GTTGCCGAGAGTCGT; 5.8I2 GCCTGGGCGTCACGC

Materials and methods

The plant materials used in this experiment are presented in Table 1. Specimens of the samples used are maintained at the Department of Agronomy, Kansas State University. Young leaves were collected from greenhouse-grown plants and frozen at –70°C until DNA was extracted. Total DNA was purified following procedures described previously (Hulbert and Bennetzen 1991).

Based on the reported ribosomal DNA sequence of rice (Takaiwa et al. 1985), carrot, and broad bean (Yokota et al. 1989), 20-bp conserved regions in 17*s* and 26*s* ribosomal genes were chosen as primers (Fig. 1). An extra 8 bp including an *Eco*R I site was added to the 5′ end of each primer. The DNA fragment amplified using these two primers is approximately 800 bp long and included ITS1, ITS2 and the 5.8*s* ribosomal gene.

Table 1 Plant materials used

Acession	PI#[a]	Source	Origin
Section *Sorghum:*			
S. bicolor (2n=20)			
subsp. *bicolor* cv. Shanqui red;		J. Axtell	China
subsp. *arundinaceum*			Africa
arundinaceum	302266		
verticilliflorum	365024		Africa
aethiopicum;	302105		Africa
subsp. *drummondii* sudanese;	343165		Africa
S. halepense (2n=40) halepense	302236		
S. propinguum (2n=20);	302191		
Section *Chaetosorghum:*			
S. macrospermun (2n=40)	85-229	K. Schertz	Australia
Section *Heterosorghum:*			
S. laxiflorum (2n=40);		P. Lawrence	Australia
Section *Parasorghum:*			
S. australiense (2n=20)	85-261	K. Schertz	Australia
S. nitidum (2n=10)		T. Wu	China
S. purpureo-sericeum (2n=10)		T. Wu	Africa
S. versicolor (2n=10);		T. Wu	Africa
Section *Stiposorghum:*			
S. matarankense (2n=10)	85-242	K. Schertz	Australia
S. stipoideum (2n=10);		P. Lawrence	Australia
Zea mays (2n=20) line H95.			
Cleistachne sorghoides (2n=36)	368246	MPIS[b],	Uganda[a]

[a] Plant introduction number
[b] Mid-West area plant introduction number

The amplification conditions were: 50 mM KCl, 10 mM Tris-HCl(pH 8.8), 0.1% Triton X-100, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$, 8% DMSO, 0.1 μg/0.1 ml of genomic DNA and 0.05 units of *Taq* DNA polymerase. After an initial heating step at 94°C for 3 min, samples were incubated for 35 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 2 min. The reactions were completed by incubating at 72°C for 5 min.

The amplified DNA was digested with *Eco*RI, and ligated into the *Eco*RI site of Bluescript PBSII SK- and used to transform *E. coli* strain DH5αF′ using standard techniques (Sambrook et al. 1989). *Eco*RI-digested miniprep DNAs were electrophoresed on agarose gels to confirm that the cloned inserts were in the range of about 800 bp.

Double-strand sequencing was conducted with the Sequenase 2.0 version (US Biochemicals) using either the dGTP or the 7-deaza-dGTP kit. Both gel-purified PCR products and purified clones were used as templates. Difficulties with obtaining sequence information from PCR-amplified templates from some accessions made it necessary to derive the complete sequence from cloned templates. In this case, at least two different clones were used. Besides the two PCR primers, four 15-bp primers corresponding to each strand of two regions within the 5.8*s* gene were used to prime sequencing reactions (Fig. 1).

DNA sequences from different accessions were aligned with the aid of the PC GENE program, and data was analyzed by Wagner parsimony (Swofford 1993) using PAUP 3.1 on a Macintosh Quadra 700. Deletions were considered to be single evolutionary events and were weighted the same as nucleotide substitutions. Two exact methods were used as search strategies to identify the most parsimonious tree: Branch and bound searching with TBR branch swapping and the MULPARS options on, and an exhaustive search of a data set with a reduced number of taxa. The latter was performed after deleting the data for *S. halepense*, *S. propinquum* and all of the *S. bicolors* except the cultivated line. These accessions varied by no more than one informative base. Maize was specified as an out group and the "midpoint rooting" option was selected in constructing the trees. Two equally-parsimonious trees were identified by both methods. A strict consensus tree was constructed in which any clade appearing in both of the shortest trees was included. The ITS DNA sequence of rice (*Oryza sativa*, Takaiwa et al. 1985) was compared to the other species to determine the positions of the 17*s*, 5.8*s* and 26*s* genes and for use as a possible outgroup. The inclusion of rice in the anal-

ysis increased the number of deletion/insertion events from four to 23 and also added 48 base substitutions. This suggested that the genus *Oryza* had diverged long before the progenitors of genera *Sorghum* and *Zea* diverged from each other and that its distant relationship may make it unsuitable as an outgroup. When a phylogenetic analysis was performed using the rice sequence as an outgroup, however, an identical consensus tree was identified.

The rDNA sequences of one representative of the section *Sorghum* (the cultivated line) and each of the other species were submitted to GenBank. The sequences were given the accession numbers U04788–U04798.

Results

Length variations of the ITS region

The length of the entire *S. bicolor* rDNA ITS region was 588 bp. Other species in the study had similar lengths of ITS regions. The samples in the study had GC contents varying between 60.4% to 65.8%. The use of 7-deaza dGTP

Table 2 Polymorphic base pairs used in the analysis. Numbers at the top of the columns refer to the number of the base in the sequenced fragment starting from the end of 17*s* gene. Ambiguous base pairs due to difficulties in sequence reading or differences between clones are shown by 'N', which were not counted in any case. Dashes represent gaps in the sequences due to deletion/insertion events

Taxa	1111111111111111111111111111111111111111111222222222222222222222222223 0011111112222333333445555566666667778902333334444555566777788999990 3801236793589123589041234612357891287187014791378018286789020145 60
Sorghum bicolor cv. sanqui red	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAGATGAGCCTGAGAGATCTT
S. bicolor race aethiopicum	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAGATGAGCCTGAGAGATCTT
S. propinquum	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAGATGAGCCTGAGAGATCTT
S. halepense	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAGATGAGCCTGAGAGATCTT
S. bicolor race veticiliflorum	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAGATGAGCCTGAGAGATCTT
S. bicolor race sudanense	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAGATGAGCCTGAGAGATCTT
S. bicolor race arundinaceum	CCGTGAATTTACCGTCATTTTCAAGTTTGTCAACTAAATCGTCT-AAAATGAGCCTGAGAGATCTC
S. nitidum	CC-TGAATTTATCGTCATTTTCAAGTTTGTCAACTAAATCGTGT-AAGATGAGCCTGAGAGATCTT
S. macrospermum	CA-CGGGCTCGTCGCTGTCCCCAAGTGAATCGTATGGATCACCT-TGGATAAGGACNNGAGATCTT
S. laxiflorum	TC-CGGGCTCGTCGCTGTCCCCAAGTGACTCGTATGGATCACCT-TGGATAANGCCGCGTGATCTT
S. versicolor	CC-CGAGCTTATCGCTGCTCTCACGCGAGCTGTACGAAATATCT-TGGACGCGCCCGCAAGATCTT
S. purpureo-sericeum	CC-CGAGCTTACCGCCGCTCTCATGCGAGCTGTATGAAATATAT-AGGACGTGCCCGCGAGATCTT
Cleistachne sorghoides	CC-CGAGCTTGCCGCCGCTCNNNNGNGAGCNGTNTGANATATCT-NGGACGNGCCCGCGAGATCTT
S. stipoideum	CC-CGAGCTTGCCGCCGCTCTCACGCGAGCTGTATGACATATCC-AGGACGCGCCTGAGAGATCTT
S. australiense	CC-CGAGCTTGCCGCCGCTCTCACGCGAGCTGTATGACATATCC-AGGACGTGCCCGCGAGATCTT
S. matarankense	CC-CGAGCTCGCNNCCGCTCTTGCGCGAGCCGTACGGCGTATCC-GGGGCGCGCCCGCGATCATAT
Zea mays cv. h95	CC-CNAGCATACCGCCGCTCTNGCNCGAGCTGTATGAAATATATGAGGACGTGCCAGGGAGATCTT

Taxa	4444444444444444444455555555555555555556666666666666666666666666777 1233334444688899999900111123334455677990003444444555567778888999902 1714781245957915678901167990891259936250298234678012931352679235 35
Sorghum bicolor cv. sanqui red	CGTCCCGGCACCACTCGG--TTTTTCCGATCGCCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. bicolor race aethiopicum	CGTCCCGGCACCACTCGG--TTTTTCCGATCGCCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. propinquum	CGTCCCGGCACCACTCGG--TTTTTCCGATCGCCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. halepense	CGTCCCGGCACCACTCGG--TTTTTCCGATCGCCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. bicolor race verticiliflorum	CGTCCCGGCACCACTCGG--TTTTTCCGATCGTCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. bicolor race sudanense	CGTCCCGGCACCACTCGG--TTTTTCCGATCGTCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. bicolor race arundinaceum	CGTCCCGGCACCACTCGG--TTTTTCCGATCGCCTCGGGTAGTGCCGTGCCATCGCACTTTCGC
S. nitidum	TGTCCCGGCACCACTCGG--TTTTTCTGATTCCCTAAGGTAATGTCGCGCAATTGCACTTTCGC
S. macrospermum	CGATTCGGCAGCACTTGG-- CTGTCCCGAGCGGTACGGATGGCGCTGCCTCAACGCGCTCCCGT
S. laxiflorum	CGATTCGGCAGCACTTGG-- CTGTCCCGAGCGGTACGCATGGCGCCCCCTCAACGCGCTCCCGT
S. versicolor	CGTTCTAGCAACACAAAGG- CTGTCTCTACCGCCACGGGCGGCACCGCGCCAAGACGCCCCCGT
S. purpureo-sericeum	CGTTCCGGCAACACAAAGG-CTGTCTCTACCGCCACGGGCGGCGCCGCGCCAAGACGCCCCCGT
Cleistachne sorghoides	CGTTCCGGCAACACAAAGGGAGCTGTCTCTACCGCCACGGGCGGCGCCGCGCCAAGACGCCCCCGT
S. stipoideum	CGTTCCGGCAACAAGAAGT-CTGTCTCTGCCGCCACGGGCGGTGCCGCGCCAAGACGCACCCAT
S. australiense	CGTTCCGCACACAAGAAGT- CTGTCTCTGCCGCCACGGGCGGCGCCGCGCCAAGAGGACCCTAT
S. matarankense	CGTTCCGGCAACAAGAAGG- CTGTCTCTGCCGCCACGGGCGGCGCCGCGCCGAGACGCCCCCGT
Zea mays cv. h95	CATTCGGGCAAACCCCGCGGCGCCGCCCCTACCGCCACGGGCGGCGCCGCGCCAAGACGCCCCCGT

was helpful in minimizing localized compression problems due to a high GC content.

There was a total of four deletion/insertion events in the 17 species analyzed (Table 2). Compared to the *S. bicolor* sequence, *Z. mays* and *S. bicolor* race arundinaceum had one extra base pair in the ITS1 region. In the ITS2 region, *S. bicolor* and its close relatives, including *S. nitidum*, *S. laxiflorum*, and *S. macrospermum*, appeared to have a 4-bp deletion. The other *Parasorghum* and the *Stiposorghum* species had a 3-bp deletion at the same position. *Z. mays* and *C. sorghoides* did not have the deletion. These differences in deletions may indicate the lineages of species evolution.

rDNA sequence variation

The most common polymorphisms in the ITS region were base substitutions. In addition to the insertion/deletions, 126 polymorphisms were observed as base substitutions; 78 of these were informative (Table 2). Two equally-parsimonious trees (Fig. 2) were constructed from the informative nucleotides. The only differences between the two trees were the placements of *S. purpureo-sericeum*, *S. versicolor* and *C. sorghoides* within the same clade. A strict consensus tree (Fig. 3) was constructed from the two most-parsimonious trees.

The number of nucleotide differences between taxa are shown in Table 3. There were only a few base pair differences among *S. bicolor* races (Table 3). The variations among the accessions of this group range from 0 to 5 bp and they formed a clade in the phylogenetic tree (Fig. 3). This group also included the two other species of the section *Sorghum*, *S. halepense* and *S. propinquum*, due to their identical ITS sequence with that of *S. bicolor*. A species from the section *Parasorghum*, *S. nitidum*, was the next closest species in the analysis. It had 12–16-bp differences from the members of the *S. bicolor* group. The two species in the sections *Heterosorghum* and *Chaetosorghum*, *S. laxiflorum* and *S. macrospermum*, had only 8-bp differences between them, and they formed a single clade in the phylogenetic tree (Fig. 3). Though *S. laxiflorum* and *S. macrospermum* had 53–59-bp differences from the *S. bicolor* group, they fell into the same major clade in the phylogenetic tree as *S. nitidum* and *S. bicolor*. This result supports the hypothesis that the sections *Sorghum*, *Heterosorghum*, and *Chaetosorghum* form one circle of affinity (Garber 1950; Celarier 1957; Wu 1990). However, the placement of *S. nitidum* (Fig. 3) was contradictory to the current consideration that this species is a member of the section *Parasorghum* (Snowden 1935; Garber 1950; Wu 1989).

C. sorghoides and the two *Stiposorghum* species were grouped together with the *Parasorghum* species (except *S. nitidum*) in the phylogenetic tree. The number of nucleotide differences between these species ranged from 3 to 26 bp, while their differences from the taxa in the first group were between 37 to 77 bp (Table 3). This result agrees with the statements that the sections *Parasorghum* and *Stiposorghum* are closely related to each other (Garber 1950; Celarier 1958; Wu 1989), and that *C. sorghoides* is closely related with some *Parasorghum* species (Duvall and Doebley 1990). The tree in Fig. 3 also agrees with that of Duvall and Doebley in that *S. australiense*, a *Parasorghum* species, is more closely related to the *Stiposorghum* species than other *Parasorghum* species. The phylogenetic analysis also suggested that maize is more closely related to most *Parasorghum* species than to *Eusorghum* species (Fig. 3). The number of base-pair differences between *Z. mays* and taxa in the *Parasorghum-Stiposorghum-Cleistachne* group varied between 18 and 36, while the differences varied between 54 and 67 for maize and the taxa in the first group.

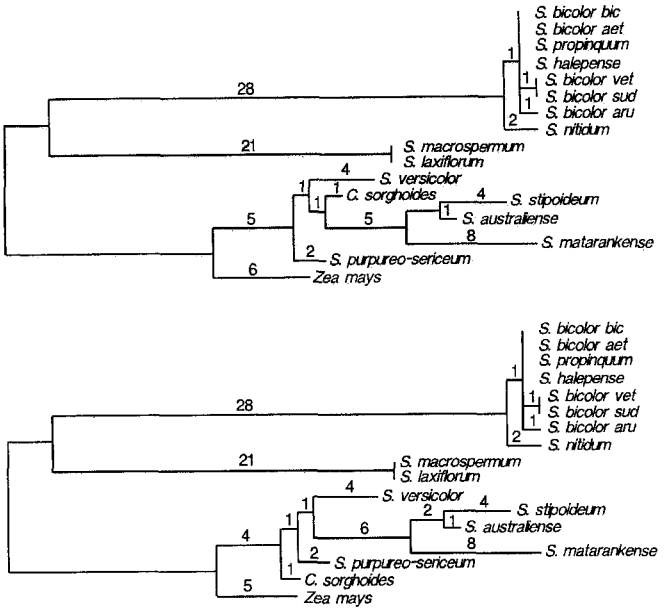

Fig. 2 The two shortest and equally-parsimonious trees constructed by analysis of the rDNA sequence data by Wagner parsimony. Numbers above the branches are the number of steps or the reconstructed number of evolutionary changes along the branch

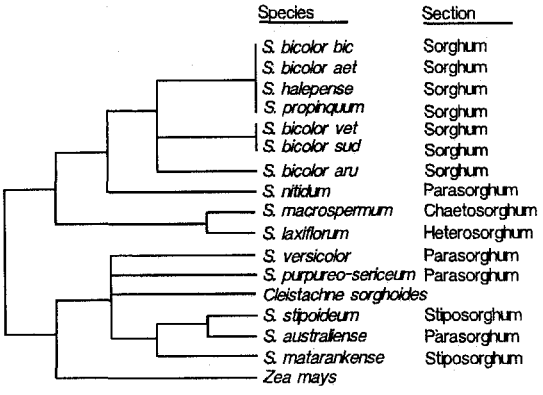

Fig. 3 Phylogenetic tree for 15 *Sorghum* accessions, *Cleistachne sorghoides*, and *Zea mays*. The tree is the strict consensus tree constructed from the two most-parsimonious trees in Fig. 2. *S. bicolor* bic=cv Sanqui red; *S. bicolor* aet=aethiopicum; *S. bicolor* vet=verticiliflorum; *S. bicolor* sud=sudanense; *S. bicolor* aru=arundinaceum

Table 3 Number of nucleotide differences between pairs of taxa

Taxa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 *Sorghum bicolor* cv. sanqui red	–																
2 *S. bicolor* race aethiopicum	0	–															
3 *S. propinquum*	0	0	–														
4 *S. halepense*	0	0	0	–													
5 *S. bicolor* race verticiliflorum	1	1	1	1	–												
6 *S. bicolor* race sudanense	1	1	1	1	0	–											
7 *S. bicolor* race arundinaceum	4	4	4	4	5	5	–										
8 *S. nitidum*	12	12	12	12	13	13	16	–									
9 *S. macrospermum*	53	53	53	5	53	53	57	63	–								
10 *S. laxiflorum*	55	55	55	55	55	55	59	65	8	–							
11 *S. versicolor*	57	57	57	57	58	58	61	66	46	47	–						
12 *S. purpureo-sericeum*	52	52	52	52	53	53	54	60	45	46	11	–					
13 *Cleistachne sorghoides*	47	47	47	47	48	48	49	56	37	38	9	3	–				
14 *S. stipoideum*	54	54	54	54	55	55	56	63	50	52	20	15	11	–			
15 *S. australiense*	63	63	63	63	64	64	65	72	54	55	23	16	13	11	–		
16 *S. matarankense*	68	68	68	68	69	69	70	77	54	55	26	24	16	21	24	–	
17 *Zea mays* cv. h	59	59	59	59	60	60	61	67	54	56	26	18	18	28	31	36	–

The heterogeneity of rDNA

The rDNA sequences were not always homogeneous in the species analyzed. Five clones of *S. nitidum* were sequenced, two of which had a deletion of about 40 bp in the ITS2 region, while the other three did not. The sequences of the clones were otherwise identical, and either sequence resulted in the same position in the phylogenetic tree since deletions, like substitutions, were considered to be single evolutionary events. Similarly, one clone of *S. australiense* had an apparent insertion of roughly 100 bp. It is possible that these differences were caused by PCR artifacts. It is also possible that these represent rDNA copies that have been modified and lost their function. However, the presence of two ribosomal arrays with sequence differences in any of the species should have been detected using PCR products directly as sequencing templates. We did not see any evidence of heterogeneity in any of the tetraploid species, i.e., *S. halepense*, *S. laxiflorum*, and *S. macrospermum*. This, however, should not be taken as evidence of autopolyploidy because ribosomal gene sequences are subject to relatively rapid rates of concerted evolution (Zimmer et al. 1988).

Discussion

Sorghum halepense, *S. propinquum*, the sorghum cultivar, and race aethiopicum of *S. bicolor* subsp. *arundinaceum* all had identical ITS sequences. The races verticilliflorum (subsp. *arundinaceum*) and sudanense (subsp. *drumondii*) of *S. bicolor* had rDNA sequences identical to each other, and differed from the cultivated sorghum by just 1 bp. The race arundinaceum of *S. bicolor* differed by 4 bp from the sorghum cultivar (Table 3). The similarity in ITS sequences between some of the races in the subspecies *arundinaceum* and the cultivated sorghum was expected since analysis of morphological traits (Snowden 1935; Liang and Casady 1966; De Wet and Huckabay 1967; De Wet and Harlan 1971; Doggett 1976; Mann et al. 1983) and chloroplast RFLP patterns (Duvall and Doebley 1990) have suggested that one or more of the races may be the direct progenitors of cultivated sorghum. *Sorghum propinquum* is generally considered to be more distantly related to cultivated sorghum, due primarily to its Asian distribution (De Wet 1978; Wei 1984) and perennial growth habit. However, its otherwise similar morphology to cultivated sorghum, its identical ITS sequence, and the fact that they intercross freely (De Wet 1978; Chittenden et al. 1994) suggest that it could be considered a rhizomatous version of *S. bicolor*.

It has been proposed that *S. halepense* arose within *S. bicolor* (Bhatti et al. 1960). Others have considered it a product of the hybridization of *S. bicolor* and *S. propinquum* (Doggett 1976). The identical rDNA sequence of these species supports either view. *S. halepense* may be considered as an autotetraploid even if the latter view is true, however, since both its probable parents are apparently closely related. This proposal is consistent with the observation made by Tang and Liang (1988) that as many as eight quadrivalents formed in a hybrid between *S. halepense* and a 40-chromosome *S. bicolor*.

Sorghum nitidum has the most extensive range of distribution among all the *Parasorghum* and *Stiposorghum* species. It can be found in southern China, India, southeast Asia, Pacific Islands and northern Australia (Garber 1950). The wide distribution has also given rise to a large amount of variation in morphology and ploidy, having varieties with n=5, 10, or 20 (Celarier 1958; Gu et al. 1984). This species maintains the hairy ring on its nodes, the identifying trait for *Parasorghum* species, but it also possesses characters which make it unique in the section. Like some *Eusorghum* species (i.e., *S. propinquum*), it is a perennial weed with rhizomes (Wu 1984). *S. nitidum* is the only species in the sections *Parasorghum* and *Stiposorghum* which has awnless varieties (Garber 1950). Chromosomes of *S. nitidum* are significantly smaller than those of other spe-

cies in *Parasorghum* and *Stiposorghum*, and resemble those of *S. bicolor* and *S. halepense* (Gu et al. 1984). From the above observations and the rDNA sequence data we obtained, it is apparent that problems exist with the sectional boundaries of the *Parasorghum* group. The close affinity of *S. nitidum* and the members of the section *Sorghum* imply it should either be included in this section or given a unique sectional status.

Sorghum macrospermum is the only species in the section *Chaetosorghum*, and is found only in the Northern Territory of Australia. *Sorghum laxiflorum* is the only species in the section *Heterosorghum*, and is distributed in the Philippine islands, New Guinea, and northern Queensland within Australia (Garber 1950). Both species have 40 chromosomes in their somatic cells. The major morphological differences between the two are that *S. macrospermum* has simple panicles and a poorly-developed awn, while *S. laxiflorum* has divided panicles and awns. Their very similar rDNA sequence constitutions (8-bp difference) suggested that it may be appropriate to combine these two sections.

Geographically-close members of the sections *Parasorghum* and *Stiposorghum* species were grouped together in Fig. 3. The three Australian species, *S. australiense*, *S. matarankense*, and *S. stipoideum,* form a clade, even though they belong to two sections. All these species were originally grouped under the section *Parasorghum* by Snowden (1935). The distinct characters for this section were nodes with rings of hair and simple panicles. Garber (1950) later advocated the division of the section into two sections, in which *Parasorghum* species have an obtuse callus and awns shorter than 65 mm, while species in *Stiposorghum* have a pointed callus and awns longer than 65 mm. However, Wu (1989) has shown that some intermediate shapes of callus exist so that we may not always be able to define them as "pointed" or "obtuse". The close affinity of *S. australiense* to the two *Stiposorghum* accessions, as indicated by rDNA sequence, implies that awn length may not be a good criterion for classification. Our data supported the conclusion of Duvall and Doebley (1990) that the section *Parasorghum* is not monophyletic. It may be appropriate to consider either combining the sections *Parasorghum* and *Stiposorghum* or else reevaluating the taxonomic status of the species in the two sections.

The rDNA of *Z. mays* showed considerable similarity to the *Parasorghum* and *Stiposorghum* species, differing by as few as 18 out of 593 bp (Table 3). The commonly-used explanation for the somatic chromosome number of 20 of maize and sorghum is that they were derived by polyploidization events involving 2n=10 species. If polyploidization events occurred in the evolution of the two genomes they clearly were not recent, however, since both species are generally considered as diploids and homoeologous chromosomes are no longer recognizable in either species (Helentjaris et al. 1988). Comparative mapping of the sorghum and maize genomes has identified numerous genomic areas in which linkage arrangements of DNA sequences are conserved (Hulbert et al. 1990; Whitkus et al. 1992; Melake Berhan et al. 1993). It is conceivable that the most recent common ancestor of maize and sorghum was a 2n=20 species. The ribosomal DNA sequence analysis, however, has indicated that both species probably have 2n=10 chromosome species to which they are more closely related. *S. nitidum* appeared closely related to cultivated sorghum while maize appeared more closely related to the major *Parasorghum/Stiposorghum* group (Fig. 3; Table 3). It seems more likely, therefore, that *S. bicolor* and *Z. mays* acquired their present chromosome number independently.

C. sorghoides, distributed in eastern tropical Africa and western India (Garber 1950), is obviously a close relative of *Sorghum* species though it has 36 somatic chromosomes. Both chloroplast DNA (Duvall and Dobeley 1990) and our rDNA sequence data support this view. It appears that the genus *Cleistachne* has a common progenitor with the *Parasorghum* and *Stiposorghum* species. Whether *Tripsacum* species (n=18) have any inherent relations with the *Cleistachne* species remains uninvestigated. Our data agree with the proposal (Duvall and Doebley 1990) that either the genus *Sorghum* should be treated as several genera or that *Cleistachne* should be included in the genus *Sorghum*.

Acknowledgements The authors express their gratitude to Dr. K. Schertz of Texas A and M University, Dr. T. Wu of National Taiwan University, Dr. P. Lawrence of the Agricultural Department of Australia and Dr. J. Axtell of Purdue University, for providing the seeds. We are also grateful to Dr. E. A. Zimmer for guidance and helpful discussion during the conception of the project.

References

Appels R, Dvorak J (1982) The wheat ribosomal DNA spacer region: its structure and variation in populations and among species. Theor Appl Genet 63:337–348

Baldwin BG (1992) Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol Phylog Evol 1:3–16

Bhatti AG, Endrizzi JE, Reeves RG (1960) Origin of Johnson grass. J Hered 51:106–110

Celarier RP (1957) Cytotaxonomic notes on the subsection halepensia of the genus *Sorghum*. Bull Torrey Bot Club 85:49–62

Celarier RP (1958) Cytotaxonomy of the Andropogoneae. III. Subtribe Sorgeae, Genus *Sorghum*. Cytologia 23:395–418

Chittenden LM, Shertz KF, Lin YR, Wing RA, Paterson AH (1994) A detailed RFLP map of *Sorghum bicolor S. propinquum* suitable for high-density mapping suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. Theor Appl Genet 87:925–933

De Wet JMJ (1978) Systematics and evolution of *Sorghum* sect. *Sorghum* (Gramineae). Am J Bot 65:477–484

De Wet JMJ, Harlan JR (1971) The origin and domestication of *Sorghum bicolor* Econ Bot 25:128–135

De Wet JMJ, Huckabay JP (1967) The origin of *Sorghum bicolor* II. Distribution and domestication. Evolution 21:787–802

Doggett H (1976) Sorghum *Sorghum bicolor* (Gramineae-Andropogoneae). In: Simmonds NW (ed) Evolution of crop plants. Longman Group Limited, London

Duvall MR, Doebley JF (1990) Restriction-site variation in the chloroplast genome of *Sorghum* (Poaceae). Systematic Bot 15:472–480

Garber ED (1950) Cytotaxonomic studies in the genus *Sorghum*. Univ Calif Publ Bot 23:283–361

Gu MH, Ma HT, Liang GH (1984) Karyotype analysis of seven species in the genus *Sorghum*. J Hered 75:196–202

Hamby RK, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Routledge, Chapman and hall, Inc., New York London, pp 50-91

Helentjaris T, Weber D, Wright S (1988) Identification of genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphism. Genetics 118:353–363

Hulbert SH, Bennetzen JL (1991) Recombination at the *RP1* locus of maize. Mol Gen Genet 226:377–382

Hulbert SH, Richter TD, Axtell JD, Bennetzen JL (1990) Genetic mapping and characterization of sorghum and related crops using maize RFLP probes. Proc Natl Acad Sci USA 87:4251–4255

Liang GH, Casady AJ (1966) Quantitative presentation of the systematic relationships among twenty-one *Sorghum* species. Crop Sci 6:76–79

Mann JA, Kimber CT, Miller FR (1983) The origin and early cultivation of sorghum in Africa. Texas Agri Exp St Bull No. 1454

Melake Berhan A, Hulbert SH, Butler LG, Bennetzen JF (1993) Structure and evolution of the genomes of *Sorghum bicolor* and *Zea mays*. Theor Appl Genet 86:598–604

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

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

Snowden JD (1935) A classification of the cultivated sorghum. Bull Misc Info, No.5. Royal Botanic Garden, Kew, England

Springer PS, Zimmer EA, Bennetzen JL (1989) Genomic organization of the ribosomal DNA of sorghum and its close relatives. Theor Appl Genet 77:844–850

Swofford DL (1993) PAUP: Phylogenetic analysis using parsimony. Computer program distributed by Illinois National History Survey. Champaign, Illinois

Takaiwa F, Oono K, Sugiura M (1985) Nucleotide sequence of the 17-25 region from rice rDNA. Plant Mol Biol 4:355–364

Tang H, Liang GH (1988) The genomic relationship between cultivated sorghum and Johnsongrass: a re-evaluation. Theor Appl Genet 76:277–284

Wei S (1984) Elementary survey of the cultivated origin of Kaoliang (Chinese sorghum)–Studies concerning unearthed samples in the historical site, Jingchun village, Wanrong County. Agri His Sinica 2:45–50

Whitkus R, Doebley J, Lee M (1992) Comparative genome mapping of sorghum and maize. Genetics 132:1119–1130

Wu TP (1984) Comparative karyomorphology of two species in *Parasorghum*. Proc Natl Sci Council Taiwan, China. Part B 6:319–325

Wu TP (1989) Cytotaxonomic studies in two polyploid species of the genus *Sorghum*. Proc Nat Sci Council Taiwan, China. Part B 13:15–22

Wu TP (1990) *Sorghum macrospermum* and its relationship to the cultivated species *S. bicolor*. Cytologia 55:141–151

Yokota YT, Lida KY, Kato A, Tanifuji S (1989) Nucleotide sequences of the 5.8 rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. J Mol Evol 29:294–301

Zimmer EA, Jupe ER, Walbot V (1988) Ribosomal gene structure, variation and inheritance in maize and its ancestors. Genetics 120:1125–1136