Restriction Site Variation in the Chloroplast Genome of Sorghum (Poaceae)

Melvin R. Duvall; John F. Doebley

Systematic Botany, Vol. 15, No. 3. (Jul. - Sep., 1990), pp. 472-480.

Stable URL:
http://links.jstor.org/sici?sici=0363-6445%28199007%2F09%2915%3A3%3C472%3ARSVITC%3E2.0.CO%3B2-6

Systematic Botany is currently published by American Society of Plant Taxonomists.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aspt.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Fri Jul 21 08:33:16 2006

Systematic Botany (1990), 15(3). pp. 472-480
© Copyright 1990 by the American Society of Plant Taxonomists

Restriction Site Variation in the Chloroplast Genome of *Sorghum* (Poaceae)

MELVIN R. DUVALL

Department of Botany, W408 #166 NHB, Smithsonian Institution, Washington, D.C. 20560

JOHN F. DOEBLEY

Department of Plant Biology, University of Minnesota, St. Paul, Minnesota 55108

ABSTRACT. Restriction site variation was analyzed in the chloroplast (cp) DNA's of six species of *Sorghum* and one species each of *Cleistachne* and *Zea*. Each cpDNA was separately digested with 14 restriction enzymes and hybridized with six maize or sorghum cpDNA probes that together comprise approximately 75% of the plastid genome. Eighty-four restriction site mutations and two deletion mutations were observed among the cpDNA's of these species. These variable character mutations were used in phylogenetic analyses to construct rooted, most-parsimonious, phylogenetic trees on which confidence limits were placed by bootstrap analysis. The estimated number of nucleotide substitutions per site between each of the cpDNA's was also calculated. These analyses 1) reveal a relatively high level of sequence divergence within the genus *Sorghum* relative to that of other genera; 2) confirm that *Sorghum* sect. *Sorghum* is a monophyletic group; 3) cast doubt on the monophyly of sect. *Parasorghum*; and 4) suggest that the genus *Sorghum* may be either paraphyletic or polyphyletic. The fact that restriction site presence/absence in the plastid genome could be easily studied between genera as divergent as *Zea* and *Sorghum* suggests that analysis of variation in cpDNA could be used to construct a phylogeny for the entire tribe Andropogoneae.

Sorghum Moench, which is classified in the tribe Andropogoneae, has been the object of extensive taxonomic inquiry (e.g., Celarier 1959; deWet 1978; Doggett 1970; Garber 1950; Snowden 1935). These investigations have disclosed considerable morphological variation provoking disagreement over the taxonomic positions of and relationships among the taxa. Five sections are generally recognized. *Sorghum* sect. *Sorghum* consists of three species native to Africa and Asia: *S. bicolor* (cultivated sorghum and its wild relatives), *S. halepense* (Johnson grass), and *S. propinquum* (Kunth) A. Hitchc. *Sorghum* sect. *Parasorghum* contains eight to ten annual and perennial species found throughout portions of the Eastern Hemisphere. *Sorghum* sect. *Stiposorghum* consists of about five species restricted to northern Australia. Sections *Chaetosorghum* and *Heterosorghum* (the latter of which was not included in this study) are monotypic and are found in the Australo-Pacific region.

This project was undertaken to assess the phylogenetic relationships among the species of *Sorghum* using restriction site variation in chloroplast (cp) DNA. In particular, our objectives were 1) to determine if sect. *Sorghum* and sect. *Parasorghum* are distinct monophyletic groups within the genus; and 2) to estimate the amount of cpDNA divergence between species.

Species of two other related genera, *Cleistachne* Benth. and *Zea* L., were included in this study to permit polarization of polymorphic characters. *Zea diploperennis* Iltis, Doebley, & Guzman is a distantly related species in the Andropogoneae. However, the plastid genome of *Zea* has been shown to be collinear to that of *Sorghum* (Dang and Pring 1986) indicating that genomic rearrangements would not complicate comparison of the two genomes. *Cleistachne sorghoides*, most notably distinguished from *Sorghum* by its solitary sessile spikelets and by the lack of pedicels or pedicellate spikelets, has been classified together with *Sorghum* in the supergenus, "Sorghastrae" (Stapf 1917) suggesting that the two are sister genera and that *Cleistachne* is a logical choice for an outgroup.

MATERIALS AND METHODS

Leaves were obtained from six species of *Sorghum* (20 accessions), *Cleistachne sorghoides* (two

TABLE 1. Accessions analyzed. [a] Collection prefixes indicate the source of each accession: IS = The International Crop Research Institute for the Semi-Arid Tropics (ICRISAT); PI = the USDA Regional Plant Introduction Station, Experiment, Georgia; KS = Dr. Keith Shertz, USDA-ARS, Texas A&M University, College Station, Texas 77843. [b] Vouchers of mature specimens have been preserved at the indicated herbaria. [c] *Sorghum almum* is considerd to be a hybrid between *S. halepense* and a diploid from *S. bicolor* subsp. *arundinaceum*.

Taxa	Collection[a]	Herbaria[b]	Origin
Sorghum sect. *Sorghum*			
1. *S. bicolor* (L.) Moench subsp. *bicolor*	IS 6964	—	Sudan
2.	IS 6705	ISC	Guinea
3.	KS TxB623	ISC	USA
4. subsp. *drummondii* (Stend.) deWet & Harlan	IS 26733	—	India
5. subsp. *arundinaceum* (Desv.) deWet & Harlan	IS 14257	MIN	Angola
6.	PI 213900	TAES	Kenya
7.	IS 18856	—	Sudan
8.	IS 18820	TAES	Egypt
9.	IS 18806	TAES	Uganda
10.	IS 21661	MIN	Malawi
11. *S. halepense* (L.) Pers.	IS 18897	ISC	—
12.	IS 18899	MIN	—
13.	IS 14263	ISC	Angola
14. (*S.* × *almum* L. Parodi)[c]	IS 18853	TAES	Thailand
15. (*S.* × *almum*)[c]	IS 14246	ISC	Angola
16.	IS 144212	TAES	Angola
Sorghum sect. *Parasorghum*			
17. *S. nitidum* Pers.	KS 85-248	MIN	Australia
18. *S. australiense* Garber & Snyder	KS 85-261	MIN	Australia
Sorghum sect. *Stiposorghum*			
19. *S. matarankense* Garber & Snyder	KS 85-242	MIN	Australia
Sorghum sect. *Chaetosorghum*			
20. *S. macrospermum* Garber	KS 85-229	MIN	Australia
21. *Cleistachne sorghoides* Benth.	IS 14340	ISC	Malawi
22.	IS 14346	—	Malawi
23. *Zea diploperennis* Iltis, Doebley, & Guzman	Guzman 777	MIN	Mexico

accessions), and one accession of *Zea diploperennis* (table 1). Vouchers of flowering specimens of 18 surviving specimens of the 23 accessions were preserved (table 1). Fresh leaf material was pulverized in liquid nitrogen. Total cellular DNA was extracted using the protocol of Saghai-Maroof et al. (1984) with a slightly modified extraction buffer (100 mM Tris pH 7.5, 1.4 M NaCl, 20 mM EDTA, 2% mixed alkyltrimethylammonium bromide, 1% 2-mercaptoethanol). Approximately 2–4 μg aliquots of each DNA preparation were digested with one of 14 restriction enzymes (Bethesda Research Laboratories or New England Biolabs) including 13 that cut DNA at six-base nucleotide recognition sequences, and one (Cfo I) that cut DNA at a four-base recognition sequence. These DNA digests were electrophoresed in 0.8% agarose gels with a running buffer of 100 mM Tris-acetate, 1 mM EDTA (pH 8.1). After denaturation and neutralization the DNA fragments were transferred to nylon hybridization membranes (Zetabind) in a bidirectional fashion. Membranes were sequentially probed with six non-overlapping cloned portions of the plastid genome (see Doebley et al. 1987 for specific locations of probes) including: 1) a 32 kilobase (kb) cosmid clone, cB9, of maize cpDNA (D. M. Lonsdale, unpubl.); 2–5) Charon 4A clones λ-5 (14.7kb), λ-9 (16.4kb), λ-11 (18.7kb), and λ-12 (13.1kb) of maize cpDNA (Larrinua et al. 1983); and 6) a mixture of a 5.4 kb plasmid

TABLE 2. Restriction site loss and gain mutations observed on autoradiographs after digestion of DNAs with restriction enzymes and hybridization with cpDNA probes. Underscored numbers represent the sizes of fragments that were not visualized. See table 1 for a key to sample numbers.

Mutation	Enzyme	Probe	Fragment losses (kb)	Fragment gains (kb)	Samples mutated
1.	Bam HI	CB9	13.66	9.33 + 3.59	1–16, 23
2.	Dra I		6.11	3.04 + 2.89	1–16, 23
3.	Eco RI		2.43	1.98 + $\underline{0.45}$	17, 18, 19
4.			2.18	1.58 + $\underline{0.60}$	23
5.	Hind III		7.71	4.92 + 2.65	17
6.	Nco I		11.78	7.37 + 4.21	23
7.	Xmn I		3.65	3.51 + $\underline{0.14}$	18, 19
8.			2.22	1.95 + $\underline{0.27}$	18, 19
9.			3.21	1.81 + 1.45	20
10.	Bam HI	λ-5	4.35 + 3.68	8.66	18, 19
11.	Bgl II		6.18 + 1.17	6.95	20
12.	Cfo I		2.47 + 1.50	3.62	17
13.	Cla I		6.36 + 1.50	7.25	23
14.			3.62	3.09 + $\underline{0.53}$	23
15.			1.55 + $\underline{0.26}$	1.81	23
16.	Dra I		5.07	4.21 + 1.17	1–3, 5, 6, 8–10, 12
17.			5.07	3.01 + 2.20	18, 19
18.			2.40	1.87 + $\underline{0.53}$	17
19.			2.40	1.17 + 1.25	20
20.	Eco RI		1.57	1.37 + $\underline{0.20}$	21, 22
21.			2.38	2.26 + $\underline{0.12}$	23
22.	Hind III		2.68 + 1.49	3.80	23
23.	Nco I		12.18 + 7.42	18.57	18, 19
24.			7.28 + $\underline{1.27}$	8.55	20
25.	Ssp I		3.49	2.96 + $\underline{0.53}$	23
26.	Stu I		3.03	1.58 + $\underline{0.60}$	23
27.	Bgl II	λ-9	4.32	3.22 + 1.55	21, 22
28.	Cfo I		4.19 + 1.29	5.26	20
29.	Dra I		3.80	2.53 + 1.54	18, 19
30.			2.58 + $\underline{0.65}$	3.23	10
31.			2.24	2.09 + $\underline{0.15}$	1–6, 9
32.			3.64	3.41 + $\underline{0.23}$	23
33.	Eco RI		3.74	3.01 + 1.13	23
34.	Nru I		29.84	20.16 + 9.35	1–17, 20
35.	Nsi I		8.62 + 1.42	9.95	17–23
36.	Ssp I		2.05	1.47 + 1.05	17
37.			2.14	1.37 + 1.23	20
38.			2.25	1.64 + $\underline{0.61}$	23
39.	Bam HI	λ-11	8.56 + 1.89	10.45	2, 7, 11, 13–16
40.	Bgl II		4.67 + 2.19	7.01	20
41.	Eco RI		7.52 + $\underline{1.26}$	9.64	18, 19
42.	Hind III		5.64 + 1.25	6.97	23
43.	Nco I		7.85	6.65 + 1.36	18, 19
44.	Nru I		14.03	7.59 + 5.25	23
45.	Sac II		32.73	27.01 + 4.17	1–7, 9–11, 13–16
46.	Ssp I		2.33	1.87 + $\underline{0.46}$	1–3, 5, 9
47.			4.28	2.00 + 2.25	18
48.	Stu I		7.33 + 1.88	9.49	8, 12

TABLE 2. Continued.

Mutation	Enzyme	Probe	Fragment losses (kb)	Fragment gains (kb)	Samples mutated
49.			7.18 + 6.78	14.34	23
50.			4.98 + 1.80	6.78	23
51.	Xmn I		1.34	1.30 + 0.04	1–16
52.	Bam HI	λ-12	3.05	3.02 + 0.03	21, 22
53.			7.13 + 3.40	11.58	23
54.	Bgl II		4.42 + 0.20	4.62	18
55.	Cla I		11.01	8.98 + 1.91	18, 19
56.			3.43 + 0.36	3.79	1–17, 20–22
57.	Cfo I		2.42 + 2.42	4.79	23
58.	Dra I		6.20	5.30 + 1.18	1–17, 20–22
59.			3.46	2.32 + 1.16	21, 22
60.			1.25 + 0.12	1.37	21, 22
61.			3.46	2.10 + 1.54	23
62.	Nco I		6.38 + 2.17	8.08	18, 19
63.	Nsi I		14.51	7.59 + 7.36	18, 19
64.			7.36	4.86 + 2.50	18, 19
65.	Sac II		17.56 + 2.26	21.93	18, 19
66.	Ssp I		2.27	1.47 + 0.80	1–3, 5, 9
67.	Xmn I		4.09 + 1.31	5.44	23
68.	Bam HI	pLD7 + pB7	5.23	4.09 + 1.46	1–16
69.			4.09	3.85 + 0.20	18, 19
70.			4.20	3.45 + 1.15	23
71.	Bgl II		15.19	10.98 + 3.55	23
72.	Cfo I		1.42	1.31 + 0.11	18, 19
73.			1.53 + 0.76	2.29	18
74.	Cla I		2.97	2.88 + 0.09	18, 19
75.	Dra I		3.71 + 0.22	3.93	1–17, 20
76.			4.33	3.16 + 0.86	18
77.			0.79 + 0.14	0.93	20
78.			3.71	1.68 + 1.68	23
79.	Eco RI		2.39	2.15 + 0.24	18, 19
80.	Nru I		15.76	12.90 + 1.72	23
81.	Nsi I		3.63	3.53 + 0.10	18, 19
82.	Stu I		21.88 + 12.13	32.63	18, 19
83.	Xmn I		2.76 + 1.28	3.88	18, 19
84.			1.18 + 0.13	1.31	18, 19

clone pLD7 of sorghum cpDNA (Dang and Pring 1986) and pB7, a 3.5 kb Pst I fragment cloned from maize cpDNA (Doebley 1989). Together these regions account for slightly more than 75% of the plastid genome excluding part of the inverted repeat and the extremity of the large single copy region containing the 32-kilodalton thylakoid membrane protein (psbA).

Phylogenetic trees were rooted using *Zea diploperennis* as the outgroup. *Cleistachne sorghoides*, originally selected for outgroup rooting, proved to be an unsuitable candidate (see Discussion).

The "Phylogenetic Analysis Using Parsimony" (PAUP version 2.4.1) software package, written by David Swofford, was used to construct a set of rooted, equally parsimonious, phylogenetic Wagner trees using the "branch-and-bound" function. The "Phylogeny Inference Package" (PHYLIP, version 3.1) computer software system, written by Joseph Felsenstein, was used to construct a majority-rule consensus Wagner parsimony tree by 200 subsamplings of the original data matrix in a "bootstrap" analysis (Felsenstein 1985) and to place confidence intervals

TABLE 3. Number of bands scored from the autoradiographs of DNA digests hybridized against chloroplast probes for each probe with each restriction enzyme. The number of variable sites observed, if any, is listed in parentheses. The number of variable sites per kilobase for each probe is given in the last row.

	Probe					
Enzyme	cB9	λ-5	λ-9	λ-11	λ-12	pLD7 + pB7
Bam HI	10 (1)	6 (1)	3	7 (1)	7 (2)	8 (3)
Bgl II	8	6 (1)	8 (1)	7 (1)	4 (1)	3 (1)
Cfo I	13	9 (1)	8 (1)	8	5 (1)	5 (2)
Cla I	7	9 (3)	5	6	7 (2)	5 (1)
Dra I	8 (1)	11 (4)	10 (4)	9	10 (4)	9 (4)
Eco RI	10 (2)	11 (2)	7 (1)	8 (1)	5	5 (1)
Hind III	9 (1)	8 (1)	3	10 (1)	2	2
Nco I	6 (1)	5 (2)	2	5 (1)	4 (1)	1
Nru I	4	3	5 (1)	6 (1)	1	3 (1)
Nsi I	8	2	9 (1)	5	5 (2)	4 (1)
Sac II	5	1	1	4 (1)	3 (1)	1
Ssp I	8	10 (1)	9 (3)	9 (2)	7 (1)	1
Stu I	6	3 (1)	2	8 (3)	1	3 (1)
Xmn I	15 (3)	6	6	9 (1)	6 (1)	5 (2)
Mutations per kb	0.28	1.2	0.73	0.70	1.2	1.9

on monophyletic groups. The log-likelihood of selected alternative trees was tested using a maximum likelihood method for restriction sites data (program RESTML in PHYLIP) excluding the five Cfo I mutations because of a program limitation permitting only a single length of enzyme recognition site.

Estimates of the number of nucleotide substitutions per site (p) between each of the cpDNA's were calculated by the maximum likelihood method of Nei and Tajima (1983) using equations 19, 21, and 23 computed by the program MAXLIKE that was provided by Nei. Note that these estimates of p were calculated from restriction site polymorphisms in only 75% of the plastid genome and may not be strictly comparable to those reported in other studies.

RESULTS

Restriction site polymorphisms (losses/gains) were identified by the additive pattern of DNA fragment sizes that were visualized on the autoradiographs (table 2). Fragments shorter than about 1 kb in length could not be routinely visualized by the techniques used in this study (although size differences of as little as 50 base pairs (bp) between DNA fragments larger than 1 kb could be resolved). Thus, restriction site losses or gains that produced fragments smaller than 1 kb presupposed the existence of small unseen fragments to account for the observed differences. Such suppositions are reasonable, especially since the fragment patterns resulting from restriction site mutations are distinguishable from those resulting from small insertion/deletion events.

In two instances (mutations 24 and 41, table 2) slightly larger fragments of lengths 1.26 kb and 1.27 kb that were not visualized were hypothesized to account for observed differences in fragment patterns. Fragments of this length are not small enough to escape detection by the methods used in this study. For this reason, we hypothesize that mutation 24 involves a site change that lies outside the region covered by the probe λ-5 used to detect these fragments. If this is true, then it would explain the inability to detect the 1.27 kb fragment. A similar explanation could also account for the inability to detect the 1.26 kb fragment associated with mutation 41 as there is a gap between the two probes λ-11 and λ-12.

The numbers of DNA fragments scored for each restriction enzyme with each probe for all taxa are given in table 3. The highest frequency of restriction site mutations, 1.2–1.9 mutations per kb, were observed in the probes corresponding to the termini of the large single copy region of the genome (λ-5, λ-12, and pLD7+pB7). The lowest frequency of mutations was observed with the cB9 probe, which corresponds to the entire small single copy region and adjoining parts of the two inverted repeats. A similar distribution of variation within the plastid genome has been observed in previous studies (Doebley et al. 1987; Sytsma and Gottlieb 1986).

An estimated 484 restriction sites were observed among the 22 accessions (excluding *Zea diploperennis*) representing 2.0% of the plastid genome. Sixty of the 484 sites (12.4%) were variably present among the cpDNA's, over three times the variability found among five species of *Zea* and *Tripsacum* L. (Doebley et al. 1987) and nearly six times that found in 14 species of *Triticum* L. and *Aegilops* L. (Bowman et al. 1983). An additional 24 site mutations distinguished *Zea diploperennis* from the other taxa.

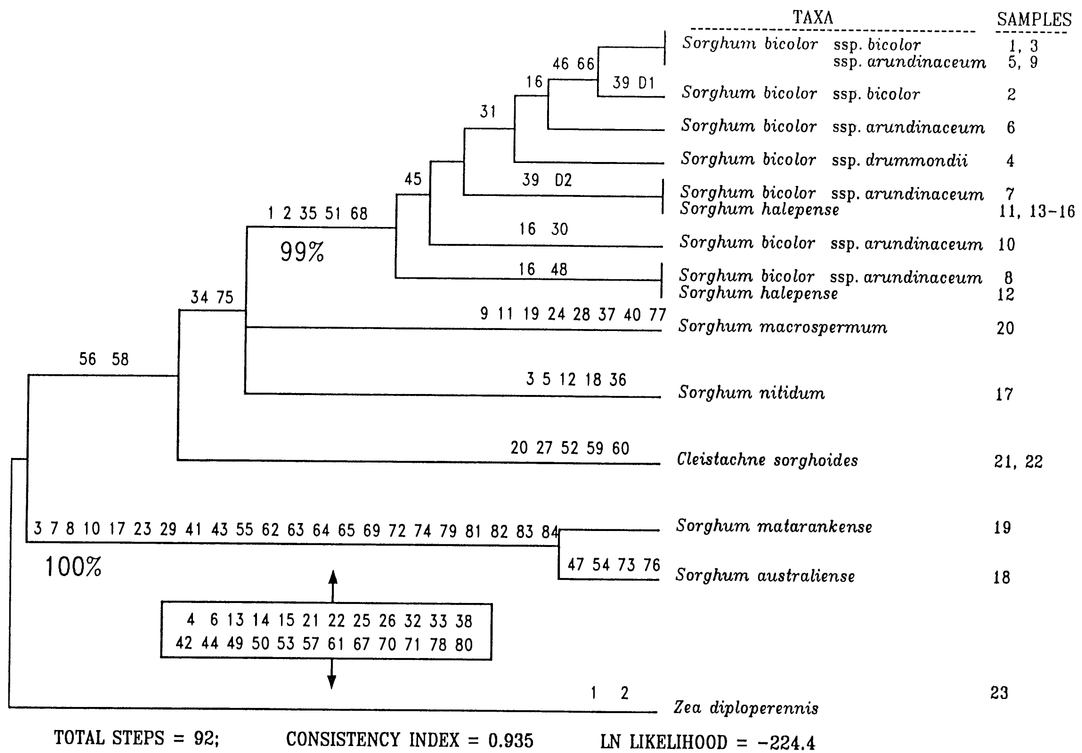

FIG. 1. Phylogenetic tree for six species of *Sorghum, Cleistachne sorghoides,* and *Zea diploperennis* (see table 1 for sample numbers). The tree was constructed using 84 cpDNA restriction site mutations and two deletion mutations (D1 and D2). Identification numbers (table 2) for the restriction site mutations appear above the branch segments. The tree is the majority rule consensus tree generated by 200 bootstrap subsamplings of the original data set executed in the PHYLIP software system, and also one of a set of most parsimonious trees generated by PAUP analysis using the "branch and bound" function. The two monophyletic groups for which bootstrap confidence levels exceed 95% are identified by the percent confidence level at the branch point. Branch lengths are arbitrary.

Thirteen unique restriction site profiles were found among the 23 cpDNA's analyzed in this study. Within *Sorghum* sect. *Sorghum,* four accessions of *S. bicolor* (samples 1, 3, 5, and 9, table 1) exhibited identical profiles. Sample 7 of *S. bicolor* and five accessions of *S. halepense* (samples 11, and 13–16, table 1) also shared an identical restriction profile, as did samples 8 and 12 of *S. bicolor* and *S. halepense,* respectively. Thus, the species of this section show a relatively high level of intraspecific cpDNA variation and this variation does not appear to be very well correlated with taxonomic boundaries.

Two insertion/deletion events were observed, among the taxa of sect. *Sorghum,* as consistent decreases for some samples in the size of one fragment with nine of the 14 restriction digests probed with λ-11. The most parsimonious interpretations of these observations are that the fragments of reduced size represent 1) a deletion of about 200 bp in the cpDNA's of samples 7, 11, and 13–16; and 2) a deletion of about 290 bp in the cpDNA of one accession of *S. bicolor* (sample 2).

Of the 60 variable restriction sites, 37 were phylogenetically informative, i.e., shared by two or more cpDNA's. Phylogenetic analysis of restriction site variation by the bootstrap algorithm of PHYLIP produced the consensus tree seen in figure 1. In this analysis, the monophyly of the 16 accessions of *Sorghum* sect. *Sorghum,* and that of *S. matarankense* Garber & Snyder and *S. australiense* Garber & Snyder, are both supported at a confidence level of 99% or above. This tree was one of a set of equal length (92 steps), most parsimonious trees constructed by

TABLE 4. Estimates of the number of nucleotide substitutions per site, p (given as $100p$) and their standard errors in parentheses (upper triangle) and the number of mutations (lower triangle) among the cpDNA's of six species of *Sorghum, Cleistachne sorghoides,* and *Zea diploperennis.* Letters refer to groups of samples with identical cpDNA's. A = samples 1, 3, 5, and 9; B = samples 7, 11, and 13–16; C = samples 8 and 12; D = samples 21 and 22. See table 1 for identification of sample numbers.

	A	2	4	6	B	C	10	17	18	19	20	D	23
A	—	0.062 (0.036)	0.189 (0.063)	0.146 (0.056)	0.316 (0.082)	0.360 (0.088)	0.252 (0.073)	0.945 (0.015)	2.586 (0.259)	2.296 (0.241)	1.107 (0.160)	1.122 (0.161)	2.216 (0.237)
2	1	—	0.252 (0.073)	0.209 (0.067)	0.253 (0.074)	0.425 (0.096)	0.316 (0.082)	1.012 (0.152)	2.662 (0.264)	2.371 (0.246)	1.176 (0.166)	1.191 (0.166)	2.283 (0.240)
4	3	4	—	0.126 (0.052)	0.126 (0.052)	0.297 (0.080)	0.190 (0.064)	0.750 (0.130)	2.426 (0.250)	2.159 (0.233)	0.955 (0.148)	0.882 (0.141)	2.058 (0.227)
6	2	3	1	—	0.210 (0.067)	0.254 (0.074)	0.147 (0.056)	0.836 (0.137)	2.468 (0.252)	2.180 (0.234)	0.997 (0.151)	0.968 (0.148)	2.148 (0.233)
B	5	4	1	3	—	0.298 (0.080)	0.190 (0.064)	0.752 (0.130)	2.432 (0.250)	2.144 (0.232)	0.957 (0.148)	0.884 (0.142)	2.015 (0.225)
C	5	6	4	3	4	—	0.233 (0.071)	0.798 (0.134)	2.439 (0.251)	2.149 (0.233)	0.960 (0.149)	0.931 (0.146)	1.973 (0.222)
10	4	5	3	2	3	3	—	0.817 (0.136)	2.453 (0.251)	2.165 (0.234)	0.978 (0.150)	0.950 (0.147)	2.084 (0.229)
17	15	16	12	13	12	12	13	—	2.165 (0.233)	1.881 (0.215)	0.800 (0.135)	0.773 (0.132)	2.037 (0.226)
18	40	41	37	38	37	37	38	33	—	0.422 (0.096)	2.544 (0.258)	2.165 (0.233)	3.443 (0.310)
19	36	37	33	34	33	33	34	29	4	—	2.252 (0.240)	1.881 (0.215)	3.135 (0.292)
20	18	19	15	16	15	15	16	13	38	34	—	0.978 (0.150)	2.415 (0.251)
D	17	18	13	14	14	14	14	12	33	29	15	—	2.037 (0.226)
23	36	37	33	34	33	33	34	35	51	47	38	33	—

PAUP analysis. The topologies of the trees in this set differed from each other mainly because of unresolved trifurcations among species within sect. *Sorghum.* Five of the 84 restriction site mutations (6.0%) demonstrate homoplasy. The consistency index of each tree in this set was 0.935.

The number of nucleotide substitutions per site, p, between each pair of the 13 cpDNA restriction site profiles is given in table 4. The mean estimate of p over all intrageneric pairwise values for the six species of *Sorghum* is 0.0111 (range: 0.0006–0.0266), over seven times that of *Zea* (Doebley et al. 1987). The mean and range of p among species of *Zea* is comparable to those for just the 16 accessions of *Sorghum* sect. *Sorghum,* for which $p = 0.0023$ (range: 0.0006–0.0036).

The mean intergeneric estimate of all pairwise values of p between *Sorghum* and *Zea* is 0.0235 (range: 0.0197–0.0344), comparable to the values of 0.0155 and 0.0316 previously obtained for those genera based on comparative restriction maps (Doebley et al. 1987) and nucleotide sequence data (Doebley et al., unpubl. data), respectively. The mean intergeneric estimates of p between *Cleistachne* and *Sorghum* is 0.0116 (range: 0.0077–0.0216), two to four times the intergeneric values calculated for closely related grass genera including *Pennisetum* Rich. ex Pers. and *Cenchrus* L. (Clegg et al. 1984) and *Zea* and *Tripsacum* (Doebley et al. 1987).

DISCUSSION

Our results (fig. 1) suggest several phylogenetic conclusions regarding sect. *Sorghum.* First, cpDNA data clearly indicate that the species of

sect. *Sorghum* belong to a monophyletic group distinguished from the other sections of *Sorghum* by the synapomorphic mutations 35, 51, and 68. The monophyly of this section is confirmed by bootstrap analysis at a 99% confidence level. Second, although five of the six accessions of *S. halepense* possessed the same cpDNA restriction site profile, *S. halepense* and *S. bicolor* are not clearly distinguished as monophyletic taxa within the section (fig. 1). This situation could result from reciprocal introgression between these two species. *Sorghum bicolor* and *S. halepense* can form fertile hybrids when an unreduced pollen grain of *S. bicolor* fertilizes *S. halepense* (Celarier 1958). Introgression could also be achieved through partially fertile triploid intermediates. Third, our data show that cultivated sorghum is identical in its cpDNA restriction site profile to its near wild relative (*S. bicolor* subsp. *arundinaceum* (Desv.) deWet & Harlan) supporting the hypothesis that it is a cultivated form of this wild taxon.

The relatively high degree of cpDNA variation that we discovered among the accessions of *S. bicolor* is noteworthy. In all, seven cpDNA types defined by 10 restriction site mutations and two insertion/deletion events were observed. This is remarkable considering that only 10 accessions of this species were analyzed and that only 14 restriction enzymes were employed. These data suggest that *S. bicolor* possesses cpDNA variation in excess of that which has been reported in other angiosperm species (Soltis et al. 1989). As mentioned above this may result in part from introgression between *S. bicolor* and *S. halepense*.

Sorghum australiense and *S. matarankense* share 22 synapomorphic restriction site mutations (fig. 1) and can be considered a monophyletic group with a confidence of 100% by bootstrap analysis. Further, these two species are highly diverged from the remaining species of *Sorghum* with an overall mean p value of 0.0232. These results suggest that the ancestor of *S. australiense* and *S. matarankense* has undergone more rapid cytoplasmic evolution than other species of *Sorghum*. *Sorghum australiense* and *S. matarankense* have traditionally been placed in separate sections of the genus, sect. *Parasorghum* and sect. *Stiposorghum*, respectively.

Sorghum nitidum Pers. and *S. australiense*, traditionally classified together in sect. *Parasorghum*, exhibit numerous cpDNA restriction site differences reflected in a large divergence value ($p = 0.0216$) between them. Thus, our data do not support the monophyly of sect. *Parasorghum*. However, the log-likelihood of a rearrangement of the consensus tree so that *S. nitidum* and *S. australiense* are monophyletic is not significantly different (log-likelihood difference of only -13.6) from that of the tree of fig. 1.

The cpDNA's of the two accessions of *Cleistachne sorghoides* exhibit five synapomorphic mutations (fig. 1, table 2) but also share two restriction site mutations (56 and 58, table 2) with the cpDNA's of four of the six species of *Sorghum* placing *Cleistachne* within the cytoplasmic lineage of *Sorghum*. While these data suggest that *Sorghum* is not a monophyletic lineage, we must emphasize that our data do not constitute proof of this point. The tree of figure 1 is not significantly different in log-likelihood (a difference of only -2.7) from one in which *Cleistachne* is forced into an outgroup position. The latter tree is slightly longer, requiring two additional steps.

The relatively large degree of variation observed between the cpDNA's of the species of *Sorghum* and between *Sorghum* and *Cleistachne* suggests either that the cpDNA's of these taxa are evolving more rapidly than those of other grass genera or that the genus is rather ancient. Note that hybridization and introgression are unlikely to have affected our results since fertile intersectional hybrids are unknown in *Sorghum* and species from different sections differ substantially in chromosome morphology (Garber 1950).

Our data suggest that the genus *Sorghum* is either paraphyletic or polyphyletic (discriminating between these conditions requires identification and analysis of ancestral groups) and that sect. *Sorghum* is more closely related to *Cleistachne* than it is to sect. *Stiposorghum*. Thus, our data could be used to support arguments that either *Sorghum* should be treated as several distinct genera or that *Cleistachne* should be submerged within *Sorghum*. However, which changes, if any, should be made must await more detailed taxonomic and phylogenetic analyses.

Finally, in these analyses we have been able to compare readily the cpDNA restriction site profiles of *Zea* and *Sorghum*, two genera which probably represent opposite ends of the tribe Andropogoneae. This is not surprising given the high degree of conservation of the plastid

genome between these two genera (Dang and Pring 1986). This result indicates that it should be possible to employ the analytical approach used in this study to construct a phylogeny for the entire tribe.

ACKNOWLEDGMENTS. We thank Dr. Keith Schertz (Department of Soil and Crop Sciences, Texas A&M University) for contributing leaf material from four accessions of *Sorghum*, and Drs. Kay Klier (Department of Botany, Iowa State University) and Patricia Davila (Departamento de Botanica, Universidad Nacional de Mexico) for several verifications of taxonomic identity. This research was supported in part by a grant from the Graduate School of The University of Minnesota and a grant (BSR-8508490) from the National Science Foundation.

LITERATURE CITED

BOWMAN, C. M., G. BONNARD, and T. A. DYER. 1983. Chloroplast DNA variation between species of *Triticum* and *Aegilops*: Location of the variation on the chloroplast genome and its relevance to the inheritance and classification of the cytoplasm. Theoret. Appl. Genet. 65:247–262.

CELARIER, R. P. 1958. Cytotaxonomy of the Andropogoneae. III. Subtribe Sorgheae, genus *Sorghum*. Cytologia 23:396–421.

———. 1959. Cytotaxonomy of the Andropogoneae. IV. Subtribe Sorgheae. Cytologia (Tokyo) 24:285–303.

CLEGG, M. T., J. R. RAWSON, and K. THOMAS. 1984. Chloroplast DNA variation in pearl millet and related species. Genetics 106:449–461.

DANG, L. H. and D. R. PRING. 1986. A physical map of the sorghum chloroplast genome. Pl. Molec. Biol. 6:119–123.

DEWET, J. M. J. 1978. Systematics and evolution of *Sorghum* sect. *Sorghum* (Gramineae). Amer. J. Bot. 65:477–484.

DOEBLEY, J. F. 1989. Molecular evidence for a missing wild relative of maize and the introgression of its chloroplast genome into *Zea perennis*. Evolution 43:1555–1559.

———, W. RENFROE, and A. BLANTON. 1987. Restriction site variation in the *Zea* chloroplast genome. Genetics 117:139–147.

DOGGETT, H. 1970. *Sorghum*. London: Longmans, Green and Co.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

GARBER, E. D. 1950. Cytotaxonomic studies on the genus *Sorghum*. Univ. Calif. Publ. Bot. 23:283–362.

LARRINUA, I. M., K. M. T. MUSKAVITCH, E. J. GUBBINS, and L. BOGORAD. 1983. A detailed restriction endonuclease site map of the *Zea mays* plastid genome. Pl. Molec. Biol. 2:129–140.

NEI, M. and F. TAJIMA. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genetics 105: 207–217.

SAGHAI-MAROOF, M. A., K. M. SOLIMAN, R. A. JORGENSEN, and R. W. ALLARD. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. U.S.A. 81: 8014–8018.

SNOWDEN, J. D. 1935. A classification of the cultivated sorghums. Bull. Misc. Inform. 1935:221–255.

SOLTIS, D. E., P. S. SOLTIS, T. A. RANKER, and B. D. NESS. 1989. Chloroplast DNA variation in a wild plant, *Tolmiea menziesii*. Genetics 121:819–826.

STAPF, O. 1917. Gramineae. Pp. 1–154 in *Flora of tropical Africa* IX. Ashford, England: Reeve & Co.

SYTSMA, K. J. and L. D. GOTTLIEB. 1986. Chloroplast DNA evolution and phylogenetic relationships in *Clarkia* sect. *Peripetasma* (Onagraceae). Evolution 40:1248–1261.