Plant Syst. Evol. 249: 233–246 (2004)
DOI 10.1007/s00606-004-0210-7

Plant Systematics and Evolution
Printed in Austria

Sorghum laxiflorum and *S. macrospermum*, the Australian native species most closely related to the cultivated *S. bicolor* based on ITS1 and *ndh*F sequence analysis of 25 *Sorghum* species

S. L. Dillon[1], **P. K. Lawrence**[1], **R. J. Henry**[2], **L. Ross**[3], **H. J. Price**[4], and **J. S. Johnston**[5]

[1]Australian Tropical Crops and Forages Collection, Queensland Department of Primary Industries and Fisheries, Biloela, QLD, Australia
[2]Centre for Plant Conservation Genetics, Southern Cross University, Lismore, NSW, Australia
[3]Dallas, TX, USA
[4]Department of Soil & Crop Sciences, Texas Agricultural Experiment Station, Texas A&M University, College Station TX, USA
[5]Department of Entomology, Texas Agricultural Experiment Station, Texas A&M University, College Station TX, USA

Received March 8, 2004; accepted June 9, 2004
Published online: October 18, 2004

Abstract. Australian species make up seventeen of the world's twenty-five recognised species of *Sorghum*, with the genus separated into five sections: *Eu-sorghum, Chaetosorghum, Heterosorghum, Para-sorghum* and *Stiposorghum*. Whereas the genetic relationships within section *Eu-sorghum* are well known, little is known about the genetic relationships and crossabilities outside the primary genepool. We made a detailed investigation of phylogenetic relationships within *Sorghum* to identify wild species most closely related to cultivated sorghum (with outgroups *Zea mays* and *Saccharum officinarum*). The ribosomal ITS1 gene of ten species and the chloroplast *ndh*F gene from nineteen species were sequenced. Independent and combined analyses of the ITS1 and *ndh*F data sets were undertaken. The *Eu-sorghum* species were resolved into a strongly supported lineage by all three analyses, and included the Australian natives *S. laxiflorum* and *S. macrospermum* in the ITS1 and combined analyses. All remaining sorghum species were resolved into a second well-supported lineage in the combined analyses, although some internal relationships within this second lineage remain unresolved. We identified *S. laxiflorum* and *S. macrospermum* as the Australian species most closely related to cultivated sorghum. Our data support a reduction in the number of subgeneric sections from five to three: *Eu-sorghum* (unchanged); a combined *Chaetosorghum*/*Heterosorghum* to reflect the very close relationship between these two species; and a combined *Para-sorghum*/*Stiposorghum* section, thereby removing the unclear taxonomic and genetic boundaries between these species.

Key words: Wild sorghum, phylogeny, ITS1, *ndh*F, Poaceae.

Introduction

Australian species of *Sorghum* make up seventeen out of the world's twenty-five recognised *Sorghum* species, with the genus separated into five taxonomic sections: *Eu-sorghum,*

Chaetosorghum, Heterosorghum, Para-sorghum and *Stiposorghum*. The six *Eu-sorghum* species consist of the cultivated *Sorghum bicolor* (L.) Moench, *S.* x *almum* Parodi, *S. arundinaceum* (Desv.) Stapf, *S.* x *drummondii* (Steud.) Millsp. & Chase, *S. halepense* (L.) Pers., and *S. propinquum* (Kunth) Hitchc. The *Eu-sorghums* originate from Africa and Asia with $2n = 20$ and 40 chromosomes, and the known progenitor of cultivated *S. bicolor* is *S. arundinaceum* (DeWet and Harlan 1971, Doggett 1976, Duvall and Doebley 1990). *Sorghum drummondii* originated from a natural cross between *S. bicolor* and *S. arundinaceum*, while *S. propinquum* is reputedly a perennial rhizomatous form of *S. bicolor* (Doggett 1976, Chittenden et al. 1994, Sun et al. 1994). *Sorghum halepense* (Johnson grass) is derived from a doubling of the chromosomes in a natural cross between *S. arundinaceum* and *S. propinquum*, while a natural cross between *S. bicolor* and *S. halepense* gave rise to *S. x almum* (Doggett 1976). The close genetic relationships and inter-crossability between the *Eu-sorghum* species are well known (Magoon and Shambulingappa 1961, Wu 1979, Chittenden et al. 1994, Paterson et al. 1995, Stenhouse et al. 1997).

The section *Chaetosorghum* contains the single species *S. macrospermum* E. D. Garber that is endemic to a small area in the Northern Territory of Australia. *Sorghum laxiflorum* F. M. Bailey forms the section *Heterosorghum*, is native to northern Australia and Papua New Guinea, and is geographically more diverse than its close relative *S. macrospermum* (Garber 1950, Lazarides et al. 1991, Dillon et al. 2001). The section *Para-sorghum* contains the seven species. *S. grande* Lazarides, *S. leiocladum* (Hack.) C. E. Hubb., *S. matarankense* E. D. Garber & Snyder, *S. nitidum* (Vahl) Pers., *S. purpureo-sericeum* (Hochst. ex. A. Rich.) Asch. & Schweinf., *S. timorense* (Kunth) Buse and *S. versicolor* Andersson with these species native to northern monsoonal Australia, Africa and Asia (Garber 1950, Lazarides et al. 1991, Phillips 1995). *Sorghum amplum* Lazarides, *S. angustum* S. T. Blake, *S. brachypodum* Lazarides, *S. bulbosum* Lazarides, *S. ecarinatum* Lazarides, *S. exstans* Lazarides, *S. interjectum* Lazarides, *S. intrans* F. Muell. ex Benth., *S. plumosum* (R. Br.) P. Beauv., and *S. stipoideum* (Ewart & Jean White) C. A. Gardner & C. E. Hubb. form the section *Stiposorghum*, all being endemic to northern Australia (Garber 1950, Lazarides et al. 1991).

The gene pool concept is useful to describe the total pool of different genes within a genus. The primary gene pool consists of species that readily cross producing viable hybrids with chromosomes that freely recombine. The secondary gene pool consists of species with a certain degree of hybridisation barriers due to ploidy differences, chromosome alterations, or incompatibility genes making gene transfer difficult. The tertiary gene pool consists of species from which gene transfer is very difficult due to strong sterility barriers (Harlan and de Wet 1971). In *Sorghum*, the *Eu-sorghum* species form both the primary and secondary gene pools, while *Chaetosorghum, Heterosorghum, Para-sorghum* and *Stiposorghum* species form the tertiary gene pool. Advances in biotechnology are increasing the use of these tertiary genepool species in breeding programs.

Direct evaluation of pest and disease resistances in non-primary wild sorghum species (sections *Chaetosorghum, Heterosorghum, Para-sorghum* and *Stiposorghum*) have been undertaken over the last twenty years (Bapat and Mote 1982, Karunakar et al. 1994, Franzmann and Hardy 1996, Sharma and Franzmann 2001, Kamala et al. 2002, Komolong et al. 2002). These studies show that many Australian native species of *Sorghum* contain valuable genes conferring resistances to pests (Australian species are non-hosts to the sorghum midge [*Stenodiplosis sorghicola* Coquillett]) and diseases already in Australia. Importantly, these species have resistances to some pests (shoot fly [*Atherigona soccata* Rondani], greenbug [*Schizaphis graminum* Rondani], stem borer [*Chilo partellus* Swinhoe] etc) and diseases (sorghum downy mildew [*Peronosclerospora sorghi* (Weston & Uppal) C.G. Shaw] etc) currently not present in

Australia. These non-primary wild *Sorghum* species are therefore valuable sources of genes that could confer resistance to many pests and diseases currently affecting cultivated sorghum production. However, the phylogenetic relationships that exist between the cultivated and wild species must be established to act as a guide to using these wild species for breeding.

Whereas the phylogenetic relationships within section *Eu-sorghum* (primary and secondary gene pools) are well known, little is known about the phylogenetic relationships and crossabilities of the tertiary genepool species. Recent research has attempted to determine the phylogenetic relationships between primary and non-primary *Sorghum* species, however many relationships within the genus remain unresolved (Sun et al. 1994, Spangler et al. 1999, Dillon et al. 2001, Spangler 2003).

The ribosomal ITS1 work by Dillon et al. (2001) attempted to determine the phylogenetic relationships between all 25 *Sorghum* species, and obtained two distinct lineages. The *Eu-sorghum* species were all closely related with strong statistical support, with most Australian native species in a separate lineage. Many of the relationships between the Australian natives remained unresolved based on ITS1 alone (Dillon et al. 2001). Research on a subset of *Sorghum* species led Spangler (2003) to suggest that three distinct lineages exist within *Sorghum*. Although many relationships both between and within these lineages remain unresolved, Spangler (2003) suggests that each lineage should be recognised as distinct genera *Sorghum*, *Sarga* and *Vacoparis*. However, the order of evolution of species cannot be fully resolved without evaluating all species of a genus, and hence, the recognition of three distinct genera for *Sorghum* is premature.

To resolve the phylogenetic relationships between the Australian *Sorghum* species and further strengthen known relationships, the chloroplast *ndh*F gene was sequenced. The *ndh*F gene is located at one end of the small single copy region of the chloroplast genome, is one of eleven subunits (*ndh*A-K) and encodes the ND5 protein of the chloroplast NADH dehydrogenase involved in chloroplast respiration (Ferguson 1999, Kim and Jansen 1995, Bohs and Olmstead 1997). The *ndh*F has a relatively high rate of molecular evolution and provides three times more phylogenetic information than other chloroplast genes because of its length and sequence divergence (Olmstead and Sweere 1994, Kim and Jansen 1995). The highest rate of variability has been shown in the 3′ end of the gene, with 60% of the nucleotide substitutions and all of the alignment gaps positioned there (Olmstead and Sweere 1994, Clark et al. 1995, Catalán et al. 1997). The *ndh*F gene has delineated closely related species within many plant groups including sunflower, solanum and Poaceae (Clark et al. 1995, Kim and Jansen 1995, Neyland and Urbatsch 1996, Bohs and Olmstead 1997, Catalán et al. 1997, Spangler et al. 1999, Catalán and Olmstead 2000, Spangler 2003). Because *ndh*F gene sequences were able to elucidate relationships between closely related Poaceae species, we decided to use this gene to further resolve the phylogenetic relationships between the Australian native sorghums.

To fully determine the phylogenetic relationships between all 25 *Sorghum* species, new ITS1 sequences were generated for nine *Sorghum* species plus *Cleistachne sorghoides* Benth., and *ndh*F gene sequences were determined for 19 species (including the two outgroup species *Zea mays* (L.) and *Saccharum officinarum* (L.)). This new evaluation confirmed two lineages within *Sorghum,* and showed that two endemic species *S. laxiflorum* and *S. macrospermum* are the Australian species most closely related to cultivated *Sorghum bicolor*.

Materials and methods

Seed material and accession catalogue information. All seed material for the *Sorghum* and *Cleistachne* species were obtained from the Australian Tropical Crops and Forages Collection, Queensland Department of Primary Industries

and Fisheries, Australia. The catalogue information on each accession used for DNA extraction is shown in Table 1. ITS1 and *ndh*F sequence data were available from Genbank for some *Sorghum* species and the outgroup *Zea mays* and are shown in Table 2. The two outgroups *Zea mays* and *Saccharum officinarum* were chosen to represent species from different subtribes within Andropogoneae that have recently diverged from *Sorghum* (approx. 15 mya), and that have been resolved to separate clades from *Sorghum* in previous phylogenetic studies (Al-Janabi et al. 1994, Sun et al. 1994, Spangler et al. 1999).

DNA Extraction from fresh leaf tissue. For DNA extraction, seed from each accession was germinated and grown for four to six weeks. Total genomic DNA was extracted from fresh leaf tissue of *Sorghum* and *Cleistachne* using the hexadecyltrimethylammonium bromide (CTAB) method described in Dillon et al. (2001). Dr Giovanni Cordeiro (Centre for Plant Conservation Genetics, Southern Cross University, Lismore NSW Australia) kindly provided genomic DNA for *Saccharum officinarum* cultivars Fiji38 and Q124. At least two individuals per accession were used for sequencing, which gave a minimum of four sequences per species from which concensus sequences were generated.

Ribosomal ITS1 PCR amplification and sequencing. The ribosomal ITS1 gene was amplified from ten species (Table 1) following the protocol detailed in Dillon et al. (2001), and using the primers described by White et al. (1990). All amplified ITS1 products were visualised on 1% agarose gels and products purified using Qiaquick PCR purification columns (QIAGEN) following set protocols. Sequencing reactions for the ITS1 were carried out by the Australian Genome Research Facility (University of Queensland, St Lucia) using ABI PRISM ® BigDye™ Terminator v3.0 protocols (www.agrf.org.au). Purified DNA were sequenced using an ABI PRISM 377 DNA sequencer, and tracked and analysed using ABI sequencing software.

Chloroplast *ndh*F PCR amplification and sequencing. The chloroplast *ndh*F gene was amplified and sequenced from 19 species (Table 1) using primers described by Olmstead and Sweere (1994). Two overlapping fragments of length 1318 bp (fragment A) and 1138 bp (fragment B) were amplified using primers 1F and 1318R, and 972F and 2110R respectively. PCR reactions were carried out in a Perkin Elmer Geneamp 9700 thermocycler in a total reaction volume of 50µl and contained 0.5µM each primer, 0.2 mM equimolar dNTPs, 2 units *Taq* DNA Polymerase (Roche Diagnostics, Mannheim, Germany), *Taq* DNA reaction buffer (to 10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl_2, pH 8.3) and approximately 40 ng DNA template. PCR cycling conditions followed 93 °C for 1 min, 53 °C for 1 min and 72 °C for 2 min, for 35 cycles, with a final extension phase of 7 minutes at 72 °C. All amplified *ndh*F fragments were visualised on 1.0% agarose gel with ethidium bromide. The amplified *ndh*F fragments were purified using QIAGEN Qiaquick PCR purification columns and followed set protocols. Fragment A was sequenced using the internal primers 1F, 536F, 536R, 972F, 972R and 1318R, while the Fragment B was sequenced using the primers 972F, 1318F, 1318R, and 2110R. Two extra primers 1821F and 1821R (described by Clark et al. 1995) were required to completely sequence Fragment B. Sequencing reactions and gel separations for the *ndh*F gene fragments A and B of *S. angustum* and *S. macrospermum* were performed by the Australian Genome Research Facility, St Lucia, Queensland, Australia following set protocols. All other sequencing reactions were performed by Mr Larry Ross using ABI PRISM ® BigDye™ Terminator v3.0 protocols.

ITS1 and *ndh*F sequence alignment and analyses. New ITS1 and chloroplast *ndh*F sequences were manually checked to ensure no errors in base calling were present. The two overlapping sections of the chloroplast *ndh*F gene were aligned using ClustalX (Thompson et al. 1997), and a consensus sequence for each species generated. Very few differences between accessions within species occurred, and where observed differing bases were allocated the appropriate International Union of Biochemistry (IUB) nucleotide code (www.sequenceanalysis.com/code.html). Multiple sequence alignments were obtained separately using ClustalX, and verified and modified where inconsistencies were located. The chloroplast *ndh*F sequences generated do not contain approximately 100 bp from the 3′ end of the gene (between primer 2110R and the stop codon). Sequence alignment matrices for both the ITS1 and *ndh*F data sets are available upon request from the corresponding author.

Table 1. Catalogue information including Australian Tropical Crops and Forages Collection accession numbers, Herbarium taxonomic voucher numbers, and the Genbank accession numbers assigned to the new sequences for both the ITS1 and *ndh*F analyses

Species & Subgeneric section	Ribosomal ITS1		
Section Eu-sorghum	Accession No.	Voucher No.[a]	Genbank
S. drummondii (Steud.) Millsp. & Chase	PI 330272	PI 330272	AY048866
S. propinquum (Kunth) Hitchc.	AusTRCF 302546	BRI AQ 773674	AY282488
Section Chaetosorghum			
S. macrospermum E. D. Garber	AusTRCF 302367	DNA C867	AY048867
Section Para-Sorghum			
S. grande Lazarides	AusTRCF 302580	BRI AQ585960	AF302914
S. matarankense E. D. Garber & Snyder	AusTRCF 302521	BRI AQ 773673	AY282490
	AusTRCF 302637	DNA D129480	
S. timorense (Kunth) Buse	AusTRCF 302532	BRI AQ 773672	AY048871
	AusTRCF 302660	DNA D129474	
Section Stiposorghum			
S. angustum S. T. Blake	AusTRCF 302596	BRI AQ 585973	AF302924
	AusTRCF 302604	BRI AQ 585980	
S. brachypodum Lazarides	AusTRCF 302480	CANB 480297	AF302925
	AusTRCF 302670	DNA D133019	
S. ecarinatum Lazarides	AusTRCF 302648	DNA D129449	AF302926
	AusTRCF 302661	DNA D129486	
Cleistachne sorghoides Benth.	AusTRCF 317661	IS 14340	AY282494
	Chloroplast ***ndh*F**		
Section Eu-sorghum	Accession No.	Voucher No.[a]	Genbank
S. drummondii (Steud.) Millsp. & Chase	PI 330272	PI 330272	AY282483
S. propinquum (Kunth) Hitchc.	AusTRCF 302546	BRI AQ 773674	AY282484
S. x almum Parodi	AusTRCF 302385	[–]	AY282482
Section Chaetosorghum			
S. macrospermum E. D. Garber	AusTRCF 302367	DNA C867	AY048873
Section Heterosorghum			
S. laxiflorum F. M. Bailey	AusTRCF 302503	BRI AQ 773670	AY282470
Section Para-Sorghum			
S. grande Lazarides	AusTRCF 302580	BRI AQ585960	AY282476
S. matarankense E. D. Garber & Snyder	AusTRCF 302517	BRI AQ 773676	AY282480
S. nitidum (Vahl) Pers.	AusTRCF 302539	CANB 479893	AY282471
S. purpureosericeum (Hochst. ex. A. Rich.) Asch. & Schweinf.	AusTRCF 318068	IS 18945	AY282472
Section Stiposorghum			
S. amplum Lazarides	AusTRCF 302623	DNA D129461	AY282473
S. angustum S. T. Blake	AusTRCF 302604	BRI AQ 585980	AY048874
S. brachypodum Lazarides	AusTRCF 302670	DNA D133019	AY282479
S. ecarinatum Lazarides	AusTRCF 302648	DNA D129449	AY282481
S. exstans Lazarides	AusTRCF 302577	BRI AQ 586005	AY282475

Table 1 (continued)

S. interjectum Lazarides	AusTRCF 302563	BRI AQ 585985	AY282478
S. intrans F. Muell. ex Benth.	AusTRCF 302390	BRI AQ 773629	AY282477
S. plumosum (R. Br.) P. Beauv.	AusTRCF 302489	BRI AQ 773634	AY282474
Cleistachne sorghoides Benth.	AusTRCF 317661	IS 14340	AY282469
Saccharum officinarum (L.)	Fiji 38	[–]	AY282485

[a] Herbarium voucher specimen prefixes: DNA = Northern Territory Herbarium, Darwin, NT Australia; CANB = Australian National Herbarium, Canberra, ACT Australia; BRI = Queensland Herbarium, Mt Coot-tha, QLD Australia.

Table 2. Genbank accession numbers for the ITS1 and *ndh*F of *Sorghum* species

Species & Subgeneric section	ITS1 Genbank	*ndh*F Genbank
Section Eu-sorghum		
S. bicolor (L.) Moench	AF302921	U21981
S. arundinaceum (Desv.) Stapf	AF302923	AF117429
S. halepense (L.) Pers.	AF302920	AF117424
S. x almum Parodi	AF302922	[–]
Section Heterosorghum		
S. laxiflorum F. M. Bailey	AF302927	[–]
Section Para-Sorghum		
S. leiocladum (Hack.) C. E. Hubb.	AF302915	AF117426
S. nitidum (Vahl) Pers.	AF302917	[–]
S. purpureosericeum (Hochst. ex. A. Rich.) Asch. & Schweinf.	U04793	[–]
S. timorense (Kunth) Buse	[–]	AF117431
S. versicolor Andersson	U04795	AF117432
Section Stiposorghum		
S. amplum Lazarides	AF302912	[–]
S. bulbosum Lazarides	AF302910	AF117423
S. exstans Lazarides	AF302918	[–]
S. interjectum Lazarides	AF302911	[–]
S. intrans F. Muell. ex Benth.	AF302916	[–]
S. plumosum (R. Br.) P. Beauv.	AF302913	[–]
S. stipoideum (Ewart & Jean White) C. A. Gardner & C. E. Hubb.	AF302909	AF117430
Zea mays (L.)	U46612	U21985
Saccharum officinarum (L.)	AF302919	[–]

The new ITS1 and *ndh*F sequences determined in our study were combined with existing data available from Genbank (Tables 1 and 2). The *S. angustum, S. brachypodum* and *S. ecarinatum* ITS1 accessions (Genbank AF302924 to AF302926) generated by Dillon et al. (2001) have been updated with new verified sequences.

Phylogenetic analyses were carried out on the separate ITS1 and *ndh*F data sets and in combination using PAUP*4.0b10 (Swofford 2002). Maximum parsimony branch and bound searches were performed with all characters having equal weight and gaps treated as missing. Branches with minimum length of zero were collapsed to create polytomies, and duplicate trees were eliminated from the set of most parsimonious trees. Full heuristic bootstrap analyses were generated using 10,000 replicates with TBR branch swapping and

multrees option in effect. Bootstrapping is the percentage of times the clade is recovered from 10,000 repetitions of tree construction. Prior to the combined analysis of the ITS1 and *ndh*F data set, congruence between the nuclear and chloroplast datasets was tested using the partition-homogeneity test using PAUP with a significance level of $P < 0.01$ (also called the incongruence length difference test; Farris et al. 1995). This test compares the sums of lengths of the most parsimonious trees of the data analyses to the distribution of the sums of lengths of the most parsimonious trees (we used random partition of the characters, 100 repetitions, and TBR branch swapping). The partition-homogeneity test provided $P = 0.01$ indicating that the ITS1 and ndhF datasets were congruent and could be used for combined analysis.

Results

ITS1 analysis showing a single resolved lineage in *Sorghum*. Aligned ITS1 sequences were 252 bp in length, and contained 47 parsimony informative characters that on maximum parsimony analysis generated 60 trees of length 127 and consistency index (CI) of 0.764. The strict consensus of the 60 trees is shown in Fig. 1a, with the bootstrap support for each clade shown above the branches. A single lineage (A) was resolved with moderate bootstrap support (70%) that contained the *Eu-sorghum* species (clade B, 100% bootstrap) and the Australian natives *S. laxiflorum* and *S. macrospermum* (clade C, 100% bootstrap).

***Ndh*F consensus tree showing four lineages.** The sequence length of the aligned chloroplast *ndh*F data set was 2014 bp with only 26 parsimony informative characters. One hundred trees of length 92 and CI of 0.628 were generated by maximum parsimony analysis, with the strict consensus tree generated from these 100 trees showing four lineages (Fig. 1b). Clade B retains the *Eu-sorghum* species as in Fig. 1a. *Sorghum laxiflorum* and *S. macrospermum* form clade C and includes the Australian native species *S. nitidum* (bootstrap = 79%). The African grass *Cleistachne sorghoides* and *S. versicolor* (of African origin) form lineage D with 81% bootstrap support. All remaining *Sorghum* species form lineage E that is very strongly supported by bootstrap data (98%). Internal relationships within lineage E are either weakly supported by bootstrap data (55–59%) or remain unresolved (Fig. 1b).

Combined analysis of ITS1 and *ndh*F data showing partially unresolved lineages. The sequence data for ITS1 and *ndh*F had a combined length of 2266 bp with 73 parsimony informative characters that generated 96 trees of length 230 and with a consistency index of 0.752 using maximum parsimony analysis. The strict consensus tree of the 96 most parsimonious trees is shown in Fig. 2. Lineage A is again resolved including the *Eu-sorghum* species, *S. laxiflorum* and *S. macrospermum,* but the bootstrap support for the lineage has dropped to a moderate 63% compared with individual ITS1 and *ndh*F analysis. Internal clades within lineage A are strong, with 100% support for clade B consisting of *Eu-sorghum* species, and 100% support for clade C containing *S. macrospermum* and *S. laxiflorum*. All remaining *Sorghum* species plus *Cleistachne sorghoides* form a second lineage (lineage H) that has strong bootstrap support (85%) for the grouping of all of these species together. Support for internal relationships within lineage H vary. Very strong support (98%) exists for clade E containing *S. brachypodum* through to *S. purpureo-sericeum*, while many relationships within this clade remain unresolved. The African species *S. versicolor* and *Cleistachne sorghoides* again form clade D with strong bootstrap support (90%). The combined analysis resolves some relationships within *Sorghum* more strongly than either the ITS1 or *ndh*F analyses alone, however, many relationships within the genus remain unresolved.

Discussion

Two lineages are resolved by our combined analysis of ribosomal ITS1 and chloroplast *ndh*F data. Lineage (A) consists of the six

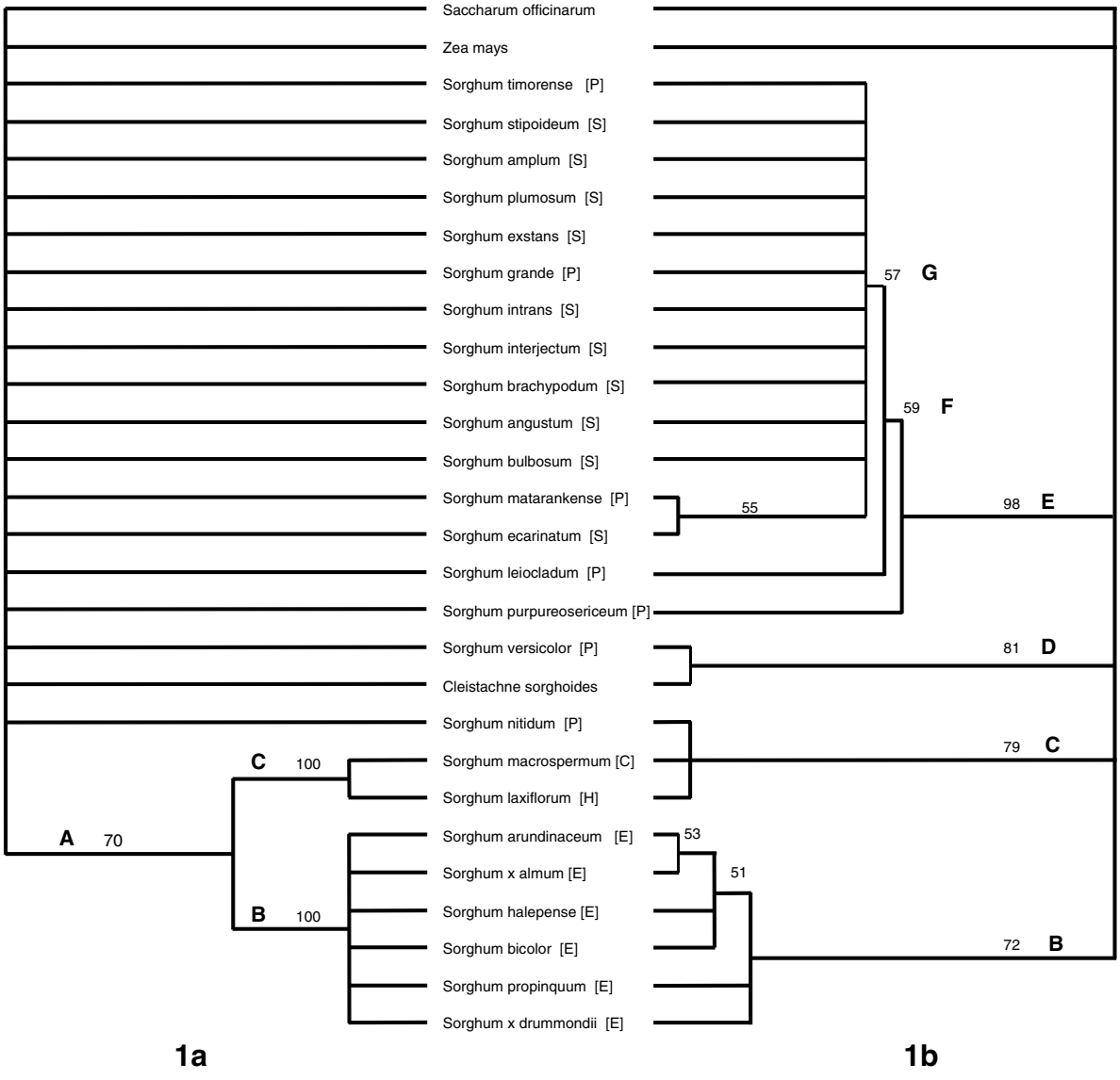

Fig. 1. The strict concensus trees generated using PAUP branch and bound maximum parsimony analysis. Letters A–D designate clades discussed in the text. **1a**. ITS1 data: strict concensus of 60 equally parsimonious trees of 127 steps and consistency index (CI) of 0.764. **1b.** *ndh*F data: strict consensus of 100 equally most parsimonious trees of 92 steps and consistency index (CI) of 0.826. Numbers above branches are percentages of 10,000 bootstrap replicates in which each clade was recovered. Trees were rooted using *Zea mays*. Letters in parenthesis indicate taxonomic sections within *Sorghum* where P = Para-sorghum, S = Stiposorghum, C = Chaetosorghum, H = Heterosorghum and E = Eu-sorghum

Eu-sorghum species (clade B) plus the natives *S. laxiflorum* and *S. macrospermum* (clade C), making these the two Australian species most closely related to the cultivated species *S. bicolor*. All remaining Australian *Sorghum* species plus the exotic species *S. purpureosericeum, S. versicolor* and *Cleistachne sorghoides* are strongly resolved to a second lineage H (Fig. 2).

The very close relationships between the *Eu-sorghum* species are well documented (Magoon and Shambulingappa 1961, Wu

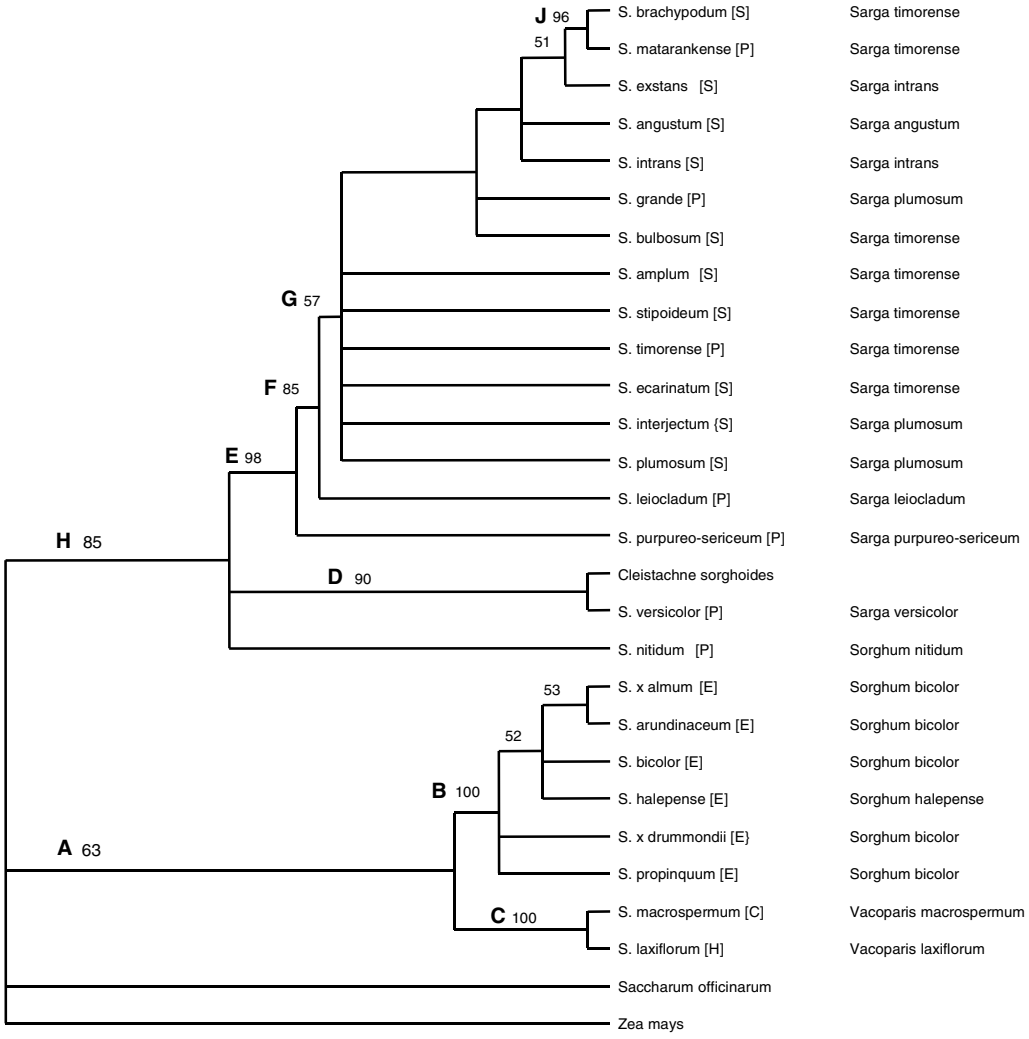

Fig. 2. The strict consensus tree of 96 equally parsimonious trees of 230 steps (CI = 0.752) for the combined ITS1 *ndh*F data under maximum parsimony analyses. Numbers above branches are percentages of 10,000 bootstrap replicates in which the clade was recovered. Letters A and E designate clades discussed in the text. Trees were rooted using *Zea mays*. Letters in parenthesis indicate taxonomic sections within *Sorghum* where P = Para-sorghum, S = Stiposorghum, C = Chaetosorghum, H = Heterosorghum and E = Eu-sorghum. Taxa on right indicate suggested reclassification of *Sorghum* species into the three genera *Sorghum, Sarga* and *Vacoparis* described by Spangler (2003)

1979, Chittenden et al. 1994, Paterson et al. 1995, Stenhouse et al. 1997, Spangler et al. 1999, Dillon et al. 2001, Spangler 2003), and are again reiterated in all our analyses where the *Eu-sorghums* formed clade B with very strong bootstrap support.

Sorghum laxiflorum (*Heterosorghum*) and *S. macrospermum* (*Chaetosorghum*) are very similar in taxonomy, both annuals and $2n = 40$ (Lazarides et al. 1991). In our analyses, *S. laxiflorum* and *S. macrospermum* formed clade C with 100% bootstrap support situated within the same lineage (A) as the *Eu-sorghum* species with moderate bootstrap of 63%. This indicates two things: that *S. laxiflorum* and *S. macrospermum* are very closely related, and that they are more closely associated with the *Eu-sorghum* species than they are with other

Australian native species. It is well established that *S. laxiflorum* and *S. macrospermum* are cytologically, morphologically and genetically distinct from the other Australian native *Sorghum* species (Garber 1950, Celarier 1958, Wu 1990, Sun et al. 1994, Spangler et al. 1999, Dillon et al. 2001, Spangler 2003, Price et al. 2004). Our analyses show that *S. laxiflorum* and *S. macrospermum* have a closer genetic association with the *Eu-sorghum* rather than the Australian species which supports the theory of Garber (1950) and Wu (1990), and that the *Eu-sorghum, Heterosorghum,* and *Chaetosorghum* species form one circle of affinity, and the *Para-sorghum* and *Stiposorghum* another.

Recent molecular and taxonomic evaluation led Spangler (2003) to suggest that *S. laxiflorum* and *S. macrospermum* are distinct enough from all other *Sorghum* species to form their own new genus, *Vacoparis*. Our data does not support the suggestion by Spangler (2003) to describe these species to the new genus *Vacoparis* as *V. laxiflorum* and *V. macrospermum*. Our data does support the removal of the subgeneric boundary between *Chaetosorghum* and *Heterosorghum* to create a unified subgeneric section for *S. macrospermum* and *S. laxiflorum*.

Para-sorghum and *Stiposorghum* species are characterised by bearded culm nodes, with subgeneric separation of these species traditionally based on the length and shape of the sessile spikelet callus (Garber 1950). In several instances, the differences in callus morphology identify taxa to the species level (Lazarides et al. 1991). The callus length and shape are not good indicators of subgeneric taxa as they are continuous across many *Para-sorghum* and *Stiposorghum* species (Spangler 2003). The subjective nature of callus morphology is reflected in all morphological and genetic phylogenies of *Sorghum* where no clear resolution of the *Para-sorghum* and *Stiposorghum* into discreet clades has occurred (Sun et al. 1994, Spangler et al. 1999, Dillon et al. 2001).

In our study, the *Para-sorghum* and *Stiposorghum* species are resolved to varying degrees depending on the gene sequence used. Under our ITS1 analysis, all *Para-sorghum*/*Stiposorghum* species are an unresolved polytomy (Fig. 1a). This differs from the ITS analysis of both Sun et al. (1994) and Dillon et al. (2001) where most of the *Para-sorghum* and *Stiposorghum* were resolved into a lineage separate from the *Eu-sorghum*, *Heterosorghum* and *Chaetosorghum* species. Although our ITS1 had a relatively high proportion of phylogenetically informative characters (18.7%), it was unable to resolve the very close relationships within the Australian *Sorghum* species.

Using chloroplast *ndh*F analysis, most of the *Para-sorghum*/*Stiposorghum* species were resolved into lineage E (Fig. 1b) with very strong bootstrap support (98%). *Sorghum purpureo-sericeum* is the basal species on this lineage followed by *S. leiocladum*, which are both sister to the large unresolved polytomy of 13 *Para-sorghum*/*Stiposorghum* species (clade G, 58% bootstrap). Our *ndh*F analysis produced a phylogeny significantly different from that produced by Spangler et al. (1999) who resolved three lineages of *Sorghum*. Lineage 1 contained *S. bicolor, S. halepense, S. arundinaceum* and *Cleistachne sorghoides;* lineage 2 *S. laxiflorum, S. nitidum* and *S. macrospermum*; and lineage 3 with *S. angustum, S. bulbosum, S. stipoideum, S. timorense, S. leiocladum* and *S. sp*. The *S. sp* accession has since been identified as *S. versicolor* (Kamala et al. 2002). Our *ndh*F analysis resolved lineage B that roughly correlates to Spangler lineage 1 (minus *Cleistachne sorghoides*), our lineage E correlates to Spangler lineage 3 less *S. versicolor*. Also in E were most of the *Para-sorghum* and *Stiposorghum* analysed by Spangler et al. (1999). Our lineage C correlates with Spangler lineage 2. Lineage D resolved a close relationship between *S. versicolor* (African origin) and *Cleistachne sorghoides* (African, Asian origin) that supports the earlier study by Dillon et al. (2001), but contradicts the *ndh*F relationship of *Cleistachne sorghoides* being closest to *S. arundinaceum* (*Eu-sorghum*) depicted by Spangler et al. (1999).

In contrast to the findings of Sun et al. (1994) and Spangler (2003), *S. nitidum* is not resolved within the same lineage as the *Eu-sorghum* species. In our ITS1 analysis it remains unresolved (Fig. 1a); under our *ndh*F analysis, *S. nitidum* is most closely associated with *S. laxiflorum* and *S. macrospermum* (lineage C, Fig. 1b); and in our combined analysis, it is a basal species of the *Para-sorghum*/*Stiposorghum* lineage H (Fig. 2). The *S. nitidum* ITS1 sequence published by Sun et al. (1994) was generated from an accession of Chinese origin. Our *S. nitidum* sequence was obtained from six taxonomically verified Australian accessions, and was significantly different (11.6%) from the sequence generated by Sun et al. (1994). Despite every effort, we were unable to obtain any Chinese *S. nitidum* seed for any taxonomic or genetic verification of the accession used by Sun et al. (1994). As there was no within species differences in our ITS1 sequences from six taxonomically verified plants, we can assume until proven otherwise that the inclusion of all *Sorghum* species in our analysis has depicted the more distant genetic relationship between *S. nitidum* and the *Eu-sorghum* species.

A comparison of our *S. nitidum ndh*F sequence and that generated by Spangler et al. (1999) showed no differences except that we resolved a large block of ~60 bp of N bases around the 930bp mark in their sequence. This led us to also resolve the blocks of N bases in three other Spangler et al. (1999) *ndh*F sequences. The resolution of these N bases and the inclusion of all *Sorghum* species in our *ndh*F analysis produced an aligned dataset with 26 phylogenetically informative characters (1.3% of total length) that was able to resolve *S. nitidum* as more closely related to the *Para-sorghum* and *Stiposorghum* species (bootstrap 85%).

In 1950, Garber determined that *S. nitidum* and *S. leiocladum* shared at least one genome with very high homology, confirming that these two species share a common ancestor. Gu et al. (1984) showed that the chromosomes of *S. nitidum* were smaller than the chromosomes of other *Para-sorghum* species, and suggested they were like those of *S. bicolor* and *S. halepense*. Recent studies of chromosome numbers and DNA content showed that *S. nitidum* chromosomes resemble those of other *Para-sorghum* and *Stiposorghum* species rather than the much smaller chromosomes of *S. bicolor* (Price et al. 2004). The evidence of Garber (1950) and our resolution of the closer relationship of *S. nitidum* to the *Para-sorghum*/*Stiposorghum* species, refutes the suggestion of Spangler (2003) that *S. nitidum* is closer to the *Eu-sorghum* species.

In his most recent publication, Spangler (2003) suggests that the *Para-sorghum* and *Stiposorghum* species (excluding *S. nitidum*) should be reclassified into the resurrected genus *Sarga*, collapsing the sixteen species to only seven. Our combined ITS1 *ndh*F analysis (Fig. 2) shows a clear lineage of the *Para-sorghum* and *Stiposorghum* species with very strong bootstrap support (clade E, 98%) although many internal relationships are weakly supported or unresolved (lineage H, Fig. 2). The suggested reclassification of the *Para-sorghum* and *Stiposorghum* species as *Sarga* species by Spangler (2003) is depicted on the right of Fig. 2. Spangler (2003) suggests collapsing *Sorghum amplum, S. brachypodum, S. bulbosum, S. ecarinatum, S. matarankense, S. stipoideum* and *S. timorense* into the single species *Sarga timorense*. *Sorghum exstans* and *S. intrans* are collapsed into *Sarga intrans*, while *Sorghum grande, S. interjectum* and *S. plumosum* are collapsed into *Sarga plumosum*. The taxa *Sorghum angustum, S. leiocladum, S. purpureosericeum* and *S. versicolor* are maintained as discreet species renamed to *Sarga*.

The proposed *Sarga timorense* species are resolved to clade G with relatively weak bootstrap support in our analysis. *Sorghum brachypodum* and *S. matarankense* are strongly resolved to clade J and are distinct from *Sorghum timorense* (Fig. 2). The $x = 5$ genome sizes reported in Price et al. (2004) also show that *Sorghum timorense* is distinct from the other proposed *Sarga timorense* species in terms of chromosome and genome size. These

species are closely related, but unless it can be demonstrated that there are no reproductive barriers between the species, and regular pairing of chromosomes occurs in hybrids, it is illogical to reduce them into the single species *Sarga timorense* at this time.

While not as well supported at the branch point, collapsing *Sorghum exstans* and *S. intrans* to form *Sarga intrans*, and reducing *Sorghum grande*, *S. interjectum* and *S. plumosum* to *Sarga plumosum* is questionable given the current published data from ITS, *ndh*F and chromosomal studies. If the Spangler system is adapted as currently proposed, it creates more problems than it solves. Taxonomic reclassification must be done after sufficient data involving combined genetic, cytogenetic, ecological, classical systematic, and molecular phylogenies have accumulated to justify the change, which to date has not been achieved.

Conclusions

Based on our combined analysis, it is evident that *Sorghum laxiflorum* and *S. macrospermum* are very closely related and are closer to the *Eu-sorghum* species rather than the Australian native species. Sufficient evidence has accumulated to create a single subgeneric section within *Sorghum* to encompass both *S. laxiflorum* and *S. macrospermum* (combine sections *Chaetosorghum* and *Heterosorghum*). All the *Para-sorghum* and *Stiposorghum* species were resolved into a strongly supported lineage, however, discreet clades for each of these subgeneric sections was not achieved using our dataset. Given the continuous nature of the morphological characters (callus shape and length) used to separate the *Para-sorghum* and *Stiposorghum* species, and given the many genetic studies that have failed to resolve these species into two discreet clades, it would be practical to combine them into a single subgeneric section.

Funding for this research was provided in part by the Australian Grains Research Development Corporation Project DAQ00015; a minigrant from the Texas A&M Office of University Research; the Queensland Department of Primary Industries and Fisheries; and the Centre for Plant Conservation Genetics, Southern Cross University. The authors would also like to thank Robert Brown, Prof. Pauline Ladiges and two reviewers for their valuable comments on the draft manuscript.

References

Al-Janabi S. M., McClelland M., Petersen C., Sobral B. W. S. (1994) Phylogenetic analysis of organellar DNA sequences in the Andropogoneae: Saccharinae. Theor. Appl. Genet. 88: 933–944.

Bapat D. R., Mote U. N. (1982) Sources of shootfly resistance in Sorghum. J. Maharashtra Agric. Univ. 7: 238–240.

Bohs L., Olmstead R. G. (1997) Phylogenetic relationships in *Solanum* (Solanaceae) based on *ndh*F sequences. Syst. Bot. 22: 5–17.

Catalán P., Kellogg E. A., Olmstead R. G. (1997) Phylogeny of Poaceae subfamily Pooideae based on chloroplast *ndh*F gene sequences. Mol. Phylogenet. Evol. 8: 150–166.

Catalán P., Olmstead R. G. (2000) Phylogenetic reconstruction of the genus *Brachypodium* P. Beauv. (Poaceae) from combined sequences of chloroplast *ndh*F gene and nuclear ITS. Plant Syst. Evol. 220: 1–19.

Celarier R. P. (1958) Cytotaxonomy of the Andropogoneae III. Subtribe Sorgheae, genus *Sorghum*. Cytologia 23: 395–418.

Chittenden L. M., Schertz K. F., Lin Y.-R., Wing R. A., Paterson A. H. (1994) A detailed RFLP map of *Sorghum bicolor* x *S. propinquum*, suitable for high-density mapping, suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. Theor. Appl. Genet. 87: 925–933.

Clark L. G., Zhang W., Wendel J. F. (1995) A phylogeny of the grass family (Poaceae) based on *ndh*F sequence data. Syst. Bot. 20: 436–460.

DeWet J. M. J., Harlan J. R. (1971) The origin and domestication of *Sorghum bicolor*. Economic Botany 25: 128–135.

Dillon S. L., Lawrence P. K., Henry R. (2001) The use of ribosomal ITS to determine phylogenetic relationships within *Sorghum*. Plant Syst. Evol. 230: 97–110.

Doggett H. (1976) Sorghum: *Sorghum bicolor* (Gramineae-Andropogoneae). Evolution of

Crop Plants. N. W. Simmonds. London, Longman, pp. 112–117.

Duvall M. R., Doebley J. F. (1990) Restriction site variation in the chloroplast genome of *Sorghum* (Poaceae). Syst. Bot. 15: 472–480.

Farris J. S., Källersjö M., Kluge A. G., Bult C. (1995) Constructing a significance test for incongruance. Syst. Biol. 44: 570–572.

Ferguson D. M. (1999) Phylogenetic analysis and relationships in Hydrophyllaceae based on *ndh*F sequence data. Syst. Bot. 22: 253–268.

Franzmann B. A., Hardy A. T. (1996) Testing the host status of Australian indigenous sorghums for the sorghum midge. Third Australian Sorghum Conference, Tamworth, Australia, pp. 365–367.

Garber E. D. (1950) Cytotaxonomic studies in the genus *Sorghum*. University of California Publications in Botany. Berkeley, University of California Press. 23: 283–361.

Gu M.-H., Ma H.-T., Liang G. H. (1984) Karyotype analysis of seven species in the genus *Sorghum*. J. Hered. 75: 196–202.

Harlan J. R., de Wet J. M. J. (1971) Toward a rational classification of cultivated plants. Taxon 20: 509–517.

Kamala V., Singh S. D., Bramel P. J., Rao D. M. (2002) Sources of resistance to downy mildew in wild and weedy Sorghums. Crop Sci. 42: 1357–1360.

Karunakar R. I., Narayana Y. D., Pande S., Mughogho L. K., Singh S. D. (1994) Evaluation of wild and weedy sorghums for downy mildew resistance. International Sorghum and Millets Newsletter 35: 104–106.

Kim K.-J., Jansen R. K. (1995) *Ndh*F sequence evolution and the major clades in the sunflower family. Proc. Natl. Acad. Sci. USA 92: 10379–10383.

Komolong B., Chakraborty S., Ryley M., Yates D. (2002) Identity and genetic diversity of the sorghum ergot pathogen in Australia. Aust. J. Agric. Res. 53: 621–628.

Lazarides M., Hacker J. B., Andrew M. H. (1991) Taxonomy, cytology and ecology of indigenous Australian Sorghums (*Sorghum* Moench: Andropogoneae: Poaceae). Austral. Syst. Bot. 4: 591–635.

Magoon M. L., Shambulingappa K. G. (1961) Karyomorphology of *Sorghum propinquum* and its bearing on the origin of 40-chromosome *Sorghum*. Chromosoma 12: 460–465.

Neyland R., Urbatsch L. E. (1996) The *ndh*F chloroplast gene detected in all vascular plant divisions. Planta 200: 273–277.

Olmstead R. G., Sweere J. A. (1994) Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43: 467–481.

Paterson A. H., Schertz K. F., Lin Y. R., Liu S. C., Chang Y. L. (1995) The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. Proc. Natl. Acad. Sci. USA 92: 6127–6131.

Phillips S. (1995) Poaceae: *Sorghum purpureosericeum* description. Flora of Ethiopia and Eritrea Volume 7. I. Hedberg and S. Edwards. Addis Ababa. 7: 301.

Price H. J., Dillon S. L., Hodnett G., Rooney W. L., Ross L., Johnston J. S. (2004) Genome evolution in the genus *Sorghum* (Poaceae). Ann. Bot. (in press).

Sharma H. C., Franzmann B. A. (2001) Host-plant preference and oviposition responses of the sorghum midge, *Stenodiplosis sorghicola* (Coquillett) (Dipt., Cecidomyiidae) towards wild relatives of sorghum. J. Appl. Ent. 125: 109–114.

Spangler R., Zaitchik B., Russo E., Kellogg E. (1999) Andropogoneae evolution and generic limits in *Sorghum* (Poaceae) using *ndh*F sequences. Syst. Bot. 24: 267–281.

Spangler R. E. (2003) Taxonomy of Sarga, Sorghum and Vacoparis (Poaceae: Andropogoneae). Austral. Syst. Bot. 16: 279–299.

Stenhouse J. W., Prasada Rao K. E., Gopal Reddy V., Appa Rao S. (1997) Sorghum. In: Fuccillo D., Sears L., Stapleton P. (eds.) Biodiversity in trust. Cambridge University Press, Cambridge, pp. 292–308.

Sun Y., Skinner D. Z., Liang G. H., Hulbert S. H. (1994) Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. 89: 26–32.

Swofford D. L. (2002) PAUP. Phylogenetic Analysis Using Parsimony (and other methods), Sinauer Associates, Sunderland Massachusetts.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., Higgins D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by

quality analysis tools. Nucl. Acids Res. 25: 4876–4882.

White T. J., Bruns T., Lee S., Taylor J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M. A., Gelfand D. H., Sninsky J. J., White T. J. (eds.) PCR Protocols: A guide to methods and applications. Academic Press, San Diego, pp. 315–322.

Wu T.-P. (1979) Cytological and morphological studies on *Sorghum roxburghii, S. propinquum* and their F1 hybrid. Proc. Natl. Sci. Counc. ROC. 3: 291–298.

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

Addresses of the authors: Sally L. Dillon (e-mail: sally.dillon@dpi.qld.gov.au), Peter K. Lawrence, Australian Tropical Crops and Forages Collection, Queensland Department of Primary Industries and Fisheries, Biloela, QLD, Australia). Robert J. Henry, Centre for Plant Conservation Genetics, Southern Cross University, Lismore, NSW, Australia, Larry Ross, 4060 Preferred Place #720, Dallas, TX 75237, USA. H. James Price, Department of Soil & Crop Sciences, Texas Agricultural Experiment Station, Texas A&M University, College Station TX 77843-2474, USA. J. Spencer Johnston, Department of Entomology, Texas Agricultural Experiment Station, Texas A&M University, College Station TX 77843-2475, USA.