ISSN 1479-2621
Plant Genetic Resources 3(1); 19–28
DOI: 10.1079/PGR200454

The new use of *Sorghum bicolor*-derived SSR markers to evaluate genetic diversity in 17 Australian *Sorghum* species

Sally L. Dillon[1]*, Peter K. Lawrence[1] and Robert J. Henry[2]

[1]*Australian Tropical Crops and Forages Centre, Queensland Department of Primary Industries and Fisheries, LMB 1, Biloela, QLD 4715, Australia and* [2]*Grain Foods CRC Limited, Centre for Plant Conservation Genetics, Southern Cross University, PO Box 157, Lismore, NSW 2480, Australia*

Received 29 June 2004; Accepted 6 October 2004

Abstract

The *Sorghum* genus is extremely diverse both morphologically and geographically, however, relatively few of the 25 recognized species have been evaluated genetically. The apparent lack of basic knowledge pertaining to the levels of genetic diversity both within and between the 17 Australian wild species is a major obstacle to both their effective conservation and potential use in breeding programmes. Twelve *Sorghum bicolor*-derived simple sequence repeat (SSR) markers were evaluated for cross-species amplification in all 25 *Sorghum* species. The SSR markers were highly polymorphic, with diversity indices ranging from 0.59 to 0.99 with mean of 0.91. Five markers combined were able to differentiate 24 of the 25 *Sorghum* species, with intra-species polymorphism apparent. *Sorghum bicolor*-derived SSRs have proven to be an efficient source of markers for genetic diversity studies of the relatively poorly characterized Australian indigenous *Sorghum* species.

Keywords: cross-species amplification; germplasm conservation; microsatellites; sorghum breeding

Introduction

Australian indigenous *Sorghum* species comprise 17 of the world's 25 recognized *Sorghum* species, with the genus separated into five subgeneric sections based on taxonomic differences: *Eu-sorghum*, *Chaetosorghum*, *Heterosorghum*, *Para-sorghum* and *Stiposorghum* (Table 1) (de Wet, 1978; Lazarides *et al*., 1991; USDA, ARS, 2004). *Sorghum bicolor*, *S. arundinaceum*, *S.* × *drummondii*, *S. halepense*, *S. propinquum* and *S.* × *almum* form section *Eu-sorghum*, the primary and secondary gene pools of *Sorghum*. They originate from Africa, Asia and South America, consist of the cultivated species, their progenitors, and some serious weed pests, with their close genetic relationships and inter-crossabilities well known (de Wet and Harlan, 1971; Doggett, 1976; Duvall and Doebley, 1990; Chittenden *et al*., 1994; Sun *et al*., 1994; Spangler *et al*., 1999; Dillon *et al*., 2001, 2004). All species outside of section *Eu-sorghum* form the tertiary gene pool.

Section *Chaetosorghum* contains the single species *S. macrospermum* that is endemic to a small area in the Northern Territory of Australia. It is closely related to *S. laxiflorum* (section *Heterosorghum*), a geographically more diverse species indigenous to Australia and Papua New Guinea (Garber, 1950; Lazarides *et al*., 1991; Dillon *et al*., 2001, 2004). Section *Para-sorghum* contains the *seven* species *S. grande*, *S. leiocladum*, *S. matarankense*, *S. nitidum*, *S. purpureo-sericeum*, *S. timorense* and *S. versicolor* which are indigenous to Australia, Africa and Asia (Garber, 1950; Lazarides *et al*., 1991; Phillips, 1995). *Sorghum amplum*, *S. angustum*, *S. brachypodum*, *S. bulbosum*, *S. ecarinatum*, *S. exstans*,

* Corresponding author. E-mail: sally.dillon@dpi.qld.gov.au

S. interjectum, *S. intrans*, *S. plumosum* and *S. stipoideum* are the Australian endemic species that form section *Stiposorghum* (Garber, 1950; Lazarides *et al.*, 1991).

Sorghum is the world's fifth most important cereal crop producing more than 100,000 million tonnes annually from 39 countries on six continents (Smith and Frederiksen, 2000). *Sorghum bicolor* has been selected over hundreds of years to produce the short, high-yielding varieties grown today, and this has led to a significant reduction in the genetic diversity within this cultivated species (Tao *et al.*, 1993; Cui *et al.*, 1995; Ahnert *et al.*, 1996; Jordan *et al.*, 1998). The wild relatives and progenitor species of this important crop represent a broad genetic base rich in traits for pest/disease resistances, tolerance for abiotic stresses, as well as quality and yield traits.

The Australian indigenous *Sorghum* species are extremely diverse both morphologically and geographically. They predominate in monsoonal northern Australia, and form a major component of understorey vegetation over extensive areas in higher rainfall tropical areas. These regions appear to be the centre of diversity for the Australian species, with the widespread polyploid species *S. nitidum* and *S. plumosum* considered parent species from which the localized perennial *S. leiocladum* and the annual species were derived (Lazarides *et al.*, 1991). Over the past decade, the Australian Tropical Crops and Forages Collection team (www.dpi.qld.gov.au/auspgris/) have conducted seed-collecting missions throughout northern Australia to conserve the wide range of genetic diversity existing both within and between these indigenous *Sorghum* species. To date, 378 seed collections have been made representing the 17 Australian species. A small number of Australian *Sorghum* species have been evaluated in morphological and cytological studies (Garber, 1950; Lazarides *et al.*, 1991; Huelgas *et al.*, 1996; Price *et al.*, 2004). Pest and disease evaluations have identified Australian *Sorghum* species with resistance to diseases such as ergot and downy mildew, and pests such as shootfly, spotted stem borer, with many of the Australian *Sorghum* species also being non-hosts to the sorghum midge (Bapat and Mote, 1982; Karunakar *et al.*, 1994; Franzmann and Hardy, 1996; Sharma and Franzmann, 2001; Kamala *et al.*, 2002; Komolong *et al.*, 2002; Kameswara Rao *et al.*, 2003). More recently, some of these accessions have been used to determine genetic relationships via DNA gene sequencing (Sun *et al.*, 1994; Spangler *et al.*, 1999; Dillon *et al.*, 2001, 2004). The value of these collections would be more readily recognized if a comprehensive assessment of their genetic diversity were conducted (both within and between species). Enormous potential exists for widening the genetic base of sorghum, however, these Australian indigenous species still remain an untapped source.

Microsatellite, or simple sequence repeat (SSR), markers occur throughout the plant genome and are highly polymorphic among closely related cultivars and species due to mutations causing variation in the number of repeat units (Morgante and Olivieri, 1993; Saghai-Maroof *et al.*, 1994; Brown *et al.*, 1996). High rates of polymorphism ensure that SSR markers are efficient for germplasm analysis and in marker-assisted breeding programmes. However, the development of SSRs for new species remains a costly and time-consuming process, especially if the target species have little or no commercial value, as is the case for the Australian indigenous *Sorghum* species. An alternative approach that has been successful in many plant groups is to use existing SSRs from one species to amplify alleles in related taxa (Peakall *et al.*, 1998; Hernández *et al.*, 2001; Rossetto, 2001; Chen *et al.*, 2002; Clauss *et al.*, 2002; Scott *et al.*, 2003; González-Martínez *et al.*, 2004; Rohrer *et al.*, 2004; Sudupak, 2004). This is possible due to conservation of the repeat sequence and flanking regions (containing the primer binding sites) across plant genomes. However, the success of cross-species PCR amplification will depend on the evolutionary distance between the source and target species (Westman and Kresovich, 1998; Rossetto, 2001).

A significant effort has been made to identify SSRs in the cultivated *S. bicolor*, with more than 150 markers now available (Brown *et al.*, 1996; Taramino *et al.*, 1997; Kong *et al.*, 2000; Schloss *et al.*, 2002). Only the work by Schloss *et al.* (2002) has evaluated these SSRs in species other than the cultivated *S. bicolor*, but to date, there has not been an SSR cross-species evaluation outside of the *Eu-sorghum* species. In order to determine whether these SSRs are valuable markers in Australian indigenous *Sorghum* species, 19 markers derived from *S. bicolor* and one from *Zea mays* (L.) were screened across the 25 *Sorghum* species to evaluate the level of cross-species amplification, and to estimate the level of genetic diversity across the genus.

Materials and methods

Plant materials and DNA extraction

Forty-four accessions from 25 *Sorghum* species were analysed in this study, with Table 1 showing the species, chromosome numbers, accession numbers and countries of origin for each. All *Sorghum* seed was obtained from the Australian Tropical Crops and Forages Collection, Queensland Department of Primary Industries and Fisheries, Biloela, Australia (www.dpi.qld.gov.au/auspgris/). Genomic DNA was extracted from fresh leaf tissue of five individuals per accession using the CTAB method

Table 1. Taxonomic section, species, ploidy, accession number and country of origin of the 44 *Sorghum* accessions

Species and subgeneric section	Ploidy[a] $2n =$	AusTRCF number	Country of origin	Herbarium voucher[b]
Section *Eu-sorghum*				
S. bicolor (L.) Moench	20	312827	Mozambique	Macia
		314746	USA	Tx 623A
S. arundinaceum (Desv.) Stapf	20	PI 302118	Ethiopia	PI 302118
S. drummondii (Steud.) Millsp. & Chase	20	PI 330272	Ethiopia	PI 330272
		PI 255739	Turkey	PI 255739
S. halepense (L.) Pers.	40	300167	Australia	BRI AQ773626
		300188	Australia	BRI AQ773627
S. propinquum (Kunth) Hitchc.	20	302546	USA	BRI AQ773674
S. × *almum* Parodi	40	302386	Ethiopia	BRI AQ773678
		302387	Unknown	BRI AQ773628
Section *Chaetosorghum*				
S. macrospermum E. D. Garber	40	302367	Australia	DNA C867
Section *Heterosorghum*				
S. laxiflorum F. M. Bailey	40	302503	Australia	BRI AQ773670
		302510	Australia	BRI AQ773635
Section *Para-sorghum*				
S. grande Lazarides	30, 40	302580	Australia	BRI AQ585960
S. leiocladum (Hack.) C. E. Hubb.	10[c], 20	300170	Australia	DNA D0155521
		300180	Australia	DNA D0155683
S. matarankense E. D. Garber & Snyder	10	302517	Australia	BRI AQ773676
		302521	Australia	BRI AQ773673
S. nitidum (Vahl) Pers.	10, 20	302539	Australia	CANB 479893
		302543	Australia	CANB 479881
S. purpureo-sericeum (Hochst. ex. A. Rich.) Asch. & Schweinf.	10[c]	IS 18945	Sudan	IS 18945
S. timorense (Kunth) Buse	10, 20	302532	Australia	BRI AQ773672
		302660	Australia	DNA D129474
S. versicolor Andersson	10[c]	IS 14262	Angola	IS 14262
Section *Stiposorghum*				
S. amplum Lazarides	10, 30[c]	302623	Australia	DNA D129461
		302455	Australia	CANB 480260
S. angustum S. T. Blake	10	302596	Australia	BRI AQ585973
		302604	Australia	BRI AQ585980
S. brachypodum Lazarides	10	302670	Australia	DNA D133019
		302480	Australia	CANB 480297
S. bulbosum Lazarides	10	302417	Australia	BRI AQ773630
		302646	Australia	DNA D129481
S. ecarinatum Lazarides	10	302648	Australia	DNA D129449
		302661	Australia	DNA D129486
S. exstans Lazarides	10	302401	Australia	CANB 479848
		302473	Australia	CANB 479831
S. interjectum Lazarides	30, 40	302563	Australia	BRI AQ585985
S. intrans F. Muell. ex Benth.	10	302390	Australia	BRI AQ773629
		302668	Australia	DNA D133021
S. plumosum (R. Br.) P. Beauv.	10, 20, 30	302462	Australia	BRI AQ778820
		302489	Australia	BRI AQ 773634
		302533	Australia	CANB 479828
S. stipoideum (Ewart & Jean White) C. A. Gardner & C. E. Hubb.	10	302625	Australia	DNA D129494
		302644	Australia	DNA D129466

[a] Lazarides *et al.* (1991) and Sun *et al.* (1994).
[b] 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.
[c] Price *et al.* (2004).

of Dillon *et al.* (2001). DNA from the five individuals was pooled to form a bulk sample, and DNA concentrations for each bulk sample were measured with a Perkin Elmer Lambda BIO10 UV/VIS Spectrometer and adjusted to a concentration of ~20 ng/μl.

Primer optimization in Australian Sorghum species

Nineteen SSR markers developed for *S. bicolor* and one for *Z. mays* were selected from the literature (Brown *et al.*, 1996; Taramino *et al.*, 1997; Kong *et al.*, 2000). Primer pairs for each marker were synthesized by Invitrogen Life Technologies, Brisbane, QLD, Australia, with their characteristics and expected allele sizes in *S. bicolor* shown in Table 2.

The 20 SSR primer pairs were optimized for annealing temperature (T_{ann}) and magnesium chloride ($MgCl_2$) concentration in *S. bicolor* (as a positive control), *S. exstans* and *S. laxiflorum* to ensure optimal primer performance across species. Touchdown PCR resulted in significant levels of non-specific amplification, and was not used further in this study. Optimal PCR amplification across the three species was achieved using $T_{ann} = 55$°C for all primer pairs with a published annealing temperature between 50 and 55°C, and $T_{ann} = 65$°C for all primers with published annealing between 60 and 65°C (Table 2). The 1.5 mM $MgCl_2$ present in the *Taq* buffer resulted in the most stringent amplification of SSR products across the three test species. Following this general optimization, three primers failed to amplify product, while five continued to amplify non-specific product in *S. bicolor* and were excluded from further analysis.

PCR amplification and PCR product visualization

Amplification was carried out in a total volume of 10 μl in a Perkin Elmer GENEAMP 9700 PCR thermocycler. Each reaction contained 20–40 ng genomic DNA, 200 μM equimolar dNTPs, 1.0 U *Taq* DNA polymerase (Roche Diagnostics), *Taq* buffer (to 10 mM Tris-HCl, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$) and 0.25 μM of each primer. Cycling conditions included an initial denaturation step of 5 min at 94°C followed by 30 cycles of 30 s at 94°C, 30 s at the optimized T_{ann} for each primer pair, and 1 min extension at 72°C. This was followed by a 10 min extension phase at 72°C. Reactions lacking DNA were included as negative controls, and amplifications were confirmed in replicate reactions. PCR products were visualized on 2.5% agarose

Table 2. Microsatellite locus, repeat characteristics, published annealing temperature and expected size ranges in *Sorghum bicolor* (the diversity index and the observed size range across all species are also listed)

Locus	Repeat unit	Annealing temp. (°C)	Primer GC content (%)	Expected size[a] (bp)	Observed size (bp)	Diversity index (*D*)
Xtxp3[b]	(CT)8 + (CT)36	50	50.0	232	–[e]	–
Xtxp4	(GA)23	55	48.0	173	–	–
Xtxp6	(CT)33	50	52.5	120	–	–
Xtxp8	(TG)31	60	48.5	148	123–210	0.98
Xtxp24	(TC)21	60	50.0	145	113–152	0.98
Xtxp25	(CT)12	55	47.5	139	124–204	0.94
Xtxp30	(AAT)25	60	55.5	273	236–312	0.98
Xtxp33	(TC)20C(TG)5 + (CT)9CC(TG)7	55	50.5	221	–	–
Xtxp37	(TC)23	55	43.5	189	–	–
Xtxp43	(CT)28	60	50.0	171	129–170	0.97
SbAGAO1[c]	(AG)33	54	35.0	88–116	63–107	0.96
SbAGDO2	(AG)32	54	53.0	134	–	–
SbAGEO1	(AG)30	54	47.5	208–240	191–244	0.99
SbAGFO8	(AG)34	54	54.5	134–176	96–166	0.59
SbAGGO2	(AG)41	54	41.0	190	–	–
SbAGHO4	(AG)39	54	53.0	110–170	99–148	0.97
SvPEPCAA	(AT)10	54	51.5	206–250	184–244	0.80
SBKAFGKI[d]	(AAC)9	60	50.0	280–320	249–410	0.87
ZMADH2N	(AG)7	60	46.0	110–120	–	–
Sb1-10	(AG)27	65	57.5	350–400	251–490	0.95
Mean						0.91
SD						0.11

[a] In *Sorghum bicolor* from Brown *et al.* (1996), Taramino *et al.* (1997) and Kong *et al.* (2000).
[b] All markers starting with 'Xtxp' sourced from Kong *et al.* (2000).
[c] All markers starting with 'SbAG', and including SvPEPCAA, sourced from Taramino *et al.* (1997).
[d] SBKAFGK1, ZMADH2N and Sb1-10 sourced from Brown *et al.* (1996).
[e] A dash shows markers that were not used across species.

gels with ethidium bromide to determine whether SSR alleles were amplified in the wild species.

Microsatellite allele sizing and verification

SSR alleles were separated in a 5 or 6% polyacrylamide gel stained with ethidium bromide using a Corbett GS2000. Products were run at 1200 V for between 30 and 60 min with a molecular marker size standard, and their relative sizes determined using Gene Profiler (Scanalytics Inc.). Allele sizes were compared across species and the level of polymorphism of each marker was calculated using the diversity index (D) with $D = 1 - \Sigma p_i^2$, where p_i equals the frequency of the ith microsatellite allele (Brown *et al.*, 1996).

Amplification in the wild species was tested for all loci using the conditions optimized in *S. bicolor*, *S. exstans* and *S. laxiflorum*. Selected loci that produced robust alleles were purified using a Qiagen PCR purification column and direct sequenced in both directions to determine whether the amplicons contained the expected SSR repeat. Sequences were obtained using BigDye Terminator chemistry and analysed on ABI 377 automatic sequencers by AGRF (Brisbane, Australia).

Results

Twelve SSR loci produced clear bands within the expected size range in *S. bicolor* and amplified robust alleles in *S. exstans* and *S. laxiflorum*. These 12 SSRs were then screened across the remaining *Sorghum* accessions. Four SSRs amplified alleles in at least 15 *Sorghum* species, with the remaining eight SSRs able to amplify alleles in five to nine species. Eight of the 12 SSRs amplified alleles within 50 bp of the expected size range, with the remaining four SSR markers amplifying alleles within 100 bp of the expected size range (Tables 2 and 3). Direct sequencing of amplified alleles confirmed SSR repeat content for all 12 markers in *S. bicolor*. Sequencing of selected alleles in other species showed the expected SSR repeat motif in alleles for markers: Xtxp8 in *S. halepense*; Xtxp43 in *S.* × *almum*; SbAGAO1 in *S. timorense* and *S.* × *drummondii*; SvPEPCAA in *S. stipoideum*, *S. nitidum*, *S. versicolor* and *S. propinquum*; SBKAFGK1 in *S. timorense*; and Sb1-10 in *S.* × *drummondii*. SbAGEO1 consistently amplified alleles around the 650 bp size in most of the *Para-sorghum* and *Stiposorghum* species, however, these alleles did not contain the $(AG)_n$ microsatellite repeat and were not used in any analysis. Other than SbAGEO1, all other sequencing reactions confirmed microsatellite repeat motifs. The allele sizes for each SSR amplified in each accession are available from the authors upon request.

The total number of polymorphic alleles amplified for each locus over all accessions ranged from six for Xtxp30 to 23 for SBKAFGK1, with a mean of 11.8 per SSR locus (Table 3). The level of polymorphism (diversity index) for each SSR locus across the 25 species ranged from 0.59 to 0.99 with a mean of 0.91 (Table 3). Eleven of the 12 SSRs used in this study had diversity index values between 0.80 and 0.99 (Table 2). A combination of the five loci, Xtxp8, Xtxp25, SbAGFO8, SvPEPCAA and SBKAFGKI, was able to differentiate between 24 of the 25 *Sorghum* species. *Sorghum purpureo-sericeum* did not amplify any SSR alleles using the general optimization used in this study. The SSR SbAGFO8 had a lower diversity index of 0.59, as it was monomorphic among most of the *Para-sorghum* and *Stiposorghum* species.

When comparing allele amplification across subgeneric sections of *Sorghum*, alleles within ~50 bp of the expected size range (in *S. bicolor*) were amplified in the six *Eu-sorghum* species (Tables 2 and 3). Seven SSR markers were amplified in sections *Para-sorghum* and *Stiposorghum*, two SSR markers were amplified in *Chaetosorghum*, and four in *Heterosorghum*. Most alleles were within ~65 bp of the expected size in *S. bicolor* (Tables 2 and 3).

Each accession we analysed in this study was the bulk DNA of five plants, leading to potential multiple alleles for each accession. There were one to four alleles amplified per marker per species indicating allelic diversity within some species (between accessions) was observed for the 12 SSR markers. This indicates that population diversity within species may exist, or that some of the Australian native *Sorghum* species may be heterozygotes.

Discussion

The success of cross-species amplification depends on the evolutionary distance between the target species (Rossetto, 2001), so it was expected that the *S. bicolor*-derived SSRs would transfer preferentially to section *Eu-sorghum* species as they are closely related (Sun *et al.*, 1994; Spangler *et al.*, 1999; Dillon *et al.*, 2001, 2004). As the six *Eu-sorghum* species are so close, they have fewer mutations (greater conservation) in the flanking regions containing the primer binding sites within their genomes, which is reflected by the SSR loci amplifying alleles in at least four of these species with seven of the markers amplified in all six *Eu-sorghum* species.

The *Chaetosorghum* and *Heterosorghum* species are the Australian species most closely related to the cultivated *S. bicolor* (Dillon *et al.*, 2001, 2004), so it would be expected that they too would amplify a larger number of the SSR markers than the *Para-sorghum*

Table 3. Number of species, number of polymorphic alleles and size ranges of alleles amplified within each taxonomic subsection

Taxonomic subsection	Total no. of species	Microsatellite marker												
		Xtxp8	Xtxp24	Xtxp25	Xtxp30	Xtxp43	SbAGAO1	SbAGEO1	SbAGFO8	SbAGHO4	SvPEPCAA	SBKAFGK1	Sb1-10	Mean
Eu-sorghum	6	4 / 7[a]	6 / 7	6 / 9	5 / 6	6 / 8	6 / 8	5 / 8	6 / 9	5 / 7	5 / 8	6 / 13	6 / 12	
		123–210	113–152	124–200	236–312	133–170	68–107	191–244	111–166	106–148	188–220	249–384	251–490	
Chaetosorghum/ Heterosorghum	2	0	0	0	0	1 / 1	2 / 2	0	0	0	2 / 3	1 / 3	0	
		–	–	–	–	138	63–65	–	–	–	186–218	274–406	–	
Para-sorghum	7	0	0	2 / 4	0	1 / 1	1 / 1	0	5 / 4	0	5 / 7	3 / 5	2 / 2	
		–	–	200–204	–	140	63	–	96–102	–	206–244	274–406	274–280	
Stiposorghum	10	1 / 2	0	7 / 7	0	0	0	0	10 / 4	1 / 1	10 / 12	9 / 8	1 / 1	
		142–146	–	161–204	–	–	–	–	96–102	99	194–227	272–410	258	
Total	25	5 / 8	6 / 7	15 / 17	5 / 6	8 / 10	9 / 10	5 / 8	21 / 13	6 / 8	22 / 19	19 / 23	9 / 13	11.8
Diversity index (*D*)		0.98	0.98	0.94	0.98	0.97	0.96	0.99	0.59	0.97	0.80	0.87	0.95	0.91

[a] Within each cell the data are listed as: species/polymorphic loci, with allele size range below. Some alleles are shared across subsections and the values shown in the total polymorphic loci are therefore less than the combined total from all subsections for some markers.

and *Stiposorghum* species. However, only four of the 12 SSRs were amplified using our general PCR optimization. Cross-species transfer to the *Para-sorghum* species was more successful with seven SSRs consistently amplified, although no SSR was amplified by all species in the section. The marker SvPEPCAA was highly polymorphic and could differentiate five of the seven *Para-sorghum* species (Table 3). SSR transfer to section *Stiposorghum* was higher, with seven markers again consistently amplified, with the two loci SbAGFO8 and SvPEPCAA amplified in all 10 species (Table 3). The higher rate of transfer to the *Stiposorghum* species' could be associated with the species' more advanced features and hence genomes when compared with the *Para-sorghum* species (Lazarides *et al.*, 1991).

SSR allele size variation can be caused by slip strand mispairing during DNA replication, resulting in mutations in multiples of the core microsatellite repeat unit (Levinson and Gutman, 1987). Under this stepwise distribution, size differences between alleles are a multiple of the SSR core repeat unit. Two-thirds of the polymorphic loci amplified by our SSR markers fit this stepwise size distribution, however, 33.1% of our polymorphic loci showed non-stepwise variation in allele size. Much of the length variation in SSR alleles in maize is due to insertions/deletions (indels) in the flanking region of the core repeat unit rather than mutations in the repeat number (Matsuoka *et al.*, 2002). Alleles amplified with large differences between expected size (in source species) and observed size have been shown to contain large indels within the flanking region of the repeat unit, and are not due to the amplification of duplicated loci (Matsuoka *et al.*, 2002). The alleles we sequenced in some of the wild species to confirm SSR repeat showed variation in repeat number, with indels also present in some of the flanking regions. As a relatively small number of these alleles were sequenced, they can only give an indication that the non-stepwise size allele distribution, and the amplification of our larger allele sizes in the wild *Sorghum* species, could be due to variation in SSR repeat number and indels in the flanking regions, rather than the amplification of duplicated loci.

Our data suggest that the level of polymorphism (diversity) shown by these markers over all *Sorghum* species is higher than the levels previously described within *S. bicolor*. Brown *et al.* (1996) developed SSRs with a mean diversity index of 0.56, and Taramino *et al.* (1997) developed SSRs with a mean diversity index of 0.80 when screened across a small number of *S. bicolor* accessions originating from Africa, Asia and the USA. The SSRs developed by Kong *et al.* (2000) were screened across 208 accessions representing 16 races and 12 subraces of *S. bicolor* with a mean diversity index of 0.89. A small number of these available microsatellites were used by Dean *et al.* (1999), Djè *et al.* (1999, 2000), Grenier *et al.* (2000) and Smith *et al.* (2000) in *S. bicolor*, with diversity index values ranging from 0.14 to 0.93, however, none of these available SSRs has been used in wild *Sorghum* species.

The higher diversity index found using a subset of these SSR markers across the 25 *Sorghum* species is not unexpected, as the wild relatives of crop species are generally considered more diverse (McLauchlan *et al.*, 2001). In *Sorghum*, the section *Eu-sorghum* species are more genetically diverse than the cultivated *S. bicolor* (using restriction fragment length polymorphisms (RFLPs) and allozymes) (Morden *et al.*, 1990; Aldrich and Doebley, 1992; Cui *et al.*, 1995). A majority of the Australian *Sorghum* species are largely outcrossing due to high levels of self-incompatibility (Lazarides *et al.*, 1991). They would therefore have higher levels of genetic diversity than the cultivated *S. bicolor*, which is reflected by the amplification of more polymorphic, and also multiple alleles, in our heterogeneous DNA samples. However, levels of genetic diversity have been overestimated in other plant taxa when pooled DNA samples were used due to competition between primer binding sites within the PCR reaction (Halldén *et al.*, 1996). Multiple alleles have also been amplified from individual DNA samples in the Australian indigenous species using the same SSRs, however, the frequencies of each amplified allele are yet to be determined, so it is possible that the level of genetic diversity found across the Australian indigenous species could be overestimated due to the use of pooled DNA samples. The genome specificity of microsatellite markers has previously been shown in wheat (Bryan *et al.*, 1997; Röder *et al.*, 1998; Stephenson *et al.*, 1998; Harker *et al.*, 2001), however, as it is unknown how many genomes are present within the *Sorghum* genus, no conclusions can be drawn regarding the genome specificity of these markers in the Australian native *Sorghum* species.

The repeat type of SSRs can affect cross-species amplification. SSRs with AG repeats are more abundant and widely dispersed throughout the *Sorghum* genome than other repeat types (Taramino *et al.*, 1997; Kong *et al.*, 2000). In our study, although SSRs with AG repeats were amplified in more species, they were more monomorphic and therefore less informative. Repeat length can also affect cross-species amplification, with longer repeats experiencing potential repeat interruptions resulting in failed amplification across species (Weber, 1990; Taramino *et al.*, 1997; Rossetto, 2001). We found no relationship between repeat length and degree of polymorphism using this set of *S. bicolor* SSRs across all *Sorghum* species. The GC content of SSR primer pairs can also affect cross-species amplification. Primers with higher GC contents (~50%) are more likely to amplify

across species using original/existing PCR protocols, while primer pairs with lower GC content (~30%) require greater modification of protocols for successful amplification to occur (Dayanandan *et al.*, 1997). The primer pairs tested across *Sorghum* ranged in GC content from 35 to 57.5% (Table 2), with no direct relationship between primer GC content and improved amplification evident in *Sorghum*.

As mentioned previously, the optimized annealing temperatures (T_{ann}) used in our PCR conditions were up to 5°C higher than the previously published T_{ann} for the *S. bicolor* primers. This increase in T_{ann} is in contrast to previous reports where a 2–5°C decrease was found to improve cross-species amplification as the evolutionary distance between target species increased (Rossetto, 2001, for review). Decreases in T_{ann} (lowered primer binding stringency) compensates for potential mutations between the primers and the primer-binding site in the target species. As we were able to achieve successful cross-species amplification to the Australian *Sorghum* species with increased T_{ann}, we can suggest that within *Sorghum* the primer binding sites are relatively strongly conserved compared to other plant taxa.

The genomes of related taxa may differ by rearrangements, mutations and duplications, and therefore vary for the presence, copy number, sequence repeat and flanking region for SSR markers. Amplification across taxa may therefore generate both false positives (where products contain no repeats) and false negatives (where repeats are actually present, but there is no amplification) (Fisher *et al.*, 1998; Westman and Kresovich, 1998; Devos *et al.*, 2000). It is quite possible that there have been base changes near the 3′ end of the corresponding primer annealing sites that result in no primer annealing and no allele amplification within some of the wild *Sorghum* genomes where SSR repeats may actually be present (null alleles). Equally, some of the alleles amplified in the wild *Sorghum* species that have not yet been sequenced may be false positives. Further sequencing of amplified SSR alleles across the *Sorghum* genus will identify any false positives, however, sequencing has shown that alleles of the same size in other wild *Sorghum* species do contain the expected repeat unit, and it is therefore reasonable to assume that alleles in other species contain the expected microsatellite repeat. Within-species optimization of specific SSR markers would improve the amplification of existing null alleles in the wild *Sorghum* species.

In conclusion, the relatively high transfer rate of *S. bicolor*-derived SSRs to the wild species and their high level of diversity suggests that these SSRs are an efficient, highly informative source of molecular markers in the Australian *Sorghum* species. Most of the SSR markers were polymorphic across species where amplification occurred, and the multiple alleles amplified between populations within species indicate that heterozygosity and population diversity does exist. These SSR markers can now be used as tools for population and quantitative genetic studies in the Australian *Sorghum* species.

Acknowledgements

The Queensland Department of Primary Industries and Fisheries, the Australian Grains Research Development Corporation Project DAQ00015, and the Centre for Plant Conservation Genetics provided funding for this research. The authors would like to thank Ms Julie Edwards for her assistance with the many DNA extractions, Dr Paul Grundy, Tonia Grundy and the two anonymous reviewers for their comments on the draft manuscript.

References

Ahnert D, Lee M, Austin DF, Livini C, Woodman WL, Openshaw SJ, Smith JSC, Porter K and Dalton G (1996) Genetic diversity among elite sorghum inbred lines assessed with DNA markers and pedigree information. *Crop Science* 36: 1385–1392.

Aldrich PR and Doebley J (1992) Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild *Sorghum bicolor*. *Theoretical and Applied Genetics* 85: 293–302.

Bapat DR and Mote UN (1982) Sources of shootfly resistance in Sorghum. *Journal of the Maharashtra Agricultural University* 7: 238–240.

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F and Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics* 93: 190–198.

Bryan GJ, Collins AJ, Stephenson P, Orry A, Smith JB and Gale MD (1997) Isolation and characterisation of microsatellites from hexaploid bread wheat. *Theoretical and Applied Genetics* 94: 557–563.

Chen X, Cho YG and McCouch SR (2002) Sequence divergence of rice microsatellites in *Oryza* and other plant species. *Molecular Genetics and Genomics* 268: 331–343.

Chittenden LM, Schertz KF, Lin YR, Wing RA and 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. *Theoretical and Applied Genetics* 87: 925–933.

Clauss MJ, Cobban H and Mitchell-Olds T (2002) Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaeae). *Molecular Ecology* 11: 591–601.

Cui YX, Wu GW, Magill CW, Schertz KF and Hart GA (1995) RFLP-based assay of *Sorghum bicolor* (L.) Moench genetic diversity. *Theoretical and Applied Genetics* 90: 787–796.

Dayanandan S, Bawa KS and Kesseli R (1997) Conservation of microsatellites among tropical trees (*Leguminosae*). *American Journal of Botany* 84: 1658–1663.

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE and Kresovich S (1999) Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. *Crop Science* 39: 1215–1221.

Devos KM, Pittaway TS, Reynolds A and Gale MD (2000) Comparative mapping reveals a complex relationship between the pearl millet genome and those of foxtail millet and rice. *Theoretical and Applied Genetics* 100: 190–198.

de Wet JMJ (1978) Systematics and evolution of *Sorghum* section Sorghum (Gramineae). *American Journal of Botany* 65: 477–484.

de Wet JMJ and Harlan JR (1971) The origin and domestication of *Sorghum bicolor*. *Economic Botany* 25: 128–135.

Dillon SL, Lawrence PK and Henry RJ (2001) The use of ribosomal ITS to determine phylogenetic relationships within *Sorghum*. *Plant Systematics and Evolution* 230: 97–110.

Dillon SL, Lawrence PK, Henry RJ, Ross L, Price HJ and Johnston JS (2004) *Sorghum laxiflorum* and *S. macrospermum*, the Australian species most closely related to *S. bicolor*. *Plant Systematics and Evolution* 249: 233–246.

Djè Y, Forcioli D, Ater M, Lefèbvre C and Vekemans X (1999) Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. *Theoretical and Applied Genetics* 99: 157–163.

Djè Y, Heuertz M, Lefèbvre C and Vekemans X (2000) Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. *Theoretical and Applied Genetics* 100: 918–925.

Doggett H (1976) Sorghum: *Sorghum bicolor* (Gramineae-Andropogoneae). In: Simmonds NW (ed.) *Evolution of Crop Plants*. London: Longman, pp. 112–117.

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

Fisher PJ, Richardson TE and Gardner RC (1998) Characteristics of single- and multi-copy microsatellites in *Pinus radiata*. *Theoretical and Applied Genetics* 96: 969–979.

Franzmann BA and Hardy AT (1996) Testing the host status of Australian indigenous sorghums for the sorghum midge. In: *Proceedings of the Third Australian Sorghum Conference*. Occasional Publication 93. Melbourne: Australian Institute of Agricultural Sciences, pp. 365–367.

Garber ED (1950) Cytotaxonomic studies in the genus Sorghum. *University of California Publications in Botany* 23: 283–361.

González-Martínez SC, Robledo-Arnuncio JJ, Collada C, Díaz A, Williams CG, Alía R and Cervera MT (2004) Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines. *Theoretical and Applied Genetics* 109: 103–111.

Grenier C, Deu M, Kresovich S, Bramel-Cox PJ and Hamon P (2000) Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs non-random sampling procedures B. Using molecular markers. *Theoretical and Applied Genetics* 101: 197–202.

Halldén C, Hansen M, Nilsson NO, Hjerdin A and Säll T (1996) Competition as a source of errors in RAPD analysis. *Theoretical and Applied Genetics* 93: 1185–1192.

Harker N, Rampling L, Shariflou MR, Hayden M, Holton TA, Morell M, Sharp P, Henry RJ and Edwards KJ (2001) Microsatellites as markers for Australian wheat improvement. *Australian Journal of Agricultural Research* 52: 1121–1130.

Hernández P, Dorado G, Laurie DA, Martin A and Snape JW (2001) Microsatellites and RFLP probes from maize are efficient sources of molecular markers for the biomass energy crop *Miscanthus*. *Theoretical and Applied Genetics* 102: 616–622.

Huelgas VC, Lawrence P, Adkins SW, Mufti MU and Godwin ID (1996) Utilisation of the Australian native species for sorghum improvement. *AIAS Occasional Publication* 93: 369–375.

Jordan DR, Tao YZ, Godwin ID, Henzell RG, Cooper M and McIntyre CL (1998) Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. *Euphytica* 102: 1–7.

Kamala V, Singh SD, Bramel PJ and Rao DM (2002) Sources of resistance to downy mildew in wild and weedy sorghums. *Crop Science* 42: 1357–1360.

Kameswara Rao N, Reddy LJ and Bramel PJ (2003) Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution* 50: 707–721.

Karunakar RI, Narayana YD, Pande S, Mughogho LK and Singh SD (1994) Evaluation of wild and weedy sorghums for downy mildew resistance. *International Sorghum and Millets Newsletter* 35: 104–106.

Komolong B, Chakraborty S, Ryley M and Yates D (2002) Identity and genetic diversity of the sorghum ergot pathogen in Australia. *Australian Journal of Agricultural Research* 53: 621–628.

Kong L, Dong J and Hart GE (2000) Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). *Theoretical and Applied Genetics* 101: 438–448.

Lazarides M, Hacker JB and Andrew MH (1991) Taxonomy, cytology and ecology of indigenous Australian Sorghums (*Sorghum* Moench: Andropogoneae: Poaceae). *Australian Systematic Botany* 4: 591–635.

Levinson G and Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* 4: 203–221.

Matsuoka Y, Mitchell SE, Kresovich S, Goodman M and Doebley J (2002) Microsatellites in Zea—variability, patterns of mutations, and use for evolutionary studies. *Theoretical and Applied Genetics* 104: 436–450.

McLauchlan A, Henry RJ, Isaac PG and Edwards KJ (2001) Microsatellite analysis in cultivated hexaploid wheat and wild wheat relatives. In: Henry RJ (ed.) *Plant Genotyping: The DNA Fingerprinting of Plants*. Wallingford: CABI, pp. 147–159.

Morden CW, Doebley J and Schertz KF (1990) Allozyme variation among the spontaneous species of *Sorghum* section Sorghum (Poaceae). *Theoretical and Applied Genetics* 80: 296–304.

Morgante M and Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* 3: 175–182.

Peakall R, Gilmore S, Keys W, Morgante M and Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology and Evolution* 15: 1275–1287.

Phillips S (1995) Poaceae: *Sorghum purpureosericeum* description. In: Hedberg I and Edwards S (eds) *Flora of Ethiopia and Eritrea*, Vol. 7. Sweden: Swedish Science Press, p. 301.

Price HJ, Dillon SL, Hodnett G, Rooney WL, Ross L and Johnston JS (2004) Genome evolution in the genus *Sorghum* (Poaceae). *Annals of Botany* (in press).

Röder MS, Korzun V, Gill BS and Ganal MW (1998) The physical mapping of microsatellite markers in wheat. *Genome* 41: 278–283.

Rohrer JR, Ahmad R, Southwick SM and Potter D (2004) Microsatellite analysis of relationships among North American plums (Prunus sect. Prunocerasus, Rosaceae). *Plant Systematics and Evolution* 244: 69–75.

Rossetto M (2001) Sourcing SSR markers from related plant species. In: Henry RJ (ed.) *Plant Genotyping: The DNA Fingerprinting of Plants*. Wallingford: CABI, pp. 211–224.

Saghai-Maroof MA, Biyashev RM, Yang GP, Zhang Q and Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. *Proceedings of the National Academy of Science USA* 91: 5466–5470.

Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Paterson AH and Kresovich S (2002) Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theoretical and Applied Genetics* 105: 912–920.

Scott LJ, Shepherd M and Henry RJ (2003) Characterization of highly conserved microsatellite loci in *Araucaria cunninghamii* and related species. *Plant Systematics and Evolution* 236: 115–123.

Sharma HC and Franzmann BA (2001) Host-plant preference and oviposition responses of the sorghum midge, *Stenodiplosis sorghicola* (Coquillett) (Dipt., Cecidomyiidae) towards wild relatives of sorghum. *Journal of Applied Entomology* 125: 109–114.

Smith CW and Frederiksen RA (2000) *Sorghum: Origin, History, Technology, and Production*. New York: John Wiley and Sons.

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M and Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Science* 40: 226–232.

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

Stephenson P, Bryan G, Kirby J, Collins A, Devos K, Busso C and Gale M (1998) Fifty new microsatellite loci for the wheat genetic map. *Theoretical and Applied Genetics* 97: 946–949.

Sudupak MA (2004) Inter and intra-species Inter Simple Sequence Repeat (ISSR) variations in the genus *Cicer*. *Euphytica* 135: 229–238.

Sun Y, Skinner DZ, Liang GH and Hulbert SH (1994) Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics* 89: 26–32.

Tao Y, Manners JM, Ludlow MM and Henzell RG (1993) DNA polymorphisms in grain sorghum (*Sorghum bicolor* (L.) Moench). *Theoretical and Applied Genetics* 86: 679–688.

Taramino G, Tarchini R, Ferrario S, Lee M and Pé ME (1997) Characterisation and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theoretical and Applied Genetics* 95: 66–72.

USDA, ARS (2004) National Genetic Resources program. Germplasm Resources Information Network (GRIN) [online database]. Beltsville, MD: National Germplasm Resources Laboratory, http://www.ars-grin.gov/cgi-bin/npgs/html.

Weber JL (1990) Informativeness of human (dC-dA)n·(dG-dT)n polymorphisms. *Genomics* 7: 524–530.

Westman AL and Kresovich S (1998) The potential for cross-taxa simple-sequence repeat (SSR) amplification between *Arabidopsis thaliana* L. and crop brassicas. *Theoretical and Applied Genetics* 96: 272–281.