Theor Appl Genet (1993) 86:679–688

THEORETICAL AND APPLIED GENETICS

DNA polymorphisms in grain sorghum (*Sorghum bicolor* (L.) Moench)

Y. Tao, J. M. Manners, M. M. Ludlow, R. G. Henzell*

CSIRO, Division of Tropical Crops and Pastures, Cunningham Laboratory, 306 Carmody Road, St. Lucia, Brisbane, 4067, Australia

Received: 18 August 1992 / Accepted: 9 December 1992

Abstract. Molecular markers [random amplified polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP)] were used to determine the frequency of DNA polymorphism in grain sorghum (*Sorghum bicolor* (L.) Moench). Twenty-nine oligonucleotide primers were employed for RAPDs, generating a total of 262 DNA fragments, of which 145 were polymorphic in at least one pairwise comparison between 36 genotypes. Individual primers differed significantly in their ability to detect genetic polymorphism in the species. The overall frequency of polymorphisms was low with a mean frequency of 0.117 polymorphisms per RAPD band being obtained from all pairwise comparisons between genotypes, with maximum and minimum values of 0.212 and 0.039, respectively. Results from phenetic analysis of bandsharing data were consistent with current sub-specific groupings of the species, with clusters of Durra, Zerazera, Caud-Nig, Caud-Kaura and Caffrorum being discernible. The results also indicated that individuals of a similar taxonomic grouping but different geographic origin may be genetically less identical than previously considered. Similar frequencies of polymorphism to that obtained with RAPDs were obtained with RFLPs. Results from these experiments indicated that a high level of genetic uniformity exists within *S. bicolor*.

Key words: Random amplified polymorphic DNA – Restriction fragment length polymorphisms – Genetic diversity – Genome mapping – Sorghum (*Sorghum bicolor*)

Introduction

The application of molecular markers in plant breeding programs will facilitate the improvement of many crop species (Paterson 1991; Tanksley 1989). There are two genetic marker systems based on the direct analysis of DNA that are in common use in genetic improvement programs in plants. These are restriction fragment length polymorphisms (RFLPs; Beckmann and Soller 1983) and random amplified polymorphic DNA (RAPD) markers (Williams et al. 1990). Most applications of molecular markers in plants have so far utilized RFLPs, and several high density genetic maps have been developed for many important crop species (Paterson et al. 1991). In addition, RFLPs have been used for obtaining markers linked to important agronomic traits and to study plant phylogenetics and evolution (Paterson et al. 1991). Recently, many of these important applications of DNA markers have also been undertaken using RAPD with greater convenience and speed than with RFLPs (Williams et al. 1990, 1991; Welsh et al. 1991; Klein-Lankhorst et al. 1991; Carlson et al. 1991; Martin et al. 1991; Michelmore et al. 1991).

The ease with which a genetic map can be developed and applied to a target crop species depends on the genetic complexity of the species and the extent of DNA polymorphism present in the species. Some plant species, such as maize (Smith 1988), potato (Gebhardt et al. 1989) and Brassicas (Figdore et al. 1988) exhibit a high degree of DNA polymorphism even within com-

Communicated by J. W. Snape

* *Present address:* Queensland Department of Primary Industries, Hermitage Research Station, Warwick, Queensland, 4370, Australia

Correspondence to: Y. Tao

mercially exploitable germ plasm pools (Lee et al. 1989; Dudley et al. 1991; Messmer et al. 1991; Melchinger 1991). However, other plants such as soybean (Apuya et al. 1988; Keim et al. 1989) and tomato (Miller and Tanksley 1990) exhibit relatively low frequencies of RFLPs. Thus, the first step in developing genetic maps with DNA markers has been to examine the frequency of DNA polymorphism within species and to identify suitable parents showing sufficient DNA polymorphism for efficient mapping. Genetic mapping in genetically monomorphic species has usually been achieved by using wide crosses between highly divergent parental genotypes, sometimes even using different species (Paterson et al. 1991). The low frequency of DNA polymorphism within a species can also limit the utilization of mapped DNA markers in crosses that are of agronomic importance, but involve more genetically monomorphic parents. For these reasons it is important to establish the frequency of DNA polymorphism within a species before engaging in a plant improvement program using molecular markers.

Grain sorghum is grown as a dryland crop in sub-tropical and tropical Australia and many other regions in the world. In Australia, important objectives for plant breeding programs are increased resistance to drought, midge, lodging and leaf diseases, and high grain yield (Henzell 1992). The application of molecular markers may enhance this breeding effort, particu-

Table 1. A list of the *S. bicolor* genotypes used for the DNA polymorphism investigation

Code	Genotype[a]	Important agronomic trait	Group[c]	Origin/source[d]
1.	Dabar*	High osmotic adjustment (OA)[b]	38: Caud-Kafir	Sudan
2.	Kulum*	Moderate OA	39: Caud-Nig	Sudan
3.	CSM63	Low OA	3: Consp	Mali
4.	CSM387*	High OA	3: Consp	Mali
5.	SC29C*	Moderate OA	41: Durra	Ethiopia
6.	SC31C*	High OA	41: Durra	Ethiopia
7.	SC33C*	Moderate OA	41: Durra	Ethiopia
8.	SC36C*	Moderate OA	41: Durra	Ethiopia
9.	SC56C	Non-senescence, deep roots	39: Caud-Nig	Sudan
10.	SC402	Moderate OA	34: Caud-Kaura	Nigeria
11.	R9188*	Moderate non-senescence	Nig/Nig-Bicolor	Sudan
12.	Tx2813*	Moderate midge resistance	39(1): Zerazera (derivative)	USA
13.	QL27*	Low-level non-senescence, low OA	34: Caud-Kaura (derivative)	Australia
14.	B1877*	High-level non-senescence	Nig/Durra/Docline (derivative)	USA
15.	QL12*	Moderate non-senescence	34: Caud-Kaura (derivative)	Australia
16.	TAM422*	High OA	51: Subgi-Milo (derivative)	USA
17.	KS4*	Senescent, high yield	22: Caffrorum	USA
18.	ES7+*	Drought resistance	F_1 hybrid	USA
19.	40019	Drought resistance	Unknown complex	Australia
20.	QL39	High-level midge resistance	Complex	Australia
21.	QL41	High-level non-senescence	Durra/Complex (derivative)	Australia
22.	AF28*	High-level midge resistance	30(2): Nig-Guin	Brazil
23.	Tx635	Human food type	39(1): Zerazera (derivative)	India/USA
24.	SC35C*	High non-senescence	41: Durra	Ethiopia
25.	M35-1*	Drought resistance, food type	41: Dura	India
26.	SC173C*	Low-level midge resistance, high yield	39(1): Zerazera	Ethiopia
27.	SC108C*	High-level midge resistance, high yield	39(1): Zerazera	Ethiopia
28.	Short Kaura*	Yellow endosperm, non-senescent	34: Caud-Kaura	Nigeria
29.	SC165C*	Moderate midge resistance	39: Caud-Nig	Nigeria
30.	SC170*	High yield	39(1): Zerazera	Ethiopia
31.	IS12736*	Dessication susceptibility	2: Roxburgha	China
32.	IS13590*	Dessication resistance	33: Caudatum	Uganda
33.	IS183*	High temperature resistance	22: Caffrorum	Africa
34.	SA281 (Early Hegari)		38: Caud-Kafir	USA
35.	Tx2737*	Greenbug resistance	34: Caud-Kaura	USA
36.	QL38	High-level midge resistance	Complex	Australia

[a] All genotypes were assessed for RAPDs whilst those marked * were also assessed for RFLPs
[b] Ludlow and Muchow (1990)
[c] Murty et al. (1967)
[d] Genotypes whose source is either the USA or Australia are either of unkown geographic origin or are derivatives of a breeding program

Table 2. A list of primers, number of scored bands produced from each primer and mean 1 – F values

Number	Code[a]	Primer 5′ to 3′	Bands scored			Mean[b]
			Polymorphic	Monomorphic	Total	1 – F
1	H02	TCGGACGTGA	9	4	13	0.153
2	H03	AGACGTCCAC	7	5	12	0.104
3	H04	GGAAGTCGCC	8	3	11	0.210
4	H05	AGTCGTCCCC	5	7	12	0.065
5	H06	ACGCATCGCA	4	5	9	0.109
6	H07	CTGCATCGTG	6	4	10	0.147
7	H08	GAAACACCCC	6	3	9	0.138
8	H09	TGTAGCTGGG	7	3	10	0.213
9	H11	CTTCCGCAGT	6	4	10	0.150
10	H12	ACGCGCATGT	7	4	11	0.126
11	H13	GACGCCACAC	5	2	7	0.204
12	H14	ACCAGGTTGG	7	5	12	0.120
13	T02	GGAGAGACTC	6	4	10	0.149
14	A01	CAGGCCCTTC	4	4	8	0.158
15	A04	AATCGGGCTG	4	5	9	0.057
16	A05	AGGGGTCTTG	1	9	10	0.003
17	A06	GGTCCCTGAC	5	4	9	0.099
18	A08	GTGACGTAGG	2	3	5	0.100
19	A11	CAATCGCCGT	4	4	8	0.116
20	A17	GACCGCTTGT	4	3	7	0.111
21	A20	GTTGCGATCC	1	3	4	0.059
22	No. 2	TGAGTCACCG	8	2	10	0.134
23	No. 5	TCCGTCTAGC	4	2	6	0.092
24	G	TCGTTAGCCAA	2	4	6	0.077
25	Y	TCTCGATGCA	1	7	8	0.014
26	K02	CCGAATTCGCC	7	4	11	0.155
27	K04	ACTCAGCATG	3	5	8	0.091
28	K05	GCTTCGATACG	4	2	6	0.133
29	K19	AGTTCAGGC	8	3	11	0.178
Total bands			145	117	262	

[a] Primers 1–21 are Operon Technologies catalog, numbers 22–29 are CSIRO primer bank numbers
[b] Mean of the 1 – F value (see text) from pairwise comparisons of 36 genotypes

larly for traits such as midge resistance and drought resistance, for which screening is difficult. At present, there is no published genetic map of sorghum incorporating DNA markers, but mapping projects are underway (Hulbert et al. 1990; Godwin et al. 1992). The future success of using mapped RFLP and RAPD markers in breeding programs will greatly depend on the degree of genetic variation in the germ plasm under investigation. At present, there is very little information available on genetic diversity within *S. bicolor*, and no comprehensive survey has been reported. In this investigation, we have screened 36 genotypes of diverse origin for their frequency of DNA polymorphism using RAPDs and RFLPs. This population includes genotypes carrying traits of agronomic importance, and the information on DNA polymorphisms presented provides an essential basis on which to plan future marker-facilitated breeding programs.

Materials and methods

Plant material

A total of 36 genotypes were used (Table 1). These genotypes were chosen because they represent a wide range of geographic origin, a range of sub-specific taxonomic groups, a range of agronomic characteristics and material under development in the Australian and North American plant breeding programs. Plants were grown in a glasshouse, and leaves were removed for DNA analysis at about 5 weeks after emergence.

RAPD analysis

DNA for polymerase chain reaction (PCR) analysis was isolated by the method of Sharp et al. (1989). A total of 30 oligonucleotide primers were used (Table 2). These were mainly 10-base oligonucleotides with some 11- and 9-base primers also included. The oligonucleotides were obtained either from Operon Technologies (Alameda, Calif.) or from the Queensland Institute of Medical Research (Brisbane, Australia). Amplification reactions

were carried out in a 25 µl volume containing 67 m*M* TRIS-HCl (pH 8.8), 16.6 m*M* $(NH_4)_2SO_4$, 0.45% (v/v) Triton-X-100, 200 µg/ml gelatine, 3.5 m*M* $MgCl_2$, 125 µ*M* dNTPs, 0.2 µ*M* primer, 25–50 ng genomic DNA and 1.4 units of *Taq* polymerase (Biotech Int, Australia). Reactions were performed in a Perkin Elmer-Cetus thermal cycler programmed for 40 cycles of 1 min at 94 °C, 1 min at 37 °C and 2 min at 72 °C. Amplification products were analyzed by electrophoresis in 1.3% agarose gels using TAE buffer and visualized by staining with ethidium bromide.

RFLP analysis

DNA extractions, restriction enzyme digests, electrophoresis, alkaline blotting to nylon membranes, DNA probe preparation and labelling, hybridization and autoradiography were all carried out as described previously (Sharp et al. 1989) with the following modifications: Hybond N^+ was used as a membrane, the concentration of sodium dodecyl sulphate (SDS) in the wash buffer was 0.5% (w/v) and plasmid inserts were amplified by PCR using forward and reverse sequencing primers (New England Biolabs, USA). The DNA probes used included maize probes from the maize RFLP map kindly supplied by Drs. B. Burr (Brookhaven Natl Laboratory, USA) and E. Coe (University of Missouri, USA). Probe pTA71 (Gerlach and Bedbrook 1979) containing a wheat ribosomal DNA repeat unit (rDNA) was obtained from Dr. W. Gerlach (CSIRO, Plant Industry, Canberra, Australia). The cDNA clones, M3 and B19, were kindly provided by Dr. P. Chandler (CSIRO, Plant Industry, Canberra, Australia) and represent dehydrin (dehydration-induced protein) sequences (Close et al. 1989) of the maize *dhn* gene and the barley *dhn4* gene, respectively.

Analysis of banding patterns

For both RAPD and RFLP analysis the fraction of bands (F) in common between two genotypes was estimated using the formula of Nei and Li (1979), i.e. $F = 2M_{xy}/M_x + M_y$, where M_{xy} was the number of bands shared by two genotypes and M_x and M_y were the numbers of bands in each genotype. The SAS (Similarity Analysis System) program was used for cluster analysis of the dissimilarity values (1 − F). The McQuitty procedure of SAS, based on the unweighted pair group method with arithmetic mean (UPGMA) and the Proc Tree procedure of SAS were used to produce a phenogram. Phenetic analysis of bandsharing data was chosen in preference to a cladistic clustering analysis because it makes no assumptions on the molecular basis of the genetic polymorphism.

Results

RAPD analysis

Initially five individual plants of each genotype (Table 1) were assayed for RAPDs using 4 primers. No polymorphisms were detected between individuals within a genotype. The DNA samples from five individuals of each genotype were then pooled for a larger scale study. Subsequently, 30 PCR primers of arbitrary sequences were tested on 36 genotypes of sorghum (Tables 1 and 2). All of the 30 primers tested except 1

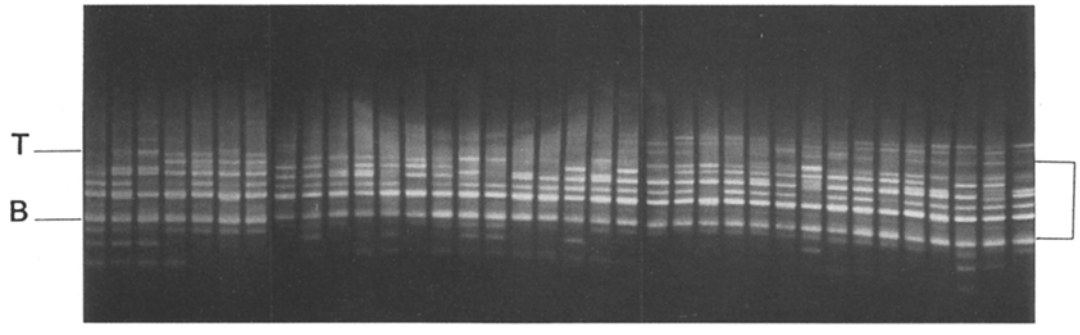

Fig. 1. Ethidium bromide-stained agarose gel of PCR products from DNA samples of 36 sorghum genotypes generated using primer H08. The window of gel length over which the bands were scored is indicated on the *right*. Positions of the bands are indicated on the *left* from top (*T*) to *bottom* (*B*) of the window. The presence (+) or absence (−) of each particular band in each lane (in *descending order* from T to B) is indicated below:

Lane number

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
−	+	+	−	−	+	−	−	−	−	−	−	−	−	−	−	−	+	−	+	+	−	−	−	−	+	+	−	−	−	−	−	−	−	−	−	T
+	−	−	+	+	−	+	−	+	+	+	+	+	+	+	+	−	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
−	−	−	+	+	+	+	+	+	−	+	+	+	−	−	−	−	+	−	+	−	+	+	+	+	+	+	+	+	+	+	−	+	+	−		
+	+	+	−	+	−	+	−	−	+	+	+	−	−	+	−	+	−	−	+	−	−	−	−	−	+	−	+	−	−	−	+	+	−	+	+	Scored
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	bands
−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	+	−	
+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+	
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	B

Table 3. Matrix of 1 – F values between 36 sorghum genotypes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1	0.000																																			
2	0.109	0.000																																		
3	0.096	0.087	0.000																																	
4	0.165	0.129	0.152	0.000																																
5	0.097	0.118	0.111	0.153	0.000																															
6	0.101	0.112	0.115	0.159	0.085	0.000																														
7	0.099	0.115	0.123	0.156	0.082	0.096	0.000																													
8	0.102	0.108	0.106	0.158	0.081	0.079	0.076	0.000																												
9	0.108	0.109	0.102	0.171	0.128	0.127	0.129	0.116	0.000																											
10	0.095	0.101	0.104	0.168	0.120	0.121	0.123	0.106	0.112	0.000																										
11	0.115	0.105	0.093	0.150	0.113	0.127	0.109	0.102	0.118	0.078	0.000																									
12	0.119	0.120	0.117	0.166	0.118	0.121	0.103	0.091	0.128	0.102	0.098	0.000																								
13	0.122	0.128	0.095	0.185	0.111	0.122	0.130	0.076	0.119	0.103	0.099	0.082	0.000																							
14	0.123	0.124	0.096	0.160	0.101	0.098	0.089	0.088	0.105	0.109	0.095	0.120	0.089	0.000																						
15	0.129	0.114	0.091	0.144	0.107	0.126	0.128	0.095	0.138	0.117	0.103	0.112	0.092	0.099	0.000																					
16	0.130	0.135	0.112	0.172	0.118	0.132	0.103	0.096	0.118	0.117	0.093	0.108	0.113	0.105	0.112	0.000																				
17	0.125	0.120	0.097	0.141	0.119	0.135	0.127	0.109	0.116	0.126	0.081	0.132	0.112	0.103	0.095	0.116	0.000																			
18	0.140	0.130	0.118	0.166	0.134	0.134	0.116	0.109	0.126	0.150	0.116	0.136	0.141	0.118	0.115	0.106	0.099	0.000																		
19	0.146	0.131	0.117	0.184	0.123	0.129	0.116	0.098	0.115	0.125	0.141	0.136	0.126	0.112	0.115	0.120	0.124	0.092	0.000																	
20	0.122	0.128	0.090	0.175	0.141	0.145	0.116	0.129	0.136	0.110	0.106	0.100	0.111	0.134	0.130	0.126	0.114	0.128	0.150	0.000																
21	0.085	0.106	0.099	0.162	0.089	0.113	0.074	0.083	0.099	0.094	0.080	0.104	0.100	0.070	0.109	0.089	0.087	0.097	0.122	0.087	0.000															
22	0.131	0.116	0.098	0.168	0.124	0.123	0.120	0.107	0.104	0.103	0.130	0.124	0.115	0.111	0.113	0.119	0.138	0.147	0.127	0.127	0.116	0.000														
23	0.096	0.104	0.079	0.159	0.118	0.129	0.113	0.112	0.111	0.070	0.090	0.116	0.122	0.118	0.109	0.129	0.103	0.116	0.123	0.087	0.086	0.093	0.000													
24	0.097	0.103	0.111	0.149	0.071	0.077	0.064	0.072	0.114	0.108	0.109	0.099	0.120	0.096	0.134	0.129	0.133	0.117	0.116	0.122	0.085	0.100	0.089	0.000												
25	0.135	0.120	0.133	0.178	0.119	0.119	0.116	0.114	0.126	0.120	0.126	0.131	0.137	0.113	0.130	0.121	0.135	0.149	0.123	0.149	0.102	0.101	0.108	0.102	0.000											
26	0.097	0.119	0.106	0.171	0.138	0.129	0.105	0.108	0.130	0.088	0.120	0.089	0.110	0.112	0.145	0.130	0.144	0.153	0.154	0.086	0.096	0.116	0.088	0.100	0.123	0.000										
27	0.096	0.112	0.105	0.185	0.116	0.132	0.098	0.112	0.119	0.103	0.114	0.093	0.104	0.105	0.149	0.124	0.153	0.146	0.137	0.085	0.084	0.120	0.095	0.094	0.116	0.037	0.000									
28	0.105	0.116	0.089	0.168	0.099	0.125	0.097	0.080	0.127	0.131	0.122	0.127	0.102	0.104	0.086	0.127	0.136	0.114	0.093	0.135	0.108	0.118	0.109	0.098	0.130	0.129	0.128	0.000								
29	0.102	0.102	0.090	0.165	0.116	0.126	0.128	0.125	0.096	0.096	0.102	0.117	0.112	0.114	0.132	0.148	0.115	0.140	0.146	0.099	0.098	0.102	0.074	0.108	0.115	0.087	0.081	0.116	0.000							
30	0.103	0.109	0.086	0.181	0.112	0.134	0.111	0.103	0.110	0.109	0.111	0.079	0.100	0.117	0.135	0.120	0.149	0.153	0.144	0.087	0.102	0.106	0.095	0.096	0.113	0.055	0.039	0.114	0.167	0.000						
31	0.140	0.146	0.143	0.182	0.128	0.117	0.119	0.122	0.124	0.118	0.134	0.134	0.130	0.121	0.149	0.134	0.153	0.157	0.142	0.137	0.120	0.094	0.119	0.104	0.080	0.115	0.109	0.148	0.128	0.111	0.000					
32	0.105	0.122	0.130	0.190	0.115	0.111	0.108	0.100	0.123	0.086	0.123	0.091	0.076	0.131	0.127	0.133	0.147	0.161	0.141	0.105	0.104	0.113	0.107	0.103	0.115	0.066	0.071	0.132	0.101	0.083	0.102	0.000				
33	0.136	0.121	0.098	0.152	0.124	0.131	0.133	0.120	0.127	0.121	0.092	0.148	0.138	0.130	0.121	0.127	0.079	0.120	0.135	0.130	0.113	0.123	0.096	0.123	0.125	0.145	0.149	0.141	0.095	0.135	0.128	0.127	0.000			
34	0.109	0.126	0.129	0.212	0.145	0.152	0.133	0.141	0.122	0.095	0.117	0.127	0.128	0.124	0.174	0.138	0.147	0.155	0.151	0.109	0.102	0.118	0.117	0.123	0.130	0.102	0.080	0.157	0.088	0.098	0.122	0.100	0.131	0.000		
35	0.137	0.117	0.125	0.174	0.120	0.121	0.128	0.116	0.122	0.112	0.123	0.133	0.149	0.135	0.132	0.122	0.152	0.125	0.083	0.161	0.134	0.129	0.109	0.113	0.110	0.161	0.144	0.096	0.132	0.150	0.138	0.138	0.121	0.142	0.000	
36	0.113	0.108	0.081	0.170	0.122	0.113	0.136	0.103	0.109	0.094	0.100	0.089	0.079	0.106	0.104	0.109	0.103	0.127	0.132	0.071	0.091	0.111	0.091	0.111	0.128	0.085	0.084	0.118	0.082	0.076	0.136	0.093	0.108	0.108	0.124	0.000

(Operon primer T10 with sequence CCTTCGGAAG) successfully amplified DNA fragments from sorghum DNA samples (Table 2). A total of 262 fragments were visualized across all 36 genotypes of sorghum investigated, and each primer produced aproximately eight to nine intensely staining fragments from each genotype (Fig. 1). Bands of high (> 2 kb) and low (< 0.2 kb) molecular weight were not scored because they could not be reproducibly amplified in replicate experiments. An example of the gel window and bands scored is shown in Fig. 1. Among the 262 fragments, 145 were polymorphic in at least one pairwise comparison between sorghum genotypes, and the remaining 117 were monomorphic.

The number of polymorphic fragments generated by specific primers varied from 1 to 9. Pairwise comparisons were made between all the 36 genotypes with the 29 primers, and the average 1 – F values obtained with each primer were calculated. This provided an indication of the degree of genetic polymorphism detected by each primer (Table 2). There was considerable variation in the ability of individual primers to detect genetic polymorphism, and 1 – F values ranged from 0.003 up to 0.213. The primers which detected the greatest amount of polymorphism (primers nos. 3, 8 and 11; Table 2) did not show any sequence homology or complementarity.

Similarity between sorghum genotypes assessed by RAPDs

The 1 – F values obtained for each pairwise comparison of RAPD bands among the 36 genotypes are shown in Table 3. The mean 1 – F value from all of

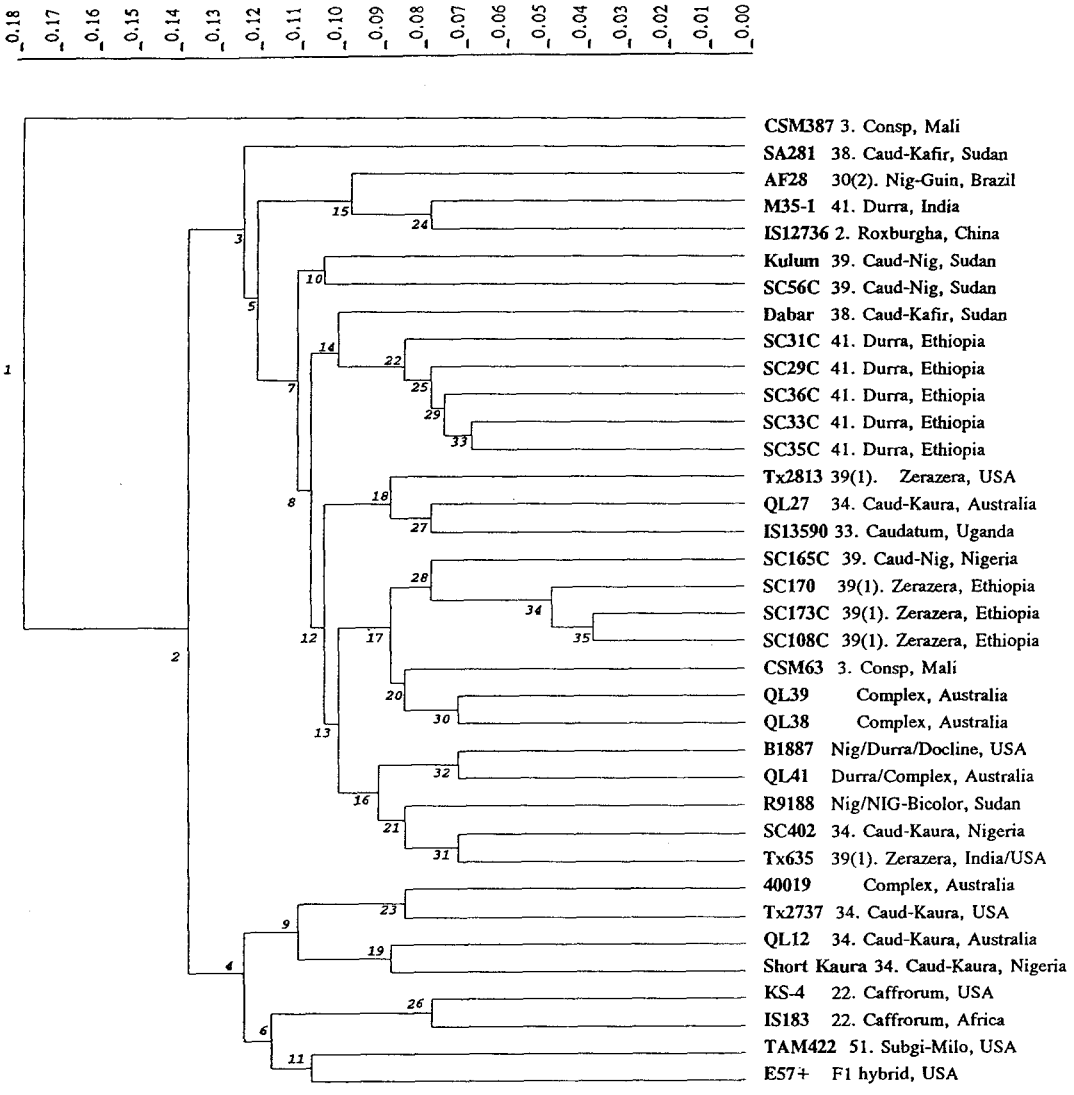

Fig. 2. Phenogram of 36 sorghum genotypes based on RAPD data

these comparisons was 0.117. The highest 1 – F value, 0.212, was observed between the genotype CSM387 (group 3: Consp) and genotype SA281 (group 38: Caud-Kaffir). The genotype CSM387 showed high 1 – F values (mean 1 – F is 0.166) when compared to all of the other genotypes (mean 1 – F value is 0.117) and appeared to be the most divergent genotype. The lowest 1 – F value recorded was 0.039 and was observed between genotypes SC108C and SC170, which are from the same Zerazera group and both originate from Ethiopia.

Cluster analysis of the 1 – F data shown in Table 3 was carried out using the UPGMA method and resulted in the phenogram shown in Fig. 2. The highly divergent nature of the genotype CSM187 is clearly shown on this phenogram. The remaining genotypes fit into two major sub-groups (3 and 4 in Fig. 2) the root of which is indicated at point 2 in Fig. 2. The composition of these sub-groups is highly heterogeneous, but some distinct clusters consistent with established classification and geographic origin can be discerned. Within sub-group 3, cluster 10 contains 2 Caud-Nig genotypes from the Sudan, but this cluster omits another Caud-Nig genotype (SC165C) from another geographical region, Nigeria. Cluster 22 contains all of the Ethiopian Durra genotypes, but omits the Durra genotype from India (M35-1). This Indian genotype shows the greatest similarity to genotype IS12736 (cluster 24), which is the only other genotype of Asian origin tested. The Ethiopian Zerazera genotypes formed a cluster of high genetic similarity (cluster 34). Two Australian developed cultivars (QL38 and QL39), which have the same ancestor in their pedigree, were clustered (cluster 30). Sub-group 3 contained all of the Durra and Zerazera genotypes and other lines derived from these important groups. In sub-group 4, two clusters of known taxonomic groups were identified. Cluster 9 contained 3 Caud-Kaura genotypes although 1 Caud-Kaura genotype (SC402) remained unclustered and was present in sub-group 3. Cluster 26 contained 2 genotypes of Caffrorum designation. In general, the cluster analysis is consistent with previous taxonomic grouping, but some significant exceptions were present. The most striking discrepancy was between CSM387 and CSM87, which were previously classified in the same group and have the same geographic origin in Mali.

RFLP analysis

A total of 27 genotypes (Table 1) were assessed for RFLPs using 11 heterologous probes originating from either maize, barley or wheat (Table 4). All of these DNA probes detected single (Fig. 3) or low numbers of bands. The barley dehydrin cDNA (B19) hydridized to up to five bands, suggesting a multi-gene family.

Three restriction enzymes, viz. *Hin*dIII, *Eco*RI, *Bam*HI, were used to detect RFLPs. From the total of 33 probe-enzyme combinations used, 43 bands were observed, and 9 probe-enzyme combinations detected RFLP. Of the probes 4 were monomorphic in all 27 genotypes with all three enzymes. Mean 1 – F values for the DNA probes detecting polymorphism varied from 0.041 to 0.261 when all probe-enzyme combinations were included. The mean 1 – F value obtained for 33 probe-enzyme combinations in all pairwise comparisons of the 27 genotypes was 0.153. This mean 1 – F value from RFLPs (0.153) is slightly higher than that obtained from RAPDs (0.126) for the same 27 genotypes. However, a total of only 43 fragments were produced from the RFLP analysis compared to 262 fragments from the RAPD analysis. The small number of bands detected and the corresponding low number of RFLPs precluded a meaningful cluster analysis of the genotypes.

Table 4. A list of RFLP probes hybridized with 27 sorghum genotypes after *Bam*HI, *Eco*RI and *Hin*dII digestion

Probe	Source	Useful enzyme(s)[a]	Number of bands[b]	Mean 1 – F
pTa71	Wheat rDNA clone	*Bam*HI	5	0.041
M3	Maize dehydrin gene	*Bam*HI/*Hin*dIII	4	0.261
B19	Barley dehydrin gene	*Hin*dIII	7	0.197
BNL8.29	Maize genomic DNA	*Eco*RI	4	0.166
BNL9.11	Maize genomic DNA	*Bam*HI	3	0.081
UMC44	Maize genomic DNA	*Eco*RI/*Hin*dII	3	0.243
UMC55	Maize genomic DNA	*Hin*dIII	5	0.241
UMC76	Maize genomic DNA	–	3	0
UMC94	Maize genomic DNA	–	3	0
UMC137	Maize genomic DNA	–	3	0
BNL7.71	Maize genomic DNA	–	3	0

[a] Enzymes which produced RFLPs
[b] Total bands observed from hybridization of each probe to DNA samples from the 27 genotypes restricted with all three enzymes
[c] Values were calculated on the basis of data obtained from all of the probe-enzyme combinations

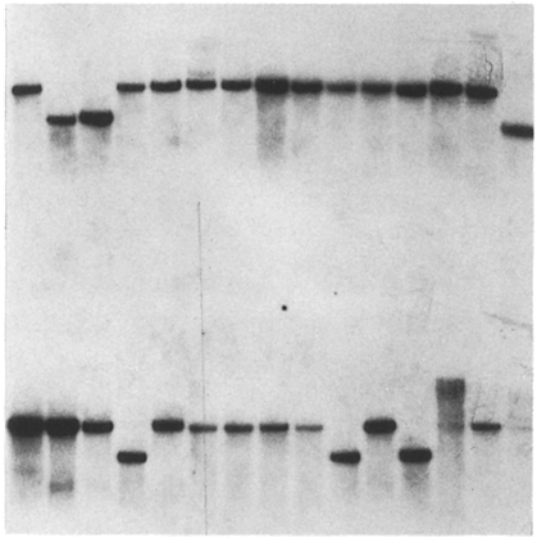

Fig. 3. Polymorphisms in *S. bicolor* revealed by hybridization of maize genomic DNA clone UMC44 to *Bam*HI-restricted DNA from 27 sorghum genotypes (Table 1). The *bottom right* three lanes were not included in this survey

To provide a more significant comparison of the ability of RFLPs and RAPDs to detect DNA polymorphisms in sorghum, 2 genotypes (QL27 and TAM422) were closen for a detailed comparison with a greater number of RFLP probes. These 2 genotypes were from the distinct sub-groups 3 and 4 in the RAPD classification (Fig. 2) and were of interest to us because they differ substantially in the important physiological trait of osmotic adjustment (Ludlow and Muchow 1990) and represent potential parents for studying the inheritance of this trait. An additional 18 RFLP probes from the maize RFLP map (University of Missouri, Columbia) were used to compare these two genotypes, making a total of 29 DNA probes. The same restriction enzymes were used to test for RFLPs, resulting in a total of 87 probe-enzyme combinations. From the 29 probes tested with each enzyme, 4, 3 and 4 probes detected RFLPs with *Bam*HI, *Eco*RI and *Hin*dIII, respectively, indicating that there was no preference for any enzyme to detect polymorphism. A total number of 106 fragments were included in this comparison of 2 genotypes, and a $1-F$ value of 0.103 was obtained. This value obtained with RFLPs compares favorably with the value of 0.113 obtained with RAPDs (Table 3).

Discussion

The genotypes studied in this work represent germ plasm with important agronomic characteristics for potential commercial exploitation. The incorporation of some of these traits into elite commercial lines could be enhanced by the application of molecular markers. However, solving the initial problem of locating a molecular marker linked to a trait and, secondly, being able to detect the marker in other genetic backgrounds depends upon the degree of genetic polymorphism within the germ plasm to be exploited. Even though our investigation included a number of genotypes of wide taxonomic distinction within the species, only a low level of genetic polymorphism was detected when compared to that observed in some other crop species, e.g. maize. Future programs aimed at using molecular markers in sorghum will need to take this level of genetic variation into consideration.

In associated work analyzing the heritability of RAPDs and osmotic adjustment in an F_2 population of a cross between the genotypes QL27 and Tam422, we have demonstrated that RAPD bands, which are polymorphic in the parent, are inherited as dominant genetic markers (Y. Tao, unpublished data). Due to the heritable nature of RAPD markers, it follows that the data displayed in Table 3 can be used as a basis for the rational calculation of the numbers of primers that will need to be used to generate a useful number of RAPD polymorphisms either for genetic mapping or screening F_2 and backcross families, near-isogenic lines (Martin et al. 1991) or bulked segregants (Michelmore et al. 1991) for markers linked to traits of interest. Obviously, a careful selection of mapping parents needs to be made to ensure a high level of genetic polymorphism for efficient mapping. From our data, the best choice of parents for a intraspecific mapping polulation would be the genotypes CSM387 and SA281, which differed in 21% of their RAPD bands. Because of the dominant nature of RAPD markers genetic mapping is best undertaken in a backcross or recombinant inbred population rather than in a F_2 population. However, the low frequency of genetic polymorphism in the genotypes tested here suggests that a genetic map based solely on RAPD markers would not be readily transferred to crosses involving other more monomorphic parents.

The comparison of $1-F$ values obtained with specific oligonucleotide primers (Table 2) demonstrated that oligonucleotide primers differ in their usefulness in detecting DNA polymorphisms within this species. This suggests that eventually it may be possible to identify a set of primers that target more variable parts of the sorghum genome. A collection of such 'hypervariable' RAPD primers would greatly facilitate the application of RAPDs in breeding and may be particularly useful for 'DNA fingerprinting' new cultivars. No clue as to why some primers were more useful in detecting polymorphism was obtained in this investigation, but further analysis of the amplified bands and their target regions in the genome is underway.

In this investigation, the frequency of DNA polymorphism detected by RAPD and RFLP markers was found to be similar. This may be because DNA polymorphisms detected on RFLPs and RAPDs are produced by similar processes. Small differences in DNA sequence may occur at either the site of cleavage by a restriction enzyme for a RFLP or at the binding site of a primer for a RAPD. Alternatively, insertions and deletions may occur between these sites. However, the molecular basis of RAPD remains to be fully elucidated. The speed and convenience of RAPD analysis permits a larger number of DNA fragments to be compared than is possible with RFLPs.

With RAPDs it is possible to assess genetic similarity amongst a range of genotypes on the basis of band sharing. Quantitative phenetic analysis of RAPDs in other plant species has produced matrices of genetic similarity that are consistent with classifications based on morphological and agronomic criteria, e.g. peanut (Halward et al. 1991) and *Stylosanthes* spp. (Kazan et al. 1992, 1993). The phenogram produced on the basis of RAPDs for the genotypes of sorghum showed several consistencies with previous sub-specific taxa defined on the basis of morphological criteria. Clusters of genotypes corresponding to the recognized Durra, Zerazera, Caud-Nig, Caud Kaura and Caffrorum groups were discerned. Interestingly, the RAPD data did not cluster an Indian Durra genotype with the Ethiopian Durra group and a Nigerian Caud-Nig genotype with the Sudanese Caud-Nig group. This suggests that geographical origin may also have an important effect on the genetic similarity of genotypes regardless of apparent morphological similarities. Some genotypes were not clustered within supposedly related groups. The most obvious of these apparent anomalies was the finding that CSM387 was highly divergent from the other genotypes. The reason for the level of genetic dissimilarity of this particular genotype is not known. With the exception of CSM387, the remaining genotypes were clustered into two principal groups, 3 and 4. As far as we are aware these two groupings, which appear to separate the Caud-Kaura and Caffrorum into sub-group 4 and others into sub-group 3, have not been recognized in the past. Clearly, there is considerable scope for a more comprehensive analysis and perhaps a reassessment of the sub-specific taxonomy of *S. bicolor* using RAPDs.

It will be of interest to determine whether genetic similarity determined by RAPDs can provide a guide to distinguishing heterotic groups of sorghum. Evidence in maize (Smith and Smith 1989; Stuber 1992) has indicated that DNA polymorphism frequency measured by RFLPs strongly correlates with heterotic effect. Thus, the matrix of DNA polymorphisms presented in Table 3 and Fig. 2 may have potential for commercial exploitation in hybrid seed production as well as taxonomic significance.

In summary, the data presented here represent the first major assessment of DNA polymorphism in *S. bicolor* and further demonstrate the utility of RAPD markers for assessing genetic similarity in plants. These results will lay the foundation for future research aimed at using molecular markers to exploit the useful traits of some of the genotypes that we have assessed.

Acknowledgments. We are very grateful to Prof. F. R. Miller for much of the information contained in Table 1, to Dr. Peter N. Jones for his kind help on programming and computing, to Dr. Kemal Kazan for advice on RAPD analysis and to Mr. Jaya Basnayake for assistance in growing the plants.

References

Apuya NR, Frazier BL, Keim P, Roth EJ, Lark KG (1988) Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merrill. Theor Appl Genet 75:889–901

Beckmann JS, Soller M (1983) Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theor Appl Genet 67:35–43

Carlson JE, Tulsieram LK, Glaubitz JC, Luke, VWK, Kauffeldt C, Rutledge R (1991) Segregation of random amplified DNA markers in F_1 progeny of conifers. Theor Appl Genet 83:194–200

Close TJ, Kortt AA, Chandler PM (1989) A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. Plant Mol Biol 13:95–108

Dudley JW, Saghai Maroof MA, Rufener GK (1991) Molecular markers and grouping of parents in maize breeding programs. Crop Sci 31:718–723

Figdore SS, Kennard WC, Song KM, Osborn TC (1988) Assessment of the degree of restriction fragment length polymorphism in *Brassica*. Theor Appl Genet 75:833–840

Gebhardt C, Blomendahl U, Debener T, Salamini F, Ritter E (1989) Identification of 2n breeding lines and 4n varieties of potato (*Solanum tuberosum* ssp. *tuberosum*) with RFLP fingerprints. Theor Appl Genet 78:16–22

Gerlach WL, Bedbrook JR (1979) Cloning and characterisation of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7:1869–1885

Godwin ID, Tao YZ, Manners JM, Ludlow MM (1992) Cellular and molecular approaches to sorghum improvement. In: Foale MA, Henzell RG, Vance PN (eds) Proc 2nd Aust Sorghum Conf. Australian Institute of Agricultural Science, Melbourne, Occasional Publ No. 68, pp 83–96

Halward TM, Stalker HT, Larue EA, Kochert G (1991) Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome 34:1013–1020

Henzell RG (1992) Grain sorghum breeding in Australia: its current status and future prospects. In: Foale MA, Henzell RG, Vance PN (eds) Proc 2nd Aust Sorghum Conf. Australian Institute of Agricultural Science, Melbourne, Occasional Publ No. 68, pp 70–82

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

Kazan K, Manners JM, Cameron DF (1993) Genetic relationships and variation in the *Stylosanthes guianensis* species

complex assessed by random amplified polymorphic DNA. Genome 36:43–49
Kazan K, Manners JM, Cameron DF (1993) Genetic variation in agronomically important species of *Stylosanthes* determined using random amplified polymorphic DNA markers. Theor Appl Genet 85:882–888
Keim P, Shoemaker RC, Palmer RG (1989) Restriction fragment length polymorphism diversity in soybean. Theor Appl Genet 77:786–792
Klein-Lankhorst RM, Vermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83:108–114
Lee M, Godshalk EB, Lamkey KB, Woodman WW (1989) Association of restriction fragment length polymorphisms among maize inbred with agronomic performance of their crosses. Crop Sci 29:1067–1071
Ludlow MM, Muchow RC (1990) A critical evaluation of traits for improving crop yield in water limited environments. Adv Agron 43:107–153
Martin GB, Williams JG, Tanksley SD (1991) Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. Proc Natl Acad Sci USA 88:2336–2340
Melchinger AE, Messmer MM, Lee M, Woodman WL, Lamkey KR (1991) Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. Crop Sci 31:669–678
Messmer MM, Melchinger AE, Lee M, Woodman WL, Lee EA, Lamkey KR (1991) Genetic diversity among progenitors and elite lines from the Iowa Stiff Stalk Synthetic (BSSS) maize population: comparison of allozyme and RFLP data. Theor Appl Genet 83:97–107
Michelmore RW, Paran I, Kessel RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832
Miller JD, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448
Murty BR, Arunachalam V, Saxena MBL (1967) Classification and catalogue of a world collection of sorghum. Indian J Genet Plant Breed 27:1–312
Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273
Paterson AH, Tanksley SD, Sorrells ME (1991) DNA markers in plant improvement. Adv Agron 46:39–90
Sharp PJ, Chao S, Desai S, Gale MD (1989) The isolation, characterization and application in the Triticeae of a set of wheat RFLP probes identifying each homeologous chromosome arm. Theor Appl Genet 78:342–348
Smith JSC (1988) Diversity of United States hybrid maize germ plasm; isozymic and chromatographic evidence. Crop Sci 28:63–69
Smith JSC, Smith OS (1989) The use of morphological, biochemical and genetic characteristics to measure distance and to test for minimum distance between inbred lines of maize (*Zea mays* L.). (Mimeo of paper presented at UPOV Workshop, Versailles, France, October 1989). Pioneer Hi-Bred Int, Johnston, Iowa
Stuber CW (1992) Biochemical and molecular markers in plant breeding. In: Janick J, Wiley J and Sons (eds) Plants breeding reviews vol 9. John Wiley and Sons, Chichester Brisbane, pp 37–61
Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnology 7:257–264
Welsh J, Honeycutt RJ, McClelland M, Sobral BW (1991) Parentage determination in maize hybrids using the arbitrarily primed polymerase chain reaction (AP-PCR). Theor Appl Genet 82:473–476
Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18: 6531–6535
Williams MN, Pande N, Nair S, Mohan M, Bennett (1991) Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa* L.) genomic DNA. Theor Appl Genet 82:489–498