Theor Appl Genet (2000) 101:438–448 © Springer-Verlag 2000

L. Kong · J. Dong · G.E. Hart

Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs)

Received: 13 September 1999 / Accepted: 2 December 1999

Abstract Fifty one clones isolated from a size-fractionated genomic DNA library of *Sorghum bicolor* (L.) Moench, that had been probed with four radiolabeled di- and tri-nucleotide oligomers, were sequenced. Fifty of the clones contained one or more simple-sequence repeats (SSRs) [72% of which were $(AG/TC)_n$ SSRs] and, following analysis of the clones, polymerase-chain-reaction primer sets that amplify 38 unique SSR loci were developed. Genotyping of the 38 loci in 18 sorghum accessions, including the parents of a recombinant inbred (RI) mapping population, revealed polymorphism at 36 of the loci among the 18 accessions and at 31 of the loci (not including null alleles at two loci) between the parents of the RI population. All of the latter 31 loci were mapped. The genotypes at 17-mapped SSR loci were assayed in 190 *S. bicolor* accessions in order to determine δ^*_T, the estimated level of allelic differentiation (the estimated probability that two members of a population, chosen at random and without replacement, differ in allelic composition), at each of the loci. The mean δ^*_T value determined for *S. bicolor* overall was 0.89, the range of mean δ^*_T values for ten *S. bicolor* races was from 0.88 to 0.83, and the range of mean δ^*_T values for ten working groups (= sub-races) of the race caudatum, with only two exceptions, was from 0.87 to 0.79. The lowest δ^*_T values for six of the loci among the ten race-caudatum working groups ranged from 0.86 to 0.70; thus, the probability that different alleles will be present at one or more of these loci in two accessions chosen at random from a working group is > 0.996 when three of the loci are genotyped, and >0.9999 when all six of the loci are genotyped. The results of this study confirm that most *S. bicolor* SSR loci are sufficiently polymorphic to be useful in marker-assisted selection programs and they indicate that the levels of polymorphism at some loci are high enough to allow the vast majority of *S. bicolor* accessions, even accessions within working groups, to be distinguished from one another by determining the genotypes at a small number, perhaps as few as a half-dozen, SSR loci.

Key words Fingerprinting · Genetic differentiation · Linkage mapping · *Sorghum bicolor* · SSRs

Communicated by M.A. Saghai Maroof

L. Kong, J. Dong, G.E. Hart (✉)
Department of Soil and Crop Sciences,
Texas A&M University, College Station, TX 77843, USA
e-mail: g-hart@tamu.edu, Fax: +1979-845-6483

Present address: L. Kong
5 Leavey Court, North York, Ontario, M2H 1E5, Canada

Introduction

DNA simple-sequence repeats (SSRs, also called microsatellites) are numerous and highly polymorphic in plants (Morgante and Olivieri 1993; Saghai Maroof et al. 1994; Wang et al. 1994; Rongwen et al. 1995; Yang et al. 1995). Although costly to develop relative to some other classes of genetic markers, once developed, analysis of them is both easy and inexpensive. Their high information content and other favorable characteristics make them excellent genetic markers for many types of investigations, including marker-assisted selection and fingerprinting of germplasm collections. Large numbers of SSR loci have been genetically mapped in several agronomically important plant species, including wheat (Röder et al. 1998), rice (Cho et al. 1998), and soybean (Cregan et al. 1999).

Sorghum is fifth in acreage among the world's cereals (Doggett 1988). A diverse genus belonging to the tribe Andropogoneae, it consists of cultivated and wild species, many of which are interfertile. *Sorghum bicolor* ssp. *bicolor* ($2n = 20$) is the most important taxon agronomically in that it includes the cultivated grain races. It is a diploid, is highly self-pollinated, and possesses considerable diversity, both morphologically and in agronomic traits, such as adaptive pest resistance. It consists of five basic races (bicolor, caudatum, durra, guinea, and kafir) and several hybrid races (for a review see Doggett 1988). Also, numerous working-groups (= sub-races) have been identified. Several RFLP linkage maps of *S. bicolor* have been constructed (Hulbert et al. 1990;

Binelli et al. 1992; Whitkus et al. 1992; Berhan et al. 1993; Chittenden et al. 1994; Pereira et al. 1994; Ragab et al. 1994; Xu et al. 1994; Dufour et al. 1996, 1997; Peng et al. 1999). Combined, these maps include over 800 markers (Bennetzen et al. 2000). Only a small number of SSR primer sets have been developed for sorghum, however (Brown et al. 1996; Taramino et al. 1997), and linkage map locations have been published for only seven sorghum SSR loci (Taramino et al. 1997).

This paper reports the results of a study designed to: (1) isolate and sequence SSR-containing clones from a size-fractionated *S. bicolor* genomic DNA library, (2) develop and test primer sets for SSRs contained in the sequenced clones, (3) determine linkage-map positions of polymorphic SSR loci, and (4) estimate the levels of allelic differentiation, δ_T (Gregorius 1987), at selected mapped SSR loci. Brief accounts of some of this research were reported in earlier abstracts (Dong et al. 1996; Kong et al. 1997). The characteristics and map positions of a large number of additional sorghum SSRs isolated from BAC and enriched-genomic DNA libraries will be described in a later paper.

Materials and methods

Plant materials

S. bicolor (L.) Moench accession BTx623 was the source of the DNA used to construct a genomic DNA library from which SSRs were isolated. Each SSR locus that was amplified with the primer set produced for it was genotyped in five related kafir male-sterile lines (B-lines), including BTx623, six related fertility restorer lines (R-lines), six photoperiodism-converted zerazera ('caudatum-like') lines, and IS3620C, a guinea line (Table 1). These lines were obtained from Dr. F.R. Miller. Linkage mapping was performed in a population of 137 F_{6-8} recombinant inbred (RI) lines developed by Dr. K.F. Schertz from a cross between BTx623 and IS3620C (Peng et al. 1999). One-hundred and ninety *S. bicolor* accessions (Table 2), representing 16 races and 12 working groups (sub-races) within one of the races (caudatum), were genotyped at selected SSR loci. Within the races and working groups, the accessions analyzed were chosen at random from a much larger number of accessions obtained from Dr. Jeff Dahlberg.

Isolation of SSRs and design and production of primers

Genomic DNA was purified by the method of Murray and Thompson (1980) as modified by Saghai Maroof et al. (1984). To produce a size-fractionated sorghum genomic DNA library, DNA purified from BTx623 was digested with *Sau*3AI, electrophoresed in a low-melting-point agarose gel, and the gel zone containing DNA fragments 100–500 bp in length was excised and digested with agarase. DNA fragments released by the digestion were ligated into the *Bam*HI site of pGEM1 and the vector was transformed into *Escherichia coli* strain DH10B by electroporation. Transformed cells were plated onto LB+agar medium with ampicillin, IPTG, and X-gal. Following overnight growth of cells at 37°C, white colonies were selected, re-plated, transferred to nylon membranes, and probed with a radiolabeled nucleotide mixture containing two dinucleotide 20-mers [$(AG)_{10}$ and $(AC)_{10}$] and two trinucleotide 21-mers [$(CCT)_7$ and $(ATT)_7$]. Hybridization of membranes was performed overnight at 48°C in 0.5 M monobasic/dibasic sodium phosphate buffer, pH 7.2, containing 1% BSA, 7% SDS, and 1 mM EDTA. Following hybridization, the membranes were washed twice in a 2×SSC+0.1% SDS solution at 48°C for 30 min and once in a 1×SSC+0.1% SDS solution at 48°C for 30 min. Clones that gave a high signal-level were sequenced in an Applied Biosystems 373 DNA sequencer. Initially, colonies that displayed a high signal-level were re-plated (three samples/colony) and re-probed, but this was found to be unnecessary and was discontinued. Polymerase-chain-reaction (PCR) primer sets for amplification of SSRs were designed using the Oligo 5.0 software program (purchased from National Biosciences, Plymouth, Minn., and now obtainable from Molecular Biology Insights, Cascade, Colo.) and were obtained from Gibco/BRL, Gaithersburg, Md.

Table 1 Sorghum lines that were genotyped at each SSR locus that was amplified with the primer set produced for it

Guinea line	B-lines	Zerazera lines	R-lines
IS3620C[a]	BTx378	SC108-14E	RTx430
	BTx623[a]	SC110-14E	RTx435
	BTx631	SC120-14E	RTx2536
	BTx3197	SC170-14E	RTx2737
	BTam618	SC173-14E	RTx2767
		SC175-14E	RTx2908

[a] A parent of the recombinant inbred mapping population

Amplification of SSRs

PCR reactions were conducted in either a PE9600 or a PE9700 Perkin-Elmer (Norwalk, Conn.) DNA Thermal Cycler in a 12.5-μl reaction mixture containing 200 μM of each dNTP, 1×PCR buffer (BRL), 1 mM $MgCl_2$, 25 ng of each primer, 0.25 units of *Taq* polymerase (BRL) and 10 ng of DNA. The PCR program consisted of an initial denaturation for 2 min at 94°C and then 30 cycles of denaturation for 1 min at 94°C, annealing for 30 s, and extension at 72°C for 1 min. Annealing was performed at either 55°C or 60°C with most of the primer pairs (see Results). The last PCR cycle was followed by a 7-min extension at 72°C. PCR products were separated in 6% denaturing polyacrylamide gels and silver-stained using the procedure of Fritz et al. (1999). The ability of each primer set to amplify the SSR locus for which it was designed was tested on replicated DNA samples of BTx623 and IS3620C, the parents of the RI mapping population. Primer sets that did not amplify the SSR locus from at least one of the parental lines were not used further.

Data analyses

SSRs were placed on a framework RFLP linkage map of the BTx623 × IS3620C RI population. This framework map, composed of a subset of the RFLP loci mapped in the population by Peng et al. (1999), was constructed using the computer program MAPMAKER V2.0 for Mcintosh. Recombination frequencies were converted to centiMorgans (cM) using the Kosambi mapping function (Kosambi 1944). Criteria for the framework map included spacing of markers at intervals of 5–15 cM to the maximum extent possible, a minimum LOD score ≥ 3.0 for terminal triplets, and a minimum LOD score ≥ 5.0 for all non-terminal triplets. SSR loci were placed on the framework map using the MAPMAKER 'Near' and 'Try' commands, and the LOD score for each adjacent triplet was determined using the 'Ripple' command.

The level of allelic differentiation, δ_T (Gregorius 1987), was estimated as:

$$\delta_T^* = \left(1 - \sum_i p_i^2\right)[N/(N-1)]$$

where p_i is the frequency of the ith allele of each locus analyzed and N is the sample size. δ_T^* is the estimated probability that two members of a population, chosen at random and without replace-

Table 2 Race, working group, plant Introduction (P.I.) number, and local name of 190 *S. bicolor* accessions that were genotyped at 17 SSR loci

Race	Working group	P.I. #	Local name (if any)	Race	Working group	P.I. #	Local name (if any)
Bicolor	Bicolor	563146	–	Caudatum	Caudatum-bicolor	570222	–
Bicolor	Bicolor	570860	–	Caudatum	Caudatum-bicolor	570947	–
Bicolor (?)	Bicolor	570884	Dori Dark Brown	Caudatum	Caudatum-durra	562775	Feseika
Bicolor	Bicolor	570992	Ankolib	Caudatum	Caudatum-durra	569085	WM 10103
Bicolor	Bicolor	571245	A29	Caudatum	Caudatum-durra	569090	WM 10115
Bicolor	Bicolor-kafir	570502	–	Caudatum	Caudatum-durra	569232	WM 10454
Bicolor	Dochna	570372	–	Caudatum	Caudatum-durra	569283	WM 10591
Bicolor	Caffrorum-bicolor	570821	Zera Zera 3	Caudatum	Caudatum-durra	569387	WM 11090
Bicolor	Caudatum-bicolor	569340	WM 11017	Caudatum	Caudatum-durra	570393	–
Bicolor	Caudatum-bicolor	571195	Wad Akar 5	Caudatum	Caudatum-durra	570855	Culum
Bicolor	Durra-bicolor	217683	Shidakia Q2–2-53	Caudatum	Caudatum-durra	570898	DS 16
Bicolor	Sudanense	570498	–	Caudatum	Caudatum-durra	571001	Tabroro
Bicolor-caudatum	Caudatum-bicolor	569264	WM 10550	Caudatum	Caudatum-durra	571260	A 75
Bicolor-caudatum	Caudatum-bicolor	569287	WM 10596	Caudatum	Caudatum-guineense	152615	MN753 Culum Brick CSC52
Bicolor-caudatum	Caudatum-bicolor	570841	Klor	Caudatum	Caudatum-guineense	217694	Maga Abiad Q2–2-68
Bicolor-caudatum	Caudatum-bicolor	571202	Yt 69	Caudatum	Caudatum-guineense	570391	Atari
Bicolor-caudatum	Caudatum-bicolor	571380	Feterita	Caudatum	Caudatum-guineense	570402	Puchalla White
Bicolor-caudatum	Nigricans-bicolor	568674	–	Caudatum	Caudatum-guineense	570714	Beid Elkebish
Bicolor-durra	Durra-bicolor	568316	–	Caudatum	Caudatum-guineense	570747	Thok Brown
Bicolor-durra	Durra-bicolor	568630	–	Caudatum	Caudatum-guineense	571184	Nagad Grey
Bicolor-durra	Durra-bicolor	569148	WM 10231	Caudatum	Caudatum-kaura	570265	–
Bicolor-durra	Durra-bicolor	570426	–	Caudatum	Caudatum-kafir	569108	WM 10143
Bicolor-durra	Durra-bicolor	571156	El Safra	Caudatum	Caudatum-kafir	569141	WM 10214
Bicolor-durra	Durra-dochna	570415	–	Caudatum	Caudatum-kafir	569191	WM 10370
Bicolor-guinea	Guinea-bicolor	152727	MN 855 (Malwal soni)	Caudatum	Caudatum-kafir	569235	WM 10465
Bicolor-guinea	Guinea-bicolor	570686	Addar Abu Shar	Caudatum	Caudatum-kafir	570778	UM El. Teiman
Bicolor-kafir	Caffrorum-bicolor	568572	–	Caudatum	Caudatum-kafir	571186	Query 2
Bicolor-kafir	Caffrorum-bicolor	570564	–	Caudatum	Caudatum-kafir	571274	A 132
Caudatum	Caudatum	152591	MN 730 (Akur-Gob)	Caudatum	Caudatum-nigricans	568996	Ligalo
Caudatum	Caudatum	152611	MN 749 (Budy)	Caudatum	Caudatum-nigricans	570705	Kemurit White
Caudatum	Caudatum	152965	–	Caudatum	Caudatum-nigricans	570716	Lwel
Caudatum	Caudatum	217672	Ajak Q2–2-47	Caudatum	Caudatum-nigricans	570836	Fet Gondal
Caudatum	Caudatum	217718	Wad Beshir Q2-2-75	Caudatum	Caudatum-nigricans	571065	Kano
Caudatum	Caudatum	217797	El Roble Q2-3-10	Caudatum	Caudatum-nigricans	571150	Abu Chorshan
Caudatum	Caudatum	563310	Bolichingan	Caudatum	Durra-nigricans	568544	–
Caudatum	Caudatum	570734	Fet Suki	Caudatum	Durra-nigricans	568583	–
Caudatum	Caudatum	570744	Bahana	Caudatum	Durra-nigricans	568603	–
Caudatum	Caudatum	570764	Barking	Caudatum	Durra-nigricans	568624	–
Caudatum	Caudatum	570782	Kulmuta	Caudatum	Durra-nigricans	569233	WM 10455
Caudatum	Caudatum	570800	Early Birghalli Ahmer	Caudatum	Nigricans	569972	–
Caudatum	Caudatum	571042	Gadamel Hamam	Caudatum	Nigricans	570011	–
Caudatum	Caudatum	571112	Banana 1	Caudatum	Nigricans	570016	–
Caudatum	Caudatum	571193	Wad Akar 2	Caudatum	Nigricans	570889	Addar VI
Caudatum	Caudatum	571243	A 14	Caudatum	Nigricans	571304	A 351
Caudatum	Caudatum	571285	A 212	Caudatum	Nigricans-feterita	568409	–
Caudatum	Caudatum	571300	A 315	Caudatum	Nigricans-feterita	568524	–
Caudatum	Caudatum	571334	B 205	Caudatum	Nigricans-feterita	569937	–
Caudatum	Caudatum	571336	Uganda I	Caudatum	Nigricans-feterita	570307	–
Caudatum	Caudatum-bicolor	569997	–	Caudatum	Nigricans-feterita	570509	–
Caudatum	Caudatum-bicolor	570138	–	Caudatum	Zerazera	217799	Hegari Seifi Q2-3-17

Table 2 (continued)

Race	Working group	P.I. #	Local name (if any)	Race	Working group	P.I. #	Local name (if any)
Caudatum	Zerazera	570052	Bright White	Guinea	Guineense	291027	*S. bicolor*
Caudatum	Zerazera	570717	Waramara	Guinea	Guineense	570590	–
Caudatum	Zerazera	570726	Yom	Guinea	Guineense	570748	Mugbash White
Caudatum	Zerazera	570777	Rabou	Guinea	Caudatum-guineense	152748	MN 877 (Njan Dok)
Caudatum	Zerazera	570844	Morijang	Guinea	Caudatum-guineense	570036	–
Caudatum	Zerazera	570994	Nyithin	Guinea	Caudatum-guineense	571235	SBI 100
Caudatum	#80	569482	Cross-65/18	Guinea	Roxburghii-shallu	569013	–
Caudatum	#80	569556	Cross-41/57	Guinea-caudatum	Caudatum-guineense	217881	Malwal Q2-3-36
Caudatum	#80	569563	2305	Guinea-caudatum	Caudatum-guineense	570387	Nolosungoti
Caudatum	#80	569596	–	Guinea-caudatum	Caudatum-guineense	570731	Sinidyil
Caudatum	#80	569612	–	Guinea-caudatum	Caudatum-guineense	571031	Korgi
Durra	Durra	152739	MN 868 (Mugud Yellow)	Guinea-caudatum	Caudatum-guineense	571209	Kigh
Durra	Durra	217840	Aklamoi White Q2–3-27	Guinea-caudatum	Caudatum-nigricans	569993	–
Durra	Durra	570768	Darfouri Brown	Guinea-caudatum	Caudatum-nigricans	569994	–
Durra	Durra	571221	Bhana	Guinea-caudatum	Nigricans-guineense	569323	WM 10642
Durra	Durra	571342	Wad Fahal	Guinea-caudatum	Nigricans-guineense	569413	Var-deri (White)
Durra	Caudatum-durra	569053	WM 10048	Guinea-caudatum	Nigricans-guineense	569415	Var-zeri
Durra	Caudatum-durra	570002	–	Guinea-caudatum	Nigricans-guineense	569418	Var-lotori
Durra	Caudatum-durra	570122	Typical Mayo	Guinea-caudatum	Nigricans-guineense	570711	Bari
Durra	Caudatum-durra	570267	–	Guinea-durra	Durra-membranaceum	568523	–
Durra	Caudatum-durra	571387	Korgi	Guinea-durra	Durra-membranaceum	568582	–
Durra	Cernuum	569411	Var-Makyika	Guinea-kafir	Guinea-kafir	569441	Var-bende
Durra	Dochna-durra	570431	–	Guinea-kafir	Roxburghii	570300	–
Durra	Membraneceum	152728	–	Kafir	Caffrorum-feterita	568632	–
Durra	Membraneceum	568309	–	Kafir	Caffrorum-feterita	570921	Hegiri Selfi
Durra	Membraneceum	568492	–	Kafir	Cadatum-kafir	570159	–
Durra	Membraneceum	568673	–	Kafir-caudatum	Caffrorum-birdproof	568536	–
Durra	Membraneceum	570280	–	Kafir-caudatum	Caffrorum-darso	562942	SBI 15
Durra	Nandyal	568509	–	Kafir-caudatum	Caudatum-kafir	569088	WM 10109
Durra	Nandyal	568657	–	Kafir-caudatum	Caudatum-kafir	569091	WM 10116
Durra	#90	570856	Hemaisi	Kafir-caudatum	Caudatum-kafir	569213	WM 10420
Durra-caudatum	Caudatum-durra	569030	WM 10021	Kafir-caudatum	Caudatum-kafir	570038	–
Durra-caudatum	Caudatum-durra	569045	WM 10040	Kafir-caudatum	Caudatum-kafir	570102	–
Durra-caudatum	Caudatum-durra	569046	WM 10041	Kafir-caudatum	#80	569522	Cross-36:122(122)
Durra-caudatum	Caudatum-durra	569180	WM 10320	Kafir-caudatum	#80	569673	–
Durra-caudatum	Caudatum-durra	569181	WM 10323	Kafir-caudatum	#80	569733	6041-WM 77-W-79
Durra-caudatum	Caudatum-kafir	569067	WM 10067	Kafir-caudatum	#80	569760	WM-77-W-308
Durra-caudatum	Membraneceum	571024	Sr 217	Kafir-durra	Durra-kafir	217831	Shalashali White Q2–3-24
Durra-caudatum	Nigricans-durra	568540	–	Kafir-durra	Durra-kafir	563494	*S. bicolor*
Durra-caudatum	Nigricans-durra	569004	A 218	Kafir-durra	Durra-kafir	568514	–
Guinea	Conspicuum	568505	–	Kafir-durra	Durra-kafir	569150	WM 10233
Guinea	Conspicuum	569002	Alnet	Kafir-durra	Durra-kafir	569306	WM 10618
Guinea	Conspicuum	569463	Var-akaram	Kafir-durra	Durra-kafir	570416	–
Guinea	Conspicuum	570388	Dahowgee-Ukilir	Shattercane	#90	569834	–
Guinea	Conspicuum	570687	Dura el Mak	Shattercane	#90	569847	–

ment, differ in allelic composition. As the equation indicates, it is a function of the number of alleles detected, their frequencies, and the sample size. It is equivalent to Nei's (1973, 1987): 'expected heterozygosity'

$$H_e = 1 - \sum_i p_i^2,$$

except for the important adjustment for finite sample size that is provided by $[N/(N-1)]$. δ_T^* values were derived from the 18 accessions listed in Table 1 for each SSR locus that was amplified with the primer set produced for it, and from the 190 accessions listed in Table 2 for a subset of these loci. The principal criteria used in selecting the latter loci were broad coverage of the genome and a minimum of three alleles among the 18 accessions listed in Table 1.

The probability, P, that two members of a population will have a different allele at at least one locus among the SSR loci examined was computed as:

$$P=1-[1-(\delta_T^*)_1][1-(\delta_T^*)_2]\ldots[1-(\delta_T^*)_N],$$

where $(\delta_T^*)_1,(\delta_T^*)_2\ldots(\delta_T^*)_N$ are the δ_T^* values for the N loci examined.

SSR criteria, locus symbols, and primer designations

For the purposes of this study, tri-, di-, and mono-nucleotide SSRs are defined as identical contiguous tri-, di-, and mono-nucleotides, respectively, that consist of a minimum of three, four, and seven units, respectively. SSR loci are designated with an italicized symbol consisting of a '*X*' (indicating that the function of the locus is unknown), the laboratory designator '*txp*', and an Arabic numeral. Primer-set designations consist of 'SbTXP' and the Arabic numeral that was assigned to the SSR locus that the primer set was designed to amplify.

Results

Isolation and characterization of SSRs

Fifty one clones that displayed a high signal level when probed with two radiolabeled dinucleotide 20-mers [$(AG)_{10}$ and $(AC)_{10}$] and two radiolabeled trinucleotide 21-mers [$(CCT)_7$ and $(ATT)_7$] were sequenced. Four of the clones were duplicates of other clones, one did not contain a SSR, and the other 46 contained one or more SSRs. Among recombinant (white) colonies, approximately 1 out of 2000–3000 contained a SSR.

Thirty eight of the sequenced clones contained one SSR and eight contained either two, three, or four SSRs. Thirty one of the former contained a $(AG/TC)_n$ SSR, three a $(AC/TG)_n$ SSR, one a $(AAT/TTA)_n$ SSR, and three a $(AGG/TCC)_n$ SSR. The characteristics of 30 of these 38 SSR loci are listed in Table 3. The other eight loci, in seven of which the repeating unit was AG/TC and in one AC/TG, are not included in the table because one of the two needed primers could not be designed for four of the SSRs (due to the close proximity of the SSR to one end of the clone) and the other four loci could not be amplified with the primer sets produced for them. All of the loci containing two or more SSRs are listed in Table 3 (note, however, that the *Xtxp25* primer set does not amplify one of the two SSRs that was present in the sequenced clone; see footnote #j of Table 3). Eleven $(AG/TC)_n$, five $(AC/TG)_n$, one $(AGG/TCC)_n$, two $(A/T)_n$, and one $(C/G)_n$ SSR were contained in these eight loci.

The number of repeated units per isolated SSR ranged from seven (*Xtxp22* and *Xtxp40*) to 50 (*Xtxp35*, which is not shown in Table 3 because it was not amplified by the primer set that was designed and produced for it). *Xtxp27*, an $(AG/CT)_{37}$ SSR, contains the largest number of repeated units among the amplified dinucleotide SSRs. The average number of units at the loci listed in Table 3 that contain a single dinucleotide SSR is 22.4, the average number at the six loci that contain two or more dinucleotide SSRs is 31.3, and the average number of units/SSR at the latter six loci is 12.5. The repeated unit in four of the five trinucleotide SSRs is AGG/GAA and each of the SSRs consists of either six or seven units. In contrast, *Xtxp30*, a AAT/TTA SSR, consists of 25 trinucleotides. Two of the three mononucleotide SSRs are composed of seven units and one of 12.

Primers flanking a SSR were designed and produced for 42 clones, and amplification of the targeted locus was accomplished with 38 of the primer sets. Thirty six of the primer sets amplified the SSR locus for which they were designed from the genomic DNA of both BTx623 and IS3620C and two amplified it from the genomic DNA of BTx623 only. As shown in Table 3, annealing temperatures of 55°C and 60°C were found suitable for 16 and 15 of the primer sets, respectively, and of 50°C and 45°C for six sets and one set, respectively.

Testing the potential utility of isolated SSRs

The potential utility of the aforementioned 38 SSR loci for genetic, breeding, and germplasm-analysis purposes was tested by determining the number and frequency of alleles at the loci in the 18 sorghum accessions listed in Table 1. Two loci were monomorphic among the 18 accessions, two were monomorphic among 17 of the 18 accessions, and from two to nine alleles were detected at the other 34 loci (Table 3). The average number of alleles per locus was 4.7 and the mean δ_T^* value for the 38 loci was 0.69.

Mapping of SSR loci

The size of the fragment amplified from genomic DNA was found to differ between BTx623 and IS3620C for 31 of the 38 SSR loci listed in Table 3 and the two forms of each fragment behaved as alleles. Segregation data for these loci were collected from the BTx623 × IS3620C RI population and all of the loci were placed on the framework RFLP map (Fig. 1; Table 3 lists the LG in which each SSR locus is located). Several of the loci were mapped in previously unmapped regions and four of them, *Xtxp40* and *Xtxp36* in LG E, *Xtxp6* in LG I, and *Xtxp23* in LG I, were mapped to a position distal to the most-distal RFLP locus. Two SSR loci mapped by Dr. D.

Table 3 Characteristics of 38 *S. bicolor* SSR loci and primer sets

Locus[a]	LG[b]	No. of alleles[c]	Type of SSR(s)[d]	Sequence of forward primer	Sequence of reverse primer	Size[e]	Ann. temp.
*Xtxp1**	B	5	$(AG)_{34}$	TTG GCT TTT GTG GAG CTG	ACC CAG CAG CAC TAC ACT AC	212	55
*Xtxp3**	B	7	$(CT)_8$+$(CT)_{36}$	AGC AGG CGT TTA TGG AAG	ATC CTC ATA CTG CAG GAC C	232	50
Xtxp4	B	6	$(GA)_{23}$	AAT ACT AGG TGT CAG GGC TGT G	ATG TAA CCG CAA CAA CCA AG	173	55
*Xtxp6**	I	9	$(CT)_{33}$	ATC GGA TCC GTC AGA TC	TCT AGG GAG GTT GCC AC	120	50
*Xtxp7**	B	4	$(CT)_{14}$	ACA TCT ACT ACC CTC TCA CC	ACA CAT CGA GAC CAG TTG	200	50
Xtxp8	B	6	$(TG)_{31}$	ATA TGG AAG GAA GAA GCC GG	AAC ACA ACA TGC ACG CAT G	148	60
Xtxp9	Unk[f, g]	4	$(TG)_{12}TT(TG)_{14}(AG)_{13}$	AAT AGC ACC GCC GCG CG	CAT TGT GGA GTC CCT GAT AC	160	55
Xtxp10	F	5	$(CT)_{14}$	ATA CTA TCA AGA GGG GAG C	AGT ACT AGC CAC ACG TCA C	145	50
Xtxp11	Unk[g]	1	$(AG)_{16}$	TCG AGA AAT TCA ACA TGC TG	GCT AGA CCG ACG AGA TAA G	82	55
*Xtxp12**	D	4	$(CT)_{22}$	AGA TCT GGC GGC AAC G	AGT CAC CCA TCG ATC ATC	193	55
Xtxp13	B	3	$(TG)_{13}$	TCT TTC CCA AGG AGC CTA G	GAA GTT ATG CCA GAC ATG CTG	120	55
*Xtxp14**	J	4	$(GA)_{15}$	GTA ATA GTC ATG ACC GAG G	TAA TAG ACG AGT GAA AGC CC	149	50
*Xtxp15**	J	3	$(TC)_{16}$	CAC AAA CAC TAG TGC CTT ATC	CAT AGA CAC CTA GGC CAT C	215	55
Xtxp16	Unk[h]	3	$(AG)_{35}$	TAG GGA AGA GCA AGT GCA GAC	AAG AAA GGG CCC AGA GTT TC	159	60
*Xtxp17**	I	3	$(TC)_{16}$+$(AG)_{12}$	CGG ACC AAC GAC GAT TAT C	ACT CGT CTC ACT GCA ATA CTG	164	55
*Xtxp18**	H	5	$(AG)_{21}$	ACT GTC TAG AAC AAG CTG CG	TTG CTC TAG CTA GGC ATT TC	231	55
Xtxp19	B	3	$(AG)_5$+$(AG)_{10}$	CTT TCA ATC GGT TCC AGA C	CTT CCA CCT CCG TAC TC	206	55
*Xtxp20**	G	5	$(AG)_{21}$	TCT CAA GGT TTG ATG GTT GG	ACC CAT TAT TGA CCG TTG AG	217	60
*Xtxp21**	D	5	$(AG)_{18}$	GAG CTG CCA TAG ATT TGG TCG	ACC TCG TCC CAC CTT TGT TG	179	60
Xtxp22	Unk[g]	1	$(AGG)_7$	TGC TAT GAT TGA GCC CAC TG	ATC ACA AGC AAA ACG CCA G	211	60
*Xtxp23**	J	3	$(CT)_{19}$	AAT CAA CAA GAG CGG GAA AG	TTG AGA TTC GCT CCA CTC C	182	60
*Xtxp24**	D	6	$(TC)_{21}$	TTG TGT AGT CCA TCC GAT GC	TTC TAA GCC CAC CGA AGT TG	145	60
Xtxp25	B	7	$(CT)_{12}$[j]	CCA TTG AGC TTC TGC TAT CTC	CAT TTG TCA CCA CTA GAA CCC	139	55
Xtxp26	Unk[h]	6	$(TG)_9(AG)_{12}$	AAG TGT AGT AGC AGT TTA GTC TC	TAG GTA TCA AAG GAC CAA GG	184	50
Xtxp27	D	5	$(AG)_{37}$	AAC CTT GCC CTA TCC ACC TC	TAT GAT GAA TCA AGG GAG AGG	332	45
Xtxp28	Unk[g]	2	$(GGA)_6$	TGT CGG CAT TGG CTA AAT AG	AAG CAA TGA CGG AGG TGG	123	60
Xtxp29	Unk[g]	3	$(GA)_{23}$	TAG GGC AGT GGT TAG TCG TG	TAC AAG TGG TGG TCC GAG G	338	60
Xtxp30	J	7	$(AAT)_{25}$	AAA AAG GAC GCG CAG CTG	CTG GTC TCC ACC ATC CGT AG	273	60
*Xtxp31**	C	5	$(CT)_{25}$	TGC GAG GCT GCC CTA CTA G	TGG ACG TAC CTA TTG GTG C	222	60
*Xtxp32**	A	5	$(AG)_{16}$	AGA AAT TCA CCA TGC TGC AG	ACC TCA CAG GCC ATG TCG	133	60
Xtxp33	C	8	$(TC)_{20}$ $C(TG)_5$+$(CT)_9CC(TG)_7$	GAG CTA CAC AGG GTT CAA C	CCT AGC TAT TCC TTG GTT G	221	55
Xtxp34	C	3	$(CT)_{29}$	TGG TTC GTA TCC TTC TCT ACA G	CAT ATA CCT CCT CGT CGC TC	365	55
Xtxp36	E	2[i]	$(GGA)_7GTA(T)_7$+$(A)_7$	ATG GGA CGG AAA TGC AGG AG	TTA TGC CTG CCA GCA ACT TG	185	60
*Xtxp37**	A	7	$(TC)_{23}$	AAC CTA AGA GGC CTA TTT AAC C	ACG GCG ACT ATG TAA CTC ATA G	189	55
Xtxp38	C	4	$(AG)_{17}$	ACA AAC CGC GAC GAA GTA AC	ACA AGG CAA AGC ACA AAG C	437	60
Xtxp40	E	2[i]	$(GGA)_7$	CAG CAA CTT GCA CTT GTC	GGG AGC AAT TTG GCA CTA G	138	55
Xtxp41	D	5	$(CT)_{19}$	TCT GGC CAT GAC TTA TCA C	AAA TGG CGT AGA CTC CCT TG	278	55
*Xtxp43**	A	7	$(CT)_{28}$	AGT CAC AGC ACA CTG CTT GTC	AAT TTA CCT GGC GCT CTG C	171	60

[a] Asterisks designate loci that were genotyped in the strains listed in Table 2
[b] LG = linkage group
[c] Number of alleles detected among the 18 strains listed in Table 1
[d] A '+' sign separates SSRs that are more than five bases apart
[e] Predicted number of bases in the BTx623 amplification product
[f] Unk = unknown
[g] No BTx623/IS3620C SSRP
[h] Null allele in IS3620C
[i] Seventeen of the eighteen strains listed in Table 1 are monomorphic for one allele
[j] The clone from which *Xtxp25* was isolated contained a $(G/C)_{12}$ SSR at a position distal to the site chosen for the reverse primer

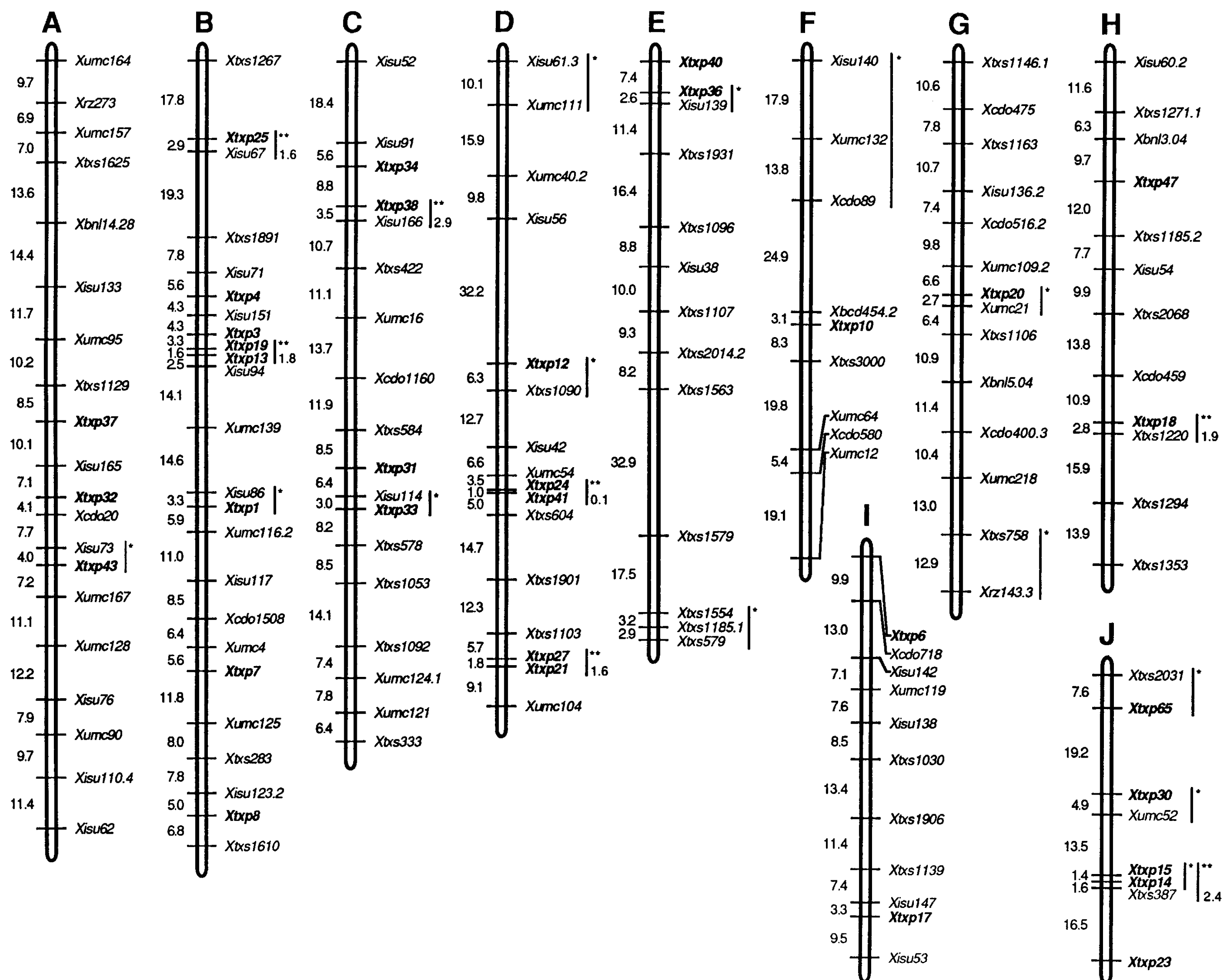

Fig. 1 Combined RFLP framework linkage map and SSR linkage map of the *S. bicolor* BTx623 × IS3620C recombinant inbred population. Symbols for SSR loci are in *bold face* and those for RFLP loci are in *plain text*. Distances between loci in cM are shown to the left of the linkage groups. All loci, except those with a *vertical line* to their right, were mapped at a LOD score > 5.0. *Double and single asterisks* designate loci mapped at LOD scores < 3.0 and between 3.0 and 5.0, respectively. The LOD scores for the former are shown *beside* the vertical lines

Bhattramakki (personal communication) that are located in what otherwise would be significant gaps in linkage groups H *(Xtxp47)* and J *(Xtxp65)* are also included on the map. Markers to fill the other gaps in the linkage groups are not available.

Relative to the RFLP loci on the framework map, 20 SSR loci were mapped at a LOD score ≥ 5.0 and eight at a LOD score between 3.0 and 5.0. Among these 28 loci, *Xtxp13* and *Xtxp19* (LG B), *Xtxp24* and *Xtxp41* (LG D), and *Xtxp21* and *Xtxp27* (LG D) were mapped at a LOD score < 3.0 relative to each other and ≥ 5.0 relative to all of the other markers. The other 22 loci were mapped at a LOD score ≥ 5.0 relative to all other SSR loci. Three of the remaining five loci, *Xtxp25* (LG B), *Xtxp38* (LG C) and *Xtxp18* (LG H), were ordered at a LOD score of 1.6 to 2.9 relative to one RFLP locus and ≥ 5.0 relative to all of the other markers; and the other two, *Xtxp14* and *Xtxp15* (LG J), were ordered at a LOD score between 3.0 and 5.0 relative to each other, of 2.4 relative to one RFLP locus, and ≥ 5.0 relative to all other markers.

Genotyping of 190 *S. bicolor* accessions at 17 SSR loci

In order to estimate the levels of allelic differentiation at SSR loci in *S. bicolor* sub-races and races, and in *S. bicolor* overall, the genotypes at 17-mapped SSR loci were assayed in the 190 accessions listed in Table 2. Sixteen of the seventeen loci contain one dinucleotide SSR and one locus, *Xtxp3,* contains two dinucleotide SSRs (see Table 3). The number of alleles detected in *S. bicolor* at each locus and the estimated levels of allelic differentiation for *S. bicolor* overall and for ten races of *S. bicolor*, as expressed by δ_T^* values, are shown in Table 4. δ_T^* values for ten working groups of race caudatum are shown in Table 5.

Table 4 Number of alleles at 17 SSR loci in *S. bicolor* and estimated levels of allelic differentation, δ_T^*, at the loci in *S. bicolor* and races of *S. bicolor*

SSRs	*S. bicolor*		*S. bicolor* races[a]									
	No. of alleles detected	– 190[b]	Caudatum 83	Durra 20	Bicolor 12	Guinea 12	Guinea-caudatum 12	Kafir-caudatum 11	Durra-caudatum 9	Bicolor-caudatum 6	Bicolor-durra 6	Kafir-durra 6
Xtxp1	20	0.92	0.89	0.87	0.92	0.67	0.88	0.96	0.97	0.93	0.80	1.00
Xtxp3	22	0.91	0.94	0.91	0.85	0.94	0.85	0.82	0.83	0.80	0.73	0.80
Xtxp6	22	0.94	0.93	0.85	0.97	0.94	0.91	0.98	0.97	–	0.93	1.00
Xtxp7	14	0.87	0.80	0.85	0.86	0.79	0.76	0.89	0.84	0.87	0.93	0.94
Xtxp12	20	0.91	0.90	0.87	1.00	0.91	0.91	0.93	0.97	–	0.73	0.80
Xtxp14	12	0.81	0.80	0.77	0.89	0.56	0.80	0.84	0.81	0.60	1.00	1.00
Xtxp15	9	0.86	0.80	0.90	0.77	0.88	0.88	0.87	0.92	–	0.93	0.93
Xtxp17	12	0.83	0.80	0.85	0.79	0.83	**0.30**	**0.40**	0.81	0.73	0.80	**0.00**
Xtxp18	15	0.88	0.83	0.92	0.88	0.92	0.80	0.93	0.89	0.93	0.93	0.93
Xtxp20	15	0.88	0.88	0.81	0.80	0.93	0.90	0.91	0.92	0.87	0.93	0.80
Xtxp21	15	0.84	0.85	0.83	0.85	0.80	0.86	0.89	0.58	0.87	0.80	0.60
Xtxp23	15	0.89	0.81	0.95	0.91	0.83	0.89	0.86	–	–	–	1.00
Xtxp24	19	0.92	0.92	0.87	0.94	0.92	0.85	0.93	0.92	0.93	–	1.00
Xtxp31	18	0.91	0.86	0.94	0.89	0.88	0.95	0.89	0.94	1.00	0.93	0.93
Xtxp32	20	0.89	0.81	0.94	0.94	0.93	0.93	0.89	0.86	0.93	–	0.87
Xtxp37	16	0.86	0.79	0.90	0.80	0.58	0.67	0.87	0.89	0.80	0.93	0.93
Xtxp43	19	0.94	0.94	0.92	0.94	0.89	0.92	0.85	0.89	0.93	0.93	1.00
Mean δ_T^*		0.89	0.86	0.88	0.88	0.84	0.83	0.87	0.88	0.86	0.88	0.85

[a] The criteria for including a δ_T^* value for a race in the table were the analysis of six or more accessions, and the obtaining of data from five or more accessions and from at least 70% of the accessions analyzed. δ_T^* values < 0.69 and ≥0.50 are underlined and those < 0.50 are in bold type

[b] Number of accessions studied

Table 5 Estimated levels of allelic differentiation, δ_T^*, at 17 SSR loci in working groups (sub-races) of *S. bicolor* race caudatum

SSRs	Working groups[a]										Race Caudatum
	Caudatum 20[b]	Caudatum-durra 11	Caudatum-guinea 7	Caudatum-kafir 7	Zerazera 7	Caudatum-nigricans 6	Durra-nigricans 5	Nigricans 5	Nigricans-feterita 5	Number 80 5	– 83
Xtxp1	0.82	0.93	0.95	0.95	0.95	1.00	0.70	0.90	0.90	–	0.89
Xtxp3	0.92	0.95	0.95	0.88	0.86	1.00	1.00	0.90	1.00	0.90	0.94
Xtxp6	0.93	0.95	–	0.90	–	0.80	0.90	0.90	1.00	–	0.93
Xtxp7	0.79	0.61	0.71	0.67	0.90	–	0.80	–	0.90	0.90	0.80
Xtxp12	0.92	0.96	–	0.86	0.90	–	0.90	0.70	1.00	0.70	0.90
Xtxp14	0.79	0.82	0.86	0.90	0.67	0.80	0.70	**0.40**	0.90	**0.00**	0.80
Xtxp15	0.72	0.78	0.52	0.67	**0.48**	**0.40**	–	0.70	0.90	0.90	0.80
Xtxp17	0.68	0.73	0.81	0.90	**0.00**	0.73	–	0.90	0.90	0.90	0.80
Xtxp18	0.85	0.89	0.72	0.90	–	–	1.00	0.70	0.80	–	0.83
Xtxp20	0.89	0.80	–	0.90	0.81	1.00	–	–	0.70	**0.00**	0.88
Xtxp21	0.85	0.87	0.81	0.86	0.52	0.93	0.90	0.90	0.90	0.70	0.85
Xtxp23	0.79	0.78	0.90	0.90	0.86	0.87	0.90	0.60	0.70	0.90	0.81
Xtxp24	0.94	0.77	1.00	0.86	0.90	–	0.90	1.00	0.70	0.60	0.92
Xtxp31	0.81	0.88	0.81	0.90	0.95	0.93	1.00	0.90	0.70	**0.40**	0.86
Xtxp32	0.68	0.93	0.52	0.67	0.95	0.87	–	1.00	0.70	0.70	0.81
Xtxp37	0.70	0.89	0.71	0.81	**0.48**	0.60	0.70	**0.40**	0.70	0.90	0.79
Xtxp43	0.87	0.91	0.90	0.91	0.85	1.00	0.90	0.90	1.00	0.90	0.94
Mean δ_T^*	0.82	0.85	0.80	0.85	0.74	0.84	0.87	0.79	0.85	*0.67*	0.86

[a] The criteria for including a δ_T^* value for a working group in the table were the obtaining of data from five or more accessions and from 70% or more of the accessions analyed. δ_T^* values < 0.69 and ≥ 0.50 are underlined and those < 0.50 are in bold type. For comparative purposes, the δ_T^* values for each SSR locus in race caudatum are shown in the far right-hand column (see also Table 4)

[b] Number of accessions studied

Two-hundred and eighty three alleles were identified in *S. bicolor* at the 17 loci (Table 4), an average of 16.6 alleles per locus. The number of alleles detected per locus ranged from nine (*Xtxp15*) to 22 (*Xtxp3* and *Xtxp6*), with fewer than 15 alleles detected at only four loci. Two-hundred and thirty four (82.7%) of the 283 alleles found in *S. bicolor* were detected in race caudatum (83 accessions examined) and 122 (43.1%) of the alleles were detected in the working-group caudatum (20 accessions examined) of race caudatum. The average frequen-

cies of the two most-common alleles were 0.24 and 0.17, and the range of frequencies for the two most-common alleles combined was from 0.57 for the two most common *Xtxp14* alleles to 0.23 for the two most-common *Xtxp43* alleles. The number of alleles per locus is positively correlated (correlation coefficient = 0.68, which is significant at the 1% level) with the number of repeated units at the loci in BTx623, the strain from which the SSRs were originally isolated (data not shown).

The mean δ_T^* value for *S. bicolor*, 0.89, is only marginally larger than the mean δ_T^* values for the races listed in Table 4. These range from 0.88 to 0.83 and have a mean of 0.86. Among the races, only 10 (6.2%) of the 162 δ_T^* values for individual SSR loci are <0.69 (data are missing for eight locus/race combinations) and only one δ_T^* value of 0.00 was detected (for *Xtxp17* in kafir-durra). Thus, while the estimated levels of differentiation at some SSR loci vary considerably among the races, the overall levels differ by only small amounts.

The mean δ_T^* values for the ten working groups of race caudatum listed in Table 5 range from 0.88 to 0.67, and the overall mean δ_T^* for the working groups is 0.81. The average δ_T^* values for five of the working groups, namely, caudatum-durra (0.85), caudatum-kafir (0.85), caudatum-nigricans (0.84), durra-nigricans (0.87), and nigricans-feterita (0.85), are closely similar to that of race caudatum (0.86) and are within the range of 0.88 to 0.83 observed for the other races, while the δ_T^* values for three of the other five groups, namely, caudatum (0.82), caudatum-guinea (0.80) and nigricans (0.79), are only slightly below this range. The δ_T^* values of 0.74 and 0.67 for zerazera and number 80, respectively, are markedly lower, however. δ_T^* values < 0.69 were detected for 13.3% (20) of the 150 locus/working-group combinations (data are missing for 20 locus/working-group combinations), more than twice the frequency that was detected at the racial level. Nine of the twenty low δ_T^* values are for zerazera and number-80 loci. It is important to note, however, that a $\delta_T^*<0.80$ was not detected in any working group for three loci, *Xtxp3*, *6* and *43*) and a δ_T^* <0.70 was not detected in any working group for three other loci (*Xtxp1*, *12* and *18*). The significance of these findings with regard to the use of sorghum SSRs for fingerprinting is discussed below.

Discussion

Isolation and characterization of SSRs

Although the size-fractionated DNA library produced in this study was probed with four radiolabeled oligomers, namely, $(AG)_{10}$, $(AC)_{10}$, $(CCT)_7$ and $(ATT)_7$ oligomers, 42 (72%) of the 58 SSRs that were isolated contained AG/TC as the repeating unit. Next highest in frequency were AC/TG (eight, 14%) and AGG/TCC (four, 7%) SSRs. Only one AAT/TTA SSR was isolated. In addition, two A/T SSRs and one C/G SSR, present in clones that contained another SSR, were isolated. AG/TC repeats also predominated among the SSRs isolated by Brown et al. (1996), who probed a size-fractionated sorghum genomic DNA library with a radiolabeled mixture containing two dinucleotides, including $(AG)_{10}$, two trinucleotides and one tetranucleotide.

Thirty six (70.6%) of the fifty one sorghum SSR loci that were sequenced were amplified from genomic DNA of both of the parents of the RI mapping population with the primer sets that were designed and produced for them, and 31 (60.8%) of the 51 loci were genetically mapped. These percentages, which are markedly higher than reported in previous studies of plant species, including the study of Brown et al. (1996) on sorghum, indicate that conditions can be defined whereby highly efficient isolation of amplifiable SSRs from a size-fractionated genomic DNA library can be accomplished.

Distribution of SSR loci in the sorghum genome

The number of SSR loci mapped in this study in linkage groups A through J is 3, 8, 4, 5, 2, 1, 1, 1, 2 and 4, respectively. This is an insufficient number of loci to allow a definitive conclusion to be made regarding whether or not SSR loci are randomly distributed in the *S. bicolor* genome. The distribution of the mapped SSR loci is suggestive of clustering, however, because the members of four pairs of loci are located 1.8 cM or less from one another, and the ratio between the number of SSR loci located in a LG and the cM length of the LG is 1:14 for LG J and 1:22 for LG B (not counting *Xtxp65*) but only 1:116 for LGs F, G, and H combined (not counting *Xtxp47* in LG H).

Utility of SSRs for fingerprinting and marker-assisted selection

Among the criteria for genetic markers that are to be used for fingerprinting and marker-assisted selection is a high level of polymorphism. Clearly, sorghum SSRs meet this criterion. The δ_T^* values for the 17 SSR loci that were genotyped in 190 *S. bicolor* accessions range from 0.83 to 0.94 (Table 4) and they have a mean of 0.89. Even *Xtxp15*, the locus at which the fewest alleles were detected (nine), has a δ_T^* of 0.86. The significance of these and the other δ_T^* values reported in this paper for fingerprinting purposes becomes apparent when the probability of distinguishing accessions using two or more loci is considered. Both *Xtxp6* and *Xtxp43* have a δ_T^* of 0.94 and both *Xtxp1* and *Xtxp24* have a δ_T^* of 0.92. Consequently, the probability that two *S. bicolor* accessions chosen at random will have a different allele at at least one of any two of these four loci is >0.99 [$P = 1-(1-0.92)^2$]. Also, the probability that two accessions will have a different allele at at least one of any three of the loci is >0.999. Furthermore, the three lowest *S. bicolor* δ_T^* values are 0.81 (*Xtxp14*), 0.83 (*Xtxp17*), and 0.84 (*Xtxp21*), and the probability that two *S. bicolor*

accessions chosen at random will have different alleles at at least one of these three loci is approximately 0.995.

As noted in the Results section, the mean δ_T^* values for the ten working groups of race caudatum listed in Table 5 range from 0.88 to 0.67, and the overall mean δ_T^* for the working groups is 0.81. These values indicate that, for the SSR loci studied, the average probability that two accessions selected at random from a working group will have different alleles at a SSR locus ranges from 0.88 to 0.67, depending upon the working group to which the accessions belong. Average δ_T^* values markedly lower than those of the ten races that were examined were detected for only two of the working groups, namely, zerazera and number 80 (Tables 4 and 5). Furthermore, a δ_T^* value <0.69 was not detected for *Xtxp1*, *3*, *6*, *12*, *18* and 43 in these two or any of the other working groups. The significance of this is that, given δ_T^* values of only 0.69 at several SSR loci, the probability that two accessions chosen at random will have different alleles at one or more loci is approximately 0.99 when four loci are genotyped and approximately 0.999 when six loci are genotyped. Indeed, the lowest δ_T^* values for *Xtxp3*, *6*, and *43* are 0.86, 0.80, and 0.85, respectively. Consequently, the probability that two accessions chosen at random from a working group of race caudatum will have different alleles at one or more of these three loci is ≥ 0.996. Furthermore, the lowest δ_T^* values for *Xtxp1*, *12*, and 18 are 0.70, 0.70 and 0.72, respectively, and thus the probability that one or more of the six loci will have different alleles is ≥ 0.9999.

Also to be noted is that the number of *Xtxp1*, *3*, *6*, *12*, *18* and *43* alleles detected in race caudatum (17, 18, 18, 15, 16 and 17, respectively; data not shown) provides an enormous number of potential genotypic combinations at the six loci; more than 22 million. (Linkage has been disregarded in these calculations because, with the exception of *Xtxp1* and *Xtxp3*, the six loci are unlinked, and the *Xtxp1* and *Xtxp6* linkage is 39 cM.) In addition, it should be noted that allelic differences *will* be present *at most* of the loci examined when two accessions are compared in the manner just described. For example, comparison of the *Xtxp3*, *6*, *12*, and *43* genotypes for all of the 190 possible pairwise combinations of the 20 accessions of group caudatum disclosed allelic differences at either three or four of the loci for 149 (78.4%) of the combinations, at two loci for 40 (21.1%) of the combinations, and at one locus (0.5%) for one combination (furthermore, genotypes were available for both members of the latter combination for only three of the four loci; data not shown). Based on the group-caudatum δ_T^* values for these loci, the probability that two group-caudatum accessions chosen at random will differ at at least one of the four loci is > 0.9999.

A further consideration is the degree to which findings derived from the working groups of one race can be applied to the working groups of other races. Given that the average δ_T^* for the 17 SSR loci in race caudatum, 0.86, is in the middle of the range of the δ_T^* values for the ten races studied (see Table 4), it seems likely that the findings are broadly applicable.

It was noted in the Results section that the δ_T^* values for SSR loci for the working groups of race caudatum have a mean of 0.81, i.e., that, for the 17 SSR loci studied, the estimated average probability that two accessions in a working group will have different alleles at a locus is 0.81. Clearly then, most sorghum SSR loci are sufficiently polymorphic to be useful in marker-assisted selection programs. The analyses presented in the paragraphs above indicate that the genotypes at as few as three or four SSR loci are capable of distinguishing most accessions from one another in *S. bicolor* working groups, even in working groups that are at the lower end of the differentiation scale. A range of degrees of genetic differentiation at SSR loci is to be expected among the members of working groups, of course, and a larger number of SSR loci should be genotyped in order to increase the probability of distinguishing closely related accessions. Also, fingerprinting of germplasm accessions is most efficiently performed using a high-throughput multiplex system wherein several SSR loci are analyzed with a single reaction and in a single lane of an electrophoresis gel (Brown et al. 1996; Dean et al. 1999). Attributes desirable for such a system include primer pairs that anneal at the same temperature, primer sequences that lack significant homology with each other and with the fragments produced by amplification, and amplified fragments that differ significantly in size. The group of SSR loci composed of *Xtxp1*, *3*, *6*, *12* and *18* possess some, but not all, of these attributes, and isolation, characterization, and mapping of other sorghum SSR loci may be required in order to develop a multiplex system for fingerprinting sorghum accessions at loci that are highly polymorphic. On the other hand, electrophoresing the pooled products of several uniplex reactions in a single lane of an electrophoresis gel can also be highly efficient, and developing a system to accomplish this may be considerably easier than developing a full multiplex system. In any case, the fundamental point remains that the results of this study indicate that the vast majority of *S. bicolor* accessions, even closely related accessions that are members of the same working group, can be distinguished from one another by determining the genotypes at a small number, perhaps as few as a half-dozen, SSR loci.

Acknowledgments We are grateful to Drs. J. Dahlberg, F.R. Miller, and K.F. Schertz for supplying us with seed stocks of the sorghum lines used in this investigation. This research was supported by Texas Agricultural Experiment Station and by grants from the United States Department of Agriculture National Research Initiative Competitive Research Grants Program and the Rockefeller Foundation.

References

Bennetzen JL, Subramanian V, Xu J, Salimath SS, Subramanian S, Bhattramakki D, Hart GE (2000) A framework genetic map of sorghum containing RFLP, SSR and morphological markers. In: Phillips RL, Vasil IK (eds) DNA-based markers in plants, 2nd edn. Kluwer Academic Publishers, Boston, (in press)

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

Binelli GL, Gianfranceschi L, Pe ME, Taramino G, Busso C, Stenhouse J, Ottaviano E (1992) Similarity of maize and sorghum genomes as revealed by maize RFLP probes. Theor Appl Genet 84:10–16

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

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

Cho YG, McCouch SR, Kuiper M, Kang M-R, Pot J, Groenen JTM, Eun MY (1998) Integrated map of AFLP, SSLP and RFLP markers using a recombinant inbred population of rice (*Oryza sativa* L.). Theor Appl Genet 97:370–380

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An intergrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE, Kresovich S (1999) Genetic redundancy and diversity among “Orange” accessions in the US National Sorghum Collection as assessed with simple sequence repeat (SSR) markers. Crop Sci 39: 1215–1221

Doggett H 1988. Sorghum, 2nd edn. John Wiley and Sons, Inc., New York

Dong J, Mullet JE, Hart GE (1996) Identification and characterization of *Sorghum bicolor* (L.) Moench simple-sequence-repeat loci. Plant Genome IV Conference Program and Abstracts, p 46

Dufour L, Grivet A, Deu M, Trouche G, Glaszmann JC, Hamon P (1996) Comparative genetic mapping between duplicated segments of maize chromosomes 3 and 8 and homoeologous regions in sorghum and sugarcane. Theor Appl Genet 92: 1024–1030

Dufour L, Deu M, Grivet A, D’Hont A, Paulet F, Bouet A, Lanaud D, Glaszmann JC, Hamon P (1997) Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid. Theor Appl Genet 94:409–418

Fritz AK, Caldwell S, Worrall WD (1999) Molecular mapping of Russian wheat aphid resistance from triticale accession PI 386156. Crop Sci 39: 1707–1710

Gregorious H-R (1987) The relationship between the concepts of genetic diversity and differentiation. Theor Appl Genet 74: 397–401

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

Kong L, Dong J, Hart GE (1997) Linkage-mapping of simple-sequence repeat loci in *Sorghum bicolor* L. Moench. Plant Genome V Conference Program and Abstracts, p 100

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

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

Murray MG, Thompson WF (1980) Rapid isolation of high-molecular-weight plant DNA. Nucleic Acids Res 8:4321–4325

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Peng Y, Schertz KF, Cartinhour S, Hart GE (1999) Comparative genome mapping of *Sorghum bicolor* (L.) Moench using an RFLP map constructed in a population of recombinant inbred lines. Plant Breed 118:225–235

Pereira MG, Lee M, Bramel-Cox P, Woodman W, Doebley J, Whitkus R (1994) Construction of an RFLP map in sorghum and comparative mapping in maize. Genome 37:236–243

Ragab RA, Dronavalli S, Saghai Maroof MA, Yu YG (1994) Construction of a sorghum RFLP linkage map using sorghum and maize DNA probes. Genome 37:590–594

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Rongwen J, Akkaya MS, Bhagwar AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90:43–48

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

Saghai Maroof MA, Biyashev RB, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc Natl Acad Sci USA 91:5466–5470

Taramino G, Tarchini R, Ferrario S, Lee M, Pe ME (1997) Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor Appl Genet 95:66–72

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem repeats. Theor Appl Genet 88:1–6

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

Yang GK, Saghai Maroof MA, Xu CG, Zhang Q, Biyashev RM (1995) Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol Gen Genet 245:187–194

Xu GW, Magill CW, Schertz KF, Hart GE (1994) An RFLP linkage map of *Sorghum bicolor* (L.) Moench. Theor Appl Genet 89:139–145