American Journal of Botany

American Journal of Botany: e245–e250. 2012.

AJB Primer Notes & Protocols in the Plant Sciences

A reference microsatellite kit to assess for genetic diversity of *Sorghum bicolor* (Poaceae)[1]

Claire Billot[2,6,7], Ronan Rivallan[2,7], Mbaye Ndoye Sall[3,7], Daniel Fonceka[2], Monique Deu[2], Jean-Christophe Glaszmann[2], Jean-Louis Noyer[2], Jean-François Rami[2], Ange-Marie Risterucci[2], Patrick Wincker[4], Puna Ramu[5], and Charles T. Hash[5]

[2]CIRAD, UMR AGAP, F-34398 Montpellier, France; [3]CERAAS, Thies Escale, BP 3320, Thies, Senegal; [4]Genoscope, 2 rue Gaston, Cremieux 91057 Evry, France; and [5]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad 502 324, Andhra Pradesh, India

- *Premise of the study:* Discrepancies in terms of genotyping data are frequently observed when comparing simple sequence repeat (SSR) data sets across genotyping technologies and laboratories. This technical concern introduces biases that hamper any synthetic studies or comparison of genetic diversity between collections. To prevent this for *Sorghum bicolor*, we developed a control kit of 48 SSR markers.
- *Methods and Results:* One hundred seventeen markers were selected along the genome to provide coverage across the length of all 10 sorghum linkage groups. They were tested for polymorphism and reproducibility across two laboratories (Centre de Cooperation Internationale en Recherche Agronomique pour le Developpement [CIRAD], France, and International Crops Research Institute for the Semi-Arid Tropics [ICRISAT], India) using two commonly used genotyping technologies (polyacrylamide gel–based technology with LI-COR sequencing machines and capillary systems with ABI sequencing apparatus) with DNA samples from a diverse set of 48 *S. bicolor* accessions.
- *Conclusions:* A kit for diversity analysis (http://sat.cirad.fr/sat/sorghum_SSR_kit/) was developed. It contains information on 48 technically robust sorghum microsatellite markers and 10 DNA controls. It can further be used to calibrate sorghum SSR genotyping data acquired with different technologies and compare those to genetic diversity references.

Key words: diversity kit; Poaceae; *Sorghum bicolor*; standardization; simple sequence repeat markers.

Combining simple sequence repeat (SSR) molecular data obtained from different laboratories or different genotyping platforms is often a challenge, especially for large-scale analyses and for highly polymorphic markers (for example, George et al., 2004). Although work can be partitioned with markers assigned only to one platform in one laboratory, this is rarely the case. Thus, one way to avoid troubles in merging data sets, and provide standardization, is to use similar controls as allelic references.

Sorghum (*Sorghum bicolor* (L.) Moench) is a major staple crop in semiarid environments, traditionally used for forage and food consumption. Five major races (bicolor, caudatum, durra, guinea, and kafir) and 10 intermediates were determined using panicle and spikelet morphologies and were confirmed by genetic markers (Deu et al., 2006). This crop is gaining attention in a context of climatic changes for its adaptation to water scarcity and its uses as biofuel. It is a reference for C_4 metabolism plants because the sequencing of its genome is now available (Paterson et al., 2009; http://genome.jgi-psf.org/Sorbi1/Sorbi1.info.html). Although directly sequence-derived markers may be routinely used, multiallelic markers such as SSR will continue to be a reference, especially for small-scale diversity studies. We present here a work on designing a robust technical reference SSR kit for sorghum from previously published and unpublished SSR markers to be used to combine information from different studies.

METHODS AND RESULTS

Forty-eight sorghum accessions (Appendix 1) were selected among a core collection of cultivated sorghum (Deu et al., 2006) to represent all races and geographic origins. DNA extraction from a single representative plant was carried out at the Centre de Cooperation Internationale en Recherche Agronomique pour le Developpement (CIRAD; France) following Deu et al. (2006). An aliquot quantity of DNA samples was sent to the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT; India) for SSR genotyping.

One hundred seventeen SSR markers (http://sat.cirad.fr/sat/sorghum_SSR_kit/) were selected among published markers (Brown et al., 1996; Taramino et al., 1997; Bhattramakki et al., 2000; Kong et al., 2000; Schloss et al., 2002; Ramu et al., 2009; Mutegi et al., 2011) and specifically developed markers. These were obtained from $(GA)_n$, $(GT)_n$, or $(CCG)_n$ microsatellite-enriched libraries, following a hybridization-based method in which DNA was previously digested with *Pst*1 or *Rsa*1 as described in Billotte et al. (2005). Seven hundred

[1]Manuscript received 17 November 2011; revision accepted 6 April 2012.

New SSR development was performed as part of a 2000 Genoscope National Sequencing Project. Kit identification was performed in association with the Generation Challenge Program (GCP; Project 2005-01—Genotyping of a Sorghum Composite Collection). M.N.S. and D.F. were supported by a grant of the GCP under the GCP Subprogram 5. All kit-characterizing experiments were performed on the Montpellier Languedoc-Roussillon Genotyping platform (GPTR, http://www.gptr-lr-genotypage.com/), Montpellier, France.

[6]Author for correspondence: claire.billot@cirad.fr

[7]These authors contributed equally to the work.

doi:10.3732/ajb.1100548

TABLE 1. Characteristics of the 48 SSR markers used in the *Sorghum bicolor* technical diversity kit.

Marker[a]	Primer sequences (5′–3′)	Repeat motif	GenBank accession no.	Chr[b]	Position[c]	T_a (°C)
gpsb067[1]	F: TAGTCCATACACCTTTCA R: TCTCTCACACACATTCTTC	$(GT)_{10}$	JQ031014	8	13744996	49
gpsb069[1]	F: CCCATAATACTTGACCTTC R: ACTTACTCCCTCTGTCCC	$(TC)_{12}$	JQ031015	6	—	50
gpsb089[1]	F: ATCAGGTACAGCAGGTAGG R: ATGCATCATGGCTGGT	$(TG)_9$	JQ031016	1	43884079	50
gpsb123[2]	F: ATAGATGTTGACGAAGCA R: GTGGTATGGGACTGGA	$(CA)_7(GA)_5$	JQ031017	8	52281926	50
gpsb148[1]	F: CAACCACAAACCAAGAG R: ATAGAAATGGGGTGGAG	$(TC)_3(CA)_5$	JQ031018	7	327117	50
gpsb151[1]	F: ATACCAAGTTTCCTTTACCT R: GTTGGGGGAGAGTTTT	$(CT)_{12}$	JQ031019	4	2048028	50
mSbCIR223[1]	F: CGTTCCAATGACTTTTCTTC R: GCCAATGTGGTGTGATAAAT	$(AC)_6$	JQ031020	2	4657470	55
mSbCIR238[2]	F: AGAAGAAAAGGGGTAAGAGC R: CGAGAAACAATTACATGAACC	$(AC)_{26}$	JQ031021	2	14746509	55
mSbCIR240[2]	F: GTTCTTGGCCCTACTGAAT R: TCACCTGTAACCCTGTCTTC	$(TG)_9$	JQ031022	8	4467743	55
mSbCIR246[2]	F: TTTTGTTGCACTTTTGAGC R: GATGATAGCGACCACAAATC	$(CA)_7$	JQ031023	7	56279794	55
mSbCIR248[2]	F: GTTGGTCAGTGGTGGATAAA R: ACTCCCATGTGCTGAATCT	$(GT)_7$	JQ031024	5	4746082	56
mSbCIR262[2]	F: GCACCAAAATCAGCGTCT R: CCATTTACCCGTGGATTAGT	$(CATG)_3$	JQ031025	10	55324102	57
mSbCIR276[2]	F: CCCCAATCTAACTATTTGGT R: GAGGCTGAGATGCTCTGT	$(AC)_9$	JQ031026	3	55555298	53
mSbCIR283[1]	F: TCCCTTCTGAGCTTGTAAAT R: CAAGTCACTACCAAATGCAC	$(CT)_8(GT)_8$	JQ031027	10	18099884	54
mSbCIR286[1]	F: GCTTCTATACTCCCCTCCAC R: TTTATGGTAGGATGCTCTGC	$(AC)_9$	JQ031028	1	57452822	55
mSbCIR300[2]	F: TTGAGAGCGGCGAGGTAA R: AAAAGCCCAAGTCTCAGTGCTA	$(GT)_9$	JQ031029	7	58286012	61
mSbCIR306[1]	F: ACATGGGGAGGAAGATGA R: GCTATTCAGGAGCCATGC	$(CATG)_3(GT)_7$	JQ031030	1	—	56
mSbCIR329[1]	F: GATCTTCACCAGGAACAGG R: ATGAGAGGAAAACATTGCTG	$(AC)_9$	JQ031031	5	1763243	55
sb4-72[3]	F: TGCCACCACTCTGGAAAAGGCTA R: CTGAGGACTGCCCCAAATGTAGG	$(AG)_{16}$	NA	6	41440341	55
sb5-206[3]	F: ATTCATCATCCTCATCCTCGTAGAA R: AAAAACCAACCCGACCCACTC	$(AC)_{13}(AG)_{20}$	NA	9	59162303	55
sb6-84[3]	F: CGCTCTCGGGATGAATGA R: TAACGGACCACTAACAAATGATT	$(AG)_{14}$	NA	2	—	55
SbAGB02[4]	F: CTCTGATATGTCGTTGTGCT R: ATAGAGAGGATAGCTTATAGCTCA	$(AG)_{35}$	NA	7	62506705	55
Xcup02[5]	F: GACGCAGCTTTGCTCCTATC R: GTCCAACCAACCCACGTATC	$(GCA)_6$	pSB0069	9	8143767	54
Xcup11[5]	F: TACCGCCATGTCATCATCAG R: CGTATCGCAAGCTGTGTTTG	$(GCTA)_4$	pSB1889	3	1992880	54
Xcup14[5]	F: TACATCACAGCAGGGACAGG R: CTGGAAAGCCGAGCAGTATG	$(AG)_{10}$	pSB1802	3	72459931	54
Xcup53[5]	F: GCAGGAGTATAGGCAGAGGC R: CGACATGACAAGCTCAAACG	$(TTTA)_5$	pSB0508	1	72905425	54
Xcup61[5]	F: TTAGCATGTCCACCACAACC R: AAAGCAACTCGTCTGATCCC	$(CAG)_7$	pSB0581	3	2576698	54
Xcup62[5]	F: CGAGAAGATCGAGAGAACCC R: TGAAGACGACGACGACAGAC	$(GAA)_6$	pSB0600	1	68743248	54
Xcup63[5]	F: GTAAAGGGCAAGGCAACAAG R: GCCCTACAAAATCTGCAAGC	$(GGATGC)_4$	pSB0600	2	59104626	54
Xisep0107[6]	F: GCCGTAACAGAGAAGGATGG R: TTTCCGCTACCTCAAAAACC	$(TGG)_4$	AW744864	3	3209015	59
Xisep0310[6]	F: TGCCTTGTGCCTTGTTTATCT R: GGATCGATGCCTATCTCGTC	$(CCAAT)_4$	AW286133	2	77623085	60
Xtxp10[7]	F: ATACTATCAAGAGGGGAGC R: AGTACTAGCCACACGTCAC	$(CT)_{14}$	NA	9	47916807	50
Xtxp12[7]	F: AGATCTGGCGGCAACG R: AGTCACCCATCGATCATC	$(CT)_{22}$	NA	4	48576873	55
Xtxp15[7]	F: CACAAACACTAGTGCCTTATC R: CATAGACACCTAGGCCATC	$(TC)_{16}$	NA	5	42049815	55

Table 1. Continued.

Marker[a]	Primer sequences (5′–3′)	Repeat motif	GenBank accession no.	Chr[b]	Position[c]	T_a (°C)
Xtxp21[7]	F: GAGCTGCCATAGATTTGGTCG R: ACCTCGTCCCACCTTTGTTG	$(AG)_{18}$	NA	4	67961876	60
Xtxp40[7]	F: CAGCAACTTGCACTTGTC R: GGGAGCAATTTGGCACTAG	$(GGA)_7$	NA	7	860855	55
Xtxp57[7]	F: GGAACTTTTGACGGGTAGTGC R: CGATCGTGATGTCCCAATC	$(GT)_{21}$	NA	6	57418801	55
Xtxp114[7]	F: CGTCTTCTACCGCGTCCT R: CATAATCCCACTCAACAATCC	$(AGG)_8$	NA	3	60794263	50
Xtxp136[7]	F: GCGAATAGCATCTTACAACA R: ACTGATCATTGGCAGGAC	$(GCA)_5$	NA	5	—	55
Xtxp141[7]	F: TGTATGGCCTAGCTTATCT R: CAACAAGCCAACCTAAA	$(GA)_{23}$	NA	10	58245266	55
Xtxp145[7]	F: GTTCCTCCTGCCATTACT R: CTTCCGCACATCCAC	$(AG)_{22}$	NA	6	—	55
Xtxp265[7]	F: GTCTACAGGCGTGCAAATAAAA R: TTACCATGCTACCCCTAAAAGTGG	$(GAA)_{19}$	NA	6	51179303	55
Xtxp273[7]	F: GTACCCATTTAAATTGTTTGCAGTAG R: CAGAGGAGGAGGAAGAGAAGG	$(TTG)_{20}$	NA	8	156965	55
Xtxp278[7]	F: GGGTTTCAACTCTAGCCTACCGAACTTCCT R: ATGCCTCATCATGGTTCGTTTTGCTT	$(TTG)_{12}$	NA	7	51120645	50
Xtxp295[7]	F: AAATCATGCATCCATGTTCGTCTTC R: CTCCCGCTACAAGAGTACATTCATAGCTTA	$(TC)_{19}$	NA	7	61172112	55
Xtxp320[8]	F: TAAACTAGACCATATACTGCCATGATAA R: GTGCAAATAAGGGCTAGAGTGTT	$(AAG)_{20}$	NA	1	55381359	54
Xtxp321[7]	F: TAACCCAAGCCTGAGCATAAGA R: CCCATTCACACATGAGACGAG	$(GT)_4(AT)_6(CT)_{21}$	NA	8	50508795	55
Xtxp339[7]	F: CCGCACTCTCCACTCT R: CGGAACACAGGGAAGG	$(GGA)_7$	NA	9	—	55

Note: NA = not available; T_a = annealing temperature.

[a] Publication information for the markers is as follows: [1]this paper; [2]Mutegi et al., 2011; [3]Brown et al., 1996; [4]Taramino et al., 1997; [5]Schloss et al., 2002; [6]Ramu et al., 2009; [7]Kong et al., 2000; [8]Bhattramakki et al., 2000.

[b] The chromosome number was defined by genetic mapping and BLAST of the primer on sorghum sequence using Primer-BLAST at the National Center for Biotechnology Information (NCBI).

[c] Position on the chromosome pseudo-molecule (when a reliable Primer-BLAST result was obtained).

fifty-two enriched sequences were clustered, and PCR primer pairs for microsatellite amplification were designed using SAT pipeline (Dereeper et al., 2007). The 10 sorghum chromosomes were each partitioned into five bins and between one and four SSR markers were chosen in each bin. Genotyping was conducted in two laboratories (CIRAD and ICRISAT) following usual practices (Appendix 2). At CIRAD, the forward primer was designed with a 5′-end M13 extension (5′-CACGACGTTGTAAAACGAC-3′), and an IR-labeled IRDye M13 primer was added to the PCR. Electrophoresis and DNA fragment detection used polyacrylamide gel–based system sequencers (LI-COR, Lincoln, Nebraska, USA). At ICRISAT, PCR was performed using fluorescent-labeled primers, and detection used a capillary electrophoresis (Applied Biosystems, Carlsbad, California, USA).

Data acquired in both laboratories were compared taking into consideration exact scoring and heterozygosity levels of each DNA sample for all SSR pairs. Markers were considered reliable when more than 90% of the individuals yielded a scorable genotype and when results were congruent for more than 95% of the data points, i.e., presented identical scoring or homogeneous allelic size shifts and similar detection of heterozygosity. In bins presenting more than one reliable marker, selection was based on minimizing missing data. A set of 48 SSR markers was defined based on these criteria (Table 1). It includes nine markers designed in genes (Schloss et al., 2002; Ramu et al., 2009) and 39 markers designed in noncoding sequences (Brown et al., 1996; Taramino et al., 1997; Kong et al., 2000; Bhattramakki et al., 2000; Mutegi et al., 2011; this paper). Among the selected markers, different types of SSR motifs are represented: dinucleotide (61.2%), trinucleotide (26.5%), tetranucleotide (8.2%), pentanucleotide (2%), and hexanucleotide (2%), as well as perfect (89.8%) and imperfect (10.2%) repeats. The 48 selected markers are distributed across all 10 linkage groups (LG), with a range of three to seven markers per LG.

Among the 48 DNA samples analyzed, 10 samples presenting the largest allelic ranges available were chosen to be technical controls (IS929, IS2156, IS2807, IS7889, IS11119, IS12531, SSM275, SSM379, SSM546, and SSM1284). No particular focus was given to reflect sorghum races or geographic origins. Allele numbers ranged between two and 10 per marker. They were associated into three control pools (C1 including IS2807, SSM1284, SSM275; C2 including IS11119, IS12531, IS929; and C3 including IS2156, IS7889, SSM379, SSM546) (Table 2). PCR amplifications were run a second time for these DNAs, in order to provide both absolute size information by allelic sequencing and clear picture references. PCR was carried out in a total reaction volume of 10 μL containing 25 ng of DNA, 1× buffer, 200 μM dNTP, 2.5 mM $MgCl_2$, 0.10 μM reverse primer, 0.08 μM forward-tailed primer, 0.10 μM M13-tailed primer labeled with IRDye 700 or IRDye 800, and 1 U of *Taq* polymerase (Life Technologies, Carlsbad, California, USA), and performed with an Eppendorf Mastercycler 384-well cycler (Eppendorf, Hamburg, Germany). PCR cycling profiles consisted of 4 min at 94°C; followed by nine cycles of 45 s at 94°C, annealing at decreasing temperatures (60°C for 1 min; −0.5°C/cycle), and 72°C for 1 min 15 s; followed by 24 cycles at 94°C for 45 s, 55°C for 1 min, and 72°C for 1 min 15 s; with a final extension of 5 min at 72°C. Five μL of PCR products was sequenced by GATC Biotech (Constance, Germany; http://www.gatc-biotech.com). The remaining 5 μL were used to provide the reference genotyping profile. They were equally pooled and diluted four times with a 1:5 formamide-water mix, and were separated and detected using a LI-COR 4300 DNA analyzer (LI-COR). Altogether, 196 alleles were sequenced and aligned against a reference sequence (Table 2).

CONCLUSIONS

A reference microsatellite kit was developed for *S. bicolor* genetic diversity assessment. It includes 10 reference accessions pooled into three groups designed to be used congruently and 48 SSR markers regularly spaced along the genome. All single results including primer design, PCR conditions, gel images, and allelic sizes are provided on the website (http://sat.cirad.fr/sat/sorghum_SSR_kit/). Control samples are available upon request (http://golo.

TABLE 2. Alleles identified for each SSR marker on reference accessions of the Sorghum Genotype Identification Kit.

SSR name	Molecular size (bp) of the alleles										Control allele sizes[a]		
											C1	C2	C3
gpsb067	170	172	174	176	178	180					170, 174, 180	172, 172, 176	170, 172, 172, 178, 180
gpsb069	191	193	195	197							191, 193, 195	193, 193, 195	191, 193, 193, 197
gpsb089	165	167	169	173							165, 167, 169	165, 165, 169	165, 169, 169, 173
gpsb123	288	290	292	294	296						290, 290, 296, 296	290, 292, 296	288, 294, 296, 296
gpsb148	135	137	143	147							135, 135, 135, 143	135, 135, 135, 147	135, 135, 135, 135
gpsb151	106	108	110	114	118	126	128				106, 108, 128	108, 114, 128	108, 110, 126, 126
mSbCIR223	108	112	114	116	118						108, 112, 118	108, 114, 118	112, 114, 116, 118
mSbCIR238	79	81	83	85	91	95	105	107			81, 85, 91	81, 95, 107	79, 81, 83, 105
mSbCIR240	108	112									108, 112, 112	108, 108, 112	112, 112, 112, 112
mSbCIR246	98	100									98, 100, 100	100, 100, 100	100, 100, 100, 100
mSb, IR248	89	95	99	101							89, 95, 99, 101, 101	89, 89, 101, 101	89, 89, 89, 101
mSbCIR262	208	214	216	220							208, 214, 216	214, 216, 216	216, 216, 216, 220
mSbCIR276	230	231	232	234							230, 230, 234	230, 230, 232	230, 230, 232, 232, 232
mSbCIR283	113	115	117	119	121	135	137	139			119, 121, 139	113, 113, 117	117, 119, 135, 137
mSbCIR286	112	114	128	134							112, 114, 114	114, 128, 134	114, 114, 114, 134
mSbCIR300	104	106	108	110							104, 106, 110	108, 108, 110	104, 108, 110, 110
mSbCIR306	120	122	124								120, 120, 122	120, 122, 122	122, 122, 124, 124
mSbCIR329	109	111	113	115	117						109, 115, 115	113, 115, 117	111, 113, 115, 115
sb4-72	183	187	189	191	193	195	203				187, 191, 195	187, 189, 193, 203	183, 189, 189, 193
sb5-206	106	108	110	112	114	116	128	132	142	146	110, 114, 128	106, 114, 116	108, 112, 142, 146
sb6-84	183	189	191	193	195	199	217				183, 189, 195	193, 199, 199	183, 191, 195, 217
sbAGB02	96	102	108	116	118	154					96, 102, 116	96, 96, 108, 154	96, 96, 96, 102
Xisep0107	199	205	206								199, 199, 205	199, 199, 205	199, 199, 205, 206
Xisep0310	164	204	219								204, 204, 219	164, 204, 204	164, 204, 204, 204
Xcup02	192	195	198	204							192, 195, 195	198, 198, 204	192, 198, 198, 198
Xcup11	165	172									165, 165, 172	165, 165, 172	165, 172, 172, 172
Xcup14	211	213	215	225							211, 213, 225	211, 211, 215	211, 211, 211, 211, 215
Xcup53	186	194	198								194, 194, 194	186, 194, 198	194, 194, 194, 198
Xcup61	198	201									198, 201, 201	198, 201, 201	198, 201, 201, 201
Xcup62	190	193									190, 190, 193	190, 193, 193	190, 190, 190, 193
Xcup63	133	139	145								133, 139, 139	139, 145, 145	139, 145, 145, 145
Xtxp010	135	143	145	151							143, 143, 151	135, 151, 151	143, 145, 145, 151, 151
Xtxp012	161	173	175	179	185	193	195	205			161, 173, 173	175, 185, 195	179, 185, 193, 205
Xtxp015	199	209	211	215	219	223					209, 215, 223	211, 215, 219	199, 199, 211, 211
Xtxp021	169	175	184	185	191	197	199				169, 175, 191	175, 185, 197	175, 175, 175, 199
Xtxp040	129	135	138	141							129, 138, 141	135, 138, 138	135, 138, 138, 138
Xtxp057	223	237	241	243	245	247	251	257			223, 241, 247	237, 241, 245	241, 243, 245, 257
Xtxp114	211	214	217								211, 211, 214	214, 214, 217	211, 214, 214, 214
Xtxp136	240	243									243, 243, 243	240, 240, 243, 243	240, 240, 243, 243
Xtxp141	135	141	145	151	155	157	161	163	167		151, 157, 157	135, 141, 155	145, 155, 161, 167
Xtxp145	208	210	214	224	232	238	242	244			226, 228, 246	214, 232, 242	210, 224, 238, 244
Xtxp265	186	198	209	213	216	222	234				186, 198, 213	198, 216, 234	186, 216, 222, 234
Xtxp273	169	181	187	190	193	199					169, 190, 190	181, 190, 190	187, 190, 190, 193
Xtxp278	243	249	252								243, 249, 249	243, 243, 249	243, 249, 249, 252
Xtxp295	155	163	167	169	173	175	177	183			173, 177, 183	155, 167, 169	163, 167, 175, 175
Xtxp320	257	269	272	275	278	281	287	290			275, 275, 281	257, 272, 287	269, 275, 275, 278
Xtxp321	192	198	200	202	204	206	208	218	252		202, 208, 252	198, 200, 202, 204	192, 198, 198, 218
Xtxp339	182	200	203								182, 200, 200	200, 200, 203	200, 200, 200, 200

[a]All allele sizes were assessed through sequencing of the alleles, and detection was performed by migration on a LI-COR system. Sizes of the control are indicated in term of allelic doses in each DNA pool (C1, C2, and C3), i.e., when an allele is repeated, it is present more than once in individuals of the control. The intensity of the peak/band should thus be expected to be more intense. Allele sizes of individual samples, as well as profiles, are provided in the associated website (http://sat.cirad.fr/sat/sorghum_SSR_kit/).

cirad.fr/FR/CRB_T_Collections_Sorgho.awp). This microsatellite kit will be very useful for compiling sorghum data across laboratories and techniques, as well as for comparisons of new genetic diversity studies to previous analyses.

LITERATURE CITED

BHATTRAMAKKI, D., J. DONG, A. K. CHHABRA, AND G. E. HART. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* 43: 988–1002.

BILLOTTE, N., N. MARSEILLAC, A.-M. RISTERUCCI, B. ADON, P. BROTTIER, F.-C. BAURENS, R. SINGH, ET AL. 2005. Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics* 110: 754–765.

BROWN, S. M., M. S. HOPKINS, S. E. MITCHELL, M. L. SENIOR, T. Y. WANG, R. R. DUNCAN, F. GONZALEZ CANDELAS, AND S. KRESOVICH. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics* 93: 190–198.

DEREEPER, A., X. ARGOUT, C. BILLOT, J.-F. RAMI, AND M. RUIZ. 2007. SAT, a flexible and optimized Web application for SSR marker development. *BMC Bioinformatics* 8: 465–475.

DEU, M., F. RATTUNDE, AND J. CHANTEREAU. 2006. A global view of genetic diversity in cultivated sorghums using a core collection. *Genome* 49: 168–180.

George, M. L. C., E. Regalado, W. Li, M. Cao, M. Dahlan, M. Pabendon, M. L. Warburton, et al. 2004. Molecular characterization of Asian maize inbred lines by multiple laboratories. *Theoretical and Applied Genetics* 109: 80–91.

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

Mutegi, E., F. Sagnard, K. Semagn, M. Deu, M. Muraya, B. Kanyenji, S. de Villiers, et al. 2011. Genetic structure and relationships within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. *Theoretical and Applied Genetics* 122: 989–1004.

Paterson, A. H., J. E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, et al. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551–556.

Ramu, P., B. Kassahun, S. Senthilvel, C. Ashok Kumar, B. Jayashree, R. T. Folkertsma, L. A. Reddy, et al. 2009. Exploiting rice–sorghum synteny for targeted development of EST-SSRs to enrich the sorghum genetic linkage map. *Theoretical and Applied Genetics* 119: 1193–1204.

Schloss, S. J., S. E. Mitchell, G. M. White, R. Kukatla, J. E. Bowers, A. Paterson, and S. Kresovitch. 2002. Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theoretical and Applied Genetics* 105: 912–920.

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

Appendix 1. List of 48 DNA samples chosen to test the reliability of the 117 SSR markers.[a]

Accession no.	Geographic origin[b]	Latitude	Longitude	Racial characterization[c]	Accession no.	Geographic origin[b]	Latitude	Longitude	Racial characterization[c]
IS303	CHN	NA	NA	B	IS23178	ZMB	−15.561	28.285	DC
IS929	SDN	NA	NA	D	IS23254	ZMB	NA	NA	B
IS2156	NGA	NA	NA	B	IS23644	GMB	13.660	−15.400	Gma
IS2262	SDN	NA	NA	C	IS28409	YEM	11.220	14.370	D
IS2263	SDN	NA	NA	D	IS29233	SWZ	NA	NA	GC
IS2807	ZWE	NA	NA	C	IS29407	LSO	−29.530	28.150	GC
IS3421	IND	19.260	76.770	D	IS30405	CHN	36.848	111.779	D
IS3967	IND	NA	NA	D	IS30538	KOR	NA	NA	B
IS4821	IND	22.310	73.180	DC	SSM29	CMR	NA	NA	Dmkr
IS6193	IND	26.320	80.290	G	SSM205	BFA	NA	NA	Gga
IS6745	BFA	NA	NA	Gga	SSM215	ETH	NA	NA	C
IS7889	NGA	10.430	3.820	C	SSM249	BFA	12.370	−1.525	G
IS8685	SDN	NA	NA	C	SSM275	BFA	12.267	−2.150	G
IS9597	NER	13.260	7.060	D	SSM379	MLI	NA	NA	G
IS11119	ETH	9.500	38.230	D	SSM505	NER	NA	NA	Gma
IS12531	ETH	9.300	42.130	DB	SSM546	NER	NA	NA	D
IS14331	ZAF	NA	NA	Gro	SSM547	NER	NA	NA	C
IS14414	MWI	NA	NA	Gcon	SSM557	NER	NA	NA	G
IS16186	CMR	NA	NA	D	SSM964	SEN	NA	NA	D
IS19453	BWA	NA	NA	D	SSM973	SEN	NA	NA	Gma
IS19455	BWA	NA	NA	Gma	SSM1049	SEN	14.720	−17.274	G
IS20016	SEN	14.130	−14.970	G	SSM1057	SEN	14.733	−16.567	D
IS22282	BWA	−23.100	26.830	B	SSM1284	COD	NA	NA	B
IS22294	BWA	−23.010	27.760	K	Sariaso10	BFA	NA	NA	C

[a] Accessions were all taken from previous collections (IS numbers available at ICRISAT, India; SSM numbers available at CIRAD, France). Information provided is from the associated databases, which do not always provide GPS indications; accessions without GPS information are noted as NA (not available). All the accessions are freely available and can be ordered through the System-wide Information Network for Genetic Resources (SINGER) website (http://singer.cgiar.org) or at CIRAD (http://golo.cirad.fr/FR/CRB_T_Collections_Sorgho.awp).

[b] Geographic origin: BFA = Burkina Faso; BWA = Botswana; CHN = China; CMR = Cameroon; COD = Democratic Republic of Congo; ETH = Ethiopia; GMB = Gambia; IND = India; KOR = Republic of Korea; LSO = Lesotho; MLI = Mali; MWI = Malawi; NER = Niger; NGA = Nigeria; SDN = Sudan; SEN = Senegal; SWZ = Swaziland; YEM = Yemen; ZAF = South Africa; ZMB = Zambia; ZWE = Zimbabwe.

[c] Racial characterization: B = bicolor; C = caudatum; D = durra; DB = durra bicolor; DC = durra caudatum; Dmkr = durra muskwaari; G = guinea; Gma = guinea margaritiferum; GC = guinea-caudatum; Gcon = guinea conspicuum; Gga = guinea gambicum; Gma = guinea margaritifirum; Gro = guinea roxburghii; K = kafir.

APPENDIX 2. PCR and detection conditions in the two laboratories for comparing the 117 SSR markers.

CIRAD: The forward primer was designed with a 5′-end M13 extension (5′-CACGACGTTGTAAAACGAC-3′).

PCR amplifications were performed in an Eppendorf thermocycler (Eppendorf, Hamburg, Germany) on 25 ng DNA in a 10 μL final volume of buffer (10 mM Tris-HCl [pH 8], 100 mM KCl, 0.05% w/v gelatin, and 2.0 mM $MgCl_2$) containing 0.08 μM of the M13-tailed primer, 0.1 μM of the other primer, 160 μM of dNTP, 1 U of *Taq* DNA polymerase (Life Technologies, Carlsbad, California, USA), and 0.06 μM of M13 IRDye 700 or IRDye 800-labeled primer (Biolegio, Nijmegen, The Netherlands).

For all primer pairs, the PCR program was: initial denaturation at 95°C for 1 min; 10 cycles of touchdown PCR from 60°C to 55°C with 94°C for 30 s, T_a for 60 s, and 72°C for 120 s; 25 cycles of 94°C for 30 s, 55°C for 60 s, and 72°C for 120 s; and a final elongation step at 72°C for 8 min.

IRDye 700–labeled or IRDye 800–labeled PCR products were diluted 10-fold and fourfold respectively, then subjected to electrophoresis in 6.5% polyacrylamide gels with a LI-COR IR2 system with LI-COR size standard (LI-COR, Lincoln, Nebraska, USA).

ICRISAT: Amplifications were performed using fluorescently labeled forward primers (6-FAM, HEX, and NED; Applied Biosystems, Carlsbad, California, USA).

PCR reactions were performed in 5 μL reaction volumes with final concentrations of 2.5 ng DNA, 2 mM $MgCl_2$, 0.1 mM of dNTP, 1× PCR buffer, 0.2 μM of forward and reverse primers, and 0.1 U of *Taq* DNA polymerase (AmpliTaq Gold, Applied Biosystems) in a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems) with the following cyclic conditions: initial denaturation at 94°C for 15 min (to activate *Taq* DNA polymerase), followed by 10 cycles of denaturation at 94°C for 15 sec, annealing at 61°C for 20 sec (temperature reduced by 1°C for each cycle), and extension at 72°C for 30 sec. This was followed by 35 cycles of denaturation at 94°C for 10 sec, annealing at 54°C for 20 sec, and extension at 72°C for 30 sec, with the final extension of 20 min at 72°C.

Amplified PCR products, according to their multiplexes, along with internal ROX-400 size standard, were separated by capillary electrophoresis using an ABI 3700 sequencer (Applied Biosystems).