

Characterization of AT-rich microsatellites in common bean (*Phaseolus vulgaris* L.)

Mathew W. Blair · Hector F. Buendía ·
Martha C. Giraldo · Isabelle Métails · Didier Peltier

Received: 11 January 2008 / Accepted: 23 August 2008 / Published online: 11 September 2008
© Springer-Verlag 2008

Abstract Polymorphism of microsatellite markers is often associated with the simple sequence repeat motif targeted. AT-rich microsatellites tend to be highly variable and this appears to be notable, especially in legume genomes. To analyze the value of AT-rich microsatellites for common bean (*Phaseolus vulgaris* L.), we developed a total of 85 new microsatellite markers, 74 of which targeted ATA or other AT-rich motif loci and 11 of which were made for GA, CA or CAC motif loci. We evaluated the loci for the level of allelic diversity in comparison to previously characterized microsatellites using a panel of 18 standard genotypes and genetically mapped any loci polymorphic in the DOR364 × G19833 population. The majority of the microsatellites produced single bands and detected single loci, however, 15 of the AT-rich microsatellites produced multiple or double banding patterns; while only one of the GA or CA-rich microsatellites did. The polymorphism information content (PIC) values averaged 0.892 and 0.600 for the AT and ATA motif microsatellites, respectively, but only 0.140 for the CA-rich microsatellites. GA microsatellites, which had a large average number of repeats, had high to intermediate PIC, averaging 0.706. A total of 45 loci

could be genetically mapped and distribution of the loci across the genome was skewed towards non-distal locations with a greater prevalence of loci on linkage groups b02, b09 and b11. AT-rich microsatellites were found to be a useful source of polymorphic markers for mapping and diversity assessment in common bean that appears to uncover higher diversity than other types of simple sequence repeat markers.

Introduction

Microsatellites, which are genetic markers based on the PCR amplification of simple sequence repeats, have been found to vary in the polymorphism they detect depending on their length, genomic location and the sequence of the repeat motif they contain (Hearne et al. 1992). Microsatellite markers for plant species generally have been made to target di-, tri- or tetra-nucleotide repeats with either equal GC/AT content, more AT content than GC content or more GC content than AT content, though less frequently in this last case (Powell et al. 1996). The source sequences for microsatellites have included simple sequence repeat-enriched library clones, expressed sequence tags and bacterial artificial chromosome end sequences (Edwards et al. 1996; Rota et al. 2005; Mun et al. 2006). The most valuable microsatellite markers in many settings are those that detect a high amount of variability that are easy to amplify and are single copy rather than being multiple banding (Morgante and Olivieri 1993; Powell et al. 1996). Therefore it is important to evaluate the different sources and motifs of microsatellites for their productivity in each crop in terms of producing polymorphic, amplifiable markers.

Legume crops are of primary importance to world agriculture due to their role in biological nitrogen fixation, wide

Communicated by A. Kilian.

M. W. Blair · H. F. Buendía · M. C. Giraldo
CIAT, International Center for Tropical Agriculture,
1380 N.W. 78th Ave., Miami, FL 33126, USA

M. W. Blair (✉) · H. F. Buendía · M. C. Giraldo
A. A. 6713, Cali, Colombia
e-mail: m.blair@cgiar.org

I. Métails · D. Peltier
Laboratoire de Génétique, UFR Sciences,
2 bd Lavoisier, 49045 Angers Cedex, France

environmental adaptability and utility in rotations with cereals (Graham and Vance 2003). Among the grain legumes, soybean, peanuts and common beans are the most important commercial crops, with soybean and peanuts as important for protein and oil production and common bean the most important for direct human consumption (Broughton et al. 2003).

Microsatellites have been developed for various legume crops but from different sources. In soybean, many microsatellites are available and most have been made through the development of ATA-enriched microsatellite libraries (Cregan et al. 1999; Song et al. 2004) with a large number also developed through end sequencing of bacterial artificial chromosome clones (Shultz et al. 2007). Similarly, in chickpea, TAA-based microsatellites have proven to be highly polymorphic (Huttel et al. 1999; Udupa et al. 1999; Udupa and Baum 2001; Udupa et al. 2004).

In common bean, on the other hand, microsatellites have mostly been developed through SSR enriched libraries for GA or CA repeat containing sequences (Gaitán-Solís et al. 2002; Métais et al. 2002; Yaish and Pérez de la Vega 2003; Buso et al. 2006; Campos et al. 2007; Benchimol et al. 2007) and through screening of EST collections or gene sequences (Yu et al. 1999, 2000; Blair et al. 2003; Hanai et al. 2007). The latter SSR type have generally been of lower polymorphism than genomic microsatellites developed for the crop and therefore are most useful in inter-genepool crosses but less so in intra-genepool crosses (Blair et al. 2006a). In peanuts, low polymorphism of microsatellite markers has been a serious problem especially within the cultivated accessions (Hopkins et al. 1999), but has been overcome by screening of restriction digested genomic clones for various motifs including AT-rich sequences (Ferguson et al. 2004). A very complete study of microsatellite frequency in the model legume *Medicago truncatula* has shown that AT and ATA motifs along with GA and CA motifs are frequent in genomic sequences from gene-rich BACs and BAC ends in this species (Mun et al. 2006). In summary, the screening of AT-rich motifs such as ATA and AT have proven to be useful for developing large sets of microsatellites in individual legume species but has not been widely exploited for common bean, although results from Métais et al. (2002) show the promise of the ATA motif in detecting high polymorphism within snap beans.

In plant breeding applications, microsatellites are commonly used for genetic mapping of commercial crops and for diversity assessments of germplasm accessions (Morgante and Olivieri 1993). Highly polymorphic microsatellites tend to be more useful in these studies because they can be used in any genetic cross combination or in closely related samples of genotypes from a local collection and in this sense, microsatellites can be complementary to SNP markers which tend to be bi-allelic and genepool specific

(Rafalski 2002). In common beans, we have used GA-motif microsatellites along with EST-based microsatellites to construct a genome-wide, simple-sequence framework map based on inter genepool crosses (Blair et al. 2003) and the same markers have been useful for evaluation of a broad range of diversity (Blair et al. 2006a) and for evaluating race structure (Díaz and Blair 2006; Blair et al. 2007). However, additional microsatellite markers are needed for saturation mapping, for constructing genetic maps in very narrow genetic crosses such as within a given race of common beans and for evaluating within accession variability. For these activities, highly polymorphic markers such as the AT-rich microsatellites available for some other legumes would be desirable for common beans.

The objective of this research, therefore, was to develop a set of microsatellite markers for common bean based on AT-rich motifs and ATA trinucleotide repeats and to evaluate their utility for genetic analysis of the crop. To do this, we build on the work of Métais et al. (2002) who developed an initial set of 15 ATA-based microsatellites. We specifically concentrated on characterizing microsatellites selected from ATA versus CAC/CA/GA enriched libraries and evaluated the genetic diversity detected by the different types of markers and their distribution on a common bean genetic map.

Materials and methods

Plant materials and DNA extraction

One population of recombinant inbred lines and a set of 18 diverse genotypes were used for this study: The population was based on the cross DOR364 × G19833 as described in Blair et al. (2003). The genotypes for the polymorphism survey consisted in 18 genotypes evaluated as survey I by Blair et al. (2006a), which included 7 genotypes from the Andean gene pool and 11 from the Mesoamerican gene pool, with a total of 15 cultivated and 3 wild accessions, these last genotypes representing Argentinean, Colombian and Mexican wild diversity. Among the cultivated genotypes, three were advanced breeding lines from CIAT (BAT477, BAT881 and DOR364) with the remainder germplasm accessions (G11360, G11350, G21657, G21078, G21242, G14519, G19833, G3513, G21212, G24404, G24390, G19892) or locally bred varieties (eg., Carioca and ICA Radical Cerinza, varieties from Brazil and Colombia, respectively) which have been used to develop mapping populations for common bean including recombinant inbred lines BAT881 × G21212 in Frei et al. (2005) and an advanced backcross population from Radical Cerinza × G24404 in Blair et al. (2006b). Germplasm accessions meanwhile, represented the Nueva Granada and Peru races within the Andean gene pool and the Durango,

Jalisco and Mesoamerica races within the Mesoamerican gene pool. The diversity panel included the parents of the DOR364 × G19833 mapping population, which was useful to determine parental polymorphism before genetic mapping. Total genomic DNA for these genotypes was isolated using a CTAB extraction method as described in Afanador and Hadley (1993) from bulked-leaf tissues of eight greenhouse-grown plants per line, so as to capture within accession diversity as part of the parental screening. Any markers detecting heterozygosity/heterogeneity in the parents of the DOR364 × G19833 population were not used for genetic mapping.

Microsatellite markers

The construction of the SSR-enriched libraries and microsatellite development was as described in Métais et al. (2002) with two types of markers developed: (1) those selected based on the motif ATA (AT-rich) and (2) those selected based on the motifs CAC, CA and GA (non-AT-rich). Microsatellites were named according to the convention shown in Table 1 and were amplified for the polymorphism survey using standard PCR conditions as described in Blair et al. (2003) with a hot start of 92°C for 3 min; then 30 cycles of 92°C denaturing for 30 s; 60°C annealing for 30 s and 72°C extension for 45 s; followed by a 5 min final extension at 72°C. The PCR reactions were carried out in a 12 µL final volume containing 25 ng of genomic DNA, 0.15 µM of each of the forward and reverse primers, 10 mM Tris-HCl (pH 7.2), 50 mM KCl, 1.5 MgCl₂, 200 mM of total dNTP and 1 unit of Taq polymerase (Promega, Madison WI). Amplifying primers were mapped using a touchdown profile with 1°C drop per cycle in annealing temperatures for 9 cycles followed by 35 cycles at 52°C. After amplification, the PCR reactions were concentrated in a vacuum centrifuge for 30 min upon which a volume of 4 µL of formamide, containing 0.4% bromophenol blue and 0.25% w/v xylene cyanol FF, was added to each PCR reaction and the mixture was denatured at 92°C for 2 min. The mixtures were then loaded onto 4% denaturing polyacrylamide (29:1 acrylamide: bis-acrylamide) gels that contained 5 M urea and 0.5× TBE and run in Sequi-Gen GT electrophoresis units (Biorad, Hercules, CA, USA) at approximately 120 W and 45°C. Detection of PCR amplification products was via silver staining according to Blair et al. (2003) and the allele sizes were estimated based on 10 and 25 bp MW ladders.

Data analysis

Genetic diversity detected by the individual markers was evaluated by estimating the polymorphism information content (PIC) from the allele assignments. The PIC value

refers to the relative value of each marker with respect to the amount of polymorphism it exhibits. PIC values as well as the distribution and range of allele sizes were calculated using the software program Powermarker (Liu and Muse 2005). For the genetic mapping exercise, each RIL was scored for the parental allele it contained for each microsatellite. Segregation data was used to place the microsatellites on the established genetic map for DOR364 × G19833 described in Blair et al. (2003). Linkage analysis was conducted with the Kosambi mapping function using the software application Mapmaker 2.0 for Windows (Lander et al. 1987) and linkage group orientation was drawn as for the genetic map of Freyre et al. (1998). The microsatellites were placed to the established linkage groups with the 'try' and 'compare' commands with a minimum LOD of 3.0.

Results

Marker characterization

A total of 85 common bean microsatellites were developed in this study (Table 1). Of these, 74 were based on AT-rich sequences and 11 were based on non-AT-rich sequences. The AT-rich microsatellites included 71 that were newly developed from an ATA-enriched library developed for this study and three AT-rich microsatellites from gene-based sequences (HRG, ME1 and BN). The 11 non-AT-rich microsatellites included two microsatellites from CAC selection, two from CA selection, and six from GA selection along with a GC-rich gene-based microsatellite (PG2). The four gene-based microsatellites are listed after the 81 microsatellites from the motif-enriched libraries in Table 1. The development of the non-AT-rich microsatellites is described in Métais et al. (2002) and complements that work, while the AT-rich microsatellites were developed from a further 1632 ATA-enriched clones. These were screened by hybridization using a ³²P labeled (ATA)₈ oligonucleotide probe resulting in 698 clones with a clear hybridization signal for which clear sequences were obtained from 288 clones.

After analysis, 271 sequences contained ATA repeats while 17 clones did not present microsatellites. Sequence alignment was used to determine possible contigs that grouped the 271 positive clones into 84 clusters, which were used to define the 71 markers described above. From these results, we deduce a redundancy of (187/271) 69% for the ATA library.

The microsatellites from the ATA-enriched library had 8–72 repeat motif units; however the mean number of repeats across the ATA-based markers (15.5) was closer to the smaller of these values. The microsatellite with the highest number of repeats was ATA20 (72) while ATA13,

Table 1 Design and diversity assessment of 85 common bean microsatellite loci

Marker	Motif and repeat length ^a	Expected product size	No. of alleles ^b	PIC ^b	Min. allele size	Max. allele size	Forward primer	Reverse primer
ATA2	(TAA)20 (GAA)10	129	3	0.441	110	125	CGT CTG TGT TTG CTA GTA	GAG AGA GGT GAA ATT GAA AG
ATA3	(TAA)9	133	2	0.239	130	132	CCG AAG ACA TAA CTG ATG AT	TGA GTA CGT GCT GGA ACT T
ATA4	(TTA)8	154	5	0.650	130	142	CAG ATT AGT TCT TTT GTC AT	GGC GTC ATG AAA CAA GAA
ATA5	(TTN)42 N = A, G, T	191	3	0.194	175	200	TCC GTA ATT GGA ACG GAA CA	CCG AIT TTC TTT TGG TCG CT
ATA6	(TAT)19	143	8	0.794	104	136	TGA TTT GTC TAA CAC TTC AC	GGG GAT GAT TTG CAT GTA G
ATA7	(ATA)11	136	8	0.834	132	160	ATA AAT CTA TTG AGT TCT AG	AAC AAG TCA ATA ATC TAA AG
ATA9	(AAT)12	191	10	0.803	172	192	AAG CCA CTG TAG CTG GAA GC	GCC CTT CTA CTA CCA CTC TA
ATA10	(ATT)11	106	5	0.545	104	128	TTG TCA TCC AAA GAT TAT TA	GAC AAT AAT AAA AAT GGT TT
ATA13	(TTA)68	315	NA	NA	–	–	ATT CAT TAG CAA CTC AAG C	TGA GTA GAG ATA AAC ACA TA
ATA16	(TAA)14	155	8	0.781	142	165	CAA AAT GGA AAG AAA ATG TC	TGT TCG TAA ATC AAA TCT TG
ATA20	(TAA)72	293	12	0.860	148	320	GCA TTG CAG GAA ATC AT	TGT ATC AAT TAC TTT TTT AC
ATA22	(TAA)18	152	7	0.746	118	178	GAG ATG ATT TGC ATG TAG TA	TCC TTG CTT CAT TGA TTT GT
ATA24	(TTA)18 (2)	170	7	0.820	98	134	TGA GTA ATG GTT ACT ATT ATT A	GTG CAC ATA CAA AAT CAT TC
ATA26	(TAA)12 (2)	188	3	0.489	130	142	TGA AGG ACA ACT TGT TCA C	GGC AAC TGT TTA CTG CTG
ATA27	(TTA)13 (1)	130	5	0.611	120	155	TGG AAG CCT ATT ATA ACA ATG A	GAA GAA GAA TAA TAG TAC GCC G
ATA30	(TAT)37 (16)	202	4	0.525	160	230	AAG CCT ATT ATA ACA ATG AA	GAT CAT GAG TAT AAT AAG AA
ATA32	(TAA)9	105	6	0.636	98	110	CAC CAT TTC AAC GTT TGA G	CAA AGT ATT TGA CGA CAC C
ATA39	(TA)13	180	NA	NA	–	–	GAC TGA GTA ACT AAT CTA AA	TTA TAT AAA GTT TCA TAG TG
ATA69	(TTA)24 (2)	229	5	0.404	88	122	TAT ATA ATA TTT ACA TTC GG	CAA TAG AAA AAC TCC TAT AG
ATA71	(ATT)10	193	8	0.720	182	197	TTC AAT TGA CTT GAA GGT CAG G	AAT AAC CCC AAA CTG AGA TAA A
ATA73	(TTA)11	97	1	0.000	97	97	ATA TTA TTT ATA CAA ATT TCA AAG	AT CGA GAT CTA GAG AGA AGG
ATA74	(TTA)7	145	4	0.529	145	152	AAT TAT TAG CAA AAA TAG AGG TG	TGC TAA ATT TCT ATC GTC TGA GT
ATA76	(TAA)12 (N)29 (TA)8	198	9	0.858	188	220	GAT CTA ATT TTT TGC TTC TT	TGC TAT TAT ATT CCT TTT CA
ATA77	(TAA)12	186	7	0.778	150	185	CCA GGA TTT GTA AAA TCA TCA T	TGT TTC TAG ATT TCT GGA TCT G
ATA82	(TAT)21	231	NA	NA	–	–	GTG GAG ATG TTC TTA TTT CGT C	AAC ACC CAA ACA AAA CTC ATC T
ATA84	(TAT)8	142	MB	MB	–	–	CGG ATG GAG ACG TGT AAA TA	TGG GAA ACA ATA CTG GAA AA
ATA86	(TAA)18	159	MB	MB	–	–	GAC CAT GAT AAA AAC ATG AT	GCA GTT CAA GAC AAA TAT TT
ATA90	(TTA)8	121	3	0.452	120	135	GAG ATG ATT TGC ATG TAG TA	TCC TTG CTT CAT TGA TTT GT
ATA108	(TTA)15	125	5	0.542	120	170	CAA TGT TGG TAT TTT ATT TT	ACA AAA ACC ACA CAT TAT TA
ATA112	(ATA)8	258	5	0.652	248	262	TTA TAA AAC GGA TCA TTC TG	CCA TTC TTC ACT CAA ATA AT
ATA113	(TAA)7	176	6	0.650	168	182	GGT CTC TAT TTT AGT TTC TAT TG	CTC CAT TAC ATA TTT ACT TTA CC
ATA116	(TAA)12	172	7	0.577	118	145	CAA ATT GAC AGG TCC AAA AT	CAT GAT TAC GAA TTC GAG CT
ATA118	(ATA)20 (4)	168	NA	NA	–	–	TCT CCA AAT AAT CGT TCG TG	TGT GAA TGC AGA AAC CAA AT
ATA119	(ATA)8	133	MB	MB	–	–	ATC ATG AGT AAA CCT CAT AAC	ATA TGT GCA ACA CAA CTC TTA

Table 1 continued

Marker	Motif and repeat length ^a	Expected product size	No. of alleles ^b	PIC ^b	Min. allele size	Max. allele size	Forward primer	Reverse primer
ATA121	(ATA)9	208	1	0.000	91	91	TCA ACA TCA CAT GTA AAA GTC A	GAA AGT TTC CTC TTC CAG CTC C
ATA122	(TTA)8	141	NA	NA	–	–	TTG ATC TAA ATG ACA CCA GA	TTT CGA TCA ATC CAA TTA TG
ATA133	(TAT)9	182	5	0.667	182	195	TTT GTC ATA TTA TCG AAT TT	TAA TCC GTA TAT AAT TTT GG
ATA138	(TTA)13	154	8	0.776	140	158	ACA ATT GAA ATT TAT TTT CC	GGT ATA CTG AAA CAG GTT GT
ATA143	(ATT)15	149	7	0.692	150	172	AAA TCT TAG TGA GTT ATT TAT T	AAA TAT AGT GAT TGA TTG AGA G
ATA145	(TTA)20 (4)	140	6	0.780	138	160	ATG TGT CAT TTA GTA GAC TAA G	TAT CTC CTG ATA TAT TGA ATT G
ATA150	(TAT)11	167	8	0.822	180	210	AGA GAG AAA AAA TGA AGA GA	CGT CCT CAT AAA CAA TAA TA
ATA154	(TAA)16	235	NA	NA	–	–	AAA ACA TCA AGC AAG CTT CA	ATC CAA AAA GGT GGC TTA GA
ATA157	(ATA)11	151	1	0.000	130	130	TTG ACA GGA TTC ATG ACT GA	CAA TGG TCA TTT ACA AAG CA
ATA165	(ATA)9	252	MB	MB	–	–	AAA ATA AAC TAT CGC TAA AA	AAT TTA CTG GGT ACT CTC TG
ATA168	(TAA)16 (1)	121	MB	MB	–	–	CTC AAA TAG AAT AGT CGA TG	TGC TGG GAT AGT ATG TAG TA
ATA170	(TAA)15 (1)	148	7	0.819	122	158	CTG AGA TAA TTT GTA ATA AT	TTA TAA ATT TTA TAC GTA AC
ATA172	(TTA)24 (5)	153	7	0.796	142	160	ACG AAA AAC GAT AAA TAA GA	AAA AGG GTT ACC ATA GAA GT
ATA173	(TTA)7	91	2	0.305	168	170	TAA TTC TAA TAC ATC ATG CG	AAT ACA TTG TAT CAT AAC CG
ATA175	(ATT)7	126	6	0.573	114	130	CCC CCC AGA ATT TAC TTT AT	TCC AGC AAA GGA AGG ATA TT
ATA180	(TAA)11	158	9	0.835	212	250	ATT ATT TTA TCA TGG ATG AA	AAA TTT CAA AAG AAA AAT AC
ATA205	(ATT)15	141	MB	MB	–	–	TCG CCT TGA GGA CAT AGT AA	ATC TCT TCC AGG GAA AAT TC
ATA209	(TAA)16 (3)	244	4	0.575	230	243	CAT AGT AAC TTT GTT GTC CA	TGA ACA TTG TTT TCT TCT AA
ATA216	(TTA)7	271	MB	MB	–	–	CAT GGT GAA AGA TAA TGA CG	AAT AAT AAT AAT TGT CCC GG
ATA217	(ATT)11	121	MB	MB	–	–	TCT GAG CAG GAG ACA TTG TG	CAA GCA GGT CAG ACA AAA TG
ATA220	(TTA)6(TAT)4(ATT)5	94	4	0.592	76	119	CTT TAT ATA AAG AAG ATA TA	ATG AGA TAA TTA TAA TAA TA
ATA226	(TAT)9	161	6	0.694	158	175	CAT GCC GCA CAC ATT ATA TT	TTT TTT CCA ATT CTG TGC TG
ATA227	(TAA)16	255	MB	MB	–	–	AAA CAC CTT AAA TAA GGG GC	TGG AAT GAT GGA AAC ATG AA
ATA230	(TAT)14	180	MB	MB	–	–	ATC ACC AAC ATG TGC ATC AA	TAT CTA CAT CAC CCT TGG GC
ATA234	(TAA)7	234	8	0.807	188	234	TTT TTC TTA GAA AGT GGC GT	AGT AAG GCA AAG TTT TTG GC
ATA236	(TTA)8	222	1	0.000	200	232	AGA ATT CAG TTT CAT GCC TC	CCA AGG AAA ATG AAA AAG TC
ATA241	(TTA)17	158	10	0.818	128	200	ATT GAT TTG CGA ACA CTC AC	ATG CAT GAT TAC GAA TTC GA
ATA243	(TAT)8	233	7	0.761	184	219	ACT ATA TTG GGT CTC CTA AA	CCA ATT CCA TAT GTA AGT TT
ATA244	(TAT)15 (2)	197	6	0.578	178	220	TGG CGG TAG AAA TAT GAA AA	GAG ATG GAA GAT GGG TTC AG
ATA247	(TAT)15 (1)	162	6	0.509	158	188	AAA CAA TGC ATA AAA AGT CC	AGA GTT TAT TGG ATT AAG CG
ATA248	(TAA)15 (3)	125	5	0.720	124	138	GAG AGT ATG GGT GTG TGC TG	CCT CCT CCT CAC CAC TCT AA
ATA268	(TA)9(ATA)15(ATG)8	136	2	0.091	148	195	AAA AGA AAA ACA ATT AIT GC	CAT ACA CAA ACA TGT CTT CA
ATA269	(TTA)11	107	7	0.780	102	132	AAC AGG AAA ATA TCA TCT TA	GAA TAA ATC TAT TGG AGT TC
ATA271	(TAAA)4	102	MB	MB	–	–	GAA TCA ACA TGT ATT GAG AA	CGT GTA AAA GAA TAA CAA TC

Table 1 continued

Marker	Motif and repeat length ^a	Expected product size	No. of alleles ^b	PIC ^b	Min. allele size	Max. allele size	Forward primer	Reverse primer
ATA288	(TTA) ₁₆ (3)	201	7	0.732	174	208	ATA AAA TTT ATA ATG GAT GG	GGG AAT CCT CTT TAT ATA TA
ATA289	(TAT) ₂ (AAT) ₁₂	249	7	0.693	235	265	GGA TGT ATT GCC CAA AAT TA	TTG GGG CAC TTA TAA TTT GA
ATA296	(TTA) ₁₁	145	11	0.849	135	175	TTT CAT ATT TTT CCA CGA CC	TAA CCC TTT CTT TAC CCC TT
CA3	(CA) ₁₂ (3)	128	1	0.000	128	128	AAC TGT CCA CCC ACG TAT CT	GAG GAC TAC GTA CAA GTG AC
CA5	(CA) ₁₃ (3)	144	2	0.280	140	144	ATA GAT GAG GTA TGT GAA AG	GGG ATA ATG TAA GAT GTA AG
CAC1	(GA) ₃₁ (2)	112	12	0.827	74	114	ACA AAC AAT TAT AAA GAG TG	GTT GAG TGA ACC AAT ATA AA
CAC2	(GA) ₂₉ (2)	72	6	0.691	59	93	AAA TTG GTT GTG TAA ATC TC	CAA TTA TAA AGA GTG AGA GA
GA4	(GT) ₁₂ (4)	149	1	0.000	145	145	AGG TGT AGT CGT TTC ACT TC	CTG TTC TTG AAA CCT ACA AG
GA9	(GA) ₈₀ (19)	229	MB	MB	–	–	AAA GAG AGA GAT AAA GGG G	GCA GGT TGC AGG TAG TAC
GA11	(GA) ₂₉ (2)	140	NA	NA	–	–	AAG GTT GAG TGA ACC AAT AT	CCA GCA GTA AGA AAC ATA AT
GA12	(GA) ₁₅ (1)	148	5	0.410	145	162	TAA AAC AGA CAG TTA TAA GG	TTG TAA TCT GTG AGT GAT TA
GA15	(GA) ₁₆	120	6	0.726	108	118	GGT TGA GTG AAC CAA TAT AA	CGA TTG TTT ATA GAA GCA GA
GA16	(GT) ₁₃ (GA) ₂₂ (3)	132	11	0.874	112	170	CAA CAA AAT CAA ATC AAA CAA T	GGT TGA CTG AAC TAA TAT AAA
BN	(TAT) ₉	158	5	0.567	144	162	CAG CTT AGC ATT GAA TGT TG	GAA AGA GCA GTA TAA GTC CT
HRG	(AT) ₂₂	124	13	0.905	114	230	GCT CAT ATT TGG GAG AAA TC	ATT GTT TGT TGC CTT TCC CA
ME1	(AT) ₁₂	169	11	0.879	174	198	ATA CGA GGT TAG CCT AAA TC	ATT CAC CGT CAA ACT CGC AA
PG2	(CGG) ₈	138	2	0.375	132	138	CAT TCC CTG AAG CCC GAA T	CGG TGA TTG ACC AGA GAA G

^a Number of imperfect repeats indicated in parenthesis after the motif

^b Allele number and PIC value not determined for non-amplifying (NA) or multiple banding (MB) microsatellites

ATA5 and ATA30 also had large numbers of repeats (68, 42 and 37, respectively), however, the last two with some imperfect repeats. ATA172 and ATA69, each with 24 repeats, were the highest repeat number microsatellites among the remaining markers while all the other microsatellites in this category had fewer than 21 repeats. The microsatellites developed for this motif may be biased against long-repeat lengths, since ATA repeats generally are difficult to sequence as was evidenced in the sequencing success rate of 41.2% for the original ATA containing clones.

A total of 68 of the ATA-targeted microsatellites had a single type of repeat while four were complex microsatellites having two or more repeat motifs. In these cases, ATA motifs were paired with GAA and ATG tri-nucleotide motifs (for ATA2 and ATA268, respectively) or with TA di-nucleotide repeats (for ATA76 and ATA268). While searching for ATA motif microsatellites, we also identified two markers that contained AT-rich sequences but not the precise ATA trinucleotide motif that had been enriched for; these were the TAAA-based marker ATA271 and the AT-based marker ATA39. Another marker, ATA5, targeted ATN repeats, that is to say it was AT-rich having imperfect repeats at the third position with N equal to A, G or T nucleotides. In addition, 18 other microsatellites were imperfect and had interrupted repeats in parts of their ATA-based sequences, of which ATA220 contained a motif where the repeats began with (TTA)₆, continued with (TAT)₄ and ended up with (ATT)₅; while ATA289 began with (TAT)₂ and ended with (AAT)₁₂.

Of the non-AT-rich microsatellites, the two CA-selected markers had 12 and 13 repeats, respectively; the two CAC-selected microsatellites had long GA motifs (Table 1); and the GA-selected microsatellites had either GT or GA motifs with an average repeat number of 31.7. It is not known why selection for CAC motif produced GA microsatellites but this could have been due to hybridization stringency used in

the procedure by Métais et al. (2002). Four gene-based sequences were selected to add additional AT-rich microsatellites, including two based on dinucleotide (AT)_n motifs and one based on the trinucleotide (TAT)₉, as well as a CG-rich marker PG2 based on the trinucleotide (CGG)₈. The average number of repeats in these microsatellites was 12.8.

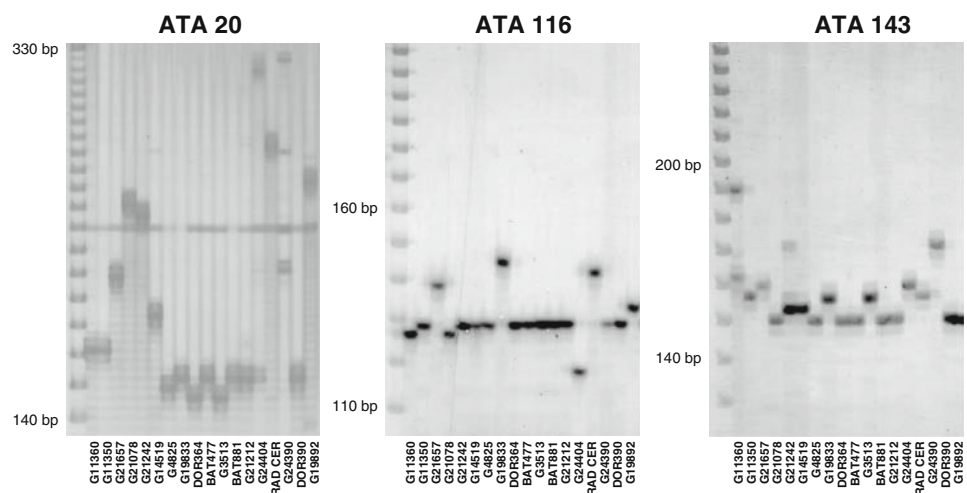
When all the AT-rich microsatellites (those with AT, ATA or ATG motifs) were compared to all the non-AT-rich microsatellites (those with CA, GT, GA or CGG motifs), we observed that the average number of repeats was lower for the AT-rich microsatellites (15.5 vs. 25.4 repeat units) although this was mainly due to the GA microsatellites that tended to be longer in repeat number (33.6) than the AT (15.7), ATA (15.6), or GT (12.3) motif markers.

Diversity assessment

PCR amplification and banding patterns of the AT-rich and non-AT-rich markers across the diversity panel were comparable, however, AT-rich microsatellites tended to be slightly more difficult to amplify but more diverse in terms of allele numbers. Figure 1 shows the amplification pattern of three of the ATA based microsatellites in the parental survey of 18 genotypes. A total of seven microsatellites (ATA13, ATA39, ATA82, ATA118, ATA122, ATA154 and GA11), which included the two with the largest number of ATA repeats produced uneven PCR amplification and these were not considered in the diversity analysis.

While a majority of the microsatellites produced single bands, multi-copy banding was observed for 11 markers as listed in Table 1 and these were also not evaluated for polymorphism information content. Several markers (ATA30, ATA108, ATA150, ATA241) produced double banding patterns, which were interpretable for polymorphism information content. Stutter bands were more common for the AT-rich microsatellites, but in these cases the top band was

Fig. 1 Amplification pattern of three polymorphic AT-rich microsatellites on the genotype survey of 18 genotypes



generally stronger than the rest of the bands and was used for allele calling.

The average number of alleles found for each class of microsatellites varied with the motif. AT motif microsatellites were the most variable showing on average 12 alleles across the 18 genotypes used in the parental survey. GA motif microsatellites were also very variable showing 8 alleles per locus on average, followed by ATA motif microsatellites with 5.8 alleles per locus on average. Meanwhile, the GT motif microsatellites were highly monomorphic, averaging only 1.5 alleles per locus.

Polymorphism information content was related to average number of alleles and again the AT motif microsatellites had the highest value (0.892), compared to GA (0.706) and ATA (0.600) motif microsatellites. The CA motif microsatellites had very low PIC values of 0.140. The highest PIC values were observed for the marker ATA76 (0.858) and the gene-based microsatellites HRG (0.905) and ME1 (0.878) that had AT motifs, as well as the complex, GA and GT motif marker GA16 (0.874). The next highest PIC values were found for the ATA motif markers, ATA7, ATA9, ATA20, ATA24, ATA150, ATA170, ATA180, ATA234, ATA241 and ATA296, all with PIC values above 0.800 and 7–11 alleles per locus. One other GA motif marker, CAC1 showed a large number of alleles (12) and a high PIC value (0.827). The remaining markers showed varying degrees of polymorphism except for four of the ATA motif markers (ATA73, ATA121, ATA157 and ATA236) and two of the GT motif markers (CA3 and GA4) that were monomorphic. Across all the microsatellites tested, there was an average of 5.89 alleles detected and an average polymorphism information content of 0.590, which was higher than the values reported in Blair et al. (2006a) for a mix of GA motif and cDNA-based microsatellites (PIC value = 0.534); in Benchimol et al. (2007) for mostly GA- and CA-based microsatellites (PIC value = 0.45); and in Métais et al. (2002) who analyzed more closely related germplasm and had PIC values averaging 0.440 and allele number averaging 5.3.

Correlations between the number of alleles and the PIC values were high across the range of all microsatellites ($r = 0.878$, $P = .0000$) and for both AT-rich ($r = 0.877$, $P = .0000$) and non-AT-rich ($r = 0.897$, $P = .0000$) microsatellites individually. Meanwhile the number of alleles and PIC values were not significantly correlated with the number of repeat motifs within the microsatellites ($r = -0.273$ and $r = -0.113$, respectively). However, the allele size range between largest and smallest allele was correlated with the number of alleles present for a locus ($r = 0.623$, $P = 0.0000$) and the PIC value ($r = 0.434$, $P = 0.001$). The correlations of allele size range with number of alleles and PIC values were also significant when considering AT-rich

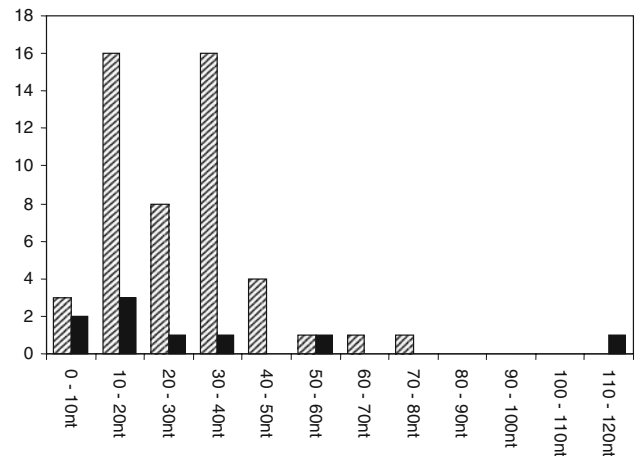


Fig. 2 Distribution of allele size ranges for AT-rich (hatched bars) and non-AT-rich (solid bars) microsatellites on the panel of 18 common bean genotypes

($r = 0.599$ and $r = 0.386$, respectively) and non-AT-rich ($r = 0.903$ and $r = 0.826$) microsatellites, separately.

Figure 2 shows that the distribution of allele size ranges for the markers developed from the AT-rich versus the non-AT-rich microsatellites. The average allele size that ranges across all the markers was 28.2 nt. Interestingly, a binomial distribution can be observed for the AT-rich microsatellites, whereas the non-AT-rich microsatellites show a skewed distribution towards low allele size ranges. The size differences between the largest and the smallest alleles were 24.4% wider on average for the AT-rich microsatellites (28.0 bp) compared to non-AT-rich microsatellites (22.5 bp). For most of the markers (71.2%), the expected allele size based on the Fin de Bagnols source (Métais et al. 2002) was within the range between the largest and the smallest allele sizes for each marker.

Genetic mapping of microsatellites

Overall polymorphism rate was 74.6% across the parents of the DOR364 × G19833 population with a total of 50 microsatellites amplified for genetic mapping. The polymorphism rate was higher in the AT-rich microsatellites (77.5%) compared to the non-AT-rich microsatellites (55.5%). Both were within the range of polymorphism for the same population as analyzed for other genomic microsatellites analyzed by Blair et al. (2006a). From this effort, a total of 45 loci could be placed on the DOR364 × G19833 genetic map (Fig. 3). All microsatellites were placed at a LOD > 3.0 and integrated with the markers mapped by Blair et al. (2003). New microsatellite loci were found on each of the 11 linkage groups of the species, however the AT-rich microsatellite loci tended to be located at relatively few genomic locations.

While each linkage group was found to have at least one new microsatellite locus; linkage groups, b02, b09 and b11 had larger numbers of microsatellite loci placed on them, (8, 6 and 7 markers, respectively) when compared to the other linkage groups where 4 or fewer loci each mapped. Clustering of the ATA microsatellite loci occurred on four linkage groups, especially on linkage group b02 where markers were mapped to two small regions in the middle of the linkage group (indicated as cluster A), as well as in regions on linkage groups b08 and b10. The marker ATA154 clustered with a group of GA motif markers from Blair et al. (2003) on linkage group b03 (indicated as cluster B); while the markers ATA150a and ATA173 clustered with another group of GA-based BM markers on linkage group b07 (indicated as cluster C).

In contrast, the microsatellite loci were better distributed on linkage groups b01, b04 and b09; while linkage group b11 had a cluster of three ATA microsatellites in the middle of the linkage group and more evenly distributed markers at the extremes. While the majority of the mapped microsatellites were mapped as single loci, duplicate loci were mapped for ATA108 on linkage groups b01 and b11; for ATA150 on linkage groups b07 and b10; and for ATA241 on linkage groups b10 and b11. Mapping of the other double banding marker, ATA30 was not possible due to monomorphism in the population.

Discussion

In this study, we first determined the utility of AT-rich microsatellites in terms of detecting allele variability and polymorphism information content. In general terms, we found the AT-rich microsatellites to be very useful markers, similar to results from several other legumes where these motifs have been targeted (Cregan et al. 1999; Huttel et al. 1999; Udupa et al. 1999; Udupa and Baum 2001; Udupa et al. 2004; Ferguson et al. 2004; Song et al. 2004). We also found that AT motif microsatellites tend to be more polymorphic than GA-based microsatellites, that ATA microsatellites are approximately equal or higher in polymorphism with GA motif markers and that CA microsatellites are of low polymorphism in common beans based on results from this study and comparisons to previous work with predominantly GA-, CA- and cDNA-based markers from Blair et al. (2006a, b) and Benchimol et al. (2007). These results are similar to those for soybean and chickpea, where AT and ATA based microsatellites were highly productive for marker development (Huttel et al. 1999; Song et al. 2004), and with the results from peanuts, where AT markers had more alleles than ATA-, GA- or GT-based markers (Ferguson et al. 2004). AT and ATA motif SSRs are also abundant and highly polymorphic in

rice (Temnykh et al. 2001). While AT and ATA markers are arguably the most valuable markers within a species, it is notable that AT-rich microsatellites are not likely to be conserved as SSR loci between different genera of legumes such as chickpea, lentil and pea (Choumane et al. 2004).

Earlier efforts for microsatellite development in plants tended to target GA and CA motif SSRs because of their utility in human populations (Hearne et al. 1992) and because of their ease of detection in hybridization or enrichment-based screening protocols. Concurrently, ATA motif microsatellites tended to be more difficult to develop due to the low-melting temperature of ATA primers used to develop enriched libraries (Métais et al. 2002) and ATA probes used to hybridize with non-enriched libraries (Cregan et al. 1999). AT motif microsatellites likewise have been difficult to develop since AT sequences are palindromic and DNA with long stretches of dinucleotide AT repeats are difficult to clone and to sequence (Powell et al. 1996). Despite these difficulties, both ATA and AT motif microsatellites have been developed for peanuts, chickpea and soybean through hybridization based selection using non-enriched libraries (Cregan et al. 1999; Huttel et al. 1999; Ferguson et al. 2004). Our study also shows that it is possible to create libraries that are enriched for ATA based microsatellites and that this works well for marker development in common beans, although a certain amount of ATA markers are difficult to amplify. These results build on those of Métais et al. (2002) who developed the first ATA microsatellites, Gaitán-Solís et al. (2002) who attempted ATA enrichment but without success and Caixeta et al. (2005) who discovered a limited number of ATA motif microsatellites based on subcloning of common bean BACs. As was found for other crops, sequencing of positive clones for ATA-motifs can also yield compound microsatellites including some with AT-motifs. Interestingly, AT motifs were common among microsatellites uncovered in the sequencing of a common bean RFLP library (Murray et al. 2002; Blair et al. 2003) and were the most prevalent motif in a previous search of common bean Genbank sequences by Yu et al. (1999) surpassing GA, ATA or other trinucleotide repeats although they were rather rare in the EST screening performed by Hanai et al. (2007) where AT and ATA motifs were 15.3 and 2.5% of the totals identified among 240 SSRs found for over 3,000 unigenes.

Another finding of our study was that while AT-rich microsatellites are comparable or better than GA motif microsatellites developed from enriched libraries there was variability in terms of the diversity uncovered by each marker. As in the study by Ferguson et al. (2004), the diversity at AT-rich microsatellite loci does not appear to be related to size of the repeat array found in the SSR loci in common bean. This could be due to the frequency of imper-

fect SSR arrays we found for ATA motif microsatellites. Interruptions in repeat arrays have been suggested to serve as anchors that prevent strand slippage that is thought to be the generator of diversity at microsatellite loci (Chambers and MacAlvoy 2000).

In general, the average number of repeats found in this study was similar to the number of repeats in previously developed microsatellites from enriched libraries, averaging 15.8 across all the newly developed markers compared to 15.4 and 16.0 for microsatellites from the GA enriched libraries made by Gaitán-Solís et al. (2002) and Buso et al. (2006), respectively. However, some very long ATA, AT and GA microsatellites (over 30 repeats) were identified in this study and some of these are highly polymorphic. Hanai et al. (2007) meanwhile found many compound microsatellites in their GA and CA motif library enrichment with repeat numbers ranging from 2 to 20 and averaging 6.8 for any single SSR stretch.

In addition to the imperfect repeat loci identified, we also found a few compound repeat loci with some of these were fairly diverse (ATA76) in terms of allele number while others were not (ATA2 and ATA268). The most polymorphic of these loci was unique in having two adjacent repeat arrays, namely (ATA)₁₂ and (TA)₈ separated by 29 base pairs. A similar ATA and AT compound locus was carefully analyzed by Udupa and Baum (2001) in 114 chickpea accessions with the finding that it was also highly variable and that the variability was due to expansion and contraction of the number of repeats. They also found, in contrast to our study, that the threshold number of repeat arrays for the two individual motifs within the locus was reciprocally correlated so that the overall size of the locus did not expand beyond a given limit. This difference may have been due to the chickpea locus having a longer repeat array than our common bean locus. For ATA motif microsatellites in common bean, variability at the long repeat marker ATA20 was studied by Métais et al. (2002) and found to be due to increases and decreases in repeat number rather than to the changes in the regions flanking the repeats.

A valuable characteristic of the AT-rich microsatellites, especially not only the AT-based markers but also the ATA markers, was that their high level of polymorphism allowed a large proportion of the markers to be genetically mapped. A total of 45 new loci could be mapped on the DOR364 × G19833 integrated map. The polymorphism of AT-rich markers in this population was over 10% higher than that for GA-based microsatellites as reported in Blair et al. (2006a, b). With the genetic mapping of microsatellites previously conducted in the species by Yu et al. (2000), Blair et al. (2003) and Grisi et al. (2007), this brings to over 160, the number of mapped loci for common bean. An interesting observation was that an intermediate number

of new loci were mapped to the middle and ends of linkage group b01, and to linkage group b09, two regions that suffer from low diversity in other crosses and was difficult to map in Blair et al. (2003) or Grisi et al. (2007).

As in the earlier study, we found that a certain number of microsatellites detected double loci although the majority was single-copy markers. Duplicate loci could have resulted from segmental duplication events or similarity in primer sequences across more than one locus. Some AT-rich microsatellites have been found to be associated with repetitive elements that could be dispersed throughout the genome (Ramsay et al. 1999; Temnykh et al. 2001), which might explain why we identified duplicate loci for some of the new microsatellites and also why we obtained a group of microsatellites that produced multiple bands and that therefore were not useful as single copy markers. Other discoveries from the genetic mapping of ATA microsatellites was that their loci cluster rather than being randomly distributed and that a greater number of loci were found on linkage groups b02, b09 and b11 compared to other linkage groups. This same phenomenon of clustering was observed for GA-based microsatellites from enriched libraries that were mapped by Blair et al. (2003), however these were found on different linkage groups, such as b03, b04, although clustering on linkage group b02 was in common with this study. Indeed our mapping shows that some ATA motif markers fall into clusters of GA based BM markers especially on linkage groups b03 and b07.

Furthermore, the ATA-based microsatellites tended to be at interstitial locations rather than at distal locations on the linkage groups. This may reflect a bias of ATA motif sequences against telomeres or towards proximal locations. Clustering of genomic microsatellites has been associated with retrotransposons and may reflect their distribution (Ramsay et al. 1999) and appears to be more common with microsatellites from enriched libraries (Ramsay et al. 2000; Tang et al. 2002). In a few cases, microsatellites from enriched genomic libraries or those associated with retrotransposons have been found to be clustered in or around centromeric regions (Areshchenkova and Ganai 1999; Jones et al. 2002; Ramsay et al. 2000). The centromeres in common bean have not been accurately mapped and therefore it is impossible to tell if the ATA microsatellite clusters are actually located near them although progress in the physical mapping of common bean may allow this in the future (Pedrosa et al. 2003). Clustering is usually not evident with microsatellites derived from gene sequences (Scott et al. 2000; Cordeiro et al. 2001) or from BAC ends (Mun et al. 2006).

In conclusion, the targeting of ATA motifs for enrichment was a valuable approach to marker development in common bean. Along with the AT motif microsatellites and additional GA markers, this study provides a set of well characterized genetic markers that can be used for diversity

assessment or genetic mapping, especially in populations derived from narrow crosses. The increase in microsatellites based on the ATA tri-nucleotide motif is useful for diversity assessment, since these markers are often easier for allele calling than GA microsatellites due to the larger size difference in repeat units of the trinucleotide motif markers which allow them to be used in automated multiplex fingerprinting. Future work will concentrate on mapping the ATA motif microsatellites in additional populations and obtaining ATA or AT motif microsatellites from BAC end sequences, given that other sources for this motif such as ESTs are not very productive (this laboratory, unpublished results). It is important likewise, to know how often AT and ATA or other motifs occur in non-enriched genomic libraries. Overall our aim has been to produce informative microsatellite markers for whole genome analysis, gene tagging, marker-assisted selection and diversity evaluation.

Acknowledgments We appreciate help with parental survey analysis from Lucy Díaz and Alma Viviana Gonzalez and with seed preparation from Agobardo Hoyos. We are especially grateful to Patrice Carreau (Vilomorin SA, La Ménittré, France) for support to the library development project and acknowledge funding from Vilmorin SA, Univ. of Angers and CIAT.

References

- Afanador LK, Hadley SD (1993) Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean. *Bean Improv Coop* 35:10–11
- Areshchenkova T, Ganai MW (1999) Long tomato microsatellites are predominantly associated with centromeric regions. *Genome* 42:536–544
- Benchimol LL, de Campos T, Carbonell SAM, Colombo CA, Chioratto AF, Formighieri EF, Gouvêa LRL, de Souza AP (2007) Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. *Genet Resour Crop Evol* 54:1747–1762
- Blair MW, Pedraza F, Buendia H, Gaitan E, Beebe S, Gepts P, Tohme J (2003) Development of a genome wide anchored microsatellite for common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 107:1362–1374
- Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, Beebe S (2006a) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 113:100–109
- Blair MW, Iriarte G, Beebe S (2006b) QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean × wild common bean (*Phaseolus vulgaris* L.) cross. *Theor Appl Genet* 112:1149–1163
- Blair MW, Díaz JM, Hidalgo R, Díaz LM, Duque MC (2007) Microsatellite characterization of Andean races of common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 116:29–43
- Broughton WJ, Hernandez G, Blair M, Beebe S, Gepts P, Vanderleyden J (2003) Bean (*Phaseolus* spp.); model food legumes. *Plant Soil* 252:55–128
- Buso GSC, Amarala ZPS, Brondani RPV, Ferreira ME (2006) Microsatellite markers for the common bean *Phaseolus vulgaris*. *Mol Ecol Notes* 6:252–254
- Caixeta ET, Borém A, Kelly JD (2005) Development of microsatellite markers based on BAC common bean clones. *Crop Breed Appl Biotechnol* 5:125–133
- Campos T, Benchimol LL, Carbonell SAM, Chioratto AF, Formighieri EF, de Souza AP (2007) Microsatellites for genetic studies and breeding programs in common bean. *Pesqui Agropecu Bras* 42:589–592
- Chambers GK, MacAlvay ES (2000) Microsatellites: consensus and controversy. *Comp Biochem Physiol Part B* 126:455–476
- Choumane W, Winter P, Baum M, Kahl G (2004) Conservation of microsatellite flanking sequences in different taxa of *Leguminosae*. *Euphytica* 138:239–245
- Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ (2001) Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to *Erianthus* and *Sorghum*. *Plant Sci* 160:1115–1123
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, Vantoai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. *Crop Sci* 39:1464–1490
- Díaz LM, Blair MW (2006) Race structure within the Mesoamerican gene pool of common bean (*Phaseolus vulgaris* L.) as determined by microsatellite markers. *Theor Appl Genet* 114:143–154
- Edwards KJ, Baker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques* 20:759–760
- Ferguson ME, Burow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, Mitchell S (2004) Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *Theor Appl Genet* 108:1064–1070
- Frei A, Blair MW, Cardona C, Beebe SE, Gu H, Dorn S (2005) QTL mapping of resistance to *Thrips palmi* Karny in common bean (*Phaseolus vulgaris* L.). *Crop Sci* 45:379–387
- Freyre R, Skroch PW, Geffory V, Adam-Blondon AF, Shirmohamadali A, Johnson WC, Llaca V, Nodari RO, Periera PA, Tsai SM, Tohme J, Dron M, Nienhuis J, Vallejos CE, Gepts P (1998) Towards an integrated linkage map of common bean. 4 Development of a core linkage map and alignment of RFLP maps. *Theor Appl Genet* 97:847–856
- Gaitán-Solis E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris*): isolation, characterization, and cross-species amplification in *Phaseolus* ssp. *Crop Sci* 42:2128–2136
- Graham PH, Vance CP (2003) Legumes: importance and constraints to greater use. *Plant Physiol* 131:872–877
- Grisi MCM, Blair MW, Gepts P, Brondani C, Pereira PAA, Brondani RPV (2007) Genetic mapping of microsatellite markers in common bean (*Phaseolus vulgaris*) population BAT93 × Jalo EEP558. *Genet Mol Res* 6:691–706
- Hanai LR, de Campos T, Camargo LEA, Benchimol LL, de Souza AP, Melotto M, Carbonell SAM, Chioratto AF, Consoli L, Formighieri EF, Siquiera MF, Tsai SM, Vieira MLC (2007) Development, characterization and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources. *Genome* 50:266–277
- Hearne CM, Ghosh S, Todd JA (1992) Microsatellites for linkage analysis of genetic traits. *Trends Genet* 8:287–294
- Hopkins MS, Casa AM, Wang T, Mitchell SE, Dean RE, Kochert GD, Kresovich S (1999) Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Sci* 39:1243–1247
- Huttel B, Winter P, Weising K, Choumane W, Weigand F, Kahl G (1999) Sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). *Genome* 42:210–217
- Jones ES, Dupall MO, Dumsday JL, Hughes LJ, Forster JW (2002) An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). *Theor Appl Genet* 102:405–415

- Lander ES, Green P, Abrahamson J, Barlow A, Daly M, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Liu K, Muse SV (2005) Powermarker: integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128–2129
- Métais I, Hamon B, Jalouzot R, Peltier D (2002) Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. *Theor Appl Genet* 104:1346–1352
- Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant J* 3:175–182
- Mun JH, Kim DJ, Choi D-H, Gish J, Debelle F, Mudge J, Denny R, Endre G, Saurat O, Duzde A-M, Kiss KB, Roe B, Young ND, Cook DR (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. *Genetics* 172:2541–2555
- Murray J, Larsen J, Michaels TE, Schaafsma A, Vallejos CE, Pauls KP (2002) Identification of putative genes in bean (*Phaseolus vulgaris*) genomic (Bng) RFLP clones and their conversion to STSs. *Genome* 45:1013–1024
- Pedrosa A, Vallejos CE, Bachmair A, Schwizer D (2003) Integration of common bean (*Phaseolus vulgaris*) linkage and chromosomal maps. *Theor Appl Genet* 106:205–212
- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1:215–222
- Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol* 5:94–100
- Ramsay L, Macaulay M, Cardle L, Morgante M, Degliivanisovich S, Maestri E, Powell W, Waugh R (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *Plant J* 17:415–425
- Ramsay L, Macaulay M, Ivanisovich S, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuvevsson S, Morgante M, Massari A, Maestri E, Marmiroli N, Sjakste T, Ganai M, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. *Genetics* 156:1997–2005
- Rota ML, Kantety RV, Yu JK, Sorrells ME (2005) Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley. *BMC Genomics* 6:23–32
- Scott KD, Egger P, Seaton G, Rossetto EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. *Theor Appl Genet* 100:723–726
- Shultz JL, Kazi S, Bashir S, Afzal JA, Lightfoot DA (2007) The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. *Theor Appl Genet* 114:1081–1090
- Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109:122–128
- Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. *Theor Appl Genet* 105:1124–1136
- Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch SR (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res* 11:1441–1452
- Udupa SM, Baum M (2001) High mutation rate and mutational bias at (TAA)_n microsatellite loci of chickpea (*Cicer arietinum* L.). *Mol Genet Genomics* 265:1097–1103
- Udupa SM, Robertson LD, Weigand F, Baum M, Kahl G (1999) Allelic variation at (TAA)_n microsatellite loci in a world collection of chickpea (*Cicer arietinum* L.) germplasm. *Mol Gen Genet* 261:354–363
- Udupa SM, Malhotra RS, Baum M (2004) Tightly linked di- and trinucleotide microsatellites do not evolve in complete independence: evidence from linked (TA)_n and (TAA)_n microsatellites of chickpea (*Cicer arietinum* L.). *Theor Appl Genet* 108:550–557
- Yaish MWF, Pérez de la Vega M (2003) Isolation of (GA)_n microsatellite sequences and description of a predicted MADS-box sequence isolated from common bean (*Phaseolus vulgaris* L.). *Genet Mol Biol* 26(3):337–342
- Yu K, Park SJ, Poysa V (1999) Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome* 42:27–34
- Yu K, Park SJ, Poysa V, Gepts P (2000) Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *J Hered* 91:429–434