Theor Appl Genet (2004) 108:1151–1161
DOI 10.1007/s00122-003-1533-4

ORIGINAL PAPER

T. Pugh · O. Fouet · A. M. Risterucci · P. Brottier · M. Abouladze · C. Deletrez · B. Courtois · D. Clement · P. Larmande · J. A. K. N'Goran · C. Lanaud

A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers

Received: 24 June 2003 / Accepted: 10 November 2003 / Published online: 4 February 2004

Abstract A linkage map of cacao based on codominant markers has been constructed by integrating 201 new simple sequence repeats (SSR) developed in this study with a number of isoenzymes, restriction fragment length polymorphisms (RFLP), microsatellite markers and resistance and defence gene analogs (Rgenes-RFLP) previously mapped in cacao. A genomic library enriched for $(GA)_n$ and $(CA)_n$ was constructed, and 201 new microsatellite loci were mapped on 135 individuals from the same mapping population used to establish the first reference maps. This progeny resulted from a cross between two heterozygous cacao clones: an Upper-Amazon Forastero (UPA 402) and a Trinitario (UF 676). The new map contains 465 markers (268 SSRs, 176 RFLPs, five isoenzymes and 16 Rgenes-RFLP) arranged in ten linkage groups corresponding to the haploid chromosome number of cacao. Its length is 782.8 cM, with an average interval distance between markers of 1.7 cM. The new microsatellite markers were distributed throughout all linkage groups of the map, but their distribution was not random. The length of the map established with only SSRs was 769.6 cM, representing 94.8% of the total map. The current level of genome coverage is approximately one microsatellite every 3 cM. This new reference map provides a set of useful markers that is transferable across different mapping populations and will allow the identification and comparison of the most important regions involved in the variation of the traits of interest and the development of marker-assisted selection strategies.

Communicated by H. Nybom

T. Pugh (✉) · O. Fouet · A. M. Risterucci · M. Abouladze · C. Deletrez · B. Courtois · D. Clement · P. Larmande · C. Lanaud
UMR 1096, CIRAD-BIOTROP,
TA 40/03, 34398 Montpellier Cedex 5, France
e-mail: pugh@cirad.fr
Tel.: +33-4-67655831
Fax: +33-4-67615605

T. Pugh
UCV-FAGRO, Av El Limón, 2101 Maracay, Venezuela

P. Brottier
GENOSCOPE, Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 Evry, France

J. A. K. N'Goran
CNRA, Centre National de Recherches Agronomiques,
BP 582 Abidjan, Ivory Coast

Introduction

Genetic mapping is a basic tool of genomic research. Molecular linkage maps provide information about the organization of the genome and may be used for genetic studies and breeding applications. Several linkage maps have already been published for *Theobroma cacao* L. (Lanaud et al. 1995; Crouzillat et al. 1996; Risterucci et al. 2000). These maps were based on codominant markers such as restriction fragment length polymorphisms (RFLPs), isoenzymes and a small number of simple sequence repeats (SSRs) and, in some of the maps, associated with dominant amplified fragment length polymorphisms (AFLP) and random amplified polymorphic DNA (RAPD) markers. They have been used to locate quantitative trait loci (QTLs) affecting traits of interest, such as disease resistance and yield factors (Crouzillat et al. 2000; Flament et al. 2001; Clement et al. 2003; Queiroz et al. 2003; Risterucci et al. 2003).

Dominant markers require a relatively small investment with respect to time and cost and are commonly used to increase the density of linkage maps due to their multilocus properties. Nevertheless, the fact that they are dominant induces a loss of information that can be detrimental for a heterozygous crop. A linkage map for cacao based on codominant markers could provide more information. PCR-based codominant markers like microsatellite markers can easily be transferred to other populations and shared between research laboratories.

Microsatellites, or SSRs, are present in the majority of eukaryotic genomes and consist of simple, short tandemly repeated di- to penta-nucleotide sequence motifs (Beckman and Soller 1990). The allelic variation in microsatellite loci can easily be detected by PCR using specific

flanking primers. Polymorphism based on variation in the number of repeated motifs is probably due to slippage during DNA replication or unequal crossing-over (Levinson and Gutman 1987). Microsatellites have been widely used in many crop species due to their abundance, high degree of polymorphism, locus specificity, reproducibility, low amount of DNA required, suitability for multiplexing on automated systems and, above all, their codominant mode of inheritance. These characteristics make SSRs an attractive option for increasing the density of the cacao linkage map. Nevertheless, the development of microsatellite markers is generally time-consuming and costly. Several procedures have been developed to produce SSR-enriched gDNA libraries. Billote et al. (1999) described an easy method for developing microsatellite markers in tropical crops that has subsequently been used with success (Billote et al. 2001; Aranzana et al. 2002; Dirlewanger et al. 2002).

In the past few years, a new generation of linkage maps based on SSR markers has been constructed for several species—sorghum (Bhattramakki et al. 2000; Haussmann et al. 2002), wheat (Gupta et al. 2002), rice (Temnykh et al. 2000; McCouch et al. 2002), peach (Dettori et al. 2001), almond (Joobeur et al. 2000), apple (Liebhard et al. 2002) and *Prunus species* (Aranzana et al. 2003). In addition, microsatellite markers have been developed and mapped in cacao (Lanaud et al. 1999; Risterucci et al. 2000; Lanaud et al. 2004).

In the investigation reported here, we developed and mapped a new set of 201 cacao microsatellite markers onto the previous cacao linkage map.

Materials and methods

Production of microsatellite markers and primer design

A genomic library enriched for $(GA)_n$ and $(CA)_n$ was constructed from the cacao (*Theobroma cacao* L.) clone Catongo based on the procedure described by Billote et al. (1999). The enrichment step was pursued as described in Kijas et al. (1994) with some modifications. Genomic DNA was digested with *Rsa*I (Invitrogen, Carlsbad, Calif.), the fragments were ligated to 5′-end phosphorylated adaptors constituted with the following two following primers: *Rsa*21: CTCTTGCTTACGCGTGGACTA and *Rsa*25: P-TAGTCCACGCGTAAGCAAGAGCACA (Edwards et al. 1996). The ligated fragments were amplified with *Rsa*21 primer. Following hybridization with a 5′-biotinylated microsatellite oligoprobe $I_5(CT)_8$ and $I_5(GT)_8$, they were selected for the presence of microsatellites with Streptavidin MagneSphere Paramagnetic Particles (PMPs) (Promega, Madison, Wis.). The selected fragments were amplified with *Rsa*21, and the partial genomic library was then constructed by ligating the PCR products into a pGEM-T plasmid (Promega). Epicurian-coli XL1-Blue supercompetent cells (Stratagene, La Jolla, Calif.) were used for the transformation of the cloned DNA fragments. Subsequently, white transformant clones were transferred onto Hybond-N+ nylon membranes (Amersham, UK) and simultaneously hybridized using [^{32}P]-labelled microsatellite oligoprobes $(GA)_{15}$ and $(GT)_{15}$. Selected microsatellite-containing clones were sequenced by Genoscope (France) using its current sequencing protocol (Artiguenave et al. 2000).

Duplicates of sequenced SSRs were removed using SEQUENCHER 4.0 (Gene codes, Ann Arbor, Mich.). The primers flanking the microsatellite repeat sequences were designed using the OLIGO 4.0 software (National Biosciences, USA). The main criteria for primer design were to produce well-matched primers that were 16–24 nucleotides long, had an average GC content ranging between 40% and 50% and an annealing temperature between 45°C and 55°C and were preferably G- or C-rich at the 3′ end.

PCR amplification protocols and detection of polymorphism

PCR reactions were performed on a MJ Research PTC Thermal cycler (MJ Research, Waltham, Mass.) in 20-µl volumes containing 10 ng of cacao DNA, 0.2 μM of a 5′-endlabelled γ-[^{33}P]ATP forward primer, 0.2 μM of reverse primer, 2 mM $MgCl_2$, 0.2 mM dNTP mix, 50 mM KCl, 10 mM Tris-HCl (pH 8.3) and 1 U *Taq* polymerase (Eurobio, France). The PCR profile was: an initial denaturation step at 94°C for 4 min, followed by 35 cycles of 30 s at 94°C, primer annealing at a specific annealing temperature (46°C or 51°C) for 1 min and 1 min at 72°C; this was completed with a final extension at 72°C for 5 min.

Twenty microliters of loading buffer (98% deionized formamide, 10 mM EDTA, bromophenol blue and xylene cyanol) was added to individual reactions. Samples were denatured at 94°C, and 5 µl of each sample was subjected to electrophoresis at 55 W on 5% denaturing polyacrylamide gels containing 75 M urea in 0.5× TBE buffer (pH 8.0). The gels were dried and exposed for 48–72 h to X-ray film (Eastman Kodak). Polymorphism revealed by the newly synthesized primer pairs was tested using DNA from the parents of the mapping population.

Mapping population

The newly developed microsatellites were mapped on 135 individuals from the same mapping population as that used to establish the map described in Risterucci et al. (2000). This progeny resulted from a cross between two heterozygous cacao clones: an Upper-Amazon Forastero (UPA 402) and a Trinitario (UF 676). Because the parents were heterozygous, there were two possibilities for single-locus segregation—those which segregated in only one of the parents (1:1) (heterozygous for one parent and homozygous for the other parent), and those which segregated in both parents (1:2:1 or 1:1:1:1) (heterozygous in both parents). Codominant markers segregating according to the latter possibility can be used as genetic bridges for aligning the linkage information from each parental dataset to produce a consensus linkage map (Grattapaglia and Sederoff 1994)

Linkage map construction

The segregation of each microsatellite marker was tested with a chi-square test for goodness-of-fit to the expected Mendelian segregation ratio function of the parental configuration. The map was produced using JOINMAP version 3.0 (Van Ooijen and Voorrips 2001) by integrating the SSR loci reported in this paper, the RFLPs, SSRs, isoenzymes and resistance and defence gene analogues previously mapped in cacao by Risterucci et al. (2000) and Lanaud et al. (2004). JOINMAP software is able to combine data of several segregation types to construct an integrated map. When the parental phase is unknown, this software chooses the best option of association between coupling/repulsion phases of a set of markers based on the recombination frequencies observed within a linkage group. A LOD score of 5.0 was used to identify linkage groups. The Kosambi mapping function was used to convert recombination frequencies into map distances (Kosambi 1944)

Marker nomenclature

All loci are designated according to the nomenclature guidelines presented by Risterucci et al. (2000) and Lanaud et al. (2004): SSRs

were denoted as mTcCIRX where m corresponds to microsatellite, Tc to *Theobroma cacao*, CIR to CIRAD and X to the microsatellite number. RFLP probes were named cTcCIR, gTcCIR and rTcCIR corresponding to cDNA, genomic and isolated RAPD genomic fragments, respectively. TELX corresponds to telomeric markers; NX/X, PTX/X and PRx/X correspond to probes homologous to the N gene (Whitham et al. 1994), Pto gene (Martin et al 1994) and PR proteins (Kauffman et al. 1987), respectively. NX corresponds to genomic probes provided by Nestlé, and CX or CaX to microsatellite markers also provided by Nestlé.

Results

Microsatellite development

Of the 705 positive clones sequenced, 154 (22%) were redundant; 463 unique sequences (66%) were found to contain a microsatellite while 88 did not contain any microsatellite. We subsequently designed 387 primer pairs (63%) that flanked the microsatellite motifs; the remaining clones were not exploited because the SSRs were too short or too close to one sequence end.

Microsatellite amplification and polymorphism

Sequenced primer pairs flanking the microsatellite motifs were tested for amplification within the parents of the mapping population. Among the 387 primer pairs tested, 227 (59%) detected polymorphism between the two parents of the mapping population, with 203 (52%) providing a clear single-locus amplification and unambiguous profile. These latter primer pairs were used for the construction of the linkage map. All three types of repeats, as defined by Weber (1990), were found among the 203 SSRs. One hundred and fifty SSRs (74%) were derived from dinucleotide motifs with a perfect microsatellite stretch. The most frequent motifs were $(CT)_n$ or $(GA)_n$ (77%), as would be expected for a library based on selection for (GA)n repeats, 21 (10%) were interrupted and 32 (16%) were compounds. Table 1 provides detailed information on the mapped markers developed in this study. The size of the amplified fragments was in agreement with that expected from sequenced data.

Linkage analysis and map construction

Ninety-one percent of all markers fitted the Mendelian ratio expected from the genotype of the parents. Only 42 of the markers (18 RFLPs, two isoenzymes, one Rgene and 21 SSRs) deviated significantly ($P<0.05$) from the expected ratio; these included 20 newly developed microsatellite markers.

If all of the markers are taken into consideration, bridge loci (heterozygous in both parents) represented 26% of the scored patterns, 66% segregated for UF 676 only, and 8% segregated for UPA 402 only. When we constructed the two parent maps separately (data not shown), good co-linearity between the two maps was observed, with only small modifications in the relative positions of closely linked markers, especially inversions—for example, mTcCIR18 and mTcCIR199 on linkage group (LG) 4. Genetic distances estimated between loci in the UF 676 or UPA 402 parent maps were generally of the same magnitude (e.g. 1.5 cM vs. 2.1 cM between mTcCIR188 and mTcCIR12, respectively or 4.1 cM vs. 5.4 cM between C104 and mTcCIR115 on LG4), even though genetic distances estimated in the UPA 402 map were slightly larger than those in the UF 676 map. This evidence allows us to use bridge loci as anchor markers to construct an integrated map.

The resulting order of the loci and map distances between markers (in centiMorgans) are shown graphically in the integrated linkage map reported in Fig. 1. The complete map contains 465 codominant markers (67 previously mapped SSRs, 201 new SSRs, 176 RFLPs, five isoenzyme loci and 16 Rgenes-RFLPs) arranged in ten linkage groups corresponding to the haploid chromosome number of cacao. Only one new SSR locus (mTcCIR146) remained unlinked at LOD 5.0. Another microsatellite marker, mtcCIR132, which produced an important rearrangement in the order of LG 2, was removed from the map.

The complete map is 782.8 cM long, with an average distance of 1.7 cM between markers. The number of mapped loci for each linkage group varies from 29 on LG8 to 61 on LG5. The most saturated linkage group in the map contains 61 markers and covers 73.2 cM with 1.2 cM between loci (LG5). Some distorted microsatellite markers appear to be clustered on LG3. Examination of the direction of segregation distortion showed that they all segregate for the male parent UF 676 only. Skewed segregations concerning both parents were also observed in loci of LGs 5 and 6.

Among the new SSRs, 116 (58%) were heterozygous for UF 676 only, 13 (6%) were heterozygous for UPA 402 only, and 72 (36%) were heterozygous in both parents (bridge markers) with two, three or four alleles (Table 2). Microsatellite markers were distributed throughout all linkage groups of the map, but their distribution was not random. The number of SSRs per group ranged from 11 in LG10 to 36 in LG3. The map length with only SSRs was 769 cM, representing 94.8% of the total length of the map. The current level of genome microsatellite coverage is approximately one microsatellite every 3 cM. Nevertheless, some SSR markers were clustered and formed dense linkage blocks in LGs 1, 3, 5 and 9. No gap longer than 10 cM was found when all of the codominant markers were considered. With respect to microsatellite markers alone, linkage groups were well covered, but there were small gaps (10–20 cM) between microsatellites on LGs 2, 6, 7 and 10. The number of loci that were heterozygous in both parents varied between linkage groups from 23 on LG9 to zero on LG8. The relative order of the markers determined on previous maps (Lanaud et al. 1995; Risterucci et al. 2000) was preserved, and only a few inversions were observed (e.g. cTcCIR49

Table 1 Characteristics of SSR loci isolated from a cocoa genomic library enriched for $(GA)_n$ and $(CT)_n$ and integrated in the codominant marker-based map: locus designation, EMBL accession number, primer sequences, linkage group (LG) location, expected PCR product size in the reference variety (Catongo), description of the repeat motif and calculated primer annealing temperature (T_m) using OLIGO 4.0 software

SSR name	EMBL accession number	5′–3′ Forward primer	5′–3′ Reverse primer	LG location	Expected size (bp)	Repeats	T_m (°C)
mTcCIR64	AJ566412	GAGAAAGTAAAAGGAGAGAG	TGTTAGAGAAATGAGAAGTG	9	167	$(GA)_{11}$	46.5
mTcCIR65	AJ566413	CATGAAAGCTAAGTGCCT	AAAAATGCGTTACAAGTGTG	3	242	$(GA)_{10}$	50.7
mTcCIR66	AJ566414	TATCCGCCAGAAACAGA	CCAACAGTAGAGTCCAGAGT	1	304	$(AG)_{20}$	50.1
mTcCIR67	AJ566415	ATAGCTCCTTTTGACACGA	TTCTCTTTTCCACCTCTTT	4	109	$(GA)_{10}$	47.9
mTcCIR68	AJ566416	ATTTAGCTGTAGCCGTTT	CCAGTTGATCTGCTTAATG	2	166	$(GA)_{17}$	49.6
mTcCIR69	AJ566417	TCGGTGTTCCATCAGTA	CATGCTATGAGATTGAAAG	5	203	$(CT)_{20}$	46.8
mTcCIR70	AJ566418	GGTATGAAGGATTGAGAG	TTCCTATTCGTATTTATGGG	8	107	$(GA)_{11}$ AA $(GA)_4$	44.4
mTcCIR71	AJ566419	CGACTAACCAGCAGAAAC	CTCCTCCTCCTCCAT	6	170	$(GA)_{10}$	47.5
mTcCIR73	AJ566420	CCAGTCAAGGAAGTATCT	AATGTCTGCAATGTTAGC	2	112	$(CT)_4$ TT $(CT)_2$ G $(TC)_8$	45.8
mTcCIR75	AJ566421	CATTCATCTTTTTCTTTCTCTC	TCCTTCTCCAACCGA	8	121	$(AG)_{10}$	48.5
mTcCIR76	AJ566422	AGCCAAAGAAAGGAT	TGAATCCGAGACAAAG	4	139	$(TC)_{21}$ TTT $(TC)_{58}$	46.2
mTcCIR77	AJ566423	GTTCTCCCCCACTCTCT	AATAAATAAATAAACAATACG	10	287	$(TC)_9$	47.8
mTcCIR78	AJ566424	TGAAAATACGTTCTGTCTGA	CAAAAAGTTTTCTGAAAGTC	3	159	$(TC)_2$ T $(TC)_9$	47.7
mTcCIR79	AJ566425	ATTTTTCTTTAGCGCACT	TAACTACCTTCCCACCTC	9	108	$(TC)_8$	48.9
mTcCIR80	AJ566426	GCTGGGGTTTCTTTGTATT	TTCTCATTTCTTTATTGGGTT	5	105	$(CT)_{10}$ CC $(CT)_2$	46.8
mTcCIR81	AJ566427	TGAAACTCCCATACTACTGA	ACAATCTGTCCATTATTCTG	3	216	$(CT)_{15}$	47.9
mTcCIR82	AJ566428	ATCATGTGCCCCTTCTAA	GGCAGCTAAGTGTTCATTC	3	174	$(AG)_6$ AA $(AG)_7$	47.9
mTcCIR84	AJ566429	CATGGGACGCTGCCT	CTCTTATTAAATTGAATTCTCT	1	136	$(GA)_{11}$	47.1
mTcCIR85	AJ566430	TTGAAGTAGAGAGTTGTAAGAA	TTATGGTGTTGGTTGTGAT	1	211	$(AG)_{16}$	46.7
mTcCIR86	AJ566431	TAACAAGGAAAATGCTCTC	GTTGAACCGAAGGAAAG	5	370	$(CT)_8$ TT $(CT)_6$	48.9
mTcCIR87	AJ566432	TAAGGGGGCAACATAAAT	CAAATAGCGCAGAGACAAT	5	145	$(AG)_{21}$	48.0
mTcCIR88	AJ566433	CTAGGATTCCATAGAAGTAA	TTGGACCTCAATTATATGT	1	187	$(CT)_{10}$	47.0
mTcCIR90	AJ566434	CCACTTCAAAAACCATTCTA	GCAACTGTCAACCATTATCTA	9	291	$(CT)_{10}$	47.6
mTcCIR91	AJ566435	TTTTGCTGAGTGTTGCTGT	ATCCGAGAAATAGAATAGGTTA	10	186	$(CT)_{10}$	47.5
mTcCIR92	AJ566436	GTTCCAAATCATCTCACTT	TCGTCTATTTCTCTTCACTCT	10	283	$(AG)_9$	46.4
mTcCIR93	AJ566437	GTTGCCACTGCTCTCGCT	CCCTTTTATTGTTCCCCATTA	7	100	$(CT)_8$	48.0
mTcCIR94	AJ566438	AATTTGTAGGGTATTTGAAGAG	CCATGCCCAGTGAGTAG	1	196	$(AG)_{16}$	51.0
mTcCIR95	AJ566439	CTCCTTCCCTTTCTCTC	CATCGTCTTCCTCTCATC	4	221	$(TC)_4$ CC $(TC)_{21}$	47.9
mTcCIR96	AJ566440	ATAGAGAAAGACCCAAATC	AGACAACAAATTTATGTAATG	9	136	$(TC)_8$	47.3
mTcCIR97	AJ566441	CTTCTTCTGGTCAATCTTCT	GCTGCATCCATCATCC	1	122	$(TC)_3$ C $(TC)_{10}$	46.9
mTcCIR98	AJ566442	CCAGTTGCTAATTTTCTTC	GCACATAGTTTTGGCAAT	9	140	$(TC)_8$	46.2
mTcCIR99	AJ566443	CGGAAAATGAAACAGAC	AATAAAAGAAAAGAACCATAC	8	249	$(GA)_9$	49.0
mTcCIR100	AJ566444	TGATGGAATAAACTAAGAACA	TAAGAAGCCAGGTCAGG	2	244	$(AG)_6$ C $(AG)_4$	48.2
mTcCIR101	AJ566445	GAACTCACCGAAGAGAAAC	GAGCCGTCACCAATGC	5	109	$(TC)_{18}$	50.6
mTcCIR102	AJ566446	TTGTGAAAAGATTGCGA	TTGCTTGTTATTGCTACTAT	1	124	$(GA)_9$	46.0
mTcCIR103	AJ566447	GAGAGATGGCTTAAGGAT	ACCATACTATTGAAACATTG	8	112	$(GA)_{10}$	46.5
mTcCIR104	AJ566448	AATAGGAAAGGGTAAGTGAAT	CAAGCATATAAAGCCAACA	10	167	$(GA)_{10}$	47.8
mTcCIR105	AJ566449	GTTTACAACTTATCGCTCTG	AATTTGTATCCCTTATTATTTA	3	201	$(CT)_8$	46.1
mTcCIR106	AJ566450	ACGAAAAATACCCTAAAAA	TGCTGTTGTTGTCTTGCT	1	143	$(GA)_9$	45.8
mTcCIR107	AJ566451	TTGCCTGGAAGAGAGA	GATGGAAAGAGAAATAATAGT	4	120	$(AG)_{11}$	46.8
mTcCIR109	AJ566452	GGAAAGTGTAGGAAAGTAGAC	GGACCAAAAAGAGCATA	5	162	$(CT)_{12}$	46.4
mTcCIR110	AJ566453	GTGAAAAGTGGGGATTG	TAAAGTAAGAGTGGTGATGGT	7	139	$(AG)_8$	48.2
mTcCIR112	AJ566454	TTACTTTGTAGGCTGCTG	CATTCCACTCATTTTGCT	10	95	$(TC)_8$	46.0
mTcCIR113	AJ566455	GGAAAGTTACAGCAAGAGAGA	ACAAGCCCGGTGAAGG	7	142	$(AG)_9$	50.7
mTcCIR114	AJ566456	CAGATGAATGGAATAACTT	GCATGAACACAAACACAC	9	207	$(TC)_9$ $(TG)_5$ G $(TG)_4$	48.7
mTcCIR115	AJ566457	GTGATTCAAATTCAAATATG	AATAGCAAGAGAGTGATGAG	4	191	$(TC)_{11}$	46.3
mTcCIR117	AJ566458	TGTGGAATAAAAGAGCAAT	CACTGGTGTAGCAAATGATA	4	168	$(TC)_{10}$	47.2
mTcCIR118	AJ566459	TCTGCCTGAAAATGTCTC	TGGGGCACTAACTTTTG	1	165	$(GA)_{10}$	47.4
mTcCIR119	AJ566460	TGGACTTGTGCTGGAAC	GCAAGAAATAAAATAGGAAC	5	123	$(AG)_{12}$	47.8
mTcCIR120	AJ566461	TGGAAAGTGCTTACTCTTATG	TCTAGTTTCAGGGGCTCT	3	95	$(AG)_{13}$	46.2
mTcCIR121	AJ566462	CATGTGCATTTAGGTGTC	TCTGGCTTCTTAGTGATAC	1	138	$(TG)_{12}$	46.8
mTcCIR123	AJ566611	ATTCCCTTAGCTTTATGTTATG	CTCGCGCCCTTTCTCT	5	172	$(CA)_4$ $(TG)_6$	51.3
mTcCIR124	AJ566463	CAGCGTCTTGGAATAAC	ACCCACACACAAGACAC	9	131	$(CT)_{12}$	46.2
mTcCIR125	AJ566464	CATGCAAATGCTTAGG	TGAACCACAGCTGACAC	9	98	$(TG)_{11}$	45.2
mTcCIR126	AJ566465	AACTCTCACTATCATCCAC	AACAAATCATCAAACACTT	9	212	$(GA)_{11}$	46.3
mTcCIR127	AJ566466	CGTTTGTCTTGCGTTC	ATGTGTTTTTGCCTCTTAC	5	130	$(TC)_8$	48.6
mTcCIR129	AJ566467	CAGTGAGGATGAGGTTC	CGACATACCAGTTTACATAA	2	129	$(TC)_{16}$	46.8
mTcCIR130	AJ566468	ACCGGCGGCTGATCTAC	CGCGCCAACCAATAAAG	1	133	$(AG)_{17}$	51.0
mTcCIR131	AJ566469	TGAGTAAGAAAAAGTAGAAAA	GATCATCGGTAAAGTAAAAT	3	204	$(GA)_9$ C $(GA)_4$	46.9

Table 1 (continued)

SSR name	EMBL accession number	5′–3′ Forward primer	5′–3′ Reverse primer	LG location	Expected size (bp)	Repeats	T_m (°C)
mTcCIR133	AJ566470	GGATCACATCCGTTTAGA	AATTTCAGCCCTCCAA	3	155	$(AG)_{11}$	49.0
mTcCIR134	AJ566471	CGTCCCAAAATCAACAC	ATAGTCTGCCCTTCCAA	8	176	$(GT)_{15}$	47.1
mTcCIR135	AJ566472	ATTAGAGAGGGGGTAGATGA	CTAGTGGGGTTGACATTTG	3	246	$(AG)_{20}$	50.0
mTcCIR136	AJ566473	GAGGAGGTGAGAGCCA	GGTTTGTATTTTTGATTGAG	6	232	$(GA)_7$ GC $(GA)_7$	49.3
mTcCIR137	AJ566474	CAGCTGTCACGGAAAC	GCCTCTTACCCTCATT	1	114	$(TC)_{10}$	46.5
mTcCIR138	AJ566475	CTGCCAAGTCAAGTAAGTTC	CTGTGGTATCAATCAATCTAAT	1	128	$(CA)_{11}$	46.8
mTcCIR140	AJ566476	GATTCATAGTGGAAACACAGT	GGAAAACAGAGAGGAAGAGT	3	104	$(CA)_7$	48.1
mTcCIR141	AJ566477	TGTTGCATAAAACACGAGTTC	CCTAAAATCCTTCCTAACAGC	7	217	$(TC)_{14}$	51.9
mTcCIR142	AJ566478	CCATTTACAACTCCCATTCATA	AACTATCCATCCACCCTACCTC	9	130	$(GA)_8$	48.4
mTcCIR144	AJ566479	CCACTGACACGCAATGAA	CTAGGACTTAGGAAAGTGTTTG	3	254	$(TC)_9$	49.6
mTcCIR145	AJ566480	CAGACTTCCAACTCAAAACT	TGAGAATAGATGGACCGAT	9	117	$(CT)_{17}$	49.7
mTcCIR147	AJ566481	TGAAGCAATTTGAAATCTGT	AACCACATCATAATGATTTAAG	7	303	$(TC)_{16}$	47.8
mTcCIR148	AJ566482	CGTCCTATCACTTCTCTTTC	TTGCCTTTACAGCCATTT	5	235	$(TC)_6$ CC $(TC)_{06}$	50.7
mTcCIR151	AJ566483	CAGGGGCTCTGTGTTT	ACCAAGAACGGGGAGA	2	139	$(GA)_{12}$	50.5
mTcCIR153	AJ566484	GCCTCTCACACCATTATCTG	TACATTCATTACTTCACTGCTG	3	217	$(TC)_9$	50.1
mTcCIR154	AJ566485	CCTTGTAAGTGTGGCAAT	TGGAACAAGAGGTTGTCA	9	167	$(GA)_{14}$	49.0
mTcCIR155	AJ566486	CTTGGACTATTGGAAAAC	AAGGATACAATAAGGTAAATAC	10	274	$(TC)_{12}$	46.5
mTcCIR156	AJ566487	GGCAGGACCAAATGAT	AAAACCAGGAACACCAG	5	216	$(CT)_{11}$	51.3
mTcCIR157	AJ566488	ACTAATGCTGTTGGCTTC	TCACTCGACTCGACTGTC	9	151	$(AG)_{09}$	49.6
mTcCIR158	AJ566489	TGTAGGTTATGCAGCGTGTTC	GATGAGGGGTGTAGCTGTTTG	4	213	$(CT)_8$	50.2
mTcCIR160	AJ566490	GATTGTTGTTTGGTATGC	GTGAAGGTGAAGGTGTG	9	288	$(GA)_8$	48.4
mTcCIR162	AJ566491	AAGATTGAGGTCACTCAGG	TAAGTTTTTTGCTTTACTCTTC	2	162	$(GA)_{19}$	48.0
mTcCIR163	AJ566492	CATAACGCAGACCAGTGT	TTTGATCATCGGCTTG	8	194	$(AG)_9$	48.3
mTcCIR164	AJ566493	AGAACGGTTCAGGACAATC	AGGACAATGATGAAGAAATAAG	3	117	$(CT)_8$	49.2
mTcCIR165	AJ566494	TTCACTTCCCTCCCCAC	CTGGGTTTGGAGTAGCTTG	2	139	$(CT)_{11}$	51.1
mTcCIR166	AJ566495	ATGAACCACTATGTAAGACC	ATTCCAAAGGATTAGCAG	9	215	$(CT)_9$ $(CA)_8$	48.2
mTcCIR167	AJ566496	GTAGAACCATAAACAACATT	ACAATCATTAAAAATACGAG	3	254	$(GA)_{16}$	46.1
mtcCIR168	AJ566497	GGTAGTATTGAGGTGCGTAT	GTGAATGAATGGATGTGAAA	4	175	$(TC)_9$	48.5
mtcCIR169	AJ566498	CTTTTGGCTGTATGTTCG	CTGCCCTCTCTTTCTCAC	5	178	$(GA)_9$ AA $(GT)_7$	51.0
mTcCIR170	AJ566499	CTCTTGCACGGCACAGGA	TTGCCCCACCCATACG	5	131	$(TG)_7$	50.0
mTcCIR172	AJ566500	CGTTCCAGTGTGGGTGA	TGTTTTCGCTCTACTGCTTC	9	127	$(TG)_9$ $(AG)_4$	51.1
mTcCIR173	AJ566501	TCCGGATGGCAATATGT	CTACCCCATGATTCTGAACT	3	148	$(CA)_7$	49.9
mTcCIR174	AJ566502	TGGCAGCAATACTTCAAA	TCCCGATGTTCCACTC	1	167	$(TG)_7$	49.5
mTcCIR175	AJ566503	TTACAATCAACAGAACCTC	TATATTGATGCGAAAGTC	3	248	$(CA)_7$	47.2
mTcCIR176	AJ566504	TCACCAATTCTCTGCTC	AATGAAATTACCTCCTTAC	2	106	$(TG)_{16}$	46.2
mTcCIR177	AJ566505	GATCCTTGAAACCACACA	TAATTTCTCTTTACACATTCTC	7	128	$(CA)_7$	47.4
mTcCIR178	AJ566506	CATCTGTTTGCACATATTTG	GCTTGGTCCCTTAACAC	9	138	$(TG)_8$	48.5
mTcCIR179	AJ566507	TTTCCATTTCTCATTTCTCAAG	ATGTTTTCATTTTCGTATCCAA	7	288	$(GT)_{16}$	50.7
mTcCIR180	AJ566508	ATGGTTTCGATTGTCTGT	CAAAATCTAAGCTGATAAAAC	3	186	$(GT)_9$	46.4
mTcCIR181	AJ566509	CTTTATGCTGTCTCTCGTA	CCAAGAATGTTTTGATACTG	7	197	$(CT)_{12}$ $(CA)_9$	47.5
mTcCIR182	AJ566510	CTAATTGTTCAAGGAGGTC	AAGTGTTTTTGGCACTATC	6	148	$(TG)_9$	46.3
mTcCIR183	AJ566511	GTTATCTTAGTTTCTAGCCAC	GTAGTCTTACACCTTGATTG	4	353	$(AC)_9$	45.9
mTcCIR184	AJ566512	GGTTTTCTAGCTCCTCC	AGGAAAGAATGACTCATACTA	1	139	$(CA)_8$ $(CT)_{13}$	48.2
mTcCIR185	AJ566513	ATCCCCCTGCCTAAAGAG	CCTGAATGAAGTAAGACCCAAT	6	142	$(CA)_{18}$	50.0
mTcCIR186	AJ566514	AAGGCTAAAGAACAAATG	CGTAGACGTCACACAATA	7	147	$(TG)_8$	46.5
mTcCIR187	AJ566515	TTCACCTAGTGTAATGGTCT	GCAGGCTTCAATTTAGAG	9	262	$(TG)_8$	49.4
mTcCIR188	AJ566516	GTCCTATTCGGTGTAACTAC	TTCTGCTTCTCTTTGCTC	4	118	$(TC)_9$ $(AC)_8$	48.0
mTcCIR189	AJ566517	GAATAGAAATTTTATGCAC	TCAAACAACATAGGTCAC	8	150	$(GT)_{12}$	45.9
mTcCIR190	AJ566518	AAGAAACTGAAGCACAAT	CACAAAGAGCATAAACTG	7	166	$(TG)_{12}$	46.7
mTcCIR192	AJ566519	TCACTTCACAATAATTTCAAG	AAATTGAATTCCAGTTGTAG	3	98	$(TG)_{16}$	46.0
mTcCIR193	AJ566520	AACCTGTGATGGACCG	AAATGGTGAATTAGGCTC	6	134	$(TG)_9$	48.9
mTcCIR194	AJ566521	ACACACAGCCTAAAACGAAA	GGGATGTGACGGATATTTAC	1	192	$(TG)_{14}$	47.7
mTcCIR195	AJ566522	CAAGTTGAATAAAGTCCTAAG	AAAATAAAGAAAATGAAGTAA	2	350	$(CA)_{10}$	46.6
mTcCIR197	AJ566523	GGATTTATTTTATTGTAAACTCC	AATGATTTCTACATTGTCACCA	5	162	$(AT)_9$ $(GT)_{18}$	46.2
mTcCIR198	AJ566524	TGGGACCATAAGGAAATC	CCCAGGTGAAGTAAGACA	3	186	$(CA)_3$ TA $(CA)_6$	46.3
mTcCIR199	AJ566525	GATTCTTATTTGATTTTCCTTA	GCACGGTTACATTTATTACA	4	211	$(TG)_{14}$ TC $(TG)_7$	47.4
mTcCIR200	AJ566526	GCCAATTTCTGACCCA	CTTAAAATAAGCCCAAATAC	8	238	$(TG)_8$	47.3
mTcCIR202	AJ566527	TCTCTCATAGCTCAAGCA	CCTGAGTCAAAGTGTCCT	3	172	$(TG)_7$ $(GA)_9$	48.3
mTcCIR203	AJ566528	GTGGATTTGGGTGGGAT	ATTGTGTTTTGGCTATGTTC	1	217	$(AC)_8$	50.7
mTcCIR204	AJ566529	ATTACCTGCCGATGAAG	TGGGTTTGGAATGATGT	3	131	$(CT)_{10}$	47.3
mTcCIR205	AJ566610	GGGGTTTGTTTTATGTAT	TGTGGGATCGTCTTCT	9	198	$(TC)_8$	47.5
mTcCIR207	AJ566530	TGGTTGACAAGGTAAAA	TGGATGTGCAAGTAAGT	4	174	$(TC)_9$	45.4
mTcCIR208	AJ566531	GCAAGCCCCTAAAACT	AAAAAGCAAAAGAAGAAGA	6	209	$(CT)_{10}$-25pb-$(AC)_{11}$	48.3
mTcCIR209	AJ566532	TACGGGCTAATGGTGA	AGGTATGCTGTATTTATGGT	6	259	$(TG)_6$ TAT $(GA)_9$	47.8
mTcCIR210	AJ566533	CAAACCCCAAACCTCAA	CAGTTATGGAAATTATTGCTCTA	1	146	$(AG)_{11}$-7pb-$(AAG)_4$	49.5
mTcCIR211	AJ566534	TGGTGCTAACTCAAATC	CAAACAAGAAGGCTAAA	8	182	$(TC)_9$	46.5
mTcCIR212	AJ566535	GAGAAACACTTCAGGATAC	GTCATTTGGCAGATTTA	9	186	$(TC)_8$	46.6

Table 1 (continued)

SSR name	EMBL accession number	5′–3′ Forward primer	5′–3′ Reverse primer	LG location	Expected size (bp)	Repeats	T_m (°C)
mTcCIR213	AJ566536	GATCTCGCAAAACTAACA	TAAGTAAAATGAAGGTGTGA	4	261	$(CT)_{26}$	47.9
mTcCIR215	-	GCTTCAACTCCAAATCAC	TAGCATCCCGTATTGTG	9	197	$(AG)_{13}$	49.1
mTcCIR216	AJ566537	ACTGCCCAGGAATCA	TCTTTGTTTCTGCCTTAT	5	158	$(GA)_{12}$	47.4
mTcCIR217	AJ566538	AGTTTCCATCTATATTGTTA	TATTGTCTACGGTTCTCT	4	132	$(CT)_{11}$	43.4
mTcCIR218	AJ566539	TGACCAAGGAAGCTCTC	GGTGGGAAAGGTGGTA	8	187	$(CT)_{11}$	48.9
mTcCIR219	AJ566540	GCGAACCAAGACAAATAC	ATGGGTGGCAATTTCT	3	187	$(AG)_8$	49.6
mTcCIR220	AJ566541	TGAAGTGTGTGTTTGTGTA	CCAATAGAGGGATGTAATA	10	201	$(TC)_{10}$	47.8
mTcCIR221	AJ566542	ATGTAGTTGGGCTGTGA	TGTTAAGAGGGAAATGAA	4	273	$(TC)_9$	48.6
mTcCIR222	AJ566543	CTACAGAAAATAGGCAATA	TCATTGTATTATCAGGTAGA	4	220	$(GA)_9$	45.2
mTcCIR223	AJ566544	GGTCCACACTCAACACT	TTATTCCATTTTCATTTACT	10	202	$(TC)_4$ GC $(TC)_2$ GC $(TC)_{15}$	45.0
mTcCIR224	AJ566545	TCAGAAAGCAATGTGGTA	AAGCAATATCAAGTGGTAAG	2	223	$(TC)_{13}$ $(AC)_8$	48.0
mTcCIR225	AJ566546	AAGACAAAGGGAAGAAGA	AGGGGAAGAGCAAATC	8	302	$(TC)_{10}$	49.8
mTcCIR226	AJ566547	TAACCCAAATTCAAAGTC	TTTCAACAGCCTCATCT	3	246	$(TC)_{11}$	47.3
mTcCIR227	AJ566548	ACATCATTAAGGAGAAACA	CAAATCACCTCAAATAATC	2	142	$(CT)_8$	46.4
mTcCIR228	AJ566549	CCCCCTGATACTGTGTG	GAAACCTAATCTCGTAATATGT	2	110	$(GA)_8$	48.7
mTcCIR229	AJ566550	ATCTCGGTAATAGCACATAA	CGCAATCCTACAACACA	10	307	$(TC)_8$	47.9
mTcCIR230	AJ566551	GTGGAAGCCTTATGATTATGT	ATTTATGCCCATGCAGAC	2	231	$(CT)_8$	49.5
mTcCIR231	AJ566552	AGGAGGAGTTGCTGAA	CAGGTTCCAATTTTGTAT	4	226	$(AG)_8$	47.3
mTcCIR232	AJ566553	GCTGTTGTCTACTTTTGAAT	CACCCTTTGAATCAGTCTA	5	205	$(TC)_{18}$	49.8
mTcCIR233	AJ566554	CCAGAAGCCAAAAGAGA	ATGGATTAAGAAGGAGGAA	4	211	$(GA)_{15}$	48.9
mTcCIR234	AJ566555	TTGTGTCGGTTTGATTC	GAAAGAGAGGGAAAGTGA	4	123	$(TC)_9$	48.1
mTcCIR235	AJ566556	TTCGGATGGCAACTAACT	AAAACAGCGGAACAGGTA	6	292	$(AG)_8$	49.1
mTcCIR236	AJ566557	GAAGTCAAAGGAAAGTCAA	TCAGAAAACGCAAATAAA	8	193	$(CT)_5$ TT $(CT)_{14}$ T $(CT)_2$	50.1
mTcCIR237	AJ566558	GAAGACAAGGATGGAGACT	GCAAAGAGAGCAGGAGA	4	103	$(TC)_{10}$	50.1
mTcCIR238	AJ566559	TTGGCTTTCTTTTAGTTA	AAATATAATCATTACTTTCCTA	6	126	$(AG)_9$	44.1
mTcCIR239	AJ566560	CTCCACAGTCAAAATAACA	TTAAATCCCGCAAAGT	5	203	$(CT)_9$	47.4
mTcCIR240	AJ566561	CATACCTACTACTGCTCTCT	AGTGATTTATGGGACTTT	2	158	$(CT)_{22}$	46.5
mTcCIR241	AJ566562	CAGTTGGAGGGCATTT	ACGAGTGAGAGAGTGAAGTT	4	146	$(CT)_{23}$	50.2
mTcCIR242	AJ566563	TTTCGGCATTCCACTA	GTAAAAACAATAATCTTCAACTA	4	287	$(CT)_9$ $(CA)_9$	48.2
mTcCIR243	AJ566564	ACAGCAGTAGACGCATTC	AAAAGGCTTGGCACAG	4	141	$(TC)_9$-20pb-$(CA)_{11}$	50.6
mTcCIR244	AJ566565	TGGCAATAACAATGAACA	ATTTTGATGATTGATGAAGA	1	264	$(TA)_4$ CATA $(CA)_{17}$ $(TA)_4$	47.2
mTcCIR245	AJ566566	GCAAATAGACAGCAAAT	TTCAAAGGAGTATAGGTAA	5	198	$(GT)_8$	45.8
mTcCIR246	AJ566567	TATCCTCTCTCTCTGTGTATC	GCAGCACTCAACCACTA	1	169	$(TG)_8$	47.6
mTcCIR247	AJ566568	CATTTTATAATTCCTTCT	ACATTCTTTATTTTCACT	3	111	$(AC)_9$ATAC $(AT)_3$	39.9
mTcCIR248	AJ566569	TGATAGATTTGCGTTTACA	CCCAGAAAAGAAGAAGAT	5	190	$(TG)_8$	48.0
mTcCIR249	AJ566570	TCTCAAGTTCAAGGGTCT	GACACAAATGCCGTTAT	1	246	$(CT)_4$ TT $(CT)_{28}$ $(AC)_{16}$	47.9
mTcCIR250	AJ566571	CCCAGAGGACCATCAC	ACTGCTCTCTCTCTACTCATC	9	237	$(AC)_{22}$	49.6
mTcCIR251	AJ566572	TCTATGGGATTTGATGAG	AGATACAGCAGGAACACA	9	188	$(CT)_7$ $(CA)_{12}$	46.8
mTcCIR252	AJ566573	AATGTGTGCTTTGTTTCTA	TTCAAGGGCGTAACTC	2	155	$(AC)_{10}$	45.8
mTcCIR253	AJ566574	TGGCTACTAAACACCTACTA	GGGAGGGGAGTAAGTT	2	155	$(AC)_7$	45.4
mTcCIR254	AJ566575	ACAACTCCAAAGAACAAG	GGTAAACCTCGTCATAAT	3	198	$(AC)_{21}$ $(AT)_9$	45.3
mTcCIR255	AJ566576	TTTACCTCCACCATCTT	TGGCACTTATCTATTACTGT	6	203	$(AC)_{11}$	47.5
mTcCIR256	AJ566577	AGAAGGGTGTCAACATTA	GAACAGTCAAACATAAGAGTA	5	185	$(AC)_{13}$ $(ATAC)_4$	46.1
mTcCIR257	AJ566578	CATACAGAAACCAGAAAAT	TATAGGGTAAAGCGAAAT	5	167	$(CT)_7$ $(CA)_{12}$	48.0
mTcCIR258	AJ566579	TAACTCACAATCCATCAT	ATGGTCATTATCAAAATC	8	116	$(TC)_7$ $(AC)_{10}$	44.5
mTcCIR259	AJ566580	TTTCCTGATTTCCATTA	AGAGGTTCCAAAATACAT	5	157	$(CA)_8$	45.2
mTcCIR260	AJ566581	TGGCAACACATACATTA	GTGTATGCCTAGATGAGA	2	112	$(AC)_{17}$	44.6
mTcCIR262	AJ566582	GTTTCTTTGCTCCGTATCT	TTTGCCAACCCTGTGT	1	165	$(TC)_{14}$	47.6
mTcCIR263	AJ566583	ACCAGGAGTTTTTCTTG	ATTAGTCAGCTCATCATTAT	3	244	$(TC)_9$	47.3
mTcCIR264	AJ566584	TGCTATCCACAACCAGT	TAACTCACTTTTGCCACTA	1	192	$(CT)_8$	47.0
mTcCIR265	AJ566585	TGAATGCTGGAAAAATGT	GTGTCTGCTTTGGTTTGT	5	246	$(AG)_{18}$	49.2
mTcCIR266	AJ566586	TCGTCGCCATCATAGA	GTCGTTATTCGGAGTTCA	9	192	$(CT)_{15}$	50.7
mTcCIR267	AJ566587	CACTACCCCTTTTCCTT	TTCATGGCTCTTTTTCTAT	5	199	$(GA)_9$	49.4
mTcCIR268	AJ566588	TGTAATCCAAATAATAAGCAT	CAGTGAAGAGGCAAGAGA	2	316	$(GA)_{17}$ GG $(GA)_9$	49.5
mTcCIR270	AJ566589	TTAGTGAGATGGTGACAAT	AATCAAGGAAAAGTTATCA	1	224	$(CT)_{10}$ $(CA)_9$	46.1
mTcCIR271	AJ566590	GCACCTTTGTTATCTTG	GGAACCCACTGAAACT	5	173	$(CT)_{12}$	45.6
mTcCIR272	AJ566591	TTTGCCTTTCCTTTCT	TTTGTCAATTTGGATAGTG	1	258	$(CT)_5$ $(TC)_4$-151pb-$(AT)_4$ $(CA)_7$	48.5
mTcCIR273	AJ566592	AGAATGATCGCAGAGAG	ACGGCATTAGAGAGAGA	1	167	$(CT)_4$ AC $(CT)_{13}$ TT $(CT)_4$	47.3
mTcCIR274	AJ566593	GAAAGGTAAATGGCTGAA	CGATCATCACGACTGCT	5	184	$(CT)_6$ CACG $(CA)_6$ $(CT)_2$	49.7
mTcCIR275	AJ566594	GGTTTGGTTTTGGTAAGAC	TAAGAGAGAGTGATGCTGACA	1	146	$(CT)_{11}$	53.0

Table 1 (continued)

SSR name	EMBL accession number	5′–3′ Forward primer	5′–3′ Reverse primer	LG location	Expected size (bp)	Repeats	T_m (°C)
mTcCIR276	AJ566595	TCCTGCTTTTTAATACAT	GTCCTATCTGCCTCACT	6	124	$(GA)_{14}$	46.3
mTcCIR277	AJ566596	ACCAAGATCAAAGTCAAGAA	GATAAGAACCAAGTGAAGAGA	7	304	$(AG)_{16}$ AA $(AG)_{11}$	50.4
mTcCIR278	AJ566597	TGGCATCTGTCTGTC	GTATATGACCGTTTGTAG	3	100	$(TCTG)_3$ $(TC)_{10}$ $(TA)_8$	41.9
mTcCIR279	AJ566598	GTCCATCTACATCATAAGC	CAGCAACAGCATCACT	5	155	$(CT)_9$	44.1
mTcCIR280	AJ566599	ATTTGTCATTTGTTGTTGT	GCCTTGGTATTGACTGT	3	89	$(CT)_8$ $(CA)_9$	44.9
mTcCIR281	AJ566600	CCGCTGTTTTGGTATTTC	GGATGAGGGGTGGTTG	2	194	$(TC)_{12}$ $(CA)_{14}$	51.4
mTcCIR282	AJ566601	TGGTGAGGGGAGAGAA	AGCAAAGGCAATAATAATG	8	172	$(GA)_2$ GG $(GA)_6$	49.4
mTcCIR283	AJ566602	ATCAATACCCACCACACA	CCCTTTTCCTCTTTTTCT	1	239	$(TC)_{11}$	49.3
mTcCIR285	AJ566603	TACTACCTCTACCCTCTTGT	ATAAATTCCTTCCCTTCT	9	216	$(AG)_{16}$	46.7
mTcCIR286	AJ566604	GTTCTGCTTCATCTGTTTA	TTCAACCCACAACCAT	1	119	$(CT)_{18}$	46.2
mTcCIR287	AJ566605	TCCTTTCTGTTTGTTCCT	TTATCCGTGTCTCCTTCT	9	301	$(TC)_9$	48.2
mTcCIR288	AJ566606	ACAACACAAGGCAAAGA	CCCATTTAGCACCAAC	5	184	$(GA)_{10}$ AA $(GA)_3$-36pb-$(GA)_{11}$	48.8
mTcCIR289	AJ566607	CTTCCGCCACTAATAAA	CTATACATAACAGCAGCCA	3	123	$(CT)_{10}$	46.8
mTcCIR290	AJ566608	AGCGAGAGACAAAGATAAT	GACTGAAATGGTGGTAAAG	6	175	$(CT)_{19}$ CACC $(CA)_{18}$	49.4
mTcCIR291	AJ566609	AGTCCCATAGGTTCCAAT	CGAGGTTATCCCCAAA	6	218	$(CT)_{12}$	50.1

Table 2 Number of polymorphic SSRs obtained from the enriched genomic library and type of segregations observed in the mapping population

	Total	UPA 402 clone	UF 676 clone
Number of primer pairs screened for polymorphism	387		
Number of polymorphic loci	223		
Number of primer pairs analysed in the progeny	201		
Segregating 1:1	129	13	116
Segregating 1:1:1:1	72		
Two alleles	8		
Three alleles	34		
Four alleles	30		

and gTcCIR122 on LG3 where two new microsatellites segregating in both parents were mapped between them).

Discussion

We describe here the development and mapping of a new set of 201 microsatellite markers which were integrated with a set of previously mapped codominant markers (SSRs, RFLPs, isoenzymes and Rgenes-RFLPs). This enabled us to obtain a new codominant marker-based cacao linkage map with 465 codominant markers arranged in ten linkage groups.

A small proportion of the new loci showed segregation distortion when compared to the expected Mendelian ratio. The reasons behind this distortion remain unclear, but if we consider the fact that RFLP markers showing distorted segregation have previously been mapped in these regions (Risterucci et al. 2000), biological mechanisms such as selection against closely linked lethal or sublethal genes, linkage with genes that are subject to direct selection or the presence of incompatibility alleles, rather than chance or error, could be hypothesized. Segregation distortions have already been observed in tree species (Viruel et al. 1995; Kijas et al. 1997; Barreneche et al. 1998; Dettori et al. 2001). Although the inclusion of distorted loci into the map increases the chance of type I errors of false linkage, these loci can be useful in increasing our knowledge of specific regions and in the mapping of QTLs.

Our map is shorter than previously published ones that include AFLP markers. It is known that AFLPs derived from certain combinations, such as *Eco*RI/*Mse*I selective primers, are often clustered in some specific regions of the genome, especially AT-rich heterochromatic regions around centromeres and at chromosome ends (Boivin et al. 1999). In the previous maps, several AFLP markers were located around telomeric regions and their removal decreased map length.

The distribution of microsatellite markers within the linkage groups was not random. The tendency of SSR loci to cluster has been reported in several species, including barley (Ramsay et al. 2000), sorghum (Bhattramakki et al. 2000), rye grass (Jones et al. 2002) and rice (Mc Couch et al. 2002). It has been suggested that this is influenced by the non-uniform distribution of recombination events in the mapping population (reduced recombination frequency) (Castiglioni et al. 1999).

The addition of new SSR markers allowed us to fill some gaps present in the previous map (especially on

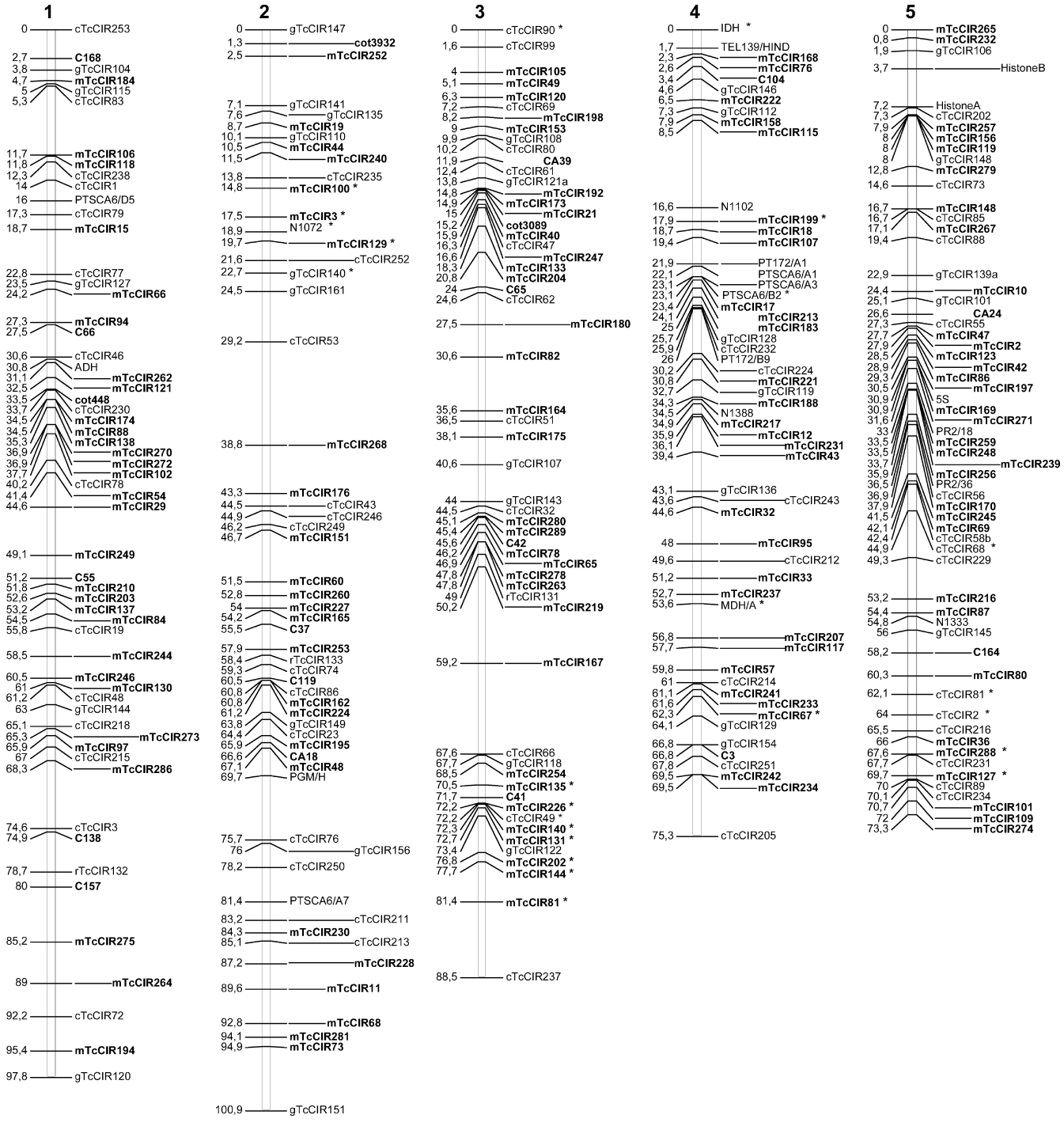

Fig. 1 Linkage map of cocoa based on the cross UPA 402 × UFA 676. The map consists of 465 codominant markers (269 SSR loci, 178 RFLPs, five isoenzymes and 16 Rgenes). Symbols for SSR loci are in *bold type*. Polymorphic markers in the gametes of UPA 402, in the gametes of UFA 676 and in the gametes of the both parents are designated on the *right*, *left* and *middle* respectively. Markers showing distorted segregation ratios are denoted with an *asterisk*

regions of LGs 3, 6 and 8) and to saturate some regions (LGs 4 and 5), but it did not enable the saturation of the distal region of LG 10, where only one new SSR was mapped in 15 cM and the average distance between markers is 3.5 cM longer than the average interval between markers (1.7 cM) in the whole map. This suggests that low polymorphism exists between the two parents of the mapping population in this region. The non-random distribution of polymorphism along chromosomes may reflect some structural or functional properties of the DNA in these parts of the genome or the dynamics of domestication and breeding process accompanied by

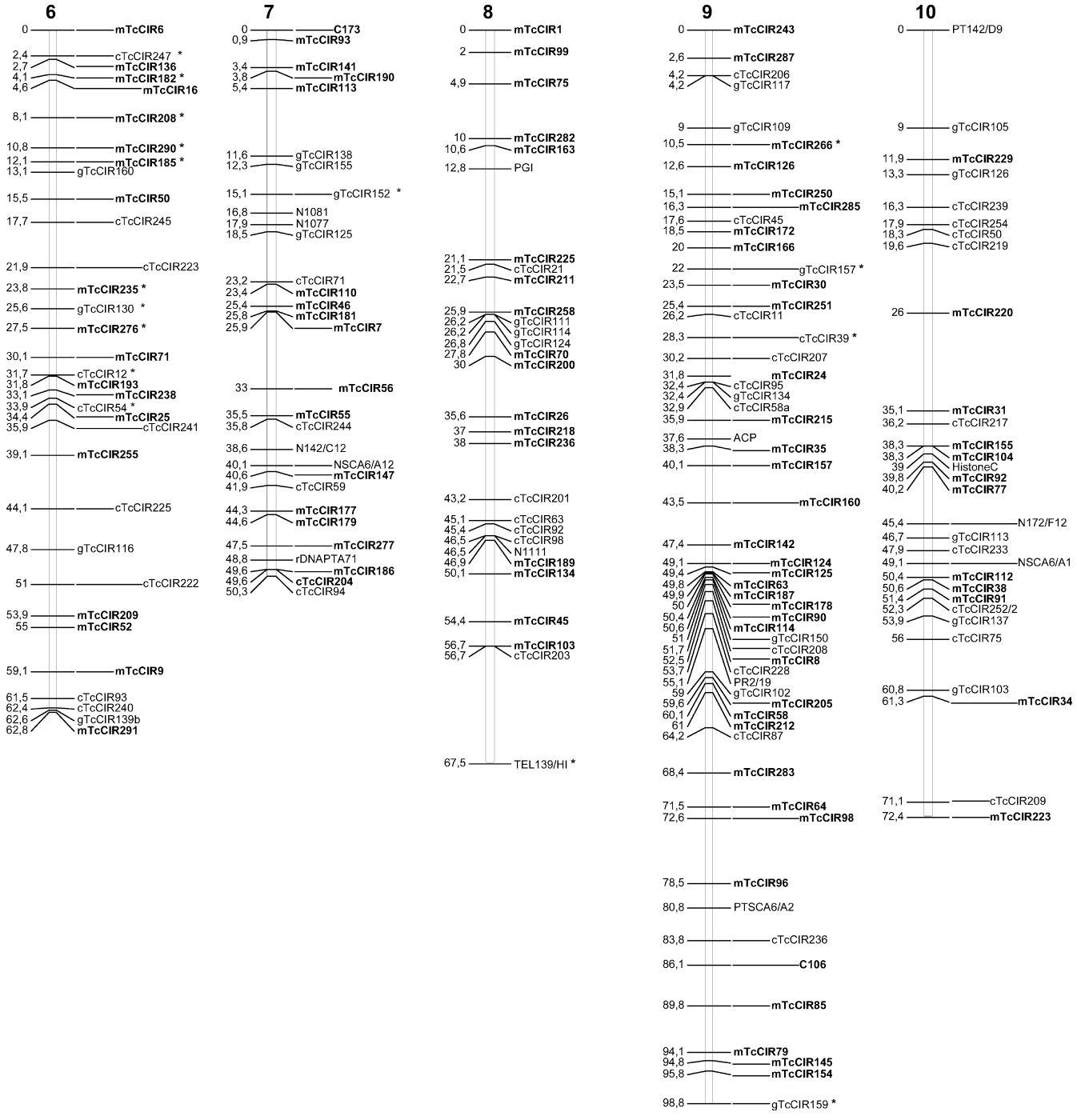

Fig. 1 (continued)

fixation of some segments of chromosome (Temnykh et al. 2000). Cregan et al. (1999) have proposed that targeted isolation of SSR loci using BAC (bacterial artificial chromosome) clones could populate these regions with a few SSRs. To this end, Bhattramakki et al. (2000) showed that BAC libraries were either equal or superior to enriched gDNA libraries as sources of microsatellites for sorghum.

Despite the high number of heterozygous markers segregating in both parents, the order of markers and their distances along LG10 of the UPA 402 map relative to those of the UF 676 map could be ambiguous. Indeed, only two closely linked markers segregated in both parents. LG8 was an exception, since all of the markers mapped there originated from UF 676 alone. We have no evidence that this is related to genes with specific functions located in this linkage group. The explanation

may lie in the pedigree of UPA 402. Indeed, UPA 402 is an Upper-Amazon Forastero clone obtained from a sib-mating involving two Forastero genotypes collected in Peru and may be highly homozygous for this linkage group. A similar situation, where a small number of markers were also assigned to LG8, was observed in other maps established by Clement et al. (2003) and Flament et al. (2001).

Nevertheless, we have produced a saturated microsatellite marker map, and the markers appear to be distributed throughout the genome. If we divide this map into 78 intervals of 10 cM each, there is at least one SSR in each interval, except for four intervals (5%) located on LGs 2, 6, 7 and 10. Macaulay et al. (2001) in barley and Aranzana et al. (2003) in peach have proposed the use of a set of single-locus, codominant and highly polymorphic markers as a framework set or 'genotyping set' that could be easily shared by different research groups. We consider that the marker density and distribution in our map could open the way for a similar approach in cacao.

The development of a saturated linkage map with the markers spaced at small intervals throughout the genome could help improve our knowledge of genome structure. The availability of 268 cacao microsatellites spread throughout the genome and with a coverage of approximately one microsatellite every 3 cM is an attempt in this direction. Studies with a similar approach have shown that the integration of SSRs into the previous linkage maps based mainly on RFLPs was also useful in extending the map length, filling gaps and improving genome coverage (Bhattramakki et al. 2000; Joobeur et al. 2000; Aranzana et al. 2003).

Due to their codominant nature, single-locus behaviour and high polymorphism, microsatellites constitute a set of useful markers that are transferable across different mapping populations, thereby allowing QTL position comparison, and easily transferable to laboratories in tropical regions. This linkage map based on codominant markers, especially SSRs associated with RFLP markers, will be used to localize the most important regions involved in the variation of the traits of interest, such as quality or disease resistance, and to develop marker-assisted selection strategies for this important crop in developing countries. Moreover, their intrinsic properties make microsatellites suitable for the analysis of linkage disequilibrium (LD). Most of the modern Criollo/Trinitario cacao varieties correspond to hybrids between a very small number of parents: a completely homozygous "ancient" Criollo and a Lower-Amazon Forastero individual. Only a few generations separate the first hybridizations from the present Criollo/Trinitario varieties, and allelic associations must have been maintained LD. Microsatellite markers could be used to evaluate the importance of this LD at the genome level in the Criollo/Trinitario varieties and to highlight the associations maintained between molecular markers and useful genes.

Acknowledgements We thank the CNS (Centre National de Séquençage) for sequencing the cacao DNA fragments needed to develop microsatellites. We thank also USDA for their financial participation in these studies and D. Crouzillat (Nestlé) for providing us with 19 microsatellites and seven genomic probes.

References

Aranzana M, Garcia-Mas J, Carbó J, Arús P (2002) Development and variability of microsatellite markers in peach. Plant Breed 121:87–92

Aranzana M, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder C, Testolin R, Abbott A, King G, Iezzoni A, Arús P (2003) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theor Appl Genet 106:819–825

Artiguenave F, Wincker P, Brottier P, Duprat S, Jovelin F, Scarpelli C, Verdier J, Vico V, Weissenbach J, Saurin W (2000) Genomic exploration of the hemiascomycetous yeast: 2. Data generation and processing. FEBS Lett 487:6–13

Barreneche T, Bodenes C, Lexer C, Trontin J-F, Fluch S (1998) A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. Theor Appl Genet 97:1090–1103

Beckman JS, Soller M (1990) Toward a unified approach to the genetic mapping of eukaryotes based on sequence-tagged microsatellite sites. Biotechnology 8:930–932

Bhattramakki D, Dong J, Chabra AK, Hart GE (2000) An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome 43:988–1002

Billote N, Lagoda PJL, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54:277–288

Billote N, Risterucci AM, Barcelos E, Noyer JL, Amblard P, Baurens FC (2001) Development, characterisation and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. Genome 44:413–425

Boivin K, Deu M, Rami JF, Trouche G, Hamon P (1999) Towards a saturated sorghum map using RFLP and AFLP markers. Theor Appl Genet 98:320–328

Castiglioni P, Ajmone-Marsan P, van Wijk R, Motto M (1999) AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor Appl Genet 99:425–431

Clement D, Risterucci AM, Motamayor JC, N'Goran JAK, Lanaud C (2003) Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. Genome 46:103–111

Cregan PB, Mudge J, Fickus EW, Marek LF, Danesh D, Denny R, Mathews BF, Jarvik T, Young ND (1999) Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. Theor Appl Genet 98:919–928

Crouzillat D, Lerceteau E, Petiard V, Morera J, Rodríguez H, Walker D, Phillips W, Ronning C, Schnell R, Osei J, Fritz P (1996) *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theor Appl Genet 93:205–214

Crouzillat D, Phillips W, Fritz J, Petiard V (2000) Quantitative trait analysis in *Theobroma cacao* L. Using molecular markers: inheritance of polygenic resistance to *Phytophtora palmivora* in two related populations. Euphytica 114:25–36

Dettori MT, Quarta R, Verde I (2001) A peach linkage map integrating RFLPs, SSRs, RAPDs, and morphological markers. Genome 44:783–790

Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet 105:127-138

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques 20:758–760

Flament MH, Kebe I, Clement D, Pieretti I, Risterucci AM, N'Goran JAK, Cilas C, Despréaux D, Lanaud C (2001) Genetic

mapping of resistance factors to *Phytophtora palmivora* in cocoa. Genome 44:79–85

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-test-cross: mapping strategy and RAPD markers. Genetics 137:1121–1137

Gupta PK, Balyan HS, Edwars KJ, Isaac P, Korzun V, Röder M, Gautier M-F, Joudrier P, Schlatter AR, Dubcovski J, de la Pena RC, Khairallah M, Penner G, Hayden MJ, Sharp P, Keller B, Wang RCC, Hardouin JP, Jack P, Leroy P (2002) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. Theor Appl Genet 105:413–422

Haussmann BIG, Hess DE, Seetharama N, Welz HG, Geiger HH (2002) Construction of a combined sorghum linkage map from two recombinant inbred populations using AFLP, SSR, RFLP, and RAPD markers, and comparison with other sorghum maps. Theor Appl Genet 105:629–637

Jones ES, Dupal MP, Dumsday JL, Hughes LJ, Forster JW (2002) An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 105:577–584

Joobeur T, Periam N, de Vicente MC, King GJ, Arús P (2000) Development of a second-generation linkage map for almond using RAPD and SSR markers. Genome 43:649–655

Kauffmann S, Legrand M, Geoffroy P, Fritig B (1987) Biological function of "pathogenesis-related" proteins: four PR proteins of tobacco have 1,3-glucanase activity. EMBO J 6:3209–3212

Kijas JMH, Fowler JCS, Garbett CA (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques 16:656–662

Kijas JMH, Thomas MR, Fowler JCS, Roose ML (1997) Integrating of trinucleotides microsatellites into a linkage map of *Citrus*. Theor Appl Genet 94:701–706

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

Lanaud C, Risterucci AM, N'Goran JAK, Clement D, Flament MH, Laurent V, Falque M (1995) A genetic linkage map of *Theobroma cacao* L. Theor Appl Genet 9:987–993

Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A, Lagoda PJL (1999) Isolation and characterization of microsatellites in *Theobroma cacao* L. Mol Ecol 8:2142–2152

Lanaud C, Risterucci AM, Pieretti I, N'Goran JAK, Fargeas D (2004) Characterisation and genetic mapping of resistance and defence gene analogs in cocoa (*Theobroma cacao* L.). Mol Breed (in press)

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Liebhard R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van De Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus* × *domestica* Borkh.). Mol Breed 10:217–241

Macaulay M, Ramsay L, Powell W, Waugh R (2001) A representative, highly informative 'genotyping set' of barley SSRs. Theor Appl Genet 102:801–809

Martin GB, Frary A, Wu R, Brommonschenkel SH, Chunwongse J, Earle ED, Tanksley SD (1994) A member of the tomato *Pto* gene family confers sensitivity to fenthion in rapid cell death. Plant Cell 6:1543–1552

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, De Clerck G, Schneider D, Cartinhour S, Ware D, Stein L (2002) Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). DNA Res 9:199–207

Queiroz VT, Guimarães CT, Anhert D, Schuster I, Daher RT, Pereira MG, Miranda VRM, Loguercio LL, Barros EG, Moreira MA (2003) Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches' broom disease. Plant Breed 122:268–272

Ramsay L, Macaulay M, degli Ivanissevich S, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuvesson S, Morgante M, Massari A, Maestri E, Marmiroli N, Sjakste T, Ganal M, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. Genetics 156:1997–2005

Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high density linkage map of *Theobroma cacao* L. Theor Appl Genet 101:948–955

Risterucci AM, Paulin D, N'Goran JAK, Lanaud C (2003) Identification of QTL related to cocoa resistances to three species of *Phytophtora*. Theor Appl Genet 108:168–174

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organisation of microsatellites sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712

Van Ooijen JW, Voorrips E (2001) JOINMAP version 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands

Viruel MA, Messeguer R, De Vicente MC, Garcia-Mas J, Puigdomènech P (1995) A linkage map with RFLP and isozyme markers for almond. Theor Appl Genet 91:964–971

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

Whitham S, Dinesh-Kumar SP, Choll D, Hehl R, Corr C, Baker B (1994) The product of the tobacco mosaic virus resistance gene N: Similarity to Toll and the interleukin-1 receptor. Cell 78:1101–1115