Mol Breeding (2008) 22:315–318
DOI 10.1007/s11032-008-9174-6

SHORT COMMUNICATION

Characterization of microsatellites from cacao–*Moniliophthora perniciosa* interaction expressed sequence tags

L. S. Lima · K. P. Gramacho · A. S. Gesteira · U. V. Lopes · F. A. Gaiotto · H. A. Zaidan · J. L. Pires · J. C. M. Cascardo · F. Micheli

Received: 24 October 2007 / Accepted: 5 March 2008 / Published online: 20 June 2008

Abstract *Theobroma cacao* L.–*Moniliophthora perniciosa* expressed sequence tags (ESTs) were converted into useful satellite markers for population analysis and genetic mapping. Forty-nine flanking primer pairs from TSH1188 (a resistant genotype) and Catongo (a susceptible genotype) ESTs were designed and screened for polymorphism analysis. Eleven were polymorphic, with an average of 3.81 alleles per locus and a total of 42 alleles. The satellite markers were tested on 21 cacao accessions and two bulked DNAs generated from 6 resistant and 6 susceptible plants from a segregating F_2 (SCA6 × ICS1) population for witches' broom resistance. These results show that EST-derived microsatellites (short sequence repeats, SSRs) in *Theobroma cacao* have many potential applications in linkage mapping and the planning of crosses.

Keywords Cacao · ESTs · Microsatellite · *Moniliophthora perniciosa* · Plant–pathogen interaction

L. S. Lima · K. P. Gramacho · U. V. Lopes · H. A. Zaidan · J. L. Pires
Centro de Pesquisas do Cacau, Cx. Postal 07, 45600-970 Itabuna, BA, Brazil

K. P. Gramacho (✉)
CEPLAC/CEPEC/SEFIT, Cx. Postal 07, 45600-970 Itabuna, BA, Brazil
e-mail: karina@cepec.gov.br

A. S. Gesteira · F. A. Gaiotto · J. C. M. Cascardo · F. Micheli
Universidade Estadual de Santa Cruz, Rodovia Ilhéus-Itabuna, km16, 45662-000 Ilheus, BA, Brasil

F. Micheli
CIRAD, UMR DAP, Avenue Agropolis, 34398 Montpellier, France

The cacao crop (*Theobroma cacao* L.) is grown by about two million producers, in more than 50 countries (Knight 2006). Witches' broom disease, caused by the fungus *Moniliophthora* (=*Crinipellis*) *perniciosa* (Stahel) (Aime and Phillips-Mora 2005), is one of the major cacao diseases in South America and the Caribbean Islands, destroying plantations and leading to important economic and social changes in disease areas such as the State of Bahia in Brazil (Albuquerque 2006). Among the strategies for disease control, the most efficient is the use of resistant cacao genotypes. The actual base of cacao resistance to *M. perniciosa* is predominantly based on Scavina 6 sources. Unfortunately the fungus is adapting to and overcoming the Scavina's resistance (Pires 2003), so the main objective of breeders is to increase resistance durability by obtaining new sources of cacao resistance and to develop a pyramidizing gene strategy. To select genotypes with new resistance genes, molecular data such as expressed sequence tags (ESTs) (Zaidan et al. 2005; Gesteira et al. 2007) and EST-short sequence repeats (SSRs) markers have been developed. Herein, 3,487 cacao–*M. perniciosa* interaction ESTs, including 2,280 from TSH1188 (a resistant genotype) and

Table 1 Primer characteristics for 11 microsatellite loci from cacao–*M. perniciosa* interaction ESTs

GenBank accession number	Name	Primer sequence (5′–3′)	T_a (°C)	Repeat type	Size	Size range (bp)	No. of alleles	H_O	H_E	SSR location
AM851097	msEstTsh-1	F: CACGAAGAAGTGGACGAT	56.8	$(AT)_9$	192	192–200	3	0.0	0.50	nd
		R: CACATGGCTTGACTGGAA								
AM851098	msEstTsh-2	F: ATTCCCTGCCCTCTTACG	62.0	$(CT)_9$	92	90–96	4	0.33	0.65	5′ UTR
		R: CCAGATGTGGATGCGGAT								
AM851096	msEstTsh-3	F: CGGGGAATCTCACACATA	56.8	$(CT)_9$	104	102–112	6	0.31	0.49	ORF
		R: ATCCTGGTTGGTGAGCTA								
AM851099	msEstTsh-4	F: ATATCTCCACCACCACAG	59.5	$(TA)_5$ $(AC)_4$	179	173–179	3	0.0	0.16	ORF
		R: CCGGAGAATGTAGAACCT								
AM851100	msEstTsh-5	F: ACGACTTTAGGAGCTGACC	TD 60–48	$(ACC)_6$	290	290–302	4	0.14	0.26	ORF
		R: AACTTCAACACCAAGACCAT								
AM851101	msEstTsh-6	F: ATGAATATTGTGGAGGAGGTT	TD 60–48	$(AGA)_7$	212	212–229	4	0.58	0.70	nd
		R: TAGCAGTGCTTACAGCTCAA								
AM851102	msEstTsh-7	F: GGAGCTGTTAGGAGAATGC	TD 60–48	$(CTT)_7$ $(CTG)_4$	158	149–161	4	0.25	0.62	nd
		R: AGACCAGGAAAGAAGAGTCC								
AM851103	msEstTsh-8	F: AACCCTTCATGAGACAATGA	TD 60–48	$(TGC)_7$	190	190–193	2	0.05	0.05	ORF
		R: CAGTCCCTTCTCTTCTGTGA								
AM851104	msEstTsh-9	F: CACTTTTGACACTTCAAGCA	TD 60–48	$(TC)_{13}$	206	206–210	3	0.67	0.57	nd
		R: TCAAATCTTGACCCCATAAC								
AM851105	msEstTsh-10	F: ACCCCTCAATCTCACACATA	TD 60–48	$(CT)_{10}$	156	156–174	5	0.47	0.70	5′ UTR
		R: GCTTGGCGCTCTTAGTATC								
AM851106	msEstTsh-11	F: GGAGAAACACCTCTCATGTC	TD 60–48	$(TAC)_{10}$	209	209–218	4	0.41	0.65	3′ UTR
		R: CTTTCTTCAAAGAAGGAAACAT								
Mean	–	–	–	–		–	3.81	0.29	0.48	–

Repeat motifs are listed as 5′–3′ with respect to the forward (F) and reverse (R) primer. T_a is the annealing temperature. Expected heterozygosity was computed according to Nei (1973). The primer sequences and size range in bp for each locus is given. H_O and H_E represent the number of observed heterozygosity and expected heterozygosity (per locus and genotypes), respectively. nd, not determined; ORF, open reading frame; UTR, untranslated region

1,207 from Catongo (a susceptible genotype), were screened for SSR detection, and 11 polymorphic SSRs were obtained for subsequent analysis of population and genetic mapping.

From the SSRs identified, we designed 49 flanking primers pairs for polymorphism analysis using either the PRIMER 3 (www.genome.wi.mit.edu/cgibin/primer/primer3_www.cgi) or the Primer Design Report software. These primers were tested in 23 accessions, being 21 genetically distinct and resistance genotypes and 2 bulked DNAs generated from 6 resistant and 6 susceptible plants from a segregating F_2 (SCA6 × ICS1) population for witches' broom resistance. Leaf samples from each genotype were harvested and used for DNA extraction according to Doyle and Doyle (1990). The polymerase chain reaction (PCR) (20 µl) was as follows: 30 ng of DNA, 0.2 mmol l^{-1} of each primer, 2.0 mmol l^{-1} $MgCl_2$, 0.2 mmol l^{-1} of each dNTP (Ludwig Biotecnologia Ltd.), 1× buffer, and 1 U *Taq* DNA polymerase (Ludwig Biotecnologia Ltd.). PCR products were first checked using 3% agarose gel and stained with ethidium bromide, then the evaluation of the polymorphism was made on 6% denaturing TBE acrylamide gel stained with silver nitrate according to Creste et al. (2001) and Gramacho et al. (2007).

Each locus was tested for Hardy–Weinberg equilibrium (HWE), and genetic diversity parameters were assessed in terms of observed number of alleles (NA), observed heterozygosity (H_O), and expected heterozygosity (H_E) using Genetix (version 4.05.2; Belkhir et al. 1999). A test for linkage disequilibrium was conducted using FSTAT software (Goudet 2001).

From the 49 microsatellite loci tested, 37 (66 %) produced robust alleles, with 13 (35%) being polymorphic (locus msEstTsh-3 was observed as the most polymorphic), and 24 (65%) being monomorphic for the accessions tested. The remaining 19 (34%) failed to amplify fragments under the various conditions tested. The analyses of 21 cacao accessions using 11 microsatellite loci revealed a total of 42 alleles. The number of alleles per locus ranged from 2 to 6, with an average of 3.81 alleles per locus (Table 1). The observed heterozygosity varied among loci from 0 to 0.67, with an average of 0.29. The expected heterozygosity varied among loci from 0.16 to 0.70, with an average of 0.48. All loci showed a significant deviation from HWE, and were independent after Bonferroni correction for multiple tests. Deviations from HWE can point either to mistyping genotypes, sampling bias or natural selection. The observed departure from HWE noted is likely due to the nature of the accessions tested: cultivated material selected and used to develop populations segregating for resistance to various diseases as well as other agronomically important traits (bean number, productivity, etc.). For each of the 11 sequences, the open reading frame (ORF) was determined using the ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and the EST-SSR was localized with respect to the ORF (Table 1). In three cases, the SSRs were located in the untranslated regions (UTR) (two in the 5′ UTR and one in the 3′ UTR), whereas in the other cases either they were located in the ORF (four cases) or the location of the SSR in the cDNA was not clearly determined (four cases). Such repeat patterns (SSR) in ORFs could reflect functional selection of amino acid reiterations in the encoded proteins. Whole-genome analyses have shown that repeat stretches of small/hydrophilic amino acids are frequent in proteins (Katti et al. 2000). Therefore, nucleotide composition might strongly affect the structures and functions of encoded proteins, and could be a determining force in the selection of SSRs in coding sequences.

The results were highly satisfactory as the polymorphic loci used herein are the first SSRs derived from ESTs derived from the cacao–*M. perniciosa* interaction. These EST-SSRs, besides allowing mapping of expressed sequences on the genetic map, also have the potential to identify parents with different resistant genes to witches' broom. Preliminary results obtained with the tested accessions showed segregation between resistant and susceptible accessions for specific genes related to the polymorphic loci, as well as the separation between the Scavina 6 genotype and the other resistant ones, indicating the involvement of different genes of resistance to witches' broom disease. It is important to note that we are leading with expressed genes from the cacao–*M. perniciosa* interaction, and for this reason, the markers identified from these genes on resistant cacao accessions may give specific information about plant resistance.

Acknowledgments The work of LSL was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). This research was supported by the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Ministère des Affaires Etrangères Français (MAE).

References

Aime MC, Phillips-Mora W (2005) The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. Mycologia 97:1012–1022

Albuquerque PSB (2006) Mapas de ligação e Identificação de locos controladores de características quantitativas (QTLs) associados à resistência a *Crinipellis perniciosa* em acessos de cacaueiro (*Theobroma cacao*) originários da Amazônia Brasileira. PhD thesis, Escola Superior de Agricultura Luis de Queiroz, Universidade de São Paulo, Piracicaba, Brasil

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1999) GENETIX 4.04, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France

Creste S, Tulmann Neto A, Figueira F (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Mol Biol Rep 19:299–306

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Gesteira A, Micheli F, Carels N, Silva AC, Gramacho KP, Schuster I, Macedo JN, Pereira GAG, Cascardo JCM (2007) Exploring the cacao–*Moniliophthora perniciosa* interaction using EST analysis. Ann Bot 100:129–140

Gramacho KP, Risterucci AM, Lanaud C, Lima LS, Lopes UV (2007) Development of microsatéllite markers for the genetic analysis of *Crinipellis perniciosa*. Mol Ecol Notes 7:153–155

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices. (version 2.9.3). 2003. Available at http://www2.unil.ch/popgen/softwares/fstat.htm

Katti MV, Sami-Subbu R, Ranjekar PK, Gupta VS (2000) Amino acid repeat patterns in protein sequences: their diversity and structural–functional implications. Protein Sci 9:1203–1209

Knight C (2006) Cocoa review: supply and demand trends. American Cocoa Research Institute. 2000. Available at: http://www.acricocoa.org/acri/index.cfm?item=publications

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

Pires JL (2003) Avaliação quantitativa e molecular de germoplasma para o melhoramento do cacaueiro com ênfase na produtividade, qualidade de frutos e resistência a doenças. Tese (Doutorado em Genética e Melhoramento). Universidade Federal de Viçosa, Viçosa, 328 p

Zaidan HA, Gesteira AS, Micheli FFL, Ceita GO, Cascardo JCM, Carels N, Braz NGR, Serra WO, Gramacho KP, Lopes UV (2005) Caracterização molecular da resistência do cacaueiro (*Theobroma cacao* L.) à vassoura de bruxa. In: XXXVIII Congresso Brasileiro de Fitopatologia, Brasilia, Brasil. Fitopatologia Brasileira. Brasília: Sociedade Brasileira de Fitopatologia, 2005, vol 30. p S88