EST-derived polymorphic microsatellites from cultivated strawberry (*Fragaria* x *ananassa*) are useful for diversity studies and varietal identification among *Fragaria* species

D. J. Gil-Ariza[*], I. Amaya[*], M. A. Botella[†], J. Muñoz Blanco[§], J.L. Caballero[§], J. M. López-Aranda[*], V. Valpuesta[†] and J.F. Sánchez-Sevilla*

[*] I.F.A.P.A.-C.I.C.E (Junta de Andalucía) Churriana, Cortijo de la Cruz, 29140 Málaga, Spain.

[†] Departamento de Biología Molecular y Bioquímica, Universidad de Málaga, 29071 Málaga, Spain.

[§]Departamento de Bioquímica y Biología Molecular. Universidad de Córdoba, 14071 Córdoba, Spain.

Correspondence: J. F. Sánchez-Sevilla, Tel.: +34 951036215; Fax: +34 951036227; E-mail: jfsevilla@uma.es

Abstract

Microsatellite or simple sequence repeat (SSR) markers derived from expressed sequence tags (ESTs) provide genetic markers within potentially functional genes, which could be very useful for breeding programs. To date, the development of microsatellite markers in the genus *Fragaria* has focused mainly on *F. vesca*. However, most of the interests of breeding programs relate to specific characteristics of cultivated strawberry. Here, we describe a set of ten EST-derived microsatellites from *Fragaria x ananassa*. These markers showed high levels of polymorphism within strawberry cultivars and among different *Fragaria* species, indicating their potential for genetic studies not only on strawberry but also in other species within the genus.

Keywords: Fragaria, cultivar identification, expressed sequence tags, simple sequence repeats.

The cultivated strawberry (*Fragaria* x *ananassa*) originated from a natural hybridization between two new world octoploid species, *F. virginiana* and *F. chiloensis*. Strawberry is an important perennial fruit crop throughout the world and annual world production in 2005 was over 3.5 million metric tons (Mt; FAO agricultural data: http://faostat.fao.org/faostat/). Despite its economic value, the polyploid constitution of strawberry has been a major barrier to the genetic characterization of the cultivated species and limited information on the genome structure have been published. The diploid *F. vesca*, with a genome size comparable to that of *Arabidopsis* and sharing a common ancestor with the cultivated strawberry, has become a model system for map development in the genus *Fragaria* (Sargent *et al*. 2004).

Microsatellites, or simple sequence repeats (SSRs) have proved to be locus-specific, co-dominant, highly reproducible and usually highly polymorphic molecular markers (Powell *et al*. 1996). Furthermore, SSRs can often be successfully amplified in related species, making them ideal for comparative mapping between diploid and polyploid species of *Fragaria*. A number of microsatellite markers have been reported to date for *F. vesca* (see Monfort *et al*., 2006 and references therein), *F. viridis* (Sargent *et al.* 2003) and *F. virginiana* (Ashley *et al*. 2003). In contrast, only a limited set of microsatellite loci have been reported for cultivated strawberry so far (Nourse *et al*., 2002; Lewers *et al*., 2005; Folta *et al*., 2005). In contrast to genomic SSRs, EST-SSRs markers have the potential of representing functional markers (Varshney *et al*., 2005). In this paper we report the development of 10 expressed sequence tag (EST)-derived microsatellite loci from *F*. x *ananassa* and analyze their variability and transferability to other *Fragaria* species.

In an ongoing strawberry genomics program we are generating ESTs from cDNA libraries of different developmental stages of strawberry fruit (unpublished results). Approximately 4000 of these ESTs were surveyed for the presence of SSR motifs using the PERL5 scrip MISA (Thiel *et al.* 2003). A set of 27 flanking primer pairs were designed for EST sequences containing dinucleotides motifs using the PRIMER3 program. Primers were designed to define product sizes ranging from 100 to 250 bp in length. SSRs were PCR-amplified using a BioRad icycler in a total volume of 15 μl, containing 1X PCR buffer (GeneCraft; 16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl pH 8.8, 0.1% Tween 20), 2 mM $MgCl_2$, 200 μM each dNTPs, 200 nM each specific primer, 0.5 U of Taq DNA polymerase (GeneCraft), and 25 ng of genomic DNA. The adopted PCR profile was as follows: 94 ºC for 5 min, 35 cycles of denaturation at 94 ºC for 1

min, annealing at 58-60 ºC for 1 min and extension at 72 ºC for 1 min, and a final extension step at 72 ºC for 5 min. Denatured PCR products were separated by electrophoresis in denaturing 6% polyacrylamide gels at a constant power of 100 W for 3 h. A 30-330 bp DNA ladder (Gibco BRL, Invitrogen) was used to define allele sizes. The fragments were visualized by silver staining following the protocol of Bassam *et al*.(1991).

Fifty strawberry cultivars of diverse origin selected from the strawberry germplasm collection at our institute were used to test for polymorphism (Sánchez-Sevilla *et al.* 2004). SSRs were named with a five-letter code (CH for Churriana, Fa for *F*. x *ananassa* and M for microsatellite) followed by a two-digit number. Out of 27 primer pairs tested, only 11 produced bands of the expected size. Ten of these primer pairs produced polymorphic patterns among the selected cultivars (Table 1). Four primer pairs (40%) identified 1-2 bands per cultivar while the other 6 primer pairs identified between 1 and 8 bands. The number of alleles detected in the sampling population ranged from 4 to 15, with an average number of alleles per locus of 8.8. Allelic variability was high for the majority of the loci, with an average PIC of 0,76, comparable to that reported for other *Fragaria* SSR markers. The similarity of these sequences with annotated proteins was analysed after a BLAST search and a putative function could be suggested for eight of them (Table 1). To ensure that they represent new strawberry SSR loci, they were also compared on 17th of March 2006 to the *Rosaceae* Genebank ESTs. ChFaM14 was 99,9% identical to *F. vesca* EST DV440160 and contained the SSR motif and possibility of flanking primer design. ChFaM8 was similar to *F. vesca* EST CX662159 although did not contain the SSR. ChFaM21 was similar to *F. vesca* ESTs CX662055 and CX309671. ChFaM23 was similar to ESTs DV438129 and CO817420 from *F.vesca* and *F. x ananassa*, respectively. ChFaM29 showed small stretches of similarity to *F. x ananassa* ESTs CO378562 and CO379499. Although some of these latest ESTs contained SSRs, they were located to close to the border of the ESTs.

To test the transferability of these *F*. x *ananassa* microsatellites to other *Fragaria* species, eight species with different levels of ploidy were analysed (Table 2). One or two alleles were detected in each of the diploid species, indicating in all cases the amplification of a single locus. Clear amplification products were detected for all analyzed species with high levels of polymorphism between them. These results are in agreement with previous reports indicating that microsatellite markers from one

Fragaria species are generally transferable within the genus *Fragaria* (Ashley *et al*., 2003; Sargent *et al*., 2003; Monfort *et al*., 2005). All the data indicate the potential use of this set of microsatellite markers for varietal identification and diversity studies not only in *Fragaria* x *ananassa* but also within the genus *Fragaria*.

Acknowledgements

We thank DJ Sargent from EMR for critically reading the manuscript. DJ Gil-Ariza has a PhD grant from Junta de Andalucía. I Amaya has a contract of INIA–CCAA. EST collection was supported by grant BIO2001-1958-C04-01.

References

Ashley MV, Wilk JA, Styan SMN, Craft KJ, Jones KL, Fedkheim KA, Lewers KS, Ashman TL (2003) High variability and disomic segregation of microsatellites in octoploid *Fragaria virginiana* Mill. (*Rosaceae*). *Theoretical Applied Genetics,* **107**, 1201-1207.

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem*., **196**, 80–83.

Folta KM, Staton M, Stewart PJ, Jung S, Bies DH, Jesdurai C, Main D (2005) Expressed sequence tags (ESTs) and simple sequence repeat (SSR) markers from octoploid strawberry (*Fragaria x ananassa*). BCM Plant Biology, **5**, 12-23.

Lewers KS, Styan SMN, Hokanson SC (2005) Strawberry GeneBank-derived and genomic simple sequence repeat (SSR) markers and their utility with strawberry, blackberry, and red and black raspberry. *Journal of the American Society for Horticultural Science*, **130**, 102-115.

Monfort A, Vilanova S, Davis M, Arús P (2005) A new set of polymorphic simple sequence repeat (SSR) markers from a wild strawberry (*Fragaria vesca*) are transferable to other diploid *Fragaria* species and to *Fragaria x ananassa*. *Molecular Ecology Notes*, 6, 197-200.

Nourse SM, Fickus EW, Cregan PB, Hokanson SC (2002) Development of simple sequence repeat (SSR) molecular markers in strawberry. In: *Strawberry research to 2001* (eds. Hokanson SC, Jamieson AR), pp. 48-53. ASHS Press Alexandria, Virginia.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*, **2**, 225-238.

Sánchez-Sevilla JF, Soria C, Villalba R, Gil D, Gálvez J, Clavero I, López-Aranda JM, Bartual R, Medina JJ (2004) Strawberry Germplasm Collection at CIFA-Malaga (Spain). *Acta Hort*. (ISHS), **649**, 119-122.

Sargent DJ, Hadonou AM, Simpson DW (2003) Development and characterization of polymorphic microsatellite markers from *Fragaria viridis*, a wild diploid strawberry. *Molecular Ecology Notes,* **3**, 550-552.

Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid Fragaria. *Theoretical Applied Genetics,* **109**, 1385-1391.

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical Applied Genetics*, **106**, 411–422.

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. *TRENDS in Biotechnology*, **23**, 48-55.

Table 1 Description of ChFaM microsatellite markers and variability evaluation in 50 *F*. x *ananassa* cultivars.

Locus	Repeat motif	Position	Primer sequence (5'-3')	T_a (ºC)	Allele size range (bp)	N_i	N_A	N_G	F_G (%)	PIC	GenBank Accesion no.	Putative function*
ChFaM1	$(GA)_{20}$	5´ÚTR	F: GGAGATTATGCACAAAATATAGAGA R: CCAGAACTCCATCAGCCTCT	58	210-260	2-7	15	37	8	0.87	DY633373	Aspartate and ornithine carbamoyltransferases
ChFaM2	$(TC)_{11}$	5´ÚTR	F: CGCAAACCCCAATCTCCT R: ATTGGGAATGTGAAAACG	59	150-190	1-2	6	8	44	0.67	DY633374	WD-repeat regulatory protein
ChFaM4	$(GA)_8$	5´ÚTR	F: CCCAGCATATACTTTGCCGTA R: TCCTTTCTTCATCCCCTCCT	58	130-210	3-6	11	10	24	0.84	DY633375	Anthocyanin regulatory C1 protein
ChFaM5	$(TC)_{19}$	5´ÚTR	F: ATCGCGTTCATTCCTTTTGA R: GACCCATATAGTCTCCAATAAAAGC	59	146-158	1-2	5	30	14	0.55	DY633376	Hypothetical protein
ChFaM7	$(TC)_{13}$	5´ÚTR	F: AATATATTACTCATCCAATCTTGTC R: AGATGGAGGGCTTGGAAGTT	60	150-180	2-5	12	32	10	0.83	DY633377	Beta 1,3-glycosyltransferase-like protein II
ChFaM8	$(GA)_{11}$	5´ÚTR	F: TTCTCCCTCACAAACCCAAT R: TCCTTCTCTGTGCCGAAGAT	60	152-158	1-2	4	7	32	0.66	DY633378	U2AF 35KDa splicing factor
ChFaM14	$(GA)_{18}$	5´ÚTR	F: GGGAGGTTCTGCTTTTAGGG R: GGTCACCTGCCTGTTGATCT	59	141-178	1-3	7	19	18	0.77	DY633379	MADS-box transcription factor
ChFaM21	$(AT)_{13}$	ORF	F: CACCTGTTGCAGTTGTGTGA R: GAATGAAAATACTTGTTTTGAGATCG	59	226-237	1-2	6	11	50	0.66	DY633380	Unknown protein
ChFaM23	$(GA)_{14}$	ORF	F: AGGAGAAGACCGGCTGTGTA R: TGCCTATAGCTGTGGCTGTG	62	141-174	3-8	12	39	8	0.88	DY633381	Ferredoxin-thioredoxin reductase
ChFaM29	$(TG)_8$	ORF	F: ACTTCATCGCCAGAATGGTC R: GCCATTCAATACACAATCCAA	62	147-182	1-4	10	26	12	0.82	DY633382	Putative Calmodulin

T_a: Annealing temperature; N_i: Number of alleles per individual; N_A: Total number of alleles; N_G: Number of unique genotypes; F_G: Frequency of the most prevalent genotype; PIC: Polymorphic information content. *Best BLASTX homology to Uniprot database.

Table 2 Cross-species amplification of ChFaM microsatellite markers in eight *Fragaria* species.

Fragaria species*	Ploidy level	Microsatellite alleles (in bp) observed for each locus†									
		ChFaM1	ChFaM2	ChFaM4	ChFaM5	ChFaM7	ChFaM8	ChFaM14	ChFaM21	ChFaM23	ChFaM29
F. vesca	2x	224	173	139	146	157	158	164	215	150, 153	166
F. viridis	2x	224	173	144, 175	144	157	153, 160	174	215	150, 153	166
F. niponica	2x	232, 241	134, 175	139	145, 162	182, 188	153, 163	160	215	147, 177	168, 171
F. gracilis	2x	239, 246	134	139, 185	148, 168	147, 159	147, 153	157, 174	215, 218	132, 163	168, 175
F. moschata	6x	226, 229	167, 187	139-175 (3)	146, 162	158-172 (5)	149-169 (3)	171-188 (3)	218, 235	126-147 (4)	159-166 (3)
F. chiloensis	8x	217, 246	164	134-175 (4)	145, 159	154-169 (6)	155	151-165 (3)	217	147-184 (4)	166-179 (4)
F. virginiana	8x	226	173	144	146	156	163	163, 165	215, 220	147-153 (3)	166
F. cuneifolia	8x	214-228 (4)	167	137-199 (4)	142	152-166 (5)	155-158	148-169 (4)	215	132-175 (6)	157-169 (4)

*One representative accession was analysed for each species. †In parentheses, number of alleles observed when more than 2.