Theor Appl Genet (2007) 115:907–916
DOI 10.1007/s00122-007-0617-y

ORIGINAL PAPER

Identification, validation and survey of a single nucleotide polymorphism (SNP) associated with pungency in *Capsicum* spp.

Ana Garcés-Claver · Shanna Moore Fellman · Ramiro Gil-Ortega · Molly Jahn · María S. Arnedo-Andrés

Received: 27 October 2006 / Accepted: 11 July 2007 / Published online: 19 September 2007

Abstract A single nucleotide polymorphism (SNP) associated with pungency was detected within an expressed sequence tag (EST) of 307 bp. This fragment was identified after expression analysis of the EST clone SB2-66 in placenta tissue of *Capsicum* fruits. Sequence alignments corresponding to this new fragment allowed us to identify an SNP between pungent and non-pungent accessions. Two methods were chosen for the development of the SNP marker linked to pungency: tetra-primer amplification refractory mutation system-PCR (tetra-primer ARMS-PCR) and cleaved amplified polymorphic sequence. Results showed that both methods were successful in distinguishing genotypes. Nevertheless, tetra-primer ARMS-PCR was chosen for SNP genotyping because it was more rapid, reliable and less cost-effective. The utility of this SNP marker for pungency was demonstrated by the ability to distinguish between 29 pungent and non-pungent cultivars of *Capsicum annuum*. In addition, the SNP was also associated with phenotypic pungent character in the tested genotypes of *C. chinense*, *C. baccatum*, *C. frutescens*, *C. galapagoense*, *C. eximium*, *C. tovarii* and *C. cardenasi*. This SNP marker is a faster, cheaper and more reproducible method for identifying pungent peppers than other techniques such as panel tasting, and allows rapid screening of the trait in early growth stages.

Communicated by I. L. Goldman.

A. Garcés-Claver · R. Gil-Ortega · M. S. Arnedo-Andrés (✉)
Technology for Plant Production Department,
Centro de Investigación y Tecnología Agroalimentaria (CITA),
Apdo 727, 50080 Zaragoza, Spain
e-mail: marnedo@aragon.es

S. M. Fellman
Department of Plant Breeding and Genetics,
Cornell University, 313 Bradfield Hall,
Ithaca, NY 14853, USA

M. Jahn
College of Agricultural and Life Sciences,
University of Wisconsin, 1450 Linden Dr,
Madison, WI 53706, USA

Introduction

Pepper (*Capsicum* spp.) is one of the most economical and agriculturally important vegetable crops due to its high consumption, as fresh or processed product, and world-wide commercial value. One of the important quality attributes of pepper is pungency, which is caused by the presence of a group of alkaloids known as capsaicinoids. These compounds have been widely studied and are currently used in the food industry for medical purposes and in defensive sprays (Henry and Emery 1986; Carmichael 1991; Andrews 1995; Lee et al. 1996; Schulzeck and Wulf 1997; Caterina et al. 2000; Daood et al. 2002).

Capsaicinoids are synthesized exclusively in placental tissue, accumulate in epidermal cells in the placenta (Iwai et al. 1979), and their synthesis begins approximately 20 days post-anthesis (dpa) (Suzuki et al. 1980; Fujiwake et al. 1982). So far, 13 enzymes have been described which putatively catalyze different reactions involved in capsaicinoids production: acyl-CoA synthetase (*ACS*), 4-coumarate CoA ligase (*4CL*), coumaroyl shikimate/quinate 3-hydroxylase (*Ca3H*), phenylalanine ammonia lyase (*PAL*), cinnamic acid 4-hydroxylase (*Ca4H*), caffeic acid *O*-methyl transferase (*COMT*), putative aminotransferase (*pAMT*), branched-chain amino acid transferase (*BCAT*), β-ketoacyl ACP synthase (*KAS*), acyl carrier protein (*ACL*), acyl-ACP thioesterase (*FAT*), acyl-transferase 3 (*AT3*) and capsaicin

synthase (*CSY*). The genes encoding the last 10 enzymes have been cloned (Lee et al. 1998; Curry et al. 1999; Blum et al. 2003; Stewart et al. 2005; Prasad et al. 2006), but only *KAS* and *CSY* have been functionally characterized (Aluru et al. 2003; Prasad et al. 2006).

Molecular research concerning capsaicinoid biosynthesis is being conducted for a complete understanding of how pungency is controlled. Several studies have focused on identifying genes which encode enzymatic steps in the capsaicinoid pathway. Kim et al. (2001) isolated 39 cDNA clones possibly related to pungency by suppression subtractive hybridization (SSH) technique. One of these, SB2-66 (GenBank accession no BF723664), showed a placenta-specific expression pattern, as well as similarities to a group of acyl-transferases. Ultimately the cloning and characterization of *AT3* (Stewart et al. 2005) gave full gene length information about SB2-66, which spans the final part of the second exon and 3′ untranslated region (UTR) of this putative acyl-transferase. *AT3* encodes *Pun1*, an important gene responsible for pungency. In *C. annuum*, the recessive gene, *pun1*, contains a 2.5 kb deletion spanning the promoter and first exon in non-pungent genotypes which result in absence of pungency. Additional putative *Capsicum* acyl-transferases, *Catf1* and *Catf-2*, have also been isolated and associated with pungency (Lang et al. 2006). Most recently, Prasad et al. (2006) have isolated and characterized an important gene in the capsaicinoid biosynthetic pathway, *csy1*, involved in the condensation of vanillylamine with a branched-chain fatty acid. In spite of extensive work concerning this topic, further studies are necessary to fully explain the biosynthesis of capsaicinoids.

To determine the presence or absence of pungency in pepper within breeding programs, different techniques have been developed ranging from panel tasting, such as the Scoville Organoleptic Test (Scoville 1912) to several analytical techniques, including high-performance liquid chromatography (HPLC) analysis (Collins et al. 1995). All of these techniques utilize fruit tissue and are limited by plant growth stage and the subsequent development of a sufficient number of fruits for screening. However, the availability of genetic markers allows determination of pungency at very early stages of plant development, reducing the amount of time and effort involved in distinguishing between pungent and non-pungent genotypes. Regarding the development of molecular markers to detect pungency in pepper, some markers have been linked to *Pun1*, such as CD35 (Tanksley et al. 1988) which is more than 10 cM away from the locus, the cleaved amplified polymorphic sequence (CAPS) marker (Blum et al. 2002), which is closer but was not able to universally distinguish between pungent and non-pungent *C. annuum* varieties, and the CAPS marker (Minamiyama et al. 2005), which is 3.6 or 12.2 cM away from the locus depend on F_2 population used in that study. Recently, five sequence-characterized amplified region (SCAR) markers (Lee et al. 2005) were designed based on the deleted region in *AT3*. So far, the only known molecular markers linked to pungency have been developed based on *Pun 1*.

Single nucleotide polymorphisms (SNPs) provide an abundant source of DNA variation in plants and animals, and possess desirable properties as molecular markers (Wang et al. 1998; Giordano et al. 1999). In the past decade, the interest in the identification and analysis of SNPs in plant species has increased (Zhu et al. 2003; Moreno-Vásquez 2003; Lopez et al. 2005). Abundance, stability and high-throughput genotyping systems employing SNPs offer enhanced fingerprinting and mapping, reliable phylogenetic analyses and novel assaying schemes for breeding programs. For example, the use of PCR-based SNP markers has been reported to be useful to construct maps in plants such as sunflower and rice (Hayashi et al. 2004; Lai et al. 2005). Other studies have identified SNPs linked to genes of interest, such as the SNP linked to a supernodulation trait in soybean, located on the sequence of the *GmNARK* gene (Kim et al. 2005).

Due to the increasing interest in SNPs, a wide range of SNP genotyping methods have been developed and a variety of platforms have been used, including gel electrophoresis, microarrays, mass spectrometry, fluorescence plate readers and flow cytometry (Paris et al. 2003; Lee et al. 2004; Giancola et al. 2006). However, SNP genotyping involving PCR is a simple and economical technique. Methodology using PCR-markers such as tetra-primer amplification refractory mutation system-PCR [tetra-primer ARMS-PCR (Ye et al. 2001)] and CAPS (Konieczny and Ausubel 1993) have already been used successfully for SNP genotyping (Chiapparino et al. 2004; Yeam et al. 2005; Kim et al. 2006). CAPS markers identify SNPs by restriction endonuclease digestion of PCR products. Alternatively, tetra-primer ARMS-PCR is based on the design of allele-specific primers with its 3′ end located at the SNP and with an additional second mismatch at −2 bp from the 3′ end. This allows for better amplification by enhancing the selectivity and eliminating the possibility of false-positive extension of the double-mismatched primers.

In this paper, we report the identification and development of an SNP marker associated with pungency. Furthermore, we describe the application of tetra-primer ARMS-PCR and CAPS analysis for SNP confirmation, and we also provide evidence of the validity of this SNP marker by genotyping of a wide range of *Capsicum* spp. This SNP marker could be used for efficient marker-assisted selection in a broad spectrum of peppers cultivars.

Materials and methods

Plant material and DNA extraction

For the analysis, we used the following pepper genotypes: 19 non-pungent and 7 pungent cultivars of *Capsicum annuum* L. and 47 pungent cultivars belonging to *C. chinense* Jacq., *C. baccatum* var. *pendulum* (Willd.) Eshbaugh, *C. baccatum* var. *praetermisum*, *C. chacoense* Hunz., *C. galapagoense* Hunz., *C. pubescens* Ruiz. & Pav., *C. tovarii* Eshbaugh, Smith & Nickrent, and *C. frutescens* L. The following *Solanaceae* species were used as controls: *Solanum pseudocapsaicum* L., *S. tuberosum* L., *S. melongena* L., *S. lycopersicum* L., *Nicotiana bethamiana* Somin, N. *clevelandii* Gray, *N. tabacum* L., *Petunia hybrida* Vilm. and *Datura stramonium* L. DNA from several genotypes belonging to *C. flexuosum* Sendtn., *C. rhombodium (ciliatum)* (H.B.K.) O. Kuntze, *C. annuum* L., *C. chinense* Jacq., *C. pubescens* Ruiz. & Pav., *C. frutescens* L., *C. eximium* Hunz., *C. cardenasii* Heiser & Smith, *C. minutiflorum* (Rusby) Hunz., *Lycianthes asaifolia* (Kunth *et* Boché) Bitter, *L. surotatensis* Gentry and *L. dejecta* (Fernald) Bitter were obtained from Professor M. Jahn of Cornell University (Ithaca, NY, USA).

To test the suitability of the allele-specific SNP marker to distinguish between homozygous and heterozygous plants, two F_1 plants and their parents, *C. annuum* cv. Serrano Criollo de Morelos-334 (SCM-334; pungent) and *C. annuum* cv. Yolo Wonder (non-pungent), were evaluated.

Plants were grown under greenhouse conditions at Zaragoza (Spain), with a temperature regime that ranged between 15 and 25°C. Total DNA was extracted from leaf tissue of each plant according to Doyle and Doyle (1987), with minor modifications and adapted to small tissue quantities as described in Arnedo-Andrés et al. (2002).

Phenotyping of *Capsicum* spp.

To evaluate pungency, mature red fruits of each genotype were harvested. Each fruit was tasted by at least two persons. Genotypes were considered phenotypically non-pungent when after tasting up to five fruits of the same genotype all were non-pungent. However, when at least one of the tasted fruits was pungent, the genotype was considered phenotypically pungent.

RT-PCR and sequencing

Pepper fruits from *C. annuum* cv. Yolo Wonder, cv. Serrano Criollo de Morelos-334, cv. Jupiter, cv. Maor and *C. chinense* cv. Habanero were harvested at 10, 20, 30 and 60 dpa. Total mRNA was isolated from placenta tissue using TriZol reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). cDNA was synthesized using ProtoScript™ first strand cDNA synthesis kit (New England Biolabs, Beverly, MA, USA) from 1 µg of RNA. Specific PCR primers were designed using the software Primer 3 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi/) to amplify clone SB2-66. The primer sequences were: SB2-66F, 5′-ATTTTGG ATGGG-GAAG-3′ and SB2-66R, 5′-CCGTAAACTTCCGTTG-3′. RT-PCR products were separated by gel-electrophoresis in 2.5% agarose in 1× TAE buffer, stained with ethidium bromide at 50 ng/µl and visualized under UV light with an image analysis system (GelDoc2000, BioRad, Hercules, CA, USA). The amplified fragments of interest were excised from the agarose gel and purified with the Agarose Gel DNA extraction kit (Roche, Basel, Switzerland). The fragments were then cloned into the pGEM-T Easy Vector System (Promega, Madison, WI, USA) and plasmids were purified with QIAprep (Qiagen, Valencia, CA, USA). Sequencing was carried out by the Secugen S.L. (Centro de Investigaciones Biológicas, CSIC, Madrid, Spain) (http://www.secugen.es). In order to confirm the reliability of the cDNA, ubiquitin primers were designed for cultivars of *C. annuum* and *C. chinense*: U1F, 5′-ATGGTCACAAGAA ACACTCC-3′, U1R, 5′-GAGGGATTTGTGTACGGA TA, U2F, 5′-GAAAACATTGAAGCTTGACC-3′ and U2R, 5′-ATGTTGCGGCTATACTCACT-3′. These primers were used to amplify the ubiquitin gene from synthesized cDNA.

SNP discovery

Sequences obtained from the earlier cited cultivars were aligned using BioEdit ver. 5.0.6 (Hall 1999), and putative SNPs were located based on sequence homology.

SNP validation

Tetra-primer ARMS-PCR method. Specific primers (Table 1) were designed from sequence containing the identified SNP using a program, which is accessible through the Internet at http://cedar.genetics.soton.ac.uk/public_html/primer1.html. Each PCR was carried out in a total volume of 10 µl,

Table 1 PCR primers used to amplify the putative SNP using tetra-primer ARMS-PCR method

Primer name	Sequence (5′→3′)	AT (°C)
Forward outer (FO)	GTGATGTTGCACAAGCAACA	59
Forward inner (FI)	TTATTGCCTAATTAATTCCAAGTCTG	59
Reverse outer (RO)	GACCGTAAACTTCCGTTGAAA	59
Reverse inner (RI)	TTCAATCAAACATCCAGTTACTTCA	59

Underlined characters indicate a deliberate mismatch at position −2 from the 3′ termini

AT annealing temperature

containing 30 ng of template DNA, 0.75 µM of inner forward and outer reverse primers, 1.75 µM of inner reverse and outer forward primers, 200 µM dNTPs, 3.5 mM $MgCl_2$, 1× buffer, and 0.5 U of *Taq* DNA polymerase (Invitrogen). PCR amplifications were performed at 95°C for 1 min, 35 cycles of 1 min at 95°C, 1 min at 54°C for annealing, and 1 min of extension at 72°C, ending with 2 min at 72°C. Annealing temperature was 72°C for the first cycle, decreasing by 1°C per cycle until the annealing temperature was reached. The PCR products were separated by gel-electrophoresis in 2.5% MetaPhor® agarose gel (FMC Corporation, Princeton, NJ, USA) and visualized as described above.

CAPS method. Specific PCR primers were used to amplify template DNA that flank the putative SNP. The primer sequences were: SB2-66F, 5′-ATTTTGGATGGGGAAG-3′ and PC-R1, 5′-CCCTCTTCAATCAAACATCCA-3′. PCR was carried out in a total volume of 25 µl, contained 40 ng template DNA, 0.2 µM dNTPs, 3.0 mM $MgCl_2$, 1× buffer, and 0.5 U of *Taq* DNA polymerase (Invitrogen). PCR amplifications were performed in a PCR system 9700 thermocycler (Perkin Elmer, Boston, MA, USA) using the following conditions: 94°C for 2 min, 35 cycles of 45 s at 94°C, 1 min at 51°C for annealing, and 2 min of extension at 72°C, ending with 5 min at 72°C. *Van91* I (Roche) enzyme recognizes the sequence CCANNNN/NTGG, which has the target SNP (/NTGG). A 10 µl aliquot of the PCR product was digested with 2.5 U of *Van91* I and 2.5 µl of 10× SuRE/Cut Buffer B, in a total volume of 25 µl and it was incubated 1 h at 37°C and 15 min at 65°C for heat inactivation. The digestion products were separated in a 2.5% MetaPhor® agarose gel and visualized as described above.

Results

Detection of the SNP

To analyze the SB2-66 pattern of expression, RT-PCR was performed using the specific primers SB2-66F and SB2-66R in the pungent genotypes, 'SCM-334' and 'Habanero', and the non-pungent genotype, 'Yolo Wonder'. After amplification with these primers two fragments of 362 and 307 bp were observed in the pungent genotypes, and only the 307 bp fragment appeared in the non-pungent genotype.

Transcript accumulations of both fragments were examined from 10 to 60 dpa on the three genotypes of interest (Fig. 1). Transcript levels of the 362 bp fragment were observed at 30 dpa in 'Habanero' and at 20 and 30 dpa in 'SCM-334', while no transcript was detected in 'Yolo Wonder'. At the same time, transcript accumulation corresponding to the 307 bp fragment was detected at 30 dpa in 'Habanero', at 20 dpa in 'SCM-334' and at 20 and 30 dpa in 'Yolo Wonder'. No transcript of either fragment was detected at 10, 40 and 60 dpa.

The 362 and 307 bp fragments were cloned and sequenced from the three listed genotypes. Alignment of these cDNA sequences showed 81% homology between the 362 and the 307 bp fragments, revealed three small deletions (Fig. 2) and nine SNPs between the fragments. The alignment between 362 bp fragment and SB2-66 showed that the 362 bp fragment corresponded exactly with the SB2-66 sequence, and therefore indicated that the 362 bp fragment was the final part of the last exon and 3′ UTR region of the putative acyl-transferase *AT3*. Sequence alignment between the 307 bp fragment and the *Catf2* sequence showed 100% homology at the final part of the last exon and 3′ UTR region of *Catf2*. Multiple alignments of the deduced amino acid sequence of the 307 bp fragment with other acyl-transferases showed a highly conserved motif DFGWG, which is characteristic in genes belonging to the BAHD superfamily of acyl-transferases (St-Pierre et al. 1998) (data not shown).

A putative SNP associated with the pungent character was revealed by alignment of the 307 bp fragment sequence between pungent and non-pungent cultivars (Fig. 2). Although the pungent cultivars had the T allele, the non-pungent cultivar showed the G allele. Presence of the SNP in two additional non-pungent genotypes, *C. annuum*

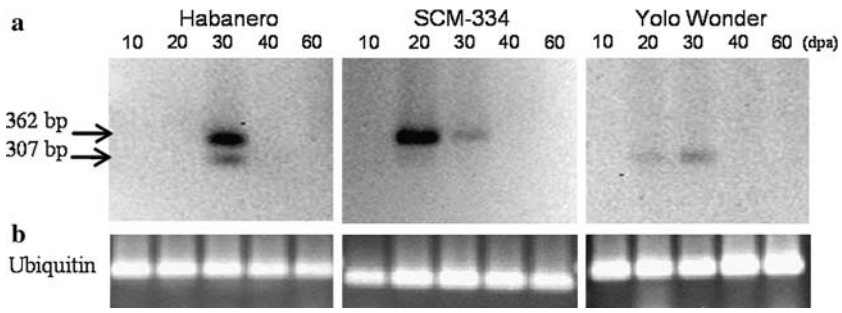

Fig. 1 **a** MetaPhor agarose gel photo of reverse transcription-polymerase chain reaction (RT-PCR) of clone SB2-66 at different developmental stages of placenta tissue from 10 to 60 day post-anthesis (dpa) of Habanero, Serrano Criollo de Morelos-334 (SCM-334) and Yolo Wonder. **b** Agarose gel photo of the ubiquitin products used as internal control in each RT-PCR

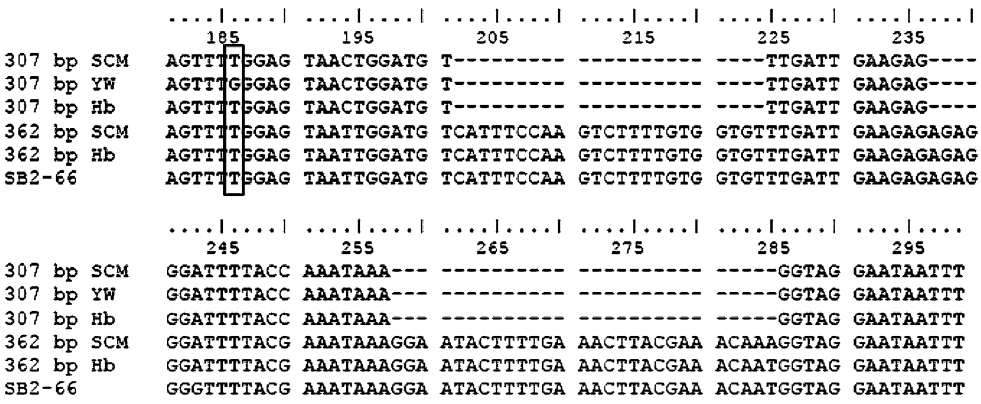

Fig. 2 Partial sequence alignments obtained after sequencing the 307 and 362 bp RT-PCR products. The SNP location is indicated by a *box* and the three deletions by *dashes*. *SCM* Serrano Criollo de Morelos-334; *YW* Yolo Wonder; *Hb* Habanero; *SB2-66* SB2-66 sequence (BF723664)

'Maor' and 'Jupiter', was confirmed by sequencing. The SNP (G/T) was located at position 253 bp of this newly expressed sequence tag of 307 bp.

Validation of the SNP

The detection of the G/T polymorphism was confirmed by tetra-primer ARMS-PCR method using the two pairs of primers described in the Materials and methods. PCR conditions including primer concentrations, annealing temperature and touchdown program, were adjusted in order to reduce non-specific amplifications. Three fragments of 191, 134, and 108 bp were amplified (Fig. 3a). The 134 bp fragment represented the G allele in non-pungent cultivars, and it was amplified with primer combination FI and RO. The 108 bp fragment, the shortest one, was the product obtained with primer combination FO and RI, and it corresponded to the pungent allele (T). The amplification efficiency of this fragment was lower than for the rest of amplification products, possibly due to competitive deprivation of the primers through the largest products. Finally, the 191 bp fragment was amplified with the two outer primers (FO/RO) and served as the template for the amplification of the two allele-specific fragments and as control.

For comparison with tetra-primer ARMS-PCR, CAPS analysis was also performed. PCR amplification of genomic DNA using SB2-66F and PC-R1 primers resulted in a 203 bp fragment. *Van91* I digestion of the PCR amplified product resulted in two restriction fragments of 174 and 29 bp in pungent genotypes, while only one undigested 203 bp fragment was observed in the non-pungent genotypes (data not shown). Of these two methods, tetra-primer ARMS-PCR was easier and more economical than CAPS; therefore, tetra-primer ARMS-PCR analysis was chosen for further experiments.

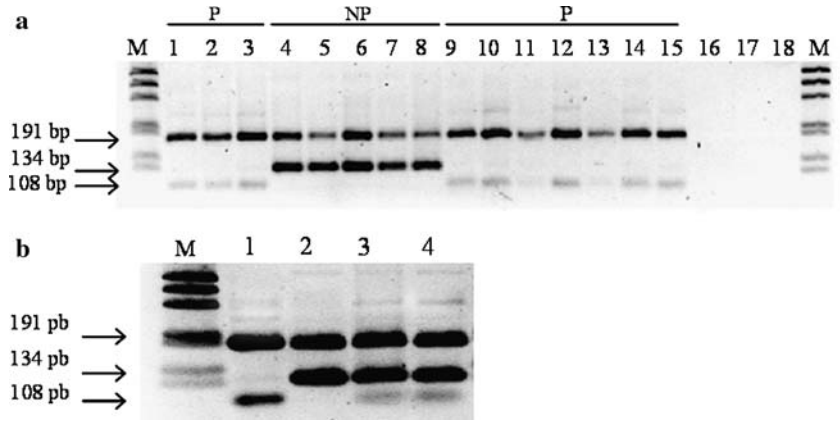

Fig. 3 **a** MetaPhor agarose gel photo of allele-specific SNP marker associated with pungency. SNP marker can distinguish the T allele (pungent), which generates a fragment of 108 bp, from G allele (non-pungent), which generates a fragment of 134 bp. 191 bp fragment is common to both alleles and it is used as internal control. Lanes: *1 C. annuum* cv. SCM-334; *2 C. annuum* cv. Lungo Dolce Sotille; *3 C. annuum* cv. Agridulce; *4 C. annuum* cv. Yolo Wonder; *5 C. annuum* cv. Jupiter; *6 C. annuum* cv. Doux des Landes; *7 C. annuum* cv. Truhart; *8 C. annuum* cv. Cherry sweet; *9 C. chinense*; *10 C. baccatum* var. *pendulum*; *11 C. baccatum* var. *praetermisum*; *12 C. galapagoens*; *13 C. pubescens*; *14 C. tovarii*; *15 C. frutescens*; *16 Datura stramonium*; *17 Solanum melongena*; *18 S. lycopersicum*; *M* 1 kb DNA ladder; *P* pungent; *NP* non-pungent. **b** Amplification of allele-specific SNP marker to distinguish between heterozygous and homozygous plants. Lanes: *M* 1 kb DNA ladder; *1 C. annuum* cv. SCM-334 (pungent); *2 C. annuum* cv. Yolo Wonder (non-pungent); *3* and *4* heterozygous F_1 plants derived from a cross between SCM-334 and Yolo Wonder

Application of tetra-primer ARMS-PCR method to SNP genotyping

The tetra-primer ARMS-PCR method was utilized to assess additional *Capsicum* genotypes. Phenotypic evaluation using a tasting panel was carried out in all assessed pepper genotypes. Phenotypes and the SNP genotyping results are presented in Table 2. Results from 15 *Capsicum* accessions and other 3 *Solanaceae* species genotyped by tetra-primer ARMS-PCR are shown in Fig. 3a.

Using the described tetra-primer ARMS-PCR method, all examined cultivars belonging to *C. annuum* showed the allele-specific products corresponding to their demonstrated phenotypic trait. The same results occurred when cultivars of *C. chinense*, *C. frutescens*, *C. galapagoense*, *C. baccatum* var. *pendulum*, *C. baccatum* var. *praetermisum*, *C. eximium*, *C. cardenasii* and *C. tovarii* were analyzed.

In the case of *C. pubescens*, it was observed that while specific alleles and common fragment were amplified precisely in some genotypes, in others there was no amplification. No fragments were amplified with the developed ARMS-PCR primers in *C. chacoense*, *C. ciliatum*, *C. flexuosum* and *C. minutifolium*. In order to verify whether the method worked correctly in these genotypes, amplification of genomic DNA with the set of primers SB2-66F and SB2-66R was carried out to confirm the presence of the 307 bp fragment containing the SNP (G/T). When the data were analyzed, the expected fragment of 307 bp did not appear, and a different pattern of PCR products was observed (data not shown). In the case of genotypes belonging to *Lycianthes*, *Solanum*, *Nicotiana*, *Petunia* and *Datura*, no 307 bp fragment was amplified, and therefore, no tetra-primer ARMS-PCR products were visualized. Considering the high level of diversity often found in 3′ end UTR region, it is not surprising that the specific 307 bp fragment identified is not present in all *Capsicum* and other related members of the Solanaceae genus.

The tetra-primer ARMS-PCR method was also used to demonstrate that this SNP marker can distinguish between homozygous and heterozygous plants. The results using the parental lines, SCM-334 and Yolo Wonder, and their corresponding pungent F_1 plants, showed that in the F_1 heterozygous plants the two allele-specific fragments, the pungent allele (T) of 108 pb and the non-pungent allele (G) of 134 pb, were amplified (Fig. 3b), whereas, only the allele-specific fragment, corresponding to the non-pungent allele (G) and the pungent allele (T), was amplified for non-pungent and pungent parental lines, respectively.

Discussion

We describe the identification, validation and survey of an SNP marker associated with pungency. During the analysis

Table 2 *Capsicum* genotypes, *Solanaceae* species, their phenotypes and their corresponding nucleotide at the SNP position, after genotyping by tetra-primer ARMS-PCR method

Genotype	Phenotype	Allele
Capsicum annuum cv. Jupiter	NP	G
C. annuum cv. Yolo Wonder	NP	G
C. annuum cv. Calatauco	NP	G
C. annuum cv. Cherry Sweet	NP	G
C. annuum cv. Antibois	NP	G
C. annuum cv. Canada Cheese	NP	G
C. annuum cv. Podorok Moldovii	NP	G
C. annuum cv. Ikeda-1	NP	G
C. annuum cv. Traba	NP	G
C. annuum cv. Cristal	NP	G
C. annuum cv. Doux D'Alger	NP	G
C. annuum cv. UF15	NP	G
C. annuum cv. Morrón de fresno	NP	G
C. annuum cv. Yolo Y	NP	G
C. annuum cv. Florida VR2	NP	G
C. annuum cv. Doux des Landes	NP	G
C. annuum cv. Maor	NP	G
C. annuum cv. Truhart	NP	G
C. annuum cv. Sweet 3575	NP	G
C. annuum cv. Hot1493	P	T
C. annuum cv. Bukeh	P	T
C. annuum cv. RNaky	P	T
C. annuum cv. Lungo dolce sottile	P	T
C. annuum cv. Agridulce	P	T
C. annuum cv. Cayenne	P	T
C. annuum cv. SCM-334	P	T
C. annuum cv. Papri Queen	P	T
C. annuum cv. Thai Hot	P	T
C. annuum cv. Chung-Yang	P	T
C. chinense cv. Orange Habanero	P	T
C. chinense '30036'	P	T
C. chinense '30080'	P	T
C. chinense cv. Peach Savina	P	T
C. chinense PI159234	P	T
C. baccatum var. *pendulum* C-130	P	T
C. baccatum var. *pendulum* C-134	P	T
C. baccatum var. *pendulum* C-135	P	T
C. baccatum var. *pendulum* C-137	P	T
C. baccatum var. *pendulum* C-138	P	T
C. baccatum var. *pendulum* C-131	P	T
C. baccatum var. *pendulum* C-209	P	T
C. baccatum var *pendulum* C-234	P	T
C. baccatum var. *pendulum* C-236	P	T
C. baccatum var. *pendulum* C-237	P	T
C. baccatum var. *pendulum* C-238	P	T
C. baccatum var. *pendulum* C-323	P	T
C. baccatum var. *pendulum* C-232	P	T
C. baccatum var. *pendulum* C-233	P	T

Table 2 continued

Genotype	Phenotype	Allele
C. baccatum var. *pendulum* C-235	P	T
C. baccatum var. *pendulum* C-057	P	T
C. baccatum var. *pendulum* C-070	P	T
C. baccatum var. *pendulum* C-117	P	T
C. baccatum var. *praetermissum* C-172	P	T
C. baccatum var. *praetermissum* C-180	P	T
C. baccatum var. *praetermissum* C-181	P	T
C. baccatum var. *praetermissum* C-182	P	T
C. galapagoense C-167	P	T
C. galapagoense C-179	P	T
C. pubescens C-139	P	ND
C. pubescens C-140	P	ND
C. pubescens C-060	P	T
C. pubescens C-228	P	T
C. pubescens C-342	P	ND
C. pubescens 80062B	P	T
C. pubescens 158277	P	ND
C. eximium	P	T
C. tovarii C-261	P	T
C. frutescens C-126	P	T
C. frutescens C-161	P	T
C. frutescens C-103	P	T
C. frutescens C-162	P	T
C. frutescens C-163	P	T
C. frutescens C-164	P	T
C. frutescens C-166	P	T
C. frutescens C-189	P	T
C. frutescens 1462	P	T
C. chacoense C-153	P	ND
C. chacoense C-154	P	ND
C. chacoense C-175	P	ND
C. chacoense C-176	P	ND
C. cardenasii	P	T
C. minutiflorum	P	ND
C. flexuosum	NP	ND
C. rhombodium (C. ciliatum)	NP	ND
Lycianthes asarifolia	NP	ND
L. surotatensis	NP	ND
L. dejecta	NP	ND
S. pseudocapsicum	NP	ND
S. tuberosum cv. King Edward	NP	ND
S. melongena	NP	ND
S. lycopersicum	NP	ND
Nicotiana benthamiana	NP	ND
N. clevelandii	NP	ND
N. tabacum cv. Samsun	NP	ND
Petunia hybrida	NP	ND
Datura stramonium	NP	ND

P pungent, *NP* non-pungent, *ND* non detected

of the EST clone SB2-66, we identified an additional sequence with 81% homology. Alignment of the sequences corresponding to this new fragment allowed us to identify an SNP (G/T) between pungent and non-pungent genotypes.

Recently, *AT3* was characterized and identified as *Pun1*. The recessive allele, *pun1*, is due to a large deletion that spans the promoter and first exon (Stewart et al. 2005). The full nucleotide sequence of *Pun1* identified SB2-66 in the 3′ end of *AT3*. A later study identified two acyl-transferases, *Catf-1* and *Catf-2* (Lang et al. 2006). *Catf-1* was described and putatively associated with capsaicin synthase. In contrast with this hypothesis, Prasad et al. (2006) have identified *csy1* and characterized it as capsaicin synthase. The full-length cDNA sequence has indicated that *csy1* sequence does not share significant homology with any of the acyl-transferase genes, including *Pun1*. In addition, to the best of our knowledge, the *Catf-1* gene corresponds to the previously described gene *Pun1*, since *Catf-1* and *Pun1* match exactly on the sequence level. Based on our results and available published data, the 307 bp fragment, where the SNP is located, shows a 100% homology with the 3′ end of the *Catf-2*. This sequence has been presented as an acyl-transferase related to pungency in *Capsicum* (Lang et al. 2006); however its role in the capsaicinoid biosynthetic pathway has not yet been demonstrated. In addition, our preliminary mapping results using a tomato population, and data obtained from published comparative maps between tomato and pepper (Livingstone et al. 1999) indicate that *AT3* and *Catf2* are not located in close proximity (data not shown).

Currently, we are focused on studying the full-length DNA sequence associated with the 307 bp fragment. In addition, alignment of the 307 bp fragment with *AT3* and other members of the BAHD superfamily of acyl-transferases has allowed the identification of the conserved domain DFGWG at the beginning of the 307 bp sequence (St-Pierre et al. 1998).

In order to confirm the putative SNP (G/T), we have used two different methods: tetra-primer ARMS-PCR and CAPS markers. Previous studies have already utilized both types of molecular markers successfully (Chiapparino et al. 2004; Acquadro et al. 2003). Our results have showed that both methods were able to discriminate alleles and were easier and cheaper than direct sequencing. However, tetra-primer ARMS-PCR was chosen due to low cost and fast results for SNP genotyping. This method, using four primers in a simultaneous reaction, is a one-step PCR reaction compared to the digestion needed for the CAPS method. Furthermore, the additional mismatch at −2 bp from the 3′-*termini* allows for better amplification by enhancing the selectivity and eliminating the false-positive extension of

the double-mismatched primers. Another advantage of this method is demonstrated by the fact that appropriate commercial restriction enzymes are not always available. In our case, *Van91* I (CCANNNN/NTGG) was found but its recognition site has an ambiguous base (N) within the SNP site. Furthermore, digestion generated small products of 174 and 29 bp, making the resolution by standard agarose gels difficult. For these reasons, ARMS-PCR was chosen for genotyping.

A wide range of species were assayed for the genotyping study, specifically a large group of pungent and non-pungent cultivars belonging to *C. annuum*. For the rest of *Capsicum* species only pungent cultivars were readily available, with the exception of *C. flexuosum* and *C. ciliatum* where non-pungent genotypes were tested. To obtain non-pungent cultivars of *Capsicum* species, apart from *C. annuum*, it is often difficult, in that non-pungent accessions belonging to wild-species have not been widely identified (Andrews 1995). However, the obtained results are quite promising in *Capsicum* germplasm used widely in research, breeding and by consumers. Use of other *Solanaceae*, including *Solanum*, *Nicotiana*, *Petunia*, and *Datura* spp., allowed us to avoid false-scoring of the presence of PCR products resulting from non-specific primer hybridization. Some *Lycianthes* genotypes were also selected because *Capsicum* spp. is considered to be derived from *Lycianthes* (Olmstead et al. 1999).

The SNP marker distinguished the pungency state of all of the 29 *C. annuum* genotypes tested: 19 non-pungent genotypes were identified with the G allele, and 10 pungent ones with the T allele. Based on the classification proposed by Walsh and Hoot (2001), all varieties in the *C. annuum* complex, consisting of *C. annuum*, *C. chinense*, *C. frutescens* and *C. galapagoense*, have been correctly genotyped with the SNP marker, showing an accurate correlation between allelic status and phenotypic character. Also, the T allele was properly identified in the pungent genotypes of *C. tovarii* and the *eximium* complex, consisting of *C. eximium* and *C. cardenasii*.

The *baccatum* complex, consisting of *C. baccatum* and *C. chacoense*, however has shown a different behaviour. While *C. baccatum*, including its two varieties *praetermissum* and *pendulum*, was associated with the T allele, no SNP was detected in any of the four examined genotypes of *C. chacoense*. Another species whose results are difficult to explain is *C. pubescens*. This species has been described as one the most problematic according to its position within *Capsicum* (Walsh and Hoot 2001). In our study, we have detected that the genotypes C-060, C-228 and 80062b were correctly identified with the T allele. However, none of the specific allele fragments were amplified in C-139, C-140, C-342 and 158277 of *C. pubescens*. Neither of the examined genotypes of *C. ciliatum, C. minutifolium* and *C. flexuosum* amplified the specific allele fragments. The higher frequency of polymorphism observed in 3′ UTRs (Bhattramakki et al. 2002) coupled with the increased diversity commonly seen between domesticated and wild species could explain the absence of specific allele fragments in those genotypes.

This study provides a new allele-specific SNP marker for the identification of the pungency trait in *Capsicum*. The results, after assessing a large number of *Capsicum* genotypes, have clearly demonstrated the feasibility of this allele-specific SNP marker, and its application for germplasm screening and marker-assisted selection. Using an allele-specific SNP marker to identify pungent or non-pungent phenotype offers many attractive features over a panel tasting or analytical techniques. First, analysis can be performed in the early growth stages of plants, requiring only small DNA quantities. Second, a large number of samples can be processed systematically, and finally, SNP genotyping using tetra primer ARMS-PCR methodology has a low cost. Furthermore, the identification of this 307 bp fragment opens new sources to continue the investigation of others potential genes involved in the capsaicinoid biosynthetic pathway.

Acknowledgments This research was financially supported by INIA-Spanish Ministry of Education and Science, grants SC00-020-C2-1 and RTA2005-00226-00-00 to R.G.-O. and M.S.A.-A., respectively, and grants A16 for consolidated DGA group. M.S.A.-A. was supported by an "INIA-CCAA" research contract from the Spanish MEC and A.G.-C. by "CITA" fellowship from CITA-DGA (Spain). We are grateful to M.R. Navarro and E. Fuertes (CITA, Spain) for technical support. We thank M. Carravedo (CITA, Spain) for providing most of the seeds of *C. annuum*. We also thank Professor Byung-Dong Kim (Seoul National University, Korea) and Professor Paul W. Bosland (New Mexico State University, NM, USA) for supplying seeds.

References

Acquadro A, Lee D, Chiapparino E, Comino C, Portis E, Donini P, Lanteri S (2003) Detection and characterisation of SNPs in *Capsicum* spp. Capsicum Eggplant Newslett 22:37–40

Aluru MR, Mazourek M, Landry LG, Curry J, Jahn M, O'Connell MA (2003) Differential expression of fatty acid synthesis genes, *Acl*, *Fat* and *Kas*, in *Capsicum* fruit. J Exp Bot 54:165–1664

Andrews J (1995) Peppers, the domesticated capsicums. University of Texas Press, Austin

Arnedo-Andrés MS, Gil-Ortega R, Luis-Arteaga M, Hormaza I (2002) Development of RAPD and SCAR markers linked to the *Pvr4* locus for resistance to PVY in pepper (*Capsicum annuum* L.). Theor Appl Genet 105:1067–1074

Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register JC, Tingey SV, Rafalski A (2002) Insertion–deletion polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. Plant Mol Biol 48:539–547

Blum E, Liu K, Mazourek M, Yoo EY, Jahn MM, Paran I (2002) Molecular mapping of the *C* locus for presence of pungency in *Capsicum*. Genome 45:702–705

Blum E, Mazourek M, O'Connell MA, Curry J, Thorup T, Liu K, Jahn MM, Paran I (2003) Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum*. Theor Appl Genet 108:79–86

Carmichael JK (1991) Treatment of herpes zoster and post herpetic neuralgia. Am Fam Physician 44:203–210

Caterina MJ, Leffler A, Malmberg AB, Martin WJ, Trafton J, Petersen-Zeitz KR et al (2000) Impaired nociception and pain sensation in mice lacking the capsaicin receptor. Science 288:306–313

Chiapparino E, Lee D, Donini P (2004) Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS-PCR. Genome 47:414–420

Collins MD, Wasmund LM, Bosland PW (1995) Improved method for quantifying capsaicinoids in *Capsicum* using high-performance liquid chromatography. HortScience 30:137–139

Curry J, Aluru M, Mendoza M, Nevarez J, Melendrez M, O'Connell MA (1999) Transcripts for possible capsaicinoid biosynthetic genes are differentially accumulated in pungent and non-pungent *Capsicum spp*. Plant Sci 148:47–57

Daood HG, Illes V, Gnayfeed MH, Meszaros B, Horvath G, Biacs PA (2002) Extraction of pungent spice paprika by supercritical carbon dioxide and sub critical propane. J Supercrit Fluids 23:143–152

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Fujiwake H, Suzuki T, Iwai K (1982) Capsaicinoids formation in the protoplast from the placenta of *Capsicum* fruits. Agric Biol Chem 46:2591–2592

Giancola S, McKhann HI, Bérard A, Camilleri C, Durand S, Libeau P, Roux F, Reboud X, Gut IG, Brunel D (2006) Utilization of the three high-throughput SNP genotyping methods, the GOOD assay, amplifluor and TaqMan, in diploid and polyploidy plants. Theor Appl Genet 112:1115–1124

Giordano M, Oefner PJ, Underhill PA, Cavalli-Sforza L, Tosi R, Richiardi PM (1999) Identification by denaturing high-performance liquid chromatography of numerous polymorphisms in a candidate region for multiple sclerosis susceptibility. Genomics 56:247–253

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hayashi K, Hashimoto N, Daigen M, Ashikawa I (2004) Development of PCR-based SNP markers for rice blast resistance genes at the *Piz* locus. Theor Appl Genet 108:1212–1220

Henry CJ, Emery B (1986) Effect of spiced food on metabolic rate. Hum Nutr Clin Nutr 40:165–268

Iwai K, Suzuki T, Fujiwake H (1979) Formation and accumulation of pungent principle of hot pepper fruits, capsaicin and its analogues, in *Capsicum annuum* var. *annuum* cv. karayatsubusa at different stages of flowering. Agric Biol Chem 43:2493–2498

Kim M, Kim S, Kim S, Kim BD (2001) Isolation of cDNA clones differentially accumulated in the placenta of pungent pepper by suppression subtractive hybridization. Mol Cells 11:213–219

Kim MY, Van K, Lestari P, Moon JK, Lee SH (2005) SNP identification and SNAP marker development for a *GmNARK* gene controlling supernodulation in soybean. Theor Appl Genet 110:1003–1010

Kim DS, Kim DH, Yoo JH, Kim BD (2006) Cleaved amplified polymorphic sequence and amplified fragment length polymorphism markers linked to the fertility restorer gene in chili pepper (*Capsicum annuum* L.). Mol Cells 21:135–40

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Lai Z, Livingstone K, Zou Y, Church SA, Knapp JS, Andrews J, Rieseberg LH (2005) Identification and mapping of SNPs from ESTs in sunflower. Theor Appl Genet 111:1532–1544

Lang Y, Yanagawa S, Sasanuma T, Sasakuma T (2006) A gene encoding a putative acyl-transferase involved in pungency of *Capsicum*. Breed Sci 56:55–62

Lee B, Choi D, Lee KW (1998) Isolation and characterization of *o*-diphenol-*O*– methyltransferase cDNA clone in hot pepper (*Capsicum annuum* L). J Plant Biol 41:9–15

Lee RJ, Yolton RL, Yolton DP, Schnider C, Janin ML (1996) Personal defense sprays: effects and management of exposure. J Am Optom Assoc 67:548–560

Lee SH, Walker DR, Cregan PB, Boerma HR (2004) Comparison of four flow cytometric SNP detection assays and their use in plant improvement. Theor Appl Genet 110:167–174

Lee CJ, Yoo EY, Shin JH, Lee J, Hwang HS, Kim BD (2005) Non-pungent *Capsicum* contains a deletion in the *Capsaicinoid synthetase* gene, which allows early detection of pungency with SCAR markers. Mol Cells 19:262–267

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the *Solanaceae*. Genetics 152:1183–1202

Lopez C, Piégu B, Cooke R, Delseny M, Tohme J, Verdier V (2005) Using cDNA and genomic sequences as tools to develop SNP strategies in cassava (*Manihot esculenta* Crantz). Theor Appl Genet 110:425–431

Minamiyama Y, Kinoshita S, Inaba K, Inoue M (2005) Development of a cleaved amplified sequence (CAPS) marker linked to pungency in pepper. Plant Breed 124:288–291

Moreno-Vázquez S, Ochoa OE, Faber N, Chao S, Jacobs JME, Maisonneuve B, Kesseli RV, Michelmore RW (2003) SNP-based codominant markers for a recessive gene conferring resistance to corky root rot (*Rhizomonas suberifaciens*) in lettuce (*Lactuca sativa*). Genome 46:1059–1069

Olmstead GR, Sweere JA, Spangler RE, Bohs L, Palmer JD (1999) Phylogeny and provisional classification of the *Solanaceae* based on chloroplast DNA. Solanaceae 4:111–137

Paris M, Potter RH, Lance RCM, Li CD, Jones MGK (2003) Typing *Mlo* alleles for powdery mildew resistance in barley by single nucleotide polymorphism analysis using MALDI-TOF mass spectrometry. Aust J Agric Res 54:1343–1349

Prasad BCN, Kumar V, Gururaj HB, Parimalan R, Giridhar P, Ravishankar GA (2006) Characterization of capsaicin synthase and identification of its gene (*csy1*) for pungency factor capsaicin in pepper (*Capsicum* sp.) Proc Natl Acad Sci USA 103:13315–13320

Schulzeck S, Wulf H (1997) Local therapy with capsaicin or ASS in chronic pain. Der Schmerz 11:345–352

Scoville WL (1912) Note *Capsicum*. J Am Pharm Assoc 1:453

Stewart C, Kang BC, Liu K, Mazourek M, Moore SL, Yoo EY, Kim BD, Paran I, Jahn MM (2005) The *Pun1* gene for pungency in pepper encodes a putative acyl-transferase. Plant J 42:675–688

St-Pierre B, Laflamme P, Alarco AM, De Luca V (1998) The terminal *O*-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl-transfer. Plant J 14:703–713

Suzuki T, Fujiwake H, Iwai K (1980) Intracellular localization of capsaicin and its analogues in *Capsicum* fruit. I. Microscopic investigation of the structure of the placenta of *Capsicum annuum* var. *annuum* cv. Karayatsubusa. Plant Cell Physiol 21:839–853

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but gene order in pepper and tomato. Proc Natl Acad Sci 85:6419–6423

Walsh BM, Hoot SB (2001) Phylogenetic relationships of *Capsicum* (*Solanaceae*) using DNA sequences from two non-coding regions: the chloroplast *atpB-rbcL* spacer region and nuclear *waxy* introns. Int J Plant Sci 162:1409–1418

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Large-scale identification, mapping and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–1082

Ye S, Dhillon S, Ke X, Collins AR, Day INM (2001) An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res 29:E88

Yeam I, Kang BC, Lindeman W, Frantz JD, Faber N, Jahn MM (2005) Allele-specific CAPS markers based on point mutations in resistance alleles at the *pvr1* locus encoding eIF4E in *Capsicum*. Theor Appl Genet 112:178–86

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in Soybean. Genetics 163:1123–1134