

Hua Xie · Yi Sui · Feng-Qi Chang · Yong Xu
Rong-Cai Ma

SSR allelic variation in almond (*Prunus dulcis* Mill.)

Received: 27 June 2005 / Accepted: 23 October 2005 / Published online: 24 November 2005
© Springer-Verlag 2005

Abstract Sixteen SSR markers including eight EST-SSR and eight genomic SSRs were used for genetic diversity analysis of 23 Chinese and 15 international almond cultivars. EST- and genomic SSR markers previously reported in species of *Prunus*, mainly peach, proved to be useful for almond genetic analysis. DNA sequences of 117 alleles of six of the 16 SSR loci were analysed to reveal sequence variation among the 38 almond accessions. For the four SSR loci with AG/CT repeats, no insertions or deletions were observed in the flanking regions of the 98 alleles sequenced. Allelic size variation of these loci resulted exclusively from differences in the structures of repeat motifs, which involved interruptions or occurrences of new motif repeats in addition to varying number of AG/CT repeats. Some alleles had a high number of uninterrupted repeat motifs, indicating that SSR mutational patterns differ among alleles at a given SSR locus within the almond species. Allelic homoplasy was observed in the SSR loci because of base substitutions, interruptions or compound repeat motifs. Substitutions in the repeat regions were found at two SSR loci, suggesting that point mutations operate on SSRs and hinder the further SSR expansion by introducing repeat interruptions to stabilize SSR loci. Furthermore, it was shown that some potential point

mutations in the flanking regions are linked with new SSR repeat motif variation in almond and peach.

Introduction

Common almond [*Prunus dulcis* (Mill.) D. A. Webb., Syn. *Prunus amygdalus* Batsch.] is economically one of the most important nut tree crops for its kernels (seeds). It has a small genome ($2n = 16$, 0.54–0.67 pg of DNA/2C) (Dickson et al. 1992) which is about twice that of *Arabidopsis thaliana* (0.30 pg of DNA/2C) (Arumuganathan and Earle 1991). Almond originated from Central Asia where it has probably been cultivated for millennia, and was introduced into Xin-Jiang, Northwest China from Iran around the first century BC (Zhu 1983). Nearly 40 almond cultivars including some important ones, such as ‘Yingzui’, ‘Zhipi’, ‘Shuangguo’, ‘Make’, ‘Wanfeng’, ‘Shuangren’, ‘Hanfeng’ and ‘Kexi’ have been described in the region to date. The study of genetic diversity of existing Chinese almond cultivars is very important to enhance their improvement in Chinese almond breeding programs.

Genomic SSR markers have been developed in almost every cultivated fruit species of *Prunus*, including apricot (Lopes et al. 2002; Messina et al. 2004), peach (Cipriani et al. 1999; Testolin et al. 2000, 2004; Sosinski et al. 2000; Aranzana et al. 2002, 2003; Bliss et al. 2002; Dirlwanger et al. 2002, 2004; Georgi et al. 2002; Wang et al. 2002), sour cherry (Downey and Iezzoni 2000), Japanese plum (Mnejja et al. 2004), and sweet cherry (Clarke and Tobutt 2003). Among these, peach genomic SSR markers have been successfully used for molecular identification and genetic similarity analysis of genotypes within other species of *Prunus* like apricot (Hormaza 2002; Zhebentyayeva et al. 2003) and sweet cherry (Dirlwanger et al. 2002; Wunsch and Hormaza 2002; Schueler et al. 2003). Peach genomic SSRs (derived from genomic libraries) (Martinez-Gomez et al. 2003; Testolin et al. 2004) and SSRs derived from almond and

Electronic Supplementary Material Supplementary material is available for this article at <http://dx.doi.org/10.1007/s00122-005-0138-5> and is accessible for authorized users.

Communicated by H. Nybom

H. Xie · Y. Sui · F.-Q. Chang · Y. Xu · R.-C. Ma (✉)
Beijing Agro-Biotechnology Research Center,
Beijing Academy of Agriculture and Forestry Sciences,
Ban-Jing Rd., Hai-Dian District, 100089 Beijing, China
E-mail: marongcai@baafs.net.cn
Tel.: +86-10-51503831
Fax: +86-10-51503980

Y. Sui · R.-C. Ma
College of Life Sciences,
Capital Normal University, 100037 Beijing, China

peach EST sequences (Xu et al. 2004) have been used in almond genetic studies. However, indels (insertions and deletions) and substitutions at SSR loci produce allelic size homoplasy in *A. thaliana* (Symonds and Lloyd 2003) and maize (Matsuoka et al. 2002). Therefore, exploration for SSR mutational patterns is important in order to interpret and use SSR data correctly.

In a previous report, we developed EST-SSR markers for a phylogenetic analysis of almond trees from China and the Mediterranean region (Xu et al. 2004). However, only amplifications and allelic size differences on PAGE were recorded; the precise mutational patterns causing allelic size variations of SSR loci have never been investigated. In the current work, genetic diversity of Chinese almond cultivars in comparison to international almond cultivars from other parts of the world was revealed using 16 SSR markers. In addition, alleles from six of the SSR loci were sequenced to document the mutational patterns.

Materials and methods

Plant materials and DNA extraction

The 43 accessions of almond and peach are listed in Table 1. Young leaves were collected in the field two weeks after flowering and transported to the lab on ice and stored at -70°C till use. Genomic DNA was extracted from young leaves using the CTAB method (Xu et al. 2004).

SSR marker selection

Eight single-locus genomic SSR markers, all of which were transferable among the *Prunus*, and eight single-locus EST-SSR markers, all of which produced well-scored fragments on silver-stained PAGE in almond were selected and used for PCR amplification. The loci BPPCT004, BPPCT007, BPPCT010, BPPCT014 and BPPCT026 (Dirlewanger et al. 2002) were from a CT-enriched genomic library of peach, Pchgms1 (Sosinski et al. 2000) was from a small-insert size genomic library of peach, and Pchgms26 (Wang et al. 2002) and Pchgms31 (Georgi et al. 2002) from a peach large insert BAC library. The loci ASSR4, ASSR17, ASSR46, ASSR54 and ASSR63 derived from almond EST sequences were previously reported (Xu et al. 2004). Three new EST-SSR markers, ASSR70, ASSR71 and ASSR72, were developed in this study and named following the convention of Xu et al. (2004). The locus ASSR70 with a (CT)₂₄ repeat was located in the 5'-UTR region of an almond *MADS*-box gene sequence (GenBank Acc. No. BU574658), the locus ASSR71 with a (GA)₁₅ repeat and the locus ASSR72 with a (TC)₁₉ repeat were, respectively, located in the 5'-UTR regions of two peach EST sequences (GenBank Acc. No. DQ102369 and BU043090). The forward (F) and reverse (R) PCR

primer sequences flanking the three loci were as following: the locus ASSR70 (F = 5'-ACTGTTACTGAGC CATGAAGAAGA-3', R = 5'-GAGAAACAAAGAGACCCCAAGAAG-3'), ASSR71 (F = 5'-AACTTTGTCTGCCTTCATCTTA-3', R = 5'-AGCAGCCCATTCTTCTCTTG-3') and ASSR72 (F = 5'-AAGTGGGAT TGGTAGGGAGGAAG-3', R = 5'-CTACGGAGCCA GTTGAGAAAAG-3') with the expected sizes of 146 bp, 168 bp and 206 bp, and annealing temperatures of 60°C, 55°C and 60°C, respectively.

PCR amplification and detection of fragments

The PCR amplifications and detection of SSR alleles were performed as described by Xu et al. (2004). To assure precision and reproducibility of fragments, DNA samples were amplified and analysed at least twice from each individual.

Cloning and sequencing of SSR alleles

Alleles of 6 of the 16 SSR loci including 2 genomic SSR loci, BPPCT010 with (AG)₄GG(AG)₁₀ and BPPCT026 with (AG)₈GG(AG)₆ (Dirlewanger et al. 2002), and 4 EST-SSR loci, ASSR4 with (TAAAAA)₃, ASSR17 with (AG)₁₅, ASSR63 with (CAT)₅ and ASSR72 with (TC)₁₅ repeat motifs (Xu et al. 2004) in their originally identified sequences, were cloned and sequenced.

Alleles from the six SSR loci were cut from the dried PAGE gels and used as templates for a new round of PCR amplifications. Each of these alleles was directly cloned into the pGEM[®]-T Easy Vector (Promega, USA) according to the manufacturer's instructions, and transformed into *Escherichia coli* DH5 α cells. DNA sequencing was carried out using CEQ8000[™] Genetic Analysis System and Dye Terminator Cycle Sequencing Chemistry Protocol (Beckman Coulter, USA). To obtain reliable sequences, at least three clones per allele were sequenced. The nucleotide sequences were aligned using Clustal W of DNASTAR software (Madison, WI, USA) with manual adjustments.

Data analysis

Number of alleles, observed heterozygosity (H_o) and power of discrimination (PD) were calculated for each locus. H_o was calculated as the number of genotypes which were heterozygous at a given locus divided by the total number of genotypes surveyed at that locus. PD was calculated as $1 - \sum G_{ij}^2$ (Kloosterman et al. 1993), where G_{ij} is the frequency of the j th genotype for the i th locus summed across all alleles at that locus. Genetic similarity (GS) between any two pairs of the 38 almond cultivars was calculated from the alleles across the 16 SSR loci using the Jaccard similarity coefficients (Sneath and Sokal 1973). A dendrogram was constructed using

Table 1 Plant materials used in this study

	Cultivars	Collection sites	Origins
Chinese almond cultivars			
1	Ayuehunzibadan	Ka-Shi County, Xin-Jiang	China
2	Tianrentaobadan	Ka-Shi County, Xin-Jiang	China
3	Kubadan	Ka-Shi County, Xin-Jiang	China
4	Gongbadan	Ka-Shi County, Xin-Jiang	China
5	Dabadan	Ka-Shi County, Xin-Jiang	China
6	Baibadan	Ka-Shi County, Xin-Jiang	China
7	Taobadan	Ka-Shi County, Xin-Jiang	China
8	Badanwang	Ka-Shi County, Xin-Jiang	China
9	Bianzuihe	Ka-Shi County, Xin-Jiang	China
10	Jianzuihuang	Ka-Shi County, Xin-Jiang	China
11	Make	Ka-Shi County, Xin-Jiang	China
12	Shuangguo	Ka-Shi County, Xin-Jiang	China
13	Duoguo	Ka-Shi County, Xin-Jiang	China
14	Amannisha	Ka-Shi County, Xin-Jiang	China
15	Zhipi	Ka-Shi County, Xin-Jiang	China
16	Yeerqiang	Ka-Shi County, Xin-Jiang	China
17	Shuangren	Ka-Shi County, Xin-Jiang	China
18	Hanfeng	Ka-Shi County, Xin-Jiang	China
19	Huangshuang	Ka-Shi County, Xin-Jiang	China
20	Kexi	Ka-Shi County, Xin-Jiang	China
21	Yingzui	Ka-Shi County, Xin-Jiang	China
22	Aifeng	Ka-Shi County, Xin-Jiang	China
23	Wanfeng	Ka-Shi County, Xin-Jiang	China
International almond cultivars			
24	Mission	Shan-Xi Province, China	USA
25	Nonpareil	Shan-Xi Province, China	USA
26	Jeffries	Shan-Xi Province, China	USA
27	Texas	Shan-Xi Province, China	USA
28	Titan	Shan-Xi Province, China	USA
29	I.X.L.	Shan-Xi Province, China	USA
30	Ferraduel	Shan-Xi Province, China	USA
31	Ferragñes	Shan-Xi Province, China	USA
32	Supernova	Shan-Xi Province, China	USA
33	Ao1	He-Bei Province, China	Unknown ^a
34	Ao2	He-Bei Province, China	Unknown ^a
35	Ao3	He-Bei Province, China	Unknown ^a
36	H ₃ -6	Shan-Xi Province, China	Unknown ^b
37	H ₁ -4	Shan-Xi Province, China	Unknown ^b
38	Fura Saco	Tras-os-Montes, Portugal	Portugal
Almond and peach hybrids			
39	Dulcis Pioneer	Shan-Xi Province, China	USA
40	Liquan	Shan-Xi Province, China	Unknown
Peach cultivars (<i>Prunus persica</i> L.)			
41	var. <i>nectarina</i>	Beijing, China	China
42	var. <i>platycarpa</i>	Beijing, China	China
43	var. <i>vulgaris</i>	Beijing, China	China

^aIntroduced from Australia in 1990s

^bIntroduced from Hungary in 1950s

data obtained from the 16 SSR loci, 6 of which were already used previously (Xu et al. 2004), with the unweighted pair group method with arithmetic averages (UPGMA; Sneath and Sokal 1973) on the basis of the similarity coefficients. All these analyses were performed with NTSYS-pc 2.10 software package (Rohlf 2002).

Results

PCR amplification of SSR loci

Due to the diploid constitution of common almond, a maximum of two reproducible alleles on silver-stained

PAGE were obtained in all the 38 almond individuals analyzed with the 8 EST-SSR loci, and these loci also yielded good amplification products within the expected size range for the three peach individuals analyzed. Similarly, the eight genomic SSR markers examined all typically gave reproducible amplification with one or two fragments in all the almond individuals, the size ranges of which were similar to those reported for these loci in peach by Sosinski et al. (2000), Dirlwanger et al. (2002) and Georgi et al. (2002). Across all the 16 loci, the allele sizes in almond were different from those of peach. However, alleles present in the two hybrids between peach and almond, 'Dulcis Pioneer' and 'Liquan', occurred in at least one of the almond or peach individuals.

Size polymorphism and discrimination

The eight EST-SSR markers produced a total of 62 alleles with an average of 7.8 alleles and $H_o = 0.678$ per locus. The eight genomic SSR markers produced a total of 61 alleles with an average of 7.6 alleles and $H_o = 0.628$ per locus (Table 2). Locus ASSR17 was the most informative (12 alleles) among the EST-SSR markers and BPPCT007, BPPCT010 and BPPCT026 were the most informative (11 alleles each) among the genomic SSR markers.

The number of genotypes per locus was 13.3 on average for the EST-SSR markers and was 12.9 for the genomic SSR markers. The 7 EST-SSR markers except for ASSR4, and 4 genomic SSR markers (BPPCT007, BPPCT010, BPPCT026, and Pchgms1) had a PD value higher than 0.800 among the 38 cultivars. The average PD value of the 8 EST-SSR markers (0.842) was higher than that of the eight genomic SSR markers (0.833). The four SSR markers, ASSR17, ASSR63, BPPCT010 and Pchgms31 could identify all the individuals except for the three pairs of international cultivars, 'Mission' and 'Texas', 'Jeffries' and 'Non Pareil', and 'Ferraduel' and 'Ferragñes'.

An average of 6.3 alleles and $H_o = 0.696$ per locus was found among the Chinese cultivars across the 16 SSR loci (Table 2), whereas the international cultivars had 5.6 alleles and $H_o = 0.583$.

The Jaccard similarity coefficients for the 16 SSR loci were used to analyze genetic similarities (GS) between all pairs of 38 almond cultivars (data not shown). GS values ranged from 0.083 ('Wanfeng' and 'Dabadan') to 0.966 ('Duoguo' and 'Zhipi') in comparisons between every pair of the Chinese cultivars. GS values in comparisons between anyone of the Chinese cultivars and all the international cultivars ranged from 0.069 ('Wanfeng' and all the 15 international cultivars) to 0.251 ('Ayuehunzibadan' and all the 15 international cultivars) with an average value of 0.174.

UPGMA cluster analysis based on GS values for comparisons among all 43 samples including 38 almond individuals, 2 hybrids between peach and almond, and 3 peach individuals was used to construct a dendrogram (Fig. 1) with the cophenetic value of 0.854, indicating a high level of reliability. The two hybrids between almond and peach appeared genetically more similar to the three peach individuals than to the almond individuals and they together formed an out-group. All the Chinese cultivars clustered together.

Allelic sequence variation of SSR loci

Substantial sequence variation was found in the 117 sequences of 6 SSR loci from almond cultivars (see Supplementary Data). For the locus ASSR4, which was

Table 2 No. of alleles, observed heterozygosity (H_o) and power of discrimination (PD) of EST- and genomic SSR markers among almond cultivars

Marker code	GenBank Acc. No.	No. of alleles			H_o			PD		
		Chinese	International	Total	Chinese	International	Total	Chinese	International	Total
EST-SSR markers										
ASSR4 ^{a,f}	BI203129	2	2	2	0.086	0.133	0.105	0.159	0.427	0.539
ASSR17 ^{a,f}	CA853978	10	8	12	0.956	0.400	0.736	0.904	0.764	0.927
ASSR46 ^a	CA854086	6	6	7	0.739	0.866	0.819	0.881	0.773	0.880
ASSR54 ^a	BU645586	7	9	10	0.913	0.933	0.921	0.885	0.880	0.928
ASSR63 ^{a,f}	CA854147	6	6	6	0.826	1.000	0.894	0.851	0.773	0.884
ASSR70	BU574658	5	7	10	0.869	0.666	0.789	0.847	0.862	0.895
ASSR71	DQ102369	6	3	7	0.956	0.533	0.789	0.873	0.560	0.847
ASSR72 ^f	BU043090	7	6	8	0.217	0.600	0.368	0.756	0.711	0.832
Average		6.1	5.9	7.8	0.695	0.641	0.678	0.770	0.719	0.842
Genomic SSR markers										
BPPCT004 ^b		3	4	4	0.478	0.200	0.368	0.639	0.800	0.754
BPPCT007 ^b		11	6	11	0.826	0.733	0.789	0.881	0.738	0.914
BPPCT010 ^{b,f}		10	8	11	0.733	0.733	0.789	0.904	0.809	0.935
BPPCT014 ^b		4	3	5	0.695	0.666	0.685	0.779	0.524	0.763
BPPCT026 ^{b,f}		10	7	11	0.782	0.400	0.631	0.892	0.853	0.918
Pchgms1 ^c		5	8	9	0.739	0.733	0.736	0.775	0.560	0.834
Pchgms31 ^d		5	4	5	0.521	0.400	0.472	0.783	0.684	0.753
Pchgms26 (F1, R1) ^e		3	3	5	0.795	0.333	0.552	0.692	0.516	0.796
Average		6.4	5.4	7.6	0.696	0.525	0.628	0.793	0.686	0.833
Average (all markers)		6.3	5.6	7.7	0.696	0.583	0.653	0.781	0.702	0.837

^aXu et al. 2004

^bDirlewanger et al. 2002

^cSosinski et al. 2000

^dGeorgi et al. 2002

^eWang et al. 2002

^fUsed for amplified sequence analysis

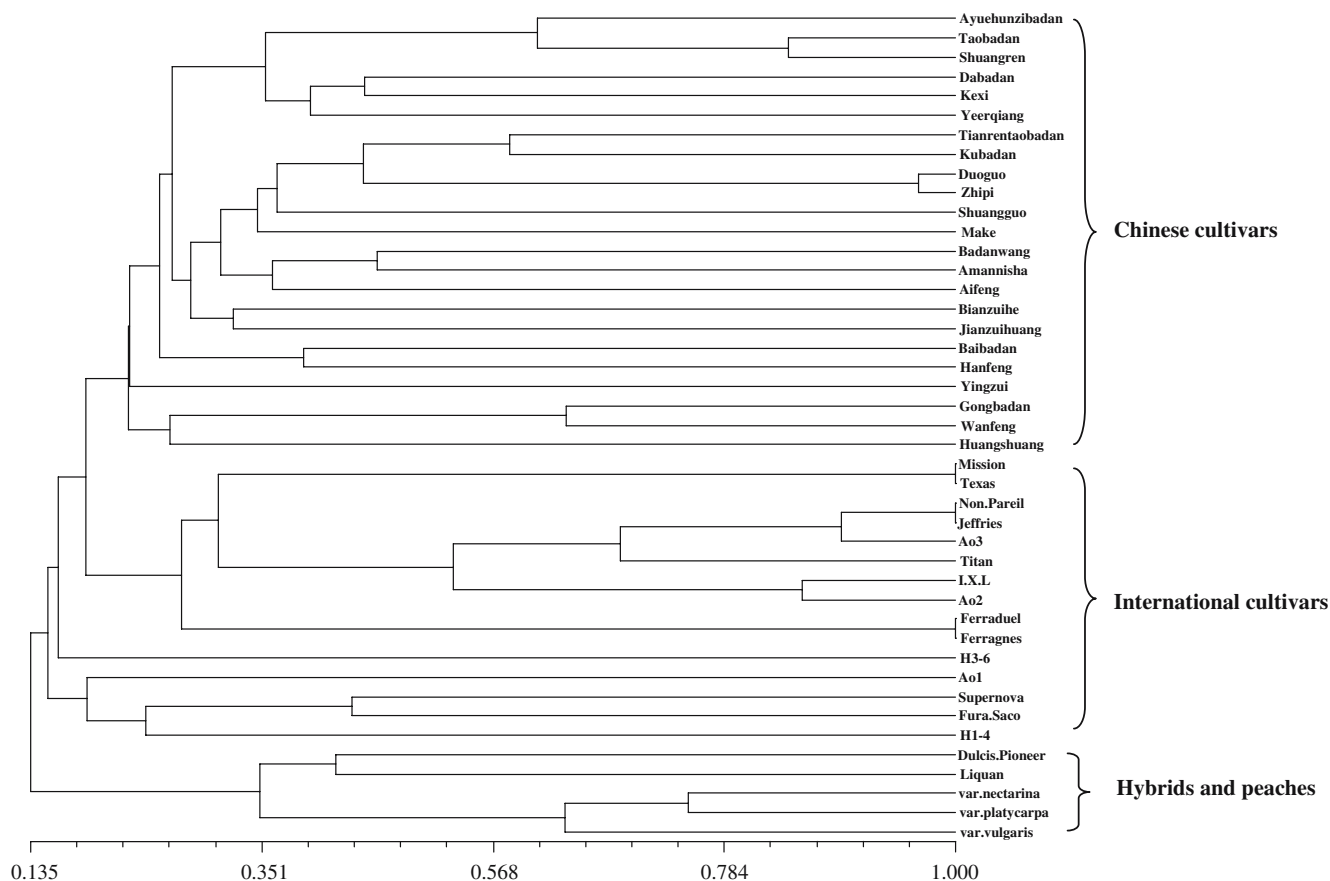


Fig. 1 UPGMA dendrogram of genetic relationship among Chinese and international cultivars of almond, and two hybrids of peach and almond as well as three peach cultivars based on analysis using eight EST- and eight genomic SSR markers. All names of the plant materials are as indicated in Table 1

originally interpreted as an imperfect AAAAAT repeat motif from an almond EST sequence BI203129 (Xu et al. 2004), different allelic sequences were produced as a result of changes in the number of a virtually perfect TAAAAA repeat motif and two single base substitutions, A/G and T/C, in the flanking regions. Similarly SSR polymorphism at the locus ASSR63 was observed due to the expansion of the CAT repeat motif in comparison with the reference sequence (GenBank Acc. No. CA854147), and only a single base insertion occurred in the flanking region of an allele from the cultivar ‘Shuangren’.

At the four SSR loci of AG/CT repeat type, ASSR17, ASSR72, BPPCT010 and BPPCT026, a striking feature of allelic variation was the occurrence of insertions, interruptions, substitutions, or composite repeat motifs in the SSR regions. The locus ASSR17 was originally shown to have a perfect AG repeat motif from the almond sequence (GenBank Acc. No. CA853978), whereas DNA sequencing in this work revealed that the locus has nearly continuous TG and AG repeat motifs in 13 of the 20 alleles sequenced. A G/T substitution interrupting the original perfect AG repeat motifs was observed in 7 of the 20 alleles, and the allelic size vari-

ation appeared only to be the result of variation in the number of TG and AG repeats.

Similarly, at the locus ASSR72 a perfect TC repeat motif was identified from a peach DNA sequence (GenBank Acc. No. BU043090), and the GA doublet created composite repeat motifs $(GATC)_{1-3}(TC)_{13-23}$ in almond cultivars. Of the 19 alleles sequenced 17 have two GATC repeats, and the expansion from two to three GATC repeats in two alleles of ‘Jeffries’ and ‘Tianrentaobadan’, was simultaneously linked with A/C substitutions in the flanking regions. The only source of size variation at the locus was the change in the number of GATC and CT repeat motifs.

Interruptions in the AG repeat regions were observed at the two loci, BPPCT010 and BPPCT026, which were originally identified in peach genomic sequences (Dirlewanger et al. 2002), such as the GG/GGGG interruptions after the fourth or sixth AG from the 5′ end of the repeat region at the locus BPPCT010 and the G/GGG interruptions before the second AG from the 3′ end of the repeat region at the locus BPPCT026. All allelic size variation of the two SSR loci was attributed to the changes in the number of motif repeat and the interruptions.

Size homoplasy, i.e. sequence variation that does not cause allelic size change, was observed at the ASSR17, ASSR72, BPPCT010 and BPPCT026 loci. This homoplasy was, however, present only in a small number of the SSR alleles in some almond cultivars, for examples, the ASSR17 alleles of 143 bp in 'Jianzuihe' and 'Ferraduel' and 149 bp in 'Baibadan' and 'Ayuehunzibadan', the ASSR72 alleles of 159 bp in 'Jeffries' and 'Wanfeng' and 163 bp in 'Supernova' and 'Tianrentaobadan', and the BPPCT026 alleles of 96 bp in 'Hanfeng' and 'Shuangguo'. The BPPCT010 locus displayed more size homoplasy, affecting the alleles of 96, 102, 108, 114 and 120 bp.

Discussion

SSR as efficient markers for genetic diversity analysis of almond

In almond, the utility of SSR markers in the study of genetic diversity has been demonstrated (Martinez-Gomez et al. 2003; Xu et al. 2004). It was further corroborated in this work, three EST-SSR markers, ASSR70, ASSR71 and ASSR72, and seven genomic SSR markers, BPPCT004, BPPCT007, BPPCT010, BPPCT014, BPPCT026, Pchgms26 (F1, R1) and Pchgms31 previously reported in peach, have been newly applied in the analysis of almond genetic diversity.

The peach genomic SSR markers yielded amplification products with the original repeat motif in all the almond individuals analyzed, exhibiting not only a very high level of polymorphism but also efficient discrimination of cultivars (Table 2). The results indicated that the genomic SSR markers previously reported in other species of *Prunus* could be excellent genetic markers for almond.

Up till now, all the genomic SSR markers together with the EST-SSR markers developed at our laboratory were still insufficient to identify the almond individuals 'Mission' and 'Texas', 'Jeffries' and 'Non Pareil', and 'Ferraduel' and 'Ferragñes' using fragment size determination. 'Mission', originally known as 'Texas Prolific' or 'Texas', was believed to be a seedling of the French cultivar Languedoc 302, and therefore, the two accessions are genetically very similar. It was interesting that the four alleles of 'Mission' and 'Texas' had different structures in the SSR region of the locus BPPCT026. 'Jeffries' is a naturally occurring somaclonal mutant taken from a 'Non Pareil' tree, and was identical with 'Non Pareil' at all loci in the present study. 'Ferraduel' and 'Ferragñes' are very closely related, as they were developed from the same breeding program (Socias i Company 1998) and seemed genetically identical in this study. It might be possible to discriminate between 'Jeffries' and 'Non Pareil' and between 'Ferraduel' and 'Ferragñes' through DNA sequence analysis of other SSR loci.

Xu et al. (2004) have identified 178 almond and 497 peach SSRs in 1,482 almond and 4,985 peach contigs and singletons assembled from 3,863 almond and 10,185 peach ESTs, respectively, implying that there could still be at least hundreds of SSR markers from the almond and peach ESTs that could be explored for genetic diversity analysis of almond.

Features of SSR loci mutations in almond

Dinucleotide repeats are the most abundant in vertebrates as well as in plants, such as rice, soybean and wheat (Gao et al. 2004). We have reported that AG/CT was the most frequent in almond and peach (Xu et al. 2004). Here we analyzed sequence variations of 117 alleles from six SSR loci, especially of 98 alleles from four loci of AG/CT repeats in almond.

In maize, a large-scale analysis of SSR loci has shown that the allelic divergence is due to the most frequent indels in their flanking regions (Matsuoka et al. 2002) while no insertions or deletions were discovered in the flanking regions of 84 sequenced alleles in *A. thaliana* (Symonds and Lloyd 2003). Similar invariance in flanking SSR regions has been observed in almond after DNA sequence analysis. Ninety-eight alleles from the four SSR loci with AG/CT repeats, ASSR17, ASSR72, BPPCT010 and BPPCT026 had no indels in the flanking regions. By contrast, the repeat motifs in the four SSR loci showed substantial size variability, involving insertions and occurrences of new motif repeats in addition to varying number of AG/CT repeats, such as insertions of GG/GGGG and G/GGG in most alleles of the loci BPPCT010 and BPPCT026, respectively, and occurrences of new TG and GATC repeat motifs at the loci ASSR17 and ASSR72, respectively.

Symonds and Lloyd (2003) found that interruptions in the repeat regions of most SSR loci were associated with shortening of the original repeat length in *A. thaliana*. In almond, substitutions in the motif repeat regions were observed in some alleles. Thus, long repeats were split into short ones, such as G/T substitution in the three alleles of the locus ASSR17, and G/C in the three alleles of the locus BPPCT026. Indels in the repeat regions and occurrences of new repeat motifs also reduced the lengths of some of the SSR regions. These changes might derive from mutational forces operating on SSR regions and hindering their further expansion.

Interestingly, expansion of the GATC repeat motif from two to three at the locus ASSR72 occurred together with single point mutations (from A to C) in the flanking region. A similar phenomenon was observed at the locus ASSR4; two point mutations (one from C to A and the other from C to T) in the flanking regions of ASSR4 were found to co-occur with the expansion of two TAAAAA repeats to three. Thus, our results suggested that point mutations in the flanking regions may be related to the birth of a new SSR repeat motif, supporting the model which assumes that a length-inde-

pendent mutation process operates on short SSR loci (Dieringer and Schlotterer 2003). As indicated at the two composite SSR loci, ASSR17 and ASSR72, we might also conclude that a SSR motif with more repeats should provide an even more efficient substrate for rapid mutation rate in comparison with SSR motifs containing fewer repeats (Symonds and Lloyd 2003).

In this work, homoplasic SSR alleles arose due to base substitutions, interruptions and/or composite repeat motifs; therefore, the actual degree of polymorphism and the genetic relationship in almond would probably be underestimated if based only on SSR allele size variation. Further work is needed to interpret SSR data correctly and to understand the molecular mechanism underlying the SSR allelic variation in almond and other *Prunus* species.

Acknowledgements This work was financially supported by the National Natural Science Foundation of China (grant No. 30170649) and the National Tenth Five-Year Program of Science and Technology (grant No. 2001BA707B02 and 2004BA713B09-05).

References

- Aranzana MJ, Garcia-Mas J, Carbo J, Arus P (2002) Development and variability analysis of microsatellite markers in peach. *Plant Breed* 121:87–92
- Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott A, King GJ, Iezzoni AF, Arus P (2003) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. *Theor Appl Genet* 106:819–825
- Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 9:208–219
- Bliss FA, Arulsekhar S, Foolad MR, Becerra V, Gillen AM, Warburton ML, Dandekar AM, Kocsisne GM, Mydin KK (2002) An expanded genetic linkage map of *Prunus* based on an interspecific cross between almond and peach. *Genome* 45:520–529
- Cipriani G, Lot G, Huang W-G, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch.]: isolation, characterization and cross-species amplification in *Prunus*. *Theor Appl Genet* 99:65–72
- Clarke JB, Tobutt KR (2003) Development and characterization of polymorphic microsatellites from *Prunus avium* 'Napoleon'. *Mol Ecol Notes* 3:578–580
- Dickson EE, Arumuganathan K, Kresovich S, Doyle JJ (1992) Nuclear DNA content variation within the Rosaceae. *Am J Bot* 79:1081–1086
- Dieringer D, Schlotterer C (2003) Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. *Genome Res* 13:2242–2251
- Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arus 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
- Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arus P (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc Natl Acad Sci* 101:9891–9896
- Downey SL, Iezzoni AF (2000) Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach and sour cherry. *J Am Soc Hortic Sci* 125:76–80
- Gao LF, Jing RL, Huo NX, Li Y, Li XP, Zhou RH, Chang XP, Tang JF, Ma ZY, Jia JZ (2004) One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. *Theor Appl Genet* 108:1392–1400
- Georgi LL, Wang Y, Yvergnaux D, Ormsbee T, Inigo M, Reighard G, Abbott AG (2002) Construction of a BAC library and its application to the identification of simple sequence repeats in peach [*Prunus persica* (L.) Batsch.]. *Theor Appl Genet* 105:1151–1158
- Hormaza JI (2002) Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. *Theor Appl Genet* 104:321–328
- Kloosterman AD, Budowle B, Daselaar P (1993) PCR-amplification and detection of the human D1S80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. *Int J Leg Med* 105:257–264
- Lopes MS, Sefc KM, Laimer M, Machado AD (2002) Identification of microsatellite loci in apricot. *Mol Ecol Notes* 2:24–26
- Martinez-Gomez P, Arulsekhar S, Potter D, Gradziel TM (2003) An extended interspecific gene pool available to peach and almond breeding as characterized using simple sequence repeat (SSR) markers. *Euphytica* 131:313–322
- Matsuoka Y, Mitchell SE, Kresovich S, Goodman M, Doebley J (2002) Microsatellites in Zea—variability, patterns of mutations, and use for evolutionary studies. *Theor Appl Genet* 104:436–450
- Messina R, Lain O, Marrazzo MT, Cipriani G, Testolin R (2004) New set of microsatellite loci isolated in apricot. *Mol Ecol Notes* 4:432–434
- Mnejja M, Garcia-Mas J, Howad W, Badenes ML, Arus P (2004) Simple-sequence repeat (SSR) markers of Japanese plum (*Prunus salicina* Lindl.) are highly polymorphic and transferable to peach and almond. *Mol Ecol Notes* 4:163–166
- Rohlf FJ (2002) NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 2.10. Exeter Software, New York
- Schueler S, Tusch A, Schuster M, Ziegenhagen B (2003) Characterization of microsatellites in wild and sweet cherry (*Prunus avium* L.)—markers for individual identification and reproductive processes. *Genome* 46:95–102
- Sneath PHA, Sokal RR (1973) Numerical taxonomy. Freeman WH and Company (eds) San Francisco, USA
- Socias i Company R (1998) Almond genetics at a turning point: the almond example. *Theor Appl Genet* 96:588–601
- Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch.]. *Theor Appl Genet* 101:421–428
- Symonds VV, Lloyd AM (2003) An analysis of microsatellite loci in *Arabidopsis thaliana*: mutational dynamics and application. *Genetics* 165:1475–1488
- Testolin R, Marrazzo T, Cipriani G, Quarta R, Verde I, Dettori MT, Pancaldi M, Sansavini S (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch.) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* 43:512–520
- Testolin R, Messina R, Lain O, Marrazzo MT, Huang WG, Cipriani G (2004) Microsatellites isolated in almond from an AC-repeat enriched library. *Mol Ecol Notes* 4:459–461
- Wang Y, Georgi LL, Zhebentyayeva TN, Reighard GL, Scorza R, Abbott AG (2002) High-throughput targeted SSR marker development in peach (*Prunus persica*). *Genome* 45:319–328
- Wunsch A, Hormaza JI (2002) Molecular characterisation of sweet cherry (*Prunus avium* L.) genotypes using peach [*Prunus persica* (L.) Batsch.] SSR sequences. *Heredity* 89:56–63
- Xu Y, Ma RC, Xie H, Liu JT, Cao MQ (2004) Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region. *Genome* 47:1091–1104
- Zhebentyayeva TN, Reighard GL, Gorina VM, Abbott AG (2003) Simple sequence repeat (SSR) analysis for assessment of genetic variability in apricot germplasm. *Theor Appl Genet* 106:435–444
- Zhu JL (1983) Almond in Xin Jiang. Xin Jiang Science Press, Xin Jiang (in Chinese)