Euphytica (2015) 206:381–391
DOI 10.1007/s10681-015-1488-y

Candidate gene prediction via quantitative trait locus analysis of fruit shape index traits in apple

Ke Cao · Yuansheng Chang · Rui Sun · Fei Shen · Ting Wu · Yi Wang · Xinzhong Zhang · Zhenhai Han

Received: 16 February 2015 / Accepted: 1 June 2015 / Published online: 16 June 2015

Abstract Fruit shape is a critical appearance quality in apple. Quantitative trait loci (QTLs) for apple fruit shape index (FSI) traits were previously mapped by our laboratory to linkage group 11 of the maternal parent in a cross population of 'Jonathan' × 'Golden Delicious' using simple sequence repeat markers. In this study, QTLs for fruit length, diameter, and FSI were identified again using a high-density single nucleotide polymorphism (SNP) genetic linkage map, and candidate genes associated with FSI were screened via whole-genome re-sequencing data for 'Jonathan' and 'Golden Delicious'. Fifteen QTLs, including four for fruit length, one for fruit diameter, and ten for FSI, were identified in three sampling years. Two overlapping year-stable QTL regions related to FSI were anchored on LG 11 of 'Jonathan'. One candidate gene (MDP0000135244) related to FSI in apple and encoding an LysM domain receptor-like kinase protein was predicted and verified in the segregated population. The nonsynonymous SNP (C11.6053728) of MDP0000135244 was present in 23 of 30 individuals with high FSI, demonstrating a close relationship between MDP0000135244 and FSI trait. These results will be useful for the application of marker-assisted selection for FSI trait in apple.

Keywords Candidate gene · Fruit shape index · *Malus* × *domestica* · QTLs

Electronic supplementary material The online version of this article (doi:10.1007/s10681-015-1488-y) contains supplementary material, which is available to authorized users.

K. Cao · Y. Chang · R. Sun · F. Shen · T. Wu · Y. Wang · X. Zhang · Z. Han (✉)
Institute for Horticultural Plants, College of Agronomy and Biotechnology, China Agricultural University, Beijing 100193, China
e-mail: rschan@cau.edu.cn

Introduction

Fruit shape is an important external quality trait for fresh market apples and, thus, is a priority breeding objective (Hazbavi 2014). However, little is known about quantitative trait loci (QTL) and candidate genes for fruit shape index (FSI), which limits the efficiency of selecting elite hybrid genotypes with positive fruit quality attributes.

With the development of apple linkage maps, many QTLs have been identified and associated with a wide range of quantitative traits, including disease and pest resistance (Gardiner et al. 2007; Durel et al. 2009; Wöhner et al. 2014), growth and tree habit (Kenis et al. 2008; Stoeckli et al. 2008), and especially fruit quality (Dunemann et al. 2009; Sarah et al. 2014; Kunihisa et al. 2014). Apple fruit shape, including height, diameter, and FSI, is of strong interest in QTL mapping analysis, but there are few reports on QTLs for FSI in apple. Using 250 F_1 individuals of a 'Telamon' × 'Braeburn' segregating population, QTLs for fruit height were mapped

on linkage groups (LG) 2, 6, 15, and 17; and QTLs for fruit diameter were mapped on LGs 2, 5, 9, 10, and 17, but the QTLs varied between two sampling years (Kenis et al. 2008). Using another hybrid population (Co-op 17 × Co-op 16), two QTLs for fruit diameter were identified on LGs 5 and 8, and three QTLs for fruit length were detected on LGs 3, 5, and 17 (Sarah et al. 2014). In 'Jonathan' × 'Golden Delicious', eight QTLs for fruit length and 11 for diameter were identified in 2 years, with most of these QTLs located on LGs 5, 8, 11, and 15 (Chang et al. 2014). These results clearly show inconsistency in QTL localisation due to differences in mapping populations and environmental factors.

Fruit shape in tomato is regulated by several major gene loci. One major gene, *ovate,* was detected on chromosome 2 of the tomato genome and was cloned and characterised as encoding a new class of nucleus-localised putative regulatory proteins (Liu et al. 2002). Although allelic variation in both *ovate* and another gene, *sun*, caused elongated fruit shape, the two loci differed in their genetic and morphological effects in tomato. For example, *sun* causes uniform elongation in both longitudinal directions, whereas *ovate* usually causes asymmetric elongation (van der Knaap and Tanksley 2001). Locule number, which has a pleiotropic effect on fruit shape and size, is controlled by the *fasciated* (*fas*) and *locule-number* (*lc*) loci. *Fas* encodes a transcription factor, and down-regulation of *fas* is caused by a large insertion in the first intron, which results in fruits with high locule number (Cong et al. 2002). The molecular nature of *lc* indicates that two single-nucleotide polymorphisms (SNPs) are located approximately 1200 bp downstream of the stop codon of a gene encoding a WUSCHEL homeodomain protein (Gustavo et al. 2011). In general, *sun*, and *ovate* control elongated shape, whereas *fas* and *lc* control fruit locule number and flat shape in tomato. In apple, the expression of two genes associated with cell proliferation, *MdANT1* (MDP0000175309) and *MdANT2* (MDP0000190889), was high from anthesis until 15 days after full bloom, a period coinciding with active cell division and rapid longitudinal fruit growth (Madhumita and Anish 2012). We previously identified five major gene loci involved in the regulation of FSI on four LGs (10, 11, 12, and 13) in a 'Jonathan' × 'Golden Delicious' mapping population (Sun et al. 2012). Recently, using the same population and simple sequence repeat (SSR) markers, the year-stable QTL for FSI, fsij08.11.2/fsij09.11 at 7.371 cM on chromosome 11 of the female parent, 'Jonathan', co-localised with a major gene locus, F11-1 (Chang et al. 2014). The consistency between QTL analysis using SSRs and major gene mapping puts fsij08.11.2/fsij09.11 at the forefront of candidate gene mining. However, because of the large coverage range of these QTLs (5.2 cM for fsij08.11.2; 5.4 cM for fsij09.11, 0.5 Mb/cM density in LG 11), exploration of QTL-based candidate genes cannot take place until fine mapping is performed.

Here, using the mapping population ('Jonathan' × 'Golden Delicious') reported previously by Chang et al. (2014) and a super-high-density genetic linkage map with SNP markers derived by restriction-associated DNA sequencing (RADseq), QTLs associated with FSI, fruit length, and fruit diameter were identified, and candidate genes were predicted.

Materials and methods

Plant materials

We used a hybrid population of 'Jonathan' × 'Golden Delicious' that consisted of 1733 seedlings (Sun et al. 2012; Chang et al. 2014). After preliminary selection for breeding in 2011, scions of all seedlings in the population were grafted onto *Malus micromalus* rootstock with a 30 cm 'SH40' dwarfing interstem and replanted at China Agricultural University (Beijing) for further genetic mapping. In total, 318 seedlings were randomly selected for RAD sequencing.

Phenotyping

The phenotypic characteristics of fruit length, fruit diameter, and FSI (ratio between height and diameter at the widest part of the fruit) were evaluated for 5 years. From 2008 to 2011, ripening fruit were harvested from the seedling populations. In 2014, the fruit were sampled from the grafted trees. Using a vernier caliper, fruit length and diameter were measured and the average values of at least five apples per tree, picked from mid-canopy, were calculated for further analysis. Differences in sample size over the 5 years were due to year-to-year variations in fruit bearing. From 2008 to 2011, only 152, 113, 21, and 66 genotypes had sufficient fruit for sampling in the

respective years, and only 122 genotypes had sufficient fruit in 2014. No fruit were picked in 2012 and 2013 because the grafted plants had not yet matured.

QTL analysis

QTL detection was carried out using the MapQTL 6.0 software package (Van Ooijen 2009). Interval mapping was performed for each trait. Genome-wide limit of detection (LOD) thresholds for QTL significance were calculated by performing 1000 iterations with MapQTL's permutation test. QTLs with LOD values higher than the genome-wide threshold (LOD $>$ 3.50) were considered significant at $P < 0.05$. The linkage maps and QTL positions were drawn using MapChart (Voorrips 2002). However, a 99 % confidence interval (LOD $>$ 4.10) was obtained using a two-LOD support interval (Van Ooijen 2009).

Prediction of candidate genes

Resequencing data for Jonathan and Golden Delicious and their SNPs were developed by our laboratory (unpublished). The apple genome reference scaffolds and predicted transcripts were downloaded from the genome database for Rosaceae (GDR) (http://www.rosaceae.org/). The minimum coverage for paired reads to call SNPs was in the range of 3–100. Only SNPs within QTL intervals with a frequency threshold of >80 % were filtered. According to the classification of SNPs in QTL intervals, annotation of all genes with nonsynonymous SNPs was performed in reference to the apple genome. After that, gene ontology (GO) annotation was performed by using the online WEGO software (Ye et al. 2006).

Gene expression analysis of candidate genes by RNAseq

In our study, the candidate gene expression data were cited, but unpublished RNA-seq results from our laboratory. 'Greensleeves', which has a similar fruit development period to 'Golden Delicious', were sampled at three stages (about 18, 67, and 132 days after anthesis, corresponding to the middle stage of cell division, cell expansion and the mature fruit stage, respectively.). Total RNA was extracted from the fruit samples and sequenced in the Genomics Facility, Institute of Biotechnology, Cornell University using Illumina HiSeq^{TM} 2000 (San Diego, CA, USA).

Verification of SNPs in candidate genes and their segregation in hybrids

Primers were designed using Primer-blast software (NCBI, Maryland, USA) based on the sequences around SNPs in candidate genes. Then, PCR products comprising SNP loci of interest were amplified in both parents and sequenced using Sanger sequencing (BGI, Beijing, China) to verify SNPs obtained from resequencing data between parents. The causal SNPs were genotyped in 30 seedlings with high FSI and in 30 seedlings with low FSI values. The sequencing results were assessed using analysis of variance (ANOVA) to detect positive relationships between SNPs and phenotype.

Results

Changes in population phenotype over time

To estimate year-to-year variation in fruit shape traits, correlation coefficients for fruit length, fruit diameter, and FSI traits between years were calculated using 3 years of sampling data from 43 ever-bearing seedlings (Table 1).

Significant correlations between fruit shape traits and year were found for 2008 and 2009. No correlation existed for fruit length and year between 2008 and 2014, or between 2009 and 2014, which reflected strong environmental variation between 2008 or 2009 and 2014. Fruit diameter in 2008 was correlated with that in 2014, but this relationship was not apparent for 2009 versus 2014. Stronger correlations were also detected for FSI between any two of the three years,

Table 1 The correlation analysis of fruit shape traits in different years

Sampling years	Fruit length	Fruit diameter	FSI
2008 versus 2009	0.677*	0.679*	0.765*
2008 versus 2014	0.314	0.433*	0.723*
2009 versus 2014	0.298	0.316	0.737*

* Significant ($P = 0.01$, $r_{0.01} = 0.389$)

indicating better year-to-year stability of FSI phenotypes. From these results, we conclude that fruit length is more sensitive to environmental conditions than fruit diameter or FSI.

QTL analysis

QTL analysis was performed using phenotypic data on fruit length, diameter, and FSI collected in 2008, 2009, 2010, 2011, and 2014 based on a high-density SNP linkage map. Fifteen QTLs were identified at the genome-wide threshold for QTL significance (LOD > 3.50) and explained genotypic variation ranging from 10.2 to 23.0 % (Table 2). The highest values in the LOD profile in 2010 and 2011, respectively, were 1.76 and 1.40 for fruit length, 2.50 and 2.94 for fruit diameter, and 1.71 and 2.50 for FSI. Therefore, no significant QTLs for the fruit shape traits were identified in 2010 and 2011 because there were fewer available individuals (Table 2).

Of the 15 identified QTLs, four were associated with fruit length, one with diameter, and 10 with FSI. Surprisingly, all of the QTLs were mapped to 'Jonathan' except for one, *flg14.1*, which was associated with fruit length and derived from 'Golden Delicious' in 2014. A year-stable QTL cluster associated with FSI was mapped to LG 11 of the maternal parent, 'Jonathan' (Table 2). Among the 10 QTLs for FSI, two overlapping regions were detected in 3 years. The first (designated 'Locus 1') was overlapped by *fsij08.11.3*, *fsij09.11.1*, and *fsij14.11.2* in the region from 9.60 to 17.90 cM in 'Jonathan', while the second (designated 'Locus 2') spanned the region from 22.21 to 22.46 cM of LG11 and was overlapped by

Table 2 Summary data for QTLs of fruit length, diameter, and FSI using an interval mapping model of the 'Jonathan' × 'Golden Delicious' population

Traits	Year	QTLs name assigned	LG	LOD	Peak position (cM)	Interval (cM, $P < 0.01$)	The nearest marker with the highest LOD	Explained variance (%)
Fruit length	2008	*flj08.7.1*	J7	4.87	6.48	5–6.48	hmC07.9626405	13.9
		flj08.7.2	J7	3.52	13.38		hmC07.1728469	10.2
	2009	*flj09.11*	J11	3.73	47.55		euLG11_008	14.4
	2014	*flg14.1*	G1	3.69	66.32		hmC01.27382288	14.4
Fruit diameter	2008	*fdj08.3*	J3	3.95	24.65		hmC03.5431144	12.0
Fruit shape index	2008	*fsij08.11.1*	J11	4.85	0.00	0–2.00	euC11.51398	14.7
		fsij08.11.2	J11	3.84	5.80		emC11.2710877	11.6
		fsij08.11.3	J11	6.09	11.48	8.60–17.90	euC11.4515793	18.2
		fsij08.11.4	J11	4.58	22.21	22.21–23.60	euC11.7408490	14.1
	2009	*fsij09.11.1*	J11	5.53	11.48	9.60–17.90	euC11.4515793	23.0
		fsij09.11.2	J11	4.92	22.46	22.21–23.60	euC11.7408490	18.3
	2014	*fsij14.5.1*	J5	3.99	75.19		huLG05_082	13.7
		fsij14.11.1	J11	3.74	1.00		euC11.51398	14.4
		fsij14.11.2	J11	5.75	9.6	8.60–18.90	euC11.4515793	22.8
		fsij14.11.3	J11	4.50	22.21	22.21–22.46	euC11.7408490	17.3

QTL names were assigned according the following rules. The first two or three letters are trait abbreviations: fruit length (*fl*), fruit diameter (*fd*), and fruit shape index (*fsi*); the following letter indicates the QTL was derived from the female (*j*) or the male (*g*) parent; the following two numbers indicate the phenotype year and are followed by a period and the number of the linkage group on which the QTL is located; the final number following the second period is the serial number of that QTL. Marker names were assigned according the following rules. The first letter indicates the restriction enzyme used: *Eco*RI (e) and *Hind*III (h); a second letter 'm' indicates that the sequencing read was blasted with a reference sequence and multiple similar locations were obtained (only sites with the highest similarity are shown); a second letter 'u' indicates that the read was uniform after blasting with the reference; the following 'C' indicates that the read was successfully blasted against the apple genome sequence; the chromosome number and detailed physical location are displayed after that; the following 'LG' indicates that the read failed to blast against the genome, in which case the linkage group number (from linkage analysis with Joinmap 4.0) and the marker order are shown

fsij08.11.4, *fsij09.11.2*, and *fsij14.11.3* (Fig. 1). Four QTLs associated with fruit length were distributed on LGs G1, J7, and J11 in 3 years without overlapping regions. Only one QTL associated with fruit diameter was detected, in LG 3 of 'Jonathan' in 2008, and it explained a small amount of variance for this trait (12.0 %).

Delineation of FSI QTLs for sequence variation analyses

Locus 1 and Locus 2 were easily anchored to the genome via the pre-anchored SNP markers at 4411077–6723403 nt and 7408490–7474945 nt, respectively, on chromosome 11 of 'Jonathan'. Then, reads in the database of whole-genome resequencing of the two parents (unpublished data) were aligned to the *Malus* × *domestica* reference genome sequence (Whole Genome ver. 1.0p, Velasco et al. 2010) to call SNPs in Locus 1 and Locus 2 between the parents. In 'Jonathan', the resequencing depth of Locus 1 and Locus 2 was 27.4× and 39.3×, respectively, and the corresponding values in 'Golden Delicious' were 28.5× and 24.2× (Table 3). Because both QTL regions for FSI trait were from the female parent, heterozygous SNPs in 'Jonathan' that could be segregated in hybrids were selected for further analysis. A total of 11,028 and 716 SNPs were detected in Locus 1 and Locus 2, respectively. In Locus 1, 2076 of these SNPs were distributed within exon regions and 913 and were nonsynonymous substitutions. In Locus 2, 91 were in exon regions and 61 were nonsynonymous. These nonsynonymous substitutions putatively affected 187 genes within Locus 1 and 12 genes within Locus 2.

Based on GO annotations in the apple genome, the 199 genes in Locus 1 and Locus 2 could be divided into nine functional groups (Fig. 2). The detailed results of GO annotation can be seen in Additional file 1 after uploading Additional file 2 to WEGO (http://wego.genomics.org.cn/). Genes whose functions included pigmentation, response to stimulus, localisation, and establishment of localisation, were not considered candidate genes for the FSI trait. Seventy-five genes involved in metabolic process, 58 genes involved in cellular processes, one gene for cellular component organization, one gene for cell death, and 10 involved in biological regulation were prioritised. Genes involved in alcohol or nitrogen compounds, lipids, pigments, organic acids, nucleotides, vitamin metabolic

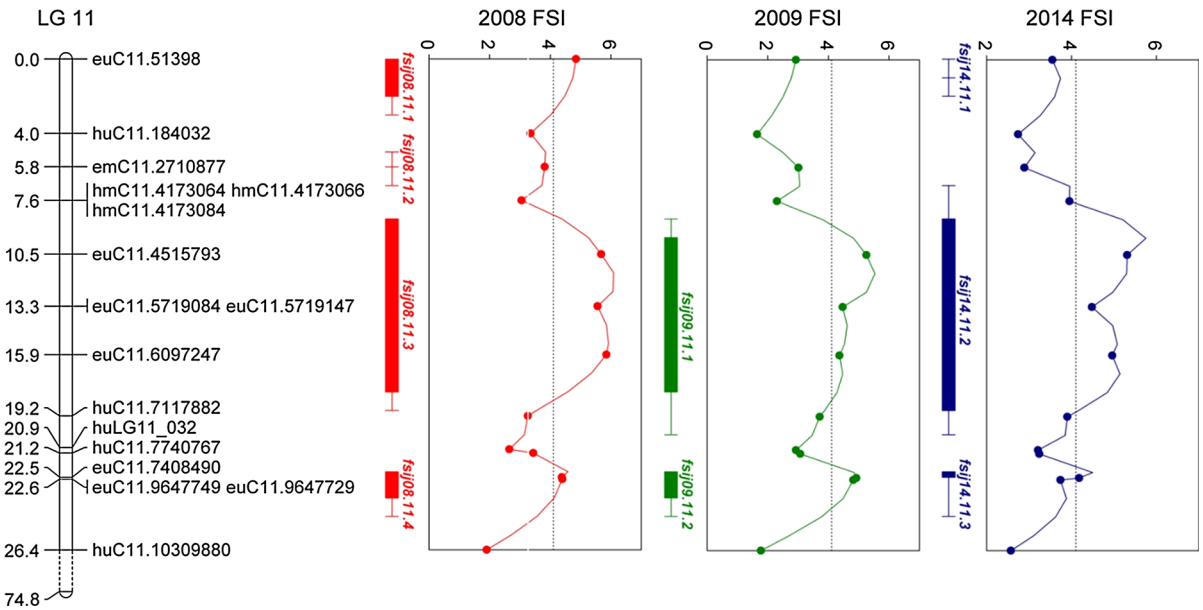

Fig. 1 Interval mapping of QTLs for FSI at a whole-genome LOD threshold ≥4.10 using the 'Jonathan' linkage map for the upper region of linkage group 11. Only the upper region of linkage group 11, which was strongly associated with FSI trait, is shown on the *left* of the figure. QTLs for FSI are shown to the *right* of the linkage group. Marker positions are listed to the *left side* of the linkage group. The LOD profile plots for simple interval mapping of FSI in 2008 (*red line*), 2009 (*green line*), and 2014 (*blue line*) are shown to the right of the linkage group. (Color figure online)

Table 3 Female parent-specific SNPs within the two FSI QTL intervals on chromosome 11

	Chromosome 11: 4411077–6723403	Chromosome 11: 7408490–7474945
Average coverage of total reads aligned to the interval of Jonathan	27.4	39.3
Average coverage of total reads aligned to the interval of Golden Delicious	28.5	24.2
The SNPs with genotype of lm × ll and hk × hk during population	11028	716
The SNPs within exons	2076	91
Nonsynonymous SNPs	913	61
Number of affected gene	187	12
Prioritized list of candidate genes based on annotation result of GDR website	3	0

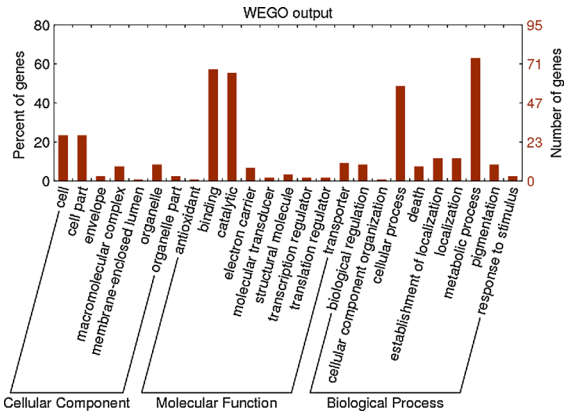

Fig. 2 Histogram of gene ontology classifications for apple of all annotated genes in the Locus 1 and Locus 2 QTL intervals of FSI trait. The genes corresponded to three main categories: biological process, cellular component, and molecular function. The right-hand y-axis indicates the number of genes and left-hand y-axis indicates percent of genes from apple in a category

processes, photosynthesis, cellular macromolecule metabolic processes, and oxidation reduction were also excluded from the grouping. Among the remaining genes, MDP0000135244 encoding peptidoglycan-binding LysM domain-containing protein was of particular interest because it involved in cell wall macromolecule metabolism. In addition, one gene without GO annotation, MDP0000896238, captured our attention because of its high similarity with *Arabidopsis thaliana ovate family protein 17* (*AtOFP17*, Wang et al. 2011). AtOFP family members function as transcription repressors, and they regulate multiple aspects of plant growth and development. However, *AtOFP17* was only highly expressed in mature roots, but not in flowers, siliques and leaves.

In addition, candidate genes were predicted based on involvement in the fruit shape formation pathway in apple and tomato to locate them on the apple genome sequence using BLASTP analysis. The five genes most correlated with *ovate* (NP_001234221) were located on chromosomes 2 (MDP0000155311 and MDP0000243940), 8 (MDP0000296199), 15 (MDP0000140421), and the unanchored chromosome (MDP0000921224). The genes most correlated with *fas* (NP_001234390) were on chromosomes 11 (MDP0000331808), 2 (MDP0000255415), 10 (MDP0000252640), 5 (MDP0000366291), and 9 (MDP0000192940). However, the detailed physical position of MDP0000331808 was at the bottom of chromosome 11, which was not consistent with the results in this study. The genes most correlated with *sun* (NP_001233793) were on chromosomes 2 (MDP0000639465), 14 (MDP0000299240), unanchored (MDP0000226363), and 1 (MDP0000185385 and MDP0000238367). There were no correlated genes for *lc* (JF284938 and JF284939) in the apple genome. In *Arabidopsis*, *ANT* and *AINTEGUMENTA-LIKE* genes regulate floral growth and ovule development, respectively (Mizukami and Fischer 2000; Krizek 2009). In apple, two *ANT* genes, *MdANT1* and *MdANT2*, were also reported to be involved in cell division during fruit growth (Madhumita and Anish 2012). One predicted apple protein (MDP0000182395, chromosome 11: 2860500–2863619), which also encoded AP2-like ethylene-responsive transcription factor (similar to the *MdANT1* and *MdANT2*), was screened and positioned at the top of chromosome 11. Unfortunately, this gene was not located in the Locus 1 interval.

Expression of two candidate genes for FSI trait in apple fruit

The results of the expression analysis of two candidate genes, MDP0000135244 and MDP0000896238, by RNAseq are shown in Additional file 3. The expression of MDP0000135244 increased during fruit development, especially at the cell expansion stage, but not during cell division (Janssen et al. 2008). The expression of MDP0000896238 was almost negligible, similar to findings in *A. thaliana* (Wang et al. 2011) that suggested MDP0000896238 was unrelated to fruit shape development.

Verification of candidate genes in parental cultivars and their hybrids

The nonsynonymous SNP in the gene of interest, MDP0000135244, was validated between the parents using Sanger sequencing. The result showed that the SNP of MDP0000135244 in the female parent 'Jonathan' was reliable. This SNP (designated C11.6053728) was polymorphic in 'Jonathan' (T/G) and monomorphic (T/T) in 'Golden Delicious' (Fig. 3). Domain motif annotation of the full-length genomic DNA sequence of MDP0000135244 revealed that the SNP was located in the first predicted exon. Amino acid translation revealed a semi-conserved substitution from asparagine (codon AAT) to lysine (codon AAG).

Then, C11.6053728 of MDP0000135244 was genotyped in 30 hybrids with high FSI ($\geq$0.9) and in 30 hybrids with low FSI ($\leq$0.8). The allele T/G was present in 23 of 30 high-FSI hybrids, and T/T was present in half of the low-FSI hybrids. The C11.6053728 genotypes were significantly correlated with FSI phenotypes, revealing a close association of C11.60537228 with FSI trait (Table 4).

Discussion

Effect of population and environment on detection of stable QTLs

Considering genotype × genotype and genotype × environment interactions, it is difficult to detect stable QTLs using different populations combined with changing environmental conditions among different planting locations and sampling years (Kenis et al. 2008). For example, QTLs for fruit length were located on LGs 2, 6, and 17 of 'Telamon' in a 'Telamon' × 'Braeburn' population in 1 year and on LGs 15 and 17 in the next year (Kenis et al. 2008). Sarah et al. (2014) detected QTLs for fruit length on LG 3 of the 'Co-op 17' linkage map in 2008 and on LGs 3 and 5 in 2009. In our study, fruit length was located on LG 7 of the 'Jonathan' linkage map in 2008 and on LG 11 in 2009. Fruit shape is a quantitative characteristic that is controlled by several genes (Sun et al. 2012). King et al. (2000) reported that QTL detection in a population is the result of single alleles derived from one parent, which must therefore be heterozygous for the locus underlying this trait. Therefore, differential segregation of alleles controlling fruit shape in parents from different populations is the primary reason stable QTLs were not detected. In addition, despite the identification of QTLs for fruit diameter in this study, *fdj08.3* was mapped on LG 3 (5.4 Mb from the top of the chromosome), the same as the LG reported by Sarah et al. (2014). However, the two QTLs were far apart, indicating that two genes were included in those QTLs.

Environmental differences between years will lead to the detection of different QTLs, even in the same population. Here, QTLs for fruit length were mapped at different LGs in 3 years, coinciding with weak correlations for the trait between 2008 or 2009 and 2014 (Table 1). However, FSI has high heritability in apple (Sun et al. 2012) and a strong correlation among sampling years, as Diaz et al. (2014) reported for melon. Huang et al. (1997) suggested that stable QTLs of traits with high heritability are easy to detect under different environmental conditions. Our findings for QTLs of FSI support that conclusion.

Two stable QTLs of FSI with highly explained variance in different environment and population were detected for high efficiency molecular breeding

The variance explained by a QTL is a key factor affecting the value of that marker for practical use. In general, a marker is considered useful when it is obtained from 'major' QTLs that explain more than 20 % of population variance (Kenis et al. 2008). In addition, marker value is dependent on population size. Sarah et al. (2014) identified a main-effect QTL for fruit length in 2008 that explained up to 46.4 % of the variance with only 46 seedlings. The variance explained decreased to 20.1 % when the population

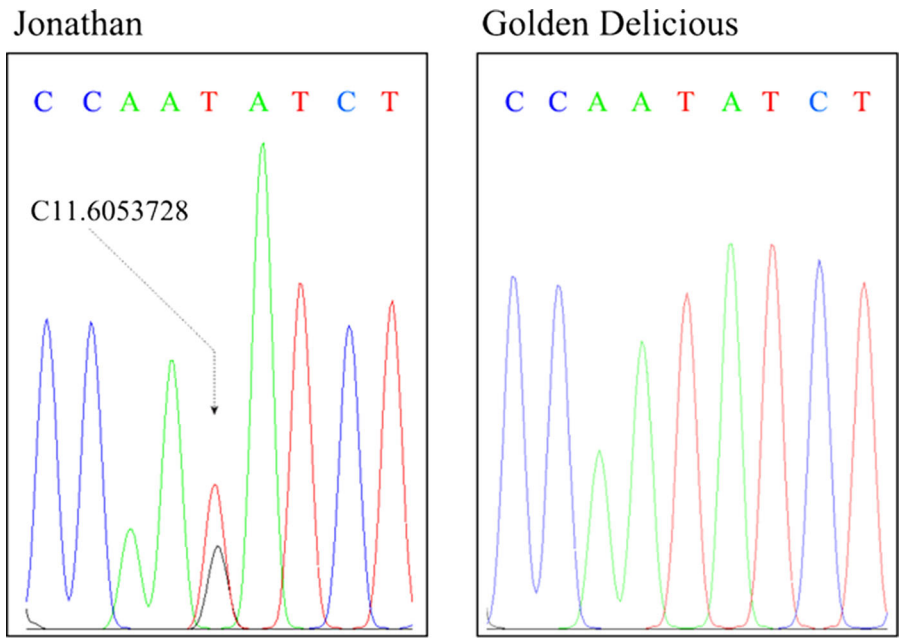

Fig. 3 SNP in MDP000013524 between 'Jonathan' and 'Golden Delicious' apples. (Color figure online)

Table 4 Genotype of nonsynonymous SNPs for MDP0000135244 in 60 hybrids

Phenotype	Genotype		$\chi 2$	*P* value
	T/G	T/T		
High FSI	23	7	4.59	0.03
Low FSI	15	15		

$\chi^2_{0.05} = 3.84$, $\chi^2_{0.01} = 6.64$

size increased to 86 in 2009. Here, the QTL fsij09.11.1 explained 23.0 % of variance in 2009 and was detected using 112 seedlings, and fsij14.11.2 was detected in 2014 using 122 seedlings and explained 22.8 % of the variance. These values indicate that these two QTLs are major determinants of FSI.

A marker is highly efficient for marker-assisted selection (MAS) when it is stable across different genetic surroundings and years. For example, Md-ACS1-2 and Md-ACO1-1 are functional genetic markers associated with fruit texture when both are homozygous in a specific variety. These markers indicate that the variety has low ethylene production and high storability (Costa et al. 2005). Md-ACS1-2 and Md-ACO1-1 have been widely applied in apple breeding after the verification of their accuracy in 255 varieties (www.rosbreed.org). In our study, two year-stable loci associated with FSI (Locus 1 and Locus 2) were mapped to 4.4–6.7 Mb and 7.4–7.5 Mb, respectively, on chromosome 11 of 'Jonathan'. Locus 2 was located near the previously reported QTL region (8.9–10.8 Mb on LG 11) using SSR linkage maps (Chang et al. 2014), and that location was also close to a major gene in a study using bulked segregant analysis with the same population (Sun et al. 2012). Although the present study did not detect significant QTLs of FSI trait in 2010 or 2011, a peak LOD was present at 5.8 cM of LG 11 of the 'Jonathan' linkage map (data not shown). Therefore, the QTLs of FSI trait identified here appear to have been stable between years and identification methods.

However, previously reported QTLs for apple shape mostly focused on fruit length and diameter, and no reports about QTLs for FSI trait are available for comparing the locations between our and other populations. We found one study that examined pear using SSR and amplified fragment length polymorphism (AFLP) markers, which reported an anchored SSR marker, NB105a, positioned on a linkage group with QTLs for pear FSI trait (Zhang et al. 2013). Surprisingly, we found that NB105a was located on LG 11 of the apple genome Celton et al. (2009). Because pear and apple chromosomes show high co-linearity, this result indicated that the QTLs identified by Zhang et al. (2013) should be located on LG 11 in

pears. Thus, QTLs of FSI trait obtained in our study and confirmed by findings in pears reflect that a stable QTL region exists among different populations.

Combining high-density linkage maps with whole-genome resequencing is helpful for screening candidate genes

High-density linkage maps provide a useful tool for the efficient identification of candidate genes. In this study, Locus 2 for FSI was mapped to a 67-kb genomic region using high-density SNP linkage maps and was available for further candidate gene prediction. Locus 1 was located in a ~2.3 Mb region, an interval that was too long for direct prediction of the candidate gene. However, the nearest marker, euC11.4515793 at 4.5 Mb from the top of chromosome 11, supplied useful information about the physical location for indirect screening. Van der Knaap et al. (2013) also maintained that SNP markers were more efficient than traditional AFLP and SSR markers for locating QTLs.

Fine mapping was formerly an efficient approach for map-based candidate gene mining. However, the use of large numbers of seedlings limits the use of this technology for fruit crops. For example, the apple scab resistance gene *Rvi12* was fine-mapped to an interval spanning 958 kb of the 'Golden Delicious' genome based on a mapping population of 1285 seedlings (Padmarasu et al. 2014). The construction of mapping populations in fruit is time-consuming and laborious. Therefore, narrowing the interval for Locus 1 in our study by increasing population size was not the best choice. Fortunately, Moriya et al. (2012) described a new method for QTL detection that is more efficient than traditional map-based cloning in fruit trees. Using 1000 F_1 seedlings from 31 population crosses, the *Co* gene, which controls the columnar growth habit of apple, was fine-mapped to a 196 kb region of LG 10 (Huang et al. 2011). Their results illustrated that the use of multiple families for QTL mapping, which considers a high number of allelic variations, showed high resolution in QTL detection (Huang et al. 2011). The same approach has been reported for other fruit crops (Rosyara et al. 2013).

Combined with whole-genome re-sequencing data from parents or hybrids with various phenotypes, candidate genes within major QTLs could be explored without using large fine-mapping populations. Using whole-genome sequencing of four sibling trees from an F_2 mapping population, candidate genes for chilling requirement traits were prioritised in peach (Zhebentyayeva et al. 2014). In our study, using whole genome resequencing data from the parental cultivars, nonsynonymous SNPs affecting 187 candidate genes within Locus 1 were predicted over a 2.3-Mb genomic region. This result provides a foundation for future research related to FSI trait. However, a disadvantage of this method is that only nonsynonymous SNPs in exons were considered. Several mutations in the promoter region, which is also involved in regulation of traits, were ignored by this bias. Therefore, the combined use of multiple families, high-density SNP maps, and resequencing information from accessions with various phenotypes will improve the efficiency and effectiveness of causal gene identification.

MDP0000135244 encoding an LysM domain receptor-like kinase protein was a key candidate gene related to FSI in apple

MDP0000135244 possesses a peptidoglycan-binding lysM domain. In previous years, this sequence motif was found in enzymes involved in cell wall degradation in bacteria (Bateman and Bycroft 2000), but whether it is involved in cell wall development in plants is unclear. Proteins including this motif in *A. thaliana* and *Oryza sativa* have antibacterial functions (Erbs et al. 2008; Shimizu et al. 2010). The protein with this motif in *A. thaliana* (AtLYK5) is the primary receptor for chitin and forms a chitin-inducible complex with CERK1 to induce plant immunity (Cao et al. 2014). It is well known that hypersensitive response (resulting in programmed cell death), callose deposition, and enlarged cells are general symptoms during immune responses (Rate and Greenberg 2001). That study may indicate that *AtLYK5* is related to cell development, with the next step in that research being to determine how CERK1 and LYK5 are able to activate plant defenses. To identify the relationship between MDP0000135244 and the FSI trait in apple, the genotype of SNP C11.6053728 in the population was analyzed. The heterozygous genotype (T/G) frequency of SNP C11.6053728 in the 30 individuals with high FSI was 23/30, which showed good correlation between the T/G genotype and high-FSI phenotype. However, the T/T genotype frequency in the 30 individuals with low FSI was 15/30, suggesting that the heterozygous and homozygous genotypes

should segregate randomly in the low-FSI group. This result was consistent with the expected segregation ratio of 1:1 in the low-FSI group because the trait was dominant for high FSI, and five segregating loci were involved in the variation in FSI (Sun et al. 2012). In general, the gene MDP0000135244 containing SNP C11.6053728 presumably is important in cell wall metabolism. In future work, a diagnostic marker designed based on these SNPs will be used to select elite hybrids for FSI trait in apple.

Acknowledgments This work was supported by the Hi-Tech Research and Development (863) Program of China (2011AA001204); National Special Funds for Scientific Research on Public Causes (Agriculture) Project (200903044); Modern Agricultural Industry Technology System (Apple) (CARS-28); and Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (Nutrition and Physiology), Ministry of Agriculture, P. R. China.

Conflict of interest The authors declare that they have no competing financial interest.

References

Bateman A, Bycroft M (2000) The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). J Mol Biol 299:1113–1119

Cao YR, Liang Y, Tanaka K, Nguyen CT, Jedrzejczak RP, Joachimiak A, Stacey G (2014) The kinase LYK5 is a major chitin receptor in *Arabidopsis* and forms a chitin-induced complex with related kinase CERK1. Elife 3:e03766

Celton JM, Tustin DS, Chagné D, Gardiner SE (2009) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. Tree Genet Genomes 5:93–107

Chang YS, Sun R, Sun HH, Zhao YB, Han YP, Chen DM, Wang Y, Zhang XZ, Han ZH (2014) Mapping of quantitative trait loci corroborates independent genetic control of apple size and shape. Sci Hortic 174:126–132

Cong B, Liu J, Tanksley SD (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proc Natl Acad Sci USA 99:13606–13611

Costa F, Stella S, Van de Weg WE, Guerra W, Cecchinel M, Dallavia J, Koller B, Sansavini S (2005) Role of the genes*Md-ACO1* and *Md-ACS1* in ethylene production and shelf life of apple (*Malus domestica* Borkh). Euphytica 141:181–190

Diaz A, Zarouri B, Fergany M, Eduardo I, Alvarez JM, Pico B, Monforte AJ (2014) Mapping and introgression of QTL involved in fruit shape transgressive segregation into 'Piel de Sapo' melon (*Cucucumismelo* L.). PLoS One 9:e104188

Dunemann F, Ulrich D, Boudichevskaia A, Grafe C, Weber WE (2009) QTL mapping of aroma compounds analysed by headspace solid-phase microextraction gas chromatography in the apple progeny 'Discovery' × 'Prima'. Mol Breed 23:501–521

Durel CE, Denancé C, Brisset MN (2009) Two distinct major QTL for resistance to fire blight co-localize on linkage group 12 in apple genotypes 'Evereste' and *Malus* floribunda clone 821. Genome 52:139–147

Erbs G, Silipo A, Aslam S, De Castro C, Liparoti V, Flagiello A, Pucci P, Lanzetta R, Parrilli M, Molinaro A, Newman MA, Cooper RM (2008) Peptidoglycan and muropeptides from pathogens *Agrobacterium* and *Xanthomonas* elicit plant innate immunity: structure and activity. Chem Biol 15:438–448

Gardiner SE, Bus VGM, Rusholme RL, Chagné D, Rikkerink EHA (2007) Genome mapping and molecular breeding. In: Kole C (ed) Fruit and Nuts, vol 4. Springer, Berlin, Heidelberg, pp 2–62

Gustavo RR, Stephane M, Claire A, Sung-Chur S, Andrew M, Mathilde C, Brian BMG, David F, van der Knaap E (2011) Distribution of *SUN*, *OVATE*, *LC*, and *FAS* in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol 156:275–285

Hazbavi I (2014) Shape and size grading of apple fruit (cv. Fuji) based on geometrical properties. IJB 4:269–273

Huang N, Angeles ER, Doming J (1997) Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. Theor Appl Genet 95:313–320

Huang XH, Paulo M, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk FA (2011) Analysis of natural allelic variation in *Arabidopsis* using a multiparent recombinant inbred line population. Proc Natl Acad Sci USA 108:4488–4493

Janssen BJ, Thodey K, Schaffer RJ, Alba R, Balakrishnan L, Bishop R, Bowen JH, Crowhurst RN, Gleave AP, Ledger S, McArney S, Pichler FB, Snowden KC, Ward S (2008) Global gene expression analysis of apple fruit development from the floral bud to ripe fruit. BMC Plant Biol 8:16

Kenis K, Keulemans J, Davey MW (2008) Identification and stability of QTLs for fruit quality traits in apple. Tree Genet Genomes 4:647–661

King GJ, Maliepaard C, Lynn JR, Alston FH, Durel CE, Evans KM, Griffon B, Laurens F, Manganaris AG, Schrevens T, TartariniS Verhaegh J (2000) Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (*Malus pumila* Mill.). Theor Appl Genet 100:1074–1084

Krizek B (2009) *AINTEGUMENTA* and *AINTEGUMENTA-LIKE6* act redundantly to regulate *Arabidopsis* floral growth and patterning. Plant Physiol 150:1916–1929

Kunihisa M, Moriya S, Abe K, Okada K, Haji T, Hayashi T, Kim H, Nishitani C, Terakami S, Yamamoto T (2014) Identification of QTLs for fruit quality traits in Japanese apples: QTLs for early ripening are tightly related to preharvest fruit drop. Breed Sci 64:240–251

Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci USA 99:13302–13306

Madhumita D, Anish M (2012) The AINTEGUMENTA genes, *MdANT1* and *MdANT2*, are associated with the regulation of cell production during fruit growth in apple (*Malus* × *domestica* Borkh.). BMC Plant Biol 12:98

Mizukami Y, Fischer RL (2000) Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. Proc Natl Acad Sci USA 97:942–947

Moriya S, Okada K, Haji T, Yamamoto T, Abe K (2012) Fine mapping of *Co*, a gene controlling columnar growth habit located on apple (*Malus* × *domestica*) linkage group 10. Plant Breed 131:641–647

Padmarasu S, Sargent DJ, Jaensch M, Kellerhals M, Tartarini S, Velasco R, Troggio M, Patocchi A (2014) Fine-mapping of the apple scab resistance locus *Rvi12* (*Vb*) derived from 'Hansen's baccata #2'. Mol Breed 34:2119–2129

Rate DN, Greenberg JT (2001) The *Arabidopsis* aberrant growth and death2 mutant shows resistance to *Pseudomonas syringae* and reveals a role for NPR1 in suppressing hypersensitive cell death. Plant J 27:203–211

Rosyara UR, Bink MCAM, van de Weg E, Zhang GR, Wang DC, Sebolt A, Dirlewanger E, Quero-Garcia J, Schuster M, Iezzoni AF (2013) Fruit size QTL identification and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of sweet cherry. Mol Breed 32:875–887

Sarah MP, Khan MA, Han YP, Kushad MM, Korban SS (2014) Identification of quantitative trait loci (QTLs) for fruit quality traits in apple. Plant Mol Biol Rep 32:109–116

Shimizu T, Nakano T, Takamizawa D, Desaki Y, Ishii-Minami N, Nishizawa Y, Minami E, Okada K, Yamane H, Kaku H, Shibuya N (2010) Two LysM receptor molecules, *CEBiP* and *OsCERK1*, cooperatively regulate chitin elicitor signaling in rice. Plant J 64:204–214

Stoeckli S, Mody K, Gessler C, Patocchi A, Jermini M, Dorn S (2008) QTL analysis for aphid resistance and growth traits in apple. Tree Genet Genomes 4:833–847

Sun HH, Zhao YB, Li CM, Chen DM, Wang Y, Zhang XZ, Han ZH (2012) Identification of markers linked to major gene loci involved in determination of fruit shape index of apples (*Malus domestica*). Euphytica 185:185–193

Van der Knaap E, Tanksley SD (2001) Identification and characterization of a novel locus controlling early fruit development in tomato. Theor Appl Genet 103:353–358

Van der Knaap E, Anderson C, Rodríguez G (2013) Diversity within cultivated tomato. In: Liedl BE, Labate JA, Stommel JR, Slade A, Kole C (eds) Genetics, genomics, and breeding of tomato. Science Publishers, New York

Van Ooijen JW (2009) MapQTL 6.0 Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV, Wageningen

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M et al (2010) The genome of the domesticated apple (*Malus* × *domestica*Borkh). Nat Genet 42:833–839

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Wang SC, Chang Y, Guo JJ, Zeng QN, Ellis BE, Chen JG (2011) *Arabidopsis* ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. PLoS One 6:e23896

WöhnerTW Flachowsky H, Richter K, Garcia-Libreros T, Trognitz F, Hanke MV, Peil A (2014) QTL mapping of fire blight resistance in *Malus* × *robusta* 5 after inoculation with different strains of *Erwinia amylovora*. Mol Breed 34:217–230

Ye J, Fang L, Zheng H, Zhang Y, Cheng J, Zhang ZJ, Wang J, Li ST, Li RQ, Bolund L, Wang J (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34:W293–W297

Zhang RP, Wu J, Li XG, Khan MA, Chen H, Korban SS, Zhang SL (2013) An AFLP, SRAP, and SSR genetic linkage map and identification of QTLs for fruit traits in pear (*Pyrus* L.). Plant Mol Biol Rep 31:678–687

Zhebentyayeva TN, Fan SH, Chandra A, Bielenberg DG, Reighard GL, Okie WR, Abbott AG (2014) Dissection of chilling requirement and bloom date QTLs in peach using a whole genome sequencing of sibling trees from an F_2 mapping population. Tree Genet Genomes 10:35–51