A *FRUITFULL-like* gene is associated with genetic variation for fruit flesh firmness in apple (*Malus domestica* Borkh.).

Volkan Cevik[1], Carol D. Ryder[1] Alexandra Popovich[1], Kenneth Manning[1], Graham J. King[2], Graham B. Seymour[3,*]

Addresses

[1]Warwick-HRI, Warwick University, Wellesbourne, Warwick, CV35 9EF, United Kingdom.
[2]Rothamsted Research, Harpenden, Hertfordshire, AL5 2QJ, United Kingdom.
[3]Division of Plant and Crop Sciences, School of Biosciences, University of Nottingham, Sutton Bonington, Loughborough, Leics. LE12 5RD. UK.

* Corresponding author. Tel: + 44 (0)115 951 6323, Fax:+ 44 (0)115 951 6334,
Email: graham.seymour@nottingham.ac.uk

Keywords: *Malus domestica*, Fruit Flesh Texture, *FRUITFUL*, Association Mapping

Abstract

The *FRUITFULL* (*FUL*) and *SHATTERPROOF* (*SHP*) genes are involved in regulating fruit development and dehiscence in Arabidopsis. We tested the hypothesis that this class of genes are also involved in regulating the development of fleshy fruits, by exploring genetic and phenotypic variation within the apple (*Malus domestica*) gene pool. We isolated and characterised the genomic sequences of two candidate orthologous *FUL*-like genes, *MdMADS2.1* and *MdMADS2.2*. These were mapped using the reference population 'Prima x Fiesta' to loci on *Malus* linkage groups LG14 and LG06, respectively. An additional MADS-box gene, *MdMADS14* shares high amino acid identity with the Arabidopsis *SHATTERPROOF1/2* genes and was mapped to *Malus* linkage group LG09. Association analysis between quantitative fruit flesh firmness estimates of 'Prima x Fiesta' progeny and the *MdMADS2.1*, *MdMADS2.2* and *MdMADS14* loci was carried out using a Mixed Model Analysis of Variance. This revealed a significant association ($p<0.01$) between *MdMADS2.1* and fruit firmness (FFF-1). Further evidence for the association between *MdMADS2.1* and fruit flesh firmness was obtained using a case-control population-based genetic association approach. For this, a polymorphic repeat, $(AT)_n$, in the 3' UTR of *MdMADS2.1* was used as a locus-specific marker to screen 168 apple accessions for which historical assessments of fruit texture attributes were available. This analysis revealed a significant association between the *MdMADS2.1* and fruit flesh firmness at both allelic ($\chi2=34$, 9 df, $P<0.001$) and genotypic ($\chi2=57$, 32 df, $P<0.01$) levels.

Introduction

The regulatory network controlling dehiscence in the dry fruits of Arabidopsis has been the subject of intensive study. Several genes have been identified that are necessary for normal silique development. These include members of the MADS-domain transcription factor family *SHATTERPROOF* (*SHP*) and *FRUITFULL* (*FUL*) (Liljegren et al. 2000; Ferrandiz et al. 2000). The *SHP* genes (*SHP1* and *SHP2*) are required both for differentiation of the separation layer and to promote lignification of the valve margins. When both *SHP* genes are knocked out the silique fails to open and the seeds are trapped inside. *FUL* negatively regulates valve margin development and *ful* mutants show 'ectopic dehiscence zone formation' throughout the valve tissue (Ferrandiz 2002; Liljegren et al. 2004). Other dry fruited species also appear to share a similar regulatory mechanism. *FUL* is a member of the *SQUAMOSA* (*SQUA*) family of MADS-box transcription factors and genes with strong homology to *FUL* are present in the 'core eudicots' which contain the *Rosaceae* and *Solanaceae* families respectively (Litt and Irish 2003). This is likely to reflect a role for these genes in floral evolution where members of the *SQUA* family are central to floral development. Genes related to *FUL* could also play a role in fruit development throughout at least the core eudicots. A central question is whether this regulatory mechanism is confined to just the dry fruited species such as Arabidopsis and *Antirrhinum*. During Angiosperm evolution, fleshy fruits are thought to have arisen from dry forms. Phylogenetic evidence indicates that in the Solanaceae fleshy berries are derived from dry types such as capsules (Knapp 2002). We postulate that genes regulating the development and ripening of dry fruits may have been recruited to new, but related functions during the evolution of fleshy fruits (Seymour et al. 2008).

In tomato the gene at the *rin* locus is a MADS-box transcription factor and has been shown to be essential for fruit ripening (Vrebalov et al. 2002). Examination of possible *FUL* orthologues in tomato (*TDR4*, *LeFUL2*) reveals that these genes are highly expressed in carpel tissue and show up-regulation during fruit ripening (Eriksson et al. 2004; also see TIGR Tomato Gene Index, www.tigr.org). In the non-ripening tomato mutants, *nor*, *rin* and *Cnr*, there are low levels of *TDR4* expression (Eriksson et al. 2004) and this again suggests a role for this potential *FUL* orthologue in ripening in a fleshy fruit. In *Malus*, a possible *FUL* orthologue, *MdMADS2*, has already been partially described (Sung et al. 1999), but its role in fruit ripening, if any, is not known.

Previously significant QTLs for fruit firmness and flesh texture were identified on apple linkage groups L01, L03, L08, L10, L15 and L16 (King et al. 2000; Maliepaard et al. 2001 and Seymour et al. 2002). Recently, candidate gene approaches placed the *Md-ACO1* on the linkage group L10 within the 5% interval border of an important QTL for fruit firmness (Costa et al. 2005), and showed that allelic variation of an SSR within the *Md-Exp7* gene, which mapped on L01, was associated with fruit softening (Costa et al. 2008).

In this paper we describe the isolation and identification of apple *MdMADS2* / *FUL*-like genomic sequences and demonstrate a genetic linkage with fruit flesh firmness through the use of genetic association studies.

Materials and Methods

Plant material and DNA isolation

DNA samples from the reference population 'Prima x Fiesta' (Maliepaard et al. 1998), were obtained from the plantation at Warwick HRI, Wellesbourne, and used for the marker and trait linkage analysis. DNA for the population-based association analysis was obtained for a 'diversity set' derived from apple variety specimen trees held in the National Top Fruit Collection at Brogdale, Kent, UK (Supplementary Table 1). For the mapping population, total DNA was isolated from young leaf tissue using a cetyltrimethylammonium bromide (CTAB) miniprep method (Doyle and Doyle 1990). DNA was isolated from the apple diversity set using the Nucleospin multi-96 Plant DNA extraction Kit (Macherey-Nagel, Germany) following manufacturer's instructions.

Trait assessment

Fruit Flesh firmness (FFF-1) measurements for the reference population 'Prima x Fiesta' were described previously (King et al. 2000). The firmness data for the apple varieties were obtained from Smith (1971). Data were scanned by optical character recognition software directly from the entries of the varieties, then hand-edited and curated within a relational database. Smith categorised each variety according to a series of multistate trait characters, with restricted vocabulary. These included the following relating to fruit texture included firmness, coarseness, crispness and hardness (Table S1).

Identification of *FRUITFUL* and *SHATTERPROOF* homologous sequences in *Malus*

Identification and characterisation of the possible *AtFUL/AGL8* orthologue *MdMADS2* (accession no. U78948) in *Malus* has already been described (Sung et al. 1999). The primers 5'-ACTGGAAGCTGGACTCTGGAACATG-3' (forward) and 5'-CTACTGCGTTCTCCTTCTCCTTCAC-3' (reverse) were designed to amplify the 277 bp 3' region of *MdMADS2* (accession no. U78948) and used for PCR amplification on first strand cDNA synthesised from 4 week old, mature green and ripe fruits using Omniscript first-strand cDNA synthesis Kit (Qiagen) following manufacturer's instructions. For PCRs following conditions were used: 50 ng of first strand cDNA was amplified by PCR in a 25 µl final reaction volume containing 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 2.5 mM $MgCl_2$, 200 µM of each dNTP, 400 nM of each primer and 1.25 U Taq DNA polymerase. The PCR conditions consisted of a preliminary denaturation step of 3 min at 94^{o}C, followed by 30 cycles of 30 s at 94^{o}C, 30 s at 55^{o}C and 1 min at 72^{o}C , with a final extension for 5 min at 72^{o}C. The PCR products were analysed on a 2% agarose gel in 1x TBE (45 m M Tris-base, 45 m M boric acid, 1 m M EDTA, pH 8.0), and bands of the expected sizes were excised and gel-purified with the QIAquick Gel Extraction kit (Qiagen, UK) and the fragments were subjected to sequencing. Sequence analyses revealed that a fragment identical to that of the *MdMADS2* gene had been amplified. The fragment was then radioactively labelled with α-[^{32}P]-dCTP and used to hybridise the *Malus* BAC library constructed from the apple variety Florina (Vinatzer et al. 1998). Following identification and subsequent confirmation of the positive BAC clones with PCR analysis, the clones were digested with the restriction enzyme *Hind* III and the digestion patterns obtained were used to

assign the BAC clones into contigs. The sequences flanking the *MdMADS2* gene were obtained by the direct BAC sequencing approach. BAC DNA was purified using a Large Construct Purification kit (Qiagen, UK) following manufacturer's instructions. The BAC sequencing reaction consisted of 1 µg of BAC DNA, 4 µl BigDye™ terminator mix (Applied Bioscience), 4 µl 5x buffer (400mM Tris-HCl, pH 9.0, 10mM $MgCl_2$) and 10µM primer. The cycle conditions consisted of a preliminary denaturation step of 5 min at 95^{o}C, followed by 45-75 cycles of 30 s at 95^{o}C, 10 s at 50^{o}C and 4 min at 60^{o}C. After obtaining 250-400 bp sequence, a new primer was designed and used for subsequent sequencing reactions until approximately 1 kb additional sequence from the initiation and stop codons of the gene were obtained.

Both nucleotide and amino acid sequences of the Arabidopsis *SHATTERPROOF1/2* (Liljegren et al. 2000) were used to identify the homologoue *Malus* sequence(s) using the tBlastx algorithm. This analysis identified *MdMADS14* (accession no. AJ251117) with 68% overall (81% positives) amino acid identity as a possible *SHP* orthologue in the *Malus* genome.

Genetic mapping of *MdMADS2.1, MdMADS2.2* and *MdMADS14* genes in the *Malus* genome

Following identification of two possible *MdMADS2* paralogoues, specific primer pairs were designed for *MdMADS2.1* (5'-GATAGGTATTGTGTTTCCATCCATGC-3' as forward and 5'-TGCTATTTGCAAGTGTTAGTGTTGTGG-3' as reverse) and *MdMADS2.2* (5'-AAGAAGGTAAGCCTGAGCTAGAAGG-3' as forward and 5'-CCTCCTCCTTTCAACCATACTTTCC-3' as reverse). Sequence analysis of

MdMADS2.1 amplicons from both parents identified a SNP on one of the alleles present in the female parent Prima. This corresponded to the *Tsp*45 I restriction site and enabled the development of a cleaved amplified polymorphic sequence (CAPS) assay marker. A SNP present on one of the Prima alleles was identified following comparison of *MdMADS2.2* sequences of both parents. This corresponded to a restriction site for the enzyme *Ava*II and enabled the development of a CAPS assay.

MdMADS14 amplification was carried out with the primers MADS14B-F (5'-TAACCATTACCCTCGCCAAG-3') and MADS14B-R (5'-GAATTGGGGTCCAT CATCAC-3') designed specific to 3' end of the gene using DNA from 'Prima' and 'Fiesta'. The amplified samples were then sequenced. One of the SNPs identified was specific to Fiesta allele and corresponded to the restriction enzyme *HpyCH*4III recognition site. This enabled the development of a CAPS assay.

Genetic linkage analysis was performed using JoinMap V.3.0 (Van Ooijen and Voorrips 2001). Distances (in centi-Morgans) were calculated with the Kosambi (1944) function.

Association analysis using mixed model analysis of variance (ANOVA)

In order to test if the genes *MdMADS2.1*, *MdMADS2.2* and *MdMADS14* were associated with the fruit flesh firmness (FFF-1), previously reported by King et al. 2000; 2001, in the reference population 'Prima x Fiesta', a mixed model analysis of variance (ANOVA) was carried out. For this, *MdMADS2.1*, *MdMADS2.2* and *MdMADS14* were used as fixed effects and the molecular markers corresponding to the previously identified fruit firmness QTL regions (i.e. the markers OPAM-19-2200 and OPC-09-0900 on LG01, OPA092000 on LG03, OPAC110780 on LG08, ACO1 on LG10,

NZ02b1 and LY26a on LG15 (King et al. 2000; 2001)) as random effects using the REML estimates of the reference population 'Prima x Fiesta' fruit flesh firmness data (FFF-1) (King et al. 2000; 2001). Post-hoc comparison between the alleles of the *MdMADS2.1* was carried out using Tukey's test. All statistical tests were carried out using SAS v8.2 (SAS Institute).

Population-based association analysis

A polymorphic repeat $(AT)_n$ in the 3' UTR of the *MdMADS2.1* was identified and used as a marker to screen 192 (96 Firm and 96 Soft) apple accessions obtained from the National Top Fruit Collection at Brogdale, Kent, UK. The PCR amplification was carried out using a fluorescently tagged forward primer 5'-FAM-GATGCTCCGCCACCTTAAT-3' and 5'-AGGGTTTTCATCATGCACATT-3' as reverse primer. Twelve un-linked SSR markers (Supplementary Table 2) distributed on 12 linkage groups, which do not correspond to previously identified fruit flesh firmness QTL regions (King et al. 2000; 2001), were used for the assessment of possible sub-structure within the apple accessions.

The 10 µl PCR reaction mixture consisted of 20 ng of genomic DNA, 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 2.5 mM $MgCl_2$, 200 µM of each dNTP, 400 nM of each primer and 1.25 U Taq DNA polymerase. The PCR conditions consisted of a preliminary denaturation step of 3 min at 94^{o}C, followed by 30 cycles of 30 s at 94^{o}C, 30 s at 50^{o}C and 30 s at 72^{o}C, with a final extension for 7 min at 72^{o}C. The amplified fragments were prepared for analysis on an ABI Prism 3100 Genetic Analyser (Applied Biosystem) by mixing 1 µl of a 1:10 dilution of the PCR products with 10 µl of Hi-Di Formamide (Applied Biosystems) and 0.15 µl a ROX GS500 internal standard (Applied

Biosystems). GeneMarker Version 1.1 (SoftGenetics LLC) software was then used to analyze the PCR products.

The genetic structure of the apple accessions was investigated with the Bayesian model-based clustering algorithm implemented in STRUCTURE v. 2.1 (Pritchard et al. 2000; Falush et al. 2003). Each analysis was performed without any prior population assignment with a burn-in of 100,000 and a run length of 1000,000 iterations. We carried out five runs of clusters from K=1 to K=8 allowing for the basic admixture model with uncorrelated allele frequencies. The run showing the highest posterior probability of data was considered for each K value. Basic statistics, including the number of alleles, observed heterozygosity and gene diversity (or expected heterozygosity) for the markers were calculated using PowerMarker V3.23 software (Liu 2002).

Allele and genotype distributions between the soft and firm populations were then compared using the chi-square method calculated by PowerMarker V3.23 software (Liu 2002).

Results

Characterisation of *MdMADS2.1* and *MdMADS2.2* genomic sequences

Hybridisation experiments and subsequent PCR analysis identified 11 candidate BAC clones containing *MdMADS2* sequences. Further fingerprinting of these BAC clones revealed two distinct contigs. This suggested that there were two possible paralogous sequences and loci within the *Malus* genome. We named these sequences *MdMADS2.1* (accession no.DQ205652) and *MdMADS2.2* (accession no.DQ205651). Using the direct BAC sequencing and primer walking approaches, we obtained 8,250 bp

and 6,167 bp of sequence for *MdMADS2.1*and *MdMADS2.2,* respectively (Fig. 1b). Alignment of the predicted amino acid sequence of the *MdMADS2.1* differs from that of the previously identified *MdMADS2* (U78948) cDNA sequence by only two amino acids (99% identity). These are therefore likely to be two different allelic forms of the same gene. However, the predicted amino acid sequence of *MdMADS2.2* showed only 84% identity with *MdMADS2* suggesting that *MdMADS2.2* is a possible paralogue of the *MdMADS2.1* (Fig. 1a). A schematic diagram of the predicted introns and exons of the *MdMADS2.1* and *MdMADS2.2,* as well as examples of the nucleotide differences between these two genes are given in Fig. 1b.

Linkage mapping of the *MdMADS2.1*, *MdMADS2.2* and *MdMADS14*

The *MdMADS2.1*, *MdMADS2.2* and *MdMADS14* specific primers were used to amplify sequences from Prima and Fiesta genomic DNA. For each of these genes only a single product was amplified from genomic DNA of both parents. Sequence analysis of *MdMADS2.1* amplicons from both parents identified a SNP in one of the alleles present in the female parent Prima. This enabled the development of a CAPS assay to map *MdMADS2.1* to 41cM on linkage group LG14 (LOD>10), located between the markers 7B9a and E35/M47/P4. (Fig. 2)

A SNP for the sequenced region of the *MdMads2.2* gene was identified. The SNP was used to design a CAPS assay, which allowed mapping of the *MdMADS2.2* gene to 20 cM on linkage group LG06 (LOD>10) between the markers 7B9b and UBC219-1000 (Fig. 2).

A SNP specific to one of the Fiesta alleles of *MdMADS14* was used to develop a CAPS assay which allowed the *MdMADS14* gene to be mapped to 14 cM on linkage

group LG09 (LOD>10) (Fig. 2) between the markers OPAB-14-0800 and OPAE-01-1190.

Association analysis using mixed model ANOVA

The three genes were located on linkage groups, which had not been associated with significant QTL accounting for variation in fruit flesh firmness (FFF-1) previously identified in the Prima x Fiesta population (King et al, 2000; 2001). In order to determine whether weak effects existed between the gene loci *MdMADS2.1*, *MdMADS2.2* and *MdMADS14* and fruit flesh firmness, we reanalysed the segregating trait data using a mixed model ANOVA. In order to reduce the number of parameters, first all the genes and the markers were included in the test, and then the genes and markers with a *P* value of higher than 0.20 were removed in subsequent analyses in order to simplify the model. At this stage *MdMADS2.2* was also excluded as the probability value was higher than the threshold value of 0.20 (*P*= 0.60). *MdMADS2.1* and *MdMADS14* were then used as fixed effects with the molecular markers corresponding to the previously identified fruit firmness QTL regions (i.e. markers on the linkage groups LG01, LG08, LG10, LG15, LG03) as random effects. In this analysis, *ACO1* (LG10), *MdMADS2.1* (LG14), OPC-09-0900 (LG01) and LY26a (LG15) all had highly significant effects on fruit flesh firmness (*P*<0.01) (Table 1). The model used explained 50 % of the variation in fruit flesh firmness in the reference population ‘Prima x Fiesta (Table 1). Prima was shown to produce fruit with softer texture than Fiesta in a previous study (King et al. 2001). The progeny with the Prima allele of *MdMADS2.1* had significantly lower fruit flesh firmness.

Population-based association analysis

In order to further test the contribution of the *MdMADS2.1* gene to apple fruit flesh firmness, a population-based association analysis was carried out (Mackay and Powell, 2007). 192 apple accessions (96 Soft and 96 Firm) with fruit firmness data were selected from the National Top Fruit Collection at Brogdale, Kent, UK. In order to avoid spurious association as a result of population sub-structure, twelve unlinked SSR markers (Supplementary Table 2) were used to investigate population structure in the apple diversity set using the software STRUCTURE v. 2.1 (Pritchard et al. 2000) implementing a Bayesian clustering approach. This analysis revealed that there was no significant population structure in the overall sample. Under the model tested, the smallest probability is associated with a prior K of 1 (Posterior probability of 1). The proportion of individuals assigned to each cluster is approximately the same (Supplementary Fig. 1). Therefore, a possible effect of population structure on the subsequent association analysis was found to be unlikely.

The microsatellite sequence $(TA)_n$ located within the 3' UTR region of the *MdMADS2.1* gene was used as a molecular marker in a fluorescence-based genotyping of the apple accessions. Good quality data were obtained from 168 apple accessions. A total of 10 alleles and 33 different genotypes were observed, with the allele sizes ranging from 174 to 196 bp. The distributions and frequency of the alleles obtained within firm and soft apple accessions are given in Fig. 3. Using this data set, a case-control population-based association analysis was carried out. This analysis identified a significant association between the sensory firmness data and *MdMADS2.1* 3'UTR SSR marker both at allelic ($\chi 2$=34, 9 df, P<0.001) and genotypic level ($\chi 2$=57, 32 df,

P<0.01) (Table 2). We conclude from this that there is a significant association between the fruit flesh firmness and the *MdMADS2.1* gene.

Discussion

The *FUL* and *SHP* genes have been shown to be involved in development and 'ripening' in the dry fruits of Arabidopsis (Dinneny and Yanofsky 2005). Related sequences are also found to be expressed in fleshy fruits including tomato and peach (Seymour et al., 2008). In this study we identified genomic sequences of two *FUL*-like genes, *MdMADS2.1* and *MdMADS2.2*, present in the *Malus* genome. These two paralogous genes are highly similar in their predicted coding sequences (Fig. 1a) but show divergence in their non-coding regions (Fig. 1b). In a previous study (Sung et al. 1999) expression analysis revealed that *MdMADS2* (corresponding to *MdMADS2.1*) was transcribed through all stages of flower development and, its transcription could be detected in the four floral organs. In our study, gene specific primers were used to identify the expression patterns of *MdMADS2.1* and *MdMADS2.2* in developing apple fruit cortex. Both *MdMADS2.1* and *MdMADS2.2* transcripts could be detected in 4 week-old, mature green and ripe apple fruits, with *MdMADS2.1* being more highly expressed than *MdMADS2.2* (data not shown). We also identified *MdMADS14* as a likely apple SHP orthologue based on sequence homology.

To test whether these genes were associated with fruit texture we adopted two strategies. The first involved evaluating the association between cortex texture and *MdMADS2.1, MdMADS2.2* and *MdMADS14* in the progeny obtained from the cross of 'Prima' and 'Fiesta' (Malieepard et al. 1998). This analysis located the *MdMADS2.1*

gene on LG14 and the *MdMADS2.2* and *MdMADS14* on, LG06 and LG09, respectively. However, none of these linkage groups coincided with the previously identified fruit firmness QTL regions (King et al. 2000; 2001). In a recent study, Liebhard et al. (2003) carried out QTL analyses for a number of physiological traits in apple using a progeny derived from 'Fiesta' and 'Discovery'. In this study, fruit flesh firmness QTLs were also identified on the LG06 and LG14. Recently, the SSR corresponding to *MdMads2* (U78948) was also mapped using the Fiesta x Discovery cross (Silfverberg-Dilworth et al. 2006). This marker was placed on LG14 at 45.6 cM, which shows close proximity to the QTL region (at 47 cM) identified by Liebhard et al. (2003). The QTL located on LG06 however, does not coincide with the location of the *MdMADS2.2*. Many fruit quality traits are affected by the inheritance of alleles at multiple loci and their possible interactions. The identification of these loci is still a key challenge of modern fruit genetic research. Conventional QTL analysis could not detect any association between the fruit flesh firmness and the *MdMADS2.1, MdMADS2.2* and *MdMADS14* genes in the 'Prima x Fiesta' population. This might be because the sizes of individual locus effects are modest, owing to genetic interactions between loci as suggested by Carlborg and Haley (2004) or the different alleles inherited from the two parents do not show significant functional difference. A range of approaches can be used to explore genetic networks for various QTLs (Kulwal et al. 2005; Li et al. 1997; Mei et al. 2003; Xing et al. 2002; Yu et al. 1997; Zhuang et al. 2002; Paran and Zamir 2003; Schadt et al. 2005; Brem et al. 2005). In this study we used mixed model ANOVA to test the association between the fruit flesh firmness and the genes of interest. This was carried out by controlling for the effect of previously identified QTLs (King et al. 2000; 2001). We obtained a significant association between the *MdMADS2.1* and the fruit flesh firmness

trait using mixed model ANOVA approach. Although there seems to be some correspondence in map location between the association and QTL studies, care has to be taken over the levels of significance associated with the latter, especially where QTL have been identified as putative. The association we obtained between the *MdMADS2.1* and the fruit flesh firmness should therefore be further tested by additional genetic as well as functional studies. The effect of the *MdMADS14*, a possible *SHP* orthologue, was close to significance. There are likely to be paralogues of *MdMADS14* in the *Malus* genome, some of which could be functionally redundant, as is the case for the *SHP* genes in Arabidopsis (Pinyopich et al. 2003).

We obtained further evidence for the association between *MdMADS2.1* and fruit flesh firmness in apple using a population-based association approach. Population-based genetic association studies have been extensively used to identify genes or genetic networks underlying complex phenotypes in medical genetics (Risch 2000; Hirschhorn and Daly 2005; Cordell and Clayton 2005). Such approaches have also been successful for the identification of genes/alleles affecting various plant traits (Neale and Savolainen 2004; Krutovsky and Neale 2005; Breseghello and Sorrells 2006; Andersen et al. 2005; Rafalski 2002; Gupta et al. 2005). In population-based association studies, spurious association could arise from subdivisions in the sample, referred to as population structure (Pritchard et al. 2000). Therefore, we investigated the genetic structure of the apple accessions used in this study with the Bayesian model-based clustering algorithm implemented in STRUCTURE (Pritchard et al. 2000). This analysis showed no evidence for significant population substructure that may affect our association analysis.

The data showed that specific alleles of *MdMADS2.1* were significantly associated with qualitative assessments of fruit texture. Although our studies link the *MdMADS2.1* locus with aspects of texture characteristics in apple, the biological basis of these effects is still to be resolved. Functional and genetic analysis in other species may, however, provide some important clues.

In tomato the *TDR4* gene shares significant homology with *MdMADS2* and *FUL*, it is up-regulated in wild type fruits and its expression is restricted in non-ripening mutants. This supports a role for these types of genes in the ripening of fleshy fruits. In Arabidopsis, the documented mode of action of *FUL* involves the repression or negative regulation of the genes controlling valve-margin development (Ferrandiz et al. 2000; Liljegren et al. 2004). Therefore, the *FUL* orthologues in other plant species with fleshy fruits could possibly play an analogous role to *FUL* in Arabidopsis and may repress cell separation and prevent a total loss of cell-to-cell adhesion, which could in turn lead to firm/crisp fruit phenotypes. Recently, Tani et al. 2007 compared the differential expression of peach *FUL*-like gene (PpMADS6), which is 83 % identical to *MdMads2.1*, between a split-pit sensitive and a split-pit resistant peach, and established an association between the expression level and the split-pit phenotype. In addition, Xu et al. 2008 found that constitutive expression of *PpMADS6* in Arabidopsis could prevent the pod shattering.

Previously, it has been show that there is likely to be a high level of synteny between the *Prunus* genome and the two component genomes of apple (Dirlewanger et al. 2004). A likely orthologue of the Malus *MdMADS2.1* (also known as *MdMADS2*) gene in *Prunus* is EST *PrpAP1* (BU039475) (Silva et al. 2005). This gene maps to linkage group ‘G05’of the ‘Texas’x’Earlygold’ *Prunus* reference map, only 1 cM away

from the molecular marker AG108A. Interestingly, in another study AG108A locus was found to harbour QTL acting on fruit size (Quilot et al. 2004), but no measurements of fruit firmness were reported. Fruit size could have an effect on texture characteristics depending on density of cells and cell wall material within the tissue.

Fruit ripening is the process of biochemical and physiological changes involving complex gene interactions. Functional analyses of genes in woody perennials such as apple are, however, very hard to pursue because of long generation times and space requirements for experimental material. Therefore, genetic studies are still the choice for the identification of genetic components underlying agronomic traits. The work described here demonstrates that information on the regulatory network controlling fruit development in the siliques of the model plant Arabidopsis can be used to nominate gene candidates involved in the texture of apple fruits.

Acknowledgements

This project was carried out with the financial support from the UK Department for Environment, Food and Rural Affairs, and the Commission of the European Communities (contract N° QLK5-CT-2002-01492), Directorate-General Research - Quality of Life and Management of Living Resources Programme. We also would like to thank Plant Research International, Wageningen, The Netherlands for making the reference population ‘Prima x Fiesta’ available to us.

References

Andersen JR, Schrag T, Melchinger AE, Zein I, Lubberstedt T (2005) Validation of Dwarf8 polymorphisms associated with flowering time in elite European inbred lines of maize (*Zea mays* L). Theor Appl Genet **111**: 206-17

Brem RB, Storey JD, Whittle J, Kruglyak L (2005) Genetic interactions between polymorphisms that affect gene expression in yeast. Nature **436**: 701-703

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L) cultivars. Genetics 172(2):1165-77

Carlborg O, Haley CS (2004) Epistasis: too often neglected in complex trait studies? Nat Rev Genet **5**: 618–625

Cordell HJ, Clayton DG (2005) Genetic Association Studies. The Lancet 366:1121-1131

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

Costa F, Van de Weg WE, Stella S, Dondini L, Pratesi D, Musacchi S, Sansavini S (2008) Map position and functional allelic diversity of *Md-Exp7*, a new putative

expansin gene associated with fruit softening in apple (*Malus x domestica* Borkh) and pear (*Pyrus communis*). Tree Genetics and Genomes 4:3 575-586

Dinneny JR, Yanofsky MF (2005) Drawing lines and borders: how the dehiscent fruit of Arabidopsis is patterned. Bioessays 27:42-49

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 USA **101**: 9891-9896

Doyle JJ, Doyle Jl (1990) Isolation of plant DNA from fresh tissue. Focus **12**: 13-15

Eriksson EM, Bovy A, Manning K, Harrison L, Andrews J, De Silva J, Tucker GA, Seymour GB (2004) Effect of the Colorless non-ripening mutation on cell wall biochemistry and gene expression during tomato fruit development and ripening. Plant Physiol **136**: 4184-97

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4):1567-87

Ferrandiz C, Pelaz S, Yanofsky MF (1999) Control of carpel and fruit development in Arabidopsis. Annu Rev Biochem **99**: 321–54

Ferrandiz C, Liljegren SJ, Yanofsky MF (2000) Negative regulation of the *SHATTERPROOF* genes by *FRUITFULL* during Arabidopsis fruit development. Science **289**: 436–38

Ferrandiz C (2002) Regulation of fruit dehiscence in Arabidopsis. J Exp Bot **53**: 2031-2038

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: Present status and future prospects. Plant Mol Biol **57**: 461-485

Hirschhorn JN, Daly MJ (2005) Genome–wide association studies for common diseases and complex traits. Nat Rev Genet **6**: 95–108

King GJ, Maliepaard C, Lynn JR, Alston FH, Durel CE, Evans KM, Griffon B, Laurens F, Manganaris AG, Schrevens E, Tartarini S, 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

King GJ, Lynn JR, Dover CJ, Evans KM (2001) Resolution of quantitative trait loci for mechanical measures accounting for genetic variation in fruit texture of apple (Malus pumila Mill). Theor Appl Genet **102**: 1227–1235

Knapp S (2002) Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the Solanaceae. J Exp Bot **53**: 2001-2022

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen **12**: 172–175

Krutovsky KV, Neale DB (2005) Nucleotide Diversity and Linkage Disequilibrium in Cold Hardiness and Wood Quality Related Candidate Genes in Douglas-fir. Genetics 171:2029-2041

Kulwal P, Kumar N, Kumar A, Gupta RK, Balyan HS, Gupta PK (2005) Gene networks in hexaploid wheat: interacting quantitative trait loci for grain protein content. Funct Integr Genomics **5**: 254–259

Li ZK, Pinson SRM, Park WD, Paterson AH, Stansel JW (1997) Epistasis for three yield components in rice *Oryza sativa* L. Genetics **145**: 453–465

Liebhard R, Kellerhals M, Pfammatter W, Jertmini M, Gessler C (2003) Mapping quantitative physiological traits in apple (*Malus x domestica* Borkh). Plant Mol Biol **52**: 511-26

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) *SHATTERPROOF* MADS-box genes control seed dispersal in Arabidopsis. Nature **404**: 766–770

Liljegren SJ, Roeder AH, Kempin SA, Gremski K, Ostergaard L, Guimil S, Reyes DK, Yanofsky MF (2004) Control of fruit patterning in Arabidopsis by INDEHISCENT. Cell **116**: 843-53

Litt A, Irish VF (2003) Duplication and diversification in the *APETALA1*/ *FRUITFULL* floral homeotic gene lineage: implications for the evolution of floral development. Genetics **165**: 821-33

Liu J (2002) Powermarker - A powerful software for marker data analysis. North Carolina State University Bioinformatics Research Center Raleigh North Carolina (www.powermarker.net)

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops. Trends Plant Sci 12(2):57-63

Maliepaard C, Alston FH, Van Arkel G, Brown LM, et al (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill) using multi-allelic markers. Theor Appl Genet **97**: 60–73

Maliepaard C, Sillanpää MJ, van Ooijen JW, Jansen RC, Arjas E (2001) Bayesian versus frequentist analysis of multiple quantitative trait loci with an application to an outbred apple cross. Theor Appl Genet 103:1243-1253

Mei HW, Luo LJ, Ying CS, Wang YP, Yu XQ, Guo LB, Paterson AH, Li ZK (2003) Gene actions of QTLs affecting several agronomic traits resolved in a recombinant inbred rice population and two testcross populations. Theor Appl Genet **107**: 89–101

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci **9**: 325–330

Paran I, Zamir D (2003) Quantitative traits in plants: beyond the QTL. Trends Genet **19**: 303–306

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature **424**: 85–88

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics **155**: 945-959

Quilot B, Wu BH, Kervella J, Genard M, Foulongne M, Moreau K (2004) QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild relative species *P davidiana.* Theor Appl Genet **109**: 884-97

Rafalski JA (2002) Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci **162**: 329-333

Risch NJ (2000) Searching for genetic determinants in the new millennium. Nature **405**: 847-856

Schadt EE, Lamb J, Yang X, et al (2005) An integrative genomics approach to infer causal associations between gene expression and disease. Nat Genet **37**: 710-717

Seymour GB, Manning K, Eriksson EM, Popovich AH, King GJ (2002) Genetic identification and genomic organization of factors affecting fruit texture. J Exp Bot 53(377):2065-71

Seymour GB, Poole M, Manning K, King GJ (2008) Genetics and Epigenetics of Fruit Development and Ripening. Current Opinion in Plant Biology **11** 58-63

Silfverberg-Dilworth E, Matasci CL, van de Weg WE et al (2006) Microsatellite markers spanning the apple (*Malus x domestica* Borkh) genome. Tree Genetics and Genomes 2:202-224

Silva C, Garcia-Mas J, Sanchez AM, Arus P, Oliveira MM (2005) Looking into flowering time in almond (*Prunus dulcis* (Mill) D A Webb): the candidate gene approach. Theor Appl Genet **110**: 959-68

Smith MWG (1971) National Apple Register of the United Kingdom Pinner: Ministry of Agriculture Fisheries and Food.

Sung SK, Yu GH, An G (1999) Characterization of *MdMADS2*, a member of the SQUAMOSA subfamily of genes, in apple. Plant Physiol **120**: 969–978

Tani E, Polidoros AN, Tsaftaris AS (2007) Characterization and expression analysis of *FRUITFULL*- and *SHATTERPROOF*-like genes from peach (*Prunus persica*) and their

role in split-pit formation. Tree Physiol 27: 649-659

Van Ooijen JW, Voorrips RE (2001) JoinMap® 30 Software for the calculation of genetic linkage maps. Plant Research International ,Wageningen ,The Netherlands.

Vinatzer BA, Zhang HB, Sansavini S (1998) Construction and characterization of a Bacterial Artificial Chromosome library of apple. Theor Appl Genet **97**: 1183-1190

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, et al (2002) A MADS box gene necessary for fruit ripening at the tomato ripening inhibitor (*Rin*) locus. Science **296**: 343–346

Xing YZ, Tan YF, Hua JP, Sun XL, Xu CG, Zhang Q (2002) Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. Theor Appl Genet **105**: 248–257

Xu Y, Zhang L, RongCai MA (2008) Functional characterization and mapping of two *MADS* box genes from peach (*Prunus persica*). Chinese Science Bulletin 53: 853-859

Yu SB, Li JX, Xu CG, Tan YF, Gao YJ, Li XH, Zhang Q, Saghai Maroof MA (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc Natl Acad Sci USA 94: 9226–9231

Zhuang JY, Fan YY, Rao ZM, Wu JL, Xia YW, Zheng KL (2002) Analysis on additive effects and additive-by-additive epistatic effects of QTLs for yield traits in a recombinant inbred line population of rice. Theor Appl Genet **105**: 1137–1145

Table 1 Association analysis of *MdMADS2.1* and *MdMADS14* genes with the apple fruit flesh firmness in the reference population 'Prima x Fiesta' using mixed model analysis of variance

Factors	Effect	Degree of freedom	SS	MS	*F*	*P*-value
MdMADS14	Fixed	1	0.96	0.96	3.33	0.0716
MdMADS2.1	Fixed	1	2.80	2.80	9.68	0.0026
ACO1	Random	1	5.060	5.06	17.50	<0.0001
OPAC-110-780	Random	1	1.14	1.14	3.94	0.0504
OPC-09-0900	Random	1	4.814	4.81	16.65	0.0001
LY26a	Random	2	5.46	2.73	9.45	0.0002
Error		82	23.71	0.29		
R^2	0.50					

Table 2 Comparison of allele and genotype distributions between the soft and firm apple accessions using the chi-square method for the determination of population-based association between *MdMADS2.1* gene and sensory fruit firmness

Marker	Allele ChiSquare	Genotype ChiSquare	Allele df	Genotype df	Allele *P*-Value	Genotype *P*-Value
MdMADS2.1 SSR 3'UTR	34	57	9	32	<0.001	<0.01

Legend to Figures

Fig. 1 (a) Alignment of the predicted amino acid sequences of MdMADS2 (accession no U78948), MdMADS2.1 (accession no DQ205652) and MdMADS2.2 (accession no DQ205651). The asterisks represent amino acid residues identical to the corresponding residues in MdMADS2 **(b)** Schematic diagram of the predicted exons and introns of the *MdMADS2.1* and *MdMADS2.2* genes Solid boxes indicate the coding regions The lines connecting the solid boxes represent the seven introns and the 5'- and 3'-flanking regions. Nucleotide differences between *MdMADS2.1* and *MdMADS2.2* genes are shown in panels in the lower part of the figure and the different nucleotide sequences were highlighted. Positions of nucleotide deletions are shown as (a) totalling over 1 kb, (b) 1 bp, (c) 7 bp, (d) 12 bp, (e) 9 bp, (f) 22 bp and (g) 190 bp

(a)

```
MdMADS2   MGRGRVQLKRIENKINRQVTFSKRRSGLMKKAHEISVLCDAEVALIIFST
MdMADS2.1 **************************************************
MdMADS2.2 ************************T***L********************

MdMADS2   KGKLFEYSNDSCMERILERYERYSYTERQLLANDNESTGSWTLEHAKLKA
MdMADS2.1 **************************************************
MdMADS2.2 ************************HA**********N**********

MdMADS2   RVEVLQRNQRHYMGEDLQSLSLKELQNLEQQLDSALKHIRSRKNQVMYES
MdMADS2.1 **************************************************
MdMADS2.2 ******************I**F**************RR************

MdMADS2   ISELQKKDKALQEQNNLLAKKVKEKENAVAQQAQLEHVQEQRLNSSSSLL
MdMADS2.1 **************************************************
MdMADS2.2 ********************N*****K**TS****D*A*K*S*D***T**

MdMADS2   PRALQSLNFGSGSNY-QAIRSSEGIPGDNQQYGDETPTPHRPNMLLPAWI
MdMADS2.1 **********************************************M
MdMADS2.2 *QE**Y***-*R***H******N**S*****D****A******T***P*M

MdMADS2   VRHLNE
MdMADS2.1 L*****
MdMADS2.2 L*****
```

(b)

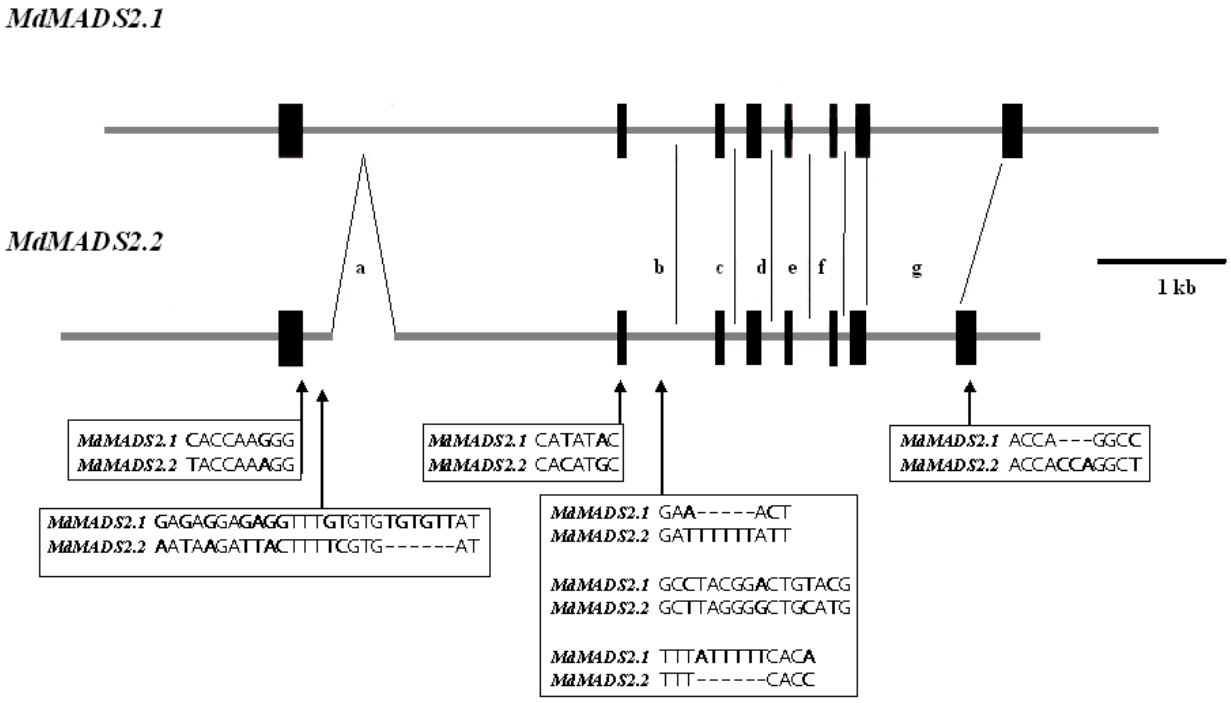
MdMADS2.1
MdMADS2.2
a
b
c
d
e
f
g
1 kb
MdMADS2.1 CACCAAGGG
MdMADS2.2 TACCAAAGG
MdMADS2.1 GAGAGGAGAGGTTTGTGTGTGTTAT
MdMADS2.2 AATAAGATTACTTTTCGTG------AT
MdMADS2.1 CATATAC
MdMADS2.2 CACATGC
MdMADS2.1 GAA-----ACT
MdMADS2.2 GATTTTTTATT
MdMADS2.1 GCCTACGGACTGTACG
MdMADS2.2 GCTTAGGGGCTGCATG
MdMADS2.1 TTTATTTTTCACA
MdMADS2.2 TTT------CACC
MdMADS2.1 ACCA---GGCC
MdMADS2.2 ACCACCAGGCT

Fig. 2 Genetic map positions of *MdMADS2.1*, *MdMADS2.2* and *MdMADS14*

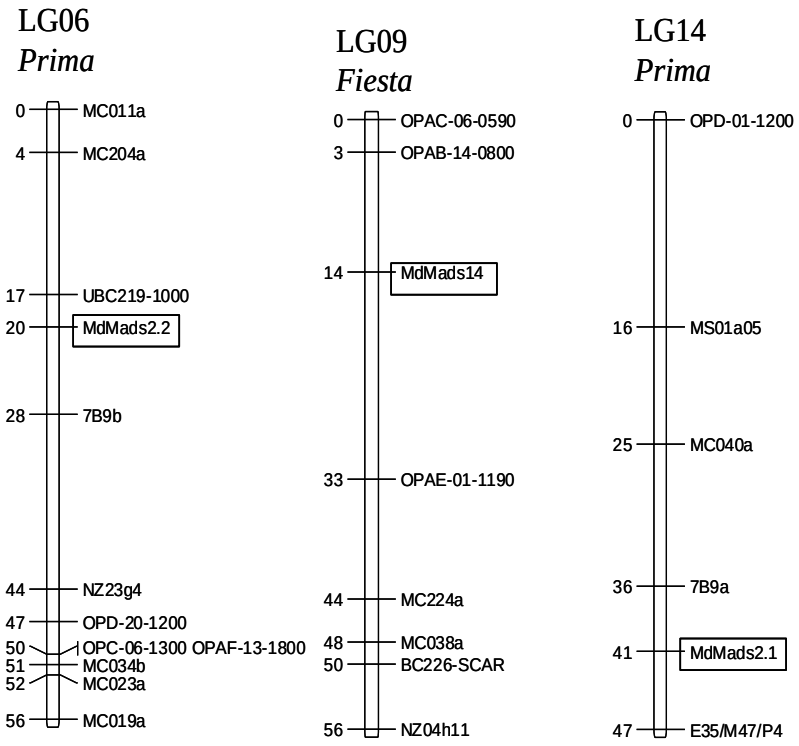

Fig. 3 Allele frequency distributions for $(TA)_n$ repeat polymorphism in the 3'UTR of the *MdMADS2.1*gene in 88 Soft and 80 Firm apple accessions