‘GGP is a key regulator of fruit vitamin C concentrations’

Running head: ‘GGP is a key regulator of fruit vitamin C concentrations’

Corresponding Author:

Mark W DAVEY
Laboratory of Fruit Breeding and Biotechnology
Department of Biosystems
Faculty of Bioscience Engineering
KU Leuven,
Willem De Croylaan 42,
B-3001, Heverlee,
Belgium

Tel. +32 – 16 322414
Fax. +32 – 16 322966
Email: mark.davey@biw.kuleuven.be

Research Category: Genetics, Genomics and Molecular Evolution
'GGP is a key regulator of fruit vitamin C concentrations'

Title:

‘Allelic variation in paralogues of GDP-L-galactose phosphorylase is a major determinant of vitamin C concentrations in apple fruit’

Authors:

1MELLIDOU, Ifigeneia
2CHAGNÉ, David
3LAING, William A
1KEULEMANS, Johan
1DAVEY, Mark W *

Addresses:

1Laboratory for Fruit Breeding and Biotechnology, Department Biosystems, Faculty of Bioscience Engineering, KU Leuven, B-3001, Heverlee, Belgium
2The New Zealand Institute for Plant & Food Research Limited (Plant & Food Research), Palmerston North Research Centre, Palmerston North 4442, New Zealand
3Plant & Food Research, Mount Albert Research Centre, Auckland 1142, New Zealand

* - corresponding author

Corresponding author; e-mail: mark.davey@biw.kuleuven.be
ABSTRACT

To identify the genetic factors underlying the regulation of fruit vitamin C (l-ascorbic acid, AsA) concentrations, quantitative trait loci (QTL) studies were carried out in an F1 progeny derived from a cross between the apple (Malus x domestica Borkh.) cultivars ‘Telamon’ and ‘Braeburn’ over three years. QTL were identified for AsA, glutathione (GSH), total antioxidant activity in both flesh and skin tissues, and for various quality traits, including flesh browning. Four regions on chromosomes 10, 11, 16 and 17 contained stable fruit AsA-QTL clusters. Mapping of AsA metabolic genes identified co-locations between orthologues of GDP-l-galactose phosphorylase (GGP), dehydroascorbate reductase (DHAR) and nucleobase-ascorbate transporter (NAT) within these QTL clusters. Of particular interest are the three paralogues of MdGGP, which all co-located within AsA-QTL clusters. Allelic variants of MdGGP1 and MdGGP3 derived from the ‘Braeburn’ parent were also consistently associated with higher fruit total AsA concentrations both within the mapping population (up to 10-fold) and across a range of commercial apple germplasm (up to 6-fold). Striking differences in the expression of the ‘Braeburn’ MdGGP1 allele between fruit from high- and low-AsA genotypes clearly indicate a key role for MdGGP1 in the regulation of fruit AsA concentrations and this MdGGP allele-specific Single Nucleotide Polymorphism (SNP) marker represents an excellent candidate for directed breeding for enhanced fruit AsA concentrations. Interestingly, co-locations were also found between MdDHAR3-3 and a stable QTL for browning in the ‘Telamon’ parent, highlighting links between the redox status of the AsA pool and susceptibility to flesh browning.

(247 words)

Keywords: l-ascorbic acid, dehydroascorbate reductase, GDP-l-galactose phosphorylase, genetic control, Malus x domestica, marker assisted breeding, nutritional enhancement, QTL, browning, SNP markers
INTRODUCTION

In plants, L-ascorbic acid (vitamin C, AsA) is essential for the detoxification of reactive oxygen species (ROS) produced under stress or following exposure to pathogens. In addition to these antioxidant functions, AsA has been shown to be involved in a range of important cellular processes, including plant development and hormone signalling, cell cycle, cell expansion, senescence, and as a cofactor for a number of important enzymes (for reviews, see Davey et al., 2000; Smirnoff et al., 2001; Noctor, 2006). Fruit AsA concentrations have also been correlated with the maintenance of quality during postharvest storage (Davey and Keulemans, 2004; Davey et al., 2007), and have been linked to susceptibility to internal browning in both apples (Davey et al., 2006; Davey and Keulemans, 2009) and pears (Veltman et al., 1999; Franck et al., 2003). Finally, AsA is clearly an essential dietary component for humans, with a protective role proposed for many disorders and diseases (Diplock et al., 1998). Given its importance for all metabolically active tissues, there is widespread interest in unravelling the mechanisms underlying the genetic control of AsA concentrations in fruits, as well as in how AsA interacts with other plant antioxidant pools.

The concentration of AsA will be determined by the net rates of biosynthesis, recycling, degradation, and/or inter- and intra-cellular transport, but the relative contribution of these various processes depends on several factors including genetics, tissue type (Bulley et al., 2009), developmental stage (Hancock et al., 2007; Bulley et al., 2009; Ioannidi et al., 2009), and light intensity (Yabuta et al., 2007; Gautier et al., 2009). The biosynthesis of AsA proceeds via L-galactose (Wheeler et al., 1998), although conclusive evidence for all steps has only relatively recently become available (Conklin et al., 2006; Laing et al., 2007). Alternative biosynthetic routes involving uronic acids (Davey et al., 1999; Agius et al., 2003), L-gulose (Wolucka and van Montagu, 2003), or myo-inositol (Lorence et al., 2004) have been proposed in several plant species including apple (Davey et al., 2004, Razavi et al., 2005) (Fig. 1), but their physiological relevance and contribution to the AsA pool is still far from clear in most plant species, with the possible exception of strawberry (Agius et al., 2003; Zorilla et al., 2011; Cruz-Rus et al., 2011).

As an antioxidant, AsA is able to accept electrons from a wide range of radical substrates, and in this process becomes oxidised first to monodehydroascorbate (MDHA) and then to dehydroascorbate (DHA). These oxidised forms of AsA can be regenerated by the ascorbate-glutathione cycle, so that glutathione (GSH) and the activities of glutathione reductase (GR), dehydroascorbate reductase (DHAR) and monodehydroascorbate reductase (MDHAR) maintain the size and redox status of the AsA pool (Noctor and Foyer, 1998) (Fig. 1). Indeed, overexpression of an Arabidopsis MDHAR (Eltayeb et al.,
GGP is a key regulator of fruit vitamin C concentrations’

...2007), and a wheat DHAR (Chen et al., 2003) have both been shown to increase foliar AsA concentrations in tobacco. MDHAR activity has also been positively correlated with both AsA and fruit firmness in tomatoes after chilling stress (Stevens et al., 2008).

...Tissue AsA concentrations can also be maintained by intercellular transport, and there is evidence for the long distance transport of AsA via the phloem from source (leaf) to sink (fruit) tissues (Franceschi and Tarlyn, 2002; Hancock et al., 2003). In apple, fruit AsA concentrations have been suggested to be partly dependent on the translocation of AsA from leaves (Li et al., 2009), but in blackcurrant, others concluded that the contribution of phloem AsA-transport to fruit AsA concentrations was negligible (Hancock et al., 2007). While the actual mechanisms of long distance transport of AsA have not been fully determined, attention has focused on the large family of Nucleobase-Ascorbate Transporters (NAT) (de Koning and Diellinas, 2000), and NAT homologues have been found to be highly expressed in vascular tissues (Maurino et al., 2006).

...Genes involved in several of these mechanisms have been proposed to be key regulators of fruit AsA concentrations, including GDP-L-galactose phosphorylase (GPP or VTC2) in kiwifruit (Bulley et al., 2009; 2012), and GDP-mannose-3,5-epimerase (GME; Gilbert et al., 2009), L-galactose-1-phosphate-phosphatase (GPP or VTC4; Ioannidi et al., 2009) and MDHAR (Stevens et al., 2007) in tomato. However, apart from GGP (Bulley et al., 2012), overexpression of these structural genes has to date had limited success in altering the fruit AsA pool (Agius et al., 2003; Bulley et al., 2009; Haroldsen et al., 2011).

...In this work, we set out to identify potential genetic determinants of fruit AsA concentrations in apple fruit using a combination of molecular and genomic approaches. Initial Quantitative Trait Loci (QTL) analyses of AsA concentrations (Davey et al., 2006) have been expanded to identify QTL for other antioxidants and fruit quality traits over three years, including results in one year comparing the concentrations of AsA in fruit and leaves. Alignments of the Malus x domestica orthologues of genes involved in AsA biosynthesis, turnover and transport, against the whole genome sequence of ‘Golden Delicious’ (Velasco et al., 2010), allowed us to identify candidate genes (CGs) co-locating with stable QTL clusters. Using Next Generation Sequencing (NGS, RNA-Seq) data, polymorphic Single Nucleotide Polymorphism (SNP) based markers were developed for these co-locating CGs, and their positions on individual linkage groups confirmed by linkage mapping in our mapping population. Finally, associations between allelic variants of these CGs, and their expression levels in cultivars with contrasting AsA concentrations, allowed us to develop allele-specific markers associated with high fruit AsA concentrations.
‘GGP is a key regulator of fruit vitamin C concentrations’
GGP is a key regulator of fruit vitamin C concentrations

RESULTS

Population variation in tissue ascorbic acid and glutathione concentrations

Fruit AsA - totAsA (total ascorbic acid) and GSH - totGSH (total glutathione) concentrations showed a high degree of variability in the ‘Telamon x Braeburn’ (‘TxB’) mapping population, with mean values for flesh totAsA and totGSH on average varying 10- and 7-fold, across the population over the three years of measurements, respectively (Table 1). Total antioxidant activity was only measured in 2006, with results indicating that the contribution of totAsA to TAA ranged from 4 to 25% in flesh and from 6 to 50% in skin tissues. Both AsA and totAsA (AsA + DHA), and GSH and totGSH (GSH + GSSG) concentrations in flesh and skin tissues showed normal distribution across progeny of the ‘TxB’ population over the years, but concentrations of the oxidized forms (DHA and GSSG) displayed a statistically normal distribution in only one of the three years of measurement (2006).

Fruit AsA, totAsA, and GSH, totGSH concentrations in individuals of the ‘TxB’ progeny were positively correlated (P<0.0001) between flesh and skin tissues, such that fruit with high skin AsA and GSH concentrations also contained high flesh AsA and GSH concentrations (Supplemental Table S1). Ascorbic acid and totAsA concentrations were also significantly correlated across the years (Supplemental Table S2; P<0.0001), however correlations for DHA, GSH, GSSG and totGSH concentrations were lower or not significant at all. Comparing totAsA concentrations across the years indicated that in 2009, values were 27% and 42% higher in fruit flesh, and 40% and 60% higher in fruit skin than the concentrations determined in 2005 and 2006, respectively. The % proportion of DHA in the totAsA pool which is an indicator of the degree of oxidative stress experienced by the tissue, was found to be significantly higher in 2006 compared to 2005 or 2009. Finally, fruit skin tissues contained approximately 40% more totGSH than flesh tissues and values in both tissues were around 50% higher in 2009 than in 2006.

These year-to-year differences in fruit AsA and GSH traits are probably related to significant differences in the climatic conditions (Supplemental Fig. S1). For example, the length of the harvest season ranged from 32 days in 2006, to 59 days in 2005, even though the first day of the harvesting period was the same each year. The highest mean population AsA and totAsA concentrations were measured in 2009 when both average daily temperature and hours of sun radiation were higher than those of the other years, suggesting an adaption to the high-light and high-temperature experienced in this year. Similarly, the lowest AsA and totAsA concentrations were measured in 2006, when hours of sunlight were lower than in either 2005 or 2009. A comparison of leaf and fruit totAsA and totGSH concentrations....
‘GGP is a key regulator of fruit vitamin C concentrations’

in 2009 indicated that apple leaves contained approximately 65-fold more totAsA and 12-fold more 
totGSH than fruit flesh tissues (Supplemental Table S3). No significant correlations were found between 
fruit and foliar AsA and totAsA concentrations in fruit of progeny from ‘TxB’, but there were significant 
correlations between foliar and both fruit flesh and skin GSH, totGSH, GSSG concentrations and % 
GSSG (Supplemental Table S4).

Population variation in flesh browning and fruit quality traits

The susceptibility of fruit flesh to browning was assessed at harvest in 2006 and 2009 by measuring the 
time taken for cut surfaces to brown (‘Br_time’), and the final colour of the affected tissue (‘Br_colour’). 
‘Br_time’ ranged from 15 to 265 minutes and was negatively correlated with the % DHA values of flesh 
tissues (Supplemental Table S5), but not with the flesh total AsA values. ‘Br_colour’, was positively 
correlated with the flesh % DHA. A strong correlation (P<0.0001) was also found between flesh AsA- 
totAsA concentrations and firmness, and to a lesser extent with fruit % dry weight (DW) and soluble 
solids content [measured as °Brix (P<0.01)] (Supplemental Table S5).

QTL mapping

We identified a total of 27 fruit AsA- and GSH-QTL which were stable for at least two of the three 
measurement years. These were clustered into four main genomic regions on linkage groups (LGs) 10, 11, 
16 and 17 (Fig. 2). The total population variability explained by these AsA-QTL (QTL for AsA, totAsA 
and DHA) ranged from 10.7% to 59.5% per trait, but was much lower for GSH-QTL (QTL for GSH, 
totGSH and GSSG) where population variability ranged from between 9.5 and 13.1% per trait 
(Supplemental Table S6).

QTL for fruit and leaf AsA concentrations

Twenty flesh AsA-QTL were identified in the ‘Telamon’ parent and 17 in ‘Braeburn’. These QTL 
individually explained from 7.6 to 30% of the total population variance (Supplemental Table S6). 
Clusters of stable (present in two or more years) and significant (LOD values > 3.5) AsA-QTL were 
identified at the bottom of LG 11 and in the middle of LG 10 in both parental maps (Table 2). Additional 
AsA-QTL on LGs 3, 6, and 16 were not stable across all years, while parent-specific QTL were located 
on LGs 3 and 17. In the majority of cases, QTL for totAsA co-located with QTL for flesh AsA
concentrations in both parents (Fig. 2; Supplemental Table S6). Of particular interest was the QTL cluster for flesh AsA located on LG 11, as these were stable over all three years and explained a relatively high proportion (up to 27.5%) of the total population variability.

Stable clusters of skin AsA-QTL were detected on LGs 9, 10 and 16, which individually accounted for up to 17.3% of population variability, while additional, non year-stable QTL were found on LGs 2, 3, 11 and 17 (Fig. 2; Supplemental Table S6). The skin AsA-QTL on LGs 11, 16 and 17 co-located with AsA-QTL of flesh, with only the QTL cluster on LG 9 specific for fruit skin AsA-totAsA concentrations.

QTL analysis of leaf AsA, totAsA and DHA concentrations (2009 only), identified two QTL clusters on LG 6 and 15 (Table 3), but only the QTL on LG 6 co-located with that of a fruit AsA-QTL derived from the ‘Telamon’ parent.

QTL for fruit and leaf GSH concentrations

From measurements in in 2006 and 2009, we identified nine skin, and nine flesh GSH-QTL, distributed across nine LGs (Fig. 2; Supplemental Table S6). LOD scores for these QTL ranged from 3.0 to 3.4, and were generally lower than those observed for AsA-QTL. Only one region in the middle of LG 16 contained QTL that were stable across parents and tissues (fruit flesh, fruit skin and leaf). In 2009, leaf GSH concentrations were measured, leading to the identification of QTL on LGs 6, 12, and 16 (Fig. 2; Table 3), with the QTL for leaf GSSG on LG 16 co-locating with fruit AsA-QTL in both parents (Fig. 2).

QTL for fruit total antioxidant activity (TAA)

In ‘Telamon’, two significant TAA-QTL for flesh were detected on LGs 6 and 12, accounting for a total of 23.1% of the population variability (Supplemental Table S6). In ‘Braeburn’, a major TAA-QTL on LG 11 (LOD 3.9, population variability 20.6%), and two minor TAA-QTL on LGs 6 and 16 were identified in flesh tissues, all partially overlapping with QTL for flesh AsA-totAsA.

QTL for flesh browning and fruit quality traits

QTL for susceptibility to flesh browning (‘Br_time’ and ‘Br_colour’) were located on LGs 3, 10, 16 and 17 (Fig. 2; Supplemental Table S7). Of these, the QTL for ‘Br_colour’ on LG 16 of ‘Braeburn’ and on LG 17 of ‘Telamon’ were stable over both years of measurements and co-located with AsA-QTL on these LGs (Fig. 2). On LG 10 (both parents) and on LG 17 (‘Telamon’), QTL for ‘Br_time’ mapped to the same position as QTL for AsA, totAsA and DHA concentrations of flesh tissues. Interestingly, stable QTL for flesh browning (‘Br_time’, ‘Br_colour’) at the bottom of ‘Telamon’ LG 17 co-located with the stable
QTL for flesh DHA concentrations, but no comparable QTL were detected in the ‘high fruit quality’ ‘Braeburn’ parent. A large cluster of fruit AsA-QTL in the middle of LG 10 in both parents co-located with a range of previously identified stable QTL for fruit traits including fresh weight, and flesh soluble solids content and firmness in this population (Supplemental Table S7; Kenis et al., 2008).

**Mapping of candidate gene orthologues and fruit QTL clusters**

From the literature, we identified 22 structural genes involved in AsA biosynthesis, recycling, degradation and transport. Using the apple reference consensus coding sequence (CDS) set (http://genomics.research.iasma.it/), 99 *Malus x domestica* orthologues of these genes were identified, based on similarity to Arabidopsis sequences and by phylogenetic analysis using published sequences from several additional plant species (Supplemental Table S8). Following an in silico analysis of their positions in the apple genome (Supplemental Table S8), SNP-based markers (Supplemental Table S9) were developed for the subset of CGs that co-located within our QTL using RNA-Seq data of both parents, allowing these CGs to be mapped onto our ‘TxB’ genetic linkage maps (Supplemental Fig. S8).

A total of 10 AsA CGs were mapped onto the ‘Telamon’ map, and 13 onto the ‘Braeburn’ map. Generally, these CGs mapped to the expected positions predicted from the ‘Golden Delicious’ genome assembly (Supplemental Fig. S2). The only exception was for *MdDHAR3-1* (MDP0000240690; LG 9), where the marker mapped to the homeologous region of LG 17, which contains two other paralogous copies of *MdDHAR3* (*MdDHAR3-2*, MDP0000175246 and *MdDHAR3-3*, MDP0000156763).

From the L-galactose biosynthetic pathway, one orthologue of *GGP* (MDP0000172222, *VTC2*, *MdGGP1*) with 71% protein sequence identity to Arabidopsis At4g26850.1, mapped to LG 11 of ‘Telamon’ within the cluster of stable flesh AsA-QTL in both parents (Fig.2). A clone of this gene from *Malus sylvestris* (GB CN915822) has been shown to have GGP activity when expressed in *Escherichia coli* (Supplemental Table S10). A paralogue of *MdGGP1*, *MdGPP2* (MDP0000288088; 69.5% identity to At4g26850.1 and 89.6% sequence identity to *MdGPP1*), mapped to LG 3 of ‘Braeburn’ within the flesh AsA-QTL cluster in ‘Telamon’ that was only detected in 2006. Again, a homologous gene from *Malus x domestica* (GB CN939721) has been demonstrated to be active (Supplemental Table S10). A third more distantly related orthologue of Arabidopsis *GPP*, *MdGPP3* (MDP0000191488; 30.1% identity to At4g26850.1 and 34% identity to *MdGGP1*), mapped to LG 10 of ‘Telamon’, and also co-located within a stable AsA-QTL. The markers developed for both *MdGGP1* (2006 and 2009) and *MdGPP3* (2005) were also the best markers describing the respective QTL (Table 2). None of the other genes involved in
'GGP is a key regulator of fruit vitamin C concentrations'

the L-galactose biosynthetic pathway co-located with any of our AsA-QTL, apart from a putative GDP-
mannose pyrophosphorylase (GMP, or VTC1; MDP0000323050, *MdGMP2-1*) with 83.5% identity to
At1g74910 which co-located with the non year-stable flesh AsA-QTL on LG 6 (Fig. 2).

From the alternative AsA biosynthetic pathways, a marker for a putative D-galacturonic acid
reductase (*MdGalUR1*, MDP0000309417; 57.5% identity with At2g37790) mapped to LG 12 of
‘Braeburn’, but did not co-locate with any QTL (Fig. 2). However, two *MdGalUR2* sequences
(MDP0000135496 and MDP0000857724) mapped in silico within the flesh AsA-QTL on LG 16 (Fig. 2;
Supplemental Table S8). It was not possible to develop markers for these others however as expression
levels were too low to identify unique SNPs from our RNA-Seq data.

We identified six *DHAR* and six *MDHAR* apple orthologues involved in the recycling of oxidised
DHA (Supplemental Table S8). Of these, *MdDHAR3-3* (65% identity with *AtDHAR3*) mapped to the
bottom of LG 17 in ‘Telamon’, and co-locates with a stable QTL for flesh DHA concentration (2005 and
2009 data) and a minor QTL for flesh AsA concentration (2005) (Fig. 2). The other paralogues of
*MdDHAR3* located in silico to LG 9 (*MdDHAR3-1*, MDP0000240690 and *MdDHAR3-4*,
MDP0000530903), also within the QTL intervals for skin AsA concentrations (2005 data). There were no
co-locations however between any of the *MdMDHAR*s and antioxidant QTL.

Twenty-one apple orthologues of the NAT protein family were identified in the reference CDS
set. Four of these are highly similar to *AtNAT7* and *AtNAT12*, which are the two members considered to
be specifically involved in AsA transport, and are located on LGs 11, 13, 15 and 16 (Supplemental Table
S8). *MdNAT7-2* (MDP0000320308; 58% identity with *AtNAT7*) mapped to LG 16 of ‘Braeburn’, co-
locating with the significant QTL for flesh AsA-totAsA concentrations in both parents (Fig. 2).

Finally, in silico alignment of genes of GSH metabolism (*GR* and *GS*; Supplemental Table S8)
identified only one potentially interesting co-location on LG 4 between *MdGS2-2* (MDP0000188669;
61% identity to At5g27380) and GSH-QTL for both flesh and skin tissues detected in ‘Braeburn’.

**Sequence variations in candidate regulatory genes for fruit AsA concentrations**

High throughput RNA sequencing (RNA-Seq) was used to detect polymorphisms in the transcript
sequences of AsA CGs, and to identify allelic SNPs linked to enhanced fruit AsA-totAsA concentrations.
Aligning the RNA-Seq reads to the ‘Golden Delicious’ reference consensus CDS set
(http://genomics.research.iasma.it/) resulted in a mean read coverage for *MdGGP1* (LG 11) of 164.6- and
123.9-fold for ‘Braeburn’ and ‘Telamon’ respectively, indicating high transcript expression and good read depth. A total of 10 different SNPs along the 1704 bp *MdGGP1* ‘Golden Delicious’ reference CDS were identified (Supplemental Fig. S3; Supplemental Table S11). This corresponds to one SNP every 170 bp, which is consistent with previous reports in apple cDNA sequences (one SNP every 149 bp; Chagné et al., 2008). These alignments were used to correct the *MdGPP1* reference gene model which was found to contain an incorrectly predicted start codon and two extra codons (Supplemental Fig. S3, S4). As a result the majority of the SNPs in *MdGGP1* shift into the 5’ UTR, and only two SNPs are found within the CDS, at reference positions 414 and 885 bp. Neither SNP leads to an amino acid change, but SNP_885 is unique to ‘Telamon’ and was developed into a diagnostic SNP marker for mapping and genotyping.

For *MdGGP2* (LG 3), mean read coverage following alignment to the ‘Golden Delicious’ reference consensus CDS set was 122.5-fold and 111.2-fold for ‘Braeburn’ and ‘Telamon’ respectively, again indicating high expression and good read depth. A total of eight SNPs were identified in the *MdGGP2* CDS of which three were heterozygous in ‘Braeburn’ and six in ‘Telamon’ (Supplemental Fig. S5; Supplemental Table S12). Both ‘Telamon’ and ‘Braeburn’ possess non-synonymous SNPs at position 74 and 169 bp. However, as both SNPs are present in both parents, they are unlikely to be responsible for the parent-specific allelic differences observed. Both also have a homozygous, but different SNP at reference position 564, which was used for mapping and genotyping.

Mean read coverage of *MdGGP3* following alignment to the ‘Golden Delicious’ CDS was 34.8-fold and 31.3-fold for ‘Braeburn’ and ‘Telamon’, respectively. Both the ‘Telamon’ and ‘Braeburn’ CDS were noticeably more divergent from the reference CDS than *MdGGP1* and *MdGGP2*, and we identified 79 SNPs in ‘Braeburn’ and 75 SNPs in ‘Telamon’ (Supplemental Fig. S6). A unique non-synonymous SNP in ‘Telamon’ at reference position 211 was developed for mapping. Both parents contain non-synonymous SNPs that could influence MdGGP3 protein function, and both also possess a homozygous 33bp deletion at reference position 272 at the beginning of intron 4, leading to the loss of 11 amino acids. In addition, in ‘Braeburn’ a heterozygous single nucleotide deletion at reference position 986 leads to a frame shift and a different and truncated N-terminal amino acid sequence (Supplemental Fig. S7). This suggests that ‘Braeburn’ contains an *MdGGP3* allele that codes for a truncated and possibly dysfunctional protein. Alignment of ‘Braeburn’ *de novo* contigs supports this conclusion, with fragments of both ‘Braeburn’ alleles apparently being present.

The mean read coverage for *MdDHAR3-1* (LG9) was 14.0-fold and 20.0-fold for ‘Braeburn’ and ‘Telamon’ respectively; however, these alignments contained a large region of non-specifically matched reads, so that only a single SNP could be reliably identified. This SNP-based marker mapped to the
homeologous region on LG17, which contains both \textit{MdDHAR3-2} and \textit{MdDHAR3-3} (MDP0000175246 and MDP0000156763, LG 17). However, the read coverage for these latter two sequences was too low to allow additional \textit{MdDHAR3}-specific markers to be developed.

For \textit{MdNAT7-2} (LG 16), the mean read coverage was 33.1 and 34.9-fold for ‘Braeburn’ and ‘Telamon’, respectively. However, coverage was not uniform and it was not possible to generate a full length consensus sequence in either case. In addition, the consensus read mapping sequences from both parents showed a considerable degree of polymorphism from the ‘Golden Delicious’ reference CDS, with 252 variants detected in ‘Braeburn’ and 151 in ‘Telamon’.

\textbf{Allele association studies}

Orthologous CG sequences which mapped within AsA-QTL, highlighted several genes with a potential role in regulating fruit AsA concentrations. Of particular interest are the paralogous copies of \textit{MdGGP}, as all three were mapped within significant flesh AsA-QTL, and two of these QTL were further stable for at least two years. In addition, \textit{MdNAT7-2} co-located with QTL for flesh antioxidant concentrations on LG 16. To further investigate the involvement of these four genes in the control of apple AsA concentrations, correlations between the allele SNP variants were examined in the ‘TxB’ progeny, as well as in a set of commercial apple cultivars with a wide range of totAsA concentrations.

TotAsA concentrations of flesh tissues in progeny of the ‘TxB’ population were significantly correlated with SNP variants for \textit{MdGGP1} (all 3 years), \textit{MdGGP2} (1 year), \textit{MdGGP3} (2 years) and \textit{MdNAT7-2} (2 years) (Table 4). Specifically, genotypes homozygous for the ‘T’ allele (T/T genotypes) of \textit{MdGGP1} at position 885 had statistically significant higher flesh totAsA concentrations than the ‘T/C’ genotypes over the three years of measurement. Similarly, ‘G/A’ genotypes for \textit{MdGGP2} (SNP564), or the ‘T/T’ genotypes for \textit{MdGGP3} (SNP211) also had higher flesh totAsA concentration than ‘G/G’ or ‘T/C’, respectively, but only in one or two years of measurements (Table 4). For \textit{MdNAT7-2} (SNP92), the ‘T/G’ genotypes had also higher fruit totAsA concentrations in 2006 and 2009, compared to ‘G/G’ genotypes. Comparing the flesh totAsA concentrations of progeny with the ‘best’ and ‘worst’ SNP variant combinations for these genes, showed that the ‘best’ SNP allelic variants resulted in fruit mean totAsA concentrations that were up to 10-fold higher than those carrying the unfavourable allelic variants (data not shown). In all cases, ‘Braeburn’ was the parent that transferred the beneficial allelic variants for high totAsA concentrations.
‘GGP is a key regulator of fruit vitamin C concentrations’

To validate these associations outside the mapping population, a selection of 22 apple commercial cultivars with known fruit totAsA concentrations (Davey et al., 2004; unpublished data) were genotyped with the same SNP-markers. In the case of *MdGPP2* (LG 3) and *MdNAT7-2* (LG 16) these markers yielded complex melting curve profiles in the High Resolution Melting (HRM) SNP-assay and so could not be used with this germplasm. This is presumably the result of additional polymorphisms present in the PCR amplicon region. However the SNP variants in *MdGGP1* and *MdGPP3* were successfully determined for 22 and for 11 of these cultivars respectively (Table 5). For the remaining 11 cultivars, the complex melting curve profiles for *MdGPP3* again precluded us from establishing SNP-genotypes. Results show that once again the *MdGGP1* ‘T’ allele and the *MdGGP3* ‘T’ allele were both associated with higher fruit totAsA concentrations in these cultivars. Furthermore, in the 11 cultivars for which we have the SNP-genotypes for both *MdGGP1* and *MdGGP3*, the favourable ‘T+T’ allele combination was always associated with an average 3-fold higher higher fruit totAsA concentration, and a maximum 6-fold higher concentration than those with the ‘C+C’ combination.

**Gene expression analysis**

To help understand the basis for the differences in totAsA contents in fruit of cultivars with different SNP-alleles, qPCR expression analysis of CGs were carried out in the flesh of fruit from ten apple genotypes with high or low AsA concentrations (Fig. 3). In all cases, the expression levels of *MdGGP1*, *MdGGP3* and *MdDHAR3-3* were positively and significantly correlated with fruit totAsA concentrations (Fig. 3B-G), with the mean difference in their relative expression among the high and low AsA groups being 2.0-, 1.5-, and 2.0-fold, respectively. In contrast, there were no statistically significant differences in the mean relative expression of *MdGGP2* and *MdNAT7-2*, although in both cases ‘Braeburn’ once again had a higher relative expression than ‘Telamon’ tissues.

Analysis of the RNA-Seq data shows that the mean expression levels of *MdGGP1* (LG 11) and *MdGGP2* (LG 3) are much higher than *MdGGP3* (LG10). Reads per kilobase exon model per million (RPKM) values for *MdGGP1*, *MdGGP2* and *MdGGP3* in ‘Braeburn’ and ‘Telamon’ respectively were 84.1 and 56.8, 61.4 and 50.4, and 13.9 and 15.3. Although these values are derived from RNA samples isolated during bud and flower development, comparison with the RPKM values of these genes in the low-AsA cultivar ‘Royal Gala’ at the mature fruit stage (132 days after full bloom) again shows that *MdGGP1* (RPKM 58) and *MdGGP2* (RPKM 52) are much more strongly expressed than *MdGGP3* (RPKM 5) (Laing – unpublished). Together with the QTL results this suggests that *MdGGP1* expression is particularly important for the regulation of tissue flesh AsA concentrations.
DISCUSSION

Antioxidant concentrations and QTL in ‘TxB’ mapping population: genetics versus environment

The absolute concentration of antioxidants varied considerably from year to year, with totAsA and totGSH concentrations being greatest in 2009. This is presumably due to the higher light and temperatures experienced prior to harvest in 2009, factors which are known to influence totAsA concentrations (Davey et al., 2000; 2007; Dumas et al., 2003). However there was no clear correlation between totAsA concentrations and mean daily temperatures or sunlight hours across the years, underlying the fact that it is difficult to dissociate the individual contributions of temperature, light, development and genetics to the regulation of antioxidant pools. None the less, the fact that stable (present in at least two measurement years) fruit flesh AsA-QTL (LGs 10, 11, and 17) were found demonstrates a strong genetic component to the control of fruit AsA concentrations in apple.

Environmental factors such as light intensity are expected to have a larger impact on AsA concentrations in the skin compared to flesh tissues and could therefore mask the underlying genetic control mechanisms. Despite this, significant QTL for skin AsA and totAsA concentrations were identified and most of these co-located with flesh AsA-QTL. Only the QTL cluster on LG 9 (both parents) and on the ‘Telamon’ LG 16 were specific for skin AsA concentrations. These regions on LG 9 and LG 16 have both been shown to contain major QTL for apple fruit polyphenolic contents (Chagné et al., 2007; 2012; Khan et al., 2012). We have previously reported that the sun-exposed (red) side of apples has higher anthocyanin and AsA concentrations and is more resistant to both biotic and abiotic stress than the shaded (green/yellow) side (Davey et al., 2004; 2007), and recently, Bulley et al. (2012) demonstrated that transgenic fruits from tomato and strawberry with elevated AsA contents also had a ~50% larger polyphenolic pool. Finally, we report here positive correlations between TAA, AsA and totAsA concentrations in both flesh and skin tissues, and co-locations between flesh AsA- and TAA-QTL on LGs 6, 11 and 16. Therefore, the skin-specific AsA-QTL on LGs 9 and 16 may be related to the interaction with the polyphenol contents in apple skin tissues. In this regard the recent report indicating that the accumulation of anthocyanins in Arabidopsis leaves is fine-tuned by the AsA redox state is of particular interest (Page et al., 2012).

For the first time we also present QTL for fruit GSH concentrations, but apart from QTL for flesh GSSG concentrations on LG 16 that co-localized with AsA-QTL in both parents, all the other QTL had low LOD scores, and no significant associations between AsA- and GSH- QTL were detected.
‘GGP is a key regulator of fruit vitamin C concentrations’

The active transport of foliar AsA to developing sink tissues has been previously demonstrated (Franceschi and Tarlyn, 2002; Hancock et al., 2003; Tedone et al., 2004). While both apple fruit and leaf tissues have been shown to be capable of AsA biosynthesis (Davey et al., 2004; Razavi et al., 2005; Li et al., 2010), there were no significant correlations between fruit and leaf totAsA concentrations, and only one co-location between leaf and fruit AsA-QTL on LG 6. This would also need to be confirmed, as it is based on only one year of measurements. Regardless, these results indicate that screening progeny on the basis of foliar AsA concentrations to identify cultivars rich in fruit vitamin C concentrations will not be highly informative.

Links between vitamin C concentrations, flesh browning and other fruit quality traits

Susceptibility to browning has been linked to flesh DHA concentrations in apple (Davey et al., 2006), and in pears the development of internal postharvest browning disorder occurs when AsA concentrations decline below a certain threshold value (Franck et al., 2003). Here, the time taken for cut apple flesh to complete browning and the final colour intensity following browning were both correlated with DHA concentrations and the % DHA of flesh tissues. In addition, stable QTL for browning co-located with stable QTL for flesh DHA concentrations on LG 17. This browning QTL was only detected in ‘Telamon’ which produces poor quality fruit, and in which the flesh % DHA was on average 3-fold higher than in ‘Braeburn’ (data not shown). Taken together, these results suggest that factors that lead to an oxidised AsA pool are associated with susceptibility to flesh browning.

In cherry tomatoes, fruit firmness and AsA concentrations have been reported to be closely correlated (Gilbert et al., 2009) and here we show that firm apples also generally contain higher flesh AsA concentrations. Flesh AsA-totAs concentrations and % DHA are also significantly correlated with fruit flesh soluble solids content, suggesting a link to the supply of sugar substrates for AsA biosynthesis. However, the significance of this is not clear, as a large cluster of QTL for fruit physiological traits, including fruit firmness, soluble solids content, fresh weight, and flesh browning all co-located with the fruit flesh AsA-QTL on LG 10 and therefore the correlation with fruit AsA-totAsA concentrations could be due to indirect effects on fruit physiology and quality.

Candidate genes regulating AsA concentrations

Candidate regulatory genes from the AsA biosynthetic pathways
‘GGP is a key regulator of fruit vitamin C concentrations’

GGP (VTC1) catalyses the conversion of D-mannose-1-P to GDP-D-mannose (Fig. 1) and we observed a
colocation between a putative MdGMP (MdGMP2-1) and the flesh AsA-QTL cluster on LG 6. However
these QTL explained only a relatively low proportion of the population variability (10%), and were
detected only in one year. In tomato (Ioannidi et al., 2009) and kiwifruit (Bulley et al., 2009), fruit AsA
concentrations were also not correlated with GMP expression.

GME has been proposed to be critical for the regulation of plant AsA concentrations by several groups
(Wolucka and Van Montagu, 2007; Gilbert et al., 2009), and in tomato, SlGME1 co-locates with a cluster
of stable fruit AsA-QTL (Stevens et al., 2007), while in apple MdGME transcript levels are highly
correlated with AsA concentrations during apple fruit development (Li et al., 2011). However, the data
presented here support results obtained in peach (Imai et al., 2009) and tomato (Ioannidi et al., 2009),
which suggest that this step is not a major point of control in mature apple fruit, at least under our field
conditions.

GGP (VTC2) (Fig. 1) catalyses the first committed step of AsA biosynthesis, and has also been
suggested to be the rate-limiting step for AsA biosynthesis in plants (Linster and Clarke, 2008; Bulley et
al., 2009, 2012). Results here provide strong evidence that MdGGP1 is the only structural gene of AsA-
metabolism that is tightly linked to flesh AsA concentrations in apple and that this is independent of the
environmental conditions, since significant AsA-QTL were detected over all three years of measurement.
While we cannot exclude the possibility that these QTL regions contain other regulatory genes, it is
interesting to note that the two other MdGGP paralogues (MdGGP2 on LG 3 and MdGGP3 on LG 10)
also mapped within fruit AsA-QTL.

The next step of the L-galactose pathway is catalysed by GPP (VTC4). Although GPP has been
suggested to be important in regulating AsA concentrations in tomato fruit (Ioannidi et al., 2009) and
apples (Li et al., 2011), these conclusions are not supported by our results. None of the other genes
involved in the main AsA biosynthetic pathway (PMI, PMM, GalDH or GLDH) mapped to positions
located within our AsA-QTL, suggesting that these have no major control over AsA homeostasis in
mature apple fruit, under Belgian field conditions.

From the alternative (uronic acid) AsA biosynthetic routes (Fig. 1), the reaction catalyzed by
GalUR (EC 1.1.1.203) in strawberry has been the most extensively studied (Agius et al., 2003). D-
galacturonate has been suggested to be an alternative substrate for the synthesis of AsA in fruit skin by
precursor feeding experiments in apple fruit tissues (Li et al., 2008), and biosynthesis proceeding via
GalUR has been suggested to play a more significant role in the later stages of fruit ripening when
GGP is a key regulator of fruit vitamin C concentrations

breakdown of cell wall pectin takes place (Badejo et al., 2011, Melino et al., 2009, Cruz-Rus et al., 2011). The in silico co-location of two MdGalUR paralogues within flesh AsA-QTL clusters on LG 16 of both parents supports a possible role for GalUR in mature fruits, but there were no other co-locations between genes of the alternative AsA biosynthetic pathways and AsA-QTL.

To summarise this section, except for paralogues of the first dedicated step of the L-galactose biosynthetic pathway (MdGGP1-3) our results do not support a key role for any of the other genes of the main L-galactose biosynthetic pathway. However, as reactions before the GGP step also supply substrates for cell wall polysaccharide biosynthesis, mutations in these genes could affect the flux of substrates through to GGP.

CGs from the AsA recycling pathway

Dehydroascorbate (DHA), the oxidized form of AsA, is unstable and can be readily degraded if not re-reduced by the activity of MDHAR (EC 1.6.5.4) and/or DHAR (EC 1.8.5.1). In tomato, an orthologue of MDHAR (SI MDHAR3) has been found to co-locate with a stable QTL for fruit AsA contents (Stevens et al., 2007), and MDHAR enzyme activity has been associated with elevated AsA concentrations under chilling stress (Stevens et al., 2008). In apple, we found no co-location between MdMDHARs and any antioxidant QTL. However, MddHAR3-3 mapped to the bottom of ‘Telamon’ LG 17 within a highly significant, stable QTL for flesh DHA concentrations, and MddHAR3-1 mapped in silico within QTL for skin AsA concentrations in LG 9, suggesting that orthologues of DHAR could be important in helping to maintain the reduced AsA pool in apple fruit. The co-location of MddHAR3-3 within a stable QTL for flesh browning suggests that genetic regulation of redox status of the AsA pool via DHAR is important for postharvest fruit quality traits in apple.

CGs related to intercellular AsA transport

Intercellular AsA transport via the nucleobase transporters (NATs) could represent an important mechanism of maintaining sink tissue AsA concentrations (Maurino et al., 2006). MdNAT7-2 mapped to the same position as significant fruit AsA-QTL on LG 16 of the ‘Braeburn’ genetic map, and another orthologue of NAT (MdNAT12-2) co-located in silico with a QTL for leaf DHA concentrations on LG 15 of both maps. As NATs are highly expressed in vascular tissues (Maurino et al., 2006), and they co-located with AsA-QTL in both fruit and leaf tissues, these results indicate they may be involved in the long-distance transport of AsA in apples.
Molecular markers for fruit AsA concentrations

Here we show that allelic variations in the nucleotide sequences of *MdGGP1*, *MdGGP3* and *MdNAT7-2* are associated with up to 10-fold differences in totAsA concentrations of fruit from individuals of our ‘Tx’B’ mapping population. The remaining *MdGGP* parologue, *MdGGP2*, also mapped within an AsA-QTL cluster on LG3, but this cluster was only detected in 2006. Importantly, fruit totAsA concentrations of 22 commercial apple cultivars were also associated with the same allelic variations in *MdGGP1* and *MdGGP3*, and the ‘T+T’ combination of alleles for *MdGGP1* and *MdGGP3*, is consistently associated with an up to 6-fold higher fruit totAsA concentration. Sequence analysis of the (corrected) *MdGGP1* CDS, showed that there are no non-synonymous SNPs (nsSNPs, Supplemental Table S11), while in *MdGGP2*, nsSNPs result in three amino acid changes (Supplemental Table S12), but none of these are predicted to alter structure or function. By comparison, *MdGGP3*, *MdDHAR3-3* and *MdNAT7-2* CDS contain multiple synonymous and nsSNPs. Further study of the impact of these sequence differences however was limited by the complexity and relatively low coverage of the *MdDHAR3-3* and *MdNAT7-2* transcript sequences. Since the SNPs in *MdGGP1* and *MdGGP2* are not associated with altered protein function, it is likely that the observed allelic associations are due to linkage with polymorphisms in the promoter region that alter allele expression. Gene expression analyses in a range of phenotypically diverse cultivars demonstrated that *MdGGP1* and *MdGGP3* as well as *MdDHAR3-3* expression levels are all strongly correlated with fruit totAsA concentrations (Fig. 3). Expression levels of *MdGGP2* (LG 3) or *MdNAT7-2* (LG 16), in contrast, were not significantly correlated, but this could be because they mapped within a non-year-stable AsA-QTL cluster and the expression analyses were carried out in fruit harvested in a different year. Conceivably therefore, these CGs might still be important only under certain environmental/developmental conditions. Finally, our RNA-Seq data demonstrate that mean expression levels of the third parologue of *MdGGP* (*MdGGP3*) are considerably lower than the expression of the other two gene copies in developing tissues as well as in mature fruit. This, in combination with the high number of SNPs and other polymorphisms, suggests that *MdGGP3* is less important for the regulation of tissue AsA concentrations and that the translation product may even be non-functional.
CONCLUSIONS

High fruit AsA concentration is considered to be a desirable trait for consumers, but in apple, as in other hard fruit species, the long juvenile period of plants means that fruit-specific traits cannot be determined until progeny reach maturity around three to five years after germination. As such there is much interest in the development of genetic markers for the early selection and/or screening of progeny for fruit AsA concentrations. Our QTL, CG mapping, and analysis of allelic variations across a selection of germplasm show that \textit{MdGGP} paralogues are consistently linked to fruit totAsA concentrations clearly indicating a major role for these genes. Gene expression analyses suggest that regulation of flesh AsA concentrations in apple is primarily through control of the expression levels of \textit{MdGGP1}, \textit{MdGGP3}, as well as of \textit{MdDHAR3-3}. The SNP-based markers developed here are thus excellent candidates for early screening of progeny for increased fruit AsA concentrations. Despite conflicting reports in the literature, our work does not support a major role for the other biosynthetic genes in apple fruit AsA homeostasis, at least under Belgian climatic conditions. However, we propose that \textit{MdDHAR3-3} helps regulate the redox state of the AsA pool in apple fruit, and that increased flesh DHA concentrations are associated with susceptibility to flesh browning.
‘GGP is a key regulator of fruit vitamin C concentrations’

MATERIAL AND METHODS

Mapping population

An F1 mapping population consisting of 257 individuals was created from a cross between the apple cultivars ‘Telamon’ and ‘Braeburn’ using ‘Telamon’ as the female parent (Kenis and Keulemans, 2005). For the 2005 and 2006 work, a population of mature trees more than 4 years old was used. Trees were grown on ‘M.9’ rootstock at the Rillaar experimental field station, Aarschot, Belgium. For the 2009 work, a regenerated copy of the same population was propagated on ‘M.27’ rootstock, grown on the same field.

Harvest and meteorological measurements

Fruit from individual progeny were harvested up to twice a week when considered commercially ripe. Ripeness was assessed by experts based on ripening parameters. A minimum of 10 healthy apples per genotype were sampled for AsA, totAsA, GSH, totGSH, and TAA analysis, and processed essentially as previously described (Davey et al., 2006). Fruit were immediately transported to the laboratory for analysis. Leaf tissue from 180 genotypes was collected before fruit harvest and stored at -80°C until further analysis. Mature apple fruit were also collected in 2011, pooled into three biological replications, each one consisting of at least three healthy apples, and processed immediately for AsA measurement or stored at -80°C for gene expression studies. Data for mean daytime air temperature (°C) and sun radiation (W/m²) during a period of five weeks before the harvest period, starting on 1 August, were obtained from the university meteorological station (Heverlee), located 25 km from the experimental field.

Fruit physiological measurements and flesh browning

For all years of QTL analysis, fruit fresh and dry weight were measured immediately after harvest. Flesh firmness was determined using a penetrometer fitted with an 11-mm diameter plug, while at the same time, the juice released by penetration of the probe was used to measure the soluble solids concentration (mainly sugars; °Brix) using a digital refractometer (Pocket PAL-1, Atago, Japan), essentially as described by Kenis et al. (2008). In 2006 and 2009, the susceptibility of the apple flesh to browning was assessed as the total time to complete visible browning (‘Br_time’) and the final colour of the affected tissue (‘Br_colour’) was evaluated (Davey and Keulemans, 2009).
AsA and GSH analysis

Samples for metabolite analysis were obtained from skin and flesh fractions from 10 individual fruit by blending using a metaphosphoric acid/EDTA/PVPP extraction, essentially as previously described (Davey et al., 2003; 2006). Extracts were filtered and injected into an HPLC system (Waters 2690 separation module) equipped with a C18 rocket column (GRACE, 53 mm x 7 mm). AsA and GSH were detected at 243 nm and 197 nm, respectively (Waters 996 photodiode array detector). AsA analysis usually occurred on the same day within a few hours of harvest, although at peak harvest time it was sometimes necessary to store apples for a maximum of up to 48 hours in a cool cell at 2°C before extraction. In total, the fruit from 136, 140 or 163 individual progeny were analysed for AsA, totAsA, and DHA concentrations in the season of 2005, 2006 or 2009, respectively, whereas GSH, totGSH, and GSSG measurements of the same samples were carried out in fruit only in 2006 and 2009.

Total antioxidant activity

The acid-soluble total antioxidant activity (TAA) of the fruit tissues harvested in 2006 was determined as the sum of the 2,2 –azinobis (3-ethylbenzthiazoline-6-sulfonic acid) (ABTS)-reactive substances present in aliquots of the same extracts used to measure AsA and GSH concentrations (Davey et al., 2007). Absorbance at 734 nm was measured in three technical replicates of all samples using the Multiskan Spectrum Microplate spectrophotometer (Thermo Labsystems). AsA was used as an external standard, and the TAA of each extract was expressed as AsA equivalents.

Linkage map analysis and development of a bin mapping set

The available genetic maps of ‘Telamon’ and ‘Braeburn’ (Kenis et al., 2008) were further saturated with 14 additional microsatellite loci and CG-based markers to improve marker density in key chromosome regions. The updated linkage maps of ‘Telamon’ and ‘Braeburn’ were constructed using the JoinMap 3.0 software (Van Ooijen and Voorrips, 2001), essentially as previously described (Davey et al., 2006; Kenis et al., 2008).
A total of 324 and 339 molecular markers were initially ordered in 17 linkage groups, corresponding to the 17 apple chromosomes. To improve the resolution of QTL analysis, markers that were less informative (scored on less than 50% of genotypes), markers with distorted segregation, or markers with two alleles heterozygous in both parents \((hk \times hk)\) segregation) were discarded, unless the gap between neighbour markers was greater than 10 cM. A subset of 298 markers was used to construct the ‘Telamon’ genetic map, with 96% of the markers highly informative \((ab \times aa; ef \times eg; ab \times cd)\), whereas 313 markers were employed for the ‘Braeburn’ genetic map which contained less than 10% less-informative markers \((hk \times hk)\). A bin set of 14 highly informative genotypes was created essentially as described by Chagné et al. (2008). Successful markers were considered those that were heterozygous in at least one parent, had a normal segregation ratio in the bin set, and mapped onto the expected position on the bin maps.

**QTL analysis**

QTL for AsA, GSH and fruit physiological traits were identified using phenotypic data from 2005, 2006 and 2009 and updated genetic maps of the ‘TxB’ population. QTL mapping was carried out using the MapQTL 4.0 software (Van Ooijen et al., 2002). QTL were detected using interval mapping (IM) in combination with restricted multiple QTL model mapping (rMQM) for normally distributed traits with a step size of 1 cM, or Kruskal-Wallis non-parametric test for the non-normally distributed ones. In rMQM mapping, the most informative marker with the highest LOD score within the intervals of a QTL was selected as a co-factor, and multiple rMQM runs were performed until the co-factors selected were stable, as described previously (Davey et al., 2006). The LOD threshold for the significance of a QTL was selected based on the genome-wide level, using 1,000 permutations at 90% for each trait (around 3.5 or 3.0 for fruit and leaf tissues, respectively). A QTL was classified as significant/non-significant and major/minor based on a LOD threshold of 3.5 and 20% of explained population variability. QTL identified were described by the marker with the highest LOD score in the corresponding QTL region, and QTL regions defined as the 1.5-LOD support interval. Linkage maps and QTL were drawn using MapChart 3.0 software (Voorrips, 2002).

**Candidate gene mapping**
The sequences of AsA-related CGs from Arabidopsis or other plant species were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/genbank/), PLAZA 2.0 (http://bioinformatics.psb.ugent.be/plaza/), Phytozome Genome Browser (http://www.phytozome.net), or SOL Genomics Network (http://solgenomics.net). Orthologous sequences were retrieved from the ‘Golden Delicious’ assembly (Velasco et al., 2010) using BLASTX versus apple consensus CDS set. Gene positions on the genome assembly and their sequences were extracted from the IASMA genome browser (http://genomics.research.iasma.it). The amino acid sequences for each CG were aligned using BLOSUM62 matrix and phylogenetic analysis was conducted using the neighbour joining method and 1000 bootstrap replicates in Geneious software version 5.4.6 (Drummond et al., 2011) in order to determine the number of copies present in Malus for each Arabidopsis accession. The positions of these homologues were aligned in silico across the 17 linkage groups of ‘TxB’ map, based on the physical position of framework SSRs.

SNP-Marker development

RNASeq data from ‘Telamon’ and ‘Braeburn’ were used to identify SNPs and polymorphisms in CGs sequences. In total 48.1 and 73.6M 75 bp paired end reads respectively were generated from RNA pooled from eight different developmental stage of tree bud development, and two tissue types. Reads were trimmed, and then mapped against the apple consensus CDS set, downloaded from the GDR website (http://www.rosaceae.org/projects/apple_genome) using default mapping parameters within CLCBio Genomics Workbench v5.5. CLCBio software was also to carry out the de novo assembly of reads into contigs for each cultivar separately, again using default alignment parameters and the automatic scaffolding function. SNPs were called relative to the ‘Golden Delicious’ reference CDS sequences on the basis of using a minimum read depth of 6-fold and a % proportion coverage of >25%. Finally, to control gene annotation, the consensus sequences from both the mapped reads and the de novo contigs were extracted and aligned against the annotated gDNA contigs used to create the apple genome assembly, again downloaded from GDR and using Geneious v5.6 software package. Functional descriptions of consensus CDS sequences and translated protein sequences were confirmed using Blast2GO software (http://www.blast2go.com/b2ghome), as well InterPro (Geneious v5.6). To validate gene function, clones of GGP were expressed in Escherichia coli, and GGP activity was measured essentially as previously described (Laing et al., 2007).
PCR primers were developed to identify length polymorphisms, or flanking SNPs to allow progeny genotyping using the HRM technique (Chagné et al., 2008; Studer et al., 2009). Gene-specific SNP primer pairs were designed using Primer3 software version 4.0 (http://frodo.wi.mit.edu/primer3/) employing the following conditions: product size between 75 to 200 bp, with optimum 120 bp spanning the identified SNPs; GC% ranging from 40-55%; self complementarity and 3’ self complementarity were set at 4 and 1, respectively; and with all the other parameters set at default. All CG-derived markers developed were screened in the ‘TxB’ bin mapping set and both parents.

Genomic DNA from leaves was extracted from ~100 mg powdered aliquots using the ‘Genomic DNA Purification kit®’ (MBI Fermentas). HRM conditions were as follows: the PCR reaction for each sample was carried out in 15 µl containing the HRM Master Mix (Type-it HRM PCR Kit, QIAGEN), 10-20 µM of each primer, and 5ng genomic DNA. Cycling conditions were incubation at 95°C for 5 min, followed by 40 cycles of 95°C (10 s) and 55°C (30 s). HRM analysis followed immediately by denaturation from 65°C to 95°C, rising by 0.1 degree in each step. The HRM analysis was carried out in a Rotorgene Q thermocycler software (QIAGEN). Markers that mapped to their expected positions and gave an expected segregation rate were further screened over 100 individuals from the ‘TxB’ population to construct new genetic maps, as previously described.

Allelic typing and gene expression studies

Gene expression studies were performed following MIQE (Minimum Information for publication of Quantitative Real-Time PCR Experiments) guidelines (Bustin et al., 2009). Total RNA was isolated from flesh fruit tissues using a modified CTAB method (Gasic et al, 2004) and RNeasy Plus Mini Kit following the manufacturer’s instructions (QIAGEN). The purity of total RNA extracted was determined as the 260/280 nm and 260/230 nm ratio with NanoDrop 2000 (Thermo Scientific), and the integrity was checked by electrophoresis in 1% agarose gel (Gel Doc EZ Imager by Bio-Rad). One µg of total RNA was reverse transcribed to cDNA with Superscript II (Invitrogen) using random oligo d(T) primers. The expression profiles of CGs considered to be key control points for AsA accumulation (MdGGP1, MdGGP2, MdGGP3, MdDHAR3-3 and MdNAT7-2; primer pair sequences listed in Supplemental Table S9 were validated by real-time quantitative qPCR (qPCR) using SYBR Green I technology on a Rotor Gene Q (Qiagen). 18S-rRNA (5’-GTT ACT TTT AGG ACT CCG CC-3’ and 5’-TTC CTT TAA GTT TCA GCC TTT-3’), Malus eIF-4A (5’-ATC AGG CTC ATC CCG TGT-3’ and 5’-AGC AAC ACC ACC CTT CCT TCC-3’; Zubini et al., 2007) and Arabidopsis actin-11 (5’-GGA CCT TGC AGG CCG TGA CC-3’
and 5'-AAC CTC CGG GCA GCG GAA TC-3') were used as reference genes. The stability of the expression of the reference genes was assessed with NormFinder software (http://www.mdl.dk/publicationsnormfinder.htm). All reactions were set up in duplicate, containing 1 μl cDNA template (50 ng), 7.5 μl Absolute™ QPCR SYBR Green Mix and 1 μl of each primer (3.75 mM) in a final volume of 15 μl. The cycling conditions were as follows: denaturation step at 95°C for 10 min, followed by 45 cycles of denaturation at 95°C for 20 s, annealing at 63°C for 20 s and extension at 72°C for 20 s. A melting curve analysis was performed, ranging from 55°C to 95°C, with temperature increasing steps of 0.5°C/s. Criteria of acceptance for reaction efficiency ranged from 0.83 to 1.03 and both R and R² > 0.97. For each run, a standard curve was included based on cDNA pooled from all samples to be analyzed in a range of different dilutions. The relative quantification of expression levels was performed using the comparative C_T method (Pfaffl, 2001). Two technical replications were performed for each of the three biological replications per sample. All expression data were calculated as an expression ratio relative to the geometric mean of the expression of 18S-rRNA and actin as these showed the highest stability in our sample set with stability values (Vandesompele et al., 2002) of 0.34 and 0.47 respectively.

Descriptive statistical analysis

Statistical analysis of all traits was carried out using the SAS 9.2 software package (SAS Institute). Mean values, standard deviations, ranges and skewness of the F_1 progeny were calculated for each year and trait, while the normality of the distributions was tested using Shapiro-Wilk test (Shapiro and Wilk, 1965). Pearson correlation coefficients were employed to compare mean values of the traits over the three years. The significance of the mean fold difference in totAsA concentration of genotypes carrying different alleles/allele combinations, as well as the significance of the mean relative expression of CGs of high/low AsA genotypes, were tested with Student’s t-test (*: P<0.05, **: P<0.01 and ***: P<0.001).

ACKNOWLEDGEMENTS

IM was supported by a FLOF fellowship from KU Leuven, as well as by a travel grant from Plant & Food Research to visit their institute in Palmerston North, New Zealand. DC and WAL received funding from the Plant & Food Research Pipfruit Research Investment Project (PRIP). The authors acknowledge stimulating discussions with Otto van Poeselaere.
‘GGP is a key regulator of fruit vitamin C concentrations’

SUPPLEMENTAL MATERIAL

Supplemental Figure S1: Climatic conditions prior to fruit harvest. Mean weekly daytime air temperature (°C) (a) and radiation (W/m²) (b) during a period of 5 weeks starting on 1 August and prior to the start of harvest period over years.

Supplemental Figure S2: Genetic linkage maps of ‘Telamon’ (TEL) and ‘Braeburn’ (BR) used for QTL analysis. Candidate genes in bold type were mapped in this study.

Supplemental Figure S3: Alignment of the MDP0000172222 (MdGGP1) consensus coding sequence (CDS) from apple cultivars ‘Braeburn’ (Br), ‘Telamon’ (Tel) and ‘Golden Delicious’ (GD) with the full length EST sequence from ‘Royal Gala’ (gi|225380879). Orange bars represent the position of SNPs relative to the ‘Golden Delicious’ reference CDS. Small green arrows indicate corrected translation start. Blue arrow indicates PolyA tail. The translation products of the corrected CDS are identical.

Supplemental Figure S4: Alignment of the MDP0000172222 (MdGGP1) consensus coding sequence (CDS) from apple cultivars ‘Braeburn’ (Br), ‘Telamon’ (Tel) and the full length EST sequence from ‘Royal Gala’ (gi|225380879) to the gDNA scaffold MDC003775.57. Yellow arrows indicate predicted exons of the reference CDS. Orange bars represent the position of SNPs relative to the ‘Golden Delicious’ reference sequence. Small green arrows indicate corrected translation start. The translation products of the corrected CDS are identical. Comparison of the alignments of the consensus CDS sequences and de novo contigs from ‘Telamon’ and ‘Braeburn’ against the gDNA contig used for genome assembly indicated that the start codon of the MdGGP1 reference sequence has been incorrectly predicted. This is because of ambiguous nucleotides and a 23 bp insertion in the reference sequence at position 327 bp, leading to a frame shift. As a result, a TGA stop codon that is present in the 5’ UTR of all ‘Telamon’ and ‘Braeburn’ sequences is missed and the correct ATG is actually located 361 bp further downstream from the predicted start codon. Alignments with available EST sequences from NCBI, and comparisons with other plant GGP peptide sequences within PLAZA v 2.5 (Van Bel et al., 2012), support these conclusions. As a result, the corrected MDP0000172222 CDS length is only 1344 instead of 1704 bp, and the translation product is 448 amino acids long instead of 568. This revised gene structure also agrees much more closely with the sequence of its parologue, MdGGP2 (MDP0000288088, Figure S3) on LG3, which has a CDS length of 1347 bp, and a translation product of 449 amino acids.

Supplemental Figure S5: Alignment of the MDP0000191488 (MdGGP2) consensus coding sequence (CDS) from ‘Braeburn’ (Br), ‘Telamon’ (Tel) and ‘Golden Delicious’ (GD). Orange bars represent the position of SNPs relative to the ‘Golden Delicious’ reference CDS. Read mapping, and alignment of
‘GGP is a key regulator of fruit vitamin C concentrations’

* de novo contigs from ‘Telamon’ and ‘Braeburn’ against the relevant gDNA assembly contig support the predicted reference gene annotation.

Supplemental Figure S6: Alignment of the MDP0000191488 (MdGGP3) consensus coding sequence (CDS) sequences from ‘Braeburn’ (Br), ‘Telamon’ (Tel) and ‘Golden Delicious’ (GD). Orange bars represent the position of SNPs relative to the ‘Golden Delicious’ reference CDS. Alignment of the MdGGP3 consensus read mapping sequences from ‘Telamon’ and ‘Braeburn’ as well as of *de novo* contigs generally supported the predicted reference gene annotation.

Supplemental Figure S7: Alignment of the MDP0000191488 (MdGGP3) consensus peptide sequences (translation products) from ‘Braeburn’ (Br), ‘Telamon’ (Tel) and ‘Golden Delicious’ (GD). ‘MDP0000191488_Br_1’ and ‘MDP0000191488_Br_2’ represent two possible alternate ‘Braeburn’ translation products arising from a single nucleotide deletion that leads to a frameshift, a premature stop codon and an N-terminal truncated peptide.

Supplemental Table S1: Correlations between fruit antioxidant metabolite concentrations. Pearson’s correlation coefficients for fruit L-ascorbic acid (AsA), dehydroascorbic acid (DHA), total ascorbic acid (AsA + DHA), glutathione (GSH), glutathione disulphide (GSSG), total glutathione (GSH + GSSG) concentrations and total antioxidant activity (TAA) of flesh (FL) and skin (SK) tissues, as well as fruit fresh weight (FW) in the F1 progeny measured in 2005, 2006 and 2009. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, n.s. not significant. For 2005, n=136; 2006, n=140; 2009, n=163.

Supplemental Table S2: Correlations between fruit antioxidant metabolite concentrations over the three years of measurements. Pearson’s correlation coefficients for fruit L-ascorbic acid (AsA), dehydroascorbic acid (DHA), total ascorbic acid (AsA + DHA), % DHA, glutathione (GSH), glutathione disulphide (GSSG), total glutathione (GSH + GSSG), % GSSG and total antioxidant activity (TAA) of flesh (FL) and skin (SK) tissues in the F1 progeny measured in 2005, 2006 and 2009. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, n.s. not significant. For 2005, n=136; 2006, n=140; 2009, n=163.

Supplemental Table S3: Overview of population mean values and distribution for foliar L-ascorbic acid (AsA), dehydroascorbate (DHA), total ascorbic acid (totAsA), glutathione (GSH), oxidized GSH (GSSG) and total GSH (totGSH) concentrations. All measurements expressed in μmol gFW⁻¹. Apple leaves from 180 individual of progeny of the ‘Telamon’ x ‘Braeburn’ mapping population were collected in 2009. Each trait value represents the mean of the results from the analysis of 10 randomly chosen mature leaves per genotype ± standard deviations (SD).
GGP is a key regulator of fruit vitamin C concentrations

Supplemental Table S4: Correlations between glutathione traits in foliar and in fruit flesh and skin tissues. Pearson’s correlation coefficients between fruit glutathione (GSH), glutathione disulphide (GSSG), total glutathione (GSH + GSSG) concentrations and % GSSG of leaf and flesh (FL) or skin (SK) tissues, in the F1 progeny measured in 2009. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, n.s. not significant.

Supplemental Table S5: Correlations between fruit l-ascorbic acid (AsA) concentrations and physiological traits. Pearson’s correlation coefficients between l-ascorbic acid (AsA), dehydroascorbic acid (DHA), and total ascorbic acid (AsA + DHA) concentrations of flesh (FL) and skin (SK), and fruit flesh browning (Br; Br_colour and Br_time), % dry weight (DW), flesh firmness and soluble solids (°Brix) in the F1 progeny measured in 2006 and 2009. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, n.s. not significant. For 2006, n=140; 2009, n=163.

Supplemental Table S6: QTL mapping for fruit antioxidant concentrations. Overview of QTL identified for apple fruit l-ascorbic acid (AsA), total AsA (totAsA), dehydroascorbate (DHA), glutathione (GSH), total GSH (totGSH), oxidized GSH (GSSG) and total antioxidant activity (TAA) in flesh (FL) and skin (SK) tissues in the ‘Telamon’ x ‘Braeburn’ mapping population. Borderline-significant QTL (3.0<LOD<3.5) are indicated with an asterisk (*). For QTL identified using the Kruskal-Wallis nonparametric mapping functions, we used significance levels of ***P=0.01, ****P=0.005, *****P=0.0005, and ******P=0.00005 and *******P=0.0001. Gene-specific markers are indicated in italics.

Supplemental Table S7: QTL mapping of fruit physiological traits. Overview of QTLs identified for fresh weight (FW), flesh browning traits (Br; Br_colour and Br_time), firmness and soluble solids (°Brix) in progeny of the ‘TxB’ mapping population. Borderline-significant QTLs (3.0 < LOD < 3.5) are indicated with an asterisk (*). For QTL obtained with Kruskal-Wallis nonparametric mapping functions, significance levels of ***P = 0.01, ****P = 0.005, *****P = 0.001, ******P = 0.0005 and *******P = 0.0001 were used. Gene-specific markers are indicated in italics.

Supplemental Table S8: Malus x domestica orthologues of candidate genes of l-ascorbic acid (AsA) metabolism. Apple candidate gene orthologues in the ‘Golden Delicious’ apple genome were identified on the basis of similarities to Arabidopsis reference sequences and multi-species phylogenetic analyses. Gene identifications, description and position (kb) on the linkage groups (LG) of the apple physical map are also displayed. Names were attributed based on their similarity to functionally characterised Arabidopsis orthologues.
‘GGP is a key regulator of fruit vitamin C concentrations’

Supplemental Table S9: List of selected candidate gene orthologues of L-ascorbic acid metabolism mapped on the ‘Telamon’ x ‘Braeburn’ genetic linkage maps. SNP-based markers for each candidate gene were developed using RNA-Seq data, and mapped using the high-resolution melting technique. Details of the gene ID, the forward and reverse primers, the length of the PCR product, the segregation type, and the position (cM) on the linkage group (LG) are shown.

Supplemental Table S10: Activities of the clones of Malus sylvestris (GB CN915822) and Malus x domestica (CN939721) with 100% identity to MdGGP1 and MdGGP2, respectively, at the transcript level when expressed in Escherichia coli. Results represent mean values of n replications (± standard deviations).

Supplemental Table S11: Polymorphisms present in ‘Braeburn’ and ‘Telamon’ MDP0000172222 (MdGGP1) consensus coding sequence (CDS), and read coverage per allele. Read coverage determined from RNA-Seq data following read mapping to the reference CDS. SNPs in bold were used for the development of HRM markers. Homozygous SNPs contain reads for only one of the two possible variants.

Supplemental Table S12: Polymorphisms present in ‘Braeburn’ and ‘Telamon’ MDP0000288088 (MdGGP2) consensus coding sequence (CDS), and read coverage per allele. Read coverage determined from RNA-Seq data following read mapping to the reference CDS. SNPs in bold were used for the development of HRM markers. Homozygous SNPs contain reads for only one of the two possible variants. Non-synonymous SNPs at reference positions 161 and 171 are part of the same codon.
GGP is a key regulator of fruit vitamin C concentrations

LITERATURE CITED


Bustin SA (2009) Why the need for qPCR publication guidelines?-The case for MIQE. Methods 50, 217-226


‘GGP is a key regulator of fruit vitamin C concentrations’


Davey MW, Keulemans J (2004) Determining the potential to breed for enhanced antioxidant status in Malus: Mean inter- and intravarietal fruit vitamin C and glutathione concentrations at harvest and their evolution during storage. Journal of Agricultural and Food Chemistry 52: 8031-8038


'GGP is a key regulator of fruit vitamin C concentrations'


Franceschi VR, Tarlyn NM (2002) L-ascorbic acid is accumulated in source leaf phloem and transported to sink tissues in plants. Plant Physiology 130: 649-656


'GGP is a key regulator of fruit vitamin C concentrations'


GGP is a key regulator of fruit vitamin C concentrations

(2012) Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. Journal of Experimental Botany 63: 2895-2908


GGP is a key regulator of fruit vitamin C concentrations


‘GGP is a key regulator of fruit vitamin C concentrations’


‘GGP is a key regulator of fruit vitamin C concentrations’


Legends to figures

Figure 1: L-ascorbic acid (AsA) biosynthetic and recycling pathways in plants. i) L-galactose pathway reactions 1-9; ii) L-gulose pathway, reactions 1 - 5, 10-13; iii) D-galacturonate pathway, reactions 14-16; iv) myo-inositol/glucuronate pathway, reactions 17-21; v) recycling pathway, reactions 22-27. Reactions with question marks yet to be identified. 1, glucose-6-phosphate isomerase; 2, mannose-6-phosphate isomerase (PMI; EC 5.3.1.8); 3, phosphomannomutase (PMM; EC 5.4.2.8); 4, GDP-D-mannose pyrophosphorylase (GMP; EC 2.7.7.13); 5, GDP-D-mannose 3’,5’-epimerase (GME; EC 5.1.3.18); 6, GDP-L-galactose-phosphorylase (VTC2 or GGP; EC 2.7.7.69); 7, L-galactose-1-P phosphatase (VTC4 or GPP; EC 3.1.3.25); 8, L-galactose dehydrogenase (GalDH; EC 1.1.1.48); 9, L-galactono-1,4-lactone dehydrogenase (GLDH; EC 1.3.2.3); 10, nucleotide pyrophosphatase or sugar-1-P guanyltransferase; 11, sugar phosphatase; 12, sugar dehydrogenase; 13, L-gulono-1,4-lactone oxidase (EC 1.1.3.8); 14, D-galacturonate-1-phosphate uridylyltransferase and D-galacturonate-1-phosphate phosphatase (possible); 15, D-galacturonic reductase (GalUR; EC 1.1.1.9); 16, aldonolactonase; 17, myo-inositol oxygenase (MIOX; EC 1.13.99.1); 18, D-glucurono-1-phosphate reductase (EC 1.1.1.19); 19, L-gulonolactonase; 20, D-glucuronate-1-phosphate uridylyltransferase; 21, D-glucurono-1-phosphate phosphatase; 22, L-ascorbate peroxidase (APX; EC 1.11.1.11); 23, L-ascorbate oxidase (AO; EC 1.10.3.3); 24, monodehydro-ascorbate reductase (MDHAR; EC 1.6.5.4); 25, dehydroascorbate reductase (DHAR; EC 1.8.5.1); 26, glutathione reductase (GR; EC 1.8.1.7).

Figure 2: Candidate gene and QTL mapping. Overview of the locations of candidate genes and QTL for mean L-ascorbic acid (AsA), total ascorbic acid (AsA + DHA), dehydroascorbic acid (DHA), glutathione (GSH), total glutathione (GSH + GSSG), glutathione disulphide (GSSG), and total antioxidant activity (TAA) of fruit flesh and skin tissues, as well as leaves, and for rate of flesh browning (Br, Br_colour and Br_time), and firmness detected in 2005 (solid dark QTL), 2006 (solid grey QTL) and 2009 (hatched QTL), on the genetic maps of the apple cultivars ‘Telamon’ (TEL) and ‘Braeburn’ (BR). Candidate genes in bold type were mapped in this study and are listed in Supplemental Table S9. Significant QTL (LOD>3.5) with 1.5 LOD support intervals are shown to the side of each linkage group (LG) and non-significant QTL (LOD<3.5) are indicated in italics.

Figure 3: Gene expression studies. Changes in flesh total ascorbic acid (total AsA) concentrations (nmol/gFW) (A) and expression levels of MdGGP1 (B), MdGGP2 (C), MdGGP3 (D), MdDHAR3-3 (E), MdNAT7-2 (F) and MdeIF1-4A (G) in the mature fruit of apple genotypes with high total AsA concentrations (dark gray bars) and low total AsA concentrations (light gray bars). qRT-PCR values were normalized against 18S-rRNA and actin. The error bars represent standard deviation of three biological
`GGP is a key regulator of fruit vitamin C concentrations`
GGP is a key regulator of fruit vitamin C concentrations

P<0.05, **: P<0.01 and ***: P<0.001). Associations were carried out in 71 (2005), 73 (2006) or 85 (2009) individuals of progeny for which both phenotypic and genotypic data were available.

Table 5: Association between fruit total ascorbic acid (totAsA) concentrations and candidate gene allelic variants in commercial cultivars. Associations between totAsA concentrations (nmol gFW$^{-1}$) of the entire (flesh and skin) fruit (Davey et al., 2004) and allelic variations of $MdGPP1$ (T$\rightarrow$C; SNP$_{885}$) and $MdGGP3$ (T$\rightarrow$C; SNP$_{211}$) across 22 apple commercial cultivars, as well as across the three ‘Telamon’ x ‘Braeburn’ individuals with highest or lowest fruit mean totAsA concentration are shown. Candidate gene SNP variants were determined by high-resolution melting analysis using the same SNP-markers developed for ‘Telamon’ and ‘Braeburn’. For $MdGPP3$, the complex melting curve profiles of 11 cultivars prevented SNP-genotyping (-). Mean fruit totAsA concentrations were determined in 2004 for the commercial cultivars (Davey et al., 2004), while mean measurement of three years (2005, 2006 and 2009) were used for the top and bottom ‘TxB’ individuals.
‘GGP is a key regulator of fruit vitamin C concentrations’

Tables

Table 1: Overview of population mean values and distribution for fruit antioxidant traits of flesh (FL) and skin (SK) tissues, as well as fruit fresh weight (FW) over years. Fruit L-ascorbic acid (AsA), total AsA (totAsA), dehydroascorbate (DHA) and % proportion of DHA were measured over a period of 3 years (2005, 2006, and 2009), fruit glutathione (GSH), total GSH (totGSH), oxidized GSH (GSSG) and % proportion of GSSG were analyzed over 2 years (2006 and 2009), whereas fruit total antioxidant activity (TAA) were measured only in 2006. Each trait value represents the mean of the results from the analysis of 10 randomly chosen apples per n genotypes (2005, n=136; 2006, n=140; 2009, n=163) ± standard deviations (SD).

<table>
<thead>
<tr>
<th>Trait</th>
<th>Year</th>
<th>Mean (± SD)</th>
<th>Lowest</th>
<th>Highest</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>AsA FL</td>
<td>2005</td>
<td>438.5 (± 166.4)</td>
<td>84.62</td>
<td>909.6</td>
<td>normal</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>313.5 (± 136.7)</td>
<td>43.70</td>
<td>828.1</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>600.2 (± 238.4)</td>
<td>155.9</td>
<td>1473</td>
<td>normal</td>
</tr>
<tr>
<td>DHA FL</td>
<td>2005</td>
<td>51.95 (± 75.38)</td>
<td>0.000</td>
<td>426.3</td>
<td>non normal</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>68.16 (± 29.69)</td>
<td>0.000</td>
<td>152.3</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>55.33 (± 38.57)</td>
<td>8.010</td>
<td>267.1</td>
<td>non normal</td>
</tr>
<tr>
<td>% DHA FL</td>
<td>2005</td>
<td>10.32 (± 11.84)</td>
<td>0.000</td>
<td>61.12</td>
<td>non normal</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>19.18 (± 8.440)</td>
<td>0.000</td>
<td>55.04</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>8.373 (± 4.054)</td>
<td>1.215</td>
<td>30.21</td>
<td>non normal</td>
</tr>
<tr>
<td>TotAsA FL</td>
<td>2005</td>
<td>490.5 (± 176.7)</td>
<td>111.8</td>
<td>1034</td>
<td>normal</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>381.7 (± 148.4)</td>
<td>97.20</td>
<td>973.0</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>654.1 (± 256.5)</td>
<td>164.8</td>
<td>1663</td>
<td>normal</td>
</tr>
<tr>
<td>AsA SK</td>
<td>2005</td>
<td>1617 (± 464.9)</td>
<td>413.7</td>
<td>2969</td>
<td>normal</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>1077 (± 332.2)</td>
<td>405.8</td>
<td>2459</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>2705 (± 796.6)</td>
<td>1052</td>
<td>5883</td>
<td>normal</td>
</tr>
<tr>
<td>DHA SK</td>
<td>2005</td>
<td>158.0 (± 188.0)</td>
<td>0.000</td>
<td>1262</td>
<td>non normal</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>115.2 (± 70.47)</td>
<td>0.000</td>
<td>580.7</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>242.0 (± 129.1)</td>
<td>44.31</td>
<td>758.0</td>
<td>non normal</td>
</tr>
<tr>
<td>% DHA SK</td>
<td>2005</td>
<td>9.004 (± 9.533)</td>
<td>0.000</td>
<td>46.02</td>
<td>non normal</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>9.885 (± 5.334)</td>
<td>0.000</td>
<td>35.81</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>8.788 (± 5.098)</td>
<td>1.771</td>
<td>42.27</td>
<td>normal</td>
</tr>
<tr>
<td>TotAsA SK</td>
<td>2005</td>
<td>1775 (± 470.6)</td>
<td>465.4</td>
<td>3377</td>
<td>normal</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>1193 (± 350.3)</td>
<td>464.1</td>
<td>2660</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>2947 (± 819.8)</td>
<td>1289</td>
<td>6260</td>
<td>normal</td>
</tr>
<tr>
<td>GSH FL</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>27.76 (± 11.09)</td>
<td>7.957</td>
<td>81.38</td>
<td>non normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>60.81 (± 20.84)</td>
<td>16.77</td>
<td>152.8</td>
<td>normal</td>
</tr>
<tr>
<td>GSSG FL</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>11.23 (± 5.267)</td>
<td>0.000</td>
<td>27.58</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>21.61 (± 14.69)</td>
<td>1.491</td>
<td>84.30</td>
<td>non normal</td>
</tr>
<tr>
<td>% GSSG FL</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>28.87 (± 10.27)</td>
<td>0.000</td>
<td>52.46</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>55.33 (± 13.97)</td>
<td>2.560</td>
<td>67.72</td>
<td>non normal</td>
</tr>
<tr>
<td>TotGSH FL</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>38.99 (± 13.70)</td>
<td>10.49</td>
<td>93.22</td>
<td>non normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>82.42 (± 26.22)</td>
<td>34.40</td>
<td>184.0</td>
<td>non normal</td>
</tr>
<tr>
<td>GSH SK</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>48.81 (± 15.88)</td>
<td>21.31</td>
<td>108.1</td>
<td>non normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>110.4 (± 38.18)</td>
<td>34.87</td>
<td>210.6</td>
<td>normal</td>
</tr>
<tr>
<td>GSSG SK</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(nmol gFW⁻¹)</td>
<td>2006</td>
<td>13.35 (± 6.179)</td>
<td>0.000</td>
<td>33.21</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>36.67 (± 22.35)</td>
<td>1.041</td>
<td>113.6</td>
<td>non normal</td>
</tr>
<tr>
<td>% GSSG SK</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>21.36 (± 7.545)</td>
<td>0.000</td>
<td>41.38</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>24.97 (± 13.26)</td>
<td>0.666</td>
<td>67.59</td>
<td>non normal</td>
</tr>
<tr>
<td>TotGSH SK</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
'GGP is a key regulator of fruit vitamin C concentrations’

<table>
<thead>
<tr>
<th>(nmol gFW⁻¹)</th>
<th>2006</th>
<th>2009</th>
<th>2009</th>
<th>138.9</th>
<th>non normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62.16 (± 19.45)</td>
<td>147.3 (± 42.58)</td>
<td>20.44</td>
<td>276.1</td>
<td>normal</td>
</tr>
<tr>
<td>TAA FL</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(μmol AsA gFW⁻¹)</td>
<td>2006</td>
<td>3.573 (± 1.381)</td>
<td>0.776</td>
<td>8.390</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TAA SK</td>
<td>2005</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(μmol AsA gFW⁻¹)</td>
<td>2006</td>
<td>11.46 (± 4.541)</td>
<td>1.861</td>
<td>25.17</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gFW</td>
<td>2005</td>
<td>150.6 (± 39.46)</td>
<td>54.49</td>
<td>265.7</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>113.7 (± 33.57)</td>
<td>19.00</td>
<td>198.6</td>
<td>normal</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>156.6 (± 47.19)</td>
<td>43.94</td>
<td>291.4</td>
<td>normal</td>
</tr>
</tbody>
</table>
‘GGP is a key regulator of fruit vitamin C concentrations’

Table 2: QTL for fruit flesh AsA concentrations. Overview of the stable (detected in at least 2 years) QTL identified for apple fruit L-ascorbic acid (AsA), total AsA (totAsA), and dehydroascorbate (DHA) in flesh (FL) tissues in the ‘Telamon’ x ‘Braeburn’ mapping population. For QTL identified using the Kruskal-Wallis nonparametric mapping functions, we used significance levels of ***P=0.01, ****P=0.005, *****P=0.001, ******P=0.0005 and ****** P=0.0001. Gene-specific markers are indicated in italics. An overview of all the QTL for antioxidant concentrations is given in Supplemental Table S6.

<table>
<thead>
<tr>
<th>Parent</th>
<th>‘Telamon’</th>
<th>Trait</th>
<th>LG</th>
<th>Year</th>
<th>LOD</th>
<th>Marker</th>
<th>% variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>AsA FL</td>
<td>10</td>
<td>2005</td>
<td>7.03</td>
<td>MdGGP3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>6.25</td>
<td>EAATMCCG246</td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>2005</td>
<td>4.02</td>
<td>CH04g07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006</td>
<td>3.5</td>
<td>CH04g07</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>8.5</td>
<td>MdGGP1</td>
<td>27.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TotAsA FL</td>
<td>10</td>
<td>2005</td>
<td>5.68</td>
<td>CH03d11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>4.5</td>
<td>CH03d11</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>2006</td>
<td>4.23</td>
<td>MdGGP1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>5.8</td>
<td>MdGGP1</td>
<td>18.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DHA FL</td>
<td>17</td>
<td>2005</td>
<td>****</td>
<td>EAAGMCAC216</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>****</td>
<td>EAAGMCAC216</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent</th>
<th>‘Braeburn’</th>
<th>Trait</th>
<th>LG</th>
<th>Year</th>
<th>LOD</th>
<th>Marker</th>
<th>% variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>AsA FL</td>
<td>10</td>
<td>2005</td>
<td>7.8</td>
<td>CH03d11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>4.5</td>
<td>CH03d11</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>2005</td>
<td>3.9</td>
<td>CH04g07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006</td>
<td>4.6</td>
<td>EAGCMCAT122</td>
<td>21.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>5.62</td>
<td>ECATMCAC190</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TotAsA FL</td>
<td>10</td>
<td>2005</td>
<td>5.98</td>
<td>CH03d11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>4.54</td>
<td>EAGAMCCA110</td>
<td>12.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>2006</td>
<td>5.72</td>
<td>EAGCMCAT122</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>4.08</td>
<td>ECATMCAC190</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DHA FL</td>
<td>17</td>
<td>2005</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>***</td>
<td>EACAMCCA102*</td>
<td>-</td>
</tr>
</tbody>
</table>
‘GGP is a key regulator of fruit vitamin C concentrations’

**Table 3: QTL mapping for leaf antioxidant concentrations.** Overview of QTL identified for apple foliar L-ascorbic acid (AsA), total AsA (totAsA), dehydroascorbate (DHA), glutathione (GSH), total GSH (totGSH), and oxidized GSH (GSSG) concentrations in 2009. Borderline-significant QTL (2.5<LOD<3.0) are indicated with an asterisk (*).

<table>
<thead>
<tr>
<th>Trait</th>
<th>QTL</th>
<th>LG</th>
<th>LOD</th>
<th>Marker</th>
<th>% variation</th>
<th>QTL</th>
<th>LG</th>
<th>LOD</th>
<th>Marker</th>
<th>% variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AsA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TotAsA</td>
<td>1</td>
<td>6</td>
<td>2.85</td>
<td>EAATMCYT119*</td>
<td>8.3</td>
<td>1</td>
<td>6</td>
<td>2.94</td>
<td>EAATMCYT147*</td>
<td>7.9</td>
</tr>
<tr>
<td>DHA</td>
<td>1</td>
<td>15</td>
<td>3.64</td>
<td>EAGAMCGC115</td>
<td>9.6</td>
<td>1</td>
<td>15</td>
<td>2.84</td>
<td>ECAAMCCT193*</td>
<td>7.4</td>
</tr>
<tr>
<td>GSH</td>
<td>1</td>
<td>12</td>
<td>4.16</td>
<td>EAAMCGC28f4</td>
<td>10.1</td>
<td>1</td>
<td>12</td>
<td>4.13</td>
<td>28f4</td>
<td>10.1</td>
</tr>
<tr>
<td>TotGSH</td>
<td>1</td>
<td>12</td>
<td>3.15</td>
<td>28f4</td>
<td>7.8</td>
<td>1</td>
<td>12</td>
<td>3.34</td>
<td>EAAGMCGG192</td>
<td>9.6</td>
</tr>
<tr>
<td>GSSG</td>
<td>1</td>
<td>16</td>
<td>3.00</td>
<td>EAGAMCC263</td>
<td>7.9</td>
<td>1</td>
<td>6</td>
<td>3.37</td>
<td>CH03d12</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
<td>2.84</td>
<td>EAATMCYT119*</td>
<td>7.0</td>
<td>2</td>
<td>16</td>
<td>2.70</td>
<td>EAGAMCGG1271*</td>
<td>6.4</td>
</tr>
</tbody>
</table>
‘GGP is a key regulator of fruit vitamin C concentrations’

Table 4: Association between fruit total ascorbic acid (totAsA) concentrations and candidate gene allelic variants in the ‘Telamon’ x ‘Braeburn’ (‘TxB’) mapping population. Associations between mean fruit flesh (FL) totAsA concentrations (nmol gFW⁻¹) in progeny of the ‘TxB’ mapping population and allelic variations in MdGGP1 (T→C; SNP_885), MdGGP2 (G→A; SNP_564), MdGGP3 (T→C; SNP_211) and MdNAT7-2 (T→G; SNP_92), across three years of measurements. Statistical significance of the mean fold difference in flesh totAsA concentrations between SNP-genotypes was tested with Student’s t-test (*: P<0.05, **: P<0.01 and ***: P<0.001). Associations were carried out in 71 (2005), 73 (2006) or 85 (2009) individuals of progeny for which both phenotypic and genotypic data were available.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allelic variation</th>
<th>TEL</th>
<th>BR</th>
<th>Year</th>
<th>Mean flesh totAsA (nmol/gFW) of SNP-variants in the 'TxB' population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T/C</td>
</tr>
<tr>
<td><strong>MdGGP1</strong></td>
<td><strong>T→C</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T/C</td>
<td>T/T</td>
<td>2005</td>
<td>446.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006</td>
<td>352.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>561.8</td>
</tr>
<tr>
<td><strong>MdGGP2</strong></td>
<td><strong>G→A</strong></td>
<td>G/G</td>
<td>G/A</td>
<td>2005</td>
<td>480.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006</td>
<td>357.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>630.2</td>
</tr>
<tr>
<td><strong>MdGGP3</strong></td>
<td><strong>T→C</strong></td>
<td>T/C</td>
<td>T/T</td>
<td>2005</td>
<td>423.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006</td>
<td>340.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>620.4</td>
</tr>
<tr>
<td><strong>MdNAT7-2</strong></td>
<td><strong>T→G</strong></td>
<td>G/G</td>
<td>T/G</td>
<td>2005</td>
<td>484.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006</td>
<td>355.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2009</td>
<td>617.9</td>
</tr>
</tbody>
</table>

<sup>a</sup>Genotypes that are homozygous for the ‘T’ allele are named as ‘T/T’ genotypes, while genotypes heterozygous for ‘T’ or ‘C’ allele are named as ‘T/C’ genotypes.

(FL) totAsA concentrations (nmol gFW⁻¹) in progeny of the ‘TxB’ mapping population and allelic variations in MdGGP1 (T→C; SNP_885), MdGGP2 (G→A; SNP_564), MdGGP3 (T→C; SNP_211) and MdNAT7-2 (T→G; SNP_92), across three years of measurements. Statistical significance of the mean fold difference in flesh totAsA concentrations between SNP-genotypes was tested with Student’s t-test (*: P<0.05, **: P<0.01 and ***: P<0.001). Associations were carried out in 71 (2005), 73 (2006) or 85 (2009) individuals of progeny for which both phenotypic and genotypic data were available.
'GGP is a key regulator of fruit vitamin C concentrations'

Table 5: Association between fruit total ascorbic acid (totAsA) concentrations and candidate gene allelic variants in commercial cultivars. Associations between totAsA concentrations (nmol gFW⁻¹) of the entire (flesh and skin) fruit (Davey et al., 2004) and allelic variations of *MdGGP1* (T→C; SNP885) and *MdGGP3* (T→C; SNP211) across 22 apple commercial cultivars, as well as across the three ‘Telamon’ x ‘Braeburn’ individuals with highest or lowest fruit mean totAsA concentration are shown. Candidate gene SNP variants were determined by high-resolution melting analysis using the same SNP-markers developed for ‘Telamon’ and ‘Braeburn’. For *MdGGP3*, the complex melting curve profiles of 11 cultivars prevented SNP-genotyping (-). Mean fruit totAsA concentrations were determined in 2004 for the commercial cultivars (Davey et al., 2004), while the average measurements of three years (2005, 2006 and 2009) were used for the ‘TxB’ individuals with the highest and lowest totAsA concentrations.

<table>
<thead>
<tr>
<th>Cultivar</th>
<th>totAsA (nmol gFW⁻¹)</th>
<th>Alleles</th>
<th>Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>MdGGP1</em> (SNP885)</td>
<td><em>MdGGP3</em> (SNP211)</td>
</tr>
<tr>
<td>1</td>
<td>'Ontario'</td>
<td>1448</td>
<td>T</td>
</tr>
<tr>
<td>2</td>
<td>'Braeburn'</td>
<td>1378</td>
<td>T</td>
</tr>
<tr>
<td>3</td>
<td>'Gravenstein'</td>
<td>1148</td>
<td>T</td>
</tr>
<tr>
<td>4</td>
<td>'Cox’s Orange Pippin'</td>
<td>1132</td>
<td>T</td>
</tr>
<tr>
<td>5</td>
<td>'Golden Delicious'</td>
<td>989.5</td>
<td>T</td>
</tr>
<tr>
<td>6</td>
<td>'Greenstar'</td>
<td>988.3</td>
<td>T</td>
</tr>
<tr>
<td>7</td>
<td>'Goldrush'</td>
<td>987.3</td>
<td>T</td>
</tr>
<tr>
<td>8</td>
<td>'Kanzi'</td>
<td>889.3</td>
<td>T</td>
</tr>
<tr>
<td>9</td>
<td>'Angold'</td>
<td>857.6</td>
<td>C</td>
</tr>
<tr>
<td>10</td>
<td>'Goldrush'</td>
<td>815.0</td>
<td>T</td>
</tr>
<tr>
<td>11</td>
<td>'Jonagold'</td>
<td>798.0</td>
<td>T</td>
</tr>
<tr>
<td>12</td>
<td>'Idared'</td>
<td>767.0</td>
<td>T</td>
</tr>
<tr>
<td>13</td>
<td>'Florina'</td>
<td>741.1</td>
<td>T</td>
</tr>
<tr>
<td>14</td>
<td>'Liberty'</td>
<td>728.8</td>
<td>C</td>
</tr>
<tr>
<td>15</td>
<td>'Pink Lady®'</td>
<td>651.1</td>
<td>C</td>
</tr>
<tr>
<td>16</td>
<td>'Fuji'</td>
<td>628.0</td>
<td>C</td>
</tr>
<tr>
<td>17</td>
<td>'Elstar'</td>
<td>545.6</td>
<td>C</td>
</tr>
<tr>
<td>18</td>
<td>'Alkmene'</td>
<td>461.7</td>
<td>C</td>
</tr>
<tr>
<td>19</td>
<td>'Prima'</td>
<td>443.8</td>
<td>C</td>
</tr>
<tr>
<td>20</td>
<td>'Merlijn'</td>
<td>442.3</td>
<td>T</td>
</tr>
<tr>
<td>21</td>
<td>'Gala'</td>
<td>437.7</td>
<td>C</td>
</tr>
<tr>
<td>22</td>
<td>'TxB 299'</td>
<td>401.4</td>
<td>C</td>
</tr>
<tr>
<td>23</td>
<td>'Priscilla'</td>
<td>400.9</td>
<td>C</td>
</tr>
<tr>
<td>24</td>
<td>'TxB 126'</td>
<td>387.8</td>
<td>C</td>
</tr>
<tr>
<td>25</td>
<td>'Telamon'</td>
<td>233.4</td>
<td>C</td>
</tr>
</tbody>
</table>

Mean totAsA of cultivars carrying T = 997.3 ± 1114
Mean totAsA of cultivars carrying C = 510.4 ± 528.9
Mean fold difference = 2
Maximum fold difference = 6

Mean totAsA of cultivars carrying T+T = 1157
Mean totAsA of cultivars carrying C+C = 437.8
Mean fold difference = 3
Maximum fold difference = 6

* totAsA concentration of 'Telamon' measured in 2006
Figure A: Total AsA flesh (nmol/gFW).

Figure B: Relative expression of MdGGP1.

Figure C: Relative expression of MdGGP2.

Figure D: Relative expression of MdGGP3.

Figure E: Relative expression of MdDHAR3.2.

Figure F: Relative expression of MdNADH-2.

Figure G: Expression of MdEF1-4A.1.