Evolutionary diversification of galactinol synthases in Rosaceae: adaptive roles of galactinol and raffinose during apple bud dormancy

Vítor da Silveira Falavigna1,*, Diogo Denardi Porto2, Yohanna Evelyn Miotto3, Henrique Pessoa dos Santos3, Paulo Ricardo Dias de Oliveira3, Márcia Margis-Pinheiro1, Giancarlo Pasquali1 and Luís Fernando Revers3,†

1 Graduate Program in Cell and Molecular Biology, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91501–970, Brazil
2 Embrapa Semiárido, Petrolina, PE, 56302–970, Brazil
3 Embrapa Uva e Vinho, Bento Gonçalves, RS, 95701-008, Brazil

* Present address: AGAP, Univ. Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France.
† Correspondence: luis.revers@embrapa.br

Received 27 September 2017; Editorial decision 27 November 2017; Accepted 27 November 2017

Editor: Robert Hancock, The James Hutton Institute, UK

Abstract

Galactinol synthase (GolS) is a key enzyme in the biosynthetic pathway of raffinose family oligosaccharides (RFOs), which play roles in carbon storage, signal transduction, and osmoprotection. The present work assessed the evolutionary history of GolS genes across the Rosaceae using several bioinformatic tools. Apple (Malus × domestica) GolS genes were transcriptionally characterized during bud dormancy, in parallel with galactinol and raffinose measurements. Additionally, MdGolS2, a candidate to regulate seasonal galactinol and RFO content during apple bud dormancy, was functionally characterized in Arabidopsis. Evolutionary analyses revealed that whole genome duplications have driven GolS gene evolution and diversification in Rosaceae speciation. The strong purifying selection identified in duplicated GolS genes suggests that differential gene expression might define gene function better than protein structure. Interestingly, MdGolS2 was differentially expressed during bud dormancy, concomitantly with the highest galactinol and raffinose levels. One of the intrinsic adaptive features of bud dormancy is limited availability of free water; therefore, we generated transgenic Arabidopsis plants expressing MdGolS2. They showed higher galactinol and raffinose contents and increased tolerance to water deficit. Our results suggest that MdGolS2 is the major GolS responsible for RFO accumulation during apple dormancy, and these carbohydrates help to protect dormant buds against limited water supply.

Keywords: Apple, bud dormancy, galactinol synthase, Malus × domestica, raffinose family oligosaccharides, tolerance to water deficit, transgenic expression, whole-genome duplication.

Introduction

Gene and genome duplication events have yielded the raw genetic material for biological evolution, with a major impact on most eukaryotes, including plants (Zhang, 2003). Whole-genome duplication (WGD) has been a common feature during angiosperm speciation and is considered to be a driving force in evolution due to the innovations associated
with these events (Jiao et al., 2011). In flowering plants, several rounds of WGD have already been traced, including an ancient event that occurred 200 million years ago (MYA) and was shared even by basal angiosperms such as Amborella trichopoda (Amborella Genome Project, 2013). In the particular case of eudicots, a triplication event that occurred about 125 MYA has shaped their genomes, and contributed to molecular and phylogenetic diversification (Jiao et al., 2012).

Apple (Malus × domestica Borkh.), one of the most important fruit crops in temperate regions worldwide, belongs to the Rosaceae family and was subject to these duplication events during its evolution. Moreover, apples also underwent a recent WGD event around 50 MYA that is restricted to the Pyreeae clade, which also encompasses pear and various other Rosaceae genera (Velasco et al., 2010; Wu et al., 2013; Xiang et al., 2017). These events triggered the emergence of specific subclades of genes, resulting in unique developmental and metabolic traits.

Temperate fruit crops such as apple are subject to alternating growth and rest periods through the year-round development cycle. The rest period or dormancy is an adaptive trait for plant survival under adverse environmental conditions, and a well-adjusted dormancy cycle is crucial for the full achievement of the plant’s genetic potential (Falavigna et al., 2015b). This process is characterized by the formation of stress-tolerant specialized bud structures that undergo a developmental program for the protection of meristems (Rohde and Bhalerao, 2007). Recent findings indicated that C-repeat binding transcription factors (CBFs), major integrators of the cold response and freezing tolerance, may also modulate dormancy in apple (Wisniewski et al., 2015; Artlip et al., 2016). Besides genes responsible for the molecular control of dormancy progression, a series of proteins and metabolites were also shown to be recruited for the protection of bud integrity during winter. Dehydrins, sugars, detoxification enzymes, chaperones, protein kinases, and other compounds stand out in the process (Maruyama et al., 2009; Eremina et al., 2016). Supporting these findings, studies carried out during bud dormancy have demonstrated the induction of many cold- and dehydration-related genes that integrate the CBF regulation pathway, with galactinol synthases (GolS; EC 2.4.1.123) being recurrent in several reports (Derory et al., 2006; Santamaría et al., 2011; Liu et al., 2012; Falavigna et al., 2014; Andersen et al., 2017).

GolS proteins integrate a subfamily of the GT8 family of glycosyltransferases that are involved in the synthesis of diverse sugar conjugates with important structural, storage, and signaling roles, as well as acting as an energy source (Sengupta et al., 2012). GolS is a key enzyme catalysing the first committed step in the biosynthesis of raffinose family oligosaccharides (RFOs), being responsible for galactinol formation using UDP-galactose and myo-inositol as substrates (Elsayed et al., 2014). Galactinol is a building block used in the synthesis of raffinose, stachyose, and other RFOs. These sugars have been described as accumulating in response to cold, freezing, and water deficit (Sengupta et al., 2015). Moreover, galactinol and RFOs are compatible solutes that play roles in carbon storage, osmotic adjustments, signal transduction, and membrane and protein stabilization, among others (Krasensky and Jonak, 2012; Elsayed et al., 2014; Sengupta et al., 2015). The accumulation of galactinol and RFOs during the last stages of seed development is responsible for protecting the embryo during seed desiccation, being necessary for proper seed germination in pea (Blöchl et al., 2007). Finally, GolS induction during bud dormancy progression likely results in the seasonal mobilization of RFOs that, together with several other proteins and metabolites, may lead to bud protection during winter (Santamaría et al., 2011; Unda et al., 2012; Ibáñez et al., 2013; Falavigna et al., 2014; Andersen et al., 2017).

In the present work, we assessed the evolutionary history of GolS genes in the Rosaceae family and identified that WGD events have driven the evolution and diversification of these genes in Rosaceae speciation. However, duplicated genes showed limited functional divergence, suggesting that differential gene expression might better define gene function rather than protein structure. Therefore, apple GolS gene expression was characterized during bud dormancy and in several developmental stages, and in fact, MdGolS2 genes presented differential transcript accumulation patterns. Given that one of the intrinsic adaptive features of bud dormancy is limited availability of free water, MdGolS2, a strong candidate to regulate seasonal galactinol and RFO content during bud dormancy, was functionally characterized in Arabidopsis. The data gathered in this work demonstrate that MdGolS2 is a major contributor to galactinol and RFO accumulation during apple bud dormancy, and that these carbohydrates act in the protection of bud integrity during adverse environmental conditions.

**Materials and methods**

**Plant material**

Apple samples were harvested in two experimental orchards located in the cities of Vacaria (~28.513777, −50.881465 and 972 m altitude) and Caçador (~26.836971, −50.975246 and 935 m altitude), Southern Brazil. In Vacaria, 3-year-old ‘Gala Baigent’ trees grafted on ‘Marubakaido’ rootstocks with ‘M.9’ as interstock were arranged in three blocks of 10 plants each. Six developmental stages were partitioned in different tissues and organs resulting in 13 samples according to the Fleckinger scale (EPPO, 1984): dormant buds (A stage; 21 July 2009); buds at initial bursting (C stage; 4 September 2009); flower buds and young leaves (E2 stage; 12 September 2009); 10 mm-diameter whole set-fruits and associated leaves (I stage; 28 October 2009); 40 mm-diameter unripe fruits divided into pulp, seed and skin, as well as leaves (J stage; 8 December 2009); Mature fruits of 70 mm diameter were also sampled and partitioned into pulp, seed, and skin (M stage; 2 February 2010). See Falavigna et al. (2015a) for images of the developmental stages. Samples were immediately frozen in liquid nitrogen in the field and stored at −80 °C until use. In Caçador, 7-year-old ‘Fuji Standard’ trees grafted on ‘Marubakaido’ rootstocks with ‘M.9’ as interstock were arranged in three blocks of 10 plants each. Forty closed terminal buds from each plant were sampled at eight time points from January 2009 to February 2010 always at 11.00 h. Samples were immediately frozen in liquid nitrogen in the field and stored at −80 °C until use. Sampling dates and corresponding chilling hours accumulated by these samples are depicted in Supplementary Table S1 at JXB online.
Gene expression analysis
DNA was isolated from 200 mg mature leaves of ‘ Gala Baigent’ apple trees using modified protocols adapted to 2 ml tubes (Lodhi et al., 1994; Lefort and Douglas, 1999). Total RNA was isolated as described in Zeng & Yang (2002) and Falavigna et al. (2014), and DNase-treated using the TURBO DNA-free Kit (Ambion, Austin, TX, USA). The GeneAmp RNA PCR Core Kit (Thermo Fisher Scientific, Waltham, MA, USA) was used for cDNA synthesis according to manufacturer’s instructions. DNA-free RNA samples from mature seeds and dormant buds were also employed in 5’ and 3’ rapid amplification of cDNA ends (RACE) reactions using the SMARTer RACE cDNA Amplification Kit (Clontech, Mountain View, CA, USA) according to the manufacturer’s protocol. RACE products were sequenced at ACTGene Ltd (Porto Alegre/RS, Brazil) using an automatic ABI-PRISM 3100 Genetic Analyzer and associated chemistry (Thermo Fisher Scientific). Real-time PCR was performed as described (Falavigna et al., 2015a). PCR efficiency and mean relative gene expression were calculated using LinRegPCR v.2012.0 (Ruijter et al., 2009) and the 2^-ΔΔCt method (Livak and Schmittgen, 2001), respectively. ARCS (accumulation and replication of chloroplast 5), MDH (malate dehydrogenase) and WD40 (transcription factor WD40-like repeat domain) or ARCS, MDH and Tmp1 (type 1 membrane protein-like) were used as reference genes for closed terminal buds or for organ/tissue samples, respectively (Perini et al., 2014).

Carbohydrate analysis
Carbohydrates were extracted using protocols modified from Amaral et al. (2007) and Eyles et al. (2013). Briefly, galactinol, raffinose and sucrose levels were determined from 100 mg pulverized closed terminal buds from ‘Fuji Standard’. Soluble carbohydrates were extracted with 1 ml 80% ethanol in an 80 °C water bath for 30 min. After centrifugation (10,000 × g for 5 min), the supernatant was stored and samples were re-extracted five times. Three independent extractions were conducted for each biological replicate. In order to measure extraction efficiency, an additional independent extraction was carried out adding 100 mg ^1° of each analyte as internal standards. After extraction, a purification step was performed to remove lesser polar compounds. Volumes of 500 µl of chloroform and 400 µl of water were added to 1 ml of total supernatant followed by vigorous shaking. After centrifugation (13,000 × g for 5 min), the supernatant was collected and filtered (0.22 µm). An UPLC (Waters, Acquity UPLC system, Milford, MA, USA) coupled to a triple quadrupole mass spectrometer (MS Waters Xevo TQ) was used for separation and quantification of carbohydrates. The chromatographic separation was performed using a combination of a pre-column (2.1 × 5 mm, 1.7 µm particles) and a column (2.1 × 100 mm, 1.7 µm particles) (Acquity UPLC BEH amide). The mobile phase consisted of (A) 0.05% ammonia in water and (B) 0.05% ammonia in acetonitrile. A linear gradient condition from 20% A–80% B to 35% A–65% B at 0.05% ammonia in water and (B) 0.05% ammonia in acetonitrile. The mobile phase consisted of (A) 0.05% ammonia in water and (B) 0.05% ammonia in acetonitrile. The column was held at 35 °C and the sample compartment at 7 °C. The flow rate was fixed at 0.3 ml min^-1. Under these conditions, retention times were: 8 min was followed by the initial condition for an additional 7 min. The column was held at 35 °C and the sample compartment at 7 °C. The flow rate was fixed at 0.3 ml min^-1. Under these conditions, retention times were: 8 min was followed by the initial condition for an additional 7 min. The flow rate was fixed at 0.3 ml min^-1. Under these conditions, retention times were: 8 min was followed by the initial condition for an additional 7 min.

Arabidopsis transformation
With the aim of generating transgenic plants expressing MgGolS2, MgGolS2’s complete coding sequence was amplified with high fidelity enzymes (Thermo Fisher Scientific) using gene-specific primers (Supplementary Table S2) and cloned into pENTR™ Directional TOPO® (Thermo Fisher Scientific). This construction was used to clone its coding DNA sequence (CDS) into the pH7WG2D.1 vector under the CaMV 35S promoter (Karimi et al., 2007) using Gateway® LR Clonase™ II Enzyme Mix (Thermo Fisher Scientific) according to the manufacturer’s instructions. The resulting vector was confirmed by sequencing and transformed into Arabidopsis Col-0 plants using the Agrobacterium tumefaciens (EHA105) floral dip method (Clough and Bent, 1998). Transformed seeds were sown in plastic pots and 3-week-old plants were evaluated for green fluorescence protein (GFP) fluorescence using the reporter EGF gene present in the vector (Karimi et al., 2007), using a Leica M165FC stereomicroscope. Plants without GFP fluorescence were discarded. Total RNA was isolated from leaves, and real-time PCR was performed using MgGolS2 gene-specific primers as previously described. The amplicon obtained from each line was sequenced to confirm its transgenic nature. Plants without the transgene were also discarded. Five independent lines were randomly chosen for further analysis. A real-time PCR (actin 2) and ACO1 (constitutive photomorphogenic 1) were employed as reference genes (Supplementary Table S2). Galactinol and raffinose levels were determined as previously described. Student’s t-test was used to evaluate the statistical significance of differences in carbohydrate levels between transgenic lines and wild-type plants.

Water withhold assay
A preliminary water deficit assay was conducted to identify the number of days needed to cause lethal damage in wild-type Arabidopsis plants under our growth conditions. Prior to the water withhold assay, all pots have had their pot capacity attained and had their weight adjusted to approximately 160 g (soil, plant, pot and water). Two-month-old transformed and non-transformed plants were transferred to vats without water for 5 weeks. After treatment, plants were rehydrated for 10 days prior to evaluation. Pots were weighed once a week to measure water loss. Photographs were taken for visual analysis.

In silico analysis
In order to identify predicted gene models coding for GolS in the apple doubled-haploid genome version 1.1 (Daccord et al., 2017), default parameters were employed in Blastp searches using a sequence previously identified by our group as query (ABD15 in Falavigna et al., 2014; Altschul et al., 1990). Resulting sequences were annotated by comparison with the NCBI non-redundant protein database using Blast2GO version 4.1.0 (E-value cutoff of 1e^-5; Conesa et al., 2005) and then screened for the presence of the GT8 domain (PF01501) in the SMART database (Letunic et al., 2012). Primers were designed using Primer3 v0.4.0 (Supplementary Table S2; Rozen & Skaltsky, 2000). The exon/intron structures of MgGolS genes were compiled into a figure using the Gene Structure Display Server 2.0 (Hu et al., 2015). The available annotated genomes from Rosaceae species in GigaDB (http://gigadb.org/), Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html), NCBI (http://www.ncbi.nlm.nih.gov/) and GDR (http://rosaceae.org/) databases were screened for the presence of GolS proteins using the same strategy described above (accessed between sampling points using Prism 5.0a (GraphPad Software, La Jolla, CA, USA). To investigate the correlation between MgGolS transcriptional level and carbohydrate content, Pearson’s correlation coefficient analysis at a 0.01 significance level was conducted using GraphPad Prism.
in June 2017). All identified sequences were trimmed as previously described. Resulting GolS amino acid sequences from apple, black raspberry, Chinese white pear, European pear, Japanese apricot, peach, strawberry, and woodland strawberry (Supplementary Table S3) were aligned using MUSCLE (Edgar, 2004) and phylogenetic relationships were inferred using MrBayes version 3.2.6 (Ronquist et al., 2012). The mixed amino acid substitution model was used in the default settings and 150 thousand generations were run, sampled every 100 generations, with the first 25% of trees discarded as burn-in. The remaining ones were summarized in a consensus tree that was visualized and edited in FigTree version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Collinear block analyses were performed by intra- and interspecific genome-wide protein sequence comparison for Rosaceous species using Blastp (E-value < 1e−10; Altschul et al., 1990). Results and gene positions were used as inputs to determine collinear blocks using MCScanX (Wang et al., 2012). The non-synonymous substitution rate (Ks) and synonymous substitution rate (Ks) of paralog genes were calculated using downstream analysis tools implemented in the MCScanX software. Ks dating was calculated from collinear blocks containing paralogous GolS genes. To this purpose, six genes flanking each GolS paralog were used. If fewer than 12 homologous genes were found, then all available genes were used instead. The type of gene duplication event was determined using the duplicate_gene_classifier algorithm implemented in MCScanX (Wang et al., 2012).

Results

GolS genes in apple

The apple genome was screened for GolS sequences by means of Blastp searches. All resulting sequences were compared with SMART and NCBI non-redundant protein databases, and the ones lacking the GT8 domain (PF01501) or not annotated as GolS-like, respectively, were excluded. This led to the identification of eight predicted gene models coding for GolS in apple (Table 1). However, only five of them were successfully amplified when using a pool of cDNAs from 13 different apple tissues and organs (see Supplementary Fig. S1). These five genes were named MdGolS1–5, and were amplified by RACE and sequenced (Table 1). MdGolS1, 2, 3, and 4 cDNA sequences showed differences in comparison with the genomic sequences: one nucleotide mismatch without amino acid change, a misprediction in the second exon that resulted in 27 additional nucleotides, three nucleotide mismatches leading to two amino acid changes, and 17

| Gene name | Genome accession | GenBank accession | Chromosomal location | Protein size (aa) |
|-----------|-------------------|------------------|----------------------|------------------|
| MdGolS1   | MD04G1190900      | KYY75539         | Chr4: 28069630–28071554 | 333              |
| MdGolS2   | MD13G1093700      | KYY75540         | Chr13: 6602655–6604528 | 337              |
| MdGolS3   | MD13G1147700      | KYY75541         | Chr13: 11543866–11546147 | 328              |
| MdGolS4   | MD17G1280400      | KYY75542         | Chr17: 34045432–34047056 | 348              |
| MdGolS5   | MD16G1095000      | KYY75543         | Chr16: 6573774–6575140 | 302              |
| MdGolS6   | MD09G1288100      | —                | Chr9: 36789422–36792275 | 363              |
| MdGolS7   | MD16G1147600      | —                | Chr16: 11523700–11526050 | 298              |
| MdGolS8   | MD11G1070000      | —                | Chr11: 5983858–5985447 | 359              |

Genome and GenBank accession codes are provided by the ‘The Apple Genome and Epigenome’ (https://iris.angers.inra.fr/gddh13/) and NCBI (http://www.ncbi.nlm.nih.gov/) databases, respectively.

The apple genome was screened for GolS sequences by means of Blastp searches. All resulting sequences were compared with SMART and NCBI non-redundant protein databases, and the ones lacking the GT8 domain (PF01501) or not annotated as GolS-like, respectively, were excluded. This led to the identification of eight predicted gene models coding for GolS in apple (Table 1). However, only five of them were successfully amplified when using a pool of cDNAs from 13 different apple tissues and organs (see Supplementary Fig. S1). These five genes were named MdGolS1–5, and were amplified by RACE and sequenced (Table 1). MdGolS1, 2, 3, and 4 cDNA sequences showed differences in comparison with the genomic sequences: one nucleotide mismatch without amino acid change, a misprediction in the second exon that resulted in 27 additional nucleotides, three nucleotide mismatches leading to two amino acid changes, and 17 mismatches leading to four amino acid changes, respectively (Supplementary Fig. S2).

The structures of MdGolS genes were retrieved from the apple genome and curated with our sequencing data, when available (Fig. 1). All MdGolS genes presented four exons, except MdGolS4 and MdGolS8 with three and five exons, respectively. All sequences contained one GT8 domain, except MdGolS7 that presented two. However, both domains in MdGolS7 were truncated. MdGolS6 also presented a truncated version of the GT8 domain.

GolS evolution among Rosaceae

GolS has already been characterized in a wide phylogenetic scope from green algae to plants (Yin et al., 2010; Sengupta et al., 2012). In this work, a deeper characterization was performed focusing on the Rosaceae family, which integrates the Rosales order and includes important commercial fruit species (Xiang et al., 2017). After an initial screening for the presence of GolS genes in the available Rosaceous genomes, retrieved sequences were trimmed, resulting in 52 genes putatively coding for GolS (Supplementary Table S3). Except for strawberry, which is octoploid, all species herein analysed are diploids. The highest number of GolS members was identified in apple and European pear, with eight and 12 members, respectively. Strawberry presented seven members, while black raspberry, Chinese pear, Japanese apricot, peach, and woodland strawberry displayed five members each. The number of GolS genes in Chinese pear should be analysed with caution, given that apple and European pear, species evolutionarily close to Chinese pear, presented a higher number of members.

A phylogenetic tree was constructed aiming to identify evolutionary relationships of GolS proteins within the Rosaceae family (Fig. 2). No apple paralogs were identified, and apple orthologs were only identified along with European pear sequences: MdGolS2 with PcGolS4, MdGolS7 with PcGolS12, and MdGolS8 with PcGolS6. MdGolS1 formed an independent cluster with European and Chinese pear orthologs. A similar result was observed for MdGolS3 and MdGolS4. MdGolS5 was present in a cluster with European pear paralogs, while MdGolS6 was present in a cluster with GolS sequences from apple, Chinese and European pear.

Table 1. Identification of GolS genes in apple

| Gene name | Genome accession | GenBank accession | Chromosomal location | Protein size (aa) |
|-----------|-------------------|------------------|----------------------|------------------|
| MdGolS1   | MD04G1190900      | KYY75539         | Chr4: 28069630–28071554 | 333              |
| MdGolS2   | MD13G1093700      | KYY75540         | Chr13: 6602655–6604528 | 337              |
| MdGolS3   | MD13G1147700      | KYY75541         | Chr13: 11543866–11546147 | 328              |
| MdGolS4   | MD17G1280400      | KYY75542         | Chr17: 34045432–34047056 | 348              |
| MdGolS5   | MD16G1095000      | KYY75543         | Chr16: 6573774–6575140 | 302              |
| MdGolS6   | MD09G1288100      | —                | Chr9: 36789422–36792275 | 363              |
| MdGolS7   | MD16G1147600      | —                | Chr16: 11523700–11526050 | 298              |
| MdGolS8   | MD11G1070000      | —                | Chr11: 5983858–5985447 | 359              |
Japanese apricot and peach sequences grouped as orthologs. The same trend was observed for sequences from black raspberry, strawberry, and woodland strawberry.

In order to trace the GolS evolutionary history within the Rosaceae family, collinearity and Ks-dating analyses were performed. Our strategy was based on the comparison of the apple genome with those from black raspberry, Chinese pear, European pear, Japanese apricot, peach, strawberry, and woodland strawberry. These comparisons generated several collinear regions containing GolS genes, allowing us to identify several paralogous pairs (Table 2) as well as several orthologs of MdGolS genes (Table 3). A graphical representation of these findings is present in Supplementary Fig. S3. Within the apple genome, eight collinear blocks containing GolS genes were identified, between chromosomes 4 and 9 (MdGolS1/MdGolS6), 4 and 17 (MdGolS1/MdGolS4), 9 and 16 (MdGolS3/MdGolS4 and MdGolS6/MdGolS7), 9 and 17 (MdGolS4/MdGolS6), 13 and 16 (MdGolS2/MdGolS5 and MdGolS3/MdGolS7), and 16 and 17 (MdGolS4/MdGolS7).

Some of these blocks are syntenic, given that they share the same Pyreae ancestors such as Mdchr9 and 17 and Mdchr13 and 16 (Velasco et al., 2010; Daccord et al., 2017). GolS paralogs were also identified for sequences from European pear, Japanese apricot, peach, and woodland strawberry (Table 2). Although no collinear blocks containing GolS genes were found between Chinese pear chromosomes, eight blocks were identified linking this genome to apple (Supplementary Fig. S3A; Table 3). Despite the European pear genome not being assembled on chromosomes, the comparison with apple rendered the highest number of collinear blocks (12), with each apple MdGolS gene having at least one counterpart in European pear (Supplementary Fig. S3B; Table 3). Japanese apricot and peach showed ten collinear blocks each with apple, besides several collinear regions containing GolS genes within each genome (Supplementary Fig. S3C, D; Tables 2 and 3). Finally, the collinearity comparison between apple and woodland strawberry rendered seven collinear blocks between these species (Supplementary Fig. S3E; Table 3). It is worth mentioning that the black raspberry and strawberry genomes are not
assembled in chromosomes, which resulted in the presence of few collinear blocks for these species.

The $K_s$-dating analysis aimed to estimate the time of gene duplications for all collinear blocks containing paralogous $Gols$ genes in the Rosaceae family (Table 2). All Rosaceous paralogs had their duplicated gene origin annotated as WGD, suggesting that segmental duplications were the main contributory force to the expansion of $Gols$ genes in these genomes. Only apple and European pear paralogs showed $K_s$ values lower than 0.25, which may correspond to the recent WGD. All $Gols$ paralogs from Japanese apricot, peach, and woodland strawberry showed $K_s$ values ranging from 1.1 to 1.8. The other five apple paralogs, $Mdgols1/Mdgols4$, $Mdgols1/Mdgols6$, $Mdgols5/Mdgols6$, and $Mdgols6/Mdgols7$, also presented $K_s$ values higher than 1. In order to identify if $K_s$ values higher than 1 corresponded to the triplication event common to all eudicots, $Gols$ genes were identified in the grapevine genome, followed by collinearity and $K_s$-dating analyses. The grapevine genome is a powerful tool for this kind of comparison considering that it has only undergone the triplication event, without subsequent polyploidies (Jaillon et al., 2007; Tang et al., 2008).

Twelve $Gols$ genes were identified in grapevine, with six collinear blocks containing $Gols$ genes that originated in WGD events. The $K_s$ dating of these collinear blocks ranged from 1 to 1.7 (Supplementary Table S4). However, a region of 41 kbp in the grapevine chromosome 14 presented six grapevine $Gols$ genes arranged in tandem. This suggests that tandem duplications can also be held accountable for the number of $Gols$ genes in grapevine, although this kind of duplication event is not present in any Rosaceous species.

Finally, the $K_s/K_s$ ratio was calculated for all Rosaceous $Gols$ paralogs. This ratio is used to identify positive selection, which is indicated by values higher than 1. A $K_s/K_s$ ratio lower than 1 indicates negative selection, whereas a ratio equal to 1 indicates neutral selection. All $K_s/K_s$ ratios were lower than 0.41 (Table 2), suggesting that these genes experienced purifying selection pressures with limited functional divergence after WGDs.

### Table 2. $K_s$-dating and $K_s/K_s$ ratios of $Gols$ paralogs from different Rosaceous species

| Duplicated genes | $K_s$ dating of the collinear block | $K_s/K_s$ ratio | Duplicated gene origin |
|-------------------|------------------------------------|----------------|------------------------|
| FvGols3/FvGols5   | 1.12                               | 0.36           | WGD                    |
| Mdgols1/Mdgols4   | 1.53                               | 0.13           | WGD                    |
| Mdgols1/Mdgols6   | 1.54                               | 0.16           | WGD                    |
| Mdgols2/Mdgols5   | 0.25                               | 0.17           | WGD                    |
| Mdgols3/Mdgols7   | 0.21                               | 0.41           | WGD                    |
| Mdgols4/Mdgols6   | 0.15                               | 0.25           | WGD                    |
| Mdgols4/Mdgols7   | 1.37                               | —              | WGD                    |
| Mdgols5/Mdgols6   | 1.98                               | 0.19           | WGD                    |
| Mdgols6/Mdgols7   | 1.73                               | 0.10           | WGD                    |
| Pcgols1/Pcgols4   | 0.18                               | 0.14           | WGD                    |
| Pcgols11/Pcgols12 | 0.18                               | 0.32           | WGD                    |
| Mpgols1/Mpgols3   | 1.16                               | 0.10           | WGD                    |
| Mpgols2/Mpgols3   | 1.46                               | 0.14           | WGD                    |
| Mpgols2/Mpgols4   | 1.57                               | —              | WGD                    |
| Mpgols3/Mpgols5   | 1.56                               | 0.15           | WGD                    |
| Ppgols1/Ppgols4   | 1.15                               | 0.12           | WGD                    |
| Ppgols2/Ppgols3   | 1.80                               | 0.05           | WGD                    |

Fv, Fragaria vesca; $K_s$, non-synonymous substitution rate; $K_s$, synonymous substitution rate; Md, Malus × domestica; Pc, Pyrus communis; Pm, Prunus mume; Pp, Prunus persica; WGD, whole-genome duplication.

### Table 3. List of collinear blocks among six Rosaceous genomes containing orthologs of Mdgols genes

| Apple          | Chinese pear | European pear | Japanese apricot | Peach | Woodland strawberry |
|----------------|--------------|---------------|------------------|-------|---------------------|
| Mdgols1        | Pbgols4      | Pcgols9       | Pmgonols1        | Ppgols4 | FvGols1             |
| Mdgols2        | —            | Pcgols1       | Pmgonols2        | Ppgols2 | —                   |
| Mdgols3        | Pbgols3      | Pcgols11      | Pmgonols1        | Ppgols1 | FvGols1             |
| Mdgols4        | Pbgols1      | Pcgols5       | Pmgonols4        | Ppgols3 | FvGols3             |
| Mdgols5        | —            | Pcgols4       | Pmgonols2        | Ppgols4 | FvGols5             |
| Mdgols6        | Pbgols1      | Pcgols5       | Pmgonols4        | Ppgols3 | FvGols3             |
| Mdgols7        | Pbgols1      | Pcgols11      | Pmgonols3        | Ppgols3 | FvGols3             |
| Mdgols8        | —            | Pcgols6       | —                | —     | —                   |

| A schematic representation of the collinear block analysis is present in the Supplementary Figure S3. |

### Expression patterns of Mdgols genes in apple tissues and organs

Transcript level analysis of $Mdgols$ genes was conducted via real-time PCR in 13 apple tissues and organs during the progression of bud dormancy, flowering and fruit ripening. Interestingly, $Mdgols$ genes showed remarkable differential expression patterns among the analysed samples. $Mdgols1$, 3, and 4 were mainly expressed in mature seeds, while $Mdgols2$ showed higher mRNA concentrations in dormant buds in comparison with other samples (Fig. 3). $Mdgols1$, 2, and 3 also presented the highest transcript levels among all samplings and genes analysed. $Mdgols3$ showed higher transcript levels in buds at initial bursting and leaves from I and J stages, although it presented the lowest transcript levels over samples in comparison with other $Mdgols$ genes. Additionally, $Mdgols4$ transcripts were not detected in dormant buds, and $Mdgols5$ could not be detected in mature fruits. No transcripts were detected for $Mdgols6$–8 in any of the tissues or organs analysed (Supplementary Fig. S1).

$Mdgols$ gene expression and sugar content analyses in apple terminal buds during an annual growth and dormancy cycle

Closed terminal buds from field-grown ‘Fuji Standard’ apple trees were sampled during a complete annual growth and dormancy cycle from summer 2009 to summer 2010 in southern Brazil. Chilling exposure, and growth cessation and resumption (50% of buds in green tip stage) were followed...
in order to determine dormancy establishment, fulfillment and release (Supplementary Table S1). The gene expression of \textit{MdGolS1}−5 was quantified by real-time PCR throughout these samplings (Fig. 4A). \textit{MdGolS1} and 2 displayed a seasonal transcript accumulation during dormancy, with higher amounts in winter buds. Additionally, both genes were down-regulated near budbreak (15 September), with lower transcript levels during spring and summer. It is worth mentioning that \textit{MdGolS2} was 8-fold more expressed during winter than \textit{MdGolS1}. \textit{MdGolS2} reached the highest expression level of all \textit{MdGolS} genes in mid-winter (30 July). Contrastingly, \textit{MdGolS3} and 5 presented low transcriptional variation along the year. Finally, \textit{MdGolS4} expression was completely abolished during dormancy entrance (27 May, autumn) and maintenance (30 June, winter), restoring its transcriptional levels during mid-winter (30 July), spring, and summer.

The seasonal expression profile observed for some \textit{MdGolS} genes prompted us to quantify carbohydrate levels during the same annual sampling cycle of ‘Fuji Standard’ apple buds (Fig. 4B). The sugar extraction efficiency was above 95% for all three analytes. Galactinol and raffinose started to accumulate during dormancy establishment (May, autumn), reaching their highest amounts during deep dormancy (June and July, winter), and restoring low levels after dormancy completion (November 2009 to February 2010, spring and summer). Raffinose levels were approximately 180-fold higher in samples harvested in July in comparison with summer samples, while galactinol levels were approximately 9-fold higher in the same comparison. Moreover, the transcript level of \textit{MdGolS1} and \textit{MdGolS2} along with galactinol and raffinose content during the year showed a strong positive correlation (Table 4). Finally, sucrose levels were higher during winter (June, July, and September) in comparison to autumn, spring and summer samples. A three-fold difference was observed in sucrose levels in the most contrasting points, September and March 2009.

Water deficit assay of transgenic Arabidopsis plants expressing \textit{MdGolS2}

The seasonal accumulation of \textit{MdGolS2} transcripts, combined with the accumulation of galactinol and raffinose during winter, suggests that these metabolites may play important roles during dormancy. Their probable function may be related to increased dormant bud tolerance to dehydration, given that one of the characteristic features of the dormancy process is limited availability of free water (de Fayé et al., 2000; Améglio et al., 2002; Kalcsits et al., 2009). Within this context, to gain insights about \textit{MdGolS2} function, Arabidopsis plants expressing this gene were generated and exposed to a water deficit assay. This assay aimed to demonstrate the functionality of the \textit{MdGolS2} enzyme \textit{in vivo} as well as to test if galactinol and RFOs are able to protect tissues exposed to limited water amounts, mimicking a situation that occurs in apple buds during dormancy (Kalcsits et al., 2009).

A preliminary water deficit assay was performed in wild-type plants to identify the number of days needed to cause lethal damage. After 24–27 days without water, wild-type plants were unable to survive even after rehydration. Five independent transgenic lines were recovered and analysed (Fig. 5A). Galactinol and raffinose contents were measured before treatment, and wild-type plants showed lower amounts of galactinol and raffinose in comparison with transgenic plants (Fig. 5B). Two-month-old transformed and wild-type plants had their water supply removed for 35 days to perform the water deficit assay. Water loss was independently measured in each pot once a week. Around 10% of water content was lost every week during the treatment, with starting values recovered after rehydration (Fig. 5C). Only the transgenic lines survived after rehydration (Fig. 5D). All transgenic lines followed their normal growth cycle after treatment, producing viable flowers, siliques, and seeds.
Galactinol synthase is a key enzyme in the biosynthetic pathway of RFOs, which are carbohydrates that play important roles in carbon storage, signal transduction, and osmoprotection, among others (Elsayed et al., 2014; Sengupta et al., 2015). Although several studies identified GolS genes and RFO levels as up-regulated during bud dormancy, suggesting that the seasonal mobilization of these carbohydrates could lead to bud integrity protection (Cox and Stushnoff, 2001; Derory et al., 2006; Santamaria et al., 2011; Liu et al., 2012; Unda et al., 2012; Ibáñez et al., 2013; Falavigna et al., 2014; Yue et al., 2015; Andersen et al., 2017), their direct relationship had still not been assessed. Within this context, evolutionary analyses of GolS genes in the Rosaceae family were performed followed by the characterization of MdGolS genes during bud dormancy, including the functional characterization of MdGolS2, a strong candidate to regulate the seasonal accumulation of galactinol and RFOs during growth arrest in apple.

The number of predicted gene models coding for GolS in apple (eight) is similar to that found in Arabidopsis (eight) and poplar (nine) genomes (Yin et al., 2010). All MdGolS proteins

Fig. 4. MdGolS gene expression and sugar content analysis during an annual growth and dormancy cycle of ‘Fuji Standard’ closed terminal buds. The vertical dashed line indicates the approximate winter solstice (21 June); the shaded area corresponds to the winter season; and the horizontal bar corresponds to the budbreak interval (50% of buds in green tip stage). Standard error bars are shown. (A) Seasonal MdGolS transcript level analysis by real-time PCR. Gene expression was plotted relative to MdGolS4 transcript levels in January 2009. nd, not-detected. (B) Galactinol, raffinose and sucrose levels in apple buds determined by UPLC. Means with the same letter are not significantly different for P<0.01 by repeated measure ANOVA followed by Tukey’s test. (This figure is available in color at JXB online.)

Discussion

Galactinol synthase is a key enzyme in the biosynthetic pathway of RFOs, which are carbohydrates that play important roles in carbon storage, signal transduction, and osmoprotection, among others (Elsayed et al., 2014; Sengupta et al., 2015). Although several studies identified GolS genes and RFO levels as up-regulated during bud dormancy, suggesting that the seasonal mobilization of these carbohydrates could lead to bud integrity protection (Cox and Stushnoff, 2001; Derory et al., 2006; Santamaria et al., 2011; Liu et al., 2012; Unda et al., 2012; Ibáñez et al., 2013; Falavigna et al., 2014; Yue et al., 2015; Andersen et al., 2017), their direct relationship had still not been assessed. Within this context, evolutionary analyses of GolS genes in the Rosaceae family were performed followed by the characterization of MdGolS genes during bud dormancy, including the functional characterization of MdGolS2, a strong candidate to regulate the seasonal accumulation of galactinol and RFOs during growth arrest in apple.

The number of predicted gene models coding for GolS in apple (eight) is similar to that found in Arabidopsis (eight) and poplar (nine) genomes (Yin et al., 2010). All MdGolS proteins
Table 4. Pearson’s correlation coefficient between MdGolS transcript levels and carbohydrate contents in apple buds of ‘Fuji Standard’

|        | MdGolS1 | MdGolS2 | MdGolS3 | MdGolS4 | MdGolS5 | Galactinol | Raffinose | Sucrose |
|--------|---------|---------|---------|---------|---------|------------|-----------|---------|
| MdGolS1| 1.00    |         |         |         |         |            |           |         |
| MdGolS2| 0.81    | 1.00    |         |         |         |            |           |         |
| MdGolS3| 0.35    | 0.10    | 1.00    |         |         |            |           |         |
| MdGolS4| 0.67    | 0.69    | −0.24   | 1.00    |         |            |           |         |
| MdGolS5| −0.27   | −0.32   | 0.60    | −0.68   | 1.00    |            |           |         |
| Galactinol | 0.91 | 0.98    | 0.19    | 0.72    | −0.33   | 1.00       |           |         |
| Raffinose| 0.90    | 0.99    | 0.20    | 0.69    | −0.31   | 1.00       | 1.00      |         |
| Sucrose | 0.67    | 0.82    | −0.29   | 0.71    | −0.47   | 0.81       | 0.80      | 1.00    |

Results from the quantification of MdGolS transcript levels in ‘Fuji Standard’ apple buds. Values in bold represent positive correlation between parameters and are significant at the 0.01 level.

showed a putative serine phosphorylation site around amino acid position 265, which is a characteristic feature of nearly all GolS proteins reported so far (Taji et al., 2002). The prevailing exon/intron structure of MdGolS genes is the same as that observed for Arabidopsis and poplar GolS sequences (Fig. 1).

Interestingly, only five MdGolS genes were expressed under the conditions analysed (see Supplementary Fig. S1), and in agreement with this organ-selective GolS expression observed in apple. Arabidopsis AtGolS5, AtGolS6, AtGolS7, AtGolS9, and AtGolS10 are almost undetectable in some organs such as leaves (Nishizawa et al., 2008). A wide range of number of GolS genes was identified in the Rosaceae family, with apple and European pear showing the highest numbers. One possible explanation is that apples and pears experienced a recent WGD event that is not shared by other Rosaceous genera such as Prunus and Fragaria (Jung et al., 2012; Zhang et al., 2012). Additionally, besides the apple, Arabidopsis, and poplar GolS gene structural similarities, all these three species were subjected to recent WGDs (Velasco et al., 2010; Jiao et al., 2011; Daccord et al., 2017), and this may be indicative that this kind of duplication event was a driving force in the evolution of this particular class of genes.

To gain insights into such assumptions, evolutionary analyses were performed across Rosaceous species. A good robustness was obtained between analyses, due to several orthologs being identified in both the phylogenetic and collinear approaches (Table 3; Fig. 2; Supplementary Fig. S3). Moreover, several collinear regions containing GolS genes provided good examples of syntenic relationships among Rosaceous genomes (Jung et al., 2012; Zhang et al., 2012; Wu et al., 2013). Interestingly, all GolS paralogs from Rosaceous genera had their duplicated gene origin annotated as WGDs. The Ks-dating estimation of these paralogs allowed us to identify two distinct group values that can be related to different duplication events (Table 2). The first one was composed of some apple and pear paralogs and displayed Ks values lower than 0.25, which corresponds to a recent WGD that is only shared by these two species in the Rosaceae family (Velasco et al., 2007; Xiang et al., 2017). The second group showed Ks values around 1.5 and was identified among GolS paralogs from Malus, Prunus, and Fragaria genomes. Similar Ks values were also obtained for GolS paralogs of grapevine (Supplementary Table S4), a species that has only undergone the triplication event without subsequent polyploidy events (Tang et al., 2008). In addition, it is estimated that Ks values around 1.5 originated after the divergence of Amborella and the rest of the angiosperms (Jiao et al., 2011). In this context, GolS paralogs from apple, Japanese apricot, peach, woodland strawberry, and grapevine with Ks values around 1.5 likely originated in the triplication event. Specifically for apple, the expansion of GolS genes was driven by both triplication and recent WGD events, with no evidences of tandem rearrangements such as the ones observed in grapevine.

In general, paleopolyploidy events may provide opportunities for positive selection, generating copies of genes that persist in the course of evolution by retaining advantageous mutations (Yang and Bielawski, 2000; Zhang, 2003). The Ks/Ka ratios of the duplicated GolS genes in the Rosaceae family indicated a strong purifying selection after WGDs (Table 2). In several plants, genes derived from older duplication events also tended to have experienced stronger purifying selection (Zou et al., 2009; Du et al., 2013; Körbes et al., 2016). The maintenance of several GolS isoforms, together with the slow mutation rate in duplicated genes, may be explained by the importance of the roles of galactinol and RFOs in different cellular functions. Within this context, one possible explanation is that the differential GolS gene expression would be more important to define gene function than the protein structure itself, and this is a common evolutionary fate of duplicate genes called subfunction-alization (Zhang, 2003). This is a potentially important trend during evolution, where two genes with identical functions are unlikely to be stably maintained in the genome unless there is a division of gene expression after duplication (Zhang, 2003).

In fact, distinct expression patterns were observed for MdGolS genes in samples from different developmental stages (Fig. 3), as well as during bud dormancy progression (Fig. 4). MdGolS1, 3, and 4 presented highest transcript levels in mature seeds, agreeing with previous reports that identified GolS accumulation during seed development (Elayed et al., 2014; Sengupta et al., 2015; Salvi et al., 2016). The proposed role of GolS would be to increase seed RFO amounts, offering desiccation tolerance and energy during germination (Blöchl et al., 2007; Elayed et al., 2014; Sengupta et al., 2015). Recently, supplementation with galactose, a carbohydrate that integrates the RFO pathway, was shown to partially rescue a delayed seed germination phenotype observed in double knockout plants for raffinose synthase 4 and 5 (Gangl and Tenhaken, 2016). It has been proposed that galactose
may repress the expression of *AtPIF6* (phytochrome interacting factor 6), an important regulator in establishing the level of primary seed dormancy, and that RFOs would be a storage complex of galactose (Gangl and Tenhaken, 2016). Interestingly, almost all *MdGolS* genes were mainly expressed in dormant seeds or buds (Fig. 3), and common pathways
between bud and seed dormancy were already reported (Leida et al., 2012; Wang et al., 2016). Our findings reinforce that bud and seed dormancy may share similar physiological mechanisms, with GolS genes playing roles in both processes.

This finding prompted us to characterize the MdGolS gene expression dynamics, as well as the galactinol and RFO contents, in dormant buds during an annual growth and dormancy cycle (Fig. 4). MdGolS1 and MdGolS2 reached highest steady-state mRNA levels during winter. In agreement, galactinol and raffinose levels showed the same seasonal pattern and a high positive correlation with MdGolS1 and MdGolS2 transcript levels (Table 4). Similar GolS transcript patterns during dormancy were identified in blackcurrant, chestnut, hybrid poplar, pear, and sessile oak (Derory et al., 2006; Santamaria et al., 2011; Liu et al., 2012; Unda et al., 2012; Ibáñez et al., 2013; Andersen et al., 2017), whereas a similar raffinose content pattern was identified in blackcurrant and trembling aspen (Cox and Stushnoff, 2001; Andersen et al., 2017). The probable function of galactinol and RFOs during bud dormancy may be related to their osmolyte properties, mainly during dehydration (Elsayed et al., 2014; Sengupta et al., 2015). Indeed, dormant buds have to cope with a limited free water content during dormancy, employing several strategies to diminish damage due to this condition (de Fay et al., 2000; Améglio et al., 2002; Kalsits et al., 2009).

The evolutionary and transcriptional data gathered for MdGolS genes suggest that rather than playing different functions, these genes may exert similar functions but in different tissues and organs. While MdGolS2 was 627-fold more expressed in dormant buds in comparison with apple seeds, MdGolS1, MdGolS3, and MdGolS4 were mainly expressed in apple seeds (Fig. 3). This suggests that MdGolS1, MdGolS3, and MdGolS4 may have retained the already described role of GolS genes in seed development (Blochl et al., 2007; Elsayed et al., 2014; Sengupta et al., 2015), while MdGolS2 might play a role during bud dormancy. The same model of subfunctionalization has already been observed in several duplicated genes from many species (Zhang, 2003; Fawcett et al., 2009). Taken together, our data suggest that the recent origin of MdGolS2 resulted in a specialized adaptive role during apple bud dormancy (Tables 2 and 4; Fig. 4). In order to evaluate if MdGolS2, galactinol, and raffinose integrate one of the mechanisms conferring protection to dormant buds against a limited water supply, this gene was functionally characterized in Arabidopsis plants.

Transgenic Arabidopsis lines expressing MdGolS2 were generated and submitted to a severe water deficit assay (Fig. 5). Transgenic plants accumulated higher amounts of galactinol and RFOs in comparison with wild-type plants, mimicking what is observed in apple buds during dormancy. After the water withdrawal assay, only the transgenic plants were capable of surviving. This provides evidence that one of MdGolS2’s functions during bud dormancy is to synthesize galactinol, leading to increased amounts of RFOs, which will act together in bud protection against limited water supply. Transgenic expression of GolS genes from several different species was also capable of conferring tolerance to water deficit, freezing, heavy metal, or salt stress (Taji et al., 2002; Nishizawa et al., 2008; Wang et al., 2009; Song et al., 2013; Sun et al., 2013; Honna et al., 2016; Wang 2016b; Yu et al., 2017; Selvaraj et al., 2017). This increased stress tolerance may be due to the ability of galactinol and RFOs to scavenge hydroxyl radicals as well as their osmolyte properties, maintaining cell turgor, and stabilizing cell proteins and membranes (Nishizawa et al., 2008; Elsayed et al., 2014; Salvi et al., 2016).

The roles of starch, soluble sugars and sugar alcohols during dormancy have already been addressed in several reports, suggesting that carbohydrate metabolism is a key feature during dormancy, especially dormancy release (Maurel et al., 2004; Anderson et al., 2005; Marafon et al., 2011; Ito et al., 2013; Zhuang et al., 2015; Hussain et al., 2016; Andersen et al., 2017). Our findings add galactinol and RFOs to the list of sugars with important roles during bud dormancy. In apple, galactinol and raffinose integrate a series of adaptive mechanisms that act together during dormancy to protect buds from the intrinsic limited water supply that occurs during winter. MdGolS2 is likely the main enzyme responsible for their seasonal accumulation, although we cannot exclude that other MdGolS, especially MdGolS1, may also contribute for RFO synthesis during bud dormancy. Moreover, the reduced levels of these carbohydrates prior to budbreak suggest that they are being used as an energy source, in a similar trend to previous observations with other sugars (Maurel et al., 2004; Zhuang et al., 2015; Hussain et al., 2016). The accumulation of metabolites that can be used as protectors as well as an energy source is a highly advantageous evolutionary feature. WGD events have driven the evolution and diversification of GolS genes in angiosperms, especially in the Rosaceae family, resulting in the generation and further maintenance of several copies of these genes. It is tempting to speculate that the appearance of new structures and developmental programs, such as buds and dormancy, required the adaptation of already established molecular pathways. Possibly, the differential gene expression patterns of GolS genes improved plant adaptation to unsuitable environments, prompting the development of complex angiosperms.

Supplementary Data

Supplementary data are available at JXB online.

Fig. S1. Agarose gel electrophoretic result of PCR analysis of MdGolS genes in ‘Gala Baigent’ apple trees.

Fig. S2. Multiple alignments of GolS CDS predicted in the apple genome with those cDNA sequences amplified by RACE.

Fig. S3. Collinearity analyses between apple and other Rosaceae species.

Table S1. Sampling dates, corresponding season and chilling hours accumulated by ‘Fuji Standard’ buds.

Table S2. List of primers employed in this work.

Table S3. Sequences used for synteny and phylogenetic analyses.

Table S4. K, dating of grapevine GolS paralogs.

Acknowledgements

This work was supported by ‘Financiadora de Estudos e Projetos’ (FINEP, 010700700) and Embrapa (02.12.12.003.00.02). VSF received a PhD.
References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Amaral L, Gaspar M, Costa P, Aidar M, Buckeridge M. 2007. Novo método enzimático rápido e sensível de extração e dosagem de amido em materiais vegetais. Hoehnea 34, 425–431.

Amborella Genome Project. 2013. The Amborella genome and the evolution of flowering plants. Science 342, 1241089.

Amélio T, Bodet C, Lacointe A, Cochard H. 2002. Winter embolism, mechanisms of xylem hydraulic conductivity recovery and springtime growth patterns in walnut and peach trees. Tree Physiology 22, 1211–1220.

Andersen UB, Kjaer KH, Erban A, Alpers J, Hincha DK, Kopka J, Zuther E, Pagler M. 2017. Impact of seasonal warming on overwintering and spring phenology of blackcurrant. Environmental and Experimental Botany 140, 96–109.

Anderson JV, Gesch RW, Jia Y, Chao WS, Horvath DP. 2005. Seasonal shifts in dormancy status, carbohydrate metabolism, and related gene expression in crown buds of leafy spurge. Plant, Cell and Environment 28, 1567–1578.

Artlip TS, Wisniewski ME, Arora R, Norelli JL. 2016. An apple rootstock overexpressing a peach CBP gene alters growth and flowering in the scion but does not impact cold hardiness or dormancy. Horticulture Research 3, 16006.

Blüchl A, Peterbauer T, Richter A. 2007. Inhibition of raffinose oligosaccharide breakdown delays germination of pea seeds. Journal of Plant Physiology 164, 1003–1006.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant Journal 16, 735–743.

Conesa A, Gómez J, García-Gómez JM, Torol J, Talón M, Robles M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.

Cox SE, Stushnoff C. 2001. Temperature-related shifts in soluble carbohydrate content during dormancy and cold acclimation in Populus tremuloides. Canadian Journal of Forest Research 31, 730–737.

Daccord N, Celton J, Linsmith G, et al. 2017. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nature Genetics 49, 1099–1106.

de Fay É, Vacher V, Humbert F. 2000. Water-related phenomena in winter buds and twigs of Picea abies L. (Karst.) until bud-burst: a biotological, histological and NMR study. Annals of Botany 86, 1097–1107.

Dorony J, Léger P, Garcia V, Schaeffer J, Hauser MT, Salin F, Luschnig C, Plomion C, Glössl J, Kremer A. 2006. Transcriptome analysis of bud burst in sessile oak (Quercus petraea). New Phytologist 170, 723–738.

Du D, Cheng T, Pan H, Yang W, Wang J, Zhang Q. 2013. Genome-wide identification, molecular evolution and expression analyses of the phospholipase D gene family in three Rosaceae species. Scientia Horticulturae 153, 13–21.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32, 1792–1797.

Elsayed Al, Rafudeen MS, Golldack D. 2014. Physiological aspects of raffinose family oligosaccharides in plants: protection against abiotic stress. Plant Biology 16, 1–8.

EPPO. 1984. EPPO crop growth stage keys: apple and pear. EPPO Bulletin 14, 291–294.

Eremina M, Rozhon W, Poppenberger B. 2016. Hormonal control of cold stress responses in plants. Cellular and Molecular Life Sciences 73, 797–810.

Eyles A, Pinkard EA, Davies NW, Corkrey R, Churchill K, O’Grady AP, Sands P, Mohammed C. 2013. Whole-plant versus leaf-level regulation of photosynthetic responses after partial defoliation in Eucalyptus globulus saplings. Journal of Experimental Botany 64, 1625–1636.

Falavigna Vda S, Miotto YE, Porto DD, Anzanello R, Santos HP, Fialho FB, Margis-Pinheiro M, Pasquali G, Revers LF. 2015a. Functional diversification of the dehydrogenase gene family in apple and its contribution to cold acclimation during dormancy. Physiologia Plantarum 155, 315–329.

Falavigna VS, Porto DD, Buffon V, Margis-Pinheiro M, Pasquali G, Revers LF. 2014. Differential transcriptional profiles of dormancy-related genes in apple buds. Plant Molecular Biology Reporter 32, 795–813.

Falavigna VS, Porto DD, Silveira CP, Revers LF. 2015b. Recent advances in genetics and molecular control of bud dormancy in pippfruits. In: Anderson JV, ed. Advances in plant dormancy. Cham: Springer International Publishing, 107–122.

Fawcett JA, Maere S, Van de Peer Y. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. Proceedings of the National Academy of Sciences, USA 106, 5737–5742.

Gangl R, Tenhaken R. 2016. Raffinose family oligosaccharides act as galactose stores in seeds and are required for rapid germination of Arabidopsis in the dark. Frontiers in Plant Science 7, 1115.

Honna PT, Fuganti-Pagliai RT, Ferreira LC, et al. 2016. Molecular, physiological, and agronomical characterization, in greenhouse and in field conditions, of soybean plants genetically modified with AtGolS2 gene for drought tolerance. Molecular Breeding 36, 157.

Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. 2015. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31, 1296–1297.

Hussain S, Teng Y, Hamid A, Zulfiqar S, Shah A, Shah H, Yaqoob A, Ali T, Ahmed M, Hussain I. 2016. Carbohydrate changes during dormancy release in floral, vegetative buds and bark tissues of pear cultivar ‘Wonhwang’ cuttings following dormancy breaking agents treatment. International Journal of Environmental & Agriculture Research 3, 59–64.

Ibáñez C, Collada C, Casado R, González-Meléndez P, Aragoncillo C, Allison I. 2013. Winter induction of the galactinol synthase gene is associated with endodormancy in chestnut trees. Trees 27, 1309–1316.

Ito A, Sugiuira T, Sakamoto D, Moriguchi T. 2013. Effects of dormancy progression and low-temperature response on changes in the sorbitol concentration in xylem sap of Japanese pear during winter season. Tree Physiology 33, 398–408.

Jailon O, Aury J-M, Noel B, et al. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449, 463–467.

Jiao Y, Leebens-Mack J, Ayyampalayam S, et al. 2012. A genome triplication associated with early diversification of the core eudicots. Genome Biology 13, R3.

Jiao Y, Wickett NJ, Ayyampalayam S, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. Nature 473, 97–100.

Jung S, Cestaro A, Troggio M, et al. 2012. Whole genome comparisons of Fragaria, Prunus and Malus reveal different modes of evolution between Prunoideae subfamilies. BMC Genomics 13, 129.

Kessisoglu R, Keindell E, Silim S, Tanino K. 2009. Magnetic resonance microscopy indicates water diffusion correlates with dormancy induction in cultured hybrid poplar (Populus spp.) buds. Tree Physiology 29, 1269–1277.

Karimi M, Depicker A, Hilson P. 2007. Recombinational cloning with plant gateway vectors. Plant Physiology 145, 1144–1154.

Körbs AP, Kulcheski FR, Margis R, Margis-Pinheiro M, Turchetto-Zolet AC. 2016. Molecular evolution of the lysophosphatic acid acyltransferase (LPAA) gene family. Molecular Phylogenetics and Evolution 96, 55–69.

Krasensky J, Jonak C. 2012. Drought, salt, and temperature stress-induced metabolic reanarrangements and regulatory networks. Journal of Experimental Botany 63, 1603–1608.

and Dr Marcus Kvitschal (Empresa de Pesquisa Agropecuária do Estado de Santa Catarina), and Dr João Fioravanço (Empresa de Pesquisa Agropecuária do Estado de Santa Catarina) for their helpful discussions in the evolution section.
Adaptive roles of galactinol and raffinose during dormancy | 1259

Lefort F, Douglas GC. 1999. An efficient micro-method of DNA isolation from mature leaves of four hardwood tree species Acer, Fraxinus, Prunus and Quercus. Annals of Forest Science 56, 259–263.

Leida C, Corrales A, Arbona V, Gómez-Cadenas A, Láscar G, Badenes ML, Rios G. 2012. Chilling-dependent release of seed and bud dormancy in peach associates to common changes in gene expression. PLoS ONE 7, e35777.

Letunic I, Doerks T, Bork P. 2012. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Research 40, D302–D305.

Liu G, Li W, Zheng P, Xu T, Chen L, Liu D, Hussain S, Teng Y. 2012. Transcriptomic analysis of ‘Suli’ pear (Pyrus pyrifolia white pear group) buds during the dormancy by RNA-Seq. BMC Genomics 13, 700.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression using real-time quantitative PCR and the 2-ΔΔCT method. Methods 25, 402–408.

Lodhi MA, Ye G-N, Weeden NF, Reisch BL. 1994. A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. Plant Molecular Biology Reporter 12, 6–13.

Marafon AC, Citadin I, Amarante L Do, Herter FG, Hawerroth FJ. 2011. Chilling privation during dormancy period and carbohydrate mobilization in Japanese pear tree. Scientia Agricola 68, 462–468.

Maruyama K, Takeda M, Kidokoro S, et al. 2009. Metabolic pathways involved in cold acclimation identified by integrated analysis of metabolites and transcripts regulated by DREB1A and DREB2A. Plant Physiology 150, 1972–1980.

Maurel K, Sakr S, Gerbe F, Guilliot A, Bonhomme M, Rageau R, Pétel G. 2004. Sorbitol uptake is regulated by glucose through the hexokinase pathway in vegetative peach-tree buds. Journal of Experimental Botany 55, 879–888.

Nishizawa A, Yabuta Y, Shigeoka S. 2008. Galactinol and raffinose constituent a novel function to protect plants from oxidative damage. Plant Physiology 147, 1251–1263.

Perini P, Pasquali G, Margis-Pinheiro M, de Oliveira PRD, Revers LF. 2014. Reference genes for transcriptional analysis of flowering and fruit ripening stages in apple (Malus × domestica Borkh.). Molecular Breeding 34, 829–842.

Rohde A, Bhalerao RP. 2007. Plant dormancy in the perennial context. Trends in Plant Science 12, 217–223.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Lartet B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61, 539–542.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Rozen S, Skaletsky H. Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ: Humana Press, 365–386.

Ruijter JM, Ramakers C, Hoogas GM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF. 2003. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research 31, e45.

Salvi P, Saxena SC, Petta BP, Kamble NU, Kaur H, Verma P, Rao V, Ghosh S, Majee M. 2016. Differentially expressed galactinol synthase(s) constitute a novel function to protect plants from oxidative damage. Plant Physiology 171, 292–298.

Sun Z, Qi X, Wang Z, Li P, Wu C, Zhang H, Zhao Y. 2013. Overexpression of TsGOLS2, a galactinol synthase, in Arabidopsis thaliana enhances tolerance to high salinity and osmotic stresses. Plant Physiology and Biochemistry 69, 82–88.

Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K. 2002. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. The Plant Journal 29, 417–426.

Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH. 2008. Synteny and collinearity in plant genomes. Science 320, 486–488.

Unda F, Canam T, Preston L, Mansfield SD. 2012. Isolation and characterization of galactinol synthases from hybrid poplar. Journal of Experimental Botany 63, 2059–2069.

Velasco R, Zharkikh A, Affourtit J, et al. 2010. The genome of the domesticated apple (Malus × domestica Borkh.). Nature Genetics 42, 833–839.

Velasco R, Zharkikh A, Troggio M, et al. 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE 2, e1326.

Wang D, Gao Z, Du P, Xiao W, Tan Q, Chen X, Li L, Gao D. 2016a. Expression of ABA metabolism-related genes suggests similarities and differences between seed dormancy and bud dormancy of peach (Prunus persica). Frontiers in Plant Science 6, 1248.

Wang Y, Liu H, Wang S, Li H, Xin Q. 2016b. Overexpression of a common wheat gene GALACTINOL SYNTHASE3 enhances tolerance to zinc in Arabidopsis and rice through the modulation of reactive oxygen species production. Plant Molecular Biology Reporter 34, 794–806.

Wang Y, Yang H, Debarry JD, et al. 2012. MCScariX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Research 40, e49.

Wang Z, Zhu Y, Wang L, Liu X, Liu Y, Phillips J, Deng X. 2009. A WRKY transcription factor participates in dehydrtation tolerance in Boea hygrometrica by binding to the W-box elements of the galactinol synthase (BhGOLS1) promoter. Planta 230, 1155–1166.

Wisniewski M, Norelli J, Artlip T. 2015. Overexpression of a peach CBF gene in apple: a model for understanding the integration of growth, dormancy, and cold hardness in woody plants. Frontiers in Plant Science 6, 85.

Wu J, Wang Z, Shi Z, et al. 2013. The genome of the pear (Pyrus bretscheidniana Rhd.). Genome Research 23, 396–408.

Xiang Y, Huang CH, Hu Y, Wen J, Li S, Yi T, Chen H, Xiang J, Ma H. 2017. Evolution of Rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. Molecular Biology and Evolution 34, 262–281.

Yang Z, Bhalerao RP. 2000. Statistical methods for detecting molecular adaptation. Trends in Ecology and Evolution 15, 496–503.

Yin Y, Chen H, Hahn MG, Mohnen D, Xu Y. 2010. Evolution and function of the plant cell wall synthesis-related glycosyltransferase family B. Plant Physiology 153, 1729–1746.

Yu X, Ohtani M, Kusano M, Nishikubo N, Uenoyama M, Umezawa T, Saito K, Shinozaki K, Demura T. 2017. Enhancement of abiotic stress tolerance in poplar by overexpression of key Arabidopsis stress response genes, AtSRK2C and AtGoS2. Molecular Breeding 37, 57.

Yue C, Cao HL, Wang L, Zhou YH, Huang YT, Hao XY, Wang YC, Yang YJ, Wang XC. 2015. Effects of cold acclimation on sugar metabolism and sugar-related gene expression in tea plant during the winter season. Plant Molecular Biology 88, 591–608.

Zeng Y, Yang T. 2002. RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. Plant Molecular Biology Reporter 20, 417a–417e.

Zhang J. 2003. Evolution by gene duplication: an update. Trends in Ecology and Evolution 18, 292–298.

Zhang Q, Chen W, Sun L, et al. 2012. The genome of Prunus mume. Nature Communications 3, 1318.

Zhuang W, Gao Z, Wu Z, Tian G, Li J, Cai B. 2012. The genome of Prunus mume (Prunus mume Sieb. et Zucc.). Ecology and Evolution 2, 833–839.

Zou C, Lehti-Shiu MD, Thibaud-Nissen F, Prakash T, Buell CR, Shiu SH. 2009. Evolutionary and expression signatures of pseudogenes in Arabidopsis and rice. Plant Physiology 151, 3–15.