



HAL
open science

Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria 9 ananassa*) fruits

Jan G. Schaart, Christian Dubos, Irene Romeo de La Fuente, Adèle M.M.L. van Houwelingen, Ric C.H. de Vos, Harry H. Jonker, Wenjia Xu, Jean-Marc Routaboul, Loic Lepiniec, Arnaud G. Bovy

► To cite this version:

Jan G. Schaart, Christian Dubos, Irene Romeo de La Fuente, Adèle M.M.L. van Houwelingen, Ric C.H. de Vos, et al.. Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria 9 ananassa*) fruits. *New Phytologist*, 2013, 197 (2), pp.454-467. 10.1111/nph.12017 . hal-01191263

HAL Id: hal-01191263

<https://hal.science/hal-01191263>

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria × ananassa*) fruits

Jan G. Schaart^{1*}, Christian Dubos^{2,3*}, Irene Romero De La Fuente¹, Adèle M. M. L. van Houwelingen⁴, Ric C. H. de Vos^{4,5}, Harry H. Jonker^{4,5}, Wenjia Xu^{2,3}, Jean-Marc Routaboul^{2,3}, Loïc Lepiniec^{2,3} and Arnaud G. Bovy^{1,5}

¹Wageningen UR Plant Breeding, Wageningen University and Research Centre, PO Box 16, 6700 AA, Wageningen, the Netherlands; ²INRA, Institut Jean-Pierre Bourgin, Saclay Plant Sciences, RD10, F-78026, Versailles, France; ³AgroParisTech, Institut Jean-Pierre Bourgin, Saclay Plant Sciences, RD10, F-78026, Versailles, France; ⁴Plant Research International, Business Unit Bioscience, Wageningen University and Research Centre, PO Box 16, 6700 AA, Wageningen, the Netherlands; ⁵Centre for Biosystems Genomics, PO Box 98, 6700 AB, Wageningen, the Netherlands

Author for correspondence:

Jan G. Schaart

Tel: +31 317480963

Email: jan.schaart@wur.nl

Received: 17 July 2012

Accepted: 23 September 2012

New Phytologist (2012)

doi: 10.1111/nph.12017

Key words: bHLH, flavonoid biosynthesis, MYB, strawberry (*Fragaria × ananassa*), WD40-repeat, yeast-two-hybrid.

Summary

- Strawberry (*Fragaria × ananassa*) fruits contain high concentrations of flavonoids. In unripe strawberries, the flavonoids are mainly represented by proanthocyanidins (PAs), while in ripe fruits the red-coloured anthocyanins also accumulate. Most of the structural genes leading to PA biosynthesis in strawberry have been characterized, but no information is available on their transcriptional regulation. In *Arabidopsis thaliana* the expression of the PA biosynthetic genes is specifically induced by a ternary protein complex, composed of AtTT2 (AtMYB123), AtTT8 (AtbHLH042) and AtTTG1 (WD40-repeat protein).
- A strategy combining yeast-two-hybrid screening and agglomerative hierarchical clustering of transcriptomic and metabolomic data was undertaken to identify strawberry PA regulators.
- Among the candidate genes isolated, four were similar to AtTT2, AtTT8 and AtTTG1 (*FaMYB9/FaMYB11*, *FabHLH3* and *FaTTG1*, respectively) and two encode putative negative regulators (*FaMYB5* and *FabHLH3Δ*). Interestingly, *FaMYB9/FaMYB11*, *FabHLH3* and *FaTTG1* were found to complement the *tt2-1*, *tt8-3* and *ttg1-1* transparent testa mutants, respectively. In addition, they interacted in yeast and activated the *Arabidopsis* BANYULS (anthocyanidin reductase) gene promoter when coexpressed in *Physcomitrella patens* protoplasts.
- Taken together, these results demonstrated that *FaMYB9/FaMYB11*, *FabHLH3* and *FaTTG1* are the respective functional homologues of AtTT2, AtTT8 and AtTTG1, providing new tools for modifying PA content and strawberry fruit quality.

Introduction

The popularity of strawberry as a fruit crop is mainly a result of its unique aroma, sweet taste, bright colour and nutritional value. These quality traits are largely determined by the metabolic composition of the fruit. For example, strawberry fruits are known as a valuable source of polyphenol compounds. The main polyphenol compounds found in strawberry fruits are flavonoids. About 70% of the total phenolics that accumulated in cv Queen Elisa ripe fruits correspond to proanthocyanidins (PAs), whereas anthocyanins, flavonols and the other phenolics (i.e. *p*-coumaric acid, ellagic acid) represent c. 20, 3 and 7%, respectively (Almeida *et al.*, 2007; Carbone *et al.*, 2009). Flavonoids have

potential human health beneficial properties. They may act as antioxidants or as signalling molecules and were shown to be able to modulate some cellular functions (Nijveldt *et al.*, 2001; Battino *et al.*, 2009). Flavonoids such as PAs are also known to impact fruit flavour and astringency (Akagi *et al.*, 2011).

The flavonoid biosynthetic pathway has been extensively studied in numerous plant species at the genetic, biochemical and molecular levels, and was recently described in strawberry (Lepiniec *et al.*, 2006; Almeida *et al.*, 2007; Carbone *et al.*, 2009). This pathway (Fig. 1a) is initiated by the condensation of one molecule of 4-coumaroyl-CoA with three molecules of malonyl-CoA by the chalcone synthase (CHS) enzyme. The successive steps catalysed by the chalcone isomerase (CHI), flavonoid 3-hydroxylase (F3H/FHT), flavonoid 3'-hydroxylase (F3'H),

*These authors contributed equally to this work.

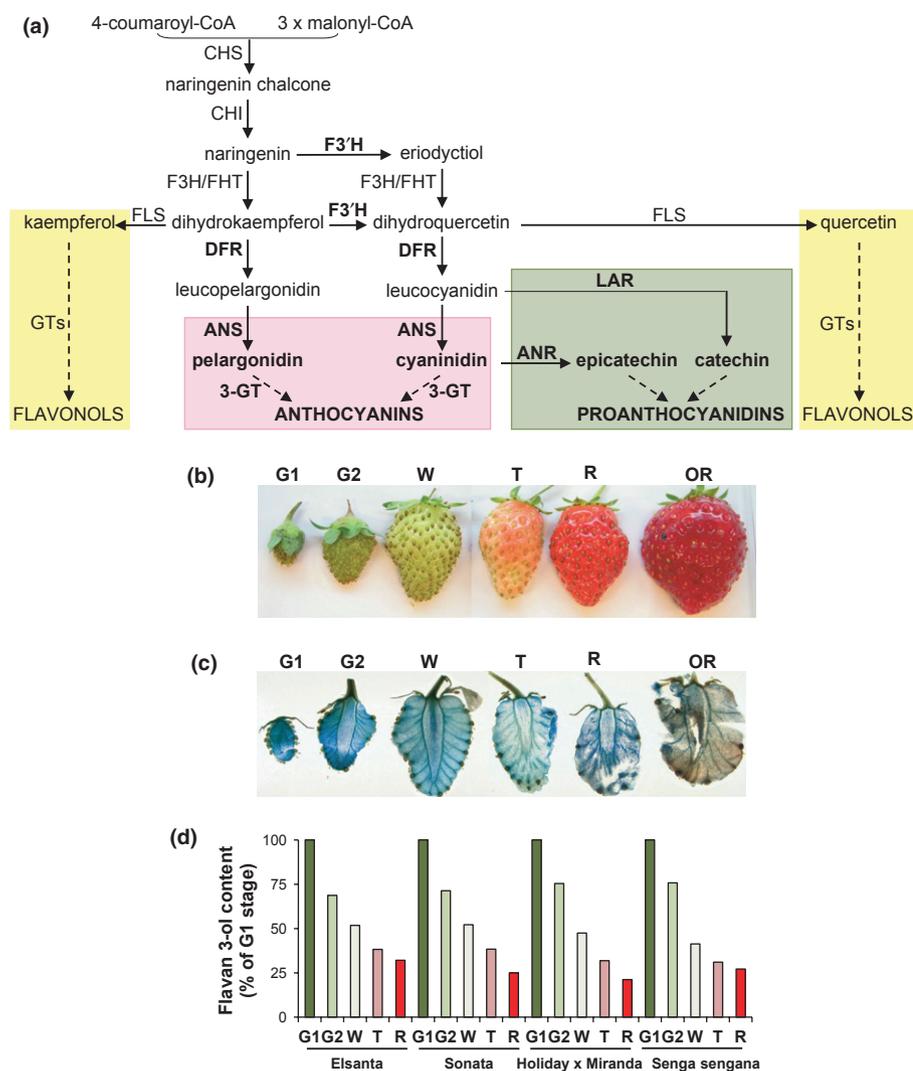


Fig. 1 Proanthocyanidin (PA) biosynthesis during fruit development. (a) Simplified model of the flavonoid biosynthetic pathway in strawberry (*Fragaria × ananassa*) fruits. CHI, chalcone isomerase; F3'H/FHT, flavonoid 3'-hydroxylase; F3'H, flavonoid 3'-hydroxylase; FLS, flavonol synthase; GTs, glycosyltransferases; DFR, dihydroflavonol-4-reductase; ANS, anthocyanidin synthase; 3-GT, 3-glycosyltransferase; LAR, leucoanthocyanidin reductase; ANR, anthocyanidin reductase. Dotted lines represent multistep reactions. (b) Strawberry fruit developmental stages (Elsanta genotype): G1, small green 1; G2, green 2; W, white; T, turning; R, red; OR, overripe. (c) 4-Dimethylaminocinnamaldehyde (DMACA) staining of cross-sections of Elsanta fruits at different developmental stages. (d) Soluble PA (flavan-3-ol) content during fruit development of four selected strawberry cultivars (relative to G1 stage).

dihydroflavonol-4-reductase (DFR), anthocyanidin synthase (ANS), and 3-glycosyltransferase (3-GT) enzymes lead to the production of the anthocyanin pigments. The F3H/FHT enzyme products dihydrokaempferol and dihydroquercetin are also the substrate of flavonol synthase (FLS) and lead, together with the involvement of several glycosyltransferases (GTs), to the production of flavonols. The products of the enzymatic reactions catalysed by DFR and ANS can also be converted into PA precursors (before polymerization/oxidation) by the leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR) enzymes, respectively. In strawberry, these two last steps produce both catechin- and epicatechin-based PAs that are found in different tissues (e.g. fruits, leaves or stems). Interestingly, *Arabidopsis* lacks a functional LAR gene and therefore produce exclusively epicatechin-based PA, which are limited to the seed coat (Abrahams *et al.*, 2003; Xie *et al.*, 2003; Routaboul *et al.*, 2006).

In *Arabidopsis*, the transcriptional regulation of PA biosynthesis has been extensively studied. Using a genetic approach, based on the study of the *transparent testa* (*tt*) phenotype observed in mutant seeds, three regulatory proteins were identified as necessary for the production of PA in the seed coat, namely AtTT2/

AtMYB123 (R2R3-MYB, Nesi *et al.*, 2001; Dubos *et al.*, 2010), AtTT8/AtbHLH042 (R/b-like bHLH; Nesi *et al.*, 2000) and AtTTG1 (WD40-repeat protein; Walker *et al.*, 1999). These three regulators were shown to act together in a ternary regulatory MYB-bHLH-WD40 (MBW) protein complex in order to regulate their target genes (e.g. *BAN/ANR*; Baudry *et al.*, 2004). For fruit crops there is limited information available on transcriptional regulation of the PA biosynthesis pathway. In grape, three MYB TF genes have been identified which show homology to *AtTT2* and promote PA biosynthesis early during grape berry development. The grape TFs VvMYB5b and VvMybPA2 are involved in the regulation of the PA pathway in skin tissue of young berries (Deluc *et al.*, 2008; Terrier *et al.*, 2009), whereas VvMybPA1 is mainly active in seeds (Bogs *et al.*, 2007). For persimmon, a MYB TF, DkMyb4, was isolated which was shown to be involved in regulation of PA biosynthesis at an early stage of fruit development (Akagi *et al.*, 2009). *DkMyb4* was expressed in fruit flesh of astringent fruit types, but its expression was reduced in nonastringent persimmon fruits, resulting in down-regulation of PA biosynthesis and the resultant nonastringency trait.

New tools specifically dedicated to strawberry have recently been developed, including the recent availability of the whole genome sequence, expressed sequence tag (EST) collections, *Agrobacterium*-mediated transformation protocols, loss-of-function mutant collections, and protocols for large-scale metabolite analysis (Bombarely *et al.*, 2010; Buendia *et al.*, 2010; Oosumi *et al.*, 2010; Shulaev *et al.*, 2011; Aaby *et al.*, 2012). Despite the fact that these tools may substantially help to improve our knowledge on strawberry fruit growth and development, little is still known about the regulation of flavonoid biosynthesis during the fruit-ripening process.

In order to get more insight into the regulation of PA biosynthesis during strawberry fruit development, a comparative genomics strategy was used, with the aim of identifying new PA-specific transcriptional regulators. We first found that *AtTT2* expression in strawberry fruit was sufficient to redirect flavonoid biosynthetic pathway from anthocyanin to PAs. This result suggests that functional homologues of the *AtTT8* and *AtTTG1* (and also *AtTT2*) transcriptional regulators are present in strawberry. A green fruit cDNA library was thus screened using both yeast- and PCR-based approaches, in order to identify new PA-specific regulatory genes. Metabolic and expression data, recorded during strawberry fruit development in 14 cultivars, were analysed using the agglomerative hierarchical clustering method (Vilo *et al.*, 2003). This approach revealed that nearly all the putative regulators identified were specifically expressed in the early phases of fruit development, when the accumulation of the PA biosynthetic gene transcripts and PA metabolites were the highest. Complementation experiments of *Arabidopsis transparent testa* null mutants (i.e. *tt2-1*, *tt8-3* and *ttg1-1*), coupled with transient expression assay in *Physcomitrella patens* protoplasts, led to the characterization of four strawberry regulatory proteins which are the functional homologues of *AtTT2*, *AtTT8* and *AtTTG1*, namely FaMYB9/FaMYB11, FabHLH3 and FaTTG1. The hypothetical function of two additional strawberry transcription factors (TFs; *FaMYB5* and *FabHLH3Δ*) that may act as negative regulators of PA biosynthesis is also discussed.

Materials and Methods

Plant material

Plants from different strawberry (*Fragaria × ananassa* Duch.) cultivars were grown in the field at the experimental farm 'the Santacker', near Elst in the fruit-growing area in the Netherlands. Fruits at different developmental stages were harvested in June in three successive years and frozen in liquid nitrogen immediately upon harvesting.

Liquid chromatography quadrupole time-of-flight tandem mass spectrometry (LC-QTOF-MS) analysis of semipolar strawberry fruit compounds

The extraction and the LC-QTOF-MS analysis of semipolar compounds were performed according to the protocol for tomato fruit, as described in Moco *et al.* (2006). In short, 0.5 g frozen

strawberry fruit powder (FW) of each of the 94 cultivars was extracted with 1.5 ml formic acid : methanol (1 : 1000, v/v) solution. The extracts were sonicated for 15 min and filtered through a 0.2 µm inorganic membrane filter (Anotop 10; Whatman, Maidstone, UK).

An LC-QTOF Ultima-MS platform was used for the profiling of the extracts. This platform consisted of a Waters Alliance 2795 HT HPLC system equipped with a Luna C18(2) pre-column (2.0 × 4 mm) and an analytical column (2.0 × 150 mm, 100 Å, particle size 3 µm; Phenomenex, Torrance, CA, USA) connected to an Ultima V4.00.00 QTOF mass spectrometer (Waters-Corporation, MS technologies, Manchester, UK). Degassed solutions of formic acid : ultra pure water (1 : 1000, v/v – eluent A) and formic acid : acetonitrile (1 : 1000, v/v – eluent B) were pumped into the HPLC system at 190 µl min⁻¹ and the gradient linearly increased from 5 to 35% eluent B over a 45 min period, followed by 15 min of washing and equilibration of the column. The column, sample and room temperatures were kept at 40, 20 and 20°C, respectively. Ionization was performed using an electrospray ionization (ESI) source and masses were detected in positive mode. A collision energy of 5 eV was used for full-scan LC-MS in the range of *m/z* = 100–1500. Leucine encephalin was used for online mass calibration (lock mass).

PA extraction and determination using phloroglucinol

The procedure was performed according to Kennedy & Jones (2001) with slight modifications in the HPLC conditions, as previously published in Almeida *et al.* (2007). Phloroglucinol adducts were analysed by reverse-phase high-performance liquid chromatography (HPLC) with photodiode array detection (280 nm) and fluorescence detection (excitation at 275 nm, emission at 310 nm). (+)-Catechin and (–)-epicatechin, purchased from Sigma, were used as standards for quantification.

Strawberry transformation

For overexpression of *AtTT2*, *AtTT8* and *AtTTG1*, the corresponding open reading frames (ORFs) were introduced into the pCaMV35S expression cassette of the binary Gateway vector pGD625 (Chalfun *et al.*, 2005) and transferred to the supervirulent *Agrobacterium tumefaciens* strain AGL0 (Lazo *et al.*, 1991). Transgenic strawberry plants of the cv Calypso were produced according to Schaart *et al.* (2002). Initially, 50–100 kanamycin-resistant clones were obtained for each construct, from which 15 good growing lines were transferred to the glasshouse. Finally, for each construct three individual transgenic lines were selected for further experimentation. Calypso nontransgenic plants were used as control. For each line, two or three plants were maintained for harvesting fruits.

Yeast-two-hybrid screening and interaction study

For the yeast-two-hybrid studies, the ORFs of the different *MYB*, *bHLH* and *TTG1* TF genes described were recombined into the pBDGAL4 bait vector (pDESTTM32; Invitrogen) and

pADGAL4 prey vector (pDESTTM22; Invitrogen) by a GatewayTM LR reaction. For the yeast-two-hybrid screening, an expressed green strawberry fruit cDNA library was recombined into the pADGAL4 prey vector. The resulting bait and prey vectors were transformed into yeast strains PJ69-4a (*MATa*) and PJ69-4α (*MATα*) (James *et al.*, 1996), respectively, and used in yeast-two-hybrid experiments as described by de Folter *et al.* (2005). GenBank accession number for the genes identified are *FaMYB5*, JQ989280; *FaMYB9*, JQ989281; *FaMYB11*, JQ989282; *FaMYC1*, JQ989283; *FabHLH3*, JQ989284; *FabHLH3Δ*, JQ989285; *FabHLH33*, JQ989286; and *FaTTG1*, JQ989287.

RNA extraction and quantitative real-time reverse transcription PCR (RT-PCR)

Total RNA isolation, cDNA synthesis and quantitative RT-PCR was performed as described previously (Schaart *et al.*, 2002) SYBR Green was used for detection of PCR products on a MyiQ Single-Color Real-Time Detection System (Bio-Rad). All the primers used in this study are described in Supporting Information, Table S1.

Arabidopsis thaliana complementation assays

FaMYB9, *FaMYB11*, *FabHLH3* and *FaTTG1* cDNAs were introduced into the binary vector pMDC32 (Curtis & Grossniklaus, 2003) by a Gateway LR recombination. *Agrobacterium*-mediated transformation of the *Arabidopsis tt2-1*, *tt8-3* and *ttg1-1* null mutants were carried out as described in Baudry *et al.* (2006). Wildtype Landsberg erecta (Ler) and Wassilewskija (WS) were used as controls where appropriate. Fourteen independent transgenic lines per construct were isolated, genotyped and checked for seed colour, from which three were chosen to measure PA accumulation in triplicate (in accordance with Routaboul *et al.*, 2006; Dubos *et al.*, 2008).

Physcomitrella patens transient expression study

FaMYB9, *FabHLH3* and *FaTTG1* cDNAs were introduced into the expression vector pBS Tpp-A by a LR recombination (Thevenin *et al.*, 2012). The reporter proBAN : GFP vector used (containing a 75-bp-long promoter) as well as all the steps from transfection to quantitative analysis, are described in detail elsewhere (Thevenin *et al.*, 2012).

Results

Large variation in flavan-3-ols within strawberry varieties

To investigate the genetic variation in soluble flavan-3-ol monomers in cultivated strawberry, aqueous-methanol extracts from red fruit of 57 different strawberry genotypes were analysed using accurate mass LC-MS. This showed that in all genotypes the monomeric flavan-3-ols catechin and epicatechin were present (Table 1, Fig. S1). In addition to these monomeric forms, two

Table 1 Proanthocyanidin (PA) composition obtained by phloroglucinolysis of selected strawberry (*Fragaria × ananassa*) cultivars

Genotypes	PAs		DP		Catechin free		Catechin end		Catechin extension		Epicatechin free		Epicatechin end		Epicatechin extension	
	Average	± SE	Average	± SE	Average	± SE	Average	± SE	Average	± SE	Average	± SE	Average	± SE	Average	± SE
Darselect × Honeoye	1322.67	170.13	5.58	0.00	96.69	22.54	220.88	24.53	183.32	33.61	nd	nd	19.36	4.51	802.41	93.96
Elsanta	1207.81	247.28	5.18	0.58	86.09	24.83	222.40	45.64	314.56	60.06	nd	nd	9.24	2.66	567.65	171.47
Florence	1150.47	165.01	7.03	0.82	51.25	11.65	161.10	15.44	243.28	29.05	nd	nd	6.47	1.98	691.51	169.78
Holiday × Miranda 1	1027.55	151.67	5.63	0.68	66.73	18.29	176.01	21.58	202.33	18.56	nd	nd	7.95	2.16	573.11	149.16
(Elsanta × Pajaro) × (Korona × Selva)	856.11	188.27	5.77	1.03	53.13	24.02	157.99	7.39	215.97	23.16	nd	nd	5.35	1.97	478.80	156.40
Korona	783.81	194.46	6.40	0.62	43.68	16.23	122.75	30.08	186.87	34.35	nd	nd	5.53	2.38	400.98	142.56
Camarosa	781.29	186.11	4.94	0.00	59.73	20.53	143.22	22.22	179.00	12.04	nd	nd	4.36	0.97	394.98	130.35
Sonata	743.27	211.18	5.02	0.81	66.81	24.61	141.62	37.71	169.19	54.01	nd	nd	7.92	1.56	389.88	116.36
Oso Grande	708.76	337.43	4.80	0.76	67.12	24.54	112.85	89.94	162.16	95.59	nd	nd	3.97	1.33	362.66	153.07
((Silveta × Holiday) × Korona) × Scott × (Bogota × (Redchief × Sivetta))	648.36	168.99	6.57	0.52	36.30	16.20	98.32	19.03	141.20	8.85	nd	nd	4.28	1.38	417.25	150.02
Figaro	620.66	100.32	6.09	0.63	27.63	7.75	103.55	15.72	159.85	15.71	nd	nd	4.41	1.22	343.85	89.97
Senga sengana	581.01	193.48	6.06	1.00	29.03	9.92	96.19	38.54	138.30	45.68	nd	nd	5.77	1.91	315.40	173.93
Holiday × Miranda 2	574.97	75.03	6.84	1.02	25.11	2.69	83.06	13.25	94.09	5.89	nd	nd	4.47	1.37	353.30	75.62
(Karina × Primella) × (Midway × Holiday)	540.87	169.96	5.83	0.70	21.44	2.87	103.58	53.88	115.22	42.55	nd	nd	4.49	1.75	311.56	108.05

Total PA, total soluble PA; mDP, mean degree of polymerization; nd, not detected.

major PA dimers (B1 and B3) were detected in all genotypes (Figs S1, S2). The concentrations of the monomeric flavan-3-ol catechin and of their dimeric PA B1 and B3 derivatives in ripe fruits seemed to be strongly correlated ($R^2 = 0.84$ and 0.91 ; Fig. S2), whereas epicatechin concentrations did not correlate with either catechin or PA dimer concentrations (not shown). Relative flavan-3-ol monomer content varied up to 8.5 times between the most extreme genotypes (Fig. S1).

Since it is difficult to extract and quantify all individual oligo- and polymeric PAs, the total PA content in red fruits of a selection of 14 contrasting strawberry cultivars (grown in three consecutive years) were determined using an HPLC photodiode array detector after hydrolysis with phloroglucinol (Table 1). Total PA concentrations varied between 0.54 and 1.32 g kg^{-1} FW among genotypes. On average, the amount of total PA is 15 times higher than the amount of free PA monomers. Catechin and epicatechin accounted, on average, for 44% ($\pm 5\%$) and 56% ($\pm 5\%$) of total PA units, respectively. Catechin was mainly present as the end unit, whereas both catechin and epicatechin were present as extension units, the latter being prominent. The mean degree of polymerization of these soluble PAs varied between 4.8 (± 0.76) and 7.03 (± 0.82).

We examined PA accumulation during fruit development by staining cross-sections of Elsanta fruits at different developmental stages with 4-dimethylaminocinnamaldehyde (DMACA), which specifically stains PA in blue (Fig. 1b,c). Fig. 1(c) shows that PAs were localized throughout the fruit tissue and seemed to be more concentrated in the green fruits compared with later ripening stages. Then we also measured soluble PA content at different fruit developmental stages with four different strawberry genotypes (i.e. Elsanta, Sonata, Holiday \times Mirande, and Senga senga) using HPLC. For all genotypes tested, a similar gradual decrease was observed during fruit ripening to a value of *c.* 25% of that measured in small green fruits (Fig. 1d).

Overexpression of the *Arabidopsis* AtTT2, AtTT8 and AtTTG1 PA biosynthesis-related transcriptional regulators in strawberry

Expression of the *Arabidopsis* PA biosynthesis regulators *AtTT2/AtMYB123*, *AtTT8/AtbHLH042* and *AtTTG1* in strawberry was carried out first to assess the potential effect of these proteins on strawberry PA biosynthesis (Walker *et al.*, 1999; Nesi *et al.*, 2000, 2001). Transgenic ripe fruits showed a loss of red coloration in the flesh (Fig. 2a) associated with an increased PA content (Fig. 2b) and a decreased anthocyanin accumulation (Fig. 2c). This increased PA content was also observed in leaves and stems of plants overexpressing simultaneously *AtTT2*, *AtTT8* and *AtTTG1* (Fig. S3). The observed anthocyanin loss was mainly the result of a strong decrease in pelargonidin accumulation (Fig. 2d), whereas cyanidins were slightly, but not significantly, increased (Fig. 2e). In order to elucidate the molecular mechanism behind this effect of *Arabidopsis* MBW complex on PA and anthocyanin biosynthesis in strawberry fruit, a transcript accumulation study of key biosynthetic genes (namely: *FaF3'H*, flavonoid 3'-hydroxylase *FaDFR*, dihydroflavonol-4-reductase; *FaANS*, anthocyanidin synthase; *FaANR*, anthocyanidin reductase;

and *FaLAR*, leucoanthocyanidin reductase; Fig. 1a) was carried out using quantitative RT-PCR (qRT-PCR). This study revealed that *FaF3'H* transcript accumulation was slightly increased, *FaANS*, *FaANR* and *FaLAR* transcript accumulation were strongly increased, and that *FaDFR* transcript accumulation remained unchanged by the *Arabidopsis* MBW complex (Fig. 2f). The effects on the transcript abundance were consistent with those observed at the metabolic level (Figs 1a, 2a–e). Interestingly, the plants overexpressing *AtTT2* alone (but not *AtTT8* or *AtTTG1* alone) showed the same changes at both metabolite and transcript levels. In ripe fruits of *AtTT2*-overexpressors, a significant increase in the amount of total PAs was also found which coincided with a decrease in pelargonidin accumulation. Moreover, these fruits also showed elevated gene expression levels of *FaF3'H*, *FaANS*, *FaANR* and *FaLAR*. This indicates that, in strawberry fruits, *AtTT2* alone is sufficient to redirect the flavonoid pathway towards the biosynthesis of PA at the expense of the pelargonidin-type of anthocyanins (Figs 2, S3).

Identification of putative strawberry homologues of AtTT2, AtTT8 and AtTTG1

Two consecutive yeast-two-hybrid screens were carried out in order to identify putative homologues of *AtTT2*, *AtTT8* and *AtTTG1* (Figs 3a, S4). The first yeast-two-hybrid screen was performed using *AtTT2*, *AtTT8* and *AtTTG1* as bait and an expressed green strawberry fruit cDNA library as prey. Because *AtTT2* showed self-activation, a series of C-terminal deletions was made, and one deleted *AtTT2* form (of 218 aa) that lost its self-activation property was selected. However, using this partial TF as bait, no interactions were identified. Using *AtTT8* as bait, three different strawberry TF cDNAs called *FaMYB5*, *FaMYB11* and *FaTTG1* were identified. When *AtTTG1* was used as bait, two proteins, *FaMYC1* and *FabHLH33*, showed interaction. In the second yeast-two-hybrid screen, two of the identified strawberry TF genes, *FaMYC1* and *FaTTG1*, were used as bait, but this did not result in new interacting partners. Screening with *FaMYC1* identified *FaTTG1* as interacting partner, while using *FaTTG1* as bait confirmed the putative interactions with *FaMYC1* and *FabHLH33*.

Next to the yeast-two-hybrid screening approach, putative candidate cDNA sequences were amplified from a mixed green and red strawberry fruit cDNA sample with degenerate primers that were designed based on the cDNA sequences of homologues of both *AtTT2* and *AtTT8*. This resulted in three additional candidate cDNAs, *FaMYB9*, *FabHLH3* and *FabHLH3A*.

Interaction of the putative flavonoid regulators in yeast

The yeast-two-hybrid strategy was also used to study the ability of the entire set of identified strawberry transcriptional regulators, including *FaMYB1*, a repressor previously described by Aharoni *et al.* (2001), to interact with each other and with *AtTT2*, *AtTT8* and *AtTTG1* (Fig. 3a). Before starting this interaction study, the self-activation property of these proteins was evaluated. We found that *FaMYB5*, *FaMYB9* and *FaMYB11* displayed a strong self-activation when fused to the GAL4 DNA-binding domain

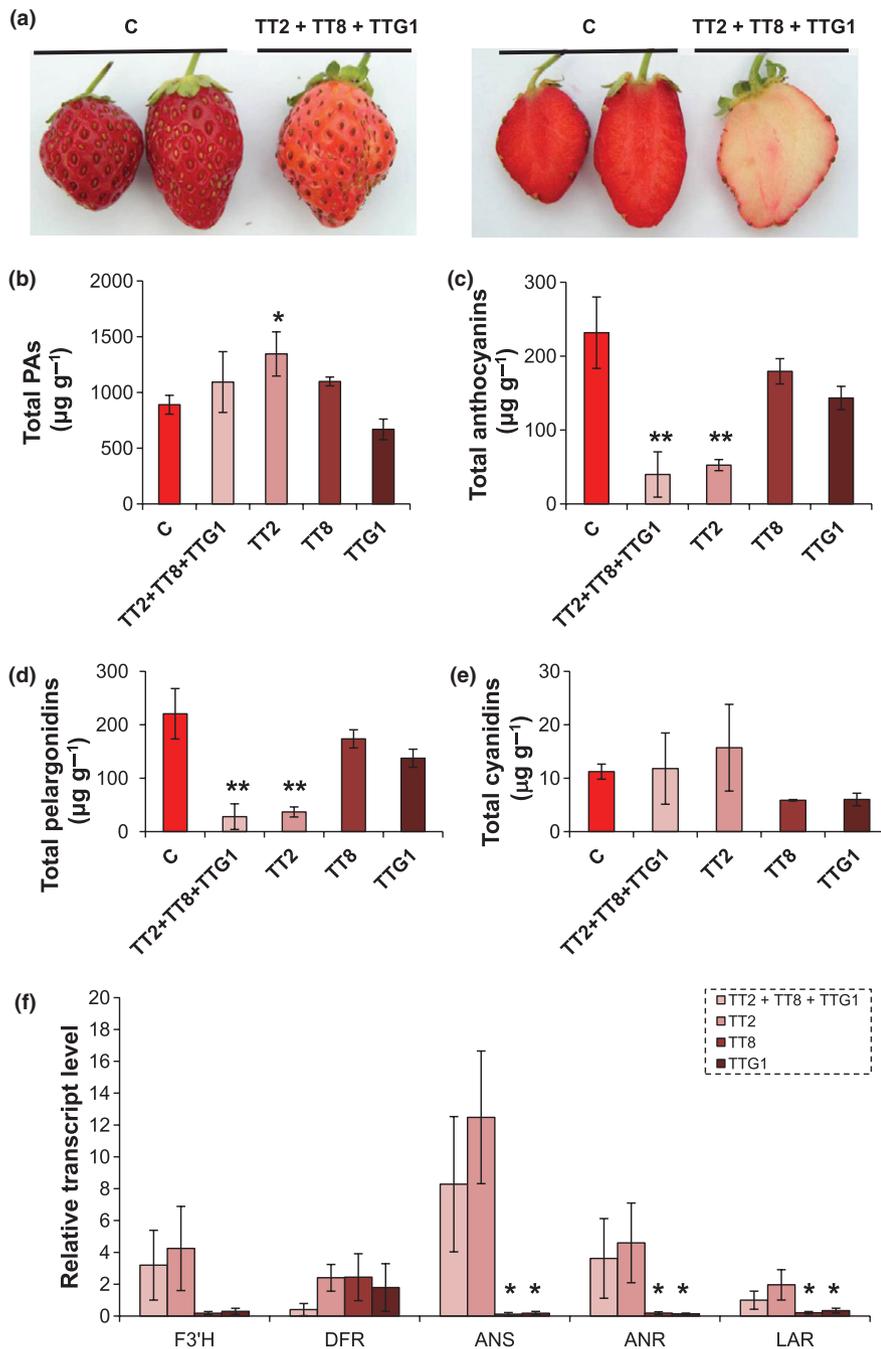


Fig. 2 Overexpressions of the *Arabidopsis* proanthocyanidin (PA)-related transcriptional regulators in strawberry (*Fragaria* \times *ananassa*) fruit result in modified PA and anthocyanin content. (a) Nontransgenic strawberry fruit (C) compared with fruits overexpressing *AtTT2*, *AtTT8* and *AtTTG1* (TT2 + TT8 + TTG1). Left panel, whole fruit; right panel, fruit longitudinal sections. (b–e) Total PA (b), anthocyanin (c), pelargonidin (d) and cyanidin content (e) in control and transgenic fruits overexpressing *AtTT2*, *AtTT8* and *AtTTG1* together or individually. (f) Transcript accumulation measurements of five key structural genes leading to PA (flavonoid 3'-hydroxylase (F3'H), anthocyanidin reductase (ANR) and leucoanthocyanidin reductase (LAR) or anthocyanin (dihydroflavonol-4-reductase (DFR) and anthocyanidin synthase (ANS)) biosynthesis in fruits from transgenic plants. Transcript abundances are relative to nontransgenic control fruits. *t*-test significant difference: **, $P < 0.01$; *, $P < 0.05$; error bars show SEs (b–f).

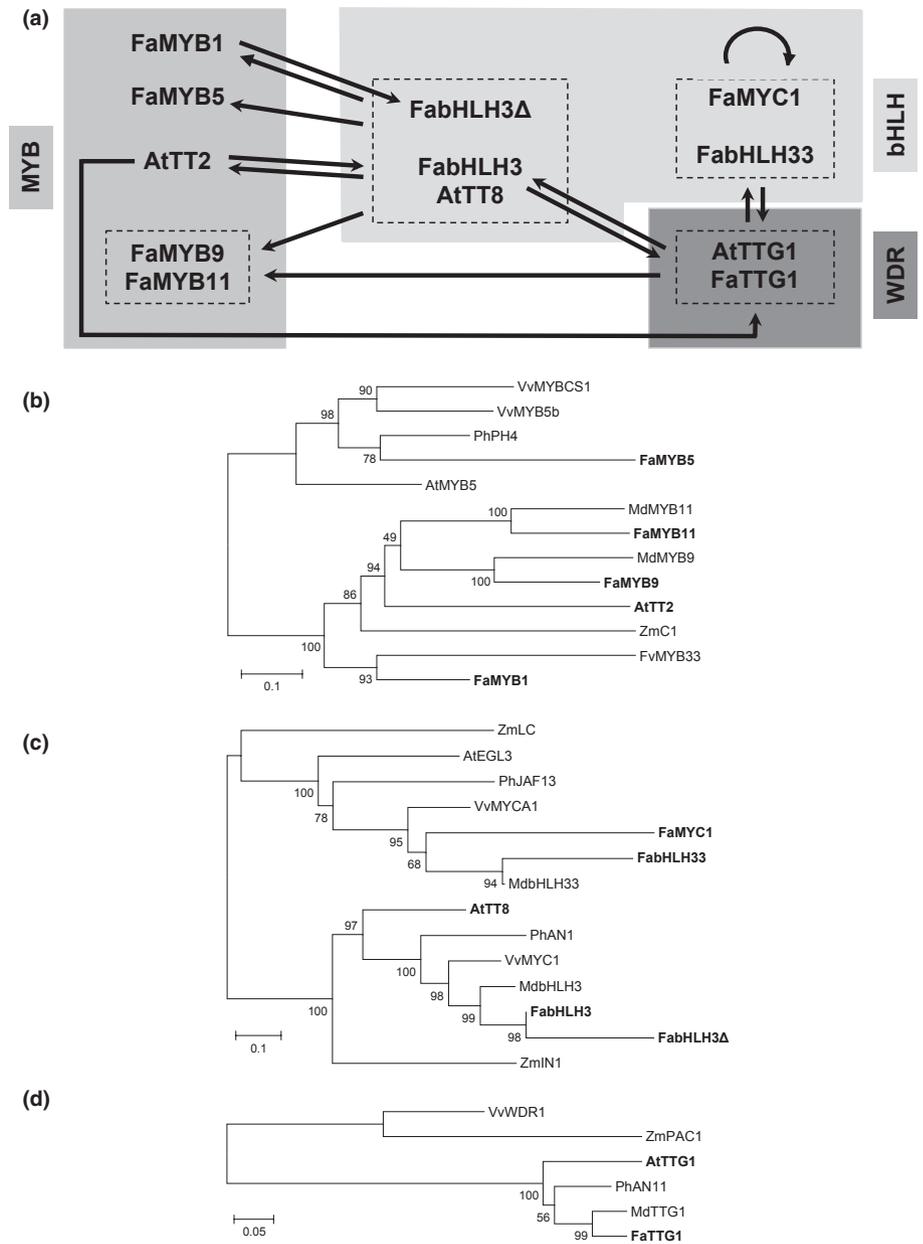
(BD); therefore, their interaction with other TFs was only assessed when fused to the GAL4 DNA-activating domain (AD). From this preliminary analysis, a total of 108 combinations were tested (12×9 matrix; Fig. S4), from which 45 interactions were found. Interestingly, most of the observed interactions in yeast were independent of the fused domain (i.e. BD or AD for the bait or prey protein, respectively), as only about one-third (37%) of the interactions were specific to the fusion to the BD domain.

When *AtTT8*, *FabHLH3* and *FabHLH3Δ* were used as bait, they were found to be able to interact with all the tested MYB TFs, whereas only *AtTT8* and *FabHLH3* (not *FabHLH3Δ*)

displayed an interaction with the *Arabidopsis* and strawberry TTG1 proteins. When *FaMYC1* or *FabHLH33* were used as bait, they were also found to interact with both TTG1 proteins, but no interaction was identified with the MYB TF used in this study. *FaMYC1* and *FabHLH33* were able to form homo- and heterodimers, a characteristic of some basic helix–loop–helix (bHLH) proteins (Feller *et al.*, 2006). When both TTG1 proteins were used as prey, interactions were found with *FaMYB9*, *FaMYB11*, *AtTT8*, *FaMYC1*, *FabHLH3* and *FabHLH33*. Finally, the BD fusion to *AtTT2* gives *AtTT8*, *FabHLH3* and *FabHLH3Δ* as well as TTG1 proteins as interacting partners,

Fig. 3 Identification of putative transcriptional regulator involved in proanthocyanidin (PA) biosynthesis in strawberry (*Fragaria × ananassa*) fruits.

(a) Schematic representation of the yeast-two-hybrid interactions identified between the strawberry putative regulators isolated in this study and AtTT2, AtTT8 and AtTTG1. Arrowheads indicate the protein that was identified as prey (i.e. fused to the GAL4 DNA-activation domain), whereas the other end indicates the bait (i.e. fused to the GAL4 DNA-binding domain). Schematic representation of the phylogenetic relationships existing between the strawberry MYB (b), basic helix-loop-helix (bHLH) (c) and TTG1-like (d) proteins isolated and used in this study (in bold) with their closest homologues. Sequences were aligned using Clustal W (opening = 10, extension = 0.2) and phylogeny reconstruction was done using the neighbour-joining method and tested using the bootstrap method with 1000 replicates. Bootstrap values are given next to each node. Both alignment and phylogenetic analysis were performed using MEGA version 4 (Tamura *et al.*, 2011). The scale bar indicates the branch length that corresponds to the number of substitutions per amino acid position (b–d). Accession IDs: ZmC1, 1613412E; FaMYB1, AAK84064; AtMYB5, Q38850; PhMYB4, ADX33331; VvMYB5b, AAX51291; VvMYBCS1, AAS68190; MdMYB11, AAZ20431; MdMYB9, ABB84757; ZmLc, NP_001105339; PhAN1, AAG25928; VvMYC1, ACC68685; MdbHHLH3, ADL36597; GL3: Q9FN69; VvMYCA1, ABM92332; MdbHHLH33, ABB84474; VvWDR1, ABF66625; ZmPAC1, AAM76742; PhAN11, AAC18914; MdTTG1, ADI58760; FvMYB33, ADU56173; PhJAF13, AAC39455; ZMIN1: AAB03841. Zm, *Zea mays*; At, *Arabidopsis thaliana*; Ph, *Petunia × hybrida*; Fa, *Fragaria × ananassa*; Fv, *Fragaria vesca*; Vv, *Vitis vinifera*; Md, *Malus × domestica*.



and when the BD was fused to FaMYB1, it only interacted with FabHLH3Δ.

Phylogenetic analysis of the identified strawberry transcriptional regulators

Phylogenetic analysis of the deduced amino acid sequences using the neighbour-joining method was employed to determine the relationships that may exist between the identified strawberry MYB, bHLH and WD40-repeat regulatory proteins and the known TF genes from *Arabidopsis*, petunia (*Petunia × hybrida*), maize (*Zea mays*), apple (*Malus × domestica*) and grape (*Vitis vinifera*). The predicted proteins were similar in length to their homologues, except for FaMYB5 and FabHLH3Δ, for which the predicted proteins were much shorter (Fig. S5).

FaMYB5 was most similar to VvMYBCS1 and VvMYB5b from grape, AtMYB5 from *Arabidopsis* (52% identity) and PhPH4 from petunia (Fig. 3b). All these MYBs form a subgroup that was previously defined when PhMYB4 was first identified. This MYB subgroup contains two conserved C-terminal motifs named ‘G20’ (Quattrocchio *et al.*, 2006). These G20 motifs are also present in FaMYB5, but not in AtTT2 (Fig. S5a). Unlike the other members of the G20 subgroup, the FaMYB5 deduced amino acid sequence missed the initial start codon, leading to a 70-amino-acid deletion at the N-terminus. The 70-amino-acid deletion results in a loss of the R2 domain that is present in all members of the R2R3-MYB family (Fig. S5a). The R2R3-MYB FaMYB9 and FaMYB11 are homologues of the apple genes MdMYB9 and MdMYB11 (67 and 68% identity, respectively), and both share some similarity with

AtTT2 (46 and 45% identity, respectively; Fig. 3a). All three identified strawberry MYB TF genes map to different locations on the genome of woodland strawberry (*Fragaria vesca*), which was published recently (Shulaev *et al.*, 2011). Although *FaMYB9* and *FaMYB11* both map to LG6, their mapping position is located 13.3 Mb apart. *FaMYB9* and *FaMYB11* are the respective homologues of the *F. vesca* TFs *FvMYB58* and *FvMYB26*, which are both part of a clade of sequences with homology to *AtTT2* of a *F. vesca* phylogenetic tree of R2R3-MYB-predicted proteins (see the Supporting Information of Shulaev *et al.*, 2011). Next to *FvMYB58* and *FvMYB26*, this clade represents four other R2R3-MYB-predicted proteins. *FvMYB33*, which is also present in this clade, was reported to be highly expressed in *F. vesca* fruits and was therefore suggested to have a key function in PA synthesis (Shulaev *et al.*, 2011). Phylogenetic analysis shows that *FvMYB33* is less related to *AtTT2* than to *FaMYB9* and *FaMYB11* (Fig. 3b).

Two different forms of the bHLH TF gene *FabHLH3* were isolated, which differed by several single nucleotide polymorphisms (SNPs) and a 152 bp insert. This insert causes a frame shift mutation, putatively leading to a short and truncated FabHLH3 protein named *FabHLH3Δ*, which is missing the typical bHLH domain and the ACT-like dimerization domain (Fig. S5b; also see Feller *et al.*, 2006). When blasting both *FabHLH3* sequences against the *F. vesca* genomic sequence, only a single homologous sequence with high similarity was found, suggesting that the two *FabHLH3* sequences are not paralogous. The FabHLH3 amino acid sequence is very similar to MdbHLH3 from

apple (80% identity) and related to *AtTT8* (57% identity) and to other *AtTT8*-like TFs such as the petunia PhAN1 or the grape *VvMYC1* (Fig. 3c). *FabHLH33* and *FaMYC1* share strong similarity at the predicted protein level, and are more closely related to *AtEGL3/AtbHLH002* (43 and 45% identity, respectively) than to *AtTT8* (Fig. 3c). Moreover, *FabHLH33* appears to be homologous to *MdbHLH33* from apple (75% identity). *FabHLH33* and *FaMYC1* map on unique positions at the *F. vesca* linkage groups LG7 and LG5, respectively.

Finally, *FaTTG1* displayed a very high similarity to its homologue from *Arabidopsis*, *AtTTG1* (80% identity protein level); however, the highest similarity was found with the apple homologue *MdTTG1* protein (91% identity; Fig. 3d). Also *FaTTG1* was mapped on a single location at the *F. vesca* genome.

Metabolomic and transcriptomic data clustering analysis during fruit development

Metabolomic (HPLC) and transcriptomic (qPCR) data from 14 genotypes at the mature stage (R), and on the four studied genotypes (Elsanta, Sonata, Holiday × Miranda, and Senga sengana) during fruit development (G1, small green; G2, green; W, white; T, turning; and R, red) were collected and organized by the agglomerative hierarchical clustering (AHC) method, using the EPCLUST tool (Fig. 4; Vilo *et al.*, 2003). The metabolites measured mainly belonged to the PA biosynthetic pathway (i.e. free or polymerized catechin and epicatechin subunits, total flavan-3-ols, and

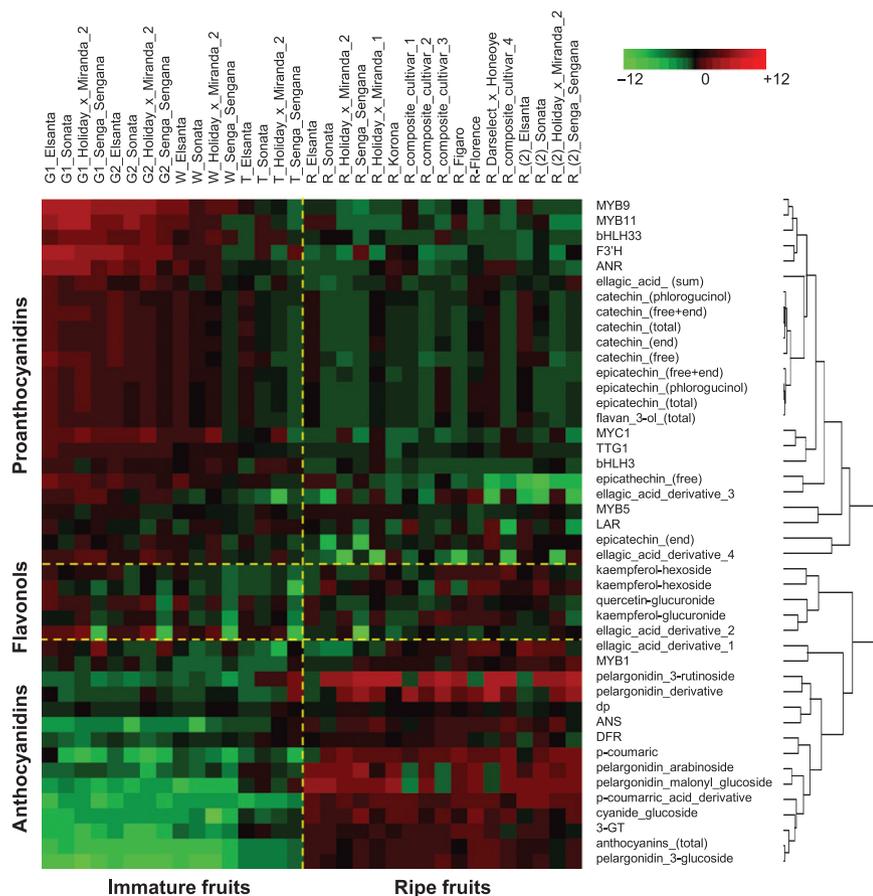


Fig. 4 Agglomerative hierarchical clustering analysis of metabolic and transcriptomic data obtained during strawberry (*Fragaria × ananassa*) fruit development. (Poly)phenolic compounds detected in strawberry fruit were clustered together with the level of transcript accumulated for key structural genes involved in proanthocyanidin (PA) and anthocyanin biosynthesis (*FaF3'H*, *FaDFR*, *FaANS*, *Fa3-GT*, *FaANR* and *FaLAR*), and for the identified transcriptional strawberry regulators (*FaMYB1*, *FaMYB5*, *FaMYB9*, *FaMYB11*, *FaMYC1*, *FabHLH3*, *FabHLH33*, and *FaTTG1*). Measurements were carried out on 14 genotypes at the mature stage (R), and on the four genotypes (Elsanta, Sonata, Holiday × Miranda, and Senga sengana) during fruit development (from G1 to R stage). Data were gathered and organized by the agglomerative hierarchical clustering method, using the EPCLUST tool. The scale indicates log₂ variations: red, increase; green, decrease.

polymerization degree), the anthocyanin pathway (i.e. pelargonidins or cyanidin conjugates and derivatives, total anthocyanins) and the flavonol pathway (i.e. kaempferol or quercetin conjugates). Some other phenolic compounds were also annotated, such as the flavonoid precursor *p*-coumaric acid (and derivative) or the antioxidant ellagic acid (as conjugate and derivative). Transcript accumulation was measured for key structural genes involved in both PA and anthocyanin biosynthesis (*FaF3'H*, *FaDFR*, *FaANS*, *Fa3-GT*, *FaANR* and *FaLAR*), and for the identified transcriptional regulators (*FaMYB1*, *FaMYB5*, *FaMYB9*, *FaMYB11*, *FaMYC1*, *FabHLH3*, *FabHLH33*, and *FaTTG1*).

This global analysis confirms that PA-related metabolites were mainly accumulating early during fruit development, and then decreased during ripening. This analysis also confirmed that the accumulation of the transcripts corresponding to the structural genes specifically involved in PA biosynthesis (i.e. *FaANR* and *FaLAR*) as well as *FaF3'H*, and all the identified regulators, except *FaMYB1*, were following the same trend.

By contrast, anthocyanin- and flavonol-related metabolites were largely accumulated in ripe fruits (R stage), concomitant with the increase in the degree of PA polymerization and the accumulation of transcripts corresponding to known anthocyanin biosynthetic genes, such as *FaDFR*, *FaANS* and *Fa3-GT*. *p*-Coumaric and ellagic acid-related compounds were also mostly accumulated in this later stage.

We observed a good correlation ($R^2 > 0.95$) between expression of *FaMYB9* and *FaMYB11*, their putative target genes *FaF3'H*, *FaANS*, *FaANR* and *FaLAR*, and total PA concentrations at the small green fruit stage. This indicates that the expression of these genes is important in determining genotypic differences in total PA content at this fruit developmental stage. This correlation was absent for *FaTTG1* or any of the *FabHLH/MYC* genes. We did not observe correlation of gene expression at the green or red stage and total PA concentrations at the red ripe fruit stage.

Complementation of *Arabidopsis* transparent testa mutants with strawberry regulators

To characterize the function of the isolated strawberry transcriptional regulator genes, we performed complementation studies by expressing strawberry R2R3-MYB, bHLH and WDR regulatory genes into *Arabidopsis* lines mutated for *AtTT2* (*tt2-1*), *AtTT8* (*tt8-3*) and *AtTTG1* (*ttg1-1*), respectively (Fig. 5a–d). The strawberry MYB and bHLH genes were expressed under the control of the double *CaMV35S* promoter, which provides a constitutive expression pattern.

Based on the results accumulated in the frame of this work, and on available data in the literature, we only selected a subset of genes susceptible to form an active MBW complex in *Arabidopsis* to be tested. *FaMYB5* and *FabHLH3Δ* were not included as they encoded truncated proteins missing structurally important domains, the R2 repeat and the typical bHLH domain, respectively (Fig. S5). As *FaMYC1* and *FabHLH33* were not able to interact with any of the studied MYB TFs (Fig. 3a), these two bHLH proteins were also not studied further.

The overexpression of *FaMYB9* and *FaMYB11* in *tt2-1*, *FabHLH3* in *tt8-3* and *FaTTG1* in *ttg1-1* mutants was able to complement the associated *transparent testa* phenotype (i.e. bright yellow seed colour), giving the mature seeds their typical brown colour (Fig. 5a–c). These observations were further confirmed by measuring PAs accumulated in the corresponding transgenic seeds (Fig. 5d).

These results showed that *FaMYB9* and *FaMYB11*, *FabHLH3* and *FaTTG1* are functional homologues of *AtTT2*, *AtTT8* and *AtTTG1*, respectively, and therefore suggest that they are probably acting together through the formation of an MBW complex.

Assay for functional interactions in *P. patens* protoplasts

As a final attempt to determine if the regulatory gene products from *FaMYB9*, *FabHLH3* and *FaTTG1* can form a MBW complex as suggested by the yeast-two-hybrid experiments, a transient expression assay in the *P. patens* protoplast (Thevenin *et al.*, 2012) was carried out. In this assay, the concerted action of the selected transcriptional regulators on the activity of the *BAN/ANR* promoter (*proBAN*) fused to green fluorescent protein (GFP) was tested.

When these genes were transfected alone or two by two, no clear induction of *proBAN* activity (i.e. increased GFP intensity) was observed (Fig. 5e). However, a strong activation of *proBAN* activity was observed when protoplasts were cotransformed simultaneously with *FaMYB9*, *FabHLH3* and *FaTTG1* (Fig. 5e), emphasizing the fact that these three proteins act together to activate their target promoters.

Discussion

PA biosynthesis decreases during strawberry fruit development

It is generally known that ripe strawberry fruits contain high concentrations of the red coloured anthocyanins, but next to this high concentrations of PAs are found in strawberry fruits (Almeida *et al.*, 2007; Buendia *et al.*, 2010). In unripe green fruits, in particular, we also found a relatively high concentration of PAs; however, this decreased considerably during fruit ripening (Fig. 1b–d), which is in agreement with previously results (Carbone *et al.*, 2009). In accordance with Buendia *et al.* (2010) we found that the concentration of PA in strawberry fruits is genotype-dependent, showing the important role played by the genetic background (Fig. S1).

Expression of *AtTT2* in strawberry fruits is sufficient to modify PA biosynthesis

To determine if the PA regulatory network identified in *Arabidopsis* was functionally conserved in strawberry, we studied the effect of expression in strawberry plants of the well-described *Arabidopsis* transcriptional regulators *AtTT2* (R2R3-MYB), *AtTT8* (bHLH) and *AtTTG1* (WD40 repeat protein), which specifically modulate the expression of genes involved in PA

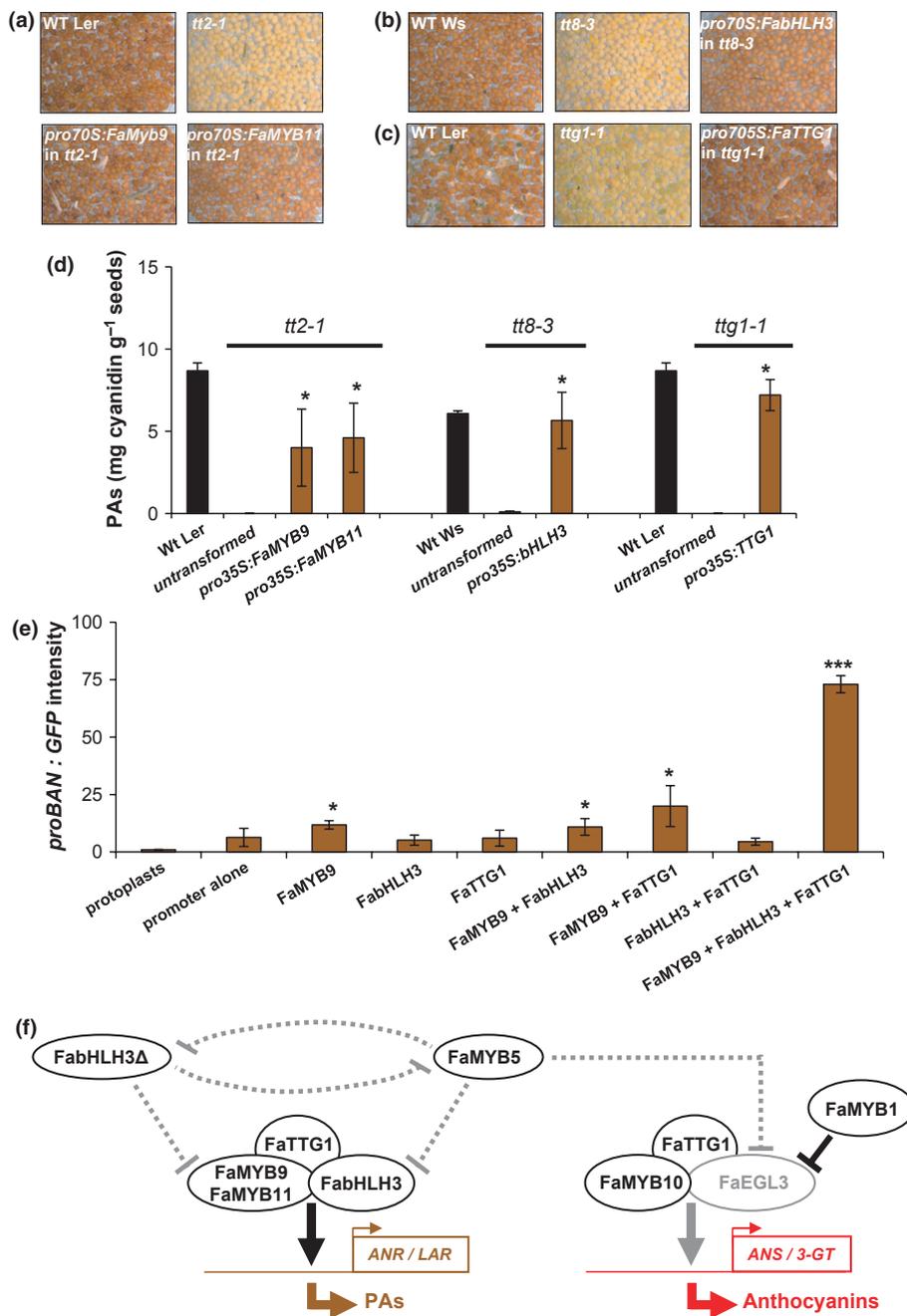


Fig. 5 Functional characterization of FaMYB9/FaMYB11, FabHLH3 and FaTTG1. Complementation of the *Arabidopsis* transparent testa null mutants: over expression of FaMYB9/FaMYB11 (a), FabHLH3 (b) and FaTTG1 (c) complement the *tt2-1*, *tt8-3* and *ttg1-1* seed colour phenotype, respectively. (d) The complemented lines accumulate proanthocyanidin (PA). Wt, wildtype; Ler and Ws, Landsberg erecta and Wassileskija ecotypes; pro35S, double CaMV35S promoter. (e) Effect of FaMYB9, FabHLH3 and FaTTG1 on *BANYULS* (*Arabidopsis* anthocyanin reductase) promoter activity (*proBAN*). *t*-test significant difference: ***, $P < 0.001$; *, $P < 0.05$; error bars show SE (d–e). Average and SE were calculated from the values obtained by measuring three independent plants per line (control and transgenics) in triplicate. (f) Putative transcriptional regulatory pathways controlling PA and anthocyanin biosynthesis in strawberry fruit. Putative proteins and interactions are indicated in grey. For example, this model hypothesized that strawberry orthologous genes of *AtPAP1/AtMYB75* and *AtEGL3/AtbHLH002* (*Arabidopsis* MBW complex regulating anthocyanin biosynthesis) remain to be identified, and that FaMYB5 and FabHLH3Δ are negative regulators. Proteins and interaction symbolized in black were identified in the frame of this study. ANR, anthocyanidin reductase; LAR, leucoanthocyanidin reductase; ANS, anthocyanidin synthase; 3-GT, 3-glycosyltransferase.

biosynthesis in seeds (Baudry *et al.*, 2004; Lepiniec *et al.*, 2006). We found that the combined overexpression of *AtTT2*, *AtTT8* and *AtTTG1* resulted in a rechannelling in the flavonoid pathway in ripe fruits, leading to increased concentrations of PAs and reduced concentrations of anthocyanins, mainly through a partial loss of pelargonidin content (Fig. 2a–e). These metabolic modifications were accompanied by an increase in transcript accumulation for some key structural genes (Figs 1a, 2f), namely *FaANS* (anthocyanidin synthase), *FaANR* (anthocyanidin reductase) and *FaLAR* (leucoanthocyanidin reductase), and, to a lesser extent, *FaF3'H* (flavonoid 3'-hydroxylase). These transcriptional modifications were in agreement with what was observed in *Arabidopsis* where the *AtTT2-AtTT8-AtTTG1* MBW complex was shown to

regulate the later steps of the PA biosynthetic pathway (Baudry *et al.*, 2004; Lepiniec *et al.*, 2006).

Interestingly, we found that these metabolic and expressional changes were also observed when *AtTT2* was expressed alone (and not when either *AtTT8* or *AtTTG1* was expressed). The simplest hypothesis for this is that, in strawberry, functional *AtTT8*-like and *AtTTG1*-like transcriptional regulators are present which are able to act as *AtTT2* interacting partners and form a functional MBW regulatory complex. As such R2R3-MYBs are considered to be highly pathway-specific (Lepiniec *et al.*, 2006), this implies that in strawberry *TT2*-like R2R3-MYB TF may have a key function in regulating PA biosynthesis. These results prompted us to isolate the strawberry homologues of *AtTT2*, *AtTT8* and *AtTTG1*.

Identification of putative AtTT2, AtTT8 and AtTTG1 homologues in strawberry

To identify the putative homologues of the *Arabidopsis AtTT2*, *AtTT8* and *AtTTG1*, two strategies were used. First, the *Arabidopsis* genes were used as bait in a yeast-two-hybrid screen of an expressed strawberry fruit cDNA library, which led to the identification of several transcriptional regulators from three different protein families, the MYB (*FaMYB5* and *FaMYB11*), the bHLH (*FaMYC1* and *FabHLH33*) and the WD40-repeat proteins (*FaTTG1*; Figs 3a, S4). As TF interactions are known to be functionally conserved (Immink & Angenent, 2002), the identified strawberry proteins may be functional counter partners in the formation of the MBW complex regulating the expression of genes involved in PA biosynthesis. Secondly, using a PCR-based strategy, additional MYB (*FaMYB9*) and bHLH (*FabHLH3* and *FabHLH3Δ*) genes were isolated from strawberry fruit. Comparison of deduced protein sequences of the identified strawberry TF genes with characterized TFs from other plant sources show that all shared homology with TFs involved in the regulation of flavonoid biosynthetic genes, and in particular with genes involved in PA biosynthesis (Fig. 3b–d).

When all the identified regulators were assayed against each other in yeast (12×9 combination matrix), some interactions were found to be specific to the domain (i.e. GAL4 DNA-binding, BD, or activation, AD, domains) to which the proteins were fused. This usual result could be explained by conformational changes of the proteins fused to AD or DB domains (Burbulis & Winkel-Shirley, 1999) and has previously been described for MYB, bHLH and WDR proteins (e.g. by Baudry *et al.*, 2004). From this study we found that *AtTT8*, *FabHLH3* and *FabHLH3Δ* were able to interact with all the studied MYBs, whereas *FaMYC1* and *FabHLH33* were not. *FaMYC1* and *FabHLH33* could form homo- and heterodimers, which was previously reported for other bHLH proteins (Feller *et al.*, 2006). Both *AtTTG1* and *FaTTG1* were found to interact with nearly all the tested MYB and bHLH TFs, with the exception of *FaMYB1*, *FaMYB5* and *FabHLH3Δ*.

The results obtained through this yeast-two-hybrid interaction matrix suggested that *FaMYB9* and *FaMYB11* could be the *AtTT2* homologues in strawberry, as they were the only R2R3-MYB found to interact with *FaTTG1*. Interestingly, as *FabHLH3* was found to interact with *FaTTG1*, as well as with *FaMYB9* and *FaMYB11*, *FabHLH3* and *FaTTG1* could likely be the strawberry orthologues of *AtTT8* and *AtTTG1*, respectively.

Coregulation study confirm the putative involvement of most of the studied transcriptional regulator in PA biosynthesis

The expression pattern of the genes encoding the enzymatic steps of the PA biosynthetic pathway have been previously studied during strawberry development; however, there was hardly any information available on their regulation (Almeida *et al.*, 2007; Carbone *et al.*, 2009). One way to identify putative regulators that are specific to a particular biosynthetic pathway is to search for

coregulation. Such coregulation can be found between TF and putative targets like structural genes, but can also include any compound related to the studied pathway. We took the advantage of the availability of a large set of strawberry genotypes (14 different) to measure in ripe fruit (all genotypes), and during fruit development (four genotypes, Fig. 1d), the accumulation of diverse flavonoid compounds (including flavonols, anthocyanins, PAs, and the *p*-coumaric acid precursor) as well as the antioxidant ellagic acid. In parallel, transcript accumulation was also measured for six key structural genes (i.e. *FaF3'H*, *FaDFR*, *FaANS*, *Fa3-GT*, *FaANR*, and *FaLAR*) and eight of the studied transcriptional regulators (i.e. *FaMYB1*, *FaMYB5*, *FaMYB9*, *FaMYB11*, *FaMYC1*, *FabHLH3*, *FabHLH33*, and *FaTTG1*). These data were clustered with the aim of identifying MYB and bHLH TFs that were specific to PA biosynthesis in strawberry fruits (Vilo *et al.*, 2003).

The results obtained from this global correlation analysis support the hypothesis that all identified transcriptional regulators are probably involved in PA biosynthesis, and were not involved either in anthocyanin or flavonol biosynthesis, or in *p*-coumaric and ellagic acid accumulation (Figs 1b–d, 4). The involvement of *FaMYB1* in the regulation of the anthocyanin biosynthetic pathway (Aharoni *et al.*, 2001) was confirmed.

FaMYB9/FaMYB11, *FabHLH3* and *FaTTG1* are the strawberry functional homologues of *AtTT2*, *AtTT8* and *AtTTG1*, respectively

During the course of this study, we found that *FaMYB9* and *FaMYB11* were able to interact with *AtTT8* and *AtTTG1*, and that *FaMYB9* and *FaMYB11* were encoding two R2R3-MYB TFs highly similar at the amino acid level to *AtTT2*. When expressed in the *Arabidopsis* yellow seeded *tt2-1* mutant, these two strawberry R2R3-MYB TFs were able to restore the wildtype brown seed colour phenotype. Because *FaMYB9* and *FaMYB11* map to different positions on the *F. vesca* genome, it is most likely that *FaMYB9* and *FaMYB11* are paralogous genes (i.e. generated by a gene duplication). However, despite the paralogous nature of *FaMYB9* and *FaMYB11*, we were not able to demonstrate a clear divergence in gene expression. Both show a similar expression pattern during fruit development (decreasing from the G1 to the R stage), independent of the studied cultivars (Fig. 4). Altogether, these data suggested that *FaMYB9* and *FaMYB11* are two potential functional homologues of *AtTT2* (Figs 3–5, S4).

Similar analysis leads us to postulate that the strawberry bHLH gene *FabHLH3* could be functionally orthologous to *AtTT8*. *FabHLH3* encodes an R/B-like (class IIIif) bHLH TF highly similar to *AtTT8*, which was found to interact in yeast with *FaMYB9*, *FaMYB11* and *FaTTG1* (Heim *et al.*, 2003; Zimmermann *et al.*, 2004). Expression of *FabHLH3* in the *Arabidopsis tt8-3* null mutant resulted in complementation of the *transparent testa* phenotype, through the accumulation of PAs in the seed coat, confirming that *FabHLH3* is a functional homologue of *AtTT8* (Figs 3, 5, S4).

FaTTG1 is the only one gene encoding a WD40 repeat protein highly similar to *AtTTG1* that was identified. TTG1-like proteins were described as being able to interact with several

proteins simultaneously, and seem to have a docking function, stabilizing the ternary TF complex, rather than a direct regulatory function (Baudry *et al.*, 2004; Lepiniec *et al.*, 2006; Hichri *et al.*, 2011). This latter specificity was in agreement with the interaction studies where FaTTG1 was found to interact with FaMYB9, FaMYB11 and FabHLH3. Complementation experiments carried out in *Arabidopsis* confirmed that *FaTTG1* was the functional homologue of *AtTTG1*.

In addition, the transactivation assays carried out in *P. patens* protoplasts also supported the idea that *FaMYB9*, *FabHLH3* and *FaTTG1* functionally interact to regulate PA synthesis during strawberry fruit development (Fig. 5f).

FaMYB5 and FabHLH3Δ: two new negative regulators of flavonoid biosynthesis?

FaMYB5, which was identified as an AtTT8-interacting partner, is more similar to AtMYB5 than to AtTT2. In *Arabidopsis*, AtMYB5 in combination with AtTT8 and AtTTG1 form a ternary complex that was shown to regulate outer seed coat development and trichome morphogenesis, and was proposed to play a role in regulating PA biosynthesis in seeds (Gonzalez *et al.*, 2009; Li *et al.*, 2009). In petunia, the *AtMYB5* homologue *PhPH4* is involved in vacuolar acidification in petal cells (Quattrocchio *et al.*, 2006), whereas the grapes homologues VvMYB5a and VvMYB5b play a role in the regulation of anthocyanin and PA biosynthesis in berries (Deluc *et al.*, 2006, 2008). Altogether, these data suggest that the R2R3-MYB that are characterized by the presence of two 'G20' motifs (Quattrocchio *et al.*, 2006) can have divergent functions (i.e. development vs secondary metabolism), even if their main role in *planta* seems to be the regulation of flavonoid biosynthesis. Beside this general sequence feature, FaMYB5 differs from the other G20-subgroup members by the lack of the R2-repeat in its DNA binding domain (Fig. S5). Recently, similar single-repeat R3-MYB TFs were identified in *Arabidopsis* as negative regulators of flavonoid biosynthesis (Dubos *et al.*, 2008; Matsui *et al.*, 2008; Zhu *et al.*, 2009). For example, ectopic expression of *AtMYBL2* in seeds inhibited both the expression of genes involved in PA biosynthesis and PA accumulation, whereas *atmybl2* loss-of-function mutants presented an enhanced anthocyanin accumulation, and transcript accumulation of the corresponding genes in vegetative tissues. Moreover, AtMYBL2 was found to interact in yeast with AtTT8, and to be able to inhibit the AtTT2–AtTT8–AtTTG1 complex activity in both *Arabidopsis* and the *P. patens* protoplast, demonstrating the repression role of AtMYBL2 through the direct interference of the MBW complex activity (Dubos *et al.*, 2010; Thevenin *et al.*, 2012). The fact that FaMYB5 interacted with AtTT8 and its strawberry homologue FabHLH3 in yeast suggests that this TF may have a similar inhibitory role in strawberry flavonoid biosynthesis (Fig. 5f). Because FaMYB5 displayed a steady amount of transcript accumulation during strawberry fruit development, it can be hypothesized that FaMYB5 may have a role in fine-tuning both PA biosynthesis during early fruit development and anthocyanin biosynthesis during fruit ripening. Interestingly, it was recently found in petunia that the R3-MYB, PhMYBx, acts in

concert with an R2R3-MYB repressor, PhMYB27, to negatively regulate the PhPHZ–PhAN1–PhAN11 MBW complex activity in order to modulate anthocyanin biosynthesis in vegetative tissues (Albert *et al.*, 2011). Based on this work, it is tempting to speculate that such a mechanism may exist in strawberry fruits, where FaMYB5 could act in combination with the negative regulator FaMYB1 (Aharoni *et al.*, 2001) to regulate the activity of the MBW complexes involved in anthocyanin and/or PA biosynthesis (Fig. 5f).

Along the same lines, *FabHLH3Δ* encodes a truncated TF that misses the bHLH and ACT-like dimerization domains (Fig. S5a). This small bHLH may act as dominant repressor because the remaining MYB-interacting domain is still present and able to bind to R2R3-MYB proteins (such as FaMYB9 or FaMYB11), thereby blocking the binding sites for functional bHLH TFs. A similar regulatory mechanism was proposed for the maize *ZmIN1* bHLH factor. The mis-splicing of *ZmIN1* results in premature termination of translation, producing a protein lacking the bHLH domain, which was proposed to be a putative anthocyanin biosynthesis negative regulator (Burr *et al.*, 1996). Because *FabHLH3Δ* was also found to interact with FaMYB5, and because both were expressed early during fruit development, it is possible that *FabHLH3Δ* may act as an inhibitor of the MYB negative regulator (or vice-versa), in a negative feedback regulatory loop (Figs 4, 5f).

The possibility that *FaMYB5* and *FabHLH3Δ* may repress flavonoid (i.e. PAs and/or anthocyanins) biosynthesis in strawberry fruits will be further investigated using both genetic and molecular approaches (e.g. loss-of-function in strawberry, heterologous mutant complementation, protein activity characterization, etc.).

Acknowledgements

We thank Dr Richard Immink and Marco Busscher for their excellent help in setting up yeast-two-hybrid screening. This research was carried out as part of the EU-funded project FLAVO (Food CT-2004-513960). The work of W.X. was generously supported by the China Scholarship Council (CSC).

References

- Aaby K, Mazur S, Nes A, Skrede G. 2012. Phenolic compounds in strawberry (*Fragaria × ananassa* Duch.) fruits: composition in 27 cultivars and changes during ripening. *Food Chemistry* 132: 86–97.
- Abrahams S, Lee E, Walker AR, Tanner GJ, Larkin PJ, Ashton AR. 2003. The *Arabidopsis* TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. *Plant Journal* 35: 624–636.
- Aharoni A, De Vos CH, Wein M, Sun Z, Greco R, Kroon A, Mol JN, O'Connell AP. 2001. The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant Journal* 28: 319–332.
- Akagi T, Ikegami A, Tsujimoto T, Kobayashi S, Sato A, Kono A, Yonemori K. 2009. DkMyb4 is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. *Plant Physiology* 151: 2028–2045.
- Akagi T, Tsujimoto T, Ikegami A, Yonemori K. 2011. Effects of seasonal temperature changes on *DkMyb4* expression involved in proanthocyanidin regulation in two genotypes of persimmon (*Diospyros kaki* Thunb.) fruit. *Planta* 233: 883–894.

- Albert NW, Lewis DH, Zhang H, Schwinn KE, Jameson PE, Davies KM. 2011. Members of an R2R3-MYB transcription factor family in *Petunia* are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant Journal* 65: 771–784.
- Almeida JRM, D'Amico E, Preuss A, Carbone F, de Vos CHR, Deiml B, Mourgues F, Perrotta G, Fischer TC, Bovy AG *et al.* 2007. Characterization of major enzymes and genes involved in flavonoid and proanthocyanidin biosynthesis during fruit development in strawberry (*Fragaria × ananassa*). *Archives of Biochemistry and Biophysics* 465: 61–71.
- Battino M, Jules B, Beatrice D-R, Laimer M, McDougall GJ, Bruno M. 2009. Bioactive compounds in berries relevant to human health. *Nutrition Reviews* 67: S145–S150.
- Baudry A, Caboche M, Lepiniec L. 2006. TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *Plant Journal* 46: 768–779.
- Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L. 2004. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant Journal* 39: 366–380.
- Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP. 2007. The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. *Plant Physiology* 143: 1347–1361.
- Bombarely A, Merchante C, Csukasi F, Cruz-Rus E, Caballero JL, Medina-Escobar N, Blanco-Portales R, Botella MA, Munoz-Blanco J, Sanchez-Sevilla JF *et al.* 2010. Generation and analysis of ESTs from strawberry (*Fragaria × ananassa*) fruits and evaluation of their utility in genetic and molecular studies. *BMC Genomics* 11: 503.
- Buendia B, Gil MI, Tudela JA, Gady AL, Medina JJ, Soria C, Lopez JM, Tomas-Barberan FA. 2010. HPLC-MS analysis of proanthocyanidin oligomers and other phenolics in 15 strawberry cultivars. *Journal of Agricultural and Food Chemistry* 58: 3916–3926.
- Burbulis IE, Winkler-Shirley B. 1999. Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. *Proceedings of the National Academy of Sciences, USA* 96: 12929–12934.
- Burr FA, Burr B, Scheffler BE, Blewitt M, Wienand U, Matz EC. 1996. The maize repressor-like gene intensifier1 shares homology with the r1/b1 multigene family of transcription factors and exhibits missplicing. *Plant Cell* 8: 1249–1259.
- Carbone F, Preuss A, de Vos RCH, D'Amico E, Perrotta G, Bovy AG, Martens S, Rosati C. 2009. Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits. *Plant, Cell & Environment* 32: 1117–1131.
- Chalfun A, Franken J, Mes JJ, Marsch-Martinez N, Pereira A, Angenent GC. 2005. ASYMMETRIC LEAVES2-LIKE1 gene, a member of the AS2/LOB family, controls proximal-distal patterning in *Arabidopsis* petals. *Plant Molecular Biology* 57: 559–575.
- Curtis MD, Grossniklaus U. 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiology* 133: 462–469.
- Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde JP, Merillon JM, Hamdi S. 2006. Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. *Plant Physiology* 140: 499–511.
- Deluc L, Bogs J, Walker AR, Ferrier T, Decendit A, Merillon JM, Robinson SP, Barrieu F. 2008. The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. *Plant Physiology* 147: 2041–2053.
- Dubos C, Le Gourrierec J, Baudry A, Huep G, Lanet E, Debeaujon I, Routaboul JM, Alboresi A, Weisshaar B, Lepiniec L. 2008. MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant Journal* 55: 940–953.
- Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. 2010. MYB transcription factors in *Arabidopsis*. *Trends in Plant Science* 15: 573–581.
- Feller A, Hernandez JM, Grotewold E. 2006. An ACT-like domain participates in the dimerization of several plant basic-helix-loop-helix transcription factors. *Journal of Chemical Biology* 281: 28964–28974.
- de Folter S, Immink RG, Kieffer M, Parenicova L, Henz SR, Weigel D, Busscher M, Kooiker M, Colombo L, Kater MM *et al.* 2005. Comprehensive interaction map of the *Arabidopsis* MADS Box transcription factors. *Plant Cell* 17: 1424–1433.
- Gonzalez A, Mendenhall J, Huo Y, Lloyd A. 2009. TTG1 complex MYBs, MYB5 and TT2, control outer seed coat differentiation. *Developmental Biology* 325: 412–421.
- Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC. 2003. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Molecular Biology and Evolution* 20: 735–747.
- Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V. 2011. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *Journal of Experimental Botany* 62: 2465–2483.
- Immink RG, Angenent GC. 2002. Transcription factors do it together: the hows and whys of studying protein-protein interactions. *Trends in Plant Science* 7: 531–534.
- James P, Halladay J, Craig EA. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144: 1425–1436.
- Kennedy JA, Jones GP. 2001. Analysis of proanthocyanidin cleavage products following acid-catalysis in the presence of excess phloroglucinol. *Journal of Agricultural and Food Chemistry* 49: 1740–1746.
- Lazo GR, Stein PA, Ludwig RA. 1991. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Biotechnology* 9: 963–967.
- Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M. 2006. Genetics and biochemistry of seed flavonoids. *Annual Review of Plant Biology* 57: 405–430.
- Li SF, Milliken ON, Pham H, Seyit R, Napoli R, Preston J, Koltunow AM, Parish RW. 2009. The *Arabidopsis* MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *Plant Cell* 21: 72–89.
- Matsui K, Umemura Y, Ohme-Takagi M. 2008. AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Plant Journal* 55: 954–967.
- Moco S, Bino RJ, Vorst O, Verhoeven HA, de Groot J, van Beek TA, Vervoort J, de Vos RCH. 2006. A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiology* 141: 1205–1218.
- Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L. 2000. The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in *Arabidopsis* siliques. *Plant Cell* 12: 1863–1878.
- Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L. 2001. The *Arabidopsis* TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* 13: 2099–2114.
- Nijveldt RJ, Nood E, van Hoorn DEC, van Boelens PG, van Norren K, van Leeuwen PAM. 2001. Flavonoids: a review of probable mechanisms of action and potential applications. *The American Journal of Clinical Nutrition* 74: 418–425.
- Oosumi T, Ruiz-Rojas JJ, Veilleux RE, Dickerman A, Shulaev V. 2010. Implementing reverse genetics in Rosaceae: analysis of T-DNA flanking sequences of insertional mutant lines in the diploid strawberry, *Fragaria vesca*. *Physiologia Plantarum* 140: 1–9.
- Quattrocchio F, Verweij W, Kroon A, Spelt C, Mol J, Koes R. 2006. PH4 of *Petunia* is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway. *Plant Cell* 18: 1274–1291.
- Routaboul JM, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, Einhorn J, Lepiniec L. 2006. Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. *Planta* 224: 96–107.
- Schaart JG, Salentijn EMJ, Krens FA. 2002. Tissue-specific expression of the beta-glucuronidase reporter gene in transgenic strawberry (*Fragaria × ananassa*) plants. *Plant Cell Reports* 21: 313–319.
- Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane, *et al.* 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nature Genetics* 43: 109–116.

- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.
- Terrier N, Torregrosa L, Ageorges A, Vialet S, Verries C, Cheynier V, Romieu C. 2009. Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiology* 149: 1028–1041.
- Thevenin J, Dubos C, Xu W, Le Gourrierec J, Kelemen Z, Charlot F, Nogue F, Lepiniec L, Dubreucq B. 2012. A new system for fast and quantitative analysis of heterologous gene expression in plants. *New Phytologist* 193: 504–512.
- Vilo J, Kapushesky M, Kemmeren P, Sarkans U, Brazma A. 2003. *Expression profiler*. In: Parmigiani G, Garrett ES, Izarray R, Zeger SL, eds. *The analysis of gene expression data: methods and software*. New York, USA: Springer Verlag, 142–162.
- Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC. 1999. The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell* 11: 1337–1349.
- Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. 2003. Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* 299: 396–399.
- Zhu HF, Fitzsimmons K, Khandelwal A, Kranz RG. 2009. CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Molecular Plant* 2: 790–802.
- Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF. 2004. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant Journal* 40: 22–34.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Liquid chromatography mass spectrometry (LC-MS) analysis of soluble flavan-3-ol in 64 different strawberry genotypes.

Fig. S2 Correlation between PA dimers and (+)-catechin.

Fig. S3 Total PA measured in control and transgenic strawberry leaves and stem.

Fig. S4 Yeast-two-hybrid interaction matrix.

Fig. S5 Detailed amino acid sequence analysis of FaMYB5 and FabHLH3Δ.

Table S1 Primers used for quantitative RT-PCR

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



New Phytologist Tansley Medal

For excellence in plant science

Full details, terms and conditions at:
www.newphytologist.org

Calling all early stage career scientists!

Deadline for submissions for 2013: 15 December 2012

Win £2000 (GBP) and have your work highlighted in *New Phytologist*, one of the world's leading plant science journals (2011 Impact Factor 6.645).

- The New Phytologist Tansley Medal is awarded annually in recognition of an outstanding contribution to research in plant science
- This is a global competition open to all plant scientists in the early stages of their career and includes both student and post-doctoral researchers with up to five years experience, excluding career breaks, since gaining/defending their PhD
- Selection is based on a two-stage process:
 - **Stage 1)** Submit your CV, a personal statement and reference: Deadline 15 December 2012
 - **Stage 2)** Submission of a single-authored minireview intended for publication: Deadline: 31 March 2013
- All competition articles that are accepted after peer review will be published in *New Phytologist* and the Tansley medal winner selected by the judges from these final papers.