Genes controlling the development and function of plant vacuoles

Li, Y.

Publication date
2017

Document Version
Other version

License
Other

Citation for published version (APA):
Li, Y. (2017). Genes controlling the development and function of plant vacuoles. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The conserved MBW complex regulates distinct pathways by activating different sets of target genes

Yanbang Li, Cornelis Spelt, Mattijs Bliek, Takuji Wada, Rumi Tominaga, Aureliano Bombarely, Gabino Sanchez Perez, Henry van de Geest, Francesca Quattrocchio and Ronald Koes
Abstract

The formation of vacuolinos, small vacuoles coexisting with the central vacuole in petunia petal epidermal cells, was shown to be controlled by a set of conserved transcription factors, AN1-PH4-PH3. These proteins are part of the WMBW transcription complex, which is highly conserved among very different plant species. Based on RNAseq data, we have identified 750 genes which expression is significantly differential between wild type and anl mutant petals, 692 genes differential in ph4 and 600 in ph3. Among these genes, 31 genes were differential in all anl, ph4 and ph3 mutants and therefore are potentially involved in the vacuolinos pathway. To date, based on trafficking of vacuolar proteins in RNA interference lines for CAC12.3 and PAT1 we identify these as genes required for different steps of the vacuolinos pathway. In addition, the comparison of the set of target genes of AN1-PH3-PH4 in petunia petals with those of the homologs in Arabidopsis leaves, resulted in the identification of a rather small group of overlapping targets. This suggests that the different functions of the WMBW complex in different developmental pathways like petals epidermal cells differentiation in petunia and trichome formation in Arabidopsis, are due to differences in the regulatory regions of their target genes, rather than in the regulatory proteins. This is consistent with the results of the analysis of the petunia MYBx gene. MYBx belongs to the R3 MYB family, like the Arabidopsis TRY and CPC factors, which are regulators of (non) hair differentiation in Arabidopsis. MYBx interacts with the AN1, participating in this way in the WMBW complex, probably as inhibitor. MYBx inhibits trichome formation and rescues root hair defects when expressed in Arabidopsis cpc mutants, while no trichome phenotype was detectable in transgenic plants expressing MYBx, which are instead affected in anthocyanin accumulation andvacuolar acidification in petals. Together, our findings provide novel insight into the function of the WMBW complexes in epidermal cells of different species and on the mechanisms by which these transcription factors acquired different functions during evolution.
Introduction

MBW complexes consisting of a tryptophan-aspartic acid (W-D) dipeptide repeat protein (WDR), a basic Helix Loop Helix (bHLH) and MYB proteins act as transcription activators of genes involved in various pathways involved in the differentiation of (sub-)epidermal plant cells. They were first identified as transcription activators of structural genes encoding enzymes of the pathway synthesizing anthocyanin pigments that color many flowers and fruits or the related proantocyanidins (also known as condensed tannins) that provide a brown color to seeds. This role of MBW complexes is widely conserved among flowering plants, suggesting that it was established early on in evolution, when anthocyanins first appeared in early angiosperms, before the separation of dicots and monocots.

Research in Arabidopsis and petunia in particular revealed that MBW complexes control a plethora of other processes affecting the differentiation of epidermal cells that seem restricted to smaller groups of species, suggesting that these functions evolved more recently (Koes et al., 2005; Ramsay & Glover, 2005). Genetic evidence shows MBW complexes control in Arabidopsis besides anthocyanin and tannin synthesis also the formation of mucilage in seeds, the formation of trichomes on aerial tissues, and the identity of non-hair cells in the root epidermis (Broun, 2005; Ishida et al., 2008). The role of MBW complexes in the specification of (non-)hair cells seems specific for Arabidopsis and presumably closely related species, and is not seen in more distantly related dicots, like petunia, or monocots, like maize. In petunia, on the other hand, MBW complexes were shown to control, besides anthocyanin and tannin synthesis (Spelt et al., 2002; Verweij et al., 2016), the stability of anthocyanins (Quattrocchio et al., 2006), the development of novel vacuole like compartments, known as vacuolinos that exist besides the central vacuole (Faraco et al., 2017), the acidification of vacuoles (Quattrocchio et al., 2006; Verweij et al., 2008, 2016; Faraco et al., 2014), and cell divisions in the seed coat epidermis (Verweij et al., 2016).

How very similar, and sometimes orthologous, MBW complexes acquired control over distinct processes during evolution or how they can activate different downstream pathways in distinct tissues is largely unclear. The highly pleiotropic phenotypes of mutants the WD repeat and bHLH components of the MBW complex(es), suggest that they are required for the activation of nearly all downstream pathways, and that the specificity of MBW complexes originates in part from the specific MYB proteins with which they hook up (Koes et al., 2005; Ramsay & Glover, 2005).

In petunia for example, the WDR protein ANTHOCYANIN11 (AN11), which is expressed everywhere in the plant and is together with the bHLH protein AN1 essential for anthocyanin production (de Vetten et al., 1997), tannin production (Zenoni et al., 2011), vacuolar hyper-acidification (Verweij et al., 2008; Faraco et al., 2014) and the formation of vacuolinos (Faraco et al., 2017; this thesis chapter 4) .The AN1 and AN11 proteins can form complexes with different MYB proteins, such as PH4 (Quattrocchio et al., 2006), which is essential for vacuolar acidification and vacuolino formations (Quattrocchio et al., 2006; Faraco et al., 2017; this thesis chapter 4) or AN2,
which is needed to activate anthocyanin synthesis (Quattrocchio et al., 1999). The formation of vacuolinos and the hyper-acidification of vacuoles in petunia petals requires in addition a fourth component, a WRKY transcription factor encoded by PH3 (Verweij et al., 2016) (Faraco et al., 2017; this thesis chapter 4), which is not essential in other pathways (like production of pigments in petals) where it only acts as intensifier (Verweij et al., 2016).

In *Arabidopsis* vegetative tissues, homologs of the AN11-AN1-PH4 complex, consisting of the AN11 homolog TRANSPARENTA TESTA GLABRA1 (TTG1) (Walker et al., 1999), the paralogous bHLH proteins, GLABRA3 (GL3) or ENHANCER OF GLABRA3 (EGL3) (Payne et al., 2000; Zhang et al., 2003) and selected MYB proteins, regulate anthocyanin production and trichome formation (Feyissa et al., 2009; Pesch & Hülskamp, 2009; Xu et al., 2014). In epidermal leaf and stem cells, TTG1 and GL3 interact with the R2R3MYB protein GL1 to control trichome formation and in epidermal root cells with the GL1-paralog WEREWOLF (WER) to specify a non-hair (atrichoblast) fate. To specify (non-) hair cells and mucilage formation in seeds this MBW complex requires the interacting WRKY protein TTG2 (Pesch et al., 2014). TTG2 is orthologous and functionally interchangeable with PH3 from petunia and acts to add specificity to the MBW complex with respect to the downstream pathways that are activated, just like PH3 in petunia (Verweij et al., 2016).

The findings that (W)MBW complexes consisting of orthologous and often functionally interchangeable WD repeat, bHLH, MYB, WRKY proteins activate such different pathways in distinct species, suggests that their roles diverged in part through changes in downstream genes. Several studies have identified putative target genes of the MBW complex involved in trichome formation in *Arabidopsis*, but little is known about MBW target genes involved in other processes in distinct species, such as vacuolino formation and trafficking of vesicles and proteins in petunia.

The TTG1-GL3/EGL3-GL1 complex induces in *Arabidopsis* leaves the expression of a homeodomain protein, GL2, which is a positive regulator of trichome formation (Rerie et al., 1994), and multiple single-repeat R3 MYB transcription factor (R3 MYB) genes, including *TRIPTYCHON* (TRY), *CAPRICE* (CPC) and *CAPRICE-LIKE MYB3* (CPL3). These R3 MYBs can move from a trichome precursor cell, where they are synthesized, to neighboring cells where they compete with GL1 for binding to GL3 or EGL3, to inhibit the formation of an active MBW complex and specification of trichome fate (Schellmann et al., 2002; Kurata et al., 2005; Digiuni et al., 2008; Zhao et al., 2008; Wester et al., 2009). The TTG1-GL3/EGL3-GL1 complex regulates, besides the transcription regulators GL2, TRY CPC and CPL3 a plethora of tother (structural) genes which function is currently unknown.

In order to identify genes regulated by the AN11-AN1-PH4-PH3 WMBW complex in petunia petals, we analyzed the transcriptomes of petunia wild type, *an1, ph3* and *ph4* petals by RNAseq analysis. Among the target genes of AN1, PH3 and PH4 transcription factors, we expect to find a group of genes required for the activation of the petal epidermal specific protein sorting pathway that involves the formation of vacuolinos. These small vacuoles, that coexist with the central vacuole where pigments accumulate, are interesting study objects because of their (yet) undefined function in the differentiation
of petal cells, their contribution to the delivery of proteins and possibly other material to the vacuole and their peculiar mechanism of fusion to the central vacuole that involves the interaction of SNAREs with structural proteins localized on the tonoplast (Faraco et al., 2017; this thesis chapter 4). Our results show that AN1, PH3 and PH4 regulate the expression of an unexpectedly large set of genes, and that PH3 and PH4 activate surprisingly large sets of genes independent from AN1, suggesting that they also operate in other transcription factor complexes than WMB(s). Together AN1, PH3 and PH4 regulate a similar set of transcription regulators, as the TTG1-GL3/EGL3-GL1-WRKY complex does in Arabidopsis leaves, but a very different set of structural genes. Among the latter we identified by functional analysis two new genes require for proteins trafficking via vacuolinos.

**Results**

**Transcriptome profiling analyses**
To identify genes controlled by the AN1-PH4-PH3 complex, we performed RNAseq analysis on petals from isogenic wild type, an1, ph4 and ph3 mutant lines. We grew these lines side by side in a normal greenhouse and harvested petal limbs from floral bud of developmental stage 4 to 7 (Koes et al., 1989). We pooled equal amount of RNA of stage 4, 5, 6 and 7 petals limb from each genotype for sequencing. The comparison of these transcriptomes identified 750 genes (327 up-regulated and 423 down-regulated), 692 genes (366 up-regulated and 326 down-regulated) and 600 genes (362 up-regulated and 238 down-regulated) significantly differentially expressed in an1, ph4 and ph3 petals, respectively, as compared to wild-type petal based on negative binomial normalization using edgeR analysis (fold-change >2, P-value < 0.05) (Fig. 1a).

To obtain more information about the function of AN1, PH4 and PH3 in petunia petal, we performed KEGG pathway (Ogata et al., 1999) enrichment analysis on these differentially expressed genes (DEGs) (Fig. 1b). We found among the genes that are differentially expressed in an1, ph4 and ph3 petals, when compared to wild-type, petals, respectively 16, 16 and 5 KEGG pathways that were significantly enriched (P-value < 0.05) compared to the transcriptome of P. axillaris. Among these pathways, genes belonging to the flavonoid biosynthesis (00941) and the petose and glucuronate interconversions (00040) were found to be down-regulated in all three mutants. Furthermore, differentially expressed genes belonging to the ascorbate and aldarate metabolism (00053), phenylpropanoid biosynthesis (00940) and circadian rhythm-plant (04712) regulation showed enrichment in an1 and ph4, while photosynthesis-antenna proteins (00196) showed enrichment only in an1 and ph3 mutants.
Fig. 1 Analysis of transcriptomes from wild type and mutant petals. (a) The numbers of up-regulated and down-regulated genes in an1, ph3 and ph4 petals compared to wild type petals. (b) Significantly (P-value < 0.05) enriched KEGG pathways that are up-regulated or down-regulated in an1, ph4 and/or ph3 petals. The KEGG pathways were categorized into 10 groups: I, Carbohydrate metabolism; II, Energy metabolism; III, Lipid metabolism; IV, Amino acid metabolism; V, Amino acid metabolism; VI, Metabolism of cofactors and vitamins; VII, Metabolism of terpenoids and polyketides; VIII, Biosynthesis of other secondary metabolites; IX, Signal transduction; X, Environmental adaptation.

**Flower pigmentation genes function as internal controls of the RNAseq**

To confirm the RNAseq results and gain further insight into the regulation of the anthocyanin pathway, we examined the expression of selected key structural anthocyanin genes in the RNAseq data (Fig. 2a). Previous studies (Quattrocchio et al., 1993, 1998; de Vetten et al., 1997; Spelt et al., 2002) revealed that the structural genes of the anthocyanin pathway in petunia can be divided into two sets: early and late genes. Early genes, which encode enzymes that catalyze the first steps of the pathway and deliver the precursors for different flavonoids (flavanones, flavonols, anthocyanins, etc.) via distinct side branches, are expressed independently from the WMBW complex (with the exception of CHS/J), while late genes, encoding enzymes specific for the anthocyanin sidebranch of the pathway, are transcriptionally regulated by the WMBW complex containing the MYB protein AN2, and, except for MF, are not affected by mutations in the PH4 MYB.

Consistent with prior results obtained by RNase protection, RNA gel blot and (q)RT-PCR (Quattrocchio et al., 2006; Provenzano et al., 2014), the RNAseq results showed that the expression of CHALCONE SYNTHASE j (CHS/J), DIHYDROFLAVONOL 4-REDUCTASE (DFR), ANTHOCYANIDIN SYNTHASE (ANS), ANTHOCYANIN ACYLTRANSFERASE (AAT), METHYLATION AT THREE (MT) and METHYLATION AT FIVE (MF) was strongly reduced in an1 petals. In addition we found that the expression of UDP GLUCOSE:FLAVONOID 3-O-GLUCOSYLTRANSFERASE (3GT), UDP GLUCOSE:FLAVONOID 5-O-GLUCOSYLTRANSFERASE (5GT), RHAMNOSYLATION AT THREE (RT), and HYDROXYLATION AT FIVE (HF1) was also reduced in an1 petal at the RNA level (Fig. 2).
Fig. 2 Regulation of structural genes involved in flower pigmentation (a) Relative expression levels of structural anthocyanin genes and PH5 in *an1*, *ph3* or *ph4* petals, as compared to wild type petals. Blue shading indicates increased expression and green shading reduced expression in mutant petals. Gray shading indicates that expression in the mutant is not significantly different from wild type. (b) Expression of target genes encoding GST-like proteins.
The early genes CHALCONE SYNTHASE a (CHS\textsubscript{a}) and FLAVONOL 3’ HYDROXYLASE (F3’H) are expressed at similar levels in wild type, an1 and ph4 petals. The expression of DFR was also reduced in ph3 petals. These results are consistent with previous data from the analysis of single transcripts (Quattrocchio \textit{et al.}, 2006; Verweij \textit{et al.}, 2016).

Surprisingly, the expression of MF, HF1 and HF2 is significantly different in ph4 petal and the expression of CHSA, CHI, HT1, MF, FLAVONOL SYNTHASE (FLS), 3GT and HF2 is significantly affected in ph3 petal. These results indicate that PH4 and PH3 might play a larger role in the regulation of anthocyanin synthesis pathway than previously thought.

After their synthesis anthocyanins are sequestered in the large central vacuole of epidermal petal cells. It has been shown glutathione S-transferase (GST)-like proteins are required for vacuolar sequestration of pigments in maize (BRONZE2, ZmBZ2) (Marrs \textit{et al.}, 1995), petunia (PhAN9) (Mueller \textit{et al.}, 2000) and \textit{Arabidopsis} (TRANSPARENT TESTA 19, AtTT19) (Kitamura \textit{et al.}, 2004). Beside PhAN9, 14 other genes encoding GSTs are differentially expressed in an1 petals (Fig. 2b), with some being up-regulated and other down-regulated.

The color of anthocyanins is affected by the pH of the central vacuole. In \textit{petunia hybrida}, the vacuole of petal epidermal cells is hyper-acidified as compared with leaf cells. PH1 and PH5 encode P-ATPase proton transporter that are essential and sufficient for vacuolar hyper-acidification (Verweij \textit{et al.}, 2008; Faraco \textit{et al.}, 2014). Consistent with prior results obtained by real-time RT-PCR (Verweij \textit{et al.}, 2016), PH5 was down-regulated in an1, ph4 and ph3 (Fig. 2a). Since PH1 was not annotated in the \textit{P.axillaris} v1.6.2 CDS annotation v4 database, the expression level of PH1 in these three mutants was missing in this analysis.

**Gene families of target genes of the MBW complexes in \textit{Arabidopsis} and petunia**

\textit{Arabidopsis} expresses in epidermal cells of leaves, roots and stems epidermal cells a complex consisting of the WD40 protein encoded by \textit{TTG1}, a bHLH protein encoded by \textit{GL3} or \textit{EGL3} and a MYB protein encoded by \textit{GL1}, \textit{MYB23} and/or the \textit{PH4} homolog \textit{AtMYB5} (Oppenheimer \textit{et al.}, 1991; Kirik \textit{et al.}, 2001, 2005; Li \textit{et al.}, 2009). These \textit{Arabidopsis} MBW complex(es) are very similar to the petunia PH4-AN1-AN11 complex, and in case it contains \textit{AtMYB5} even homologous. These partially redundant complexes regulate in \textit{Arabidopsis} leaves the differentiation of epidermal cells in unicellular structures: trichomes (Jakoby \textit{et al.}, 2008). In petunia, however, trichomes are multicellular and develop independently from MBW complexes.

To assess to what extent the appearance of new biological roles for MBW complexes during evolution, such as trichome development in Brassicaceae, relied on the acquisition of new target genes, we compared the suite of target genes regulated by GL3 in \textit{Arabidopsis} leaves (Gao \textit{et al.}, 2017), with those regulated by AN1, PH4 and PH3 in petunia petals based on gene family cluster analysis by using PlantTribes (Wall \textit{et al.}, 2008). This analysis grouped target genes of GL3, AN1, PH4 and PH3 into 547, 597, 559 and 489 gene families, respectively (Fig. 3a). Among these genes families, GL3 shared 66, 74 and 76 controlled gene families with AN1, PH4 and PH3, respectively, with less than 15.5% overlap in
any mutant (Fig. 3a,b), and surprisingly, some of these gene families were oppositely affected in Arabidopsis and petunia. In addition, 22 gene families are distributed across these four mutants, with fewer than 4.5% in any mutant (Fig. 3 and Table S1). These results indicate that homologs of MBW complexes control different set of target genes in Arabidopsis leaves and Petunia petals.

Fig. 3 Target genes of the MBW complex in Arabidopsis and petunia. (a) Venn diagram based on the gene family cluster analysis. The numbers of gene families in different mutants are shown in parentheses. (b) Heatmap showing the log2 fold-changes of the expression for the target genes of GL3, AN1, PH4 and PH3. The numbers of genes in the different mutants are shown in parentheses. Blue indicates down-regulated genes. Red indicates up-regulated genes. Gene families found in all these four mutants are highlighted in red.

Chapter 5
Target genes of WMBW complex

Target genes of the MBW complexes regulating hair determination in Arabidopsis

Based on the results of ChIP analysis for GL3-YFP and/or ChIP-chip for GL3 (Morohashi & Grotewold, 2009), 20 genes were identified as direct targets of GL3. Among the genes that were misregulated in petunia mutant petals, we found several genes, all encoding transcription regulators, which appeared related to these direct GL3 targets.

One of these (Peaxi162Scf00015g0043), which was significantly down-regulated in anl and ph4 (Fig. 4), encodes a homeodomain transcription factor with high similarity to GL2 from Arabidopsis, which is required to specify trichome fate in aerial tissues and a non-hair (atrichoblast) cell fate in the root epidermis (Rerie et al., 1994; Masucci et al., 1996; Di Cristina et al., 1996), and formation of mucilage in the seed coat epidermis (Western et al., 2001, 2004). Since mutations that eliminate AN1 or PH4 or transgenic lines constitutively expressing AN1 or PH4 have no obvious defects in trichome development, it appears that in petunia petals PhGL2 operates in different processes that remain to be identified.

Among the genes down-regulated in anl, we found three genes encoding MYB-like proteins that seem to act as repressors. Peaxi162Scf00472g00077 was significantly down-regulated in anl and ph4 (Fig. 4) and encodes a MYB protein (PhMYBL2) with high similarity to MYB-LIKE2 (MYBL2) from Arabidopsis (Fig. S1). Typical R2R3 MYB proteins, such as AN2 and PH4 act as transcription activators and contain a DNA-binding domain consisting of two MYB repeats (R2 and R3). However, in AtMYBL2 and PhMYBL2, the first MYB repeat (R2) is incomplete, suggesting that these proteins may have an inhibitory effect on transcription, for example by sequestering partner proteins into inactive complexes (Zhou et al., 2014). It has been reported that overexpression of MYBL2 can suppress Arabidopsis trichome formation (Sawa, 2002).

Fig. 4 Expression level of homologs of GL3 direct target genes in petunia. Heatmaps show the effects of mutations in AN1, PH4 and PH3 on the expression of the petunia homologs of Arabidopsis genes that are directly regulated by GL3. Blue indicates down-regulated genes. Red indicates up-regulated genes. Gray indicates P-value > 0.05 and/or fold change < 2.
The other two genes (Peaxi162Scf00521g00814 and Peaxi162Scf00786g00439) were down-regulated in *anl, ph3* and/or *ph4* petals (Fig. 4). They encode true R3MYB type proteins, as they lack the R2 repeat completely, and contain only a very short C-terminal domain, which in R2R3MYB proteins usually contains a transcription activation domain, but are phylogenetically unrelated to *AtMYBL2* or *PhMYBL2* (data not shown). Peaxi162Scf00521g00814 encodes the protein MYBx, which was previously identified in a yeast two-hybrid screen of petal cDNA library with AN1 bait (Kroon, 2004; Koes *et al.*, 2005). Due to differential splicing petals cells co-express a long and a short MYBx mRNA (Fig. S2). In yeast two-hybrid assays both the long and short MYBx isoforms interact with a similar efficiency as the R2R3 MYB proteins AN2 and PH4 with the conserved N-terminal domain of AN1 and the distantly related bHLH protein JAF13 (Fig. 5), which has a minor role in activating anthocyanin genes (Bombarely *et al.*, 2016). When constitutively expressed from the 35S promoter in stable petunia transformants, the short MYBx transcript inhibits anthocyanin synthesis and vacuolar acidification by down-regulating the expression of target genes of the MBW complex involved in both anthocyanin synthesis (e.g. *DFR*) and vacuolar acidification (e.g. *PH3*, Verweij *et al.*, 2016) (Fig. 6). Given that MYBx consists of little more than an R3 MYB repeat and an extremely short C-terminal domain, MYBx most likely acts by competing with AN2 and PH4 for binding to AN1, thereby sequestering AN1 and JAF13 into inactive MBW complexes.
Target genes of WMBW complex

In Arabidopsis seven R3 MYB proteins, including ENHANCER OF TRY AND CPC1 (ETC1) (Kirik et al., 2004a; Esch et al., 2004), CAPRICE (CPC) (Wada et al., 1997), CAPRICE-LIKE MYB3 (CPL3) (Simon et al., 2007; Tominaga et al., 2008), ETC2 (Kirik et al., 2004b), TRIPTYCHON (TRY) (Schnittger et al., 1999), TRICHOMELESS1 (TCL1) (Wang et al., 2007) and TCL2 (Gan et al., 2011), are involved in trichome and/or root hair development by negatively modulating the activity of the MBW complex in a similar way (Wang & Chen, 2014). Of these R3 MYBs, TRY, CPC, ETC1 and CPL3 were identified as direct target genes of GL3 (Morohashi & Grotewold, 2009).

To assess the evolutionary relationship between these petunia and Arabidopsis R3 MYB proteins in more detail, we identified similar proteins from a broader range of species, using CPL3 as a query, and performed phylogenetic analyses. This revealed that only one P. axillaris protein (Peaxi162Scf00786g00439.1) belongs to the CPL3 subgroup and contains all the conserved motif of R3MYB (Fig. 7). The expression of this gene is down-regulated in an1, ph4 and ph3 petal (Fig. 4). However, the expression level of four Arabidopsis R3MYBs, ETC1, CPC, TRY and CPL3, is not affected in gl3 based on the gl3 transcriptome data in two different studies (Morohashi & Grotewold, 2009; Gao et al., 2017). In addition, Brassicaceae contain many more R3 MYB proteins than other species (Fig. 7), suggesting that a series of duplications enlarged greatly the number of R3 MYB genes in this family. MYBx belongs to a subgroup different from the Arabidopsis R3 MYBs like CPC and TRY. MYBx homologs contain a conserved Fx18Wx18W motif that is typically found in the R3 repeat of R2R3 MYB proteins, instead of Mx18Wx18W, which is conserved in the proteins belonging to the CPL3 subgroups (Wang & Chen, 2014). CPL3-like MYBs contain a WxM motif, which has been shown to be required for cell-to-cell movement (Kurata et al., 2005), while the last M in this motif was replaced by [S/T] in MYBx homologs (Fig. 7b). Lastly, CPL3-like MYBs contain a
[D/E]Lx3[R/K]x3Lx6Lx3R signature motif, which has been shown to be required for the interaction between R3 MYBs and bHLH transcription factors, whereas the first amino acid in this motif is not conserved in MYBx subgroup (Zimmermann et al., 2004).

Fig. 7 Phylogenetic analysis of MYBs and CPL3 homologs. (a) Phylogenetic tree of selected R3 MYB proteins with the highest similarity to CPL3 from different species. The gray, purple and blue shaded indicate the clade containing *Arabidopsis* R3 MYBs, *Petunia* MYBx and R2R3 MYB respectively. The tree construction was done using Maximum Likelihood. The protein sequences for the different species are those with the highest similarity to CPL3. Branch support was calculated on the basis of 1000 bootstraps and bootstrap values above 50% were placed on the tree. Species names: *Arabidopsis thaliana* (*A. thaliana*); *Capsella rubella* (*C. rubella*); *Citrus sinensis* (*C. sinensis*); *Eucalyptus grandis* (*E. grandis*); *Gossypium raimondii* (*G. raimondii*); *Mimulus guttatus* (*M. guttatus*); *Nicotiana benthamiana* (*N. benthamiana*); *Petunia axillaris* (*P. axillaris*); *Solanum lycopersicum* (*S. lycopersicum*). (b) Alignment of the conserved domain of MYB proteins. R3 MYB homologs are shaded in gray or purple. Red arrowheads indicate the conserved structure of MYB repeat: M/Fx18Wx18W. Red arrowheads indicate the [D/E]Lx3[R/K]x3Lx6Lx3R signature motif. The green arrowheads indicate the WxM signature motif.
To assess whether the *Arabidopsis* and petunia R3MYB proteins are functionally similar, we performed functional analysis. In yeast two hybrid assays the *Arabidopsis* CPC protein can interact with the N-terminus of AN1 and, more weakly with the corresponding domain of JAF13, similar to MYBx (Fig. 5). Ectopic expression of AtCPC in stable petunia transformants resulted in shift of the flower color from reddish violet to a blue-violet color (Fig 8), which is similar to the color of ph mutant (Verweij et al., 2008; Faraco et al., 2014). This indicates that CPC expression down-regulates genes required for the acidification of vacuoles in petal cells, such as *PH1* and *PH5*.

![Fig. 8 Ectopic expression of p35S:CPC in petunia. Flowers of wild type line transformed with the empty vector (pGreen3K) and five transgenic plants expressing p35S:CPC. The blue color of some of the transgenics (# 15, 16 and 20) indicates a reduction of vacuolar acidification.](image)

When constitutively expressed in *Arabidopsis* CPC inhibits the formation trichomes on aerial tissues and, as judged by the formation of supernumerary root hairs, the identity the non-hair cells (atrichoblasts) in the root epidermis, consistent with previous results (Wada et al., 1997). Constitutive expression of the short MYBx isoform had essentially the same effect, when expressed in wild type *Arabidopsis* (Fig. 9), and could restore the formation of (supernumerary) root hairs in a cpc mutant (Fig. 9). Interestingly, constitutive expression of the long MYBx isoform had little or no effect on the formation of trichomes (Fig. 9) and a weak effect on root hair formation, which was most evident in the nearly root hairless cpc mutant background.

In complementary experiments we constitutively expressed CPC in petunia. Expression of CPC in a wild type petunia line shifted the color of petals from reddish violet to blue-violet, thereby mimicking the phenotype of ph mutants. This indicates that CPC can inhibit the activation MBW-regulated genes involved in vacuolar acidification, but not or much less, MBW-regulated genes involved in anthocyanin synthesis (Fig. 8). This suggests that CPC can compete with PH4 for binding to AN1, but much less with AN2.

Taken together these data show that CPC and MYBx are functionally very similar, though not fully identical.
Fig. 9 The short MYBx transcript, but not the long transcript, can replace CPC in Arabidopsis. (a) Effect of the ectopic expression of the petunia MYBx (short and long transcripts) in wild type A. thaliana. (b) Effect of the ectopic expression of shorter and long MYBx transcripts in the cpc A. thaliana mutant. (c) Root hairs and trichomes in transgenics and relative controls.

Candidate genes involved in the formation of vacuolino

To isolate the candidate genes involved in the formation of vacuolinos, we focused on overlapping differentially expressed genes (DEGs) affected by an1, ph3 and ph4, since all these three mutations abolish vacuolinos (Faraco et al. 2017). Of all the DEGs, 148 genes were found in all these three mutants (Fig. 10a). These were classified into 5 groups based on gene expression pattern (Fig. 10b). Group I consists of 5 genes encoding two cytochrome P450 superfamily proteins, two UDP-glycosyltransferases and one GST that were up-regulated in an1, ph4 and ph3 petals. Group II contains
5 genes that were down-regulated in an1 and up-regulated in ph4 and ph3 petals. These 5 genes encoded two exordium-like proteins, which are regulators of cell expansion mediated by different pathways (Schroder et al., 2011), two xyloglucan endotransglucosylases and one cytochrome P450 protein. Group III consists of 2 genes encoding an amino acid permease and a cytochrome P450 protein, both down-regulated in an1 and ph3, but up-regulated in ph4. Group IV comprised 64 genes that were down-regulated in an1 and ph4, but up-regulated in ph3. Remarkably, among these 64 genes, we found eight genes encoding pectinesterase, six genes encoding pectate lyase and two genes encoding galacturan 1,4-alpha-galacturonidase, all of which were involved in the pentose and glucuronate interconversions pathway. Group V consists of 73 genes that were down-regulated in an1, ph4 and ph3. Several membrane traffic related proteins such as RAS-related proteins-Rab5 (Cui et al., 2014), ras-related proteins-Rab6 (Bednarek et al., 1994) and Monensin sensitivity 1 (Mon1) (Cui et al., 2014) were in this group. Interestingly, the HMA5I copper transporter (Li et al., 2017; Chapter 2) was also in this group. We studied this protein in further details and shown the results in chapter 2.

To obtain more information about these overlapping genes, KEGG analysis was performed. Among these 148 genes, 62 genes were annotated to the KEGG database and arranged into 13 district protein families (Fig. 10c). Notably, 3 of these genes, two Ras-related proteins and one RING-type E3 ubiquitin transferase, were associated with membrane trafficking (ko04131).

![Fig. 10](image-url) Overlapping genes among an1, ph4 and ph3 with P-value < 0.05 and fold-change > 2 as cutoff. (a) Venn diagram showing the number of DEGs (P-value < 0.05 and 2-fold changes as a cutoff) in an1, ph4 and ph3 petals compared with wild type petals. Numbers inside the diagrams present the number of genes affected. (b) Heatmap showing the log2-fold changes in expression of AN1-PH3-PH4 regulated genes in an1, ph3 and ph4 petals, as compared to wild type petals. Red color indicates up-regulated genes, blue color indicates down-regulated genes. The numbers of genes in these groups are shown in parentheses. (c) Protein families of 148 genes overlapping all three mutants, based on the KEGG orthology ID. I, Signaling and cellular processes; II, Genetic information processing; III, Metabolism.

To further establish a minimal set of candidate genes, a cutoff with fold-change >5 was applied. This analysis results in the identification of a set of 31 genes in all these three mutants (Fig. 11), which were arbitrarily defined as potential vacuolinos genes for the further study. As expected, we found back the five mRNA fragments that were previously isolated by microarray and cDNA-AFLP analysis (Verweij, 2007), PH5 (Peaxi162Scf00177g00620.1), CAC16.5 (Peaxi162Scf00940g00412.1), MAC9F1
(Peaxi162Scf00328g00118.1), CAC12.3 (Peaxi162Scf00429g00052.1), and CAC13.10 (Peaxi162Scf00008g04614.1) within these 31 genes list (Fig. 11).

Fig. 11 Overlap among differentially expressed genes in an1, ph4 and ph3 petals, compared to wild type with P-value < 0.05 and fold-change > 5 as cutoff. Heatmap showing the log-fold changes of overlapping genes among an1, ph3 and ph4 mutant. Red indicates up-regulated genes, blue indicates down-regulated genes.

Except for these genes, we also identified two target genes by micro-array and cDNA-AFLP, that we named NorfA and CAC4.4 (Verweij, 2007). Using various PCR procedures, including 3’ and 5’ RACE, we obtained NORFa and CAC4.4 cDNAs of of 1368 and 3484 nucleotides, respectively, and this length corresponds with the size of the hybridizing band observed in RNA-gel blots (not shown). The NORFa and CAC4.4 sequences do not show any similarity to known sequences. The NorfA transcript is not translatable into a long protein, but seems to contain two small Open Reading Frames (ORF) of 153 bp (short ORF1) and 174 bp (short ORF2). The CAC4.4 transcript contains two small ORFs of 165 bp and 171 bp (Verweij, 2007).
These target genes that were found in cDNA-AFLP and micro-array analysis (Verweij, 2007) were investigated firstly.

**Functional studies on selected candidate genes related vacuolino formation**

To investigate whether these genes are involved in vacuolino pathway, we generated transgenic lines in which their expression was down-regulated by RNA interference (RNAi). To examine whether that affected the formation of or trafficking via vacuolinos, we prepared petals protoplasts from and followed the localization of transiently expressed vacuolar proteins with a GFP tag (Fig. 12).

In NorfA knockdown lines (Fig. 12a), PH5-GFP moved to the vacuolinos within 24 hours after the transfection, as in wild type petal (Fig. 12b). This implies that NorfA is not required for the formation or the function of vacuolinos.

In pat1, which encode patatin-like phospholipase, knockdown petals, Aleu-GFP proteins moved to the CV within 24 hours, as in mesophyll cells, whereas in wild type petal Aleu-GFP localized at this time point in vacuolinos (Fig. 12c). In CAC12.3 knockdown petals, Aleu-GFP moved to small dot-like compartments. These findings imply that CAC12.3 and pat1 are both required the formation of vacuolinos, however, at different points of their genesis pathway.
**Fig. 12** Phenoype and genetic analysis of transgenic plants. (a) Real-time RT-PCR analysis of \textit{CAC12.3}, \textit{pat1} and \textit{NorfA} mRNAs in petals from wild type and transgenic RNAi petunia lines. (b) Confocal micrographs of petal protoplasts from \(M_1 \times V_{30}\) and NorfA-IR lines expressing PH5-GFP, 24 hours after transformation. (c) Confocal micrographs of petal protoplasts from \(M_1 \times V_{30}\), \textit{cac12.3-IR} and \textit{pat1-IR} line transiently expressing Aleu-GFP 24 hours after transformation. Note that \textit{CAC12.3-IR} images show a pigmented epidermal cell and an unpigmented mesophyll cell. Scale bars equal 10 \(\mu\)m. GFP fluorescence is shown in green, autofluorescence of anthocyanins in blue.
**Discussion**

Here, we present the transcriptome of petals from wild-type, an1, ph4 and ph3 petunia lines. Our results provide novel insight into the role of MBW complexes in terminal differentiation processes in epidermal petal cells, and provides a valuable resource to identify genes involved in a variety of processes, such as, for example, vacuolar acidification, stabilization of anthocyanins, the formation vacuolinos and their differentiation from the central vacuole. Furthermore, MBW complexes provide a fine model for the evolution of regulatory gene networks, because the mutant phenotypes of MBW components include defects that are common for a broad range of species, such as loss of anthocyanin or tannins pigments, as well as defects that are specific for specific subsets of species. Together with the already existing research on MBW complexes, our findings now provide novel insight of how these complexes gained or lost the capacity to activate certain pathways during evolution.

To date, the WMBW complex in petunia was shown to regulate the biosynthesis of anthocyanins and proanthocyanidins, the hyperacidification of vacuoles as well as the formation of vacuolinos and their presumed fusion with the central vacuole (Verweij et al., 2008, 2016, Faraco et al., 2014, 2017). Here we have identified 148 differentially expressed genes (DEGs) in an1, ph4 and ph3 petal, compared to wild type petals. The number of DEGs is many times more than we expected and includes a large number of genes with no obvious role in the known WMBW-regulated pathways. This includes, amongst others CAC16.5 encoding a cysteine proteinase, CAC13.10 encoding a major facilitator superfamily protein, several genes encoding GSTs and several encoding membrane transporters, including PhHMA5I. PhHMA5I encodes a copper transporter that belongs to the HMA5I subgroup based on phylogenetic analysis (data not shown). To date, HMA5Is in rice and Silene localize on the tonoplast to pump Cu into the vacuole lumen (Li et al., 2017), and the function of PhHMA5I in petunia is still unknown. Together, all these findings suggesting that WMBW controls more processes than thought on the basis of the mutant phenotypes.

Although it was shown that the MYB proteins such as COLORLESS1 (C1) from maize and WER from Arabidopsis contain transcription activation domains and can in vitro bind DNA, they appear in vivo fully dependent of their HLH (and presumably their WDR) partners for reasons that remain unclear (Feller et al., 2011). Together with the observation that inactivation of the WDR and bHLH proteins affects a broad range of different processes and that inactivation of the different MYB partners affects distinct subsets of these processes, suggested that the MYB partner is a major factor determining which downstream genes will be activated (Koes et al., 2005; Ramsay & Glover, 2005), possibly together with an interacting WRKY protein, known as TTG2 in Arabidopsis or PH3 in petunia (Pesch et al., 2014; Verweij et al., 2016).

To date, only bHLHs from subgroup IIIf have been shown to interact with WDR protein and to participate in MBW complexes (Heim et al., 2003; Feller et al., 2011). Arabidopsis possesses four IIIf type bHLHs with partially redundant functions, GL3, EGL3, TT8 and MYC1 (Heim et al., 2003; Feller
et al., 2011), whereas *P. inflata*, *P. axillaris* and *P. hybrida* varieties have only two, AN1 and JAF13 (Bombarely et al., 2016). It is striking, that the expression on “only” 289 genes is affected by both *anl* and *ph4*, whereas *ph4* affects 692 genes, >2-fold as many (Fig. 10). This suggests that, contrary to our expectations based on the *anl* and *ph4* phenotypes, *PH4* regulates some 60% of its target genes independently from AN1. Given that *PH4* can interact with AN1 and JAF13 in yeast two-hybrid assays, it is possible that *PH4* regulates these AN1-independent targets in conjunction with JAF13. This is currently difficult to test as confirmed *jaf13* null alleles are not available at present (Bombarely et al., 2016).

The interaction of PH3 with the WD repeat protein AN11, and the similarity of the *ph3* and *ph4* petal phenotype, suggested that PH3 acts primarily within a WMBW complex to enhance its specificity for genes involved in vacuolar acidification and membrane trafficking via vacuolinos (Verweij et al., 2016; Faraco et al., 2017). This is at odds with the finding that PH3 regulates the expression of some 600 genes, of which only 203 genes (~33.8%) are also differentially expressed in *anl* petals. Of these 203 AN1/PH3 regulated genes, only 148 genes are also differential in *ph4* petals (Fig. 10). These findings suggest that, contrary to the expectations based on the mutants' phenotypes, the MYB protein *PH4* and the WRKY protein PH3 may operate in a combinatorial manner with transcription factors unrelated to WMBs and or that the number of WMB complexes that exist in vivo is much higher than thought.

Since PH3 participates in the MBW complex by interacting with AN11, but not with AN1 and PH4 (Verweij et al., 2016), it may participate in a variety of different MBW complexes with different bHLH (AN1 and/or JAF13) and MYB components. In yeast two hybrid assays the N-terminal domains of AN1 and JAF13 can interact with PH4 and three other MYB proteins, AN2, AN4, MYBx and MYBb1/DEEP PURPLE (DPL) that were identified as activators of anthocyanin synthesis (Quattrocchio et al., 1999, 2006; Kroon, 2004; Albert et al., 2011). It has been shown that MYBb1 is involved in the activation of structural genes of the anthocyanin pathway (Albert et al., 2011). Whether PH3 and AN11 are in the very same WMBW complexes as PH4, AN2, AN4 and/or MYBb1/DPL, still needs to be confirmed.

In the models that are commonly presented WMBW complex consist of one WRKY, one MYB, one bHLH and one WDR factor, which is probably too simple. Given that (i) the bHLH proteins AN1 and JAF13 (Kroon, 2004) and their maize homologs can in yeast form homo- and heterodimers through their C-terminal domains (Kong et al., 2012), (ii) that such a dimer may bind via the N-termini two MYB proteins and (iii) that MYB proteins are thought to bind to DNA as a dimer (Grotewold et al., 2000; Feller et al., 2011), MBW complexes may in vivo contain multiple bHLH and MYB proteins. Bimolecular Fluorescence Complementation experiments (BiFC) provided evidence that MBW complexes containing multiple (distinct) MYB proteins can indeed form in vivo (Liu et al., 2014). Hence the number of MBW complexes that exist in vivo may be considerably larger than generally assumed, and therefore also the possibilities for combinatorial control of downstream genes.
In *Arabidopsis*, WMBW complexes were shown to regulate the biosynthesis of anthocyanins, proanthocyanidins and other processes such as mucilage production in seeds and the trichome development on aerial tissues and non-hair cells in the root epidermis (Ramsay & Glover, 2005), whereas the homologs WMBW complexes in petunia are known to control the biosynthesis of anthocyanins, proanthocyanidins, vacuole hyper-acidification and vacuolinos (Verweij *et al.*, 2008, 2016, Faraco *et al.*, 2014, 2017). These findings raise the obvious question of how WMBW complexes gained or lost the capacity to activate certain pathways during evolution.

Analysis of gl3 mutants indicates that GL3 regulates a subordinate layer of transcription regulators. This includes GL2, encoding is the core activator involved in the trichome formation in *Arabidopsis* (Rerie *et al.*, 1994) and several MYB genes that partially (MYBL2) or completely lack the R2 repeat (TRY CPC and paralogs), which act as inhibitors of the MBW complex. A very similar network seems at work in petunia petals, as expression of the petunia homologs of GL2, MYBL2 and at least two R3 MYB genes, including MYBx is significantly reduced in an1 and/or ph4 petals (Fig. 4). It is very likely that GL2 and these inhibitory MYB genes regulate a very different set of target genes than their *Arabidopsis* homologs, since trichome phenotypes were never been found in any of the loss or gain (ectopic expression) mutants in petunia, not even in transgenic plants expressing MYBx, which is expected to inhibit both AN1 and JAF13. However, when expressed in *Arabidopsis* MYBx can inhibit trichome formation and rescues the root hair defects in cpc mutants (Fig. 9). This suggests that differences between *Arabidopsis* and petunia MBW complexes with regard to their function and the subordinate suite of direct and indirect target genes (Fig. 3) are to a large extent due to differences in the promoters of the subordinate genes, rather than in the regulatory proteins. This idea is consistent with the finding that the WD repeat protein PALE ALEURONE COLOR (PAC), which in maize controls only anthocyanin synthesis but not hair development, can rescue virtually all defects in *ttg1* mutants, including trichome formation (Carey *et al.*, 2004), and genetic data showing differences in the MBW regulation of early anthocyanin genes such as CHS, result form alterations in the CHS promoters, rather than the MBW proteins (Quattrocchio *et al.*, 1993, 1998).

We recently demonstrated that a WMBW complex containing the AN1, PH4 and PH3 is required in petunia epidermal petal cells for the formation of small vacuoles, vacuolinos, and the sorting of vacuolar proteins to the CV transits via vacuolinos (Faraco *et al.*, 2017; this thesis Chapter 5). Vacuolinos are absent in an1, ph3 and ph4 mutants and vacuolar proteins, like Aleu-GFP and PH5-GFP, reach in these mutants the CV via a distinct pathway without passing through vacuolinos (Faraco *et al.*, 2017; this thesis Chapter 5), suggesting that AN1, PH3 and PH4 activate genes involved in the formation of vacuolinos. AN1, PH3 and PH4 control in addition genes, such as *PH1*, that act in late steps of the vacuolino pathway, such as the fusion of vacuolinos to the CV. PH1 encodes a P3b-ATPase involved in the acidification of the CV that is needed in addition for the transport of proteins from vacuolinos to the CV, presumably by mediating their fusion through interactions with vacuolar SNARE.
proteins. Hence, we expect among the genes co-regulated by AN1, PH3 and PH4 additional genes involved in vesicle and protein trafficking along the vacuolino pathway.

Among the genes down-regulated in an1, ph3 and ph4 petals are homologs of RAB5 and RAB6. RAB proteins are membrane bound GTPases belonging to the RAS superfamily, which are found in all eukaryotes, where they function as regulators in distinct membrane trafficking pathways (Uemura & Ueda, 2014). Hence these AN1-PH3-PH4 regulated RABs are likely to control key-steps in the vacuolino pathway, which is supported by recent functional analyses showing that upon mutation of this RAB5 gene vacuolar proteins remain stuck in small structures and no longer reach the vacuolinos or the CV (S. Li et al, manuscript in preparation). Here, we show that the AN1-PH3-PH4 regulated genes PAT1 and CAC12.3 are needed for the trafficking or proteins to (and presumably the formation of) vacuolinos. PAT1 encodes a patatin-like phospholipase A2-like protein and CAC12.3 encodes a protein of unknown function, none of which were previously implicated (or suspected) to be involved in membrane trafficking or organelle formation. In wild-type petals vacuolar proteins are found 24 hrs after the introduction of encoding transgenes on vacuolinos and need 48 hours to reach the CV, whereas in an1, ph3 and ph4 petals the same proteins reach the CV within 24 hrs without passing through vacuolinos, following an apparent default (“canonical”) pathway that is highly similar to the well-studied pathway to the vacuoles in leaf cells (Faraco et al., 2017; this thesis Chapter 5). The finding that the trafficking in pat1 knock down mutants essentially phenocopy the trafficking in an1, ph3 and ph4 mutants suggest that PAT1 operates at, or close to the point where the vacuolino and canonical pathway bifurcate, whereas CAC12.3 may operate further downstream in the vacuolinos pathway, explaining why vacuolar proteins get in a cac12.3 knock-down mutant stuck in small compartments (“punctae”) and do not reach the vacuolinos or CV, while the trafficking along the canonical pathway that operates in mesophyll cells is not affected. This idea might be further tested by analysis of PAT1 CAC12.3 double mutants and identification of the identity of the puncture.

It is likely that further analysis of the remaining AN1-PH3-PH4 regulated genes will uncover roles of several additional proteins in membrane trafficking or membrane recognition, whose function could not be easily discovered in other way. This underlines the value of the vacuolino pathway as a model system for studies on membrane trafficking and recognition in a higher eukaryote, because of its amenability to genetic experiments.
Materials and methods

Plant material and growth

All petunia lines (Table 1) were grown side by side under identical greenhouse conditions as previously described (Verweij et al., 2016). Petal limbs from floral bud of developmental stage 4 to 7 were harvested and immediately frozen in liquid nitrogen.

For wild type petunia stable transformation, a hybrid between line M1 and V30, was used.

Table 1: petunia lines used

<table>
<thead>
<tr>
<th>Line</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RevW138</td>
<td>AN1</td>
</tr>
<tr>
<td>W225</td>
<td>an1n225</td>
</tr>
<tr>
<td>R162</td>
<td>PH4v2009-v2011</td>
</tr>
<tr>
<td>R162</td>
<td>ph4v2009-v2011</td>
</tr>
<tr>
<td>R167</td>
<td>PH3v49</td>
</tr>
<tr>
<td>R167</td>
<td>ph3v49</td>
</tr>
</tbody>
</table>

Arabidopsis thaliana WS ecotype was used as wild type. The cpc-2 mutant used in this study has been described previously (Wada et al., 1997). Seeds were surface-sterilized, sown on the surface of 1.5% agar plates, and grown for the observation of seedling phenotypes. Seeded plants were kept at 4°C for 2 days and then incubated at 22°C under constant white light. For each mutant and transgenic line, at least ten individual 5-day-old seedlings were assayed for root epidermis changes, and at least five 2-week-old third leaves were assayed for trichomes.

RNAseq analysis

Total RNA was isolated from equal numbers of stage 4 to 7 petal limbs for each genotype by using Trizol as previously described (de Vetten et al., 1997). Two samples were prepared for wild type, an1, ph3 and ph4 mutant petals. 500ng of total RNA (260/280 ratio +/- 2.0 and 260/230 around 1.8-2.2) for each sample were sent for sequencing (Lausanne Genomic Technologies Facility, Center for Integrative Genomics, University of Lausanne).

The results were analyzed with CLCbio.

Heatmap

Heatmaps were generated using the Bioconductor package pheatmap (https://cran.r-project.org/web/packages/pheatmap/index.html) based on log2 fold-change.

Phylogenetic tree

Peptide sequences and cDNA sequences from S. lycopersicon (ITAG3.0), N. benthamiana (V1.0.1), C. annuum_Zunla (V2.0) and C. annuum_cvCM334 (V1.55) were downloaded from the Sol Genomic Network web site (https://solgenomics.net/). The Arabidopsis annotation sequences (version TAIR10)
were downloaded from TAIR (https://www.arabidopsis.org/). Peptide sequences and cDNA sequences from *A. coerulea* (V3.1), *K. laxiflora* (V1.1), *M. guttatus* (V2.0), *V. vinifera* (version 12X), *C. papya* (V0.4), *G. raimondii* (V2.1), *T. cacao* (V1.1), *C. sinensis* (V1.1) and *E. grandis* (V2.0) were downloaded from phytozome12 (https://phytozome.jgi.doe.gov/pz/portal.html#). Homologs of R3 MYBs from different species were identified by tblastn tool from blast+ (version 2.6.0) by using the amino acid sequence of CPL3 as query.

Homologs of other GL3 direct target proteins from different species were also identified by the tblastn tool. Sequence alignments for the phylogenetic tree were generated with PhyML 3.0, and phylogenetic trees were constructed using Maximum Likelihood method based on the Le_Gascuel_2008 model (Le & Gascuel, 2008) implemented in MEGA7 with 1000 bootstrap replicates and bootstrap values above 50% were placed on the tree. A matrix of pairwise distances was estimated using a JTT model, and the tree with the highest log likelihood was chosen with Neighbor-Join and BioNJ algorithms.

**Gene family clusters.**
The gene family clusters analysis was performed by PlantTribes as previously described (Wall *et al.*, 2008).

**KEGG enrichment**
The KEGG pathway annotation was performed using the blastKOALA tool (http://www.kegg.jp/blastkoala/) against the KEGG GENES database by using the *P. axillaris* peptide sequences (version 1.6.2) as query.

**Yeast two-hybrid assay**
Constructs for yeast two-hybrid where prepared as in previous work and the yeast two-hybrid analysis was performed as described (Quattrocchio *et al.*, 2006).

**Constructs**
The BamHI to EcoRI fragment of pGreen-MYBx long or MYBx short was blunt ended and ligated into the SmaI site of the pCHF3 binary vector (Jarvis *et al.*, 1998) to create 35S:MYBx long and 35S:MYBx short constructs. The 35S:CPC used in this study has been described previously (Wada *et al.*, 1997).

The full-size coding sequences of *Pat1* and *cac12.3* and were amplified from cDNA of the *P. hybrida* line R27 with primers containing attB1 and attB2 site and used to create a Gateway Entry clone by BP reaction with pDONR P1-P2 (Gateway system; Life Technologies, Invitrogen, Carlsbad, CA, USA). Entry clones were recombined with pK7GW1WG2(I) (Plant System Biology, Gent University, Belgium) for post transcriptional gene silencing.

**P. hybrida transformation**
Petunia transformations were performed as described previously (Verweij *et al.*, 2016). RNA of transgenic petunia was extracted from flower of developmental stage 6-7. RT-PCR analysis was performed as described (Verweij *et al.*, 2016).

**Arabidopsis transformation**
Arabidopsis transformation was performed by a floral dip method (Clough & Bent, 1998), and transformants were selected on a half Murashige&Skoog (MS) medium agar plates containing 50-100 mg/l kanamycin. We isolated at least twelve T1 lines for each construct and selected at least six T2 and two T3 lines on the basis of their segregation ratios for kanamycin resistance. For each transgenic line, at least ten individual 5-day-old seedlings were assayed for root hair number, and at least five 2-week-old third leaves were assayed for trichome numbers.

The $35S:MYB_x$ long or $35S:MYB_x$ short constructs were transformed into wild type WS and cpc-1 mutant, respectively.

**pH measurements**

The pH of crude petal extracts was measured as described (Verweij et al., 2008).

**Light microscopy**

Arabidopsis root phenotypes were observed using an Olympus SZH binocular microscope. For the observation of trichomes, images were recorded with a VC4500 3D digital fine microscope (Omrón, Kyoto, Japan).

**Protoplast isolation and transformation**

Protoplast isolation and transformation with GFP-fusion proteins were performed as reported previously (Faraco et al., 2011). Confocal images were taken 24 h after the transformation with a Zeiss LSM510 using the following excitation and emission wavelengths: 488 nm and 505-550 nm for GFP; 568 nm and > 650 nm for anthocyanin autofluorescence.
References


Flower Color.
Molecular Analysis of the target anthocyanin genes.

Sets of Target Genes. Anthocyanin Pigmentation Are Functionally Conserved among Plant Species and Have Distinct helix transcription factors of the anthocyanin pathway.

MYB protein that activates vacuolar acidification through interactions with basic
Methylation.

A, Tanaka Y, Quattrocchio F, Current Opinion in Plant Bi.

26 Required for Anthocyanin Sequestration, trichome initiation by the GL3 and GL1 selectors.

GLABRA2

Arabidopsis MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. Plant Cell 21: 72–89.


Target genes of WMBW complex

Supporting information

Fig. S1 Homologs of GL3 direct target genes in various species. The tree construction was done using Maximum Likelihood. The protein sequences for the different species are those with the highest similarity to AT3G10113 (a), AT1G71030 (b) and AT3G50800 (c). Branch support was calculated on the basis of 1000 bootstraps and bootstrap values above 50% were placed on the tree. Species names: Species names: Arabidopsis thaliana (A. thaliana); Capsella rubella (C. rubella); Capsicum annuum (C. annuum); Citrus sinensis (C. sinensis); Gossypium raimondii (G. raimondii); Kalanchoe laxiflora (K. laxiflora); Mimulus guttatus (M. guttatus); Nicotiana benthamiana (N. benthamiana); Petunia axillaris (P. axillaris); Solanum lycopersicum (S. lycopersicum); Theobroma cacao (T. cacao).
Fig. S2 The MYBx gene and its transcripts. (a) structure of the MYBx gene with its two introns. The first intron can be completely spliced resulting in the MYBx short transcript, or only partially spliced (the last 24 bp are not spliced) resulting in the MYBx long transcript. (b) Amino acid sequence alignment of MYBx proteins with other R3 MYBs.

Table S1. Gene families found in gl3, an1, ph4 and ph3.

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>UDP-Glycosyltransferase superfamily protein</td>
</tr>
<tr>
<td>28</td>
<td>glutathione S-transferase</td>
</tr>
<tr>
<td>41</td>
<td>beta glucosidase 1</td>
</tr>
<tr>
<td>48</td>
<td>beta galactosidase 1</td>
</tr>
<tr>
<td>73</td>
<td>Plasma membrane ATPase 4</td>
</tr>
<tr>
<td>99</td>
<td>Glycosyl hydrolase family protein</td>
</tr>
<tr>
<td>128</td>
<td>UDP-Glycosyltransferase superfamily protein</td>
</tr>
<tr>
<td>138</td>
<td>NAC domain containing protein</td>
</tr>
<tr>
<td>165</td>
<td>Bifunctional inhibitor</td>
</tr>
<tr>
<td>261</td>
<td>Xyloglucan endotransglucosylase/hydrolase family protein</td>
</tr>
<tr>
<td>275</td>
<td>EXORDIUM like 2</td>
</tr>
<tr>
<td>362</td>
<td>PLANT CADMIUM RESISTANCE 2</td>
</tr>
<tr>
<td>585</td>
<td>glutamate decarboxylase</td>
</tr>
<tr>
<td>741</td>
<td>purple acid phosphatase 3</td>
</tr>
<tr>
<td>948</td>
<td>Calcium-dependent lipid-binding (CaLB domain) family protein</td>
</tr>
<tr>
<td>985</td>
<td>2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase</td>
</tr>
<tr>
<td>1108</td>
<td>nudix hydrolase homolog 17</td>
</tr>
<tr>
<td>1400</td>
<td>aluminum-activated malate transporter 9</td>
</tr>
<tr>
<td>2740</td>
<td>Nucleotide-diphospho-sugar transferases superfamily protein</td>
</tr>
<tr>
<td>3574</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>4066</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>4240</td>
<td>Leucine-rich repeat</td>
</tr>
</tbody>
</table>