

The conserved transcription factors, MYB115 and MYB118, control expression of the newly evolved benzoyloxy glucosinolate pathway in *Arabidopsis thaliana*

OPEN ACCESS

Edited by:

Ute Roessner,

The University of Melbourne, Australia

Reviewed by:

Rita Maria Zrenner,

Leibniz-Institute of Vegetable and

Ornamental Crops (IGZ), Germany

Tamara Givolashvili,

University of Cologne, Germany

*Correspondence:

Yongming Zhou,

National Key Laboratory of Crop

Genetic Improvement, College of Plant

Science and Technology, Huazhong

Agricultural University, No.1 Shizishan

Street, Wuhan 430070, China

ymzhou@mail.hzau.edu.cn;

Daniel J. Kliebenstein,

Department of Plant Sciences,

University of California, Davis, One

Shield Avenue, Davis, CA 95616, USA

kliebenstein@ucdavis.edu

†These authors have contributed
equally to this work.

Specialty section:

This article was submitted to

Plant Metabolism and Chemodiversity,

a section of the journal

Frontiers in Plant Science

Received: 13 March 2015

Accepted: 30 April 2015

Published: 13 May 2015

Citation:

Zhang Y, Li B, Huai D, Zhou Y and Kliebenstein DJ (2015) The conserved transcription factors, MYB115 and MYB118, control expression of the newly evolved benzoyloxy glucosinolate pathway in *Arabidopsis thaliana*. *Front. Plant Sci.* 6:343. doi: 10.3389/fpls.2015.00343

Yuanyuan Zhang^{1,2†}, Baohua Li^{2†}, Dongxin Huai¹, Yongming Zhou^{1*} and Daniel J. Kliebenstein^{2,3*}

¹ National Key Laboratory of Crop Genetic Improvement, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, ² Department of Plant Sciences, University of California, Davis, Davis, CA, USA, ³ DynaMo Center of Excellence, Copenhagen Plant Science Centre, University of Copenhagen, Copenhagen, Denmark

The evolution of plant metabolic diversity is largely driven by gene duplication and ensuing sub-functionalization and/or neo-functionalization to generate new enzymatic activities. However, it is not clear whether the transcription factors (TFs) regulating these new enzyme encoding genes were required to co-evolve with these genes in a similar fashion or if these new genes can be captured by existing conserved TFs to provide the appropriate expression pattern. In this study, we found two conserved TFs, MYB115, and MYB118, co-expressed with the key enzyme encoding genes in the newly evolved benzoyloxy glucosinolate (GLS) pathway. These TFs interacted with the promoters of the GLS biosynthetic genes and negatively influenced their expression. Similarly, the GLS profiles of these two TFs knockouts showed that they influenced the aliphatic GLS accumulation within seed, leaf and flower, while they mainly expressed in seeds. Further studies indicated that they are functionally redundant and epistatically interact to control the transcription of GLS genes. Complementation study confirmed their roles in regulating the aliphatic GLS biosynthesis. These results suggest that the newly evolved enzyme encoding genes for novel metabolites can be regulated by conserved TFs, which helps to improve our model for newly evolved genes regulation.

Keywords: neo-functionalization, sub-functionalization, glucosinolates, co-expression, R2R3-MYB

Introduction

Plants need to address and survive ever changing biotic and abiotic environmental stresses without the capacity to move to avoid these stresses. One mechanism by which plants cope with these stresses is through the production of defensive secondary metabolites. The diversity of stresses has led to an equal diversification of the secondary metabolites with a huge number of lineage and species specific compounds (Wink, 1988). A key mechanism hypothesized to generate this diversity in plant metabolism is gene duplications. Most eukaryotic genomes have undergone whole-genome duplications, especially angiosperms (Wolfe and Shields, 1997; Vision et al., 2000; Blanc et al., 2003). Over time, these duplicated genes might be lost or silenced,

maintain the ancestral function, or undergo functional divergence either through neo- or sub-functionalization (Lynch and Conery, 2000). Recent work is beginning to highlight the role of gene duplication in amplifying the biosynthetic capacity of plants but this equally requires an increase in transcriptional control over these new genes (Kliebenstein and Osbourn, 2012; Kliebenstein, 2013).

Over the past two decades, glucosinolates (GLSs) present almost exclusive in the order Brassicales including model plant *Arabidopsis thaliana* have evolved as a model system for the study of secondary metabolite diversity in plants (Sønderby et al., 2010b). GLS are sulfur-rich, nitrogen-containing, amino acid-derived compounds (Agerbirk and Olsen, 2012) involved in aiding the plant to resist a myriad of attacking herbivores and pathogens and controlling the plants fitness within the field (Kliebenstein et al., 2002; Bednarek et al., 2009; Clay et al., 2009; Fan et al., 2011; Stotz et al., 2011). There have been more than 200 different GLSs identified in plants (Clarke, 2010), with *A. thaliana* containing at least 40 GLSs. The GLS can generally be classified into three groups according to their amino acid precursors: aliphatic—methionine, valine, isoleucine or leucine, benzolic—phenylalanine or tyrosine and indolic—tryptophan GLS (Fahey et al., 2001; Kliebenstein et al., 2001c). The biosynthesis of aliphatic GLS can be divided into three steps: side-chain elongation of precursor amino acid, core structure pathway, and side-chain modifications. The side-chain modifications of aliphatic GLS greatly contributes to structural and functional diversity of GLS and appear to be the newest evolved genes within the pathway, indicating that these steps are a good model for how newly evolved biosynthetic processes may be integrated into a regulatory system (Sønderby et al., 2010b).

The near complete identification of enzymes in *A. thaliana* GLS biosynthesis has provided support for the role of gene duplication and neo- and/or sub-functionalization across multiple loci to provide the basis of the diversity in secondary metabolites (Hansen et al., 2001, 2007; Kliebenstein et al., 2001a; Chen et al., 2003; Textor et al., 2007; Kliebenstein, 2008; Li et al., 2008). For example, side-chain length variation of methionine-derived aliphatic GLS is controlled by differential expression of three tandem duplicate genes, *MAM1*, *MAM2*, and *MAM3* (Kliebenstein et al., 2001a; Kroymann et al., 2001, 2003). *MAM1* is capable of doing two elongation cycles while *MAM2* controls a single cycle and *MAM3* can catalyze through 6 elongation cycles (Kroymann et al., 2001, 2003; Field et al., 2004; Textor et al., 2007).

Similarly, neo-functionalization of two tandem 2-oxoglutarate-dependent dioxygenases: *AOP2* and *AOP3*, catalyze differential conversion of methylsulfinylalkyl-GLSs (MS-GLSs) to either the alkenyl GLS (*AOP2*) or chain-length specific hydroxyalkyl GLS (*AOP3*) (Kliebenstein et al., 2001c). A role for both tandem and whole-genome duplication in diversifying the pathway comes from five flavin-monooxygenases (*FMO_S*) (*FMO_{GS-OX1-5}*) which convert methylthioalkyl-GLSs (MT-GLSs) to methylsulfinylalkyl-GLSs (MS-GLSs). These genes differentially convert MT-GLSs to MS-GLSs and they arose from a whole-genome duplication followed by ensuing tandem duplication at each locus and sub-functionalization (Hansen

et al., 2007; Li et al., 2008). Similar evidence is also emerging from analysis of the indolic GLS pathway (Hull et al., 2000; Mikkelsen et al., 2000; Bednarek et al., 2009; Clay et al., 2009; Pfalz et al., 2009, 2011).

An understudied aspect of the evolution of new biosynthetic pathways is how their accompanying regulatory machinery arises. Are they controlled by transcription factors (TFs) that are equally derived from gene duplication or are these new pathways controlled by conserved TFs (Kliebenstein and Osbourn, 2012; Kliebenstein, 2013)? Recent observations in the GLS regulatory machinery have begun to provide evidence for both options. Two main TF families have been linked to regulating GSL production, the myeloblastosis (MYB) and myelocytomatosis (MYC) TFs (Dubos et al., 2010; Schweizer et al., 2013; Frerigmann et al., 2014; Li et al., 2014). The MYBs involve a paralogous group of genes belonging to subgroup 12 of the R2R3 MYB family, which are proved to be functionally redundant and specific involved in the transcriptional regulation of aliphatic GLSs (MYB28 family: *MYB28*, *MYB29*, and *MYB76*) (Gigolashvili et al., 2007, 2008, 2009; Hirai et al., 2007; Sønderby et al., 2007, 2010a) and indolic GLSs (MYB34 family: *MYB34*, *MYB51*, and *MYB122*) (Malitsky et al., 2008; Gigolashvili et al., 2009; Frerigmann and Gigolashvili, 2014). These MYBs are unique to the GLS containing Brassicales and arose via whole-genome and tandem duplication with ensuing neo-functionalization to focus on one branch of the GLS pathway (Bekaert et al., 2012). In contrast to MYB TFs, the MYC family shows evidence wherein conserved jasmonate signaling TFs have captured the regulation of the newly evolved GLS pathway (Dombrecht et al., 2007; Fernández-Calvo et al., 2011; Schweizer et al., 2013; Frerigmann et al., 2014; Li et al., 2014). The MYCs are encoded by four genes that arose via gene duplication and there is evidence of sub-functionalization among them (Heim et al., 2003; Fernández-Calvo et al., 2011; Niu and Figueroa, 2011; Schweizer et al., 2013; Frerigmann et al., 2014). More recently a broad set of new conserved and evolutionarily limited TFs were added to the list of potential regulators for the GLS pathway (Li et al., 2014). Thus, there is evidence for TFs controlling the GLS pathway to co-evolve with the pathway (MYBs) or to have captured the pathway (MYCs). This suggests that more work is needed to figure out if one or the other model is more prevalent.

To generate more evidence assessing which of these two models may be more likely, we focused on the transcriptional regulation of a branch of the GLS pathway that recently evolved within the *A. thaliana* lineage (Kliebenstein and Osbourn, 2012; Kliebenstein, 2013). Benzoyloxy GLS (BZ-GLS) are found in *A. thaliana* but not in its close relatives, suggesting that this pathway arose within the *A. thaliana* lineage is very young providing an opportunity to investigate how very young biochemical pathways are regulated. The production of BZ-GLS within the Col-0 accessions of *A. thaliana* requires the involvement of side-chain modifying enzymes encoded by the *AOP3*, *BZO1*, and *SCPL17* genes (Figure 1A). This pathway is transcriptionally limited to the developing seed suggesting that it is precisely regulated at the transcriptional level (Kliebenstein et al., 2001c, 2007; Lee et al., 2012). To test if this new pathway is regulated by a novel lineage specific TF or captured by an existing

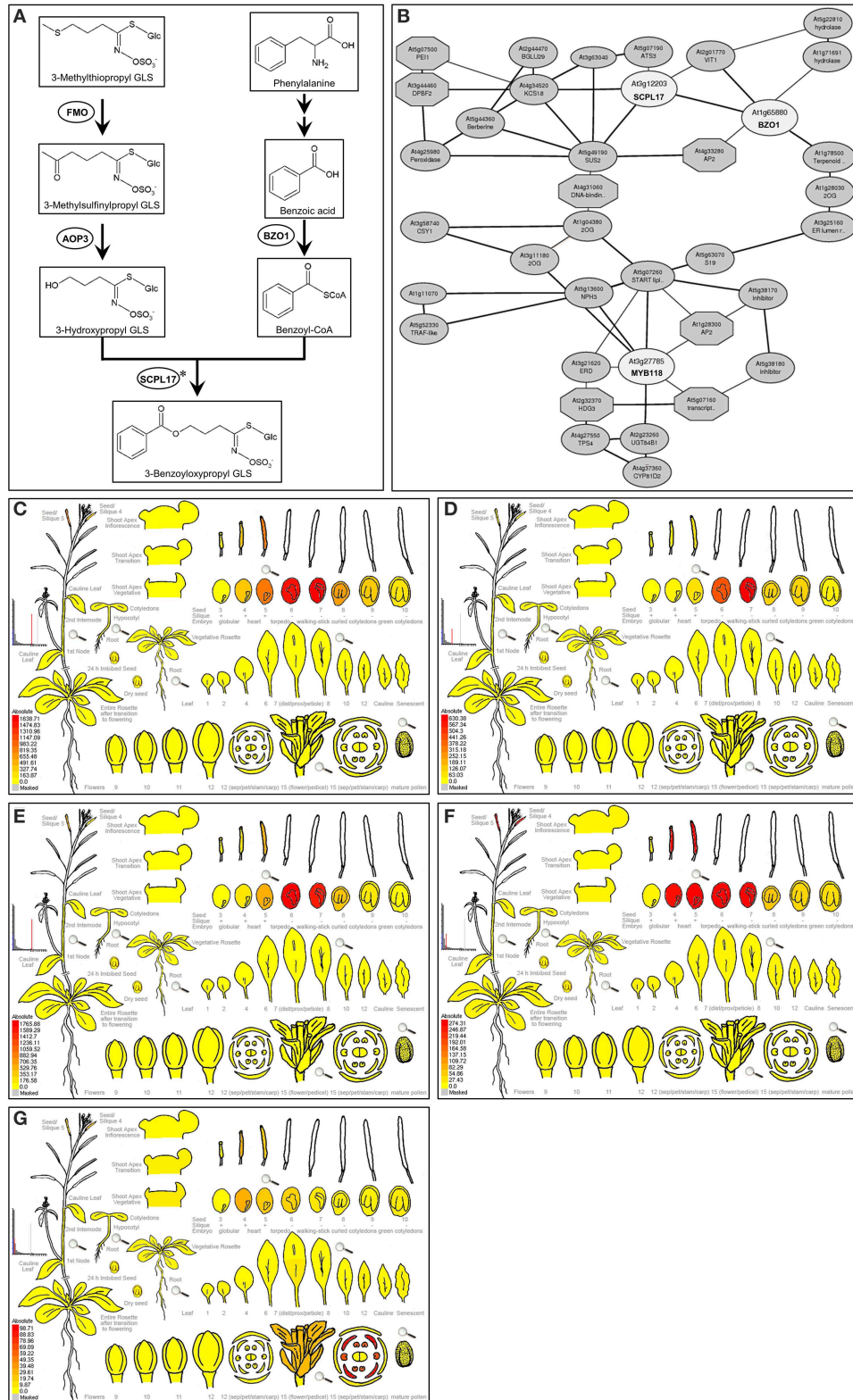


FIGURE 1 | Expression pattern of AOP3, BZO1, SCPL17, and MYB118. (A) Proposed biosynthetic pathway leading to 3-benzoyloxypropyl GLS in *A. thaliana*. The proposed biosynthesis of 4-benzoyloxybutyl GLS proceeds in a

similar manner, but with a 4-carbon aliphatic GLSs precursor. Metabolites are shown in boxes and enzymes are shown in circles. Arrows between
(Continued)

FIGURE 1 | Continued

compounds represent the number of putative enzymatic reactions. Asterisks partially characterized enzyme. **(B)** The co-expression network analysis of *BZO1*, *SCPL17*, and *MYB118* using the ATTED-II database (Obayashi et al., 2014). The TFs and genes are shown in circles. **(C–G)** e-FP display of transcript accumulation

patterns across a variety of Arabidopsis organs. Arabidopsis e-FP browser (Winter et al., 2007) presents the transcript accumulation pattern of *AOP3* **(C)**, *BZO1* **(D)**, *SCPL17* **(E)**, *MYB118* **(F)**, and *MYB115* **(G)** in the siliques. In all cases, red indicates higher levels of transcript accumulation and yellow indicates a lower level of transcript accumulation.

conserved pathway, we searched for candidate TFs regulating BZ-GLS-related genes, *AOP3*, *BZO1*, and *SCPL17* using a comprehensive strategy based on transcription co-expression analysis (Obayashi et al., 2014). This identified the homologous TFs *MYB115* and *MYB118* as likely candidates for controlling the transcription of this pathway (Dubos et al., 2010). These genes had previously been characterized as conserved TFs that regulate maturation-related genes and promote the vegetative-to-embryonic transition (Wang et al., 2009; Zhang et al., 2009; Barthole et al., 2014). They had not previously been associated with the regulation of GLS biosynthesis. Here, we functionally validate that the two R2R3-MYB TFs from *A. thaliana*, *MYB115* and *MYB118*, play important roles in transcriptional regulation of aliphatic GLS biosynthetic pathway for the production of the new BZ-GLS pathway. We further investigate the epistatic interactions between *MYB115/MYB118* and some key genes in GLS biosynthetic pathway, in order to describe the regulation model of *MYB115* and *MYB118* in aliphatic GLS biosynthesis. This shows that the newly evolved BZ-GLS pathway is regulated by conserved TFs that previously existed to provide the specific developmental context observed for BZ-GLSs.

Results

MYB115 and MYB118 Are Candidate Regulators of GLS Biosynthesis Using Co-expression Analysis and Yeast One-hybrid Assay

Using a comprehensive strategy of co-expression analysis, combined with the transcriptome co-expression profiles from ATTED-II database (Obayashi et al., 2014) and the expression pattern using Arabidopsis e-FP browser (Winter et al., 2007), a total of 300 TFs were shown to highly co-express with BZ-GLS-related genes, *AOP3*, *BZO1*, and *SCPL17*, which are the key genes involved in the side-chain modification pathway of GLS biosynthesis, especially BZ-GLS biosynthesis (Figure 1A). Focusing on R2R3-MYB TFs which are frequently key secondary metabolite regulatory TFs leads to the identification of three R2R3-MYB TFs as candidates, *MYB56*, *MYB58*, and *MYB118* (Figure 1B). From this, only *MYB118* exhibited the same seed-specific expression pattern as that of *AOP3*, *BZO1*, and *SCPL17* (Figures 1C–F). *MYB118* (AT3G27785) belongs to subgroup 25 of the R2R3-MYB TF family containing six members (Dubos et al., 2010). *MYB115* (At5g40360) is the most closely related to *MYB118* in the same subgroup, and overexpression of the two TFs showed similar phenotypes suggesting that they have similar functions (Wang et al., 2009; Zhang et al., 2009; Barthole et al., 2014). Thus, we hypothesized that *MYB118* and *MYB115* may both transcriptionally regulate the BZ-GLS pathway.

To validate the potential regulatory roles of *MYB115* and *MYB118* in the aliphatic GLS biosynthetic pathway, a yeast one-hybrid study (Y1H) was used to test their interactions with eight promoters of aliphatic GLS pathway genes which were mainly involved in the side-chain elongation and modification of GLS biosynthesis (Figure 2). Surprisingly, the result showed that *MYB115* and *MYB118* not only bound the promoters of BZ-GLS modification genes (*BZO1*, *SCPL17*, and *AOP3*) but also bound the promoters of three side-chain elongation genes (*BCAT4*, *MAM1*, and *MAM3*) and a core structure pathway gene (*CYP83A1*). Thus, the Y1H data show that *MYB118* and *MYB115* can bind a large number of promoters in the pathway and might have broader impact on the regulation of GLS biosynthesis than simply the BZ-GLS.

To validate if these Y1H predicted promoter interactions may alter the transcription of the downstream genes in planta, we analyzed the expression patterns of these genes by qRT-PCR in *MYB115* and *MYB118* gene-knockout plants, designated as *myb115* and *myb118*, respectively, and double *myb115myb118* mutants (Figures S1A,B). Total RNA was isolated from the seeds harvested 6 and 7 days after pollination from three independent mRNA extractions. The result showed that the expression levels of the genes involved in side-chain modification of GLS biosynthesis [*AOP3*, *BZO1*, *SCPL17*, and *FMO_{GS-OX2}* which co-expressed with *BZO1*, *SCPL17*, and *MYB118* in the siliques (Figure S2A), but it also have some expression in vegetative tissue (Figure S2B)] were significantly increased in *myb118* while in *myb115* they were similar to the wild-type Col-0 controls (WT) (Figures 2C,D). In addition, the double mutant *myb115myb118* lead to a further increase in expression of these genes above either single mutant genotype. In contrast to the side-chain modification genes, the chain elongation genes showed different expression patterns in single and double mutants. The expression levels of *BCAT4* and *CYP83A1* were significantly and equally increased in all *MYB115* and *MYB118* single and double knockouts in comparison with wild-type Col-0. In contrast to *BCAT4*, *BCAT3* had no altered expression in any genotype which agrees with its promoter showing no evidence of interaction with either *MYB115* or *MYB118*. For the two genes that centrally regulate side-chain elongation, *MAM1* and *MAM3*, only *myb115* showed a significant change in expression levels in comparison with wild-type Col-0. This increased expression was abolished in the double mutant. Taken together, *MYB118* and *MYB115* can regulate GLS biosynthesis in the seeds of Arabidopsis and function to genetically repress the expression of their target genes.

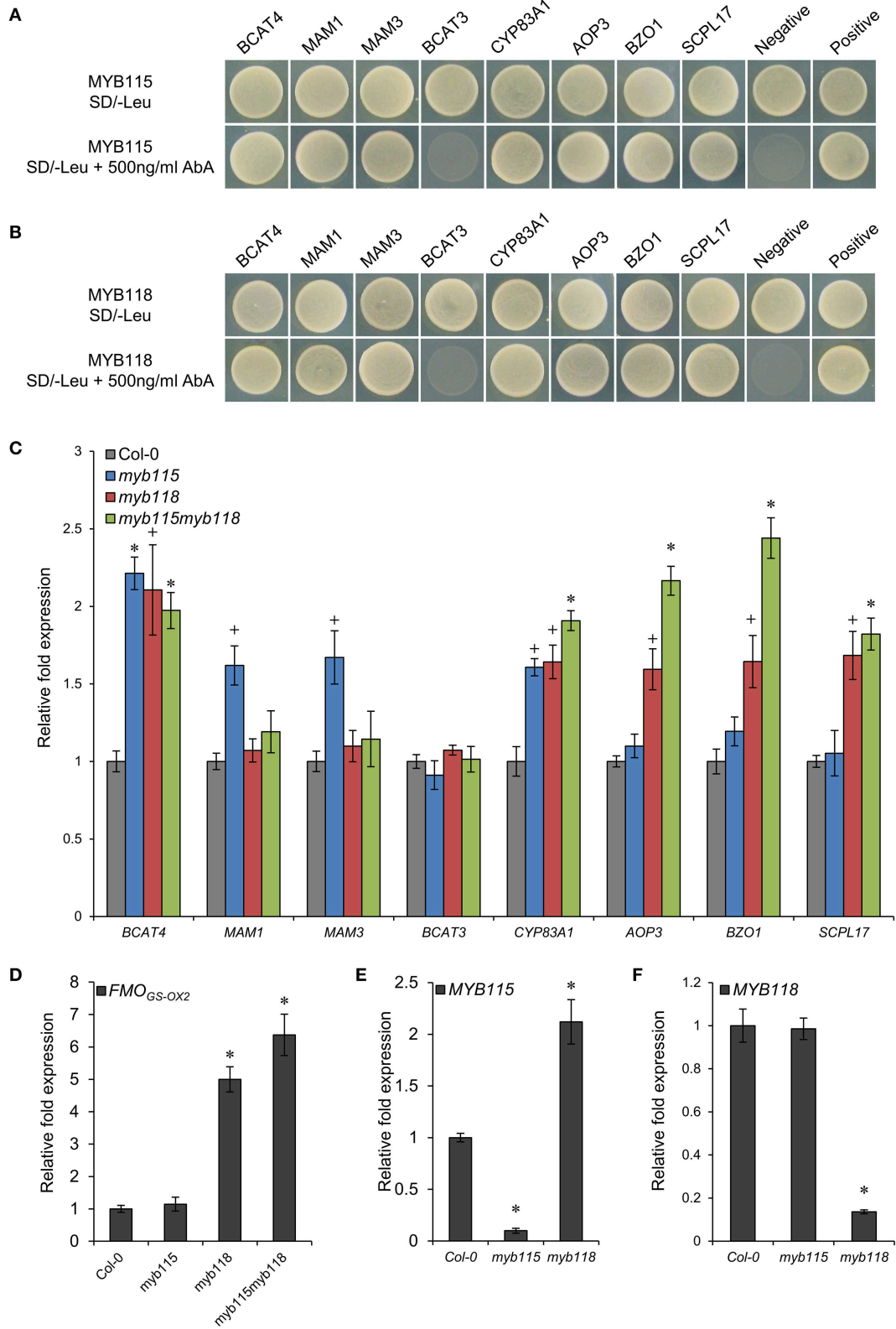


FIGURE 2 | Identification of interactions between MYB115/MYB118 and aliphatic GLS biosynthesis-related genes. (A,B) Yeast one-hybrid assays showing binding of MYB115 (A) and MYB118 (B) to the promoters of aliphatic GLS biosynthesis-related genes. The Y1H Gold yeast strains growth

in the absence (top) and presence of 500 ng ml⁻¹ Aureobasidin A (AbA) on the SD-Leu plates is shown. The mutant pABAI vector was used as the negative control and the p53 vector was used as the positive control. (C-F)

(Continued)

FIGURE 2 | Continued

qRT-PCR analysis of aliphatic GLS biosynthesis-related genes measured in the seeds of Col-0, *myb115*, *myb118*, and *myb115myb118* knockouts. The relative expression values of genes were measured on developing seeds

harvested 6 and 7 days after pollination. Ubiquitin-conjugating enzyme 21 (*UBC21*) gene expression level was used as a constitutive control. Values are the means and SE of three independent mRNA extractions. The significant differences is shown as ⁺($P < 0.05$) or * ($P < 0.01$) using *t*-tests.

Multiple GLS Phenotypes in *MYB115* and *MYB118* Single Mutant Seeds

To test if the predicted transcriptional regulatory function of *MYB115* and *MYB118* leads to altered GLS accumulation, we analyzed GLS accumulation in gene-knockout plants (*myb115* and *myb118*), both of which grew normally (for GLS abbreviations, see **Table 1** and Table S1). Because the over-expression lines show extremely stunted growth and sterility we were unable to measure GLS phenotypes in these lines (Wang et al., 2009; Zhang et al., 2009; Barthole et al., 2014). In *myb118* seeds (**Figure 3**; Table S2), the contents of MS-GLSs, OH-GLSs and BZ-GLSs were significantly increased in agreement with the induced expression of the associated genes. In contrast, the content of MT-GLSs was decreased, whereas the content of indolic GLS (I3M) did not change significantly in comparison with WT. Notably, the phenotypic changes on GLS content in *myb118* seeds are comparable, with opposite sign to the ones of *myb28* and *myb29* single mutants (Hirai et al., 2007; Sønderby et al., 2007), indicating that *MYB118* may have an equally important impact on the transcriptional network on controlling GLS production *in planta*.

In agreement with the absence of a significant transcriptional change, the phenotypic changes of GLS content in *myb115* mutants are much weaker than the ones of *myb118*. 3OHP, 4MSOB, 4MTB, and 3BZOP GLSs are statistically significantly but only slightly increased, while the levels of 4BZOB and 8MTO GLSs were similarly reduced in a significant but subtle manner (**Figure 3**; Table S2). These data indicated that *MYB115* might also control the aliphatic GLS biosynthetic genes in seed. Thus, the GLS changed in *myb115* or *myb118* in comparison with the wild type, such as MS-GLSs, OH-GLSs and BZ-GLSs, were related to the function of *AOP3* or *BZO1* in GLS biosynthesis. These findings support the result of co-expression analysis that *MYB115* and *MYB118* regulate the *AOP3/BZO1* related GLS phenotypes. Notably, both *MYB118* and *MYB115* negatively influence the regulation of most GLS accumulation in agreement with the observed transcriptional change. This is in contrast to the previous identified MYB transcriptional regulators including *MYB28* and *MYB29* which are all positive regulators for the pathway (Gigolashvili et al., 2007; Hirai et al., 2007; Sønderby et al., 2007).

MYB115 and *MYB118* Also Affect the Content of GLS in Leaf and Flower Organs

To test if *MYB115* and *MYB118* can influence GLS biosynthesis in additional tissues, we measured GLS accumulation in leaves and flowers of the single mutants. While both *MYB118* and *MYB115* are considered to be expressed specifically to reproductive tissues (Wang et al., 2009; Zhang et al., 2009; Barthole et al., 2014), they do have low residual expression in vegetative tissue (**Figures 1E,G**). Interestingly, both single

TABLE 1 | Glucosinolate (GLS) abbreviations in this study.

Abbreviation	GLS type	GLS name
3MTP	Aliphatic 3C	3-methylthiopropyl GLS
3MSOP	Aliphatic 3C	3-methylsulfanylpropyl GLS
3OHP	Aliphatic 3C	3-hydroxypropyl GLS
3BZOP	Aliphatic 3C	3-benzoyloxypropyl GLS
4MTB	Aliphatic 4C	4-methylthiobutyl GLS
4MSOB	Aliphatic 4C	4-methylsulfanylbutyl GLS
4OHB	Aliphatic 4C	4-hydroxybutyl GLS
4BZOB	Aliphatic 4C	4-benzoyloxybutyl GLS
2OH-Butenyl	Aliphatic 4C	2-hydroxy-but-3-enyl GLS
5MSOP	Aliphatic 5C	5-methylsulfanylpentyl GLS
6MTH	Aliphatic 6C	6-methylthiohexyl GLS
6MSOH	Aliphatic 6C	6-methylsulfanylhexyl GLS
7MTH	Aliphatic 7C	7-methylthioheptyl GLS
7MSOH	Aliphatic 7C	7-methylsulfanylheptyl GLS
8MTO	Aliphatic 8C	8-methylthiooctyl GLS
8MSOO	Aliphatic 8C	8-methylsulfanyloctyl GLS
4OH-I3M	Indolic	4-hydroxy-indol-3-ylmethyl GLS
I3M	Indolic	indol-3-ylmethyl GLS
4MO-I3M	Indolic	4-methoxy-indol-3-ylmethyl GLS
NMO-I3M	Indolic	N-methoxy-indol-3-ylmethyl GLS

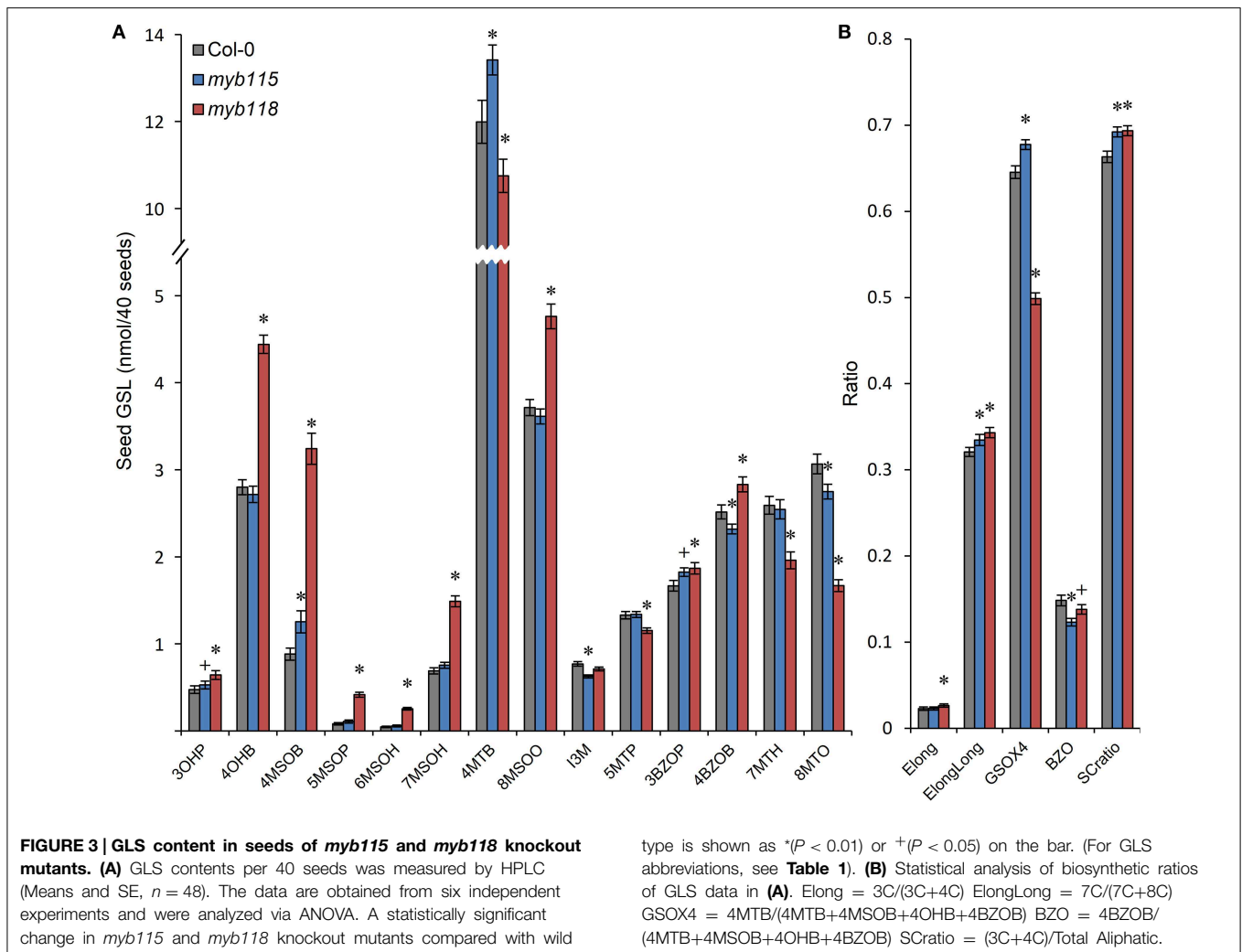
Abbreviations, type, and name of glucosinolates in this study.

mutants lead to increased aliphatic GLS accumulation in vegetative tissues with no corresponding change in indolic GLSs (**Figure 4**; Table S2). In contrast to the seeds, *myb115* showed similar GLS phenotypes to *myb118* in leaves. These data suggest that *MYB118* and *MYB115* can alter the accumulation of aliphatic GLS within the leaf even with low residual transcript accumulation.

Transcript data showed that *MYB115* shows a distinct expression pattern within the flower where it is localized to petals and stamens, but not in sepals and carpels (**Figure 1G**). To investigate the effect of this specific express pattern, we analyzed the GLSs of *myb115* and *myb118* in each flower organs (sepal, petal, stamen, and stigma). This showed that all flower organs of *myb115* and *myb118* mutants had increased levels of short chain MS-GLSs, 3MSOP, 4MSOB, and 5MSOP (**Figure 5**; Table S2). The only tissue specific change was an alteration of long chain MS-GLSs (7MSOH and 8MSOO) that was largely limited to the petal and stigmas of the single mutants. Thus, the *MYB118* and *MYB115* TFs influence GLS accumulation throughout the plant.

Complementation Study Confirmed the Phenotypes of *MYB115* and *MYB118* Single Mutants

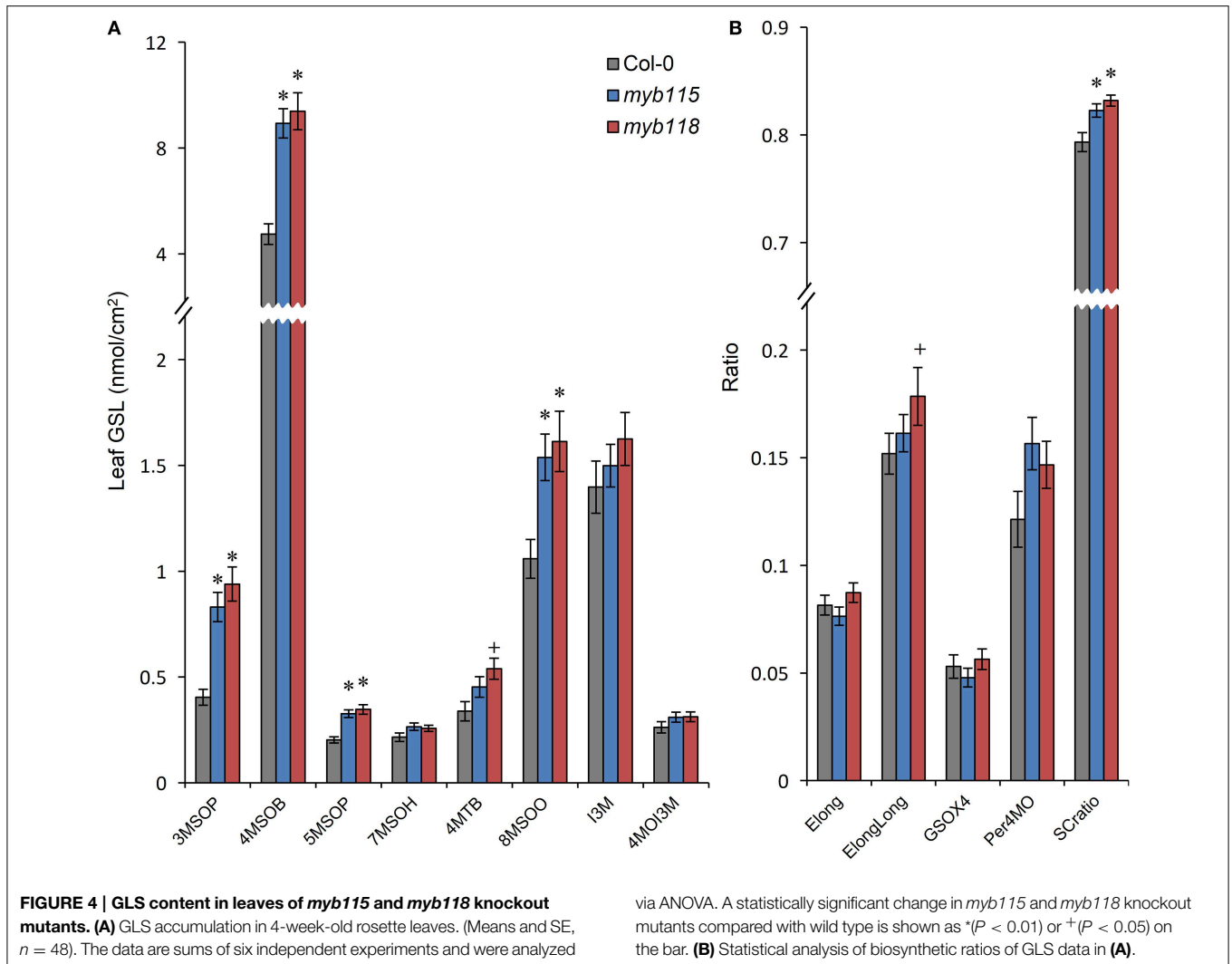
To confirm the metabolite phenotypes found in the *myb115* and *myb118* mutant, we transformed the respective T-DNA



lines mutants with the genomic region of the each MYBs respectively. We then tested the seed GLS content of three independent *myb115* and *myb118* complementation lines by nested ANOVA wherein we tested for both a difference from the WT control and between the complementation lines. No significant differences were found amongst the complementation lines and we thus focused on the genotypic differences. This showed that introducing the WT *MYB115* gene into the *myb115* mutant abolished the observed increase in short-chain GLSs and returned the genotype to a WT phenotype (Compare Figure 6 to Figure 3; Tables S2, S3). Similarly, introducing the WT *MYB118* gene into the *myb118* mutant complemented the altered short-chain GSL accumulation phenotype. Thus, the short-chain GLS phenotypes in the two *myb115* and *myb118* TDNA insertion lines were complemented by the DNA fragment harboring the intact *MYB115* and *MYB118* genes (Figure 6; Table S3). The long-chain GLS levels were not fully complemented but were closer to wild-type than the corresponding single mutant suggesting that these had been partially complemented. These results strongly support the roles of *MYB115* and *MYB118* in regulating the aliphatic GLS biosynthesis.

Double Knockouts Show an Interaction between *MYB115* and *MYB118* in Aliphatic GLS Biosynthesis

The qRT-PCR work had suggested that there may be an epistatic interaction between *MYB115* and *MYB118* in controlling aliphatic GLS accumulation (Figure 2C). To test if *MYB115* and *MYB118* may epistatically control GLS accumulation, the double mutant was grown under controlled conditions, together with wild-type control and *myb115* and *myb118* single mutants, and the content of GLS in seeds was measured. The GLS levels were then analyzed using a Two-Way ANOVA to parse out the effects of the single mutants and any interaction effects. There was a statistically significant effect of the *myb115myb118* double mutant on the short-chain GLS where the double mutant showed a higher accumulation of the OH-GLSs and MS-GLSs greater than expected from the single mutant effects (Figure 7; Table S4). Interestingly, this did not correspond to a significant double mutant increase in BZ-GLS as was expected based on the qRT-PCR (Figure 2C). This suggests that the benzoyl moiety may be limiting in these genotypes. In contrast to the GLS level in seeds, *myb115myb118* didn't show any evidence of epistatic



interaction when measuring leaf GLS accumulation (Figure S3; Table S4). In agreement with the observed epistasis between *MYB115* and *MYB118*, the expression level of *MYB115* was significantly increased in *myb118* knockouts in comparison with wild-type Col-0. In contrast, *MYB118* expression in the *myb115* genotype was similar to the wild-type Col-0 (Figures 2E,F). This result suggests that in the *myb118* knockout, the function of *MYB118* is partly replaced by elevated function of *MYB115* suggesting that *MYB118* regulates *MYB115*. Thus, *MYB115* and *MYB118* epistatically interact to control the transcription of GLS genes and GLS accumulation in Arabidopsis seeds (Figures 2, 7).

MYB115 and MYB118 Show Different Interactions with the Genes Involved in Side-chain Modification Pathway of GLS Biosynthesis

We next proceeded to test how the TF mutants interact with mutants in the AOP3 and BZO1 biosynthetic loci to test how the pathway may be diverted in case of a simultaneous regulatory and biosynthetic defect (Kliebenstein et al., 2001c, 2007). To accomplish this, we generated four double mutants, *myb115aop3*,

myb118aop3, *myb115bzo1*, and *myb118bzo1* double mutants. All the double mutants were grown together with wild-type control and all single mutants (Figure S1), and GLS content of seeds, leaves and flowers were measured and tested via Two-Way ANOVA. In seeds, *MYB115* and *MYB118* both interact with AOP3 but lead to different consequences. A *myb118aop3* double mutant leads to a large accumulation of 4MSOB GLS but no corresponding double mutant increases in 4MTB whereas the *myb115aop3* pushed the GLS to the 4MTB form (Figure 8; Table S5). Similar to the differential AOP3 interactions, the two MYBs had different effects on GLS repartitioning in the *bzo1* background (Figure 9; Table S6). *MYB115* had little to no effect in the *bzo1* background with all double mutant phenotypes appearing like the *bzo1* single mutant. In contrast, the *myb118bzo1* double mutant had a large increase in the accumulation of the short-chain OH-GSLs at the expense of the MT and MS-GSLs. Thus, *MYB115* and *MYB118* appear to lead to different shifts in GLS partitioning when abolishing parts of the BZ-GSL pathway. Further transcriptomic and metabolomic experiments are needed to better parse these changes.

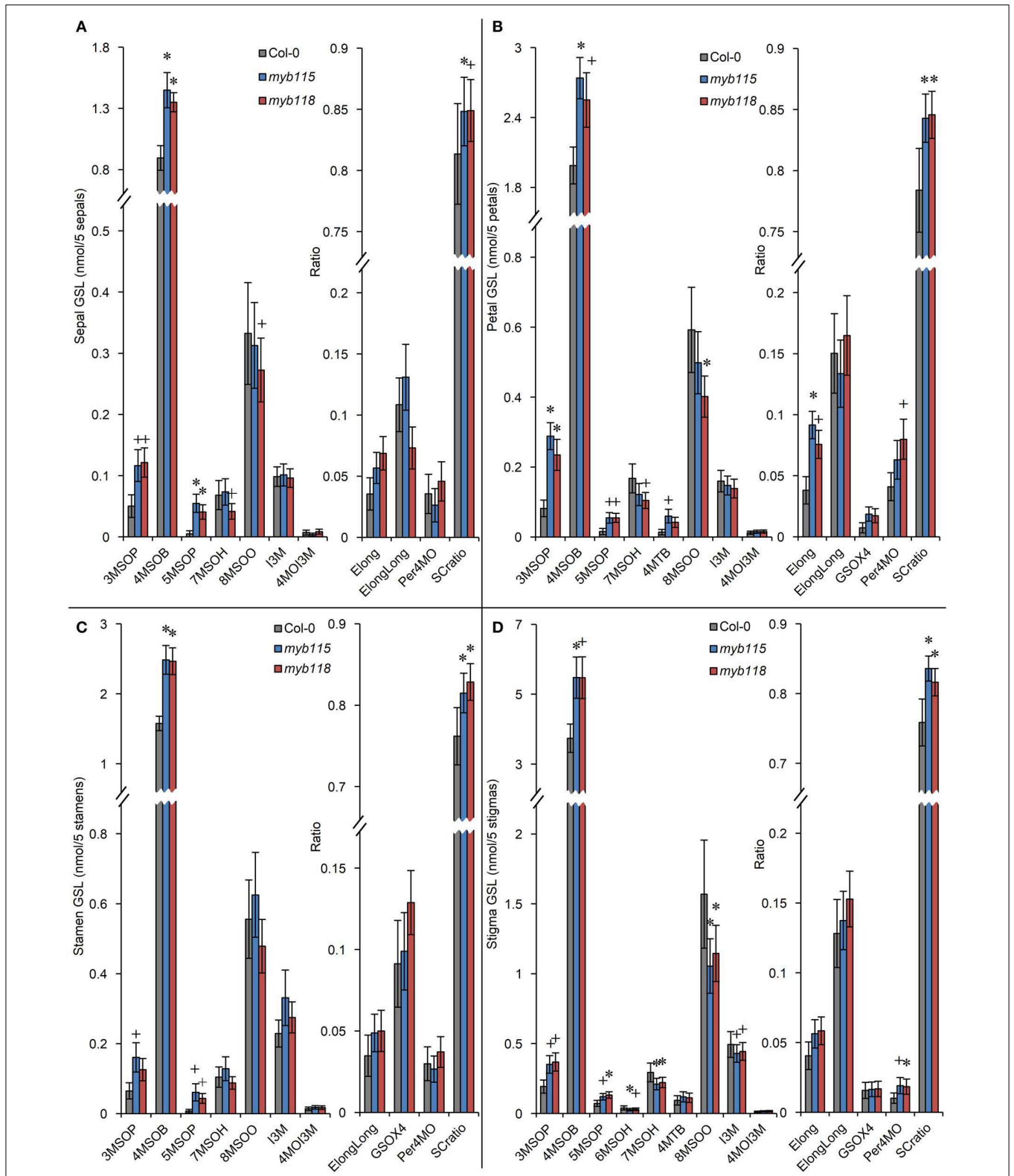
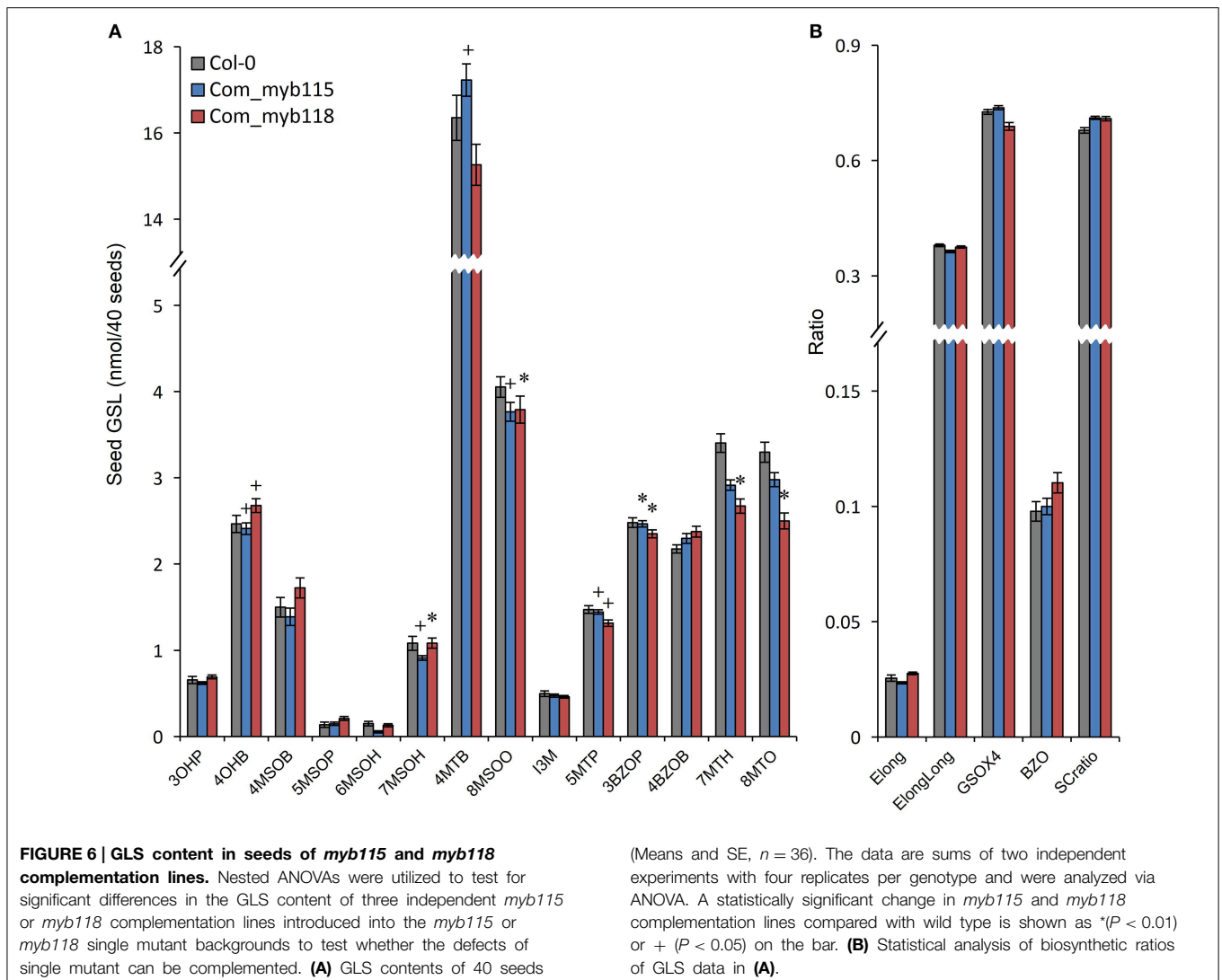


FIGURE 5 | GLS content in flower organs (sepal, petal, stamen, stigma) of *myb115* and *myb118* knockout mutants. GLS contents of sepals (A), petals (B), stamens (C), and stigmas (D) of 5 flowers and the respective statistical analysis of biosynthetic ratios. (Means and SE, $n = 24$).

The data are combined from four independent experiments and analyzed via ANOVA. A statistically significant change in *myb115* and *myb118* knockout mutants compared with wild type is shown as * ($P < 0.01$) or + ($P < 0.05$) on the bar.



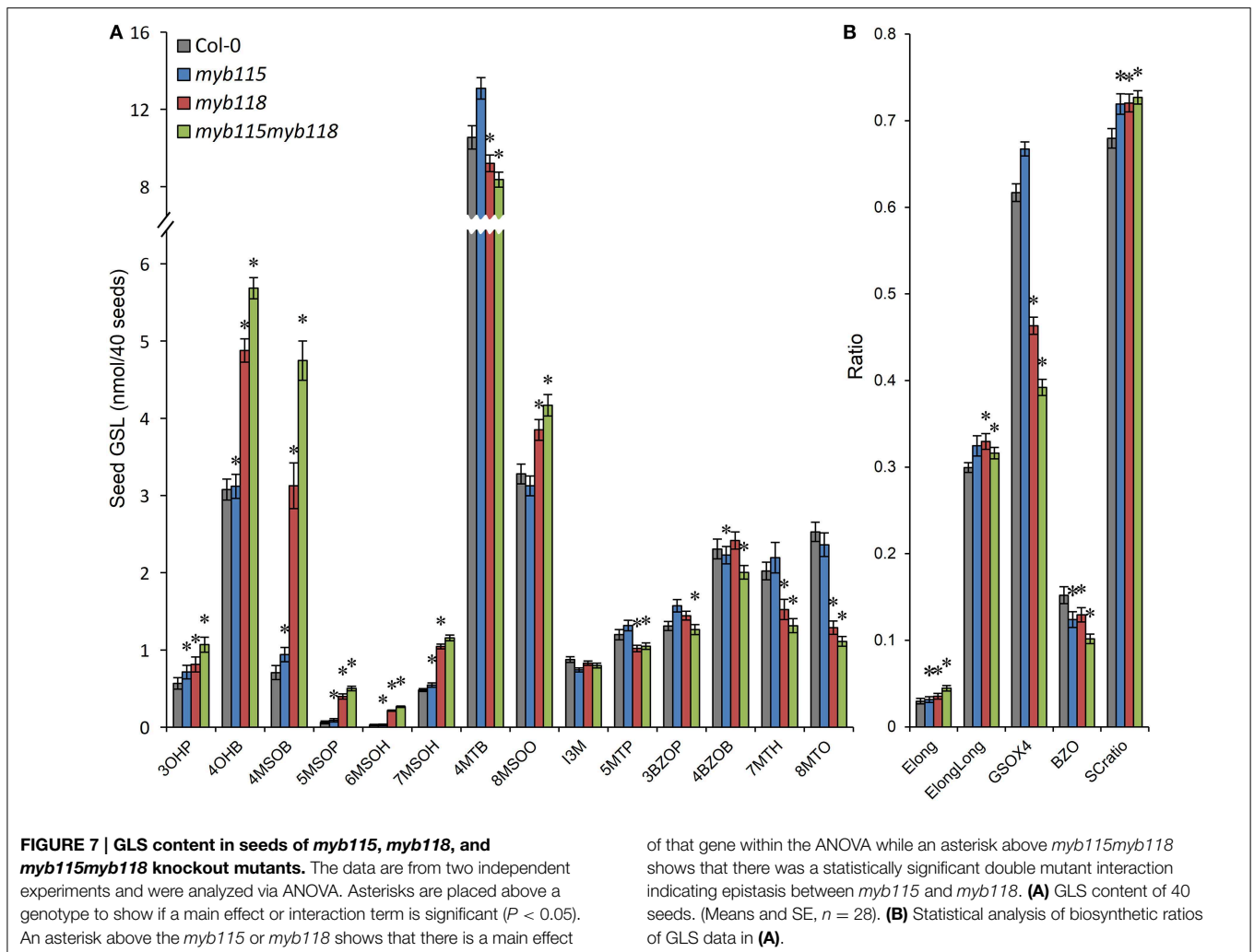
Discussion

MYB115 and MYB118 Are Conserved TFs that Regulate the Newly Evolved BZO1 and AOP3 Produced BZ-GLS in *Arabidopsis thaliana*

Plants can generate a vast diversity of metabolites, which is due to the diversified function of the biosynthetic genes encoding various metabolic enzymes, to cope with environmental changes. The evolution of this diversity is predicted to be driven by gene duplication and consequent neo- and sub-functionalization of enzymes (Kliebenstein, 2008). However, the corresponding pattern of diversification for TFs modulating secondary metabolites is less clear. For example, MYB TFs of subgroup 12 (MYB28/29/76 and MYB34/51/122) have evolved via whole-genome duplication to specifically modulate the GLS pathway genes (Bekaert et al., 2012). In contrast, the conserved JA regulators, MYC TFs of subgroup IIIe (MYC2/3/4/5), also control the GLS pathway but this is likely from regulatory capture rather than specialization for the pathway (Dombrecht

et al., 2007; Fernández-Calvo et al., 2011; Schweizer et al., 2013; Frerigmann et al., 2014; Li et al., 2014). Therefore, a question appears: do newly evolved metabolites need regulate by newly evolved TFs? To answer this question, we focused on the regulation of a pathway for a newly evolved GLS, BZ-GLSs, and identified *MYB115* and *MYB118*, two R2R3-MYB TFs co-expressed with BZ-GLS-related genes (*AOP3* and *BZO1*). *MYB115* and *MYB118* epistatically interacted to modulate the expression of the accompanying GLS genes and the respective metabolites in *Arabidopsis* seeds (Figures 7–9). *MYB115* and *MYB118* are conserved regulators of seed maturation suggesting that the new BZ-GLS biosynthetic pathway was captured by these TFs to impart the specific seed accumulation pattern for the BZ-GLS.

Interestingly, the genetic evidence suggests *MYB115* and *MYB118* function to repress the GLS genes. The identification of a repressor for defense compounds is not unexpected as these compounds are highly patterned in a spatial and ontogenic fashion which requires the combination of activators and



repressors to accomplish and to minimize any costs (Brown et al., 2003; Kliebenstein et al., 2007; Wentzell and Kliebenstein, 2008; Wentzell et al., 2008; Züst et al., 2011; Moussaieff et al., 2013). It is also not unusual to identify repressors via Yeast-1-hybrid because of the strength of the activation domain added to the TF can often overcome this activity (Gaudinier et al., 2011; Li et al., 2014; Taylor-Teeples et al., 2014). However, previous work has identified *MYB115* and *MYB118* as activators (Wang et al., 2009; Zhang et al., 2009). There are a number of hypothesis that could explain this discrepancy. First, if a number of TFs have been shown to be activators and repressors depending upon the promoter and other factors binding the promoter (Taylor-Teeples et al., 2014). Second, *MYB115* and *MYB118* may have a passive repressive role where they interfere with the interaction between the activator MYB28/29 and MYCs (Sønderby et al., 2010a; Schweizer et al., 2013; Frerigmann and Gigolashvili, 2014; Frerigmann et al., 2014). Finally, *MYB115* and *MYB118* could be involved in an incoherent feed-forward loop where they activate a unidentified repressor that overcomes any direct activation they may display. The MYB-binding sites have been classified into three types (Romero et al., 1998;

Prouse and Campbell, 2012). The MYB118-binding site was shown to belong to type I: pAACnG (where p indicates T or C, and n indicates any nucleotide) (Barthole et al., 2014) with additional influence of the flanking bases (Biedenkapp et al., 1988; Howe and Watson, 1991; Deng et al., 1996). This motif exists in *MYB118* regulated genes that are linked to seed maturation (Barthole et al., 2014). We found, the corresponding element are within the promoters of *AOP3* (TAACAG, at position -1071 to -1066 bp), *BZO1* (TAACAG, at position -345 to -340 bp), *BCAT4* (TAACCG, at position -470 to -465 bp), *CYP83A1* (CGGTTA, at position -132 to -127 bp), *MAM1* (TAACCTG, at position -2304 to -2299 bp), *MAM3* (TAACCTG, at position -1767 to -1762 bp), and *SCPL17* (CTGTGA, at position -6340 to -635 bp), which are all MYB115- and MYB118-binding promoters in Y1H assays (**Figures 2A,B**). This is similar to the observation that the conserved JA related MYC basic helix-loop-helix (bHLH) TFs also control GLS accumulation via the recruitment of the appropriate promoter elements into the enzyme encoding genes (Chini et al., 2007; Dombrecht et al., 2007; Fernández-Calvo et al., 2011; Schweizer et al., 2013; Frerigmann et al., 2014; Li et al., 2014). This

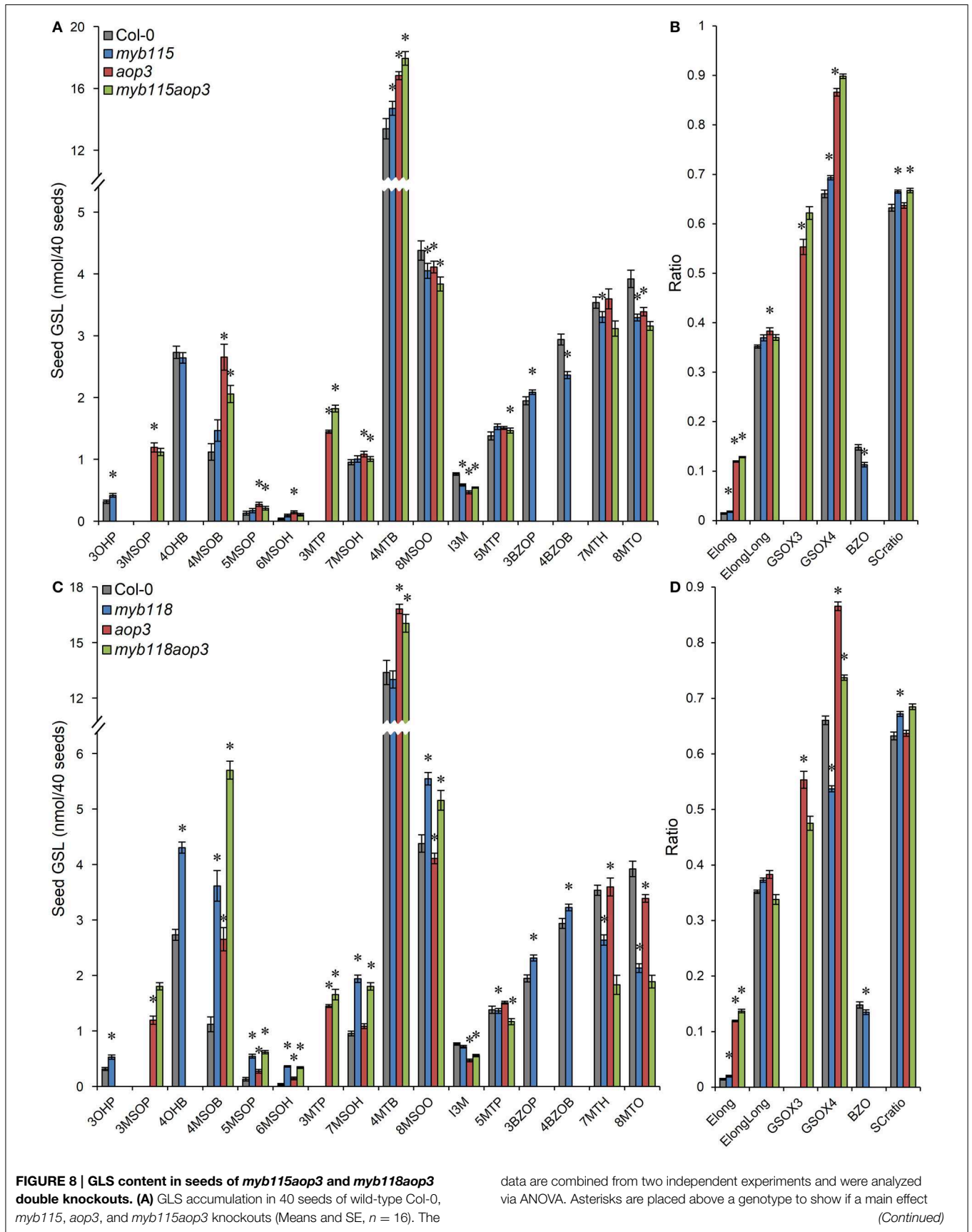


FIGURE 8 | Continued

or interaction term is significant ($P < 0.05$). An asterisk above the *myb115* or *aop3* shows that there is a main effect of that gene within the ANOVA while an asterisk above *myb115aop3* shows that there was a statistically significant double mutant interaction indicating epistasis between *myb115* and *aop3*. **(B)** Statistical analysis of biosynthetic ratios of GLS data in **(A)**. **(C)** GLS accumulation in 40 seeds of wild-type Col-0, *myb118*, *aop3*, and *myb118aop3* knockouts (Means and SE,

$n = 16$). The data are the combined results from two independent experiments and were analyzed via ANOVA. Asterisks are placed above a genotype to show if a main effect or interaction term is significant ($P < 0.05$). An asterisk above the *myb118* or *aop3* shows that there is a main effect of that gene within the ANOVA while an asterisk above *myb118aop3* shows that there was a statistically significant double mutant interaction indicating epistasis between *myb118* and *aop3*. **(D)** Statistical analysis of biosynthetic ratios of GLS data in **(C)**.

suggests that the promoters obtained elements to allow them to be regulated by these MYBs. Further studies are needed to elucidate the precise molecular function of *MYB115* and *MYB118* in the regulation of GLS biosynthesis, and to find the activators in the seed of BZ-GLS related genes as they are not regulated by the known MYBs or MYCs (Li et al., 2014). One option for these unknown components are the additional genes within the *MYB118* subgroup of R2R3-MYBs which may also play a role in regulating GLS accumulation but this remains to be tested (Dubos et al., 2010). Thus, newly evolved biosynthetic pathways can be rapidly captured by existing TFs to impart highly precise regulatory patterns without requiring the *de novo* evolution of new regulatory TFs.

Regulatory Capture of GLS Genes by *MYB115* and *MYB118* Goes beyond the New Pathway

We identified the *MYB115* and *MYB118* TFs based on co-expression analysis with BZ-GLS related genes (*AOP3*, *BZO1*, and *SCPL17*) but these co-expression networks had no other GLS genes within them. However, using the yeast one-hybrid assays, it could be demonstrated that *MYB115* and *MYB118* interact more broadly with more genes of the aliphatic GLS biosynthetic pathway, including also *BCAT4*, *MAM1*, *MAM3*, and *CYP83A1* (Figure 2). Thus, this list of potential target genes encompasses genes involved in all three steps of aliphatic GLS biosynthesis, side-chain elongation (*BCAT4*, *MAM1*, and *MAM3*), core structure pathway (*CYP83A1*) and side-chain modification genes (*BZO1*, *SCPL17*, and *AOP3*). This was however not universal as there was no observed interaction with *BCAT3*. In agreement with this broader interaction capacity, the seed GLS profiles of the knockouts had alterations in GLS phenotypes not associated with the production of BZ-GLS (Figure 3; Table S2). The combination of yeast one-hybrid analysis with seed GLS profiles of the knockouts suggests that *MYB115* and *MYB118* have captured the capacity to regulate genes in all steps of the GLS pathway including the new BZ-GLS genes (Figure 10).

MYB115 and *MYB118* Interact Genetically

Our qRT-PCR work showed that in single mutant *myb118*, the transcript level of *AOP3* and *BZO1* increased in comparison with the wild type suggesting that *MYB118* negatively regulate BZ-GLS related genes. In contrast, the single *myb115* mutant had no altered expression of the *AOP3* and *BZO1* genes (Figure 2C). However, *AOP3* and *BZO1* transcript levels of double *myb115myb118* mutants were significantly higher than

all single mutants and wild type indicating that *MYB115* had a genetic repressive activity on these genes that was redundantly in the presence of a functional *MYB118*. In agreement, the seeds GLS profiles of double mutant *myb115myb118* showed that *MYB115* and *MYB118* show a similar epistatic interaction in regulating the biosynthesis of most of GLSs. The combination of transcriptional analysis and seed GLS profiles of the knockouts suggests that there is an epistatic interaction between the two loci potentially via a regulatory interaction (Figure 10). This is not unprecedented within TFs controlling GLS genes. Work comparing single and double mutants of the *MYB28* and *MYB29* paralogs showed that they have a synergistic interaction potentially via the ability to regulate each other's gene expression (Sønderby et al., 2007, 2010a; Beekwilder et al., 2008; Gigolashvili et al., 2008). Further work is required to elucidate the molecular basis of the interaction between *MYB115* and *MYB118*.

MYB115 and *MYB118* Are Thought to Be Seed Specific, but GLS in Other Tissues Were Also Affected

Previous studies on *MYB118* and *MYB115* had suggested that these genes were largely functioning within the developing seed (Wang et al., 2009; Zhang et al., 2009; Barthole et al., 2014). This agreed with the observed expression pattern. Interesting, we could show that there was an effect of the *myb115* and *myb118* mutants on the accumulation of foliar GLSs (Figure 4; Table S2). This indicates that these genes have the capacity to influence the accumulation of GLSs in adult leaves before bolting. This however does not require these TFs to be directly GLS accumulation with in the leaves. One possibility is that because GLS can be transported in a bidirectional fashion within intact plants, there may be a unknown cell type in which they are expressed and affect global GLS patterning (Nour-Eldin et al., 2012; Madsen et al., 2014). For example, *MYB115* and *MYB118* do show some root expression under certain conditions (Gifford et al., 2008; Wang et al., 2009). Another more intriguing possibility is that by altering the accumulation of GLS in the developing seed, *MYB115* and *MYB118* alter downstream GLS regulatory decisions even when they are not present, effectively a maternal effect. Altering GLS accumulation can influence broad regulatory patterns suggesting that this is a possibility (Kerwin et al., 2011). Significant future work is required to test between these two potential indirect models of how *MYB115/MYB118* affects foliar GLS accumulation or if the residual expression of these genes in the leaf has direct consequences on the GLS accumulation in this tissue.

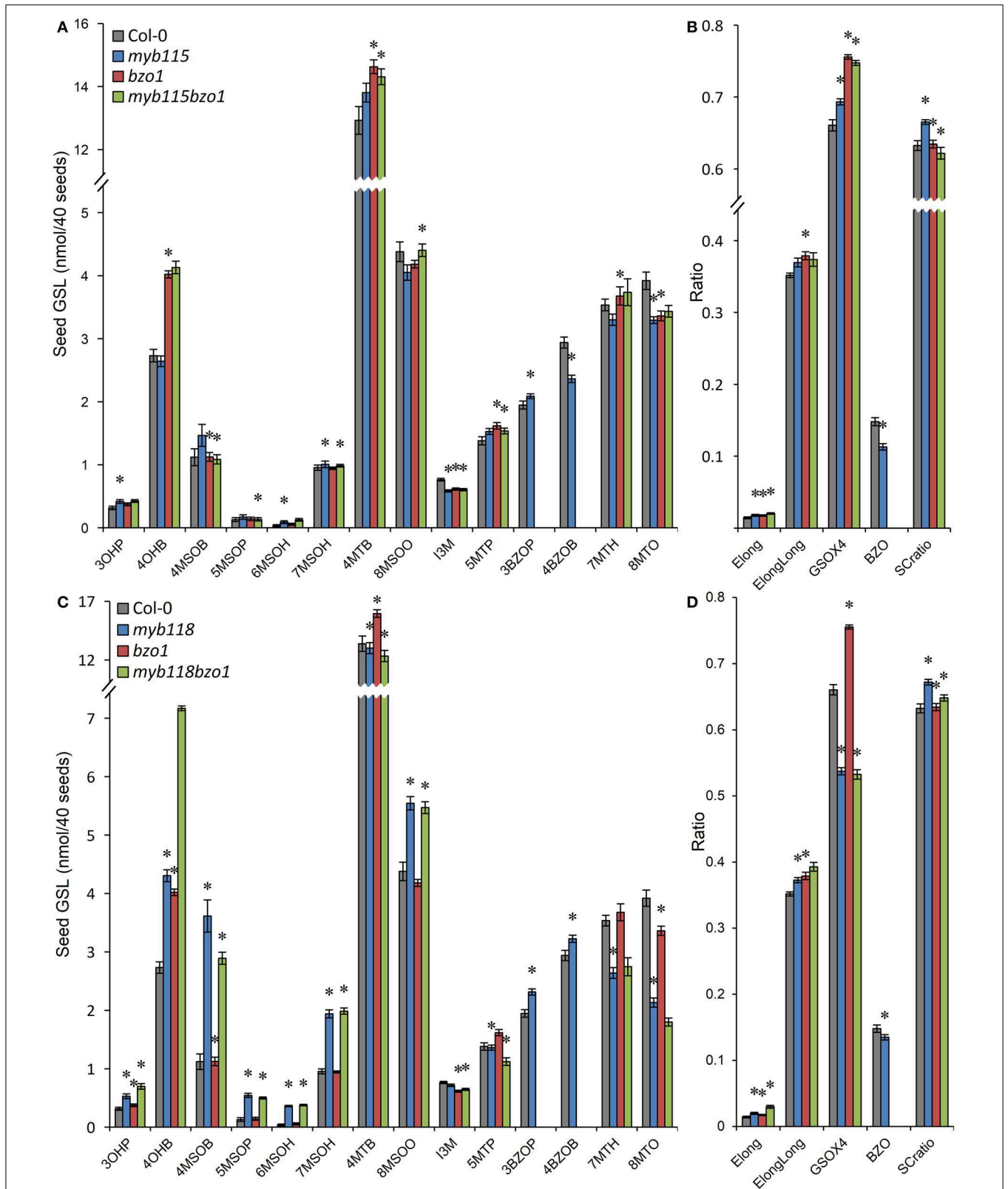


FIGURE 9 | GLS content in seeds of *myb115bzo1* and *myb118bzo1* double knockouts. (A) GLS accumulation in 40 seeds of wild-type Col-0, *myb115*, *bzo1*, and *myb115bzo1* knockouts (Means and SE, $n = 16$). The data are combined from two independent experiments and were analyzed

via ANOVA. Asterisks are placed above a genotype to show if a main effect or interaction term is significant ($P < 0.05$). An asterisk above the *myb115* or *bzo1* shows that there is a main effect of that gene within the ANOVA while
(Continued)

FIGURE 9 | Continued

an asterisk above *myb115bzo1* shows that there was a statistically significant double mutant interaction indicating epistasis between *myb115* and *bzo1*. **(B)** Statistical analysis of biosynthetic ratios of GLS data in **(A)**. **(C)** GLS accumulation in 40 seeds of wild-type Col-0, *myb118*, *bzo1*, and *myb118bzo1* knockouts (Means and SE, $n = 16$). The data are combined from two independent experiments and were

analyzed via ANOVA. Asterisks are placed above a genotype to show if a main effect or interaction term is significant ($P < 0.05$). An asterisk above the *myb118* or *bzo1* shows that there is a main effect of that gene within the ANOVA while an asterisk above *myb118bzo1* shows that there was a statistically significant double mutant interaction indicating epistasis between *myb118* and *bzo1*. **(D)** Statistical analysis of biosynthetic ratios of GLS data in **(C)**.

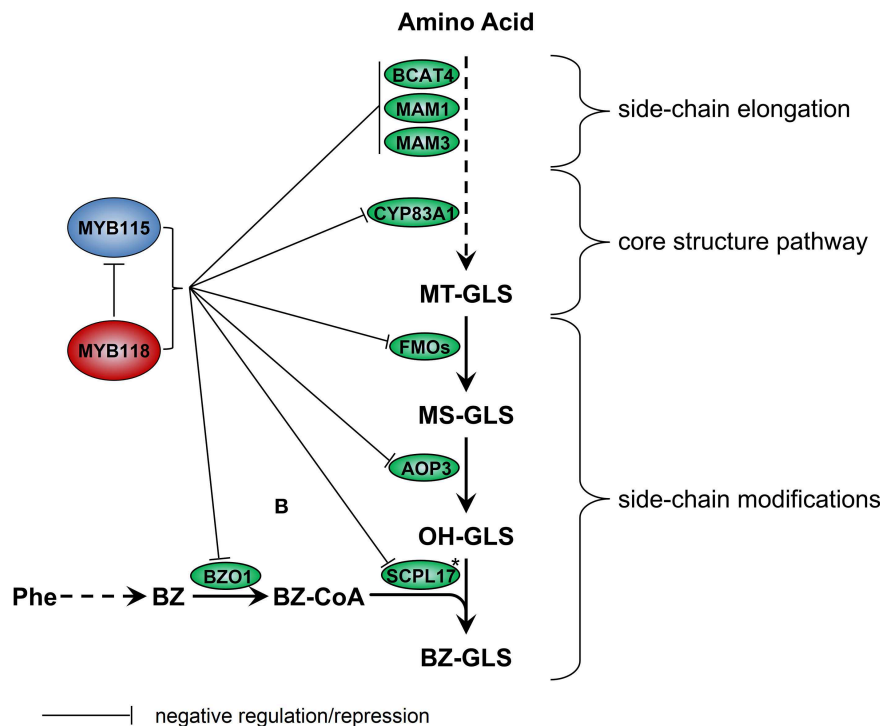


FIGURE 10 | Working model for regulation of GLS biosynthesis by MYB115 and MYB118 in *A. thaliana*. For simplicity, only genes discussed in the text have been included. Enzymes are shown in circles. Asterisks

partially characterized enzyme. Arrows show the combination of genetic and yeast one-hybrid evidence of how these MYBs influence the glucosinolate pathway transcription via an as to be yet identified molecular mechanism.

Conclusion

Using a comprehensive strategy of co-expression analysis, we found that two conserved TFs, MYB115, and MYB118, regulate the genes in the newly evolved BZ-GLS pathway within *A. thaliana*. We validated that these two TFs play important roles in transcriptional regulation of aliphatic GLS biosynthetic pathway, which are negative regulators to be identified in GLS biosynthesis. Our study suggests that newly evolved metabolites can be regulated by conserved TFs; this has led to an improved model for newly evolved genes regulation. Future experiments will be required to understand how the promoters in this pathway evolved to allow coordinated expression within the developing seed via the use of conserved TFs.

Materials and Methods

Plant Materials and Growth Conditions

All plants were grown in controlled-environment chambers with 16 h light, 100 mE light intensity, 21°C/18°C day/night

temperature, and 60% relative humidity. Wild type Col-0 controls were included in each flat to minimize any spatial aspects of the growth chamber. Before sowing, all seeds were stratified for 4 days in the dark at 4°C to break seed dormancy.

Co-expression Analysis

We performed a genome-wide co-expression analysis on BZ-GLS-related genes using ATTED-II (Obayashi et al., 2014). All these three genes were typed together (*AOP3*, AT4G03050; *BZO1*, AT1G65880, and *SCPL17*, AT3G12203) into CoExSearch box as query genes (http://atted.jp/top_search.shtml#CoExSearch). Under such a query, we got a list of the top 300 co-expressed genes for all the three genes. There are three MYB TFs among them, MYB58 (Rank#102, average correlation coefficient to query loci#0.22), MYB118 (Rank#198, average correlation coefficient to query loci#0.18) and MYB56 (Rank#291, average correlation coefficient to query loci#0.13).

T-DNA Insertion Mutants and Double Mutants

T-DNA insertion lines were obtained from the Arabidopsis Biological Resource Center stock center (*myb115*, SALK_044168;

myb118, SALK_111812; *aop3*, SALK_001655; *bzo1*, SALK_094196). Primers for genotyping are listed in Table S7. To construct the double mutant's *myb115myb118*, *myb115aop3*, *myb118zop3*, *myb115bzo1*, and *myb118bzo1*, the respective homozygous single knockouts were crossed with each other and F2 segregating progenies were genotyped to select homozygous mutations, and the F3 seeds were validated and used for the GLS phenotyping.

Construction of Complementation Vectors and Plant Transformation

For the genetic complementation study, the T-DNA insertion mutants *myb115* and *myb118* were complemented with genomic fragments spanning the genomic region of *MYB115* (3998 bp) and *MYB118* (4692 bp), respectively, which were amplified with the Phusion High-Fidelity DNA Polymerase (Biolabs) by PCR using Col-0 genomic DNA as templates. (primers for cloning are listed in Table S7). This genomic fragment was introduced in vector pKGW-RR containing a red fluorescent protein (DsRed) (Varma Penmetsa et al., 2008) by TOPO and the Gateway system. The resulting binary vectors were electroporated into *Agrobacterium tumefaciens* GV3101 strain and used for agroinfiltration of flower buds of *Arabidopsis* (Bechtold et al., 1993). The homozygous seeds were screened by DsRed fluorescence under the fluorescence microscope using dsRED-specific filter and further confirmed by PCR-based genotyping.

Yeast One-hybrid

The yeast one-hybrid assay was performed according to the manual of Matchmaker Gold Yeast One-Hybrid Library Screening System (Clontech). Briefly, the open reading frames of *MYB115* (1080 bp) and *MYB118* (1314 bp) were cloned into the pGADT7 from Col-0 cDNA of leaf material by Phusion High-Fidelity DNA Polymerase (Biolabs), as prey vectors. Cloning primers are listed in Table S7. For the promoter cloning, about 2000 bp genomic fragments of each promoter upstream of the translational start codon were fused to pAbAi as the bait vectors. All the bait vectors including p53-AbAi (positive control) were firstly transformed into the Y1H Gold yeast strain. After transformants were selected on SD/-Ura plates by PCR, the minimal inhibitory concentration of aureobasidin A (AbA; Clontech) was determined for the bait strains. All the prey vectors including a mutant W-box-AbAi (negative control) were introduced into the respective Y1H Gold strain. The co-transformed yeast cells were cultured on SD/-Leu plates with and without AbA and incubated at 30°C until colonies in the positive control (p53) were visible.

GLS Extraction and Analysis

For analysis of GLS content in all the genotypes, we planted four independent plants per genotype in a randomized design using 16 h long day length conditions. For leaf GLS analysis, we harvested one leaf from the first fully mature leaf pair of 4 week old plants. Under these conditions, the plants flower at about 7 weeks of age and as such, the plants were in mid-growth

stage. This leaf was removed and stored in 90% (v/v) methanol to inhibit enzymatic breakdown of chemical compounds and begin the extraction. Five flowers per plant were collected 6 weeks after sowing and combined to make a single biological sample per plant. The plants were then allowed to continue to develop and produce seeds. For seeds samples, we collected mature seeds from each plant individually and extracted GLSs from 40 seeds. For the dissecting flower samples, sepal, petal, stamen and stigma were dissected, and 5 organs from each genotype are used for extraction. GLSs from all tissues were extracted and analyzed by HPLC according to previously described methods (Kliebenstein et al., 2001b). The entire experiment was replicated twice and the data combined for statistical analysis.

GLS contents were analyzed via ANOVA utilizing a general linear model within R software (x64 3.1.2) (R Development Core Team, 2014). For single mutants, each of them was tested for altered GLS content in an individual ANOVA against the wild-type Col-0. For the complementation tests, three independent transgenic lines of *myb115* and *myb118* were obtained and a nested model was used wherein the independent transgene lines were nested within the higher genotype term. In all models, the separate independent experiments were combined and an experiment term was included in the model to test for effects.

Gene Expression Analysis

Developing seeds tissues were sampled at 6 and 7 days after pollination. For the expression study, nine independent biological samples spread across two experiments were utilized. Total RNA was extracted with TRIzol Reagent according to the manufacturer's instructions (Life Technologies). The extracts were treated with 2 units of Rnase-free Dnase I (New England Biolabs) to eliminate the residual genomic DNA present in the preparation and eluted with 50 µl of RNase-free Water. For reverse transcription, first-strand cDNA was synthesized from 2 µg of total RNA using the Invitrogen ThermoScript RT-PCR system (Life Technologies). qRT-PCRs was performed with three technical replicates on a Bio-Rad CFX96 Real-Time system (Bio-Rad) and DBI Bioscience Bestar-Real Time PCR Master Mix kit, following the manufacturer's instructions (DBI Bioscience). Primers for qRT-PCR are listed in Table S7. The data were analyzed with LINREG, as described by Ramakers et al. (2003). The experiment was repeated using at least three independent biological replicates.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: *AOP3*, At4g03050; *BCAT3*, At3g49680; *BCAT4*, At3g19710; *BZO1*, At1g65880; *CYP83A1*, At4g13770; *MAM1*, At5g23010; *MAM3*, At5g23020; *MYB115*, At5g40360; *MYB118*, At3g27785; *SCPL17*, At3g12203; and *UBC21*, At5g25760.

Acknowledgments

We would like to thank Dr. Douglas R. Cook (University of California, Davis) for kindly providing the vector pKGW-RR (DS-Red). This work was supported by the China Scholarship

Council grant 2010676009 to YYZ, the NSFC grant 31371659 to YMZ, the NSF DBI grant 0820580 to DK, the NSF MCB grant 1330337 to DK, the USDA National Institute of Food and Agriculture, Hatch project number CA-D-PLS-7033-H to DK and by the Danish National Research Foundation (DNRF99) grant to DK.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00343/abstract>

Figure S1 | Molecular characterization of *myb115*, *myb118*, *aop3*, and *bzo1* mutants. Structure of the *MYB115* (A), *MYB118* (B), *AOP3* (C), and *BZO1* (D) genes showing the position of T-DNA insertions in *myb115*, *myb118*, *aop3* and *bzo1* mutants are presented. Accumulation of *MYB115*, *MYB118*, *AOP3*, and *BZO1* mRNA in Col-0 and corresponding mutant backgrounds was measured by RT-PCR on developing seeds of 6 and 7 days after pollination.

References

- Agerbirk, N., and Olsen, C. E. (2012). Glucosinolate structures in evolution. *Phytochemistry* 77, 16–45. doi: 10.1016/j.phytochem.2012.02.005
- Barthole, G., To, A., Marchive, C., Brunaud, V., Soubigou-Taconnat, L., Berger, N., et al. (2014). MYB118 represses endosperm maturation in seeds of *Arabidopsis*. *Plant Cell* 26, 3519–3537. doi: 10.1105/tpc.114.130021
- Bechtold, N., Ellis, J., and Pelletier, G. (1993). In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *Comptes rendus de l'Académie des sciences. Série 3, Sciences de la vie* 316, 1194–1199.
- Bednarek, P., Pielewska-Bednarek, M., Svatoš, A., Schneider, B., Doubski, J., Mansurova, M., et al. (2009). A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science* 323, 101–106. doi: 10.1126/science.1163732
- Beekwilder, J., Van Leeuwen, W., Van Dam, N. M., Bertossi, M., Grandi, V., Mizzi, L., et al. (2008). The impact of the absence of aliphatic glucosinolates on insect herbivory in *Arabidopsis*. *PLoS ONE* 3:e2068. doi: 10.1371/journal.pone.0002068
- Bekaert, M., Edger, P. P., Hudson, C. M., Pires, J. C., and Conant, G. C. (2012). Metabolic and evolutionary costs of herbivory defense: systems biology of glucosinolate synthesis. *New Phytol.* 196, 596–605. doi: 10.1111/j.1469-8137.2012.04302.x
- Biedenkapp, H., Borgmeyer, U., Sippel, A. E., and Klempnauer, K.-H. (1988). Viral myb oncogene encodes a sequence-specific DNA-binding activity. *Nature* 335, 835–837. doi: 10.1038/335835a0
- Blanc, G., Hokamp, K., and Wolfe, K. H. (2003). A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res.* 13, 137–144. doi: 10.1101/gr.751803
- Brown, P. D., Tokuhisa, J. G., Reichelt, M., and Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* 62, 471–481. doi: 10.1016/S0031-9422(02)00549-6
- Chen, S., Glawischnig, E., Jørgensen, K., Naur, P., Jørgensen, B., Olsen, C. E., et al. (2003). CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in *Arabidopsis*. *Plant J.* 33, 923–937. doi: 10.1046/j.1365-3113X.2003.01679.x
- Chini, A., Fonseca, S., Fernandez, G., Adie, B., Chico, J., Lorenzo, O., et al. (2007). The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* 448, 666–671. doi: 10.1038/nature06006
- Clarke, D. B. (2010). Glucosinolates, structures and analysis in food. *Anal. Methods* 2, 310–325. doi: 10.1039/b9ay00280d
- Clay, N. K., Adio, A. M., Denoux, C., Jander, G., and Ausubel, F. M. (2009). Glucosinolate metabolites required for an *Arabidopsis* innate immune response. *Science* 323, 95–101. doi: 10.1126/science.1164627
- Ubiquitin-conjugating enzyme 21 (*UBC21*) gene expression was used as a constitutive control. Primers used for this study are indicated as arrows (see primers in Table S7).
- Figure S2 | Expression pattern of *FMO_{GS-ox2}*.** (A) The coexpression network analysis of *BZO1*, *SCPL17*, *MYB118*, and *FMO_{GS-ox2}* using the ATTED-II database (Obayashi et al., 2014). The transcription factor and genes are shown in circles. (B) eFP display (Winter et al., 2007) of transcript accumulation patterns of *FMO_{GS-ox2}* across a variety of *Arabidopsis* organs. In all cases, Red indicates higher levels of transcript accumulation and yellow indicates a lower level of transcript accumulation.
- Figure S3 | GLS contents in leaves of *myb115*, *myb118*, and *myb115myb118* knockout mutants.** The data are combined from three independent experiments and analyzed via ANOVA. Asterisks are placed above a genotype to show if a main effect or interaction term is significant ($P < 0.05$). An asterisk above the *myb115* or *myb118* shows that there is a main effect of that gene within the ANOVA while an asterisk above *myb115myb118* shows that there was a statistically significant double mutant interaction indicating epistasis between *myb115* and *myb118*. (A) GLS content of 4-week-old rosette leaves. (Means and SE, $n = 28$). The data are combined from three independent experiments and analyzed via ANOVA. The significant differences is shown as * ($P < 0.05$) on the bar. (B) Statistical analysis of biosynthetic ratios of GLS data in (A).
- Deng, Q.-L., Ishii, S., and Sarai, A. (1996). Binding site analysis of c-Myb: screening of potential binding sites by using the mutation matrix derived from systematic binding affinity measurements. *Nucleic Acids Res.* 24, 766–774. doi: 10.1093/nar/24.4.766
- Dombrecht, B., Xue, G. P., Sprague, S. J., Kirkegaard, J. A., Ross, J. J., Reid, J. B., et al. (2007). MYC2 differentially modulates diverse jasmonate-dependent functions in *Arabidopsis*. *Plant Cell* 19, 2225–2245. doi: 10.1105/tpc.106.048017
- Dubos, C., Stracke, R., Grotewold, E., Weissshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005
- Fahey, J. W., Zalcman, A. T., and Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56, 5–51. doi: 10.1016/S0031-9422(00)00316-2
- Fan, J., Crooks, C., Creissen, G., Hill, L., Fairhurst, S., Doerner, P., et al. (2011). *Pseudomonas sax* genes overcome aliphatic isothiocyanate-mediated non-host resistance in *Arabidopsis*. *Science* 331, 1185–1188. doi: 10.1126/science.1199707
- Fernández-Calvo, P., Chini, A., Fernández-Barbero, G., Chico, J.-M., Gimenez-Ibanez, S., Geerinck, J., et al. (2011). The *Arabidopsis* bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses. *Plant Cell* 23, 701–715. doi: 10.1105/tpc.110.080788
- Field, B., Cardon, G., Traka, M., Botterman, J., Vancanneyt, G., and Mithen, R. (2004). Glucosinolate and amino acid biosynthesis in *Arabidopsis*. *Plant Physiol.* 135, 828–839. doi: 10.1104/pp.104.039347
- Frerigmann, H., Berger, B., and Gigolashvili, T. (2014). bHLH05 is an interaction partner of MYB51 and a novel regulator of glucosinolate biosynthesis in *Arabidopsis*. *Plant Physiol.* 114:240887. doi: 10.1104/pp.114.240887
- Frerigmann, H., and Gigolashvili, T. (2014). MYB34, MYB51 and MYB122 distinctly regulate indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Mol. Plant* 7, 814–828. doi: 10.1093/mp/ssu004
- Gaudinier, A., Zhang, L., Reece-Hoyes, J. S., Taylor-Teeple, M., Pu, L., Liu, Z., et al. (2011). Enhanced Y1H assays for *Arabidopsis*. *Nat. Methods* 8, 1053–1055. doi: 10.1038/nmeth.1750
- Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., and Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 105, 803–808. doi: 10.1073/pnas.0709559105
- Gigolashvili, T., Berger, B., and Flügge, U.-I. (2009). Specific and coordinated control of indolic and aliphatic glucosinolate biosynthesis by R2R3-MYB transcription factors in *Arabidopsis thaliana*. *Phytochem. Rev.* 8, 3–13. doi: 10.1007/s11101-008-9112-6

- Gigolashvili, T., Engqvist, M., Yatusевич, R., Müller, C., and Flüggé, U. I. (2008). HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. *New Phytol.* 177, 627–642. doi: 10.1111/j.1469-8137.2007.02295.x
- Gigolashvili, T., Yatusевич, R., Berger, B., Müller, C., and Flüggé, U. I. (2007). The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant J.* 51, 247–261. doi: 10.1111/j.1365-313X.2007.03133.x
- Hansen, B. G., Kliebenstein, D. J., and Halkier, B. A. (2007). Identification of a flavin-monoxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in *Arabidopsis*. *Plant J.* 50, 902–910. doi: 10.1111/j.1365-313X.2007.03101.x
- Hansen, C. H., Wittstock, U., Olsen, C. E., Hick, A. J., Pickett, J. A., and Halkier, B. A. (2001). Cytochrome P450 CYP79F1 from *Arabidopsis* catalyzes the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic glucosinolates. *J. Biol. Chem.* 276, 11078–11085. doi: 10.1074/jbc.M010123200
- Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., and Bailey, P. C. (2003). The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* 20, 735–747. doi: 10.1093/molbev/msg088
- Hirai, M. Y., Sugiyama, K., Sawada, Y., Tohge, T., Obayashi, T., Suzuki, A., et al. (2007). Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6478–6483. doi: 10.1073/pnas.0611629104
- Howe, K. M., and Watson, R. J. (1991). Nucleotide preferences in sequence-specific recognition of DNA by c-myb protein. *Nucleic Acids Res.* 19, 3913–3919. doi: 10.1093/nar/19.14.3913
- Hull, A. K., Vij, R., and Celenza, J. L. (2000). *Arabidopsis* cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 97, 2379–2384. doi: 10.1073/pnas.040569997
- Kerwin, R. E., Jimenez-Gomez, J. M., Fulop, D., Harmer, S. L., Maloof, J. N., and Kliebenstein, D. J. (2011). Network quantitative trait loci mapping of circadian clock outputs identifies metabolic pathway-to-clock linkages in *Arabidopsis*. *Plant Cell* 23, 471–485. doi: 10.1105/tpc.110.082065
- Kliebenstein, D., Lambrix, V., Reichelt, M., Gershenzon, J., and Mitchell-Olds, T. (2001a). Gene duplication and the diversification of secondary metabolism: side chain modification of glucosinolates in *Arabidopsis thaliana*. *Plant Cell* 13, 681–693.
- Kliebenstein, D., Pedersen, D., Barker, B., and Mitchell-Olds, T. (2002). Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidopsis thaliana*. *Genetics* 161, 325–332. Available online at: <http://www.genetics.org/content/161/1/325.full>
- Kliebenstein, D. J. (2008). A role for gene duplication and natural variation of gene expression in the evolution of metabolism. *PLoS ONE* 3:e1838. doi: 10.1371/journal.pone.0001838
- Kliebenstein, D. J. (2013). New synthesis—regulatory evolution, the veiled world of chemical diversification. *J. Chem. Ecol.* 39, 349–349. doi: 10.1007/s10886-013-0274-3
- Kliebenstein, D. J., D'Auria, J. C., Behere, A. S., Kim, J. H., Gunderson, K. L., Breen, J. N., et al. (2007). Characterization of seed-specific benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. *Plant J.* 51, 1062–1076. doi: 10.1111/j.1365-313X.2007.03205.x
- Kliebenstein, D. J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., et al. (2001b). Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiol.* 126, 811–825. doi: 10.1104/pp.126.2.811
- Kliebenstein, D. J., Lambrix, V. M., Reichelt, M., Gershenzon, J., and Mitchell-Olds, T. (2001c). Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis*. *Plant Cell* 13, 681–693. doi: 10.1105/tpc.13.3.681
- Kliebenstein, D. J., and Osbourn, A. (2012). Making new molecules—evolution of pathways for novel metabolites in plants. *Curr. Opin. Plant Biol.* 15, 415–423. doi: 10.1016/j.pbi.2012.05.005
- Kroymann, J., Donnerhacke, S., Schnabelrauch, D., and Mitchell-Olds, T. (2003). Evolutionary dynamics of an *Arabidopsis* insect resistance quantitative trait locus. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14587–14592. doi: 10.1073/pnas.1734046100
- Kroymann, J., Textor, S., Tokuhisa, J. G., Falk, K. L., Bartram, S., Gershenzon, J., et al. (2001). A gene controlling variation in *Arabidopsis* glucosinolate composition is part of the methionine chain elongation pathway. *Plant Physiol.* 127, 1077–1088. doi: 10.1104/pp.010416
- Lee, S., Kaminaga, Y., Cooper, B., Pichersky, E., Dudareva, N., and Chapple, C. (2012). Benzoylation and sinapoylation of glucosinolate R-groups in *Arabidopsis*. *Plant J.* 72, 411–422. doi: 10.1111/j.1365-313X.2012.05096.x
- Li, B., Gaudinier, A., Tang, M., Taylor-Teeple, M., Nham, N. T., Ghaffari, C., et al. (2014). Promoter based integration in plant defense regulation. *Plant Physiol.* 166, 1803–1820. doi: 10.1104/pp.114.248716
- Li, J., Hansen, B. G., Ober, J. A., Kliebenstein, D. J., and Halkier, B. A. (2008). Subclade of flavin-monoxygenases involved in aliphatic glucosinolate biosynthesis. *Plant Physiol.* 148, 1721–1733. doi: 10.1104/pp.108.125757
- Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151–1155. doi: 10.1126/science.290.5494.1151
- Madsen, S. R., Olsen, C. E., Nour-Eldin, H. H., and Halkier, B. A. (2014). Elucidating the role of transport processes in leaf glucosinolate distribution. *Plant Physiol.* 166, 1450–1462. doi: 10.1104/pp.114.246249
- Malitsky, S., Blum, E., Less, H., Venger, I., Elbaz, M., Morin, S., et al. (2008). The transcript and metabolite networks affected by the two clades of *Arabidopsis* glucosinolate biosynthesis regulators. *Plant Physiol.* 148, 2021–2049. doi: 10.1104/pp.108.124784
- Mikkelsen, M. D., Hansen, C. H., Wittstock, U., and Halkier, B. A. (2000). Cytochrome P450 CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. *J. Biol. Chem.* 275, 33712–33717. doi: 10.1074/jbc.M001667200
- Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T. W., Belcher, H., et al. (2013). High-resolution metabolic mapping of cell types in plant roots. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1232–E1241. doi: 10.1073/pnas.1302019110
- Niu, Y., and Figueroa, P. (2011). Characterization of JAZ-interacting bHLH transcription factors that regulate jasmonate responses in *Arabidopsis*. *J. Exp. Bot.* 62, 2143–2154. doi: 10.1093/jxb/erq408
- Nour-Eldin, H. H., Andersen, T. G., Burrow, M., Madsen, S. R., Jørgensen, M. E., Olsen, C. E., et al. (2012). NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* 488, 531–534. doi: 10.1038/nature11285
- Obayashi, T., Okamura, Y., Ito, S., Tadaka, S., Aoki, Y., Shirota, M., et al. (2014). ATTED-II in 2014: evaluation of gene coexpression in agriculturally important plants. *Plant Cell Physiol.* 55, e6–e6. doi: 10.1093/pcp/pct178
- Pfalz, M., Mikkelsen, M. D., Bednarek, P., Olsen, C. E., Halkier, B. A., and Kroymann, J. (2011). Metabolic engineering in *Nicotiana benthamiana* reveals key enzyme functions in *Arabidopsis* indole glucosinolate modification. *Plant Cell* 23, 716–729. doi: 10.1105/tpc.110.081711
- Pfalz, M., Vogel, H., and Kroymann, J. (2009). The gene controlling the indole glucosinolate modifier1 quantitative trait locus alters indole glucosinolate structures and aphid resistance in *Arabidopsis*. *Plant Cell* 21, 985–999. doi: 10.1105/tpc.108.063115
- Prouse, M. B., and Campbell, M. M. (2012). The interaction between MYB proteins and their target DNA binding sites. *Biochim. Biophys. Acta Gene Reg. Mech.* 1819, 67–77. doi: 10.1016/j.bbagr.2011.10.010
- Ramakers, C., Ruijter, J. M., Deprez, R. H. L., and Moorman, A. F. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 339, 62–66. doi: 10.1016/S0304-3940(02)01423-4
- R Development Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Project for Statistical Computing. Available online at: <http://www.r-project.org>.
- Romero, I., Fuertes, A., Benito, M., Malpica, J., Leyva, A., and Paz-Ares, J. (1998). More than 80R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J.* 14, 273–284. doi: 10.1046/j.1365-313X.1998.00113.x
- Schweizer, F., Fernández-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., et al. (2013). *Arabidopsis* basic helix-loop-helix transcription factors MYC2, MYC3, and MYC4 regulate glucosinolate biosynthesis, insect performance, and feeding behavior. *Plant Cell* 25, 3117–3132. doi: 10.1105/tpc.113.115139

- Sønderby, I. E., Burow, M., Rowe, H. C., Kliebenstein, D. J., and Halkier, B. A. (2010a). A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in Arabidopsis. *Plant Physiol.* 153, 348–363. doi: 10.1104/pp.109.149286
- Sønderby, I. E., Geu-Flores, F., and Halkier, B. A. (2010b). Biosynthesis of glucosinolates—gene discovery and beyond. *Trends Plant Sci.* 15, 283–290. doi: 10.1016/j.tplants.2010.02.005
- Sønderby, I. E., Hansen, B. G., Bjarnholt, N., Ticconi, C., Halkier, B. A., and Kliebenstein, D. J. (2007). A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *PLoS ONE* 2:e1322. doi: 10.1371/journal.pone.0001322
- Stotz, H. U., Sawada, Y., Shimada, Y., Hirai, M. Y., Sasaki, E., Krischke, M., et al. (2011). Role of camalexin, indole glucosinolates, and side chain modification of glucosinolate-derived isothiocyanates in defense of Arabidopsis against *Sclerotinia sclerotiorum*. *Plant J.* 67, 81–93. doi: 10.1111/j.1365-313X.2011.04578.x
- Taylor-Teeple, M., Lin, L., De Lucas, M., Turco, G., Toal, T., Gaudinier, A., et al. (2014). An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* 517, 571–575. doi: 10.1038/nature14099
- Textor, S., De Kraker, J.-W., Hause, B., Gershenzon, J., and Tokuhsa, J. G. (2007). MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in Arabidopsis. *Plant Physiol.* 144, 60–71. doi: 10.1104/pp.106.091579
- Varma Penmetsa, R., Uribe, P., Anderson, J., Lichtenzweig, J., Gish, J. C., Nam, Y. W., et al. (2008). The *Medicago truncatula* ortholog of Arabidopsis EIN2, sickle, is a negative regulator of symbiotic and pathogenic microbial associations. *Plant J.* 55, 580–595. doi: 10.1111/j.1365-313X.2008.03531.x
- Vision, T. J., Brown, D. G., and Tanksley, S. D. (2000). The origins of genomic duplications in Arabidopsis. *Science* 290, 2114–2117. doi: 10.1126/science.290.5499.2114
- Wang, X., Niu, Q.-W., Teng, C., Li, C., Mu, J., Chua, N.-H., et al. (2009). Overexpression of PGA37/MYB118 and MYB115 promotes vegetative-to-embryonic transition in Arabidopsis. *Cell Res.* 19, 224–235. doi: 10.1038/cr.2008.276
- Wentzell, A. M., Boeye, I., Zhang, Z., and Kliebenstein, D. J. (2008). Genetic networks controlling structural outcome of glucosinolate activation across development. *PLoS Genet.* 4:e1000234. doi: 10.1371/journal.pgen.1000234
- Wentzell, A. M., and Kliebenstein, D. J. (2008). Genotype, age, tissue, and environment regulate the structural outcome of glucosinolate activation. *Plant Physiol.* 147, 415–428. doi: 10.1104/pp.107.115279
- Wink, M. (1988). Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. *Theor. Appl. Genet.* 75, 225–233. doi: 10.1007/BF00303957
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2:e718. doi: 10.1371/journal.pone.0000718
- Wolfe, K. H., and Shields, D. C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387, 708–712. doi: 10.1038/42711
- Zhang, Y., Cao, G., Qu, L.-J., and Gu, H. (2009). Involvement of an R2R3-MYB transcription factor gene AtMYB118 in embryogenesis in Arabidopsis. *Plant Cell Rep.* 28, 337–346. doi: 10.1007/s00299-008-0644-4
- Züst, T., Joseph, B., Shimizu, K. K., Kliebenstein, D. J., and Turnbull, L. A. (2011). Using knockout mutants to reveal the growth costs of defensive traits. *Proc. R. Soc. B Biol. Sci.* 278, 2597–2603. doi: 10.1098/rspb.2010.2475

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zhang, Li, Huai, Zhou and Kliebenstein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.