Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

Master of Science, Wan Zhang

Born in: Shijiazhuang, Hebei, China

Oral examination: 2nd March, 2020

Regulatory functions of two R2R3-MYB transcriptional repressors in *Miscanthus* phenylpropanoid pathway: Impact on the lignification process

Referees:

Prof. Dr. Thomas Rausch

Dr. Sebastian Wolf

Table of Contents

Summary

For implementation of a sustainable bioeconomy, biorefineries will play a crucial role in converting different biomasses into various platform molecules. For biorefineries using lignocellulose biomass, Miscanthus is one of the ideal sources because of high-yield potential, low-input requirements, and high energy outcome ratios. However, with respect to the lignin component, the demand for high-value products from isolated lignin requires lignin feedstocks with unique properties. Therefore, a better understanding of lignification and monolignol biosynthesis is mandatory. For several model plants (Arabidopsis, rice, poplar), the lignin biosynthesis pathway has been elucidated in extensive detail. In particular, the transcriptional regulatory network of lignin biosynthesis as well as of secondary cell wall formation has attracted attention of worldwide research. However, how transcriptional repressors are involved in regulating lignin biosynthesis in Miscanthus has remained largely unknown.

In this study, two R2R3-MYB transcription repressors, MsMYB31 and MsMYB42 were identified from *Miscanthus sinensis*. Sequence and expression analysis revealed their close structural relationship with AtMYB4, ZmMYB31 and ZmMYB42, transcription factors which have been identified as negative regulators of lignin biosynthesis and the phenylpropanoid pathway. Further characterization of both repressors has been performed via subcellular localization and functional analysis, e.g. via dual-luciferase-assays (DLA) and electrophoretic mobility shift assays (EMSA) to confirm their mode of action and specificity of binding to certain cis-elements in target gene promoters, i.e. MsC4H, MsCCR and MsCAD. Inducible expression of MsMYB31 or MsMYB42 in wild type *Arabidopsis* Col-0 further confirmed repression of phenylpropanoid metabolic pathway by both repressor proteins. Additionally, transforming the Arabidopsis myb4 mutant with MsMYB31 or MsMYB42 under control of the AtMYB4 promoter revealed that both repressors do not

complement the function of AtMYB4, indicating similar but different mechanisms of these repressors. Finally, while showing target redundancy, the differential developmental expression patterns of MsMYB31 and MsMYB42 indicate specific regulatory functions during lignification in planta. Possible physiological functions of both repressors are discussed.

Zusammenfassung

Für die Realisierung einer nachhaltigen Bioökonomie werden Bioraffenerien zur Umwandlung verschiedener Biomassen in Plattformchemikalien eine wichtige Rolle spielen. Für auf Basis von Lignozellulose arbeitende Bioraffenerien stellt Miscanthus eine ideale Ressouce da, in Folge seines hohen Ertragspotentials, der niedrigen in-put Anforderungen und der hohen potentiellen Energieausbeute. In Bezug auf die Ligninkomponente erfordert die Nachfrage nach hochwertigen Endprodukten jedoch Lignin-haltige Biomasse mit möglichst definierter Lignin-Zusammensetzung. Aus diesem Grund ist ein besseres Verständnis der Lignifizierung und der Monolignolbiosynthese erforderlich. Für verschiedene Modellpflanzen (Arabidopsis, Reis, Pappel) wurde der Lignin-Biosyntheseweg bereits weitgehend aufgeklärt. Hierbei lag der Schwerpunkt weltweiter Forschungsanstrengungen auf dem transkriptionellen Regulationbsnetzwerk, sowie auf der Bildung der sekundären Zellwand. Im Gegensatz hierzu blieb die Rolle transkriptioneller Repressoren, insbesondere für Miscanthus, bisher weitgehend unverstanden.

In dieser Arbeit wurden zwei R2R3-MYB Transkriptionsrepressoren, MsMYB31 und MsMYB42, aus *Miscanthus sinensis* identifiziert. Sequenz- und Expressionsanalysen zeigten eine nahe Verwandschaft zu den Repressoren AtMYB4, ZmMYB31 und ZmMYB42, Transkriptionsfaktoren, die bereits als negative Regulatoren der Lignin-Biosynthese bzw. des Phnenylpropanoid-Biosynthesewegs identifiziert worden waren. Die weitergehende Charakterisierung dieser Repressoren konzentrierte sich auf die Frage ihrer subzellulären Lokalisierung, sowie auf die Analyse ihrer Funktionen, d. h. über Dual Luciferase Assay (DLA) und Electrophoretic Mobility Shift Assay (EMSA), um so ihre Aktivitätsmodi und ihre Binde-Spezifitäten bezüglich bestimmter cis-Elemente in den Promotoren ihrer Zielgene zu bestimmen, z.B. MsC4H, MsCCR und MsCAD. Über induzierbare Expression von MsMYB31

und MsMYB42 in Arabidopsis Col-0 WT konnte die Repressorfunktion beider Proteine auf den Phenylpropanoid-Syntheseweg bestätigt werden. Weiterhin wurde eine Arabidopsis myb4-Mutante mit MsBYB31 oder MsMYB42, jeweils unter Kontrolle des AtMYB4-Promoters, transformiert. Hierbei zeigte sich, dass beide Miscanthus Repressoren die Funktion von AtMYB4 nicht komplementieren konnten, ein Hinweis auf ähnliche aber nicht identische Funktionen. Abschliessend konnte demonstriert werden, das beide Repressoren, MsMYB31 und MsMYB42, zwar eine gewisse Redundanz hinsichtlich ihrer Zielgene (bzw. Promotoren) aufwiesen, ihre deutlich unterschiedlichen entwicklungsabhängigen Expressionsprofile aber differentielle Funtionen im Verlauf der pflanzlichen Entwicklung indizieren. Mögliche physioogische Funktionen beider Repressoren werden diskutiert.

1 Introduction

1.1 Biorefinery - the sustainable future

1.1.1 Biomass resources

With the development of the global economy and technology, the increasing demand of energy is leading to environment deterioration and also challenging the traditional fossil fuel resources. The exploration for sustainable resources driven by this urgent situation has gained great attention.

In recent years, biorefinery approaches inspired from traditional refinery ways, are showing the possibilities to convert biomass into various bio-products in a "greener" way (Cherubini, 2010, Menon and Rao, 2012). In general, three main feedstocks for biomass are from agriculture, from forestry and from waste (Demirbaş, 2001). Among them, lignocellulosic biomass stands out with great potential for biofuel and biomaterial production. Lignocellulosic biomass is the most abundant biomass produced on Earth, the annual production is about 181.5 billion tones worldwide (Paul and Dutta, 2018). The utilization of lignocellulosic biomass for biorefinery concepts can mainly divided into two types, lignocellulosic and syngas biorefinery ways, which can be regarded as complementary approaches (Dahmen et al., 2018). The lignocellulosic way is to decompose the biomass firstly into natural intermediate units, the intermediate products are then converted into biobased chemicals and materials with further conversion methods (Figure 1-1A) (Harmsen and Hackmann, 2013, Brodin et al., 2017, Lask et al., 2019). In syngas biorefineries, the biomass will go first into gasification processes and then the hydrogen and carbon monoxide produced can be converted into fuels and chemicals (Figure 1-1B) (Dahmen et al., 2017).

Many perennial grasses such as *Miscanthus* are ideal sources for producing lignocellulosic biomass because of their high-yield potential, low-input requirements, and high energy outcome ratios (Heaton et al., 2004, Shinners et al., 2010). Figure 1-2

shows three commonly cultivated perennial grasses, recently used as lignocellulosic biomass sources.

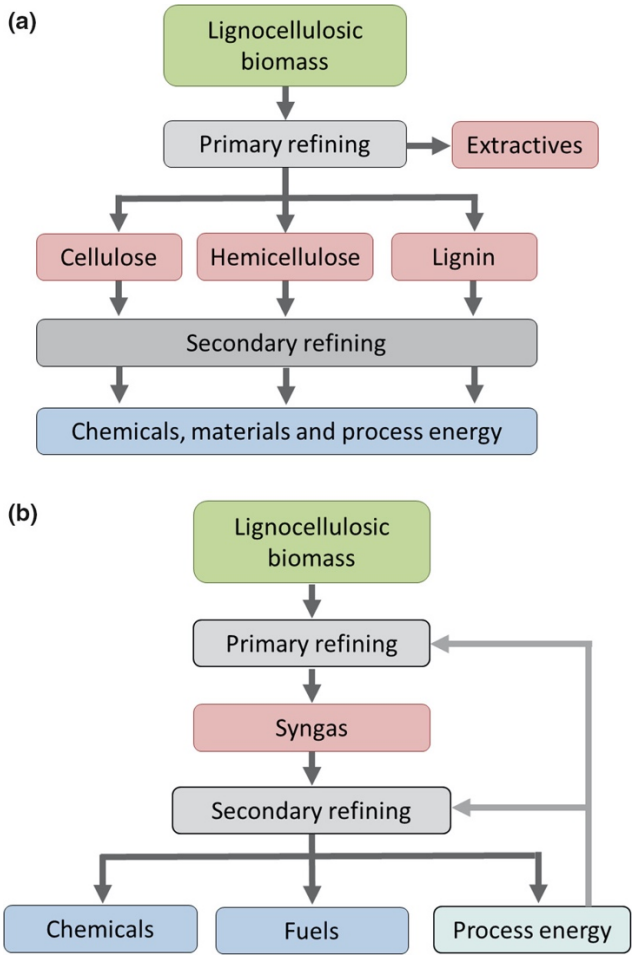

Figure 1-1 Biorefinery concepts making use of lignocellulosic biomass.
(a) The concept of lignocellulosic biorefinery; (b) The syngas biorefinery (Dahmen et al., 2018).

Figure 1-2 Examples of perennial grasses used in EU as lignocellulosic biomass sources.
Switchgrass, *Panicum virgatum L.*, picture sourse: https://jgi.doe.gov/developing-switchgrass-for-biomass-production/. *Miscanthus*, *Miscanthus x giganteus*, picture sourse: https://www.biooekonomie-bw.de. Giant reed, *Arundo donax L.*, picture sourse: http://www.freenatureimages.eu/Plants.

Table 1-1 gives an overview of perennial grasses tested in Europe as energy crops (Lewandowski et al., 2003). Perennial biomass crops (PBCs) are likely to play a vital role in biomass supply in the future (Dahmen et al., 2018). It can be concluded from the table that as energy crops, C4 plants have relatively high yield compared to C3 plants probably due to their more efficient photosynthesis mechanism (Heaton et al., 2004, Dohleman et al., 2009, Olson et al., 2012). *Miscanthus* is one of the most popular PBC planted in Europe in recent years (Brosse et al., 2012, Rivas et al., 2019).

Table 1-1 Perennial grasses grown or tested as energy crops in Europe

Common English name	Latin name	Photosynthetic pathway	Yields reported [t dry matter/ha.year]
Miscanthus	*Miscanthus spp.*	C4	5-44
Switchgrass	*Panicum virgatum L.*	C4	5-24
Giant Reed	*Arundo donax L.*	C3	5-37
Reed canarygrass	*Phalaris arundinacea L.*	C3	5-13
Meadow Foxtail	*Alopecurus pratensis L.*	C3	6-13
Big Bluestem	*Andropogon gerardii Vitman*	C4	8-15
Cypergras, Galingale	*Cyperus longus L.*	C4	4-19
Cocksfoot grass	*Dactylis glomerata L.*	C3	8-10
Tall Fescue	*Festuca arundinacea Schreb.*	C3	8-14
Raygras	*Lolium ssp.*	C3	9-12
Napier Grass	*Pennisetum purpureum Schum*	C4	27
Timothy	*Phleum pratense L.*	C3	9-18
Common Reed	*Phragmites communis Trin.*	C3	9-13
Sugar cane	*Saccharum officinarum L.*	C4	27
Giant Cordgrass	*Spartina cynosuroides L.*	C4	5-20
Prairie Cordgrass	*Spartina pectinata Bosc.*	C4	4-18

1.1.2 *Miscanthus*, one of the outstanding PBCs

The genus *Miscanthus* contains about 15 to 20 species (Brosse et al., 2012), originated from subtropical and tropical regions of Africa and Asia (Lygin et al., 2011). *Miscanthus* grows freely as weeds. Some species were used as ornamental plants before *Miscanthus* was discovered as energy crop. In previous times, its excellent fiber properties were also ideal for papermaking.

Miscanthus was gradually developed from a wild plant to an important energy crop. The most important reason is that *Miscanthus* possesses the basic characteristics that an ideal PBC should have (Lewandowski et al., 2003, Harvey, 2007, Heaton et al., 2008, Heaton et al., 2010, Zhuang et al., 2013).

High biomass yield with good quality

The photosynthetic efficiency of *Miscanthus* is higher than that of C3 plants. It is now one of the energy plants with the highest dry matter production. It takes 3-5 years for *Miscanthus* plants to be fully established. Yield reports indicated that in Europe, from the third year onwards in the spring harvest *Miscanthus x giganteus* could reach a 25 t ha^{-1} $year^{-1}$ (dry matter) yield while the highest yield obtained was 44 t ha^{-1} in Northern Greece (I. Lewandowskia, 2000). In the United States, a three-year trial of biomass production comparison between *Miscanthus* and Switchgrass revealed that within three years, the average peak biomass production of *Miscanthus* (38.2±2.3 t ha^{-1}) has reached three times of that for Switchgrass (12.5±1.8 t ha^{-1}) (Heaton et al., 2008). In China, highest yield of *Miscanthus lutarioriparius* could reach 43.8 t ha^{-1}.

Cellulose, hemicellulose and lignin are the three main components of lignocellulosic feedstocks. The composition of them determines the utilization efficiency. Compared with other PBCs, *Miscanthus* possesses high amount of cellulose (45% - 52%) and hemicellulose (24% - 33%) as well as relatively lower amount of lignin (9% - 13%) (Brosse et al., 2012), which makes it outstanding for producing biofuels. In addition, the low ash features of *Miscanthus* (2.2% compared with corn 5.2%, rice 6.3% and wheat 3.1%) indicated the high calorific value (I. Lewandowskia, 2000, Lewandowski et al., 2003, Brosse et al., 2012).

Strong environmental adaptability

Miscanthus has extensive ecological adaptability and high salt and alkali, heavy metal (Cu, Cd, Pb, Zn, Mn, etc.), drought, heat and cold resistance (Farage et al., 2006). These advantages of *Miscanthus* not only improve land use, but also play a role in protecting the environment (Yan et al., 2012).

Low planting management costs

The process of obtaining biomass including land preparation, planting, weed and pest management, fertilization, harvesting, transport and storage etc., all of which result in production costs. Compared with other energy plants, *Miscanthus* has strong vitality and high resistance, so the management cost of planting *Miscanthus* is relatively low (Fischer et al., 2005, Khanna et al., 2008, Chung and Kim, 2012).

In addition to being a promising PBC, *Miscanthus* has many other uses. For instance, the well-developed root system of *Miscanthus* plays an important role in soil fixation when planting at lake shore and the marginal lands (Xue et al., 2016). The high CO2 fixation efficiency of *Miscanthus* is beneficial to maintaining O2/CO2 in the environment (Heaton et al., 2010). The high quality of *Miscanthus* biomass is also suitable for paper production.

1.2 Lignin in the plant cell wall

The main structural polysaccharides in the plant cell walls are cellulose, hemicellulose and a few pectin (Vorwerk et al., 2004). Lignin is the most abundant component of lignocellulosic biomass except these polysaccharides, it plays an important role in the lignin-saccharide complex to enhance the strength and elasticity of plant cell walls (Foster et al., 2010, Doherty et al., 2011, Neutelings, 2011) (Figure 1-3).

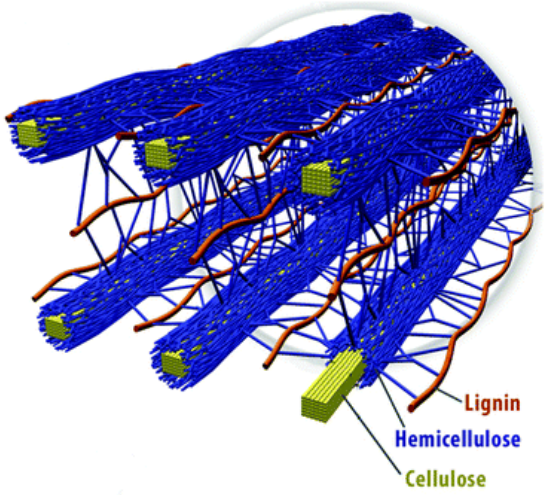

Figure 1-3 Plant cell wall structure simplified diagram.
The main polysaccharides cellulose (shown in light green) and hemicellulose (shown in blue) together with lignin (shown in red) constitute the main component of plant cell walls (Doherty et al., 2011).

On the contrary, the complex structure of biomass and the interaction of the cell wall components will affect the hydrolysis of polysaccharides, which in turn affects the efficiency of lignocellulose conversion (Hendriks and Zeeman, 2009, Alvira et al., 2010, Bhalla et al., 2013). Among them, lignin is considered to be the main physical barrier of the enzyme digestion. It not only hinders the contact of the hydrolase with polysaccharides, and its hydrolysis product is also not conducive to the progress of the fermentation reaction after saccharification (Keating et al., 2006). In view of this, by reducing the lignin content or altering lignin composition to increase its degradability, the utilization value of biomass can be improved (Ragauskas et al., 2014, Yang et al., 2013). A better understanding of lignin formation and monolignol biosynthesis will help to achieve this goal.

1.3 The phenylpropanoid metabolic pathway

Lignin forms by the polymerization of mainly three different monomers, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units. These units derived from three hydroxycinnamyl alcohol monomers *p*-coumaryl (H), coniferyl (G), and sinapyl (S) alcohols (Eudes et al., 2014). The biosynthesis of lignin monomers is derived from the phenylpropanoid metabolic pathway (Vogt, 2010).

Two main categories from the phenylpropanoid metabolic pathway are monolignols and flavonoids (Figure 1-4). Besides, a large group of phenolic compounds in plants are also derived from the phenylpropanoid metabolic pathway. These aromatic metabolites like stilbenes which are not only necessary for terrestrial plants against biotic and abiotic stresses but also have been applied in pharmaceutical preparations (Dixon et al., 2002).

1.3.1 Biosynthesis of lignin

Monolignols are synthesized through phenylpropanoid metabolism, initiated from the shikimate pathway (Douglas, 1996). The amino acid phenylalanine is catalyzed by

phenylalanine ammonia-lyase (PAL), following with cinnamic acid 4-hydroxylase (C4H), and 4-coumarate:CoA ligase (4CL) to *p*-coumaroyl-CoA, the common precursor of several different branches. From this node, *p*-coumaroyl-CoA is either transferred to feruloyl-CoA through the activation by p-hydroxycinnamoyl transferase (HCT), *p*-coumarate 3-hydroxylase (C3H) and caffeoyl-CoA O-methyltransferase (CCoAOMT), and then catalyzed by cinnamoyl-CoA reductase (CCR) that converts feruloyl-CoA to coniferaldehyde. Alternatively, the *p*-coumaroyl-CoA is directly converted to *p*-coumaraldehyde by CCR. catalyzing by ferulate-5-hydroxylase (F5H) and caffeic acid o-methyltransferase (COMT). *p*-coumaraldehyde could also be converted to sinapaldehyde. These aldehydes will then be reduced to alcohols catalyzed by cinnamyl alcohol dehydrogenase (CAD), here resulting in p-coumaroyl alcohol, sinapyl alcohol and coniferyl alcohol, which will become the H, S and G lignin subunits respectively. All these steps are conducted in the cytosol, monolignols will be then transported into cell walls. Finally, under the catalysis of peroxidase (POX) and laccase (LAC), the lignin monomers are polymerized into lignin (Bonawitz and Chapple, 2010). Additionally, in grass, tyrosine has been found recently as precursor for the biosynthesis of S lignin.

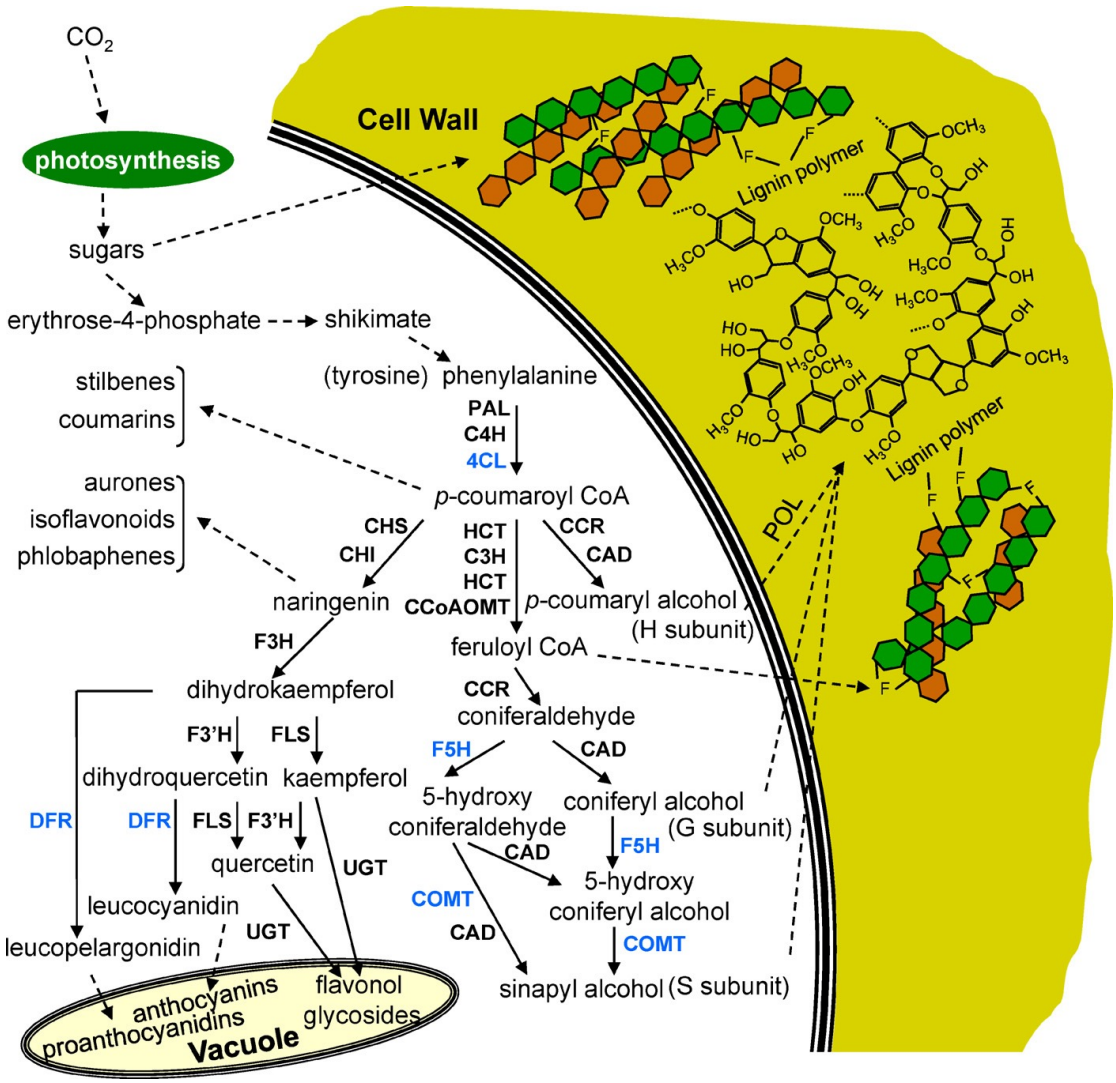

Figure 1-4 Schematic diagram of the phenylpropanoid biosynthesis pathway. (Gray et al., 2012) The main metabolic pathways showing the monolignol and flavonoids biosynthesis initiated from shikimate pathway. Enzymes, final and some of the intermediate products together with their destination are showing in the diagram.

1.3.2 Biosynthesis of flavonoids.

Flavonoids are the products of the other major branch of the phenylpropanoid metabolic pathway, the biosynthesis and accumulation of flavonoids and lignin will be inevitably influenced by each other. Flavonoids are one group of secondary metabolites with diverse biological activities widely found in the plant kingdom. Flavonoids can be divided into several types according to their structural differences, including chalcones, flavones, flavonols, flavandiols, isoflavonoids, anthocyanins, and condensed tannins

(or proanthocyanidins) (Falcone Ferreyra et al., 2012). Flavonoids play an important physiological role in plants. They are known to have control auxin transport (Besseau et al., 2007, Peer and Murphy, 2007), root development (Taylor and Grotewold, 2005), seeds germination (Kubasek et al., 1992), UV-B protection (Li et al., 1993, Höll et al., 2019), signal interaction with commensal microorganisms and plant resistance (Treutter, 2006). At the same time, they have many pharmacological effects such as anti-cancer, anti-oxidation, anti-inflammatory, and reducing blood vessel fragility (Narayana et al., 2001, Agrawal, 2011).

The main steps in the biosynthesis of flavonoids have been firmly established (Heller and Forkmann, 2017). Sharing the same precursor with lignin, flavonoids are derived from *p*-coumaroyl-CoA. Using flavonols synthesis as an example, p-coumaroyl-CoA is successively catalyzed by chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), and flavonol synthase (FLS) that finally results in flavonols. Flavonoid 3′-hydroxylase (F3′H) converts kaempferol to quercetin. Hereafter, anthocyanins and proanthocyanins are synthesized under the action of dihydroflavonol 4-reductase (DFR) (Figure 1-2).

1.4 Transcriptional regulation of the phenylpropanoid pathway

The metabolic products of the phenylpropanoid pathway play critical roles for plant growth and development as well as the adaptation of the environment. The regulation of the expression of related genes along the metabolic pathway is particularly important. Transcription factors recognize specific cis-elements on the promoters to regulate the spatiotemporal expression of genes. For the transcriptional regulation of phenylpropanoid metabolism, there are two categories of transcription factors worth mentioning here.

1.4.1 NAC transcription factors

NAC transcription factors are novel transcription factors with multiple biological functions that are specifically found in plants. There are 151 non-redundant NAC genes in rice and 117 in *Arabidopsis* (Nuruzzaman et al., 2010). The NAC domain was named after the identification of the consensus sequences from *Petunia* NAM and *Arabidopsis* ATAF1/2 and CUC2 proteins. NAC transcription factors possess a highly conserved N terminal with around 150 amino acids, which is the binding domain of the protein, while the variable C-terminal is involved in transcriptional regulation. NAC proteins play important roles in plant development, defense and abiotic stress responses (Olsen et al., 2005).

According to the phylogenetic analysis of 1,232 NAC protein sequences from 11 different organisms, NACs can be classified into eight subfamilies. NACs with the function of regulation of plant cell wall development belong to NAC-c subfamily (Shen et al., 2009)

In *Arabidopsis*, NAC domain transcription factors VASCULAR-RELATED NAC-DOMAIN 6 and 7 (VND6/7) play a crucial role in xylem vessel differentiation(Kubo et al., 2005). Overexpressing VND7 induces the expression of genes of cellulose, hemicellulose and lignin biosynthesis as well as genes related to programmed cell death (Yamaguchi et al., 2010). In addition, NAC SECONDARY WALL THICKENING PROMOTING FACTOR1 (NST1) and SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1 (SND1) function redundantly in the regulation of secondary wall synthesis in fibers (Zhong et al., 2007b).

A set of NAC transcription factors act as master regulators switching on the entire secondary wall biosynthesis in vessels and fibers by activating downstream TFs especially MYB transcription factors. This hierarchical network regulation pattern is highly conserved in vascular plants (Nakano et al., 2015, Zhong and Ye, 2015, Zhong et al., 2015).

1.4.2 MYB transcription factors

MYB transcription factors compromise one of the largest transcription factor families in plants. According to the number of the conserved N-terminal DNA-binding domain repeats, MYBs can be divided into different classes, naming 1R-MYB, R2R3-MYB, 3R-MYB and 4R-MYB respectively (Dubos et al., 2010, Jin and Martin, 1999, Ambawat et al., 2013). Among these classes, the R2R3-MYBs is the most common type in plants. There are 125 R2R3-MYBs in *Arabidopsis* (Stracke et al., 2001), 192 in Populus trichocarpa (Wilkins et al., 2009) and 157 in Maize (Du et al., 2012). The following chapter focuses mainly on the regulatory function of R2R3-MYBs related to phenylpropanoid metabolism.

1.4.2.1 MYB activators

For the activation of lignin biosynthesis, especially in the secondary wall biosynthesis regulation, MYB activators work together with NAC master switches forming a multi-tiered regulatory network (Ohtani and Demura, 2019, Zhong and Ye, 2015) (Figure 1-5). In the transcriptional regulation network, MYB46 acts as a second-layer master switch controlling the secondary wall biosynthesis. In *Arabidopsis*, AtMYB46 is a direct target of SND1 and predominantly expressed in fibers and vessels of the inflorescent stems (Zhong et al., 2007a). AtMYB83 works redundantly with MYB46 in the regulation of secondary cell wall biosynthesis (McCarthy et al., 2009). MYB46 directly activates the transcription of the xylan and lignin biosynthetic genes as well as downstream target TFs such as MYB63/58 (Zhou et al., 2009, Kim et al., 2013, Kim et al., 2014). In poplar, PtrMYB2, PtrMYB3, PtrMYB20, and PtrMYB21 were shown to be MYB46/MYB83 orthologs sharing similar functional roles (Zhong et al., 2013).

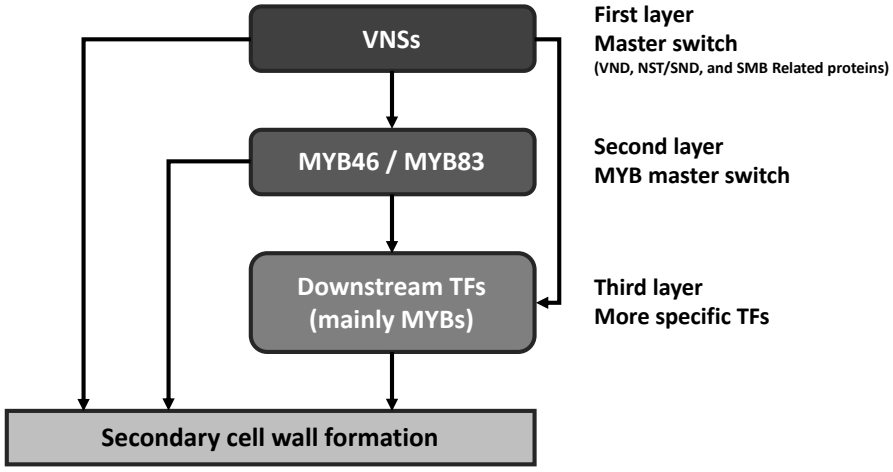

Figure 1-5 NAC-MYB-based multi-tiered regulatory network of secondary cell wall biosynthesis. The network is mainly compromised by three layers of transcription factors. First layer, also known as master switches, is the VND, NST/SND, and SMB-related (VNS) protein subfamily; the MYB master switches form the downstream of the VNS proteins; while the third layer are the MYB proteins that modulate secondary wall formation. All the three layers of transcription factors can activate specific secondary cell wall-related genes directly.

As to the regulation of flavonoids biosynthesis, MYB transcription factors work independently or cooperatively with basic Helix-Loop-Helix factors (bHLHs) and WDR proteins (WDRs), forming an MBW-complex. For instance, in *Arabidopsis*, the three closely related MYBs AtMYB11, AtMYB12 and AtMYB111 can independently activate the expression of early flavonoid biosynthesis related genes (EBGs) such as CHS, CHI, F3H and FLS (Mehrtens et al., 2005, Stracke et al., 2007), while the late flavonoid biosynthetic genes (LBGs) are activated by the MYB-bHLH-WD40 (MBW) ternary transcriptional complex (Nesi et al., 2002, Appelhagen et al., 2011). In poplar, MYB134 and MYB115 are reported to be the activator of proanthocyanidin biosynthesis by regulating key PA pathway genes (Mellway et al., 2009, James et al., 2017, Wang et al., 2017).

1.4.2.2 MYB repressors

In the past two decades, studies on the group of MYB repressors have only increased our knowledge, leaving their functionalities only partially understood. According to phylogenesis, most of the phenylpropanoid metabolism MYB repressors can be divided

into two different clusters, the phenylpropanoid/lignin group and the flavonoid group (Yoshida et al., 2015, Ma and Constabel, 2019). Figure 1-6 summarizes the to date reported MYB repression functions and interactions in the regulation the phenylpropanoid pathway to date.

AtMYB4 is one of the earliest reported R2R3-MYBs related to downregulation of phenylpropanoid pathway. It represses C4H and sinapate ester biosynthesis in *Arabidopsis*. In addition, the expression of AtMYB4 is downregulated by exposure to UV-B light, indicating a relationship to the mechanism for acclimation to UV-B (Jin et al., 2000, Zhao et al., 2007, Höll et al., 2019). In maize, the ortholog of AtMYB4, ZmMYB31 and ZmMYB42 have complementary roles on regulating lignin and phenylpropanoid metabolism and also affect cell wall structure and lignin composition (Sonbol et al., 2009, Fornale et al., 2010). In woody plants, PtoMYB156 is negatively regulating phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation (Yang et al., 2017a), while PdMYB221 represses secondary wall formation through a set of direct and indirect suppression to the related genes (Tang et al., 2015).

Many of the R2R3-MYBs are reported to have repressing function on flavonoids biosynthesis. AtMYB7 is closely related to AtMYB4 and represses several genes along the flavonoid pathway. Moreover, AtMYB7 itself is repressed by AtMYB4, proposing a reciprocal effect of the regulatory mechanism in production of UV-protecting sunscreens in *Arabidopsis* (Fornalé et al., 2014, Jin et al., 2000). Cavallini et al. characterized a set of R2R3-MYB C2 repressors in grapevine, where VvMYB4a and VvMYB4b regulate mainly the synthesis of small phenolic compounds, while MYBC2-L1 and MYBC2-L3 suppress anthocyanins and proanthocyanidins accumulation (Cavallini et al., 2015). Poplar MYB165, MYB182 and MYB194 interact with bHLH proteins and have shown repression of PA biosynthesis genes. They can also repress the activation function of MYB134, an regulator of PA biosynthesis (Yoshida et al., 2015, Ma et al., 2018). It is also reported that in peach, a negative regulator of anthocyanin and PAs, PpMYB18, can be activated by both anthocyanin and PAs related

MYB activators to form a feedback loop for balancing secondary metabolite accumulation (Zhou et al., 2019).

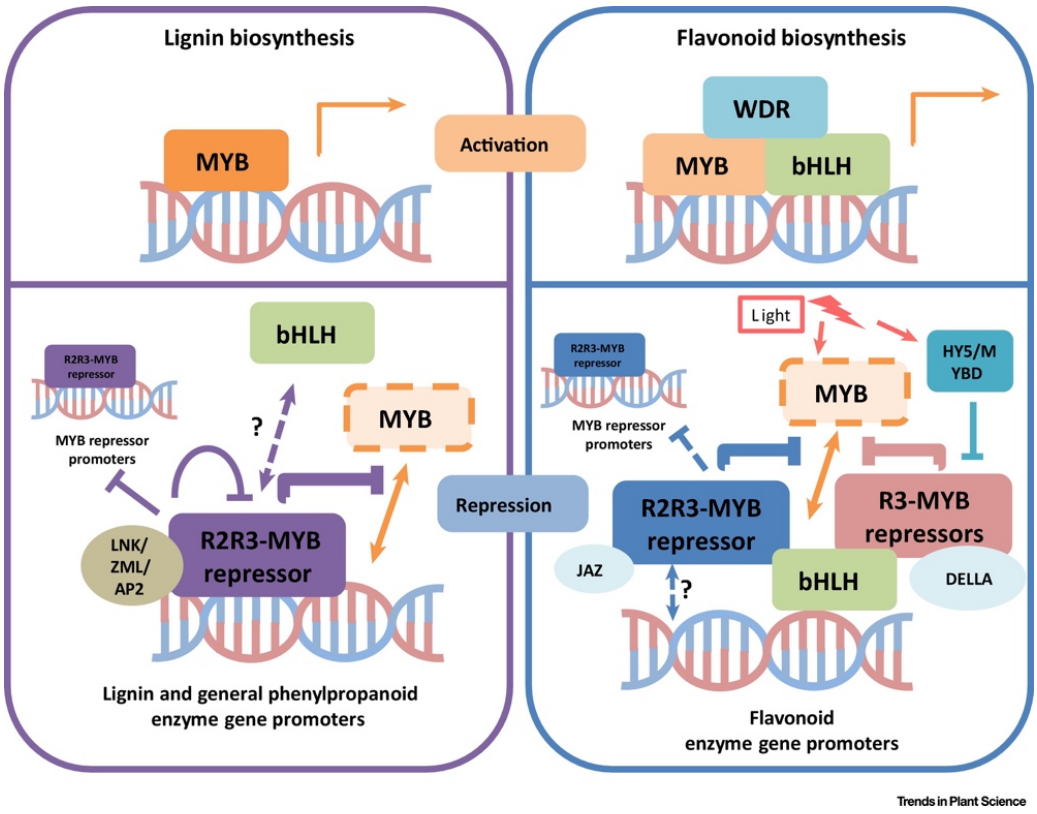

Figure 1-6 Summary of MYB repressor functions and interactions for lignin and flavonoid repressors. (Ma and Constabel, 2019)

Left part showing in purple is the regulation of lignin biosynthesis and general phenylpropanoid pathway. MYB activators recognize and bind to the AC elements on the promoters and activate gene expression (left top). Lignin MYB repressors compete with MYB activators binding to the same AC element to prevent transcription. There are co-factors that can enhance MYB repressor activity. Some of the MYB repressors have also been shown to have self-regulation functions or can inhibit the expression of other MYB repressors. Right part showing in blue indicates the regulation of flavonoid biosynthesis. For flavonoids, an MYB-bHLH-WDR (MBW) complex is required to activate gene expression. R2R3 and R3 MYB repressors interact with bHLH cofactors and disrupt the interaction of the MYB activator with the bHLH. Broken lines indicate the hypothesized interactions.

1.4.2.3 AC elements in promoters

In the promoters of many phenylpropanoid pathway genes, a type of specific 7bp sequence can be recognized and directly bound by MYB proteins thus influencing transcription. These motifs were previously named as AC elements (Lois et al., 1989, Hatton et al., 1995, Raes et al., 2003). Furthermore, in the transcriptional regulation of

secondary wall formation, these AC elements were described to have certain patterns, ACC(T/A)A(A/C)(C/T), and been named as secondary wall MYB responsive element (SMRE) in the study (Zhong and Ye, 2012). Figure 1-7 summarized the patterns of all the eight SMREs.

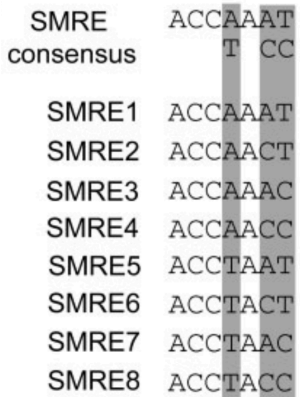

Figure 1-7 Summary of the SMREs
The eight variants based all on the possible principle ACC(T/A)A(A/C)(C/T).

1.5 Lignification modification - transgenic strategies

1.5.1 Attempts on engineering of lignin biosynthesis genes and regulators

With the growing understanding of the lignocellulosic biomass composition and cell wall formation, attempts on engineering lignin to improve the digestibility of the lignocellulosic biomass are widely applied. Transgenic downregulation of major lignin biosynthesis genes can reduce lignin content and increase the digestibility of dry matter. Extensive research has been done on the genetic engineering of various plants species to modify individual gene expression that altered lignin content and composition, thus improved the digestibility of the lignocellulosic biomass. For example, Alfalfa is one of the perennial grasses that has been developed as energy crops in the early years. In Alfalfa, Chen et al. could show independent inhibition of several enzymes in the monolignol pathway independently (C4H, HCT, C3H, CCoAOMT, F5H or COMT), resulting in decreased overall carbon flux into lignin and increased enzymatic hydrolysis efficiency in all transgenic lines. Enzymes at higher upstream positions of

the pathway had more significant effect on lowering lignin content than the downstream genes, which mainly influenced the lignin composition (S/G ratio etc.). Table 1-2 shows an overview of different examples of gene modification on various plants in recent years.

Table 1-2 Summary of attempts on transgenic downregulation of major lignin biosynthesis genes

Gene	Species	Lignin content	Lignin composition	Digestibility	References
PAL	Brachypodium	Reduced	S/G ratio ↑	Increased	(Cass et al., 2015)
C4H	Alfalfa	Reduced	S/G ratio ↓	Increased	(Chen and Dixon, 2007)
4CL	populus	Reduced	S/G ratio ↓	Unchanged	(Voelker et al., 2010)
	Switchgrass	Reduced	S/G ratio ↓	Increased	(Xu et al., 2011)
HCT	Alfalfa	Reduced	High H	Increased	(Shadle et al., 2007)
C3H	Alfalfa	Reduced	S/G ratio ↑	Increased	(Chen and Dixon, 2007)
CCR	Populus	Reduced	Unchanged	Increased	(Van Acker et al., 2014)
CCoAOMT	Alfalfa	Reduced	S/G ratio ↑	Increased	(Chen and Dixon, 2007)
F5H	Alfalfa	Unchanged	S/G ratio ↓	Unchanged	(Reddy et al., 2005)
COMT	Switchgrass	Reduced	S/G ratio ↓	Increased	(Baxter et al., 2014)
	Alfalfa	Reduced	S/G ratio ↓, 5-OH-G ↑	Increased	(Jung et al., 2013)
	Maize	Reduced	S/G ratio ↓, 5-OH-G ↑	Increased	(He et al., 2003)
CAD	Switchgrass	Reduced	S/G ratio ↓	Increased	(Fu et al., 2011)
	Tall fescue	Reduced	S/G ratio ↓	Increased	(Chen et al., 2003)
LAC	Brachypodium	Reduced	S/G ratio ↑	Increased	(Wang et al., 2015)

In general, reducing lignin content or altering lignin composition can improve the efficiency in the biomass pretreatment, it is worth mentioning that down-regulation of the expression of biosynthesis genes by knocking down/out methods does reduce the amount of lignin, however, this may be accompanied by a decrease in the overall biomass content of the plant, which limits the growth and development of plants themselves. Therefore, genetic engineering attempts gradually focused on more modest ways to improve the quality of lignocellulosic biomass.

The modification of transcription factors related to phenylpropanoid pathway provides other possibilities for improving lignocellulose biomass quality. For example, ZmMYB31 directly interact with ZmCOMT and ZmF5H promoters in vivo, and ZmMYB42 generally represses most of phenylpropanoid pathway genes. Transgenic studies in *Arabidopsis* of both the two transcription factors indicated enhanced cell wall degradability would give new chances for biotechnological manipulations of lignocellulosic biomass (Sonbol et al., 2009, Fornale et al., 2010). Additionally, the conservation or divergence of the syntelogs of MYB31 and MYB42 in maize, rice and sorghum was analyzed (Agarwal et al., 2016), revealing the functional similarities and specialized regulatory activities across the three grasses, which provide great references for understanding flux of phenylpropanoid metabolites and improving lignin engineering strategies.

1.5.2 Achievements of the research group

On the research of cell wall formation, especially the lignin biosynthesis and the transcriptional regulation of that in *Miscanthus sinensis*, transcription factors related to SECONDARY WALL-ASSOCIATED NAC DOMAIN1 (SND1) and SECONDARY CELL WALL MYBs 1-4 (SCM1-4) were identified in the *Miscanthus* transcriptome. MsSND1 acts as a master switch for the regulation of secondary cell wall formation and lignin biosynthesis. During *Miscanthus* development, expression of MsSND1 and MsSCMs coincided with the onset of secondary cell wall formation and lignification of vascular tissue and sclerenchyma fibers (Golfier et al., 2017). In addition, *Miscanthus*

laccase MsLAC1 is regulated by secondary cell wall MYB transcription factors and is involved in lignification of xylem fibers (He et al., 2019). These results have suggested promising breeding targets in *Miscanthus* for biofuel and biomaterial applications.

2 Aims

Lignocellulosic biomass has great potential for renewable biofuel and biomaterial production. Current research on *Miscanthus* has shown that it turned to be one of the most promising biomass crops nowadays. To understand more about the formation and regulation of secondary cell wall biosynthesis can help us make better use of *Miscanthus* as a sustainable resource. Previous studies in the Rausch research group mainly focused on identification and characterization of laccaces and transcriptional activators of secondary cell wall formation. In order to provide a more comprehensive understanding of the regulatory network, this thesis aimed at identification and functional study of two R2R3-MYB transcription repressors, MsMYB31 and MsMYB42.

- Firstly, being structurally related to AtMYB4, the expression patterns of the two distinct genes MsMYB31 and MsMYB42 should be evaluated to elucidate whether their functions are different.
- Secondly, the conserved domains of MsMYB31 and MsMYB42 proteins suggested their possible DNA binding functions. To confirm this hypothesis, experiments such as Dual-luciferase-assay and Electrophoretic Mobility Shift Assay (EMSA) were performed to test the repression function and DNA binding capacities of the MsMYB31 and MsMYB42 proteins. In addition, related experiments also suggested the competition between transcriptional activator MsSCM4 and repressors MsMYB31 or MsMYB42 at the same binding site.
- Lastly, to understand the effects of the two repressors *in vivo*, the model plant *Arabidopsis* was used to perform inducible overexpression and complementation experiments with MsMYB31 and MsMYB42.

3 Results

Lignin biosynthesis, together with the entire phenylpropanoid metabolic pathway in *Miscanthus* has been studied widely in recent years. Thus, it was of fundamental importance to look deep into the mechanisms of the regulatory network of phenylpropanoid metabolic pathway in *Miscanthus*. Here, repressive transcription factors play a crucial role in fine-tuning the lignification process. This work focuses on the identification and characterization of two R2R3-MYB repressors in *Miscanthus*, contributing to the understanding of lignification as well as facilitating breeding efforts towards tailored biomass to meet the requirements for sustainable biomass generation of known lignin composition.

3.1 The identification of two R2R3-MYB repressors in *Miscanthus sinensis*

3.1.1 Identification of two R2R3-MYB repressors in *Miscanthus* EST database

Using known sequences of the R2R3-MYB transcription repressors regulating the phenylpropanoid pathway in plants (AtMYB4, ZmMYB31 and ZmMYB42) as queries against the *Miscanthus* EST database, a group of contigs was identified showing high similarity to the queries. To narrow the scope of the candidates, the protein sequence of the conserved C2 motif (13aa) was used as query to do blast again in the *Miscanthus* EST database. Taking both the alignment results together, 7 candidate contigs were selected for amplifying sequences from *Miscanthus* cDNA.

Table 3-1-1 Seven candidate contigs selected from the *Miscanthus* EST database.

Contig name	Length (bp)
Sacchariflorus_TContig782	828
Goliath_TContig1672	719
Goliath_TContig30389	722
Sacchariflorus_TContig37581	2788
Sacchariflorus_TContig34586	2811
Undine_TContig8175	898
AmurSilvergrass_TContig7892	932

Young *Miscanthus* seedlings (~3 weeks) grown in the greenhouse were used for cDNA extraction. Two *Miscanthus* R2R3-MYB protein sequences were identified and named MsMYB31 and MsMYB42 respectively. Full-length CDS of MsMYB31 encodes a protein of 258 amino acids, sharing 79% identity with ZmMYB31 (95% identity for R2R3 domain only), while MsMYB42 contains 259 amino acids and shares 80% identity with ZmMYB42 (91% identity for R2R3 domain only). In the N-terminal, both MsMYB31 and MsMYB42 contain a predicted nuclear localization signal peptide, which is consistent with the potential transcription regulatory functions. In addition, there are also a conserved bHLH binding motif ([DE]Lx2[RK]x3Lx6Lx3R), indicating the potential MYB/bHLH interactions. Within the C-terminal of MsMYB31 and MsMYB42, the conserved C1, C2, C3 and C4 motifs suggest that both belong to subgroup 4 family, which functionally known as repressors.

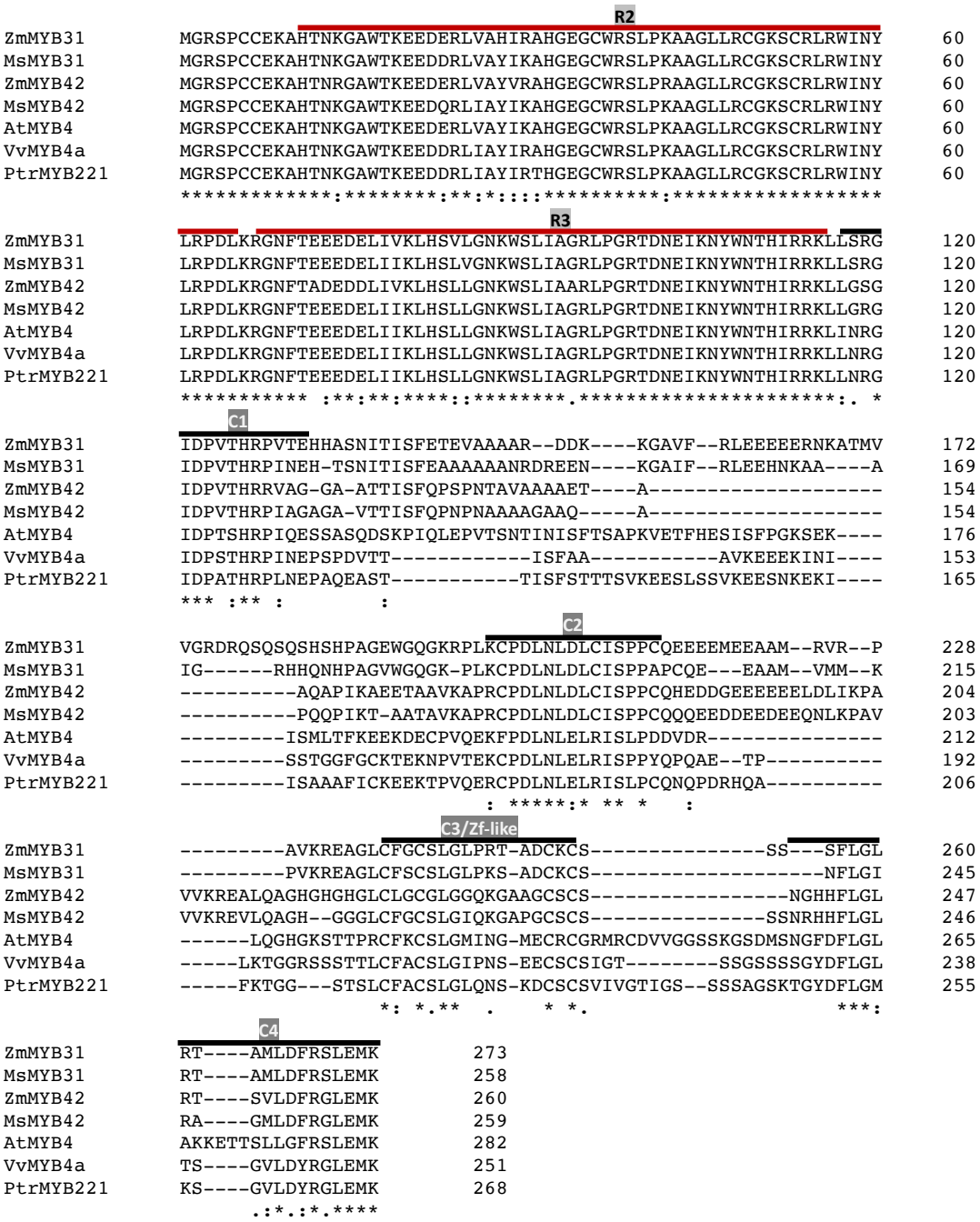

Figure 3-1-1 Multiple sequence alignment using Clustal Omega program.

Full length amino acid sequences of MYB31 and MYB42 from *Miscanthus* and *Zea mays* and their homologous genes in *Arabidopsis* thaliana, Vitis vinifera and Populus trichocarpa were analyzed by Clustal Omega. R2 and R3 MYB domains are indicated in red lines with black names, conserved C1, C2, C3 and C4 motifs are shown in black lines with white names.

In monocots, such as rice, maize, sorghum together with *Miscanthus*, the homologous genes of AtMYB4 can be divided into 2 separate groups, however, the similarities between members of the group are extremely high. Despite the almost identical R2R3

domain and conserved motifs, MYB31 and MYB42 sequences display specific parts (Figure S1). Phylogenetic analysis also showed that MsMYB31 and MsMYB42 are closely related to their own particular groups in monocots, while the R2R3-MYB repressors in dicot have no significant separation into different groups (Figure 3-1-2).

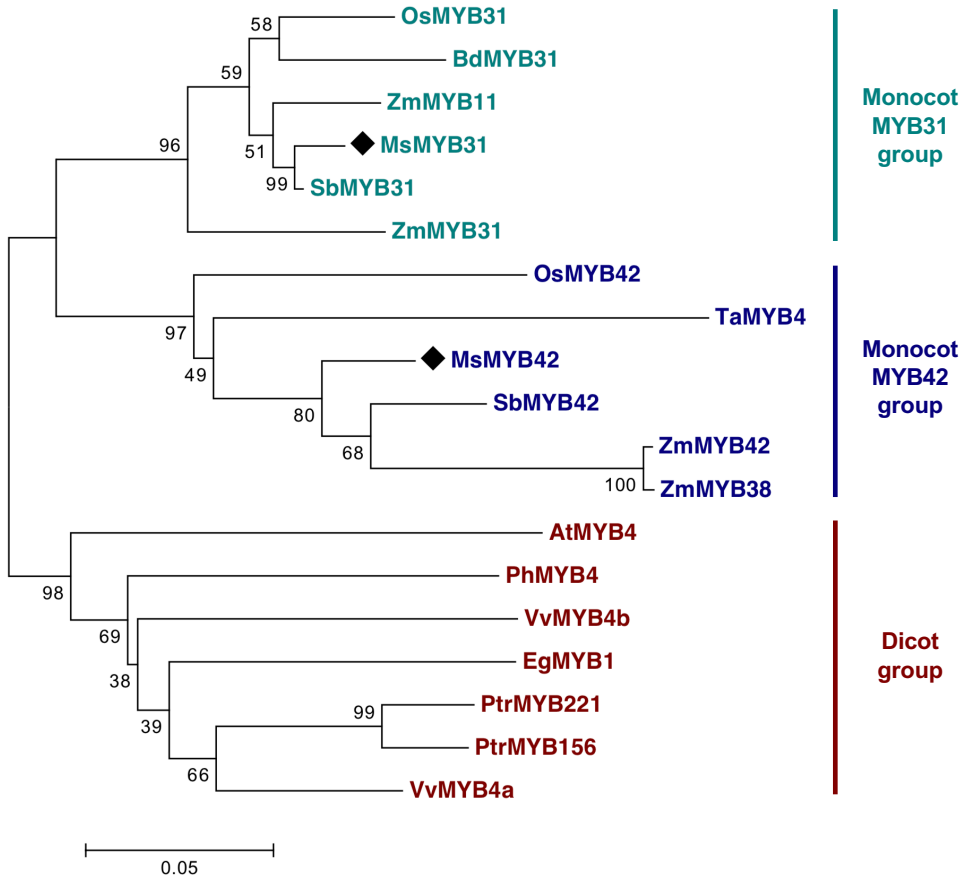

Figure 3-1-2 Neighbor-joining phylogenetic tree of MsMYB31, MsMYB42 and their homologous genes in different monocots and dicots.

In total 19 amino acid sequences were analyzed to construct the Neighbor-Joining tree (bootstrap=1000) using the MEGA7.0 software. The accession numbers of the other MYB proteins are *Zea mays* ZmMYB11 (AIB05021), ZmMYB31 (NP_001105949), ZmMYB38 (AIB04526), ZmMYB42 (NP_001106009); *Oryza sativa* OsMYB31 (XP_015612022), OsMYB42 (XP_015650911); S*orghum bicolor* SbMYB31 (XP_002462743), SbMYB42 (EES14116); *Triticum aestivum* TaMYB4 (AEG64799); *Brachypodium distachyon* BdMYB31 (XP_003578527); *Arabidopsis thaliana* AtMYB4 (AAP13410); *Vitis vinifera* VvMYB4a (ABL61515), VvMYB4b (ACN94269); *Populus trichocarpa* PtrMYB156 (AOF43278), PtrMYB221 (AOF43273); *Petunia x hybrida* PhMYB4 (ADX33331); *Eucalyptus gunnii* EgMYB1 (CAE09058).

3.1.2 Re-blast in *Miscanthus* genome database, members in subgroup 4 family

On December 20th, 2017, the first chromosome-scale assembly of *Miscanthus sinensis* doubled haploid DH1 (IGR-2011-001) was released (***Miscanthus sinensis*** v7.1 DOE-JGI, http://phytozome.jgi.doe.gov/). MsMYB31 and MsMYB42 CDS sequences were used as queries for alignments. The gene with transcript name Misin03G234000 shared 100% identity with the cloned MsMYB31 and Misin07G248300 was proposed to be MsMYB42 (99% identity). To do further sequence clustering analysis, I repeated blast search using the protein sequences of Misin03G234000 and Misin07G248300 as queries respectively. For both of them, the top 30 protein homologous sequences in the *Miscanthus* genome database were downloaded. After the two data sets of sequences were merged, in total 34 sequences were obtained. Among all 34 sequences, 14 can be clustered into the R2R3-MYB subgroup 4 family, where most of the members have been shown to act as repressors of the phenylpropanoid pathway. Another 15 sequences have closer relationship with R2R3-MYB subgroup 10/24 family, which containing members AtMYB9, AtMYB93 and AtMYB107. The remaining 5 sequences were found to be relatively far from these clusters (Figure S2).

Interestingly, when doing Neighbor-joining phylogenetic tree analysis, both monocot MYB31 and MYB42 groups contained 2 sequences with great similarity (Misin03G234000 and Misin04G262100 in MYB31 group; Misin07G248300 and Misin13G155400 in MYB42 group). This paralog phenomenon also appeared in corn (ZmMYB11 and ZmMYB31; ZmMYB38 and ZmMYB42).

3.2 MsMYB31 and MsMYB42 have distinct expression patterns in *Miscanthus*

3.2.1 The expression patterns of MsMYB31 and MsMYB42 along *Miscanthus* leaf gradient axis

In monocots, the differentiation of a single leaf from sheath bottom toward blade tip indicates the developmental gradient along the leaf. The expression profile of MsMYB31 and MsMYB42 along the leaf developmental gradient was determined in *Miscanthus* leaf, divided in 8 different segments from the leaf base to the leaf tip. The transcript level of MsMYB31 was highest at the leaf base part and decreased rapidly along the gradient to the leaf tip, whereas MsMYB42 showed the lowest transcript level in the first leaf base segment and remained higher in the rest of the segments. The diametrically opposite patterns indicated different functions at different growth stages for the two transcription factors (Figure 3-2-1A). Previous study using the same gradient tool showed that a NAC transcription factor in *Miscanthus MsSND1* expressed in growth stages associated with secondary cell wall formation, together with its potential targets (Golfier,2017) (Figure S4). Interestingly, MsMYB31 also shared similar expression pattern with MsSND1's potential targets, while MsMYB42 revealed an opposite pattern.

Further, the expression profiles of the putative orthologues from maize and rice were also checked and compared with MsMYB31 and MsMYB42 by converting the absolute value into heat maps (Figure 3-2-1B). In all three crops, MYB31 and MYB42 behaved quite similar. Possibly due to the closer genetic relationship, MYB31 and MYB42 patterns in *Miscanthus* are more similar to those in Maize.

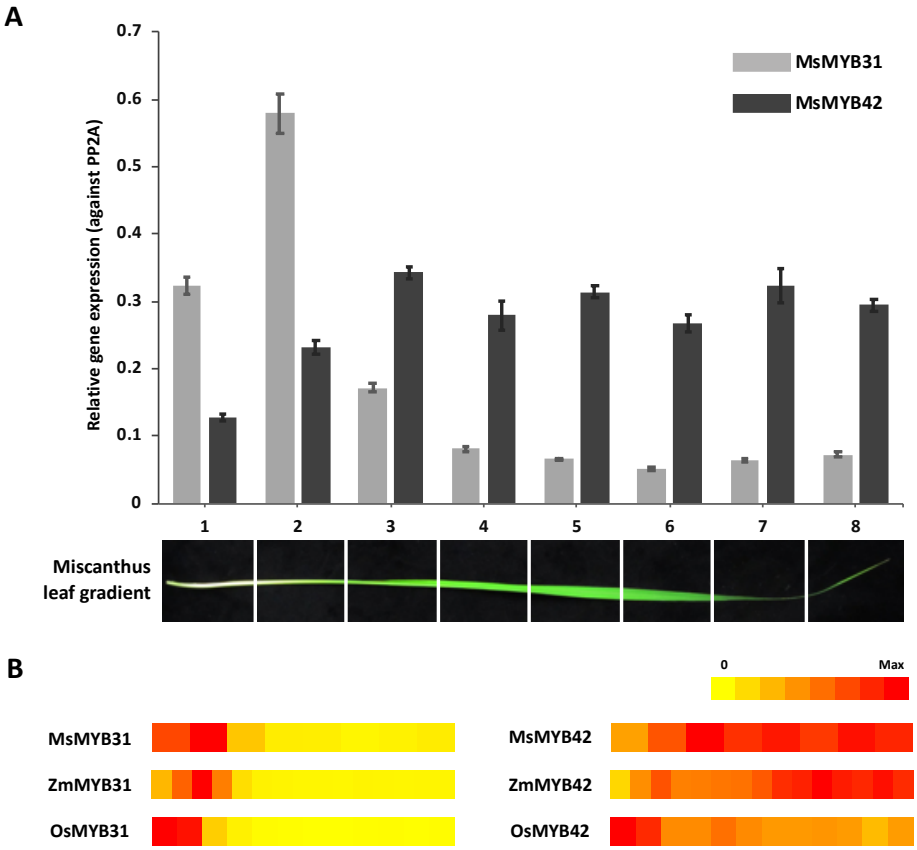

Figure 3-2-1 Expression patterns of MsMYB31 and MsMYB42 along leaf gradient samples and comparison with other homologs in monocots.

(A) Expression patterns of MsMYB31 and MsMYB42 in eight leaf gradient samples. *Miscanthus* leaves were divided into eight parts, from immature leaf base (1) to mature leaf tip (8). PP2A was used as reference gene. The expression patterns of MYB31 and MYB42 in *Miscanthus*, Maize and rice have been visualized in heat maps (B), in each independent gradient map, light yellow represent zero expression and red represent max expression of individual genes. Absolute values obtained from the database were shown in Table S1.

3.2.2 Flavonols accumulate along *Miscanthus* leaf axis

Flavonol contents were also checked in *Miscanthus* leaf gradient samples. For the determination of flavonols, leaf blades samples divided in 5 segments were prepared. Using thin layer chromatography, the flavonols extracted from leaf gradient samples could be visualized on the glass board. Figure 3-2-2 shows the comparison of flavonol contents in different samples. Firstly, flavonol contents vary quite much between different species, resulting in completely different patterns on the HPTLC separation. Moreover, in both *Miscanthus* and Sorghum leaf gradient samples, the total content accumulate from young part to mature tip part. This pattern correlates with the position

and the function of different leaf structures. For example, the accumulation of flavonols in older tissues increase the resistance of plant towards environmental stresses like UV (Winkel-Shirley, 2002). Additionally, in order to check if the same species under different environmental conditions will result in different flavonols accumulation, leaf blade samples taken from *Miscanthus* growing in different places also have been eveluated (Figure S3). There are no significant differences in compounds between samples from new plants (A) and new branches from old tillers (B); and the growing condition also did not significantly change the flavonols composition (B and C).

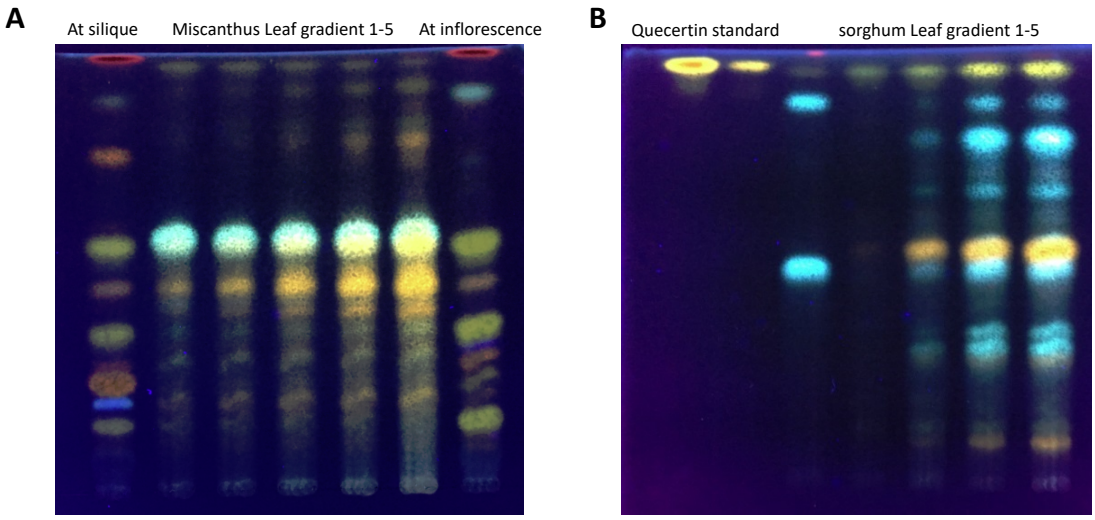

Figure 3-2-2 Thin layer chromatography pictures showing the flavonols extracted from (A) *Arabidopsis* siliques, inflorescence, *Miscanthus* leaf gradient samples and (B) quecertin standard and Sorghum leaf gradient samples.

Miscanthus and sorghum leaf samples were collected from around 2-month-old plants growing in the greenhouse from seeds. The leaf blade (around 60 cm) was divided equally into 5 parts, 1-5 means from base to tip. Each sample is a mixture of 3 leaf blades from 3 independent shoot. *Arabidopsis* silique and inflorescence samples were collected from the Col-0 WT plants.

3.3 MsMYB31 and MsMYB42 express in nucleus and regulate genes involved in lignification

3.3.1 Subcellular localization of MsMYB31 and MsMYB42

Both the open reading frames (ORFs) of MsMYB31 and MsMYB42 contain a predicted bipartite nuclear localization signal (NLS). To confirm this prediction, constructs containing fusion protein 35S:MsMYB31-GFP, 35S:MsMYB42-GFP, and 35S:GFP as control were transiently transformed into tobacco leaves via leaf infiltration. As shown in Figure 3-3-1, MsMYB31 and MsMYB42 located in nuclear-shape circles while leaves with only 35S:GFP construct had GFP signals randomly distributed everywhere. All the observation indicated that the fusion protein 35S:MsMYB31-GFP, and 35S:MsMYB42-GFP were localized in the nucleus. This is where most transcription factors are located.

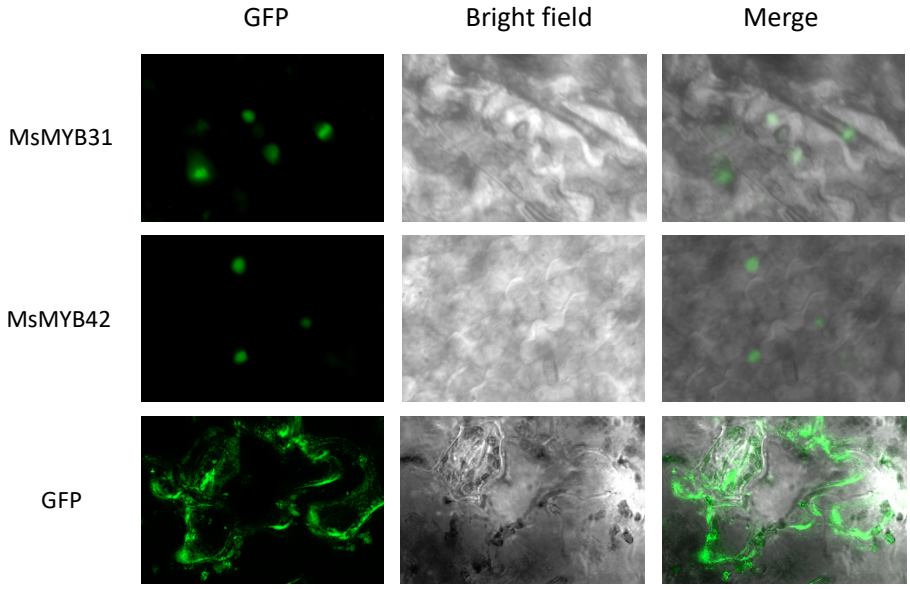

Figure 3-3-1. Subcellular localization of MsMYB31 and MsMYB42.
Transient transformation of 35S:MsMYB31-GFP, 35S:MsMYB42-GFP, and 35S:GFP in tobacco leaves via leaf infiltration. Transient transformation with DAPI staining of Sorghum protoplasts see supplement figure 13.

3.3.2 MsMYB31 and MsMYB42 repress *Miscanthus* C4H, CCR and CAD gene promoters in vivo

MYB transcription factors can bind to specific elements on the promoters to regulate the transcription of the genes. In this study, according to previous findings, we amplified several promoter fragments of genes related to the phenylpropanoid pathway. Previous study showed that in *Arabidopsis*, there were several potential MYB binding sites on the promoter of AtCHS. Interestingly, the *Miscanthus* R2R3-MYB transcription activator MsSCM4 was known to have interaction with AtCHS promoter (Golfier, 2018). So, here the promoter sequence of AtCHS was used as a positive control. Transient dual-luciferase reporter assays were conducted in Chardonnay grapevine suspension cell cultures (Höll et al., 2013).

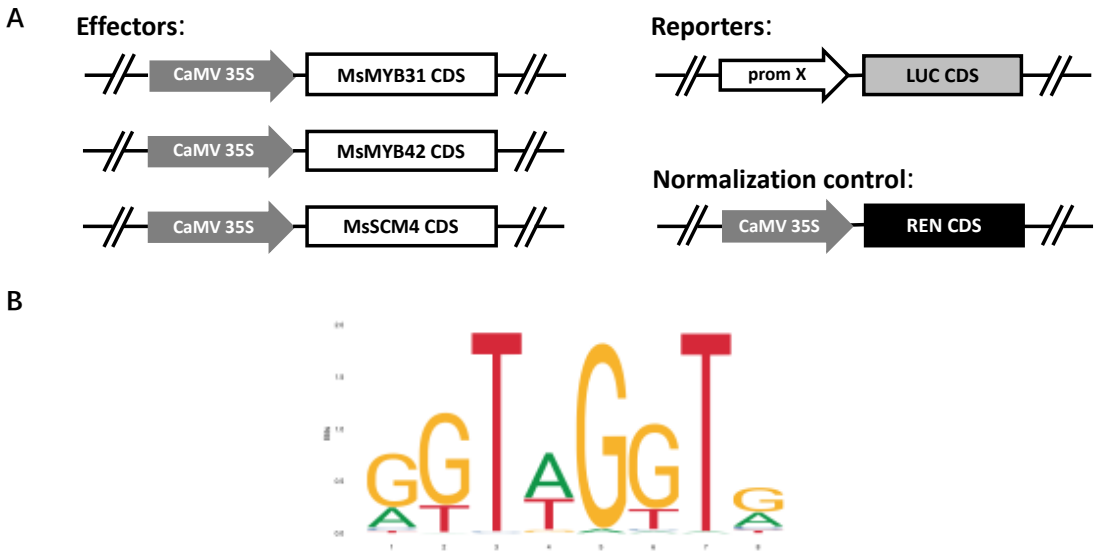

Figure 3-3-2. Introduction of constructs used in dual-luciferase-assay and the predicted binding motif of MsMYB31 and MsMYB42.

(A) Diagrams of effector, reporter and the normalization control constructs. CDS of *Miscanthus* MYB31, MYB42 or SCM4 protein were driven by CaMV 35S promoter as effectors. Promoter sequences shown in table. X fused with a firefly luciferase CDS worked as reporters. The construct including a CaMV 35S promoter linked with a renilla luciferase was introduced in all the bombardments as a normalization control. (B) The predicted binding motif (AC-element, reverse complement).

The constructs used in the dual-luciferase reporter assays were shown in figure 3-3-2A. All the transcription factors were driven by CaMV 35s promoter as effectors. Predicted

target promoters were linked to a firefly luciferase as reporters. A CaMV 35s promoter driven renilla luciferase was introduced in all reactions as a normalization control. Figure 3-3-2B shows the predicted binding motif (AC-elements) of MsMYB31, MsMYB42 as well as AtMYB4 on the online high-quality transcription factor binding profile database (http://jaspar.genereg.net). According to the Firefly/Renilla luciferase ratios, both MsMYB31 and MsMYB42 had a repression function on the AtCHS promoter, and MsMYB42 performed a stronger effect than MsMYB31 (Figure 3-3-3A).

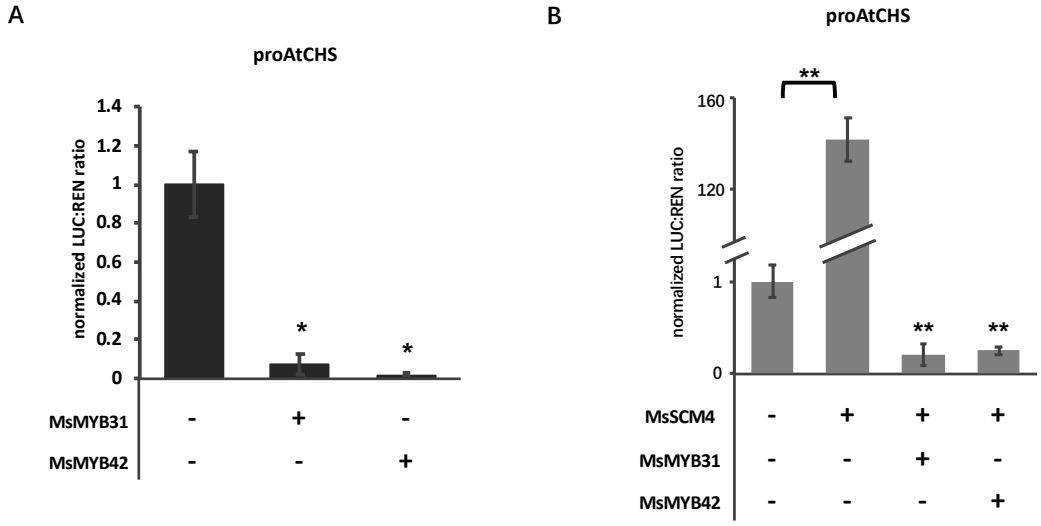

Figure 3-3-3. Repression and de-activation of *Arabidopsis* CHS by MsMYB31 and MsMYB42.
The height of the bars indicated the relative promoter activity demonstrated by LUC:REN ratio. The activity of promoter only was normalized to one. Comparing between bars within one histogram showed the fold changes of the AtCHS promoter activity thus reflecting the repression function (A) or the de-activation abilities (B) of MsMYB31 and MsMYB42. Error bars were mean ± SD of at least three biological replicates. Student's t-test was used to determine the significances: *, P<0.05; **, P<0.01.

Table 3-3-1. Predicted binding sites on the promoter sequences used in dual-luciferase reporter assay.
Potential binding sites were predicted using the online high-quality transcription factor binding profile database http://jaspar.genereg.net

Promoter	Length	Number of predicted binding site (threshold 85%)
pAtCHS	1500 bp	13
pMsC4H	1294 bp	9
pMsCCR	1218 bp	10
pMsCAD	1257 bp	8
pMsCHS	788 bp	3

According to the DLA results, comparing with promoter only, the introduction of MsMYB31 or MsMYB42 into the system reduced the promoter activity of C4H, both were significant in statistics. As to CCR, MsMYB31 and MsMYB42 had even stronger reduction to the promoter activity. However, MsMYB31 and MsMYB42 has less repression effect on CAD promoter, with only MsMYB31 statistically significant. Last, there was no effect on CHS promoter for MsMYB31 or MsMYB42 in the experiment (Figure 3-3-4). Overall, MsMYB31 and MsMYB42 had inhibitory effects on the transcriptional regulation of genes containing certain binding sites on their promoter region. The different performances of MsMYB31 and MsMYB42 towards CHS in different species might because of the limitation of the successfully spliced and cloned promoter length (1500 bp vs 788 bp), or because of the multiple isoforms of CHS in monocots.

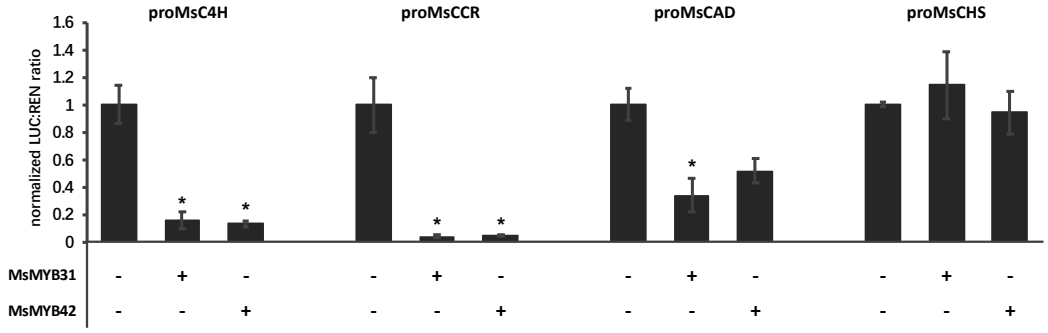

Figure 3-3-4. Repression of *Miscanthus* phenylpropanoid pathway related genes by MsMYB31 and MsMYB42.

All the 4 *Miscanthus* gene promoters tested in the analysis were shown in table 3-3-1. Error bars were mean ± SD of at least three biological replicates. Student's t-test was used to determine the significances: *, $P<0.05$.

3.3.3 MsMYB31 and MsMYB42 can block the activating function of MsSCM4

In the transcriptional regulation of genes, some of the MYB transcription activators and repressors were reported to recognize very similar binding motifs, which were named AC-elements, thus influenced the activation or repression of the genes. To test whether the activator MsSCM4 also has similar relationship with MsMYB31 or MsMYB42, DLA with MsSCM4 and MsMYB31 or MsMYB42 together was firstly tested on the

AtCHS promoter. As shown in Fig 3-3-3B, MsSCM4 indeed significantly activated AtCHS, and the activation function was strongly repressed by adding MsMYB31 or MsMYB42 together with MsSCM4, the LUC/REN ratio was even lower than promoter only.

For *Miscanthus* promoters, the activator MsSCM4 could significantly activate all the promoters tested, including MsCHS. When MsSCM4 was co-expressed with either MsMYB31 or MsMYB42, a clear deactivation effect was found in all 4 promoters (Figure 3-3-5). In addition, MsMYB31 and MsMYB42 working together had no additive effect toward the promoter activity (Figure S5).

In this regard, one possible explanation for the overall activation could be that MYB repressors and activators have preferences towards different types of AC-elements, and the binding of the activator might be affected by the presence of repressors, which caused the deactivation.

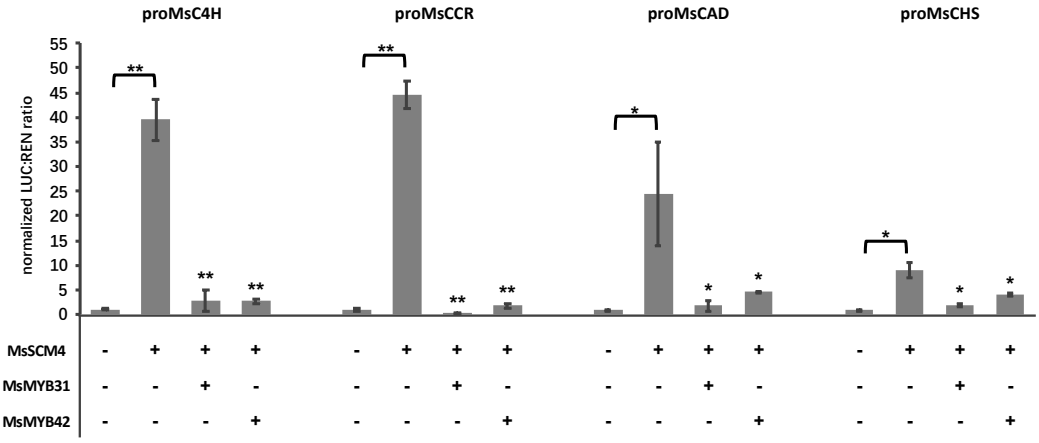

Figure 3-3-5. The deactivation function of MsMYB31 and MsMYB42 on MsSCM4 towards *Miscanthus* phenylpropanoid pathway related genes.

All the 4 *Miscanthus* gene promoters tested in the analysis were shown in table. X. Error bars were mean ± SD of at least three biological replicates. Student's t-test was used to determine the significances: *, $P<0.05$. **, $P<0.01$

3.4 *Miscanthus* MYB transcription factors bind to AC-elements on the promoters of phenylpropanoid metabolite pathway genes

3.4.1 Expression and purification of the MYB transcription factor proteins

To further determine the DNA-binding capacity of MsMYB31 and MsMYB42 protein, The CDS sequences of MsMYB31, MsMYB42 and also the activator MsSCM4 were cloned into the expression vectors pETG10A(N-His-tag) or pETG60A(N-NusA+His-tag). The recombinant proteins were expressed in E. coli strain BL21 by 4-hour 0.75M IPTG induction. The cultures were then collected and sonicated, after centrifuging the supernatants and pellets were collected separately for SDS-PAGE.

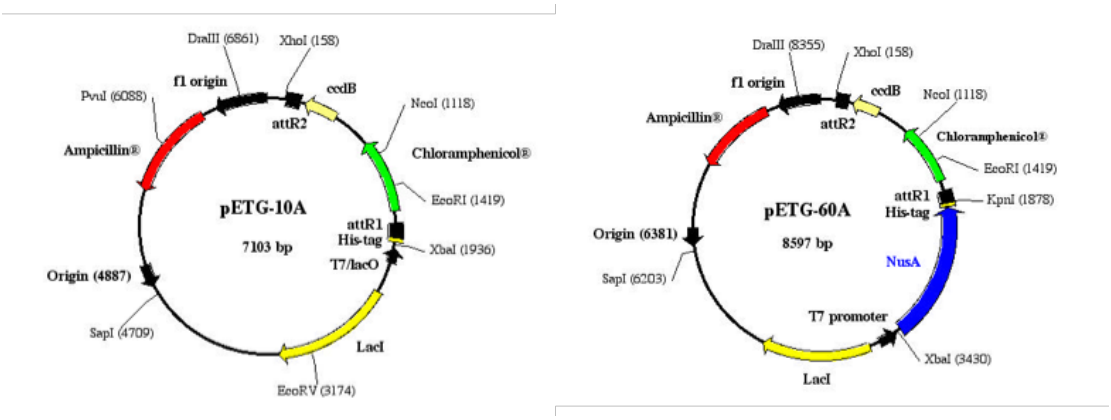

Figure 3-4-1 Vector maps of pETG-10a and pETG-60a showing the N-terminal tags.

According to the SDS-PAGE gel (Figure 3-4-2), comparing with the un-induced samples, there was successful protein expression after induction for all 3 proteins. In pETG10a vector, MsMYB42 and MsSCM4 protein were detected in both supernatant and pellet. As for MsMYB31, comparing with others, it possessed the highest yield within the same induction time, while the protein was only abundant in the pellet. So, the pETG60a vector, which contains a N-NusA-6His tag, was used to improve the protein solubility. As we could see in the SDS-PAGE gel, the N-NusA tag significantly

increased the heterologous MsMYB31 protein solubility and kept a fairly high yield, following with a great increase of protein molecular weight (predicted MW: 86.11kDa).

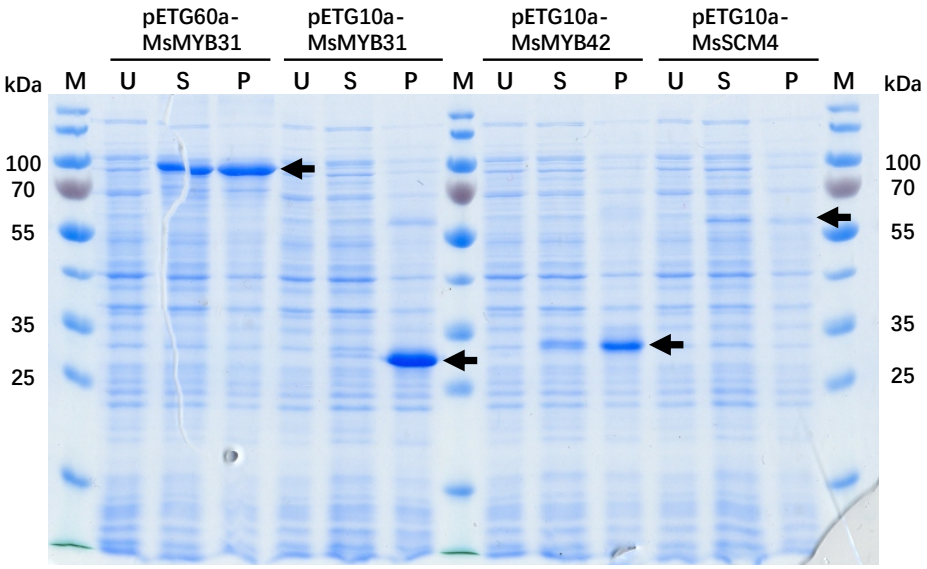

Figure 3-4-2 Expression analysis of recombinant transcription factor proteins in E. coli.
Sample proteins were detected by SDS-PAGE. After electrophoresis, the 12% polyacrylamide gel was stained in Coomassie Blue Staining Solution for 30min followed with shaking 30min in distaining solution. Every 3 lines represent one experiment. Black arrows are pointing out the position of the recombinant proteins. U, uninduced culture; S, supernatant; P, pellet; M, PageRuler Prestained Protein Ladder. The corresponding protein sizes (showing in kDa) are indicated on both sides of the gel picture.

All the 3 expressed recombinant proteins were N-6His-tagged and were purified through HisTrap™ High Performance columns following with ultrafiltration to remove imidazole. Here the SDS-PAGE gel of MsMYB42 protein is shown as an example (Figure 3-4-3). Before ultrafiltration there should be a checking SDS-PAGE gel to select the pure protein elution tubes. Elution tube 3(E3) was discarded because of the excessive non-specific bands. Only E4 to E8 were collected for the ultrafiltration step to improve the purity of the protein. The purified protein now was ready for further biochemical experiments.

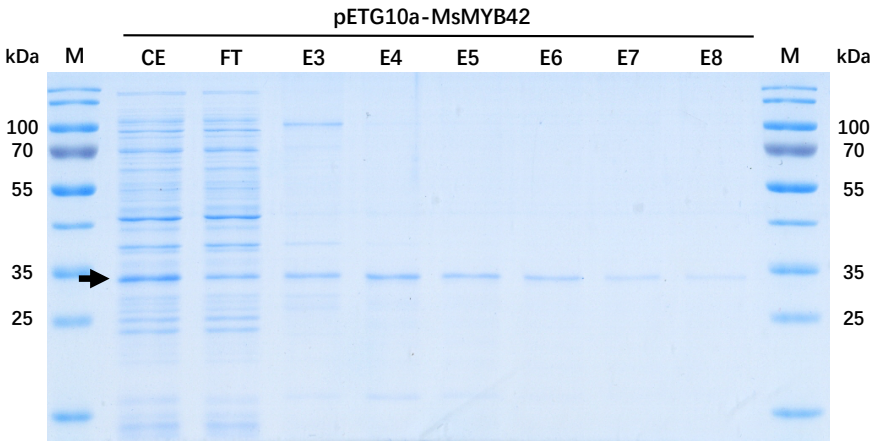

Figure 3-4-3 SDS-PAGE gel showing the purification of 6His-tagged MsMYB42 protein.
Gel condition and staining/distaining steps are the same as described in figure 3-4-2. Black arrow is pointing out the position of the recombinant proteins. CE, crude extract; FT, flow through; E3-E8, elution 3-8; M, PageRuler Prestained Protein Ladder. The corresponding protein sizes (showing in kDa) are indicated on both sides of the gel picture.

3.4.2 MsMYB31 and MsMYB42 proteins bind to the AC-elements of the promoter in vitro

R2R3-MYB transcription factors that are related to the regulation of phenylpropanoid pathway possess similar binding capabilities to the promoters. Previous studies showed that these transcription factors could recognize conserved AC-elements on the promoters, thus regulating gene transcription via DNA-protein interaction. Taking together with the dual-luciferase assay results shown in chapter 3.3, there is very strong support for MsMYB31 and MsMYB42 proteins to have direct DNA-binding abilities. Here 2 selected fragments from the promoter sequence of MsC4H and MsCCR were used as probes to perform electrophoretic mobility shift assays (EMSA) for confirming this hypothesis (Table 3-4-1).

As shown in figure 3-4-4A, MsMYB31 protein was able to recognize both the pMsC4H and pMsCCR probes and form one significant band. Incubating together with increasing amount of unlabeled probes working as competitors could lower the strength of the interaction bond. While MsMYB42, which is shown in figure 3-4-4B, behaved differently. The interaction between MsMYB42 protein and the probes revealed itself

in two significant bands. When using the mutated CY5 probes, both the bands disappeared. The binding intensity could also be obviously weakened by adding competitors (10x) into the reaction system. In addition, the competitors lost their capacities when the putative binding sites were mutated. All these results confirmed that MsMYB31 and MsMYB42 recognized and bound specifically to the AC-elements in the MsC4H and MsCCR promoters.

Table 3-4-1. DNA probes used in electrophoretic mobility shift assays

Both the selected promoter regions contain 2 predicted binding sites indicating in bold font. The total length of the double-strand probes is around 40bp. Probes were synthesized artificially with or without 5'-CY5 labels. Mut: mutated probe, both the predicted binding sites were replaced by dummy sequences.

Probe	Sequence	Location (upstream from ATG)
C4H-AC	GCAGGCCGC**ACCAACC**AAAAACCAT**ACCAACT**ACCACCGCGAT	-210 ~ -168
mut-C4H-AC	GCAGGCCGCCAAGCATAAAAACCATCAAGCAAGTTACCGCGAT	
CCR-AC	GAGAATCCT**ACCAAAC**CCAGCT**ACCAACT**CGGTCATATCAT	-193 ~ -153
mut-CCR-AC	GAGAATCCTCAAGCAGCCAGCTGAAGGAACGGTCATATCAT	

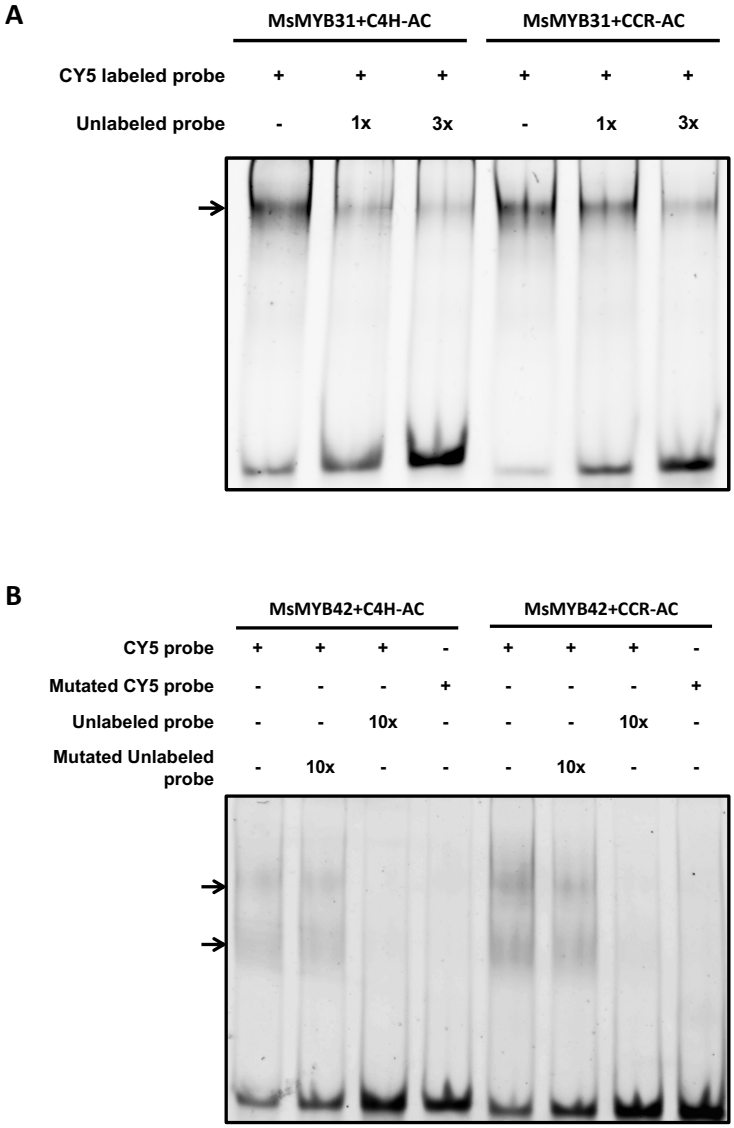

Fig.3-4-4 MsMY31 and MsMYB42 proteins binds to the AC elements in the MsC4H and MsCCR promoters *in vitro*.

Full-length proteins of MsMYB31 (with N-NusA-6His-tag) and MsMYB42 (with N-6His-tag) were used for detecting the interaction. Black arrows indicate the band of DNA-protein complex. (A) Electrophoretic mobility shift assays (EMSA) of MsMYB31 showing the direct binding to the AC elements in the MsC4H and MsCCR promoters. MsMYB31 protein showed binding capacity to both MsC4H and MsCCR promoter probes. Unlabeled probes were used as competitors. (B) EMSA of MsMYB42 showing the direct binding to the double AC elements in the MsC4H and MsCCR promoters. MsMYB42 protein was incubated together with CY5-lableled probe, unlabeled DNA with AC elements or mutated AC elements was introduced into the reaction system as competitors.

3.4.3 MsMYB42 and MsSCM4 compete in binding to the same AC-element on the promoter

Previous work has shown that in *Arabidopsis* the transcription activator MYB63 is able to bind to the AC elements in the promoters of monolignol biosynthetic genes, MsSCM4, which is a homologous gene of AtMYB63, has revealed activation capacity for MsC4H and MsCCR promoters. Based on these observations, we also performed the electrophoretic mobility shift assays with MsSCM4 using the same probes. The mobility shift pattern of MsSCM4 with the probes turned out to be very similar with MsMYB42, which displayed two bands (Figure S6). This result suggests that MsSCM4 and MsMYB42 are able to recognize both the AC-elements on one probe.

As discussed in chapter 3.3, when co-expressing MsSCM4 together with MsMYB42, the activation function of MsSCM4 was significantly suppressed. To further understand the mechanism of this deactivation, a competition electrophoretic mobility shift assay was performed by introducing both MsSCM4 and MsMYB42 in the same reaction. When the probe was incubated with both MsSCM4 and MsMYB42, more bands appeared, demonstrating a competitive combination of proteins and probes (Figure 3-4-5).

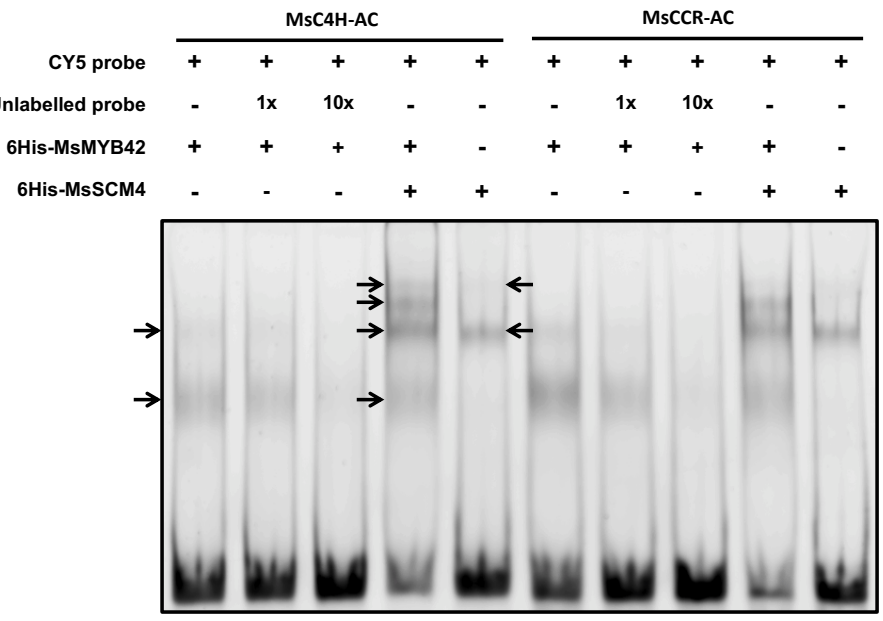

Fig.3-4-5 MsMYB42 and MsSCM4 proteins compete in binding to the AC elements in the MsC4H and MsCCR promoters.

MsC4H-AC shares the same pattern with MsCCR-AC. Using MsC4H-AC as example, the paired black arrows in column 1 and 5 indicate the interaction double bands of DNA probe and one protein. And the four-in-a-row arrows are showing the multiple interaction bands of DNA probe and two proteins.

3.5 Induced overexpression of *Miscanthus* MYB transcription repressors in *Arabidopsis* reduced the resistance of plant towards UV stress and repressed the plant stem development

3.5.1 The generation and phenotyping of the inducible lines of MsMYB31 and MsMYB42 in *Arabidopsis*

According to the dual-luciferase-assay and the electrophoretic mobility shift assay results, MsMYB31 and MsMYB42 were confirmed to function as transcriptional repressors regulating genes related to the phenylpropanoid pathway. To understand the functions of the two repressors in plants, I constructed both overexpression vectors and inducible overexpression vectors for MsMYB31 and MsMYB42 were constructed. Greengate cloning method was used to get the final constructs for transformation. Both of the constructs used CaMV 35S promoter, and in the inducible overexpression ones, a glucocorticoid receptor (GR) was fused at C-terminal of the protein that could be induced by treating with glucocorticoid hormones such as dexamethasone (DEX). Upon DEX treatment, the proteins were transferred to the nucleus and functioned as transcriptional regulator. Figure S7 shows the schematic structure of Greengate constructs used in the study.

When overexpressing MsMYB31 or MsMYB42 in tobacco leaves for subcellular localization experiments (Figure 3-3-1), the two proteins showed strong lethality. The leaves became dehydrated and shrink two days (MsMYB31) or seven days (MsMYB42) after infiltration (Figure S8). Similar phenomenon also has been observed in stable overexpression transgenic lines. All the lines with overexpressed MsMYB31 or

MsMYB42 were false positive. The stable inducible overexpression lines were successfully obtained. For MsMYB31, possibly due to the strong effect when overexpressing the protein in vivo, the positive rate was very low, finally only one line was obtained with relatively low expression. In contrast, four lines with different expression levels of MsMYB42 were selected for subsequent experiments (Figure 3-5-1).

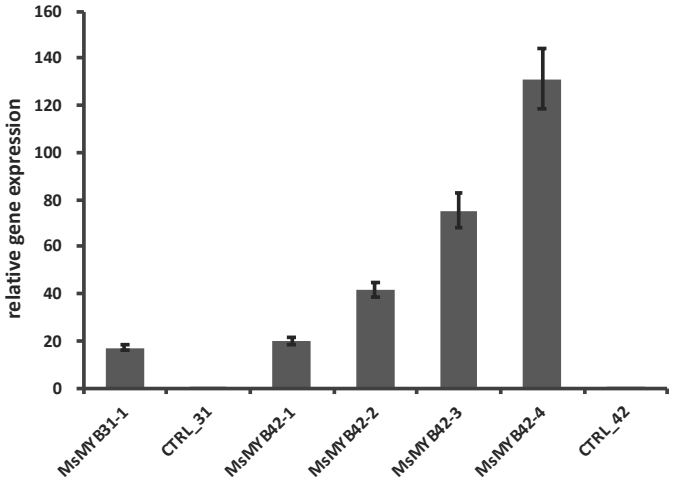

Figure 3-5-1 Relative gene expression of MsMYB31 and MsMYB42 in the inducible overexpression lines. Error bars showed ±SE of at least three biological replicates.

3.5.2 Induced overexpression of MsMYB31 and MsMYB42 in *Arabidopsis* seedlings represses the expression of genes related to phenylpropanoid pathway

Two-week-old seedlings were used for the DEX treatment. The expression of 14 genes involved in phenylpropanoid pathway were tested by quantitative RT-PCR analysis (Figure 3-5-2). In MsMYB31 line, comparing with mock treatment, after four hours' DEX induction, the expression of most of these genes was strongly suppressed except for CCR, C3H and F5H. As for MsMYB42, compared between the two lines, with higher expression of MsMYB42, the suppression of genes was stronger. However, CAD6 out of the 14 genes was not much influenced by MsMYB42.

This 4h-DEX treatment experiment showed a preliminary result for the impact of MsMYB31 and MsMYB42 in plants, more experiments were needed for revealing their real biological functions on the regulation of phenylpropanoid metabolism in vivo.

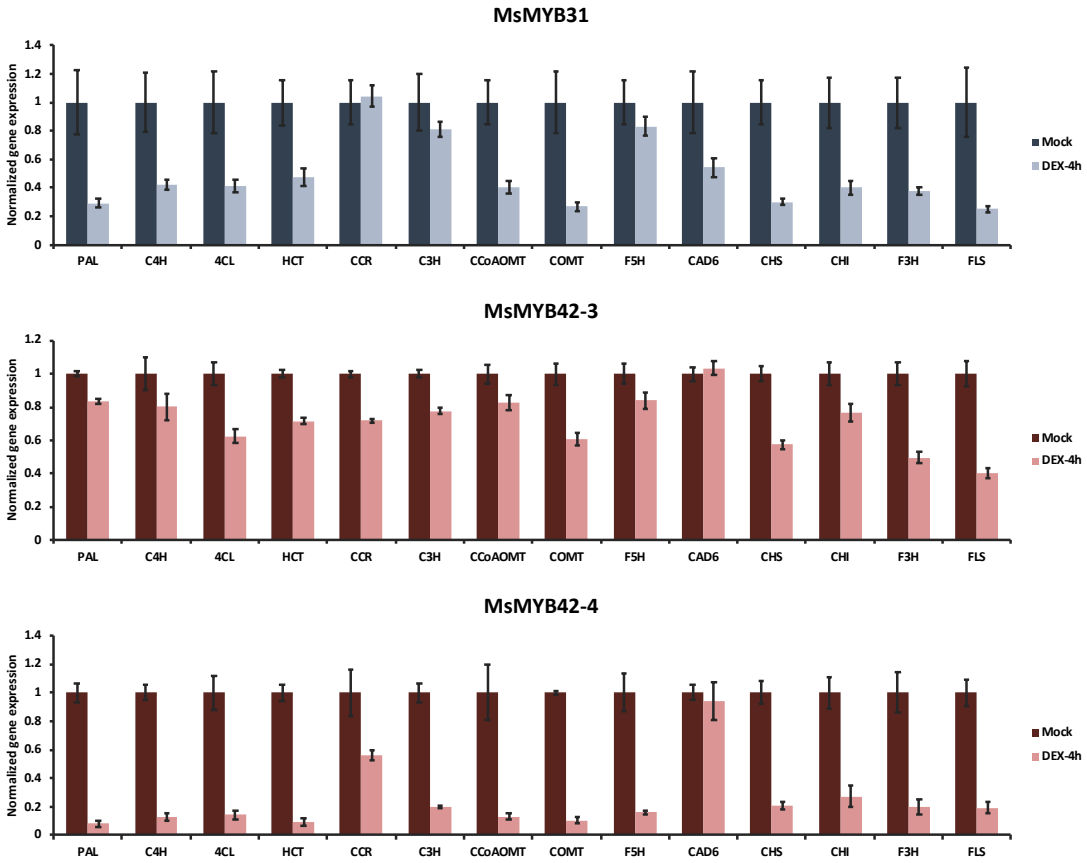

Figure 3-5-2 Induced overexpression of MsMYB31 and MsMYB42 represses the expression of genes related to phenylpropanoid pathway

Two-week-old seedings growing in 1/2MS plates were used for DEX treatment. Final concentration of DEX was 30μM. Genes selected for expression analysis were general phenylpropanoid metabolism enzymes (PAL, C4H, 4CL), monolignol biosynthesis specific genes (HCT, CCR, C3H, CCoAOMT, COMT, F5H, CAD6) and flavonoid biosynthesis specific genes (CHS, CHI, F3H, FLS). Error bars showed ±SE of at least three biological replicates.

3.5.3 Induced overexpression of MsMYB31 and MsMYB42 in Arabidopsis reduced the resistance of plant towards UV stress

In the seedling DEX treatment, *Arabidopsis* genes related to flavonoid biosynthesis such as CHS were repressed by induced overexpression of MsMYB31 or MsMYB42 (Figure 3-5-2). Dual-luciferase assay also confirmed the repression of AtCHS by MsMYB31 or MsMYB42 in suspension cell culture (Figure 3-3-3). Previous study showed there is a tight correlation between high CHS expression and increased resistance to UV, this repression suggests a decreased resistance towards UV stress. A UV resistance treatment was conducted to confirm this hypothesis.

Figure 3-5-3 shows the effects of irradiating DEX-treated Col-0, MsMYB31, MsMYB42-1 and MsMYB42-2 plants with UVB. At all four timepoints, gene expressions without UV irradiation didn't change very much. AtCHS and AtFLS had similar reaction patterns. Both of them react rapidly to the UV irradiation, reaching a peak of expression in all four lines tested at 2h UV treatment timepoint. Among them Col-0 was the highest for both genes. With continued UV treatment, the gene expression levels decreased and didn't change much at the last two timepoints. In contrast, AtF3'H expression in all lines remained at a low level in the first six hours and boosted to the highest level at the last timepoint. In general, the control line Col-0 had the strongest resistance to UV irradiation, all the three transgenic lines with induced overexpression of repressors were more or less weaker than Col-0 after 10h UV irradiation, resulting in more obvious plant dehydration. Additionally, the expression level of MsMYB31 was not higher than MsMYB42 lines, while at all timepoints genes in MsMYB31 line had the lowest expression levels.

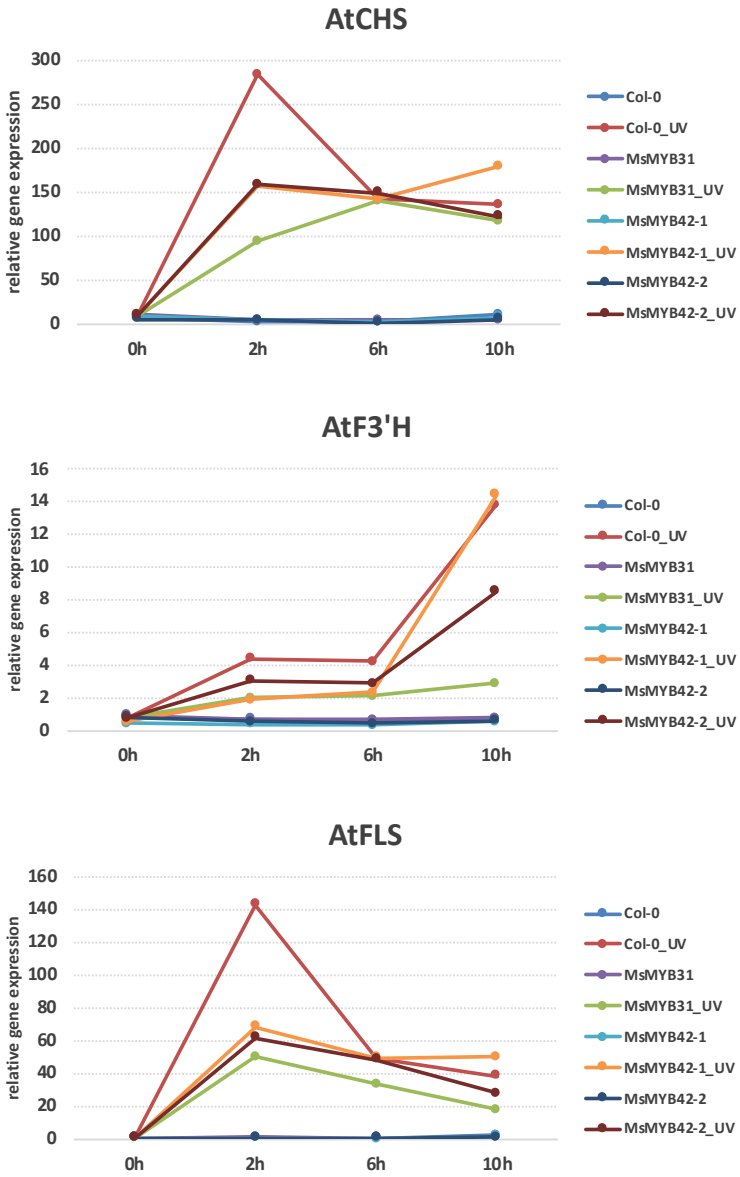

Figure 3-5-3 Induced overexpression of MsMYB31 and MsMYB42 reduces the response of the plants to UV exposure.

Four-week-old plants growing in the greenhouse were used for UV exposure experiment. One day before UV irradiation, the rosette leaves were sprayed by 30μM DEX and at the beginning of UV irradiation (time point 0h), the plants were sprayed once again by 30μM DEX. Then the treatment group was placed under UV-B plus photoactive radiation. Quantitative RT-PCR analysis of AtCHS, AtF3'H and AtFLS has been done using the treated and control samples taken at 0h, 2h, 6h, 10h timepoint.

3.5.4 The influence of continuous induction of MsMYB42 in *Arabidopsis* on inflorescence stem growth

Similar to their homologs in maize, MsMYB42 revealed a general repression function of phenylpropanoid metabolic pathway genes including monolignol biosynthesis related genes. Therefore, it is likely also for MsMYB42 to affect the development of stems in plants. Inducible lines MsMYB42-1, MsMYB42-2 and MsMYB42-3 with different MsMYB42 expression levels were used to test the effect of the continuous induction of this repressor in plant stem growth. Line MsMYB42-4 was abandoned in this experiment because the continued high expression of MsMYB42 caused plant death (Figure S9). In each line, plants were divided into two groups with average plant height. In order to keep the expression of MsMYB42 at a relatively high level while the plants could still survive, the treated group was sprayed by 30uM DEX every two days for ten days. All the stem heights measured in the last day are shown in figure 3-5-4.

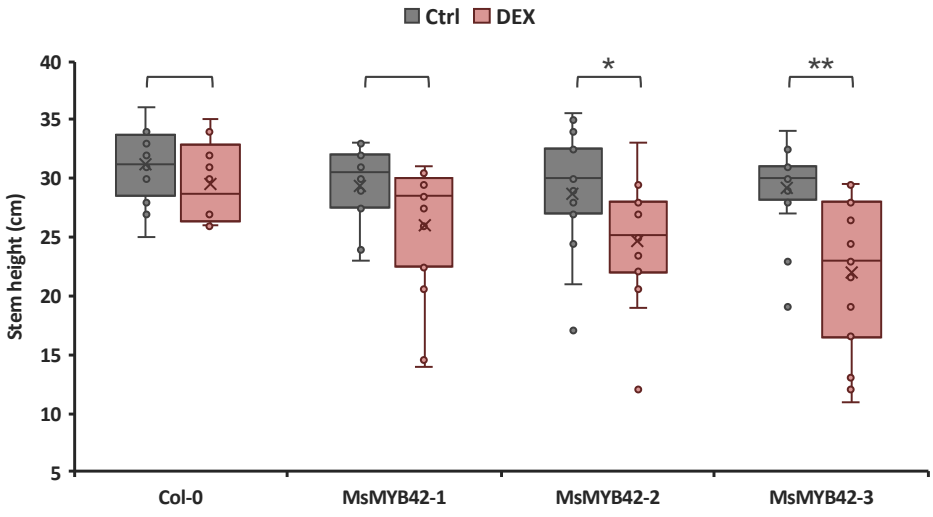

Fig 3-5-4 Continuous induced overexpression of MsMYB42 in *Arabidopsis* suppresses the inflorescence stem growth.

MsMYB42-1, MsMYB42-2 and MsMYB42-3 were tested, Col-0 was used as control. Boxes in gray indicate the height statistics of mock treatment group; boxes in light red are the height statistics of DEX treated plants. For each box plot, n=20. Statistical significance, *, $P<0.05$; **, $P<0.01$.

In mock treatment group, there are no significant differences of the stem heights in three inducible lines and Col-0. When comparing between DEX and mock treatment in each independent line, the effect of DEX on Col-0 plants as control was not significant. While in MsMYB42 lines, as the expression level of MsMYB42 increases, the effect of DEX on stem height becomes more apparent. In both MsMYB42-2 and MsMYB42-3 lines, the stem height of the DEX treatment group was significantly lower than that of the control group. This experiment confirmed the overexpression of MsMYB42 in plant repress the stem development.

3.6 Expression of *Miscanthus* MYB31 and MYB42 in *Arabidopsis* myb4 mutant under control of AtMYB4 promoter

3.6.1 Seedlings expressing MsMYB31 or MsMYB42 protein were more sensitive towards UV treatment than myb4 mutant and Landsberg WT *Arabidopsis*

In last chapter, the impacts of induced overexpression of MsMYB31 and MsMYB42 in *Arabidopsis* were evaluated in different aspects. myb4 knockout mutant (ecotype Landsberg erecta) is a nice tool to determine the function of MsMYB31 and MsMYB42 in *Arabidopsis* without any influence of native homolog protein AtMYB4. Thus, transgenic lines expressing MsMYB31 or MsMYB42 driven by 1200bp native AtMYB4 promoter were generated using myb4 mutant as background to test whether MsMYB31 and MsMYB42 can complement the function of AtMYB4. Constructs used in transformation are shown in figure S7. For each gene, two independent single insertion lines were selected and the gene expression of MsMYB31 or MsMYB42 were tested (Figure. 3-6-1). The expression level of AtMYB4 in Landsberg WT (Ler-0) and myb4 mutant were also shown in figure 3-6-1.

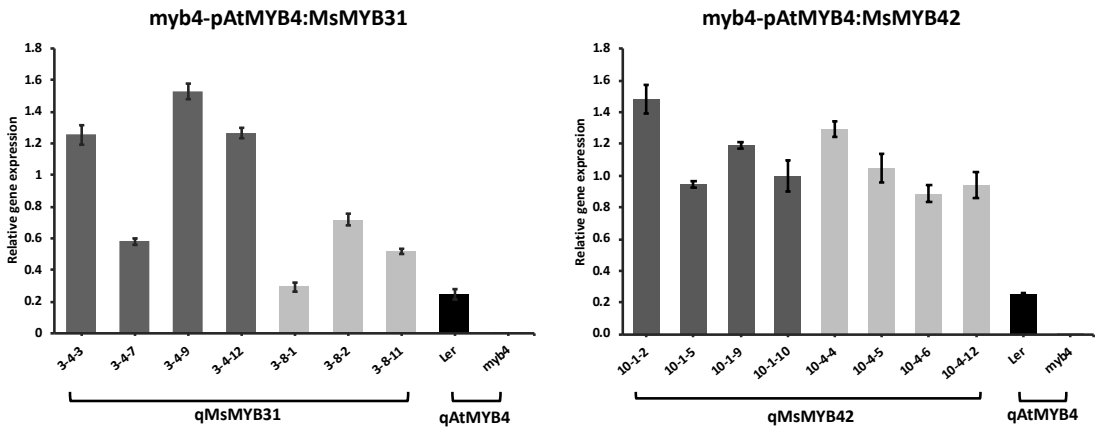

Figure 3-6-1 Quantitative RT-PCR analysis of MsMYB31 and MsMYB42 in myb4 complementation lines.

3-4 and 3-8 are two independent lines of MsMYB31, and 10-1 and 10-4 are two of MsMYB42. The gene expression level of MsMYB31 and MsMYB42 were measured respectively. *Exp2* was used as reference gene. The expression level of MYB4 in Landsberg WT (Ler) and myb4 mutant was also tested as controls. MsMYB31 or MsMYB42 was not detected in both controls (data not shown).

According to the quantitative RT-PCR analysis, for MsMYB31, average gene expression in line 3-4 (named cp-MsMYB31-1) is higher than 3-8 (named MsMYB31-2); while the two selected lines of MsMYB42, 10-1 and 10-4 (named cp-MsMYB42-1 and cp-MsMYB42-2 respectively), have relatively similar expression level. It is also confirmed via quantitative RT-PCR that myb4 mutant used in plant transformation is a knockout mutant.

It was reported in previous work that myb4 mutant line is more tolerant of UV-B irradiation than wild type. In addition, the repressor AtMYB4 expression is downregulated by exposure to UV-B light, indicating the acclimation to UV-B via de-repression mechanism (Jin et al., 2000). To determine whether MsMYB31 or MsMYB42 also have similar function comparing with AtMYB4, a UV resistance test was done with Ler-0, myb4, and MsMYB31, MsMYB42 transgenic lines. Figure 3-6-2A shows the phenotypes of different lines after 10h treatment. Comparing with normal parabolic aluminized reflector (PAR) lamp only, the seedlings in all lines treated with PAR+UV-B revealed varying degrees of weakening (Hollósy, 2002). Figure 3-6-2B shows the comparison between each line after grouping according to different degrees. Among all the six lines tested, myb4 mutant had a distinctly highest resistance towards

UV-B compared with other lines. Most of the seedlings still had normal appearance, only a few were slightly weaker, probably due to the accumulation of flavonoids. The four transgenic lines were least able to resist UV-B, most of the seedlings became shrunk after the 10h treatment. As for the ability of Ler-0 WT, it was in between the mutant and the transgenic lines, about half of the seedlings remained normal.

After treatment, the seedlings were returned to the greenhouse condition, the survival rate for all the lines were calculated after a seven days recovery. seedlings could not survive would fade and shrink, while the leaves of the surviving seedlings would gradually expand (Figure S10). myb4 line possessed the highest survival rate of 88.24%, followed by Ler-0 with 43.75%. In contrast, the survival rates of all the four transgenic lines were fairly low, less than 6%.

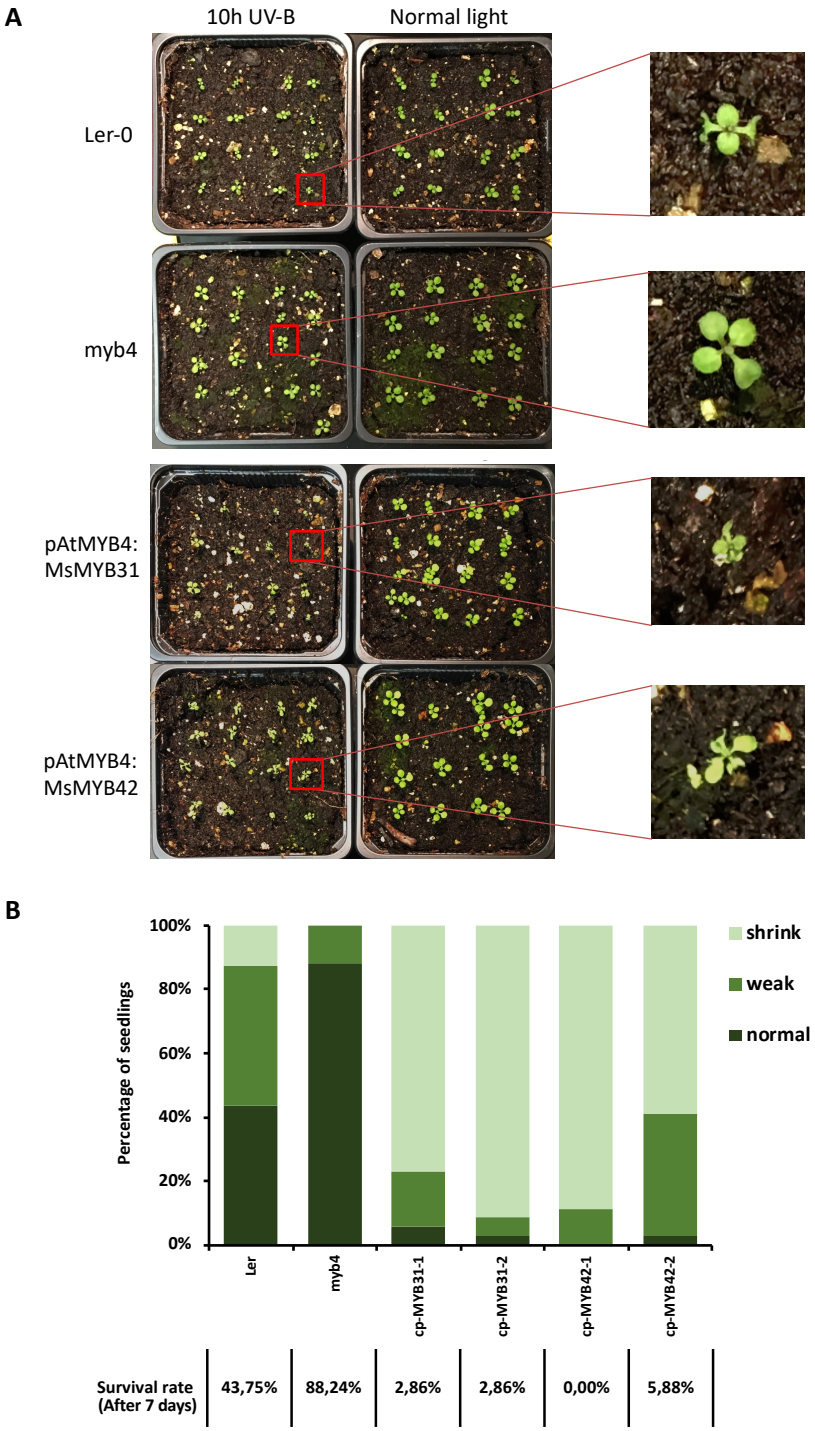

Figure 3-6-2 Phenotypes of different lines after 10h UV-B treatment.

(A) 7-day-old seedlings growing in the greenhouse were used for the UV resistance treatment. Different lines were treated with PAR+UV-B (0.1 mW/cm^2) or PAR only for 10h, pictures were taken right after irradiation. (B) Seedlings were divided into three groups of different conditions (normal, weak and shrink) after irradiation and counted, for each line in two different treatments, n=32. The survival rate of each line was calculated after recovering for seven days.

3.6.2 Plants transformed with MsMYB31 or MsMYB42 driven by AtMYB4 native promoter did not complement the gene expression changes comparing with myb4 mutant

The expression levels of fifteen different genes involved in the phenylpropanoid metabolic pathway were analyzed in Landsberg WT, myb4, and the transgenic lines. In order to reduce the influences caused by different environmental or growth conditions, when preparing cDNAs for qRT-PCR, in each line three to four separate cDNA samples were mixed together. In general, plants transformed with MsMYB31 or MsMYB42 driven by AtMYB4 native promoter did not complement the gene expression changes comparing with WT and myb4 mutant (Figure 3-6-3).

According to the qRT-PCR results, firstly, all the fifteen genes had lower expression level in myb4 mutant than in WT, among those there were still several genes (e.g. 4CL and HCT) had not very much reduction in myb4 mutant than in WT. In MsMYB31 and MsMYB42 transgenic lines, genes belonging to the general phenylpropanoid metabolic and lignin biosynthesis pathway had either comparable expression levels with those in myb4 mutant (e.g. F5H and C3H), or even lower expression levels (e.g. PAL, CCoAOMT and COMT). As for some of the genes related to flavonoids biosynthesis pathway (e.g. CHS and DFR), they had relatively higher expression in MsMYB31 and MsMYB42 transgenic lines than in myb4 mutant, but still lower than those in WT.

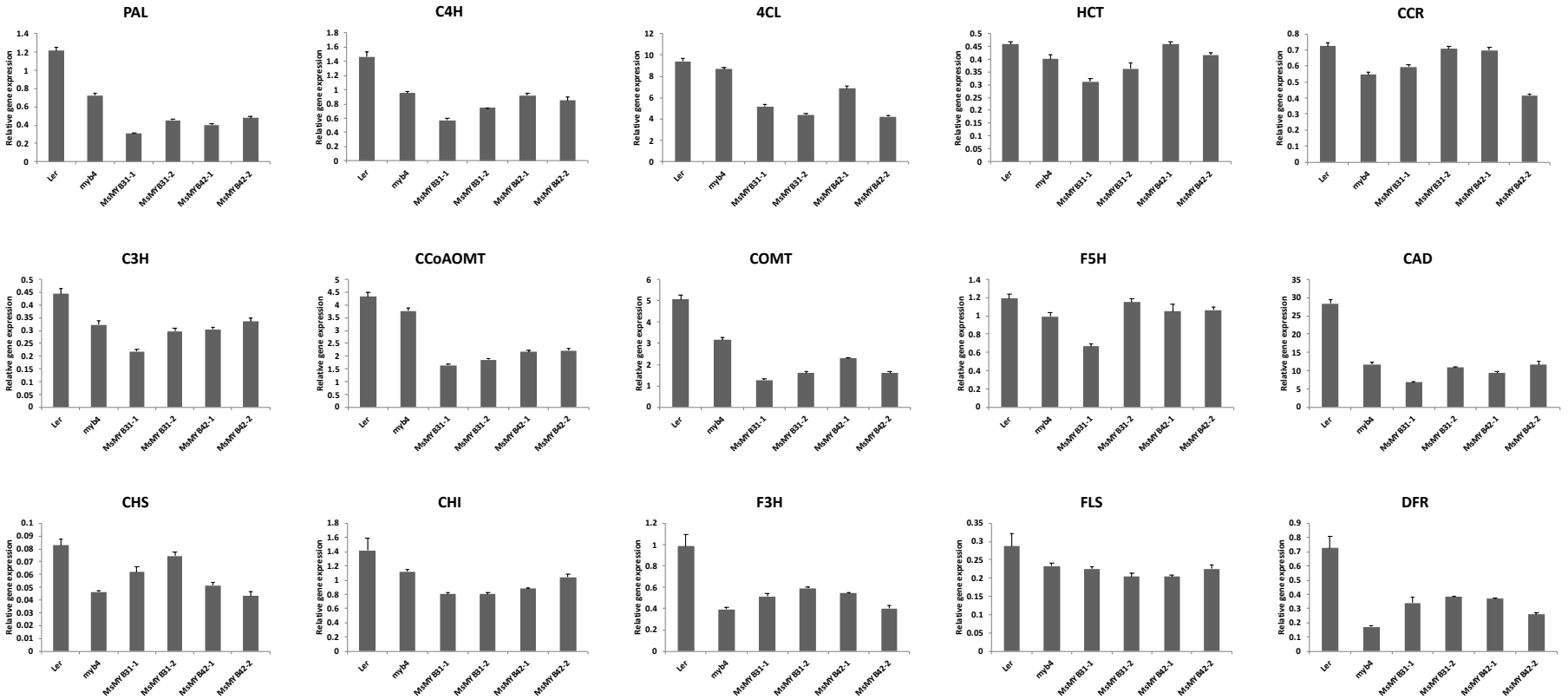

Figure 3-6-3 Quantitative RT-PCR analysis of genes involved in phenylpropanoid metabolic pathway in different *Arabidopsis* lines.

3.6.3 The analysis of cell wall components in different *Arabidopsis* lines

On the basis of the gene expression analysis of different lines, there is also the possibility for the transgenic lines to have altered lignin contents, thereby affecting cell wall components. Thus, mature stem samples of Landsberg WT, myb4, and the transgenic lines were collected for cell wall components determination experiments.

3.6.3.1 Lignin content decreased while pectin content increased in the transgenic lines

According to figure 3-6-4A, there is significant reduction of the insoluble lignin in all transgenic lines, while acid soluble lignin slightly increased in these lines. Taking both the contents together, all four transgenic lines have significantly less lignin comparing with Ler-0. Among them cp-MsMYB31-1 has the strongest reduction, both the contents in two MsMYB42 lines are slightly higher than the two MsMYB31 lines. Lignin content in myb4 is roughly the same as Ler-0. As for pectin, the transgenic lines have higher contents than Ler-0 and myb4 (Figure 3-6-4B). Corresponding to the lowest lignin content, cp-MsMYB31-1 possess the highest pectin content.

3.6.3.2 H lignin slightly increased in transgenic lines thus altered the H/G and H/S ratio

To look more in details of the influence of MsMYB31 or MsMYB42 in myb4 mutant, lignin composition in different lines was measured. Table 3-6-1 showed the monolignol composition in weight percentage, mol percentage and the ratios between then in all the six mature stems.

H lignin is the least abundant one among all the three monolignol compositions in dicots. In the four transgenic lines, the content of H lignin increased comparing with Ler-0, with an increasement range from 12.8% (cp-MsMYB42-1) to 40.8% (cp-MsMYB31-1). On the contrary, S and G lignin have no significant changes in the transgenic lines.

For myb4 mutant, it has a slight increase of H and S lignin (6.7% each) comparing with Ler-0, leading to a higher S/G ratio.

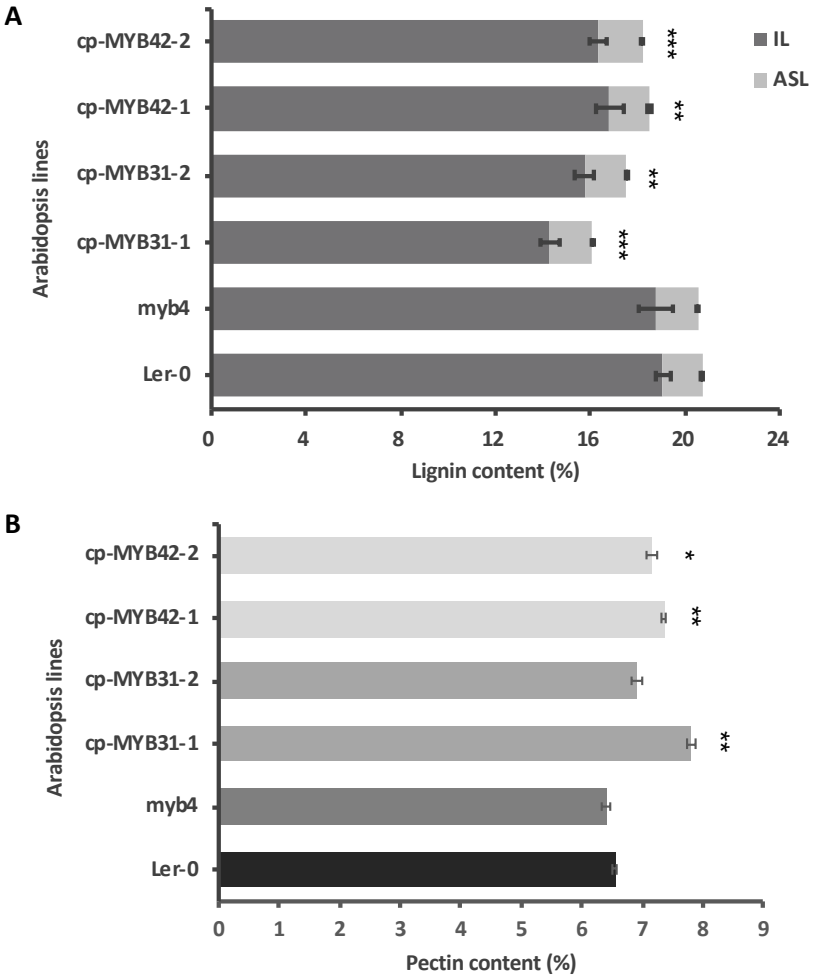

Figure 3-6-4 Total lignin and pectin contents of mature stems of different *Arabidopsis* lines.
(A) Lignin content (weight percentage) is presented by the sum up of insoluble lignin (IL) and acid-soluble lignin (ASL) contents in mature stems of different lines. (B) measurement of pectin content (weight percentage) in mature stems of different lines. Ler-0 WT is used as control. Student's t-test showing the significant differences comparing with Ler-0: *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$.

Table 3-6-1 Monolignol composition of mature stems in different *Arabidopsis* lines

	weight percentage			mol percentage			ratios		
	H	G	S	H	G	S	H/G	S/G	H/S
Ler-0	3.43	61.62	34.96	4.49	64.80	30.71	0.069	0.474	0.146
myb4	3.66	59.05	37.30	4.81	62.32	32.88	0.077	0.528	0.146
cp-MYB31-1	4.83	60.92	34.25	6.30	63.76	29.94	0.099	0.470	0.210
cp-MYB31-2	4.26	60.48	35.26	5.57	63.51	30.92	0.088	0.487	0.180
cp-MYB42-1	3.87	60.90	35.24	5.06	64.01	30.93	0.079	0.483	0.164
cp-MYB42-2	4.05	59.54	36.41	5.31	62.68	32.01	0.085	0.511	0.166

3.6.3.3 Cellulose and hemicellulose components in complementation lines did not change significantly

In addition to lignin and pectin, the major polymer structure in cell wall is cellulose and hemicellulose. The content analysis in the six lines revealed that sugar content is not strongly affected by either loss of AtMYB4 or introduction of MsMYB31 or MsMYB42. Only cp-MsMYB31-1 has a significant induction in cellulose content (Table 3-6-2).

Table 3-6-2 Sugar contents of mature stems in different *Arabidopsis* lines

Data shown as means ± SD (n=3). Student's t-test was used for significance analysis: *, P<0.05.

Lines	Cellulose (%)	Hemicellulose (%)
Ler-0	54.48±1.74	19.67±0.42
myb4	53.89±1.03	20.42±0.03
cp-MYB31-1	58.96±0.60 *	19.52±0.19
cp-MYB31-2	56.27±1.42	19.81±0.32
cp-MYB42-1	54.77±0.96	19.33±0.7
cp-MYB42-2	56.88±2.71	19.79±0.68

Taking all together, stable expression of MsMYB31 and MsMYB42 driven by native promoter in *Arabidopsis* indeed influenced the cell wall components by reducing lignin contents, while the transgenic lines have slightly higher contents of pectin. The sugar polymers did not strongly influence by the two transcription repressors. Some of the activators involved in cell wall formation could activate both lignin biosynthesis and

cellulose related genes, while in MsMYB31 and MsMYB42 lines did not show much evidences related to similar mechanisms.

3.6.4 The delayed inflorescence stem development of myb4 mutant can be complemented in the transgenic lines in greenhouse growth condition

Previously, in the process of growing the myb4 mutant material for harvesting seeds, comparing with Ler-0, a phenotype of delayed inflorescence stem development of myb4 mutant could be observed in greenhouse growth condition (Figure S11). This delayed growth phenotype was only significant at the early stem development stages. Although myb4 mutant was still not as high as the others, the height differences were not statistically significant anymore when collecting seeds from mature plants. This time, during the preparation of the mature stem tissues of all the six *Arabidopsis* lines for cell wall contents measurements, myb4 mutant showed the same delayed growth phenotype. In addition, the delayed inflorescence stem development of myb4 mutant can be reversed in the MsMYB31 or MsMYB42 transgenic lines.

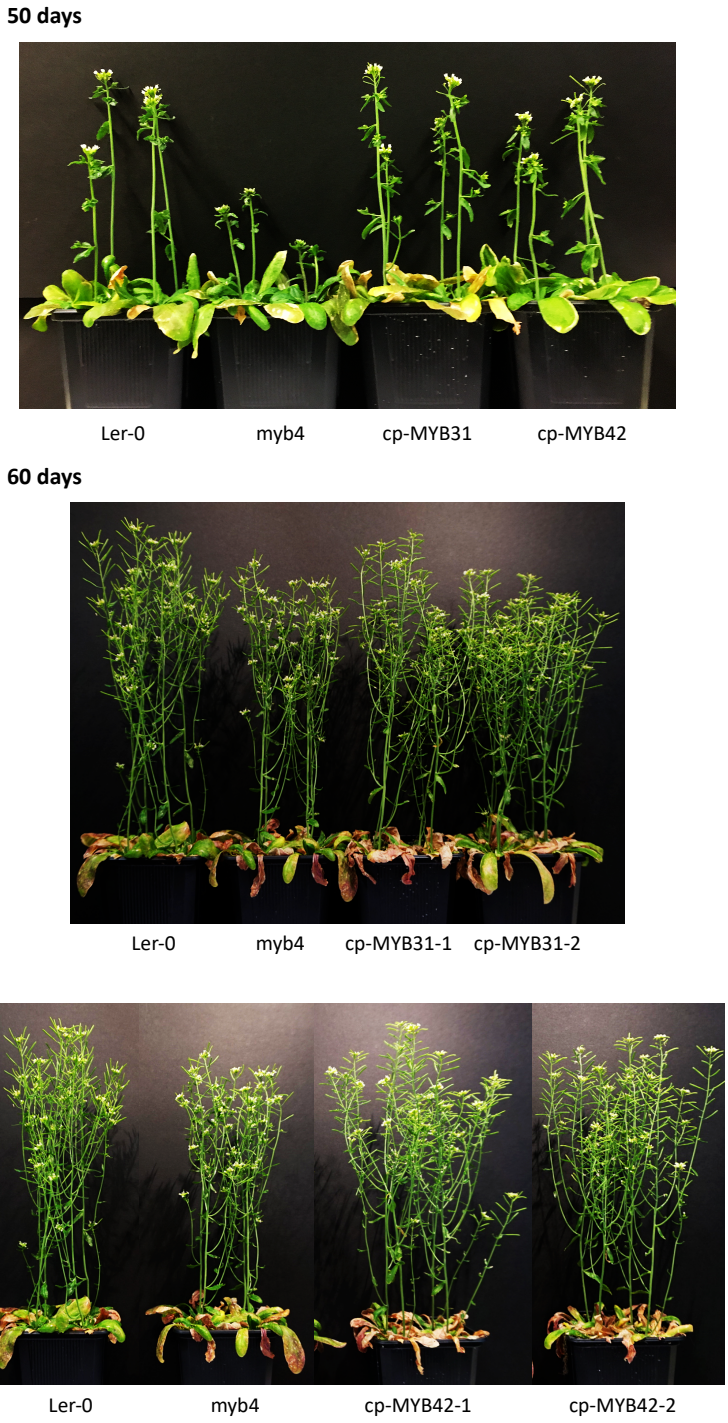

Figure 3-6-5 Phenotypes of Ler-0 WT, myb4 mutant and MsMYB31/MsMYB42 transgenic lines at 50 days and 60 days.

Plants were growing in greenhouse condition. Pots for photographs in each line generally represent the average height of different lines.

As shown in Figure 3-6-5, at day 50, the myb4 mutants were significantly shorter than the plants of all the other lines, accompanying a delayed development of inflorescence. 10 days later, the stems of myb4 mutant almost caught up to the same height of the

other lines. Generally, average height of Ler-0 plants were the highest among all the six lines in day 60, but not significantly different from the others. It's worth mentioning that the germination was not affected proving by figure 3-6-2A. Figure 3-6-6 shows the summary of the growth data, the percentage of height increase for each line is calculated and shown below the line names. The height of myb4 mutant stems increased by 76.25% in ten days. The transgenic lines complemented this phenotype of stem growth, similar to Ler-0, increasing by around 50%.

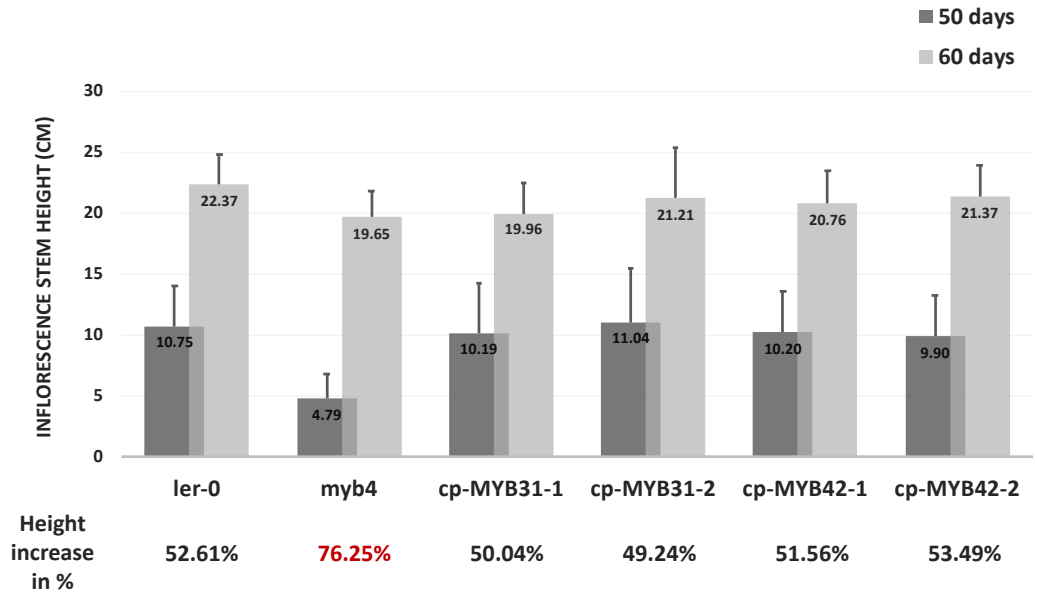

Figure 3-6-6 Measurement of the inflorescence stem heights.

The height of inflorescence stems of Ler-0 WT, myb4 mutant and cp-MsMYB31/MsMYB42 transgenic lines were measured at both 50 days (dark gray) and 60 days (light gray). N=32, experiment repeated two times in the same growth condition.

In the meantime, another set of the same plants was growing in the climate chamber, with the long-day condition, 22 degrees in 16h day time and 18 degrees at 8h night. During the stem growth stages, the delayed phenotype of myb4 disappeared. There were no significant differences between different lines anymore. This may due to the different growth environments.

3.6.5 The silique-stem angle changed in myb4 mutant and the transgenic lines

Although the growth environment differences may have strong impacts on stem development, there is a phenotype difference between WT and the others that appeared in both growth environments. In figure 3-6-7A, plants growing in the climate chamber were used for taking pictures. The silique-stem angle of Ler-0 WT is obviously smaller than the other lines; two MsMYB31 lines have even larger angle compared with myb4 mutant and MsMYB42 lines. This phenotype can also be seen in figure 3-6-5, when the plants were growing in the greenhouse. Figure S12 also shows the picked off stems in the greenhouse condition.

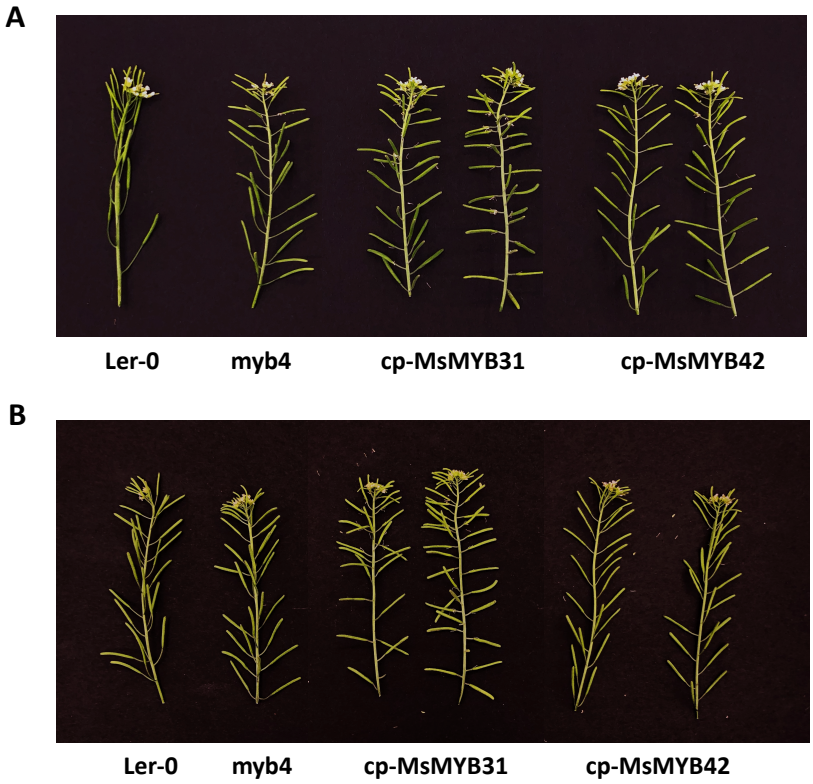

Figure 3-6-7 Phenotype of silique angle in Ler-0 WT, myb4 mutant and cp-MsMYB31/MsMYB42 transgenic lines

Around 10cm of the inflorescent stems were cut from different *Arabidopsis* lines. Photos were taken directly after picking off from the stem (A) or after placing on the bench for two hours (B).

Interestingly, after taking the picture in figure 3-6-7A, the cuts of stems were placed on the bench for two hours, resulting in a larger angle of siliques of Ler-0. So far, the angle difference shown in Ler-0 was not obvious anymore, while the evident larger angles in MsMYB31 lines were stable. Taking all together, the determining factors of the silique-

stem angles might not only be influenced by the content of supporting structures in the cell wall e.g. lignin content, water content in the cell is also an important factor.

Supplementary Materials

(A)

```
T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)
Cedric Notredame
SCORE=89
*
 BAD  AVG  GOOD
*
MsMYB31    :  89
MsMYB42    :  89
cons       :   8

MsMYB31    MGRSPCCEKAHTNKGAWTKEEDDRLVAYIKAHGEGCWRSLPKATGLLRCGKSCRLRWINYLRPDLKRGNFTE
MsMYB42    MGRSPCCEKAHTNKGAWTKEEDQRLIAYIKAHGEGCWRSLPKAAGLLRCGKSCRLRWINYLRPDLKRGNFTE

cons       **********************:**:*****************:****************************


MsMYB31    EEDELIIKLHSLVGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRGIDPVTHRPINEHTSNITISFEAAA
MsMYB42    EEDELIIKLHSLLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLGRGIDPVTHRPIAGAGAVTTISFQPNP

cons       ************:*********************************.***********    :  ****:. .


MsMYB31    -----AAANRDREENKGAIFRLEEHNKAAAIGRHHQNHPAGVWGQGKPLKCPDLNLDLCISPPAPCQEEAAM
MsMYB42    NAAAAGAAQAPQQPIKTAAT----------------------AVKAPRCPDLNLDLCISPPCQQQEEDDE

cons            .**:  ::  * *                          *. :**************.  ***


MsMYB31    ----VMMKP---VKR-------EAGLCFSCSLGLPKSA-DCKCS-----NFLGIRTAMLDFRSLEMK
MsMYB42    EDEEQNLKPAVVVKREVLQAGHGGGLCFGCSLGIQKGAPGCSCSSSNRHHFLGLRAGMLDFRGLEMK

cons            :**   ***        .****.****: *.* .*.**     :***:*:.*****.****
```

(B)

```
T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)
Cedric Notredame
SCORE=94
*
 BAD  AVG  GOOD
*
OsMYB31    :  93
SbMYB31    :  94
MsMYB31    :  94
ZmMYB31    :  93
cons       :   9

OsMYB31    MGRSPCCEKAHTNKGAWTKEEDDRLIAYIKAHGEGCWRSLPKAAGLLRCGKSCRLRWINYLRPDLKRGNFTE
SbMYB31    MGRSPCCEKAHTNKGAWTKEEDDRLVAYIKAHGEGCWRSLPKAAGLLRCGKSCRLRWINYLRPDLKRGNFTE
MsMYB31    MGRSPCCEKAHTNKGAWTKEEDDRLVAYIKAHGEGCWRSLPKATGLLRCGKSCRLRWINYLRPDLKRGNFTE
ZmMYB31    MGRSPCCEKAHTNKGAWTKEEDERLVAHIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINYLRPDLKRGNFTE

cons       **********************:**:*:*:*************:****************************


OsMYB31    EEDELIIKLHSLLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRGIDPVTHRPINDS-ASNITISFEAA
SbMYB31    EEDELIIKLHSLLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRGIDPVTHRPINEH-TSNITISFEAA
MsMYB31    EEDELIIKLHSLVGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRGIDPVTHRPINEH-TSNITISFEAA
ZmMYB31    EEDELIVKLHSVLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRGIDPVTHRPVTEHHASNITISFETE

cons       ******:****::*********************************************:.:  :********:


OsMYB31    AAAA-RD---DKAAVFRREDHPHQP-----KAVTV------AQ-EQQAAADWGHGK-PLKCPDLNLDLCISL
SbMYB31    AAA--RDREENKGAVFRLEEHNKAT-AAAAAAIGRD-----HHQNHHPAGDWGQGK-PLKCPDLNLDLCISP
MsMYB31    AAAANRDREENKGAIFRLEEHNKA------AAIGR------HH-QNHPAGVWGQGK-PLKCPDLNLDLCISP
ZmMYB31    VAAAARD--DKKGAVFRLEDEEEEERNKATMVVGRDRQSQSHS-HSHPAGEWGQGKRPLKCPDLNLDLCISP

cons       .**  ** . .*.*:** *:. .  .  .:      . . :.*. **:** ***************


OsMYB31    P---SQEE-----PM--MMKPVKRETGVCFSCSLGLPKSTDCKCS--SFLGLRTAMLDFRSLEMK
SbMYB31    PAAPCQEE-----KAMVTMKPVKREAGLCFSCSLGLPKSADCKCS--NFLGLRTAMLDFRSLEMK
MsMYB31    PA-PCQEE-----AAMVMMKPVKREAGLCFSCSLGLPKSADCKCS--NFLGIRTAMLDFRSLEMK
ZmMYB31    P---CQEEEEMEEAAMRVRPAVKREAGLCFGCSLGLPRTADCKCSSSSFLGLRTAMLDFRSLEMK

cons       *  .***      .****:*:**.******:::*****  .***:**************
```

(C)

Figure S1 Multiple sequence alignment using T-COFFEE program

Sequences comparison of (A) MsMYB31 and MsMYB42; (B) members of MYB31 group in monocots; (C) members of MYB42 group in monocots.

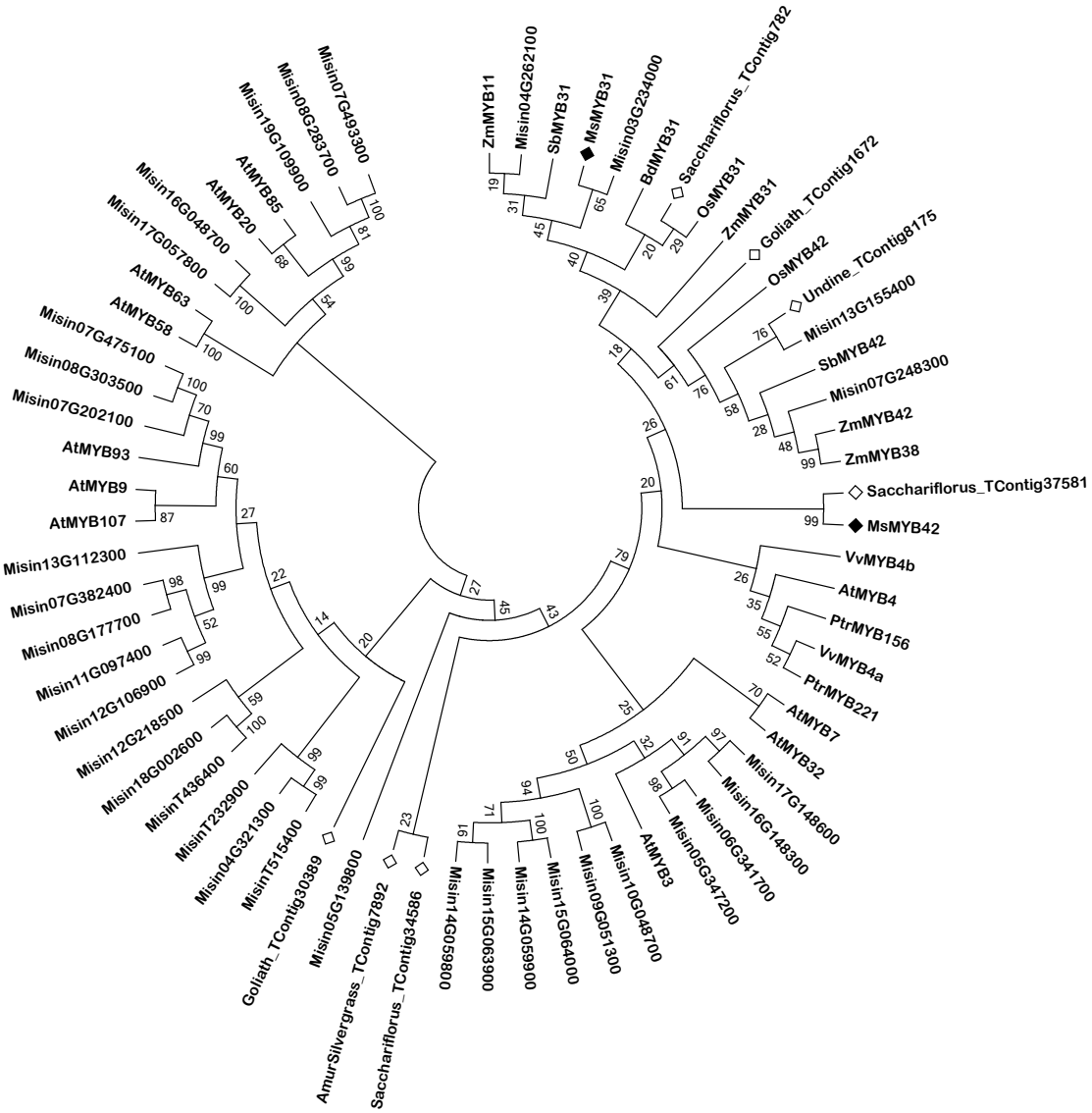

Figure S2 Neighbor-joining phylogenetic tree of in total 67 sequences

In total 63 amino acid sequences were analyzed to construct the Neighbor-Joining tree using MEGA 7.0 software. MsMYB31 and MsMYB42 are highlighted by solid diamond squares and the hollow diamond squares indicate seven contigs in the *Miscanthus* EST database. Sequences with the name starting with "Misin" are all obtained from the *Miscanthus sinensis* database (v7.1 DOE-JGI, http://phytozome.jgi.doe.gov/).

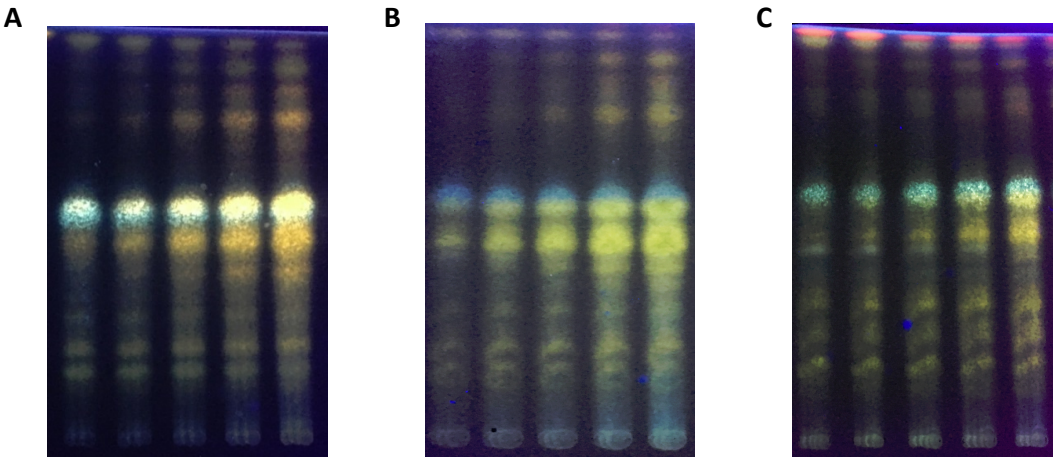

Figure S3 Thin layer chromatography pictures of different *Miscanthus* leaf gradient axis. Each sample is a mixture of 3 leaf blades from 3 independent shoot sharing the same pattern with figure 3-2-2. (A) Samples of 2-month-old new plants from seeds grown in the green house, blade length around 60 cm. (B) Samples from plants in the botanic garden field, blade length around 70 cm. (C) Samples from new *Miscanthus* shoots generated from old tillers growing in the greenhouse, blade length around 70 cm.

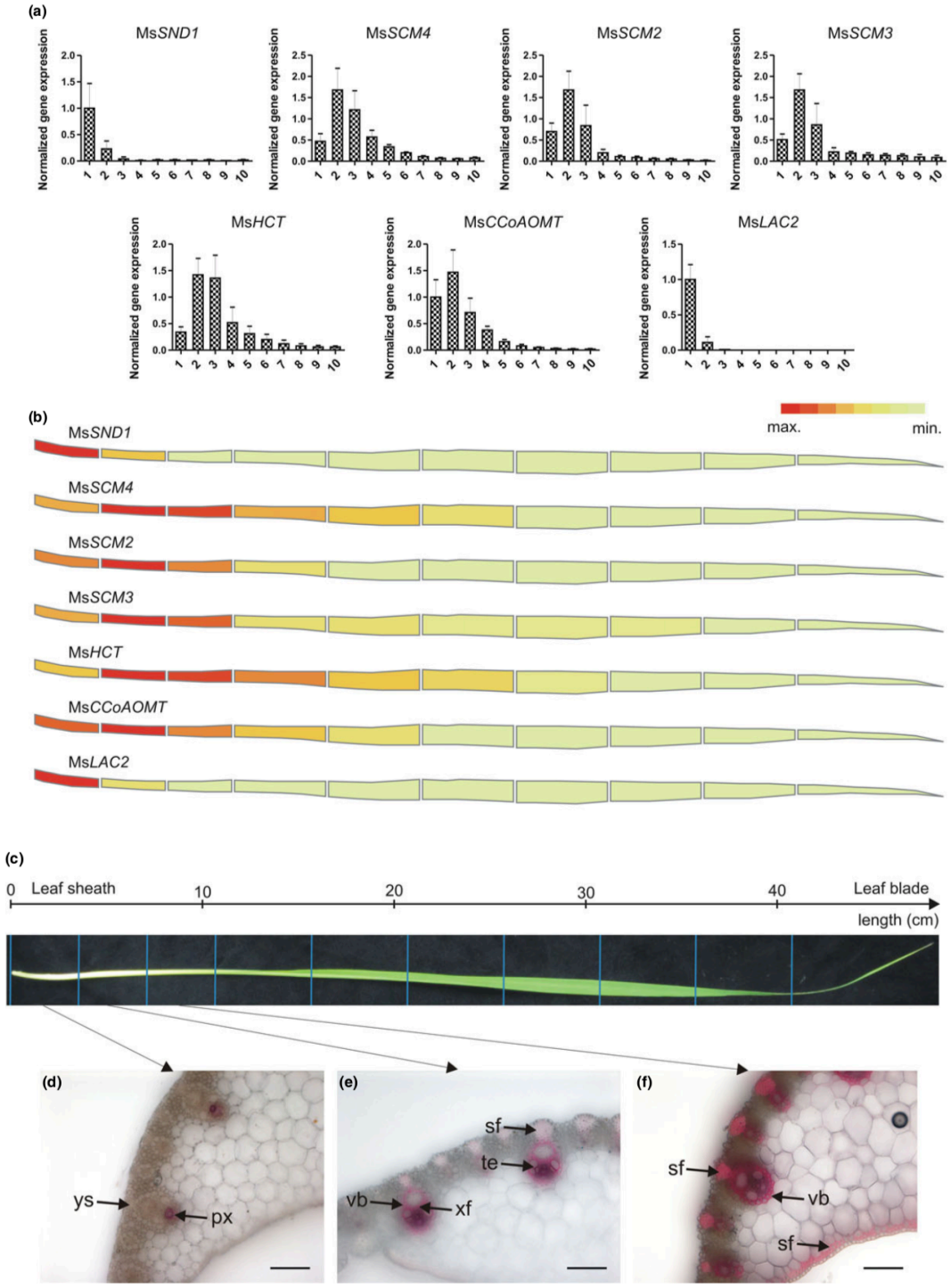

Figure S4 MsSND1 expression is correlated with vascular development and with the expression of its putative target genes *Miscanthus* (obtained from (Golfier, 2017))

Gene expression profiles of the indicated genes over ten developmental zones as obtained by quantitative real-time PCR. The results of three biological replicates were combined, normalized against two reference genes (PP2A and UBC, [a]) and visualized as heat map in (b). Cross-sections of the first three basal zones (d–f) from the *Miscanthus* leaf depicted in (c) were stained for lignin with HCl-phloroglucinol, indicating that the expression of MsSND1 and its putative targets is concomitant with the onset of vascular development. Scale bars: 100 lm. px, protoxylem; sf,

sclerenchyma fibers (extraxylary fibers); te, tracheary elements; vb, vascular bundle; xf, xylary fibers; ys, young sclerenchyma

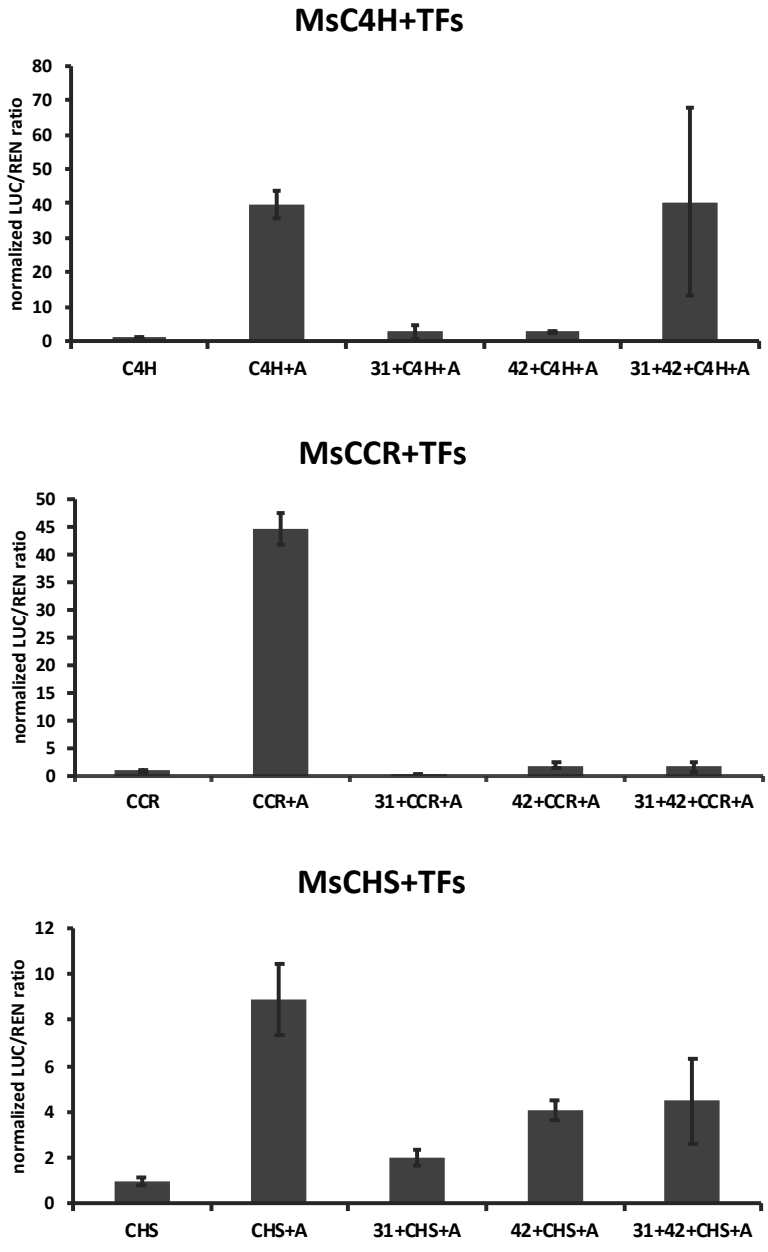

Figure S5 Deactivation function of MsMYB31 plus MsMYB42 on MsSCM4 towards *Miscanthus* C4H, CCR and CHS gene promoters.

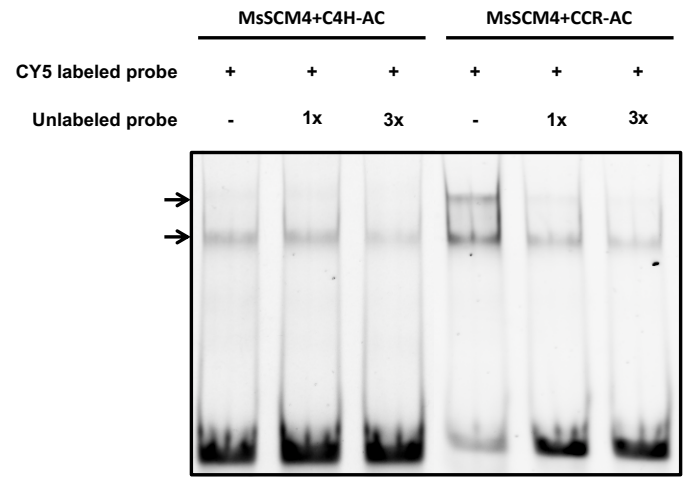

Figure S6 MsSCM4 protein binds to the AC elements in the MsC4H and MsCCR promoters in vitro.

The black arrows indicate the interaction band of DNA and protein.

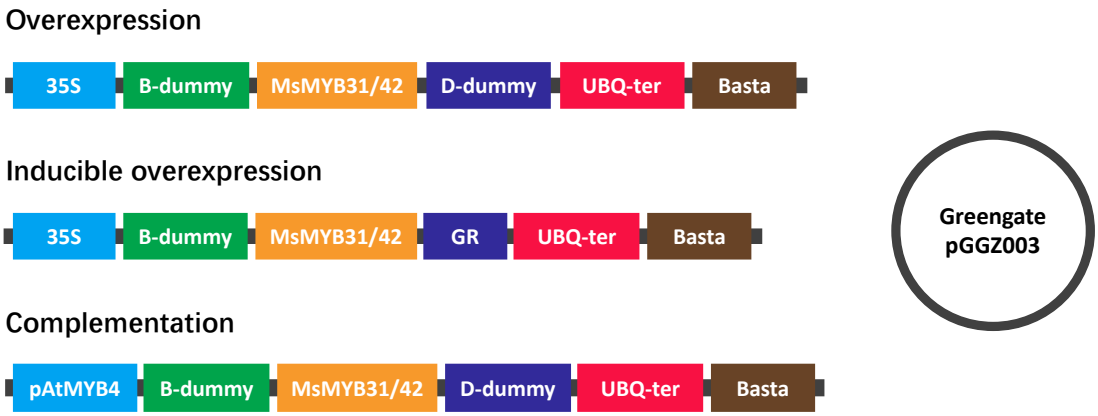

Figure S7 Schematic structure of Greengate constructs used for plant transformation.

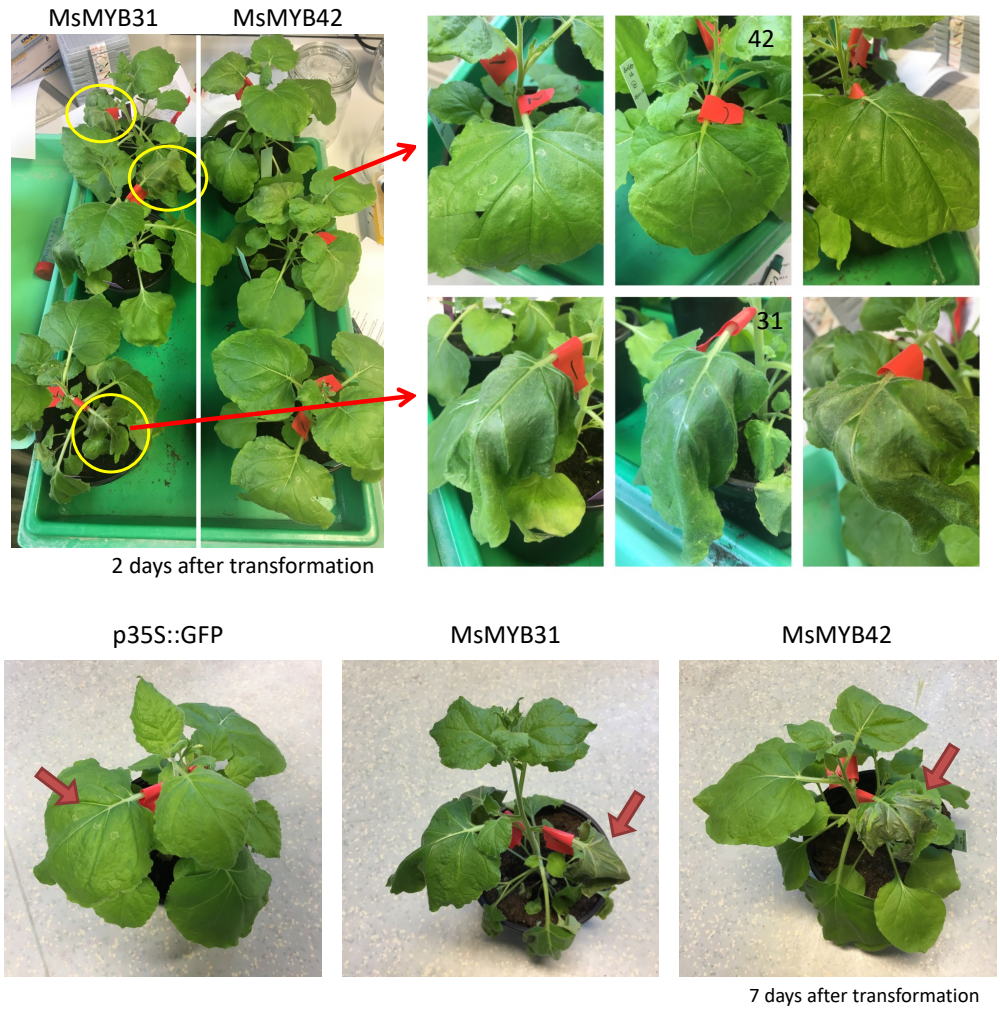

Figure S8 Phenotype of tobacco leaves transiently transformed with p35S:MsMYB31 and p35S:MYB42, respectively.

Pictures were taken two days or seven days after transformation. Yellow circles and red arrows indicate the infiltrated part of the leaves. Plant transformed with p35S:GFP was used as control.

Figure S9 Continued DEX induction of the inducible lines when selecting positive lines.
3-week-old plants of all the selected lines were sprayed with 30μM DEX once per day for three days to keep a high expression of the transformed protein in the nucleus. Red labels indicated MsMYB31 lines, yellow labels indicated MsMYB42 lines. After the treatment, plants of line MsMYB42-4 became pale and stopped growing.

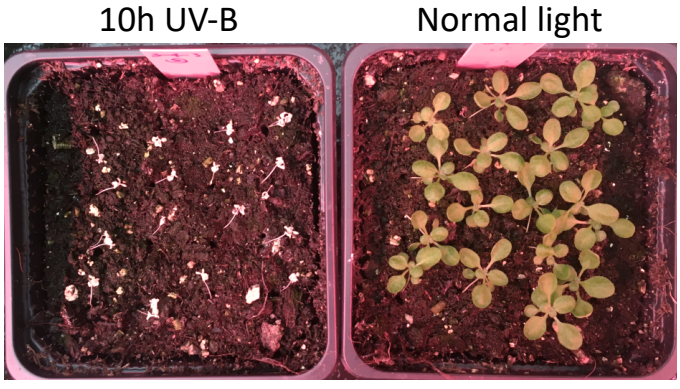

Figure S10 Example of seven days recovery for the UV treated seedlings (cp-MsMYB31-1).

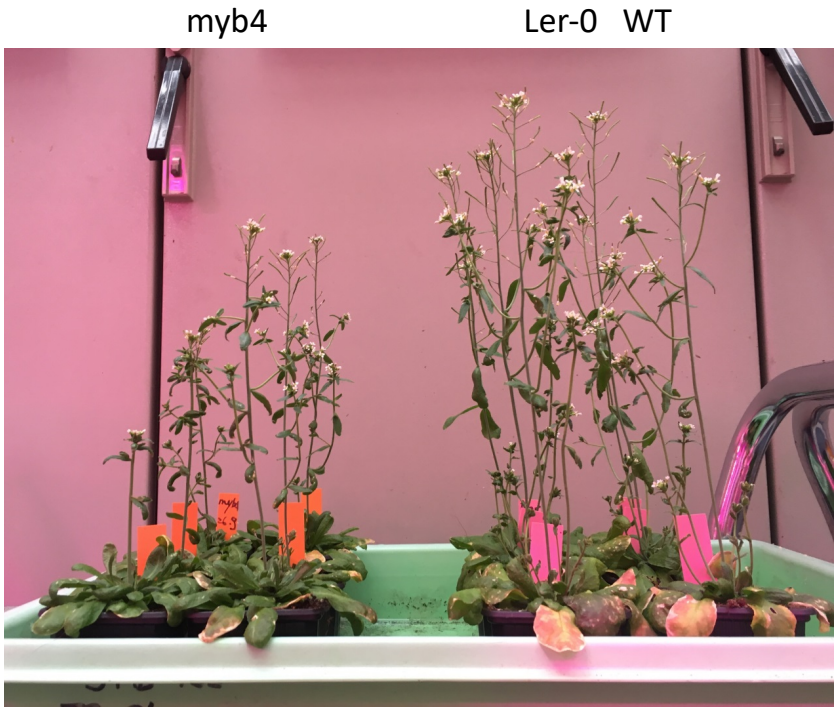

Figure S11 Phenotype of delayed stem development of myb4 compared to Ler-0 WT.

Figure S12 Phenotype of silique angles of six lines growing in the greenhouse condition.

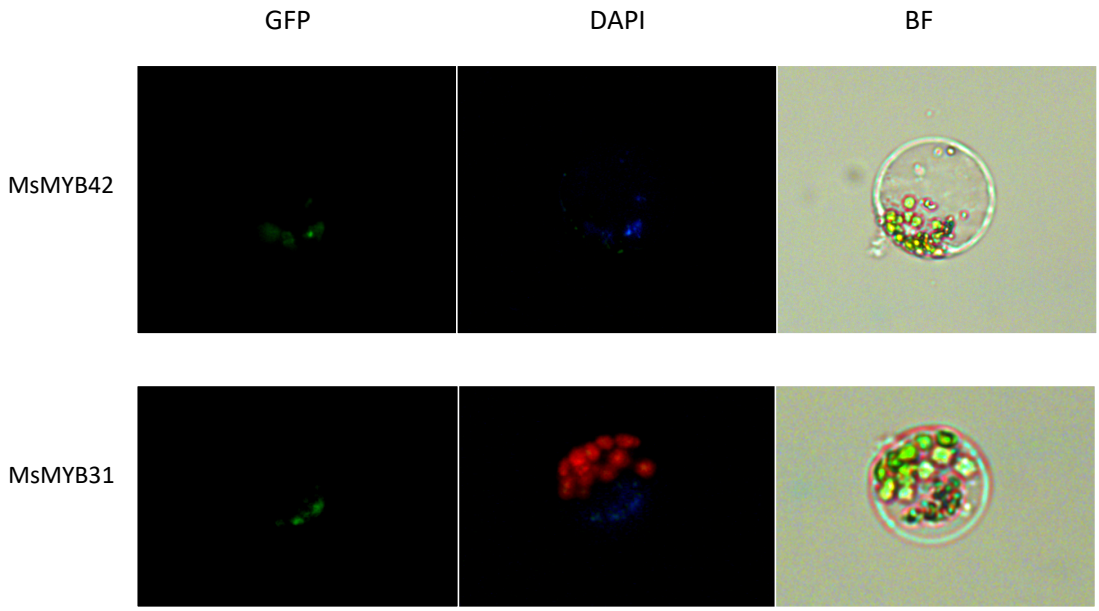

Figure S13 Subcellular localization of MsMYB31 and MsMYB42 in Sorghum protoplast. GFP signals showed the transient transformation of 35S:MsMYB31-GFP, 35S:MsMYB42-GFP, which is co-localized with DAPI staining (blue). The red color shows the autofluorescence of the chloroplasts without filtering out from the scene.

Table S1 Leaf gradient expression absolute values of MYB31 and MYB42 in Maize and rice.

Tissue	Expression Level		Tissue	Expression Level	
	ZmMYB31	ZmMYB42		OsMYB31	OsMYB42
M1	18.29	2.04	R1	151.35	35.16
M2	40.04	5.77	R2	141.34	29.57
M3	64.76	8.83	R3	29.55	16.00
M4	32.98	6.46	R4	8.62	16.01
M5	8.58	6.63	R5	2.46	19.77
M6	3.34	7.00	R6	1.40	16.18
M7	2.86	7.20	R7	0.45	14.18
M8	2.40	8.45	R7	0.45	14.18
M9	2.40	10.69	R8	1.29	14.68
M10	2.09	11.31	R9	2.17	13.85
M11	2.81	12.97	R10	0.82	9.46
M12	2.62	11.63	R11	1.96	13.68
M13	2.34	11.03			
M14	2.76	12.28			
M15	2.63	10.83			

Data obtained from

http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi

http://bar.utoronto.ca/efprice/cgi-bin/efpWeb.cgi

4 Discussion

Originated in Asia, *Miscanthus* is nowadays one of the promising energy crops widely applied in EU. The biological characteristics of *Miscanthus* enable it to provide considerable biomass and suitable for growing in lands that have relatively poor conditions. Lignin is the second largest components of the plant cell wall. For the development of plant itself, lignin can provide good protection for plant growth. On the contrary, lignin also brings inconvenience to the usage of biomass in the pretreatment procedures. An in-depth understanding of the secondary cell wall biosynthesis pathways in *Miscanthus*, especially the lignin metabolism pathway, may provide insight on how to generate biomass with better quality for more efficient utilization in the process of biomass energy generation. Lignin metabolism belongs to a branch of phenylpropanoid metabolic pathway, which provides many secondary metabolites for plants that play crucial roles in biological activities such as plant growth and development, defense to biotic and abiotic stresses. By studying the transcription factors that inhibit the pathway, we can better understand the transcriptional regulation activities of plants on phenylpropanoid metabolism. In addition, recent genome sequencing of *Miscanthus* has greatly helped to advance this research. In this study, two transcriptional repressors were identified in *Miscanthus*, naming MsMYB31 and MsMYB42 respectively, that regulate phenylpropanoid metabolic pathway and analyzed their target genes and their roles in plants.

The discussion chapter will mainly focus on the following topics: a) Sequence analysis as well as expression patterns analysis of MsMYB31 and MsMYB42. b) In-depth exploration into the inhibition mechanism of MsMYB31 and MsMYB42. c) The functional study of MsMYB31 and MsMYB42 in regulating phenylpropanoid metabolic pathway in plant, at both molecular and metabolism levels. d) The R2R3-MYB repressors of phenylpropanoid metabolic pathway in plant and their relationship between plant growth and development.

4.1 Identification of *Miscanthus* MYB31 and MYB42

Previous studies have shown that the plant R2R3-MYB proteins probably evolved from 3R-MYB ancestors by losing R1 repeat (Rabinowicz et al., 1999). This happened before the separation of monocots and dicots (Chaw et al., 2004), which may explain the conservative structure of subgroup 4 family R2R3-MYB repressors regulating the phenylpropanoid pathway. Despite the highly conserved R2R3 repeats similar to other R2R3-MYBs, all the members in subgroup 4 family harbor the C1 (lsrGIDPxT/NHR), C2 (pdLNLD/EL, EAR-repressor), and C3 (ZF-like) motifs. There is also a C4 (dFLGL and LDF/YRxLEMK) motif shared by most of the R2R3-MYB repressors in subgroup 4 family (Figure 3-1-1).

The C2 motif is also called EAR-repressor domain, presenting in a large variety of different transcription factor families such as ERF, bZIP, C2H2, Homobox and MYB families (Kagale et al., 2010). In *Arabidopsis*, the conserved pattern of EAR domain containing proteins shows in DLNxxP or LxLxL ways. For subgroup 4 family R2R3-MYB repressors, in dicots the protein of AtMYB4, VvMYB4a (Cavallini et al., 2015) and PdMYB221 (Tang et al., 2015), and in monocots the protein of ZmMYB31 and ZmMYB42 all possess the exact LxLxL pattern. Due to the highly conserved sequences, the C2 motif is used as a secondary query to narrow the scope of selection. Thus, seven contigs containing the exact “CPDLNLDL” pattern were selected out of the candidates (Table 3-1-1). Finally, two sequences were successfully amplified from *Miscanthus* cDNA, naming MsMYB31 and MsMYB42 respectively according to the sequence similarity to ZmMYB31 and ZmMYB42.

In monocots, MYB31 and MYB42 are involved dynamically together within the phenylpropanoid pathway regulatory network. Researches of the orthologs in different crops (maize, rice and sorghum) suggested that the two MYBs work as repressors and have both conserved and specialized repression activities across the three grasses (Agarwal et al., 2016). The identification of MsMYB31 and MsMYB42 also revealed the similar regulatory mechanism. Very similar sequences provide the possibility that

they function similarly with their orthologs (Figure S1). The expression patterns of MsMYB31 and MsMYB42 along the developmental leaf gradient suggest the division of work in regulation of the phenylpropanoid pathway (Figure 3-2-1). To understand the distinct spatiotemporal expression pattern, promoter studies of the two repressors may give further insights to the regulatory mechanisms.

4.2 MsMYB31 and MsMYB42 repress genes involved in phenylpropanoid pathway

Previous study shows that in *Arabidopsis*, C4H and CHS gene are repressed by overexpressing AtMYB4 (Jin et al., 2000). MsSCM4 is also known to activate the promoter of AtCHS (Golfier et al., 2017). Using 1500 bp of AtCHS promoter sequence as positive control, dual-luciferase reporter assays were conducted to test the regulation function of MsMYB31 and MsMYB42.

As expected, MsMYB31 and MsMYB42 strongly repressed the promoter of AtCHS, and also blocked the activation function of MsSCM4 to the promoter of AtCHS (Figure 3-3-3). Promoter sequences of *Miscanthus* phenylpropanoid pathway genes C4H, CCR, CAD and CHS were firstly analyzed with binding site prediction and then amplified for DLA (Table 3-3-1). MsMYB31 and MsMYB42 with *Miscanthus* C4H, CCR and CAD promoters revealed similar repression pattern to the positive control, while proMsCHS was not regulated by MsMYB31 or MsMYB42, and had relatively weak activation by MsSCM4. The de-activation of MsSCM4 towards proMsCHS by MsMYB31 or MsMYB42 remained similar to the other promoters. This no-effect result of MsMYB31 and MsMYB42 with proMsCHS could be explained by the lack of preferred binding sites, since the length of proMsCHS is only 788bp, shorter than all the others. And there are only three predicted binding sites (threshold 85%) on the cloned proMsCHS (Table 3-3-1). Besides, the complexity of *Miscanthus* gene isoforms is another possible explanation.

To date, there are more studies discussing about the transcriptional regulation by TFs of the flavonoids branch than lignin branch, and transcription activators have been studied more than repressors. The results presented have indicated that MsMYB31 and MsMYB42 act as repressors on regulating phenylpropanoid pathway genes, especially lignin biosynthesis genes. These two repressors can also block the function of MYB-type activators. Further research is needed to understand how repressors exercise inhibitory function.

Another interesting finding is that from the DLA results, along the monolignols biosynthesis pathway to lignification, the transcriptional regulation by MYBs (both repressors and activator) generally decreases in degree. In other words, the MYBs tested in the study have stronger effects on the upstream genes such as C4H and CCR, whereas the more downstream gene CAD and laccases (LAC) which focus on lignification (He et al., 2019) are less induced or inhibited by those MYBs. This may be due to the different upstream and downstream positions of different genes on the pathway that determine the effect intensity needed. But it may also because there are multiple isoforms of one gene on one position, in this experiment, the promoter of only one isoform for each gene was studied. Possible mechanisms behind this phenomenon has yet to be verified by further experiments.

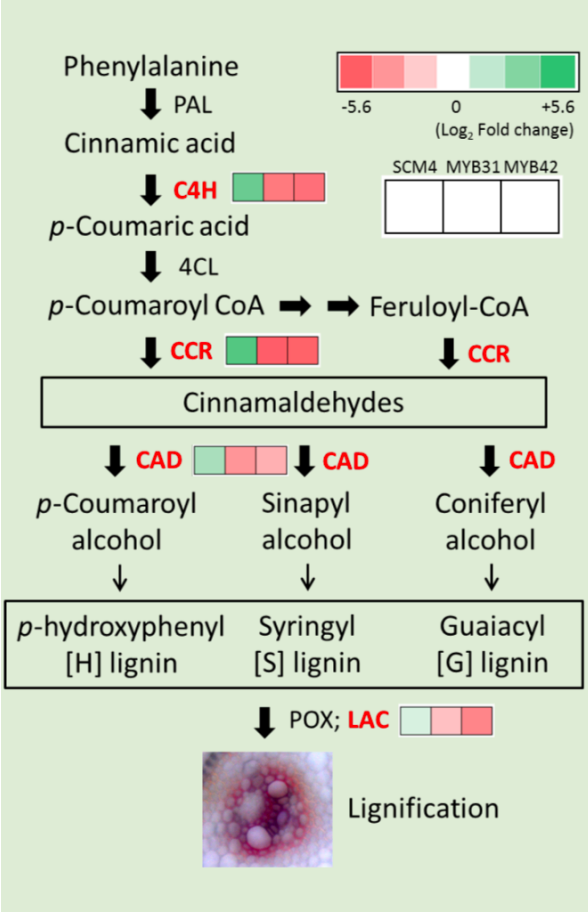

Figure 4-1 Regulation of lignin biosynthesis pathway genes by different R2R3-MYB transcription factors.

The schematic diagram displays the simplified phenylpropanoid-lignin biosynthesis pathway. Heat maps indicate log_2-transformed value of fold changes according to the Dual-luciferase-assay results. Green and red boxes represent up or down regulation by the MYB activator MsSCM4 or the repressors MsMYB31 and MsMYB42. PAL, Phenylalanine ammonia-lyase; C4H, Cinnamate 4-hydroxylase; 4CL, 4-Coumarate-CoA ligase; CCR, Cinnamoyl-CoA reductase; Cinnamyl-alcohol dehydrogenase; CAD, Cinnamyl-alcohol dehydrogenase; POX, Peroxidases; LAC, Laccases.

4.3 MsMYB31 and MsMYB42 proteins bind to AC-elements on the promoters

The R2R3-MYBs have been demonstrated to bind to motifs which are enriched in adenosine (A) and cytosine (C) residues, thus referred to as AC elements. Researches have been identified the consensus sequence pattern as ACC[A/T]A[A/C][T/C] using various methods such as cyclic amplification and selection of targets (CASTing) and systematic evolution of ligands by exponential enrichment (SELEX) ((Prouse and

Campbell, 2012, Fornale et al., 2010). Different R2R3-MYBs may have preferences towards different kinds of AC-elements, namely AC-I (ACCTACC), AC-II (ACCAACC), and AC-III (ACCTAAC), respectively. For instance, *Arabidopsis* MYB58 and MYB63 are able to bind to all the three AC elements (Zhou et al., 2009), while ZmMYB31 binds to AC-II, but not to AC-III (Fornale et al., 2010).

Furthermore, comprehensive studies of the binding motifs of NAC and MYB transcription factors involved in regulating secondary wall biosynthesis in plants have been conducted by Ye's group (Zhong et al., 2010a, Zhong et al., 2010b, Zhong et al., 2011, Zhong and Ye, 2011). These sequences are designated as secondary wall NAC-binding elements (SNBEs) and secondary wall MYB-responsive elements (SMREs). In the study, the 7-bps SMREs encompasses eight variants of sequences, including the three previously defined AC-elements that can be recognized by various MYB transcription factors (Zhong et al., 2013).

Confirming by EMSA, MsMYB31 and MsMYB42 also bind to the AC elements on C4H and CCR promoters (Figure 3-4-4). The probes used in the assays are selected from the promoter sequences of C4H and CCR. Both probes contain two closely located AC-elements (interval less than 10 bp), and from similar position upstream of the ATG (Table 3-4-1).

Interestingly, the interaction pattern between protein and probes of MsMYB31 and MsMYB42 revealed differently. As shown in figure 3-4-4A, protein NusA-6xHis-MsMYB31 interact with both the probes and form one significant band, while 6xHis-MsMYB42 form two bands with both the probes (Figure 3-4-4B). One possible explanation for the difference is according to the work of Zhong et al. (2013). In their study, they proposed a classification of eight variants of secondary wall MYB-responsive elements (SMREs) on the gene promoters, which can be differentially recognized by specific MYB transcription factors. In this study, the C4H probe used in EMSA contains one SMRE4 (ACCAACC) and one SMRE2 (ACCAACT) sequence, while that of CCR are one SMRE3 (ACCAAAC) and also one SMRE2 (Table 3-4-1). The two bands for MsMYB42 EMSA results suggested that it may recognize all the

three SMREs (SMRE2\3\4). And MsMYB31 might have preference towards these SMREs and just not tending to bind SMRE2 sequence.

Another hypothesis to explain the different band patterns is the size difference of the recombinant proteins. There is a huge NusA-tag connected to the N-terminus of the MsMYB31 protein. 6xHis-tagged MsMYB31 is insoluble thus cannot be extracted and purified through columns (Figure 3-4-2). N-utilization substance A (NusA) tag is reported to largely increase the solubility, linking with an extra affinity tag, it is still available for further protein purification through columns (Gopal and Kumar, 2013). Following with the increase of solubility, the molecular weight of the recombinant protein is more than twice as previous. Taking into account the distance limitation between the two predicted binding sites on the probes, two proteins may not bind to one probe at the same time due to the oversized volumes. The size of MsMYB42 with only 6xHis-tag is much smaller than that of NusA-6xHis-MsMYB31, resulting in doubled bands on the EMSA gel (Figure 3-4-4).

In addition, the DNA binding capacity of the protein of a transcription activator, MsSCM4, was also tested via EMSA (Figure S6). It also showed the double-bands pattern similar to MsMYB42 but with different shift distance. The exquisite design that let both MsMYB42 and MsSCM4 proteins incubate together with the probe suggested the competitive binding of DNA and different proteins (Figure 3-4-5). Figure 4-1 proposes a simplified model of different interactions describing the four bands appeared in the column.

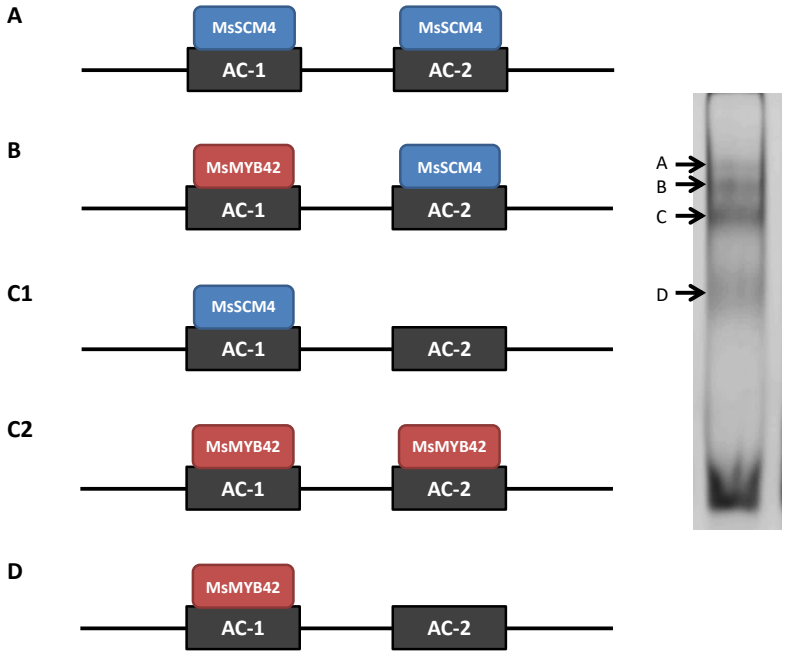

Figure 4-2 A hypothesized model for the protein-DNA interaction of MsMYB42 and MsSCM4 with EMSA probes

Black arrows point out the interaction bands on the EMSA gel and the proposed model explains the possible interactions according to the order. EMSA figure is cut out from figure 3-4-5. Patterns from A to D in turn represent the four bands pointing out by black arrows from top to bottom. Because the two proteins are different in size, band A is suggested to be one probe binding by two MsSCM4 proteins that has the slowest mobility. Following with band A, band B should be probe with one MsMYB42 and one MsSCM4 protein. Comprised by two possible patterns, band C is either probe with one single MsSCM4 protein, or with two MsMYB42 proteins. Band D is the lowest, suggested to be probe with only one MsMYB42 protein.

Taking all together, the EMSA results provide the possibility of visualization of DNA-protein interactions, which helps us with better understanding of the function of the TFs. The *Miscanthus* MYB activator MsSCM4 and repressors MsMYB31/42 recognize same motifs thereby regulate gene expression at transcript level. This gives hints to one of the de-activation patterns that repressors can compete binding to the same motifs on the promoter with activators, resulting in blocking the interaction of activators and cofactors to form the protein complex.

4.4 The function of MsMYB31 and MsMYB42 in regulating phenylpropanoid metabolic pathway

4.4.1 Functional studies at molecular level

For functional studies of transcription factors, it is important to find their target genes. As mentioned earlier, methods such as DLA, EMSA and transgenic strategies have been used to identify target genes. Along *Miscanthus* lignin biosynthesis pathway from upstream to downstream, the promoters of three C4H CCR and CAD genes can be suppressed by MsMYB31 and MsMYB42 in vivo. At the same time, the activation effect of MsSCM4 can also be inhibited by them (Figure3-3-4 and 3-3-5). This indicates that inhibitors can strongly hinder the initiation of transcription of these genes. EMSA results further illustrates the direct binding of MsMYB31 and MsMYB42 proteins and the DNA. In the *Arabidopsis* transgenic lines, inducible overexpression of these two genes can inhibit most of the enzymes along the phenylpropanoid pathway to different degrees (Figure 3-5-2). This might include both direct targets and indirect targets that are regulated by downstream regulators. What's more, there are many different factors influencing gene expression temporally and spatially. It is not comprehensive enough to draw conclusions to the regulatory function only from results in a certain growth period. Functional studies should also focus on metabolite levels.

In addition, due to the limitation of the whole genome sequencing and time constraints, no more *Miscanthus* promoter sequences have been successfully cloned, including promoters of other transcriptional activators and promoters of repressors themselves. Agarwal et al. have shown that in maize, rice and sorghum, MYB31 and MYB42 participate in auto- and cross-regulation in all three species (Agarwal et al., 2016). The promoter of AtMYB4 also possesses putative MYB4 motifs and is self-downregulated (Zhao and Dixon, 2011, Mitra et al., 2019). These mechanisms reasonably explain the realization of the feedback function in the whole process of metabolite accumulation.

Attention should be paid to the interaction between proteins and DNAs, including self-regulation functions in the upcoming studies.

4.4.2 Functional studies at metabolic level

To study the functions of these two transcription factors from the perspective of metabolites can give us better understanding to their inhibitory effects on the phenylpropanoid metabolic pathway.

In the branch of lignin metabolism, the inhibition function is assessed in both induced overexpression system and complementary system in *Arabidopsis*. Firstly, in the inducible overexpression system, at the inflorescence stem development stage continuous, high expression of MsMYB42 in Col-0 plants limits the biosynthesis of lignin (Figure 3-5-2), and may even affect the expression of other genes related to secondary cell walls, resulting in plant height differences compared to the wild type (Figure 3-5-4). Similar results have been found in other repressor overexpression systems. For instance, ectopic overexpression of PvMYB4 in switchgrass resulted in reduced lignin content and increased sugar release efficiency from cell wall residues (Shen et al., 2012). *Arabidopsis* plants overexpressing maize ZmMYB31 and ZmMYB42 had shorter and weaker stems, less lignin contents and the cell wall components have changed (Fornale et al., 2006, Sonbol et al., 2009, Fornale et al., 2010). Besides, in our complementary system, in the myb4 knock-out background, MsMYB31 and MsMYB42 were expressed driven by native AtMYB4 promoter. Compared to Ler-0, lignin content in cp-MsMYB31 and cp-MsMYB42 stems decreased by 10% to 20% (Figure 3-6-4). MsMYB31 had stronger impact than that of MsMYB42, which was consistent with the results found in tobacco leaf infiltration (Figure S8). These findings together with the gene expression tests (Figure 3-5-2 and 3-6-3) suggest that MsMYB31 and MsMYB42 can inhibit lignin metabolism.

Additionally, there was no significant difference in cell wall composition between the myb4 mutant and Ler-0 WT (Figure 3-6-4, Table 3-6-2). This also suggests that there are other mechanisms that complement MYB4 deficiency in *Arabidopsis* for the plants

to maintain normal growth and development. However, in the cp-MsMYB31 and cp-MsMYB42 transgenic lines, the introduction of MsMYB31 or MsMYB42 changed the cell wall composition (Figure 3-6-4, Table 3-6-2), which means the two repressors from *Miscanthus* cannot fully complement AtMYB4 function. Although closely related, probably due to the mechanism differences between monocot and dicot plants, the functions of AtMYB4 and MsMYB31/42 are not exactly the same. They may have different target genes and regulatory patterns, as well as being regulated by different other factors. So far, a stable *Miscanthus* transformation system has not been established in the lab, which also limits the research on the function of transcription factors of *Miscanthus in vivo*. The model plant *Arabidopsis* provides us with preliminary clues, and further studies would be carried out in *Miscanthus*.

The metabolic pathway of flavonoids is also worth exploring. Unlike lignin, which is a complex structure formed by polymerization of lignin monomers, the secondary metabolites on flavonoids branches are mostly small molecules that have strong biological activities and are beneficial to plants and even human activities. Moreover, these components possess characteristics such as having colors or aroma, which are convenient for detection and analysis. The research on flavonoids metabolism is far ahead of that on lignin metabolism. For instance, the accumulation of flavonols can help plants to resist UV stress in the environment. Under UVB treatment, flavonols in myb4 mutants accumulated more than that in the wild type, which makes myb4 mutant better at dealing with ultraviolet irradiation (Höll et al., 2019). On the other hand, MsMYB31 and MsMYB42 complementation lines were more sensitive to UV than the wild type, suggesting the stronger inhibition of MsMYB31 and MsMYB42 on the accumulation of flavonols than AtMYB4 (Figure 3-6-2).

This result was also verified in the inducible overexpression lines. As shown in figure 4-3, after 6 hours of UV treatment, the accumulation of flavonols in rosette leaves was compared between MsMYB31/42 lines and Col-0 wild type plants. It is shown that after spraying DEX in advance and followed with 6 hours placement, the flavonol accumulation in the inducible lines was already slightly less than the wild type without

ultraviolet irradiation. Flavonol contents increased in all lines after 6 hours uv treatment, however, the accumulation of flavonols in the induced expression lines were less than those in the wild type, and the deficiency of quercetin was the most obvious.

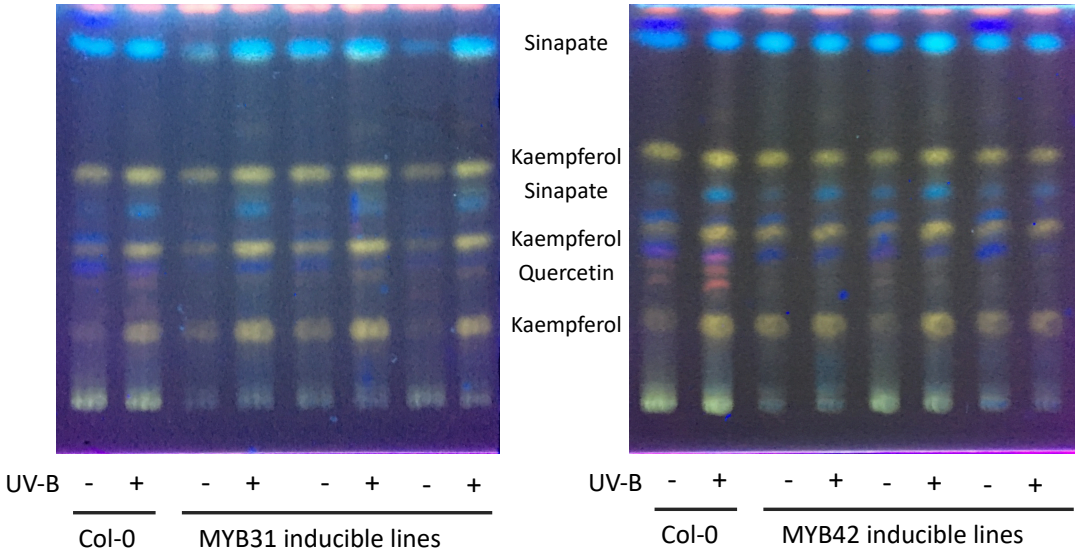

Figure 4-3 Thin layer chromatography pictures showing the flavonols content extracted from different *Arabidopsis* rosette leaf samples.

Three-week-old plants were used for the treatments, the transgenic expression in nucleus in all the plants were induced by spraying with DEX before putting under different irradiation conditions. Samples were collected after 6 hours of UV/control treatment.

Corresponding to the thin layer chromatography results, in the UV exposure treatment, quantitative RT-PCR analysis of flavonol biosynthesis genes also showed the repression function of MsMYB31 and MsMYB42 (Figure 3-5-3). For AtCHS and AtFLS, which are more promoting total flavonol accumulation, the reaction towards UV irradiation is much faster than that of AtF3'H, which is more related to composition changes (from kaempferol to quercetin). Compared to the gene expression in Col-0, in the MsMYB31 and MsMYB42 inducible lines, these genes all had less expression. However, with the irradiation treatment getting longer, at the time point of 10 hours, MsMYB31 and MsMYB42 started to behave differently. In MsMYB42 inducible lines, especially in MsMYB42-1 line, the expression of related genes began to rise, while MsMYB31 line remained relatively strong repression function. This phenomenon also suggests functional differences of the two repressors.

To sum up, the two repressors do not only specifically regulate one direction along the phenylpropanoid metabolic pathway, instead, they have inhibition function for both branches, revealed at both molecular and metabolic levels. Unlike many reported activators that specifically regulate the biosynthesis of one single product, these two inhibitors exhibit a wider range of repression functions. Although they seem to share many of the targets, MYB31 and MYB42 play different roles in plant. In particular, considering the completely different expression patterns of MsMYB31 and MsMYB42 along *Miscanthus* leaf gradient, MYB31 was expressed more in developing tissues, while MYB42 maintained a high expression level in all segments except in the base.

4.5 The number of R2R3-MYB repressors of phenylpropanoid metabolic pathway in plant, small but decisive

It is probably no coincidence that the number of MYB activators is much larger than the number of inhibitors. The R2R3-MYB activators, together with other transcription factors such as NACs, form a hierarchical regulatory pattern in the well described secondary cell wall transcriptional regulatory network (Nakano et al., 2015). Each MYB sits in a specific position in this network and performs its function by activating downstream gene expression and/or by being regulated by upstream genes. This regulatory network in *Miscanthus* is also partially elucidated. This study enriched the part of repressor function in the transcriptional regulatory network of *Miscanthus* secondary cell wall. Together with the knowledge in other plants, it is clear that the number of inhibitors in the network is small, while many activators have redundant functions to regulate the same target genes (Wang and Dixon, 2012, Hirano et al., 2013, Zhong et al., 2013, Zhong and Ye, 2014).

In looking for possible explanations, how transcription started gives us some clues. The transcriptional process takes place under the interaction of many trans and cis factors. In order to complete the final gene expression process, an essential protein complex called transcription initiation complex starts to form. The accurate formation of

transcription complexes determines the efficiency of transcription. The number of activators should be sufficient to ensure and fine-tune the process. The transcriptional repression is exactly the opposite, as long as any component of the transcription complex is blocked from gathering together, the process would be inhibited.

4.6 One possible explanation for the delayed inflorescence stem development of myb4 mutant - the inhibition of auxin transportation by flavonoids accumulation

Arabidopsis inflorescence stems grow rapidly in suitable environmental conditions. In the study, the inflorescence stems of myb4 mutant showed a delayed growth phenotype in the early stages compared to the wild type when growing in the greenhouse. Many studies have been done on the research of MYB4 in *Arabidopsis* (Jin et al., 2000, Zhao et al., 2007, Mitra et al., 2019), however, there are few reports on delayed growth phenotypes. This may be due to the fact that this is a relatively short-term phenotype throughout the entire developmental stages of *Arabidopsis*, and soon the mutant will catch up with the wild type in the height of inflorescence stems. Previous studies on myb4 mutant have focused more on its relationship with UV resistance and its effect on flavonoids metabolites (Höll et al., 2019). Much of the research has been done with the rosette leaves but not the stems, and this phenotype can only be seen when the inflorescence stem is growing. Moreover, the appearance of this phenotype also depends on the light and temperature conditions. In the greenhouse, this phenotype has been observed for many times. But in the long-day climate chamber, due to different light sources and the accurately controlled temperature, which is relatively lower than the temperature observed in the greenhouse, the delayed growth phenotype does not exist anymore. The possible reasons are described below.

Polar transport of auxin regulates the elongation of plant organs. Many studies have shown that flavonoids act as endogenous repressors of auxin transport (Brown et al., 2001, Besseau et al., 2007, Peer and Murphy, 2007). In myb4 mutant, because of

lacking the important repressor, flavonoids accumulate to higher level in the process of growth and development thus may affect the distribution of auxin. However, with the continuous growth and development of plants, the accumulated flavonoids will be reduced due to the existence of other complementary regulation factors. This allows the mutant to gradually reduce the height difference compared to the wild type. Additionally, temperature and light quantity and quality also affect the synthesis and transportation of auxin and flavonoids. Previous studies on the elongation of *Arabidopsis* hypocotyl showed that high temperature promotes auxin-mediated hypocotyl elongation by improving the auxin content (Gray et al., 1998). In the lab condition, red and blue light constitute the supplementary light in the case of insufficient light intensity in the greenhouse. Cryptochromes are photoreceptors that recognize blue light and regulate growth and developmental activities in plants (Pashkovskiy et al., 2016, Yang et al., 2017b). They can be activated by blue light and promote the accumulation of certain flavonoids such as anthocyanins. Besides, MYB4 is also found in many light-dependent regulatory pathways (Azuma et al., 2012, Zhang et al., 2014). In the climate chambers, in addition to the absence of extra blue light, the temperature in the incubator was kept at a lower level. In general, in this case, the relatively high temperature in the greenhouse and the conditions of supplementary light source made it easier to detect the phenotype of delayed inflorescence stem growth. And the conditions might cause this phenotype were weakened in the climate chamber.

4.7 Perspectives

In this study, the identification and characterization of two R2R3-MYB transcription repressors, MsMYB31 and MsMYB42 provides broader perspectives for understanding the negative transcriptional regulation of lignin biosynthesis and the phenylpropanoid pathway in *Miscanthus*. The repression function of MsMYB31 and MsMYB42 was confirmed and the possible physiological functions of both repressors were discussed.

The complementation studies revealed the similar but different mechanisms of MsMYB31, MsMYB42 and AtMYB4, which requires more in-depth analysis for exploring the differences and associations of homologous transcription factors in monocot and dicot plants. In addition, the target redundancy and the differential developmental expression patterns of MsMYB31 and MsMYB42 gives hints to specific regulatory functions of the two repressors, more studies can be carried out for further functional analysis.

Furthermore, the potential targets genes of MsMYB31 and MsMYB42 provide promising breeding strategies for the genetic development of *Miscanthus*, for meeting the increasing demands of better sustainable resources.

5 Materials and Methods

5.1 Bioinformatics

5.1.1 *Miscanthus* EST database

The gene identification in *Miscanthus* is based on the tBLASTn search against the published *Miscanthus* EST database, in which the transcriptomes were obtained from eight *Miscanthus*. spp. For the identification of MsMYB31 and MsMYB42, genes used as templates are AtMYB4 (AT4G38620, https://www.*Arabidopsis*.org), ZmMYB31 (GRMZM2G050305) and ZmMYB42 (GRMZM2G419239, https://www.maizegdb.org).

5.1.2 *Miscanthus* genome databases

For promoter analysis, we generated a partially sequenced genomic DNA database from M. sinenses var. Sin-13 cooperating with the Deep-Sequencing-Core Facility on Heidelberg Campus. All the promoter sequences were identified firstly using an around 100bp 5' UTR sequence of the gene as query to do blast search in the genomic DNA database. The upstream 100bp of the targeted contigs were used as new queries to blast again for prolongation of the possible promoter sequences. At last the spliced promoter sequence with sufficient length would be used as the last query for double checking.
In December 20th, 2017, the first chromosome-scale assembly of *Miscanthus sinensis* doubled haploid DH1 (IGR-2011-001) had been released (***Miscanthus sinensis*** v7.1 DOE-JGI, http://phytozome.jgi.doe.gov/). All the *Miscanthus* promoter sequences obtained using the previous methods were checked again in this database.

5.2 Plant materials

5.2.1 *Miscanthus*

Seeds of *Miscanthus sinensis* (var. Sin-13) collected in Honshu, Japan (Clifton-Brown & Lewandowski, 2002) were used for all the experiments.

On soil: Plants were grown on standardized soil ED-73 from seeds in the greenhouse under 8/16-hr (light/dark) photoperiod at 22-25 degrees until certain life point for harvesting.

In sterile glass cans: Sterile *Miscanthus* shoots were grown from seeds in autoclaved glass cans, MS medium was used as base.

MS medium (for *Miscanthus*)

Murashige & Skoog with vitamins (Duchefa) 4.56 g/L

Sucrose (Roth) 30 g/L

BAP 100 µl/L

Kin 100 µl/L

NAA 100 µl/L

Gelrite 3 g/L

pH5.8 (KOH)

5.2.2 *Arabidopsis*

On soil: In addition to standardized soil ED-73, xxxxxx were added for growing *Arabidopsis*. After vernalization the pots were placing in the greenhouse under 8/16-hr (light/dark) photoperiod at 22-25 degrees.

On plates:

Sterilization of *Arabidopsis* thalinana seeds:

For 1.5 ml Eppendorf tubes, the amount of the seeds should be no more than 0.5 ml. Fill the tube up with a premixed solution of 5ml sodium hypochlorite(12%) and 45ml 70% ethanol, mix well, let the seeds sink to the ground and discarding supernatant. Then wash the seeds two times with absolute ethanol (fill up and discard). After the last

time washing, remove the supernatant and flip the tube to distribute seeds on the tube wall. For drying the seeds, open lid and leave the tube overnight, diagonal lying in a petri dish, in the sterile bench. The seeds then are ready for planting onto ½ MS agar plates.

½ MS medium:

2.2 g/L Murashige & Skoog with vitamins (Duchefa)

3 % (w/v) Sucrose (Roth)

0.8 % (w/v) Agar (Duchefa)

pH 5.7 (KOH)

5.2.3 *Nicotiana benthamiana*

On soil: Plants were grown on standardized soil ED-73 from seeds in the greenhouse under continuous light at 22-25 degrees until certain life point for treatments (e.g. 4-week-old for leaf infiltration).

5.2.4 Grapevine suspension cell culture

The suspension culture from grapevine cells developed from callus was used for dual luciferase assays in this study. 50 ml GC medium was used for cultivation in 300 ml flasks. Subculture was prepared once per week with transferring 15 ml old culture into new 50 ml GC medium and then cultivated on the shaker in dark at 100 rpm, 25 °C.

GC for suspension cell culture:

3.2 g/L Gamborg B5

30 g/L Sucrose

0.25 g/L Caseinhydrolysate

94 µl/L Kin

54 µl/L NAA

pH 5.7-5.8 (NaOH)

5.3 Bacteria related techniques

5.3.1 Bacteria strains

All the bacteria strains used in this study are listed in table 5-1. Strains XL1-Blue MR, TOP10 and DB3.1 were used for GreenGate or Gateway cloning; BL21(DE3) for inducible protein expression in E. Coli; A. tumefaciens strains ASE and C58C1 were used for transient or stable transformation in plants.

Table 5-1 Description of bacterial strains used in this study.

Sourse	Strain name	Genotype description
E. Coli	XL1-Blue MR	*Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac*
	DB3.1	*gyrA462 endA1 Δ(sr1-recA) mcrB mrr hsdS20 glnV44 (=supE44) ara14 galK2 lacY1 proA2 rpsL20 xyl5 leuB6 mtl1*
	BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*
	TOP10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG*
A. tumefaciens	ASE (pSOUP+)	Chlr Tetr Kanr, pSOUP + helper plasmid that confers replicase activity for pSa replication origin on pGreen-derived plasmids.
	C58C1 (P14)	Rifr Tetr Kanr, 35S:p14 silencing suppressor from Pothos latent virus (PoLV).

5.3.2 Cultivation of bacteria

Both E. Coli and A. tumefaciens were cultivated using low salt LB medium with appropriate antibiotics in 37 °C (E. Coli) or 28 °C (A. tumefaciens and protein expression in E. Coli).

Low salt LB medium:

5 g/L NaCl

10 g/L Trypton/Pepton

5 g/L Yeast extract

pH 7

Table 5-2 List of antibiotics

Antibiotic	Stock conc.	Solvent	Final conc.
Ampicillin	100 mg/ml	Water	100 µg/ml
Carbenicillin	50 mg/ml	50% Ethanol	50 µg/ml
Chloramphenicol	34 mg/ml	Ethanol	34 µg/ml
Kanamycin	50 mg/ml	Water	50 µg/ml
Rifampicin	25 mg/ml	100% Methanol	25 µg/ml
Spectinomycin	50 mg/ml	Water	50 µg/ml
Tetracyclin	12.5 mg/ml	70% Ethanol	12.5 µg/ml

5.3.3 Transformation of bacteira

5.3.3.1 Preparation of chemically E. Coli competent cells

The production of competent E. coli cells was achieved using the Inoue method by Sambrook and Russell, 2006. Firstly, an overnight pre-culture from single colony was prepared, then expanded in flask with a 1/10 ratio until the OD600 reached 0.6. After

cooling down, the culture was then centrifuged for 10 min, 2,500 g at 4 °C. For each 250 ml culture, 100 ml pre-cooled Inoue buffer was used for careful resuspension. Hereafter the culture resuspended in 100 ml Inoue buffer was centrifuged again as the same condition above. Then a final 20 ml Inoue buffer plus 1.5 ml of DMSO was used for resuspension. The cell suspension was aliquoted into 50 µl aliquots, snap frozen in liquid nitrogen and stored at −80 °C until use.

100 ml 0.5 M PIPES (pH 6.7)

Dissolve 15g PIPES in ~80 ml of water

Adjust pH to 6.7 with KOH

Add water until 100 ml

Filter sterilize

Inoue buffer

Dissolve in 980 ml of water:

10.9 g MnCl2

2.2 g CaCl2

18.65 g KCl

Add 20 ml of 0.5 M PIPES buffer (above)

For each time, prepare fresh Inoue buffer with filter sterilization before use.

5.3.3.2 Chemically E. Coli transformation

Prior to the transformation steps, the 50 µl aliquots in Eppendorf tubes should be thawed on ice for 10 minutes. After all melted, add 1 pg-100 ng of plasmid DNA (1-5 to cells and mix without vortexing. Place the mix on ice for 30 minutes. Then the mixture was ready for a heat shock at 42 °C for 30 seconds. Place on ice for 5 minutes and then add 950 µl of room temperature SOC. Place the tube at 37 °C for 60 minutes. Shake vigorously (250 rpm) or rotate. After recovery, spread 50-100 µl of each dilution onto pre-warmed selection plates and incubate at 37°C overnight.

5.3.3.3 Chemically A. tumefaciens. transformation

Thaw the 50 µl aliquots in Eppendorf tubes on ice for 10 minutes. Add 1µg DNA to cells and mix by pipetting. Place on ice for 10 minutes. Put the Eppendorf tubes in liquid nitrogen for flash freezing for 5 min. Heat-shock at 37 ℃ for 5 min, following with placing on ice for 5 minutes. Add 950 ul of room temperature SOC to shake vigorously (250 rpm) at 28°C for 1.5 to 2 h. Cells were then ready to spread onto pre-warmed selection plates and incubate at 28°C for 2 days.

SOC medium

0.5 % (w/v) yeast extract (Roth)

2 % (w/v) tryptone (BD Biosciences)

10 mM NaCl

10 mM MgCl2

10 mM MgSO4

2.5 mM KCl

20 mM Glucose

pH 7.0 (NaOH)

5.3.4 Plant transformation

5.3.4.1 Floral dip*

Plants: The plants need about 4-7 weeks from seeding to “dipping”. Sow in a big pot, prick 4 plants per big pot, to get more shoots and buds cut the first shoot about 1-2cm above the plant. After that it takes approx. one week up to “dipping”.

Pre culture: Inoculate 3-5ml LB including the relevant antibiotics per 50ml tube. Shake at 28°C overnight.

Main culture: Inoculate 3-5ml LB including the relevant antibiotics per 50ml tube with a few drops of the pre culture. Depending on the quantity of buds and pots, you need about 6-8ml for 1 floral dip (6-8 pots). Shake at 28°C overnight.

Collect the equal cultures in one tube. Pipet 1,9-2ml culture in each 2ml Eppi (3-4 tubes per floral dip). Centrifuge 10 minutes at 3000g. Pour of supernatant of and add the equal amount of Dip Medium and vortex. The culture should have about OD 2 when harvesting.

"Dipping" procedure: Remove existing siliques. Use a disposable 6ml Transferpipette from Sarstedt and pipet the agrobacteria on the buds. Cover the tray over night with a lid so that most of the buds do not stick to this lid and put out of direct light. Next day remove the lid and put back to light. Redo the "dipping" procedure after approximately one week to increase the efficiency.

Dipping Medium (50 ml):

500 µl 1x MS/MES Mix

5 g Sucrose

0.25 µl 0.05%

Silwet100x MS/MES Mix (50 ml):

100mg MS

2.24g MES

pH 5.7

*The method is adapted from Heike Steininger's protocol.

5.3.4.2 Transient transformation of tobacco leaves

4-week-old *N. benthamiana* plants were used for transient transformation. Single colonies of A. tumefaciens were inoculated in 4 ml LB to prepare a pre-culture. Then an expand culture (20 ml) was inoculated in with the pre-culture. The cells were harvested by centrifuge and were resuspended in Agro buffer. The final suspension volume was 15 ml with 0.1 OD600 of P14 cells and 0.4 OD600 of other cells. Afterwards the suspension was incubated at RT for 2 h before infiltration was carried with a 2ml sterile needleless syringe. The infiltrated leaf was harvested after 2 days.

5.3.4.3 PEG transformation of *M. sinensis* protoplasts

The extraction of *M. sinensis* protoplasts was conducted following the protocol from Xinhui Shi's doctoral thesis. For PEG transformation, 10 µg of each plasmid was added into 100 µl of protoplasts (Protoplasts should have a concentration of $1x10^6$ protoplasts/ml) and mixed carefully. Then an equal volume of PEG/Ca solution was added in the protoplasts. After mixing well, it was incubated at RT for 3-30 min. The mix was then diluted with 0.44 ml W5 solution following with spinning at 100 g for 1 min to remove PEG (leave ~100 µl). Finally the mix was diluted with 1 ml W5 solution and placed in dark for 18-22 h until efficiency checking steps.

5.4 Nucleic acids related techniques

5.4.1 Extraction of DNA

5.4.1.1 Isolation of gDNA from plant tissues (CTAB)

For 100 mg homogenized tissue (fresh or frozen in liquid nitrogen), 500 µl 1.5x CTAB extraction buffer was used for mixing thoroughly via vortexing. The homogenate was then incubated in a heating block at 65 ℃ for 20-30 min, with a mix of the tube upside down every 5 min during the incubation. Then an equal volume of chloroform : isoamyl alcohol (24:1) was added and the tube was mixed continuously on a shaker for 10 min. Hereafter, the mix was centrifuged at 10000 rpm for 10 min at room temperature and then the supernatant was transferred to a new Eppendorf tube. 2X volume pre-cooled absolute ethanol was added to the supernatant, the tube was put in -20 ℃ for 20 min following a centrifugation at 12000 rpm for 15 min at room temperature. After discarding the supernatant, the pellet was washed with 500 µl 70% ethanol twice and then put in the clean bench with open lid for drying. The dry pellet was dissolved using 40-100 µl ddH2O or TE buffer.

5.4.1.2 Extraction of plasmid DNA

Plasmids used in the study were extracted from bacteria cultures with GeneJET Plasmid Miniprep Kit (Thermo Scientific, 2-4 ml) or JETSTAR2.0 Midi Kit (Genomed, 50ml) according to the manufacturer's instructions.

5.4.2 Extraction of RNA

The GeneMATRIX Universal RNA Purification Kit (EURx) was used for the isolation of RNA from plant tissues according to the manufacturer's instructions.

5.4.3 DNA amplification via PCR

The polymerase chain reaction (PCR) was performed using genomic DNA, complementary DNA or plasmid DNA as template for amplification of DNA fragments. DNA polymerases used in this study are Taq DNA Polymerase (Sigma-Aldrich), JumpStart™ Taq DNA Polymerase (Sigma-Aldrich), Q5® High-Fidelity DNA Polymerase (NEB) and KOD hot start DNA polymerase (Toyobo). PCR reaction procedures were performed according to the protocols from each manufacturer.

5.4.4 Cloning methods

5.4.4.1 GreenGate cloning

GreenGate cloning is an ideal tool to generate specific constructs with optional modules. In this study, all the constructs used for *Arabidopsis* floral dip and tobacco leaf infiltration were made from GreenGate cloning method. Figure 5-1 shows a brief description of the procedures, protocol in detail see Lampropoulos et al., 2013.

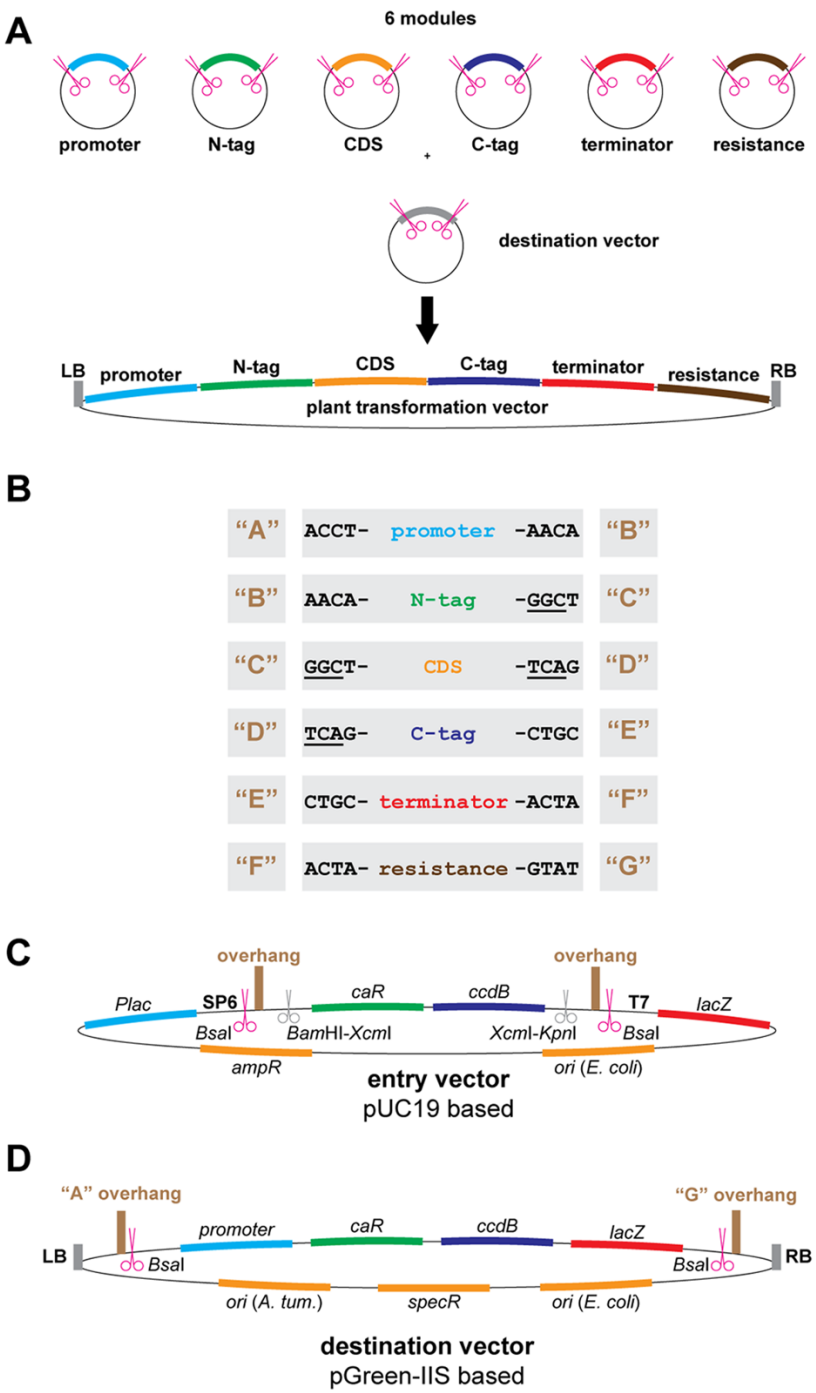

Figure 5-1 GreenGate vector design and layout. *

A) The GreenGate cloning system uses six different types of pUC19 based entry vectors into which the individual elements are inserted and a pGreen-IIS based destination vector. Magenta scissors represent BsaI recognition sites. In each GreenGate reaction, six modules are ligated between the left border (LB) and the right border (RB) sequences of the destination vector yielding a ready-to-use plant transformation vector with expression unit and resistance cassette. These six modules encompass a plant promoter, an N-terminal tag, a coding sequence (i.e. the gene of interest), a C-terminal tag, a plant terminator and a plant resistance cassette for selection of transgenic plants. The modules can only be ligated in the pre-defined order. B) The orderly assembly is enabled by a set of seven different overhangs. Each module is flanked at its 5′-end by the same overhang as the 3′-end of its preceding neighbor. The individual

overhangs all differ from each other by at least two out of the four nucleotides. The underlined nucleotides define coding triplets to which all other coding elements have to be in frame. C) Empty entry vector. The multiple cloning site of pUC19 has been replaced by two BsaI recognition sites (magenta scissors), the respective overhangs for each module type and a counter-selectable ccdB gene. DNA fragments can be cloned via the specific overhangs, via the BamHI and KpnI sites or via A-overhangs after XcmI digestion. Plac=lac promoter, SP6=SP6 promoter, caR=chloramphenicol acetyltransferase gene, T7=T7 promoter, lacZ=lacZα coding sequence, ampR=beta-lactamase gene, ori=origin of replication. D) Empty destination vector. A counter-selectable ccdB-cassette has been inserted between the LB and RB sequences of pGreen-IIS, flanked by BsaI sites, with overhangs A and G. promoter=bacterial promoter. The pSa origin of replication (ori A. tum.) requires the presence of the helper plasmid pSOUP in agrobacteria.
*The description was adapted from (Lampropoulos et al., 2013)

5.4.4.2 Gateway cloning

A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. This method was used for making protein expression constructs and Dual-Luciferase-Assay related constructs.

BP reaction:

a) Mix 75 ng of pDONR vector and 150 ng purified attB-PCR products with 1 µl BP ClonaseII enzyme mix in a 5 µl reaction.
b) Incubate the reaction mix at 25°C for 1h.
c) Add 0.5 µl Proteinase K to each sample to terminate the reaction at 37 °C for 10 min (optional).
d) 2.5 µl of the mix is used for transformation.

LR reaction:

a) Mix 150 ng of entry clone and 75 ng destination vector with 1 µl LR ClonaseII Enzyme mix in a 5 µl reaction.
b) Incubate the reaction mix at 25°C for 1h.
c) Add 0.5 µl Proteinase K to each sample to terminate the reaction at 37 °C for 10 min (optional).
d) 2.5 µl of the mix is used for transformation.

5.4.5 Determination of DNA and RNA concentration

DNA and RNA concentrations were determined on NanoDrop 2000 Spectrophotometer (ThermoFisher). The method is based on spectrophotometrically measurement of the sample absorbance at 260nm and 280 nm. Good quality DNA or RNA should have a ratio of A260/A280 at 1.8 to 2.0.

5.4.6 Reverse transcription

cDNAs used in this study were synthesized using AMV Reverse Transcriptase (Roboklon). 500 or 1000 ng RNA was used for each reaction. The steps were conducted following the reaction protocol of the manufacturer.

5.4.7 qRT-PCR

For gene expression levels determination, quantitative RT-PCR was conducted with Rotor-Gene Q (Qiagen) using qPCRBIO SyGreen Mix (PCR Biosystems). The reaction mix preparation and qPCR thermal cycling conditions are as below.

		15x	20x	30x	40x
SYBR Mix	10 µl	150	200	300	400
Primer mix	1 µl	15	20	30	40
ddH2O	7 µl	105	140	210	280
cDNA	2 µl	30	40	60	80
total	20 µl				

Cycles	Temperature	Time
1	95°C	2min
40	95°C 60°C to 65°C	5sec 20-30sec

5.5 Protein expression, purification, and determination methods

5.5.1 Protein expression

In this study, all the proteins involved in the in vitro experiments were expressed in E. Coli system.

For checking the expression level and solubility of the expressed protein, firstly, the destination constructs were transformed into BL21-DE3 competent cells. The culture was inoculated with a single colony or from small amount of glycerol stock at 37 ℃ overnight. The expand culture was then inoculated with a ratio of 1:20 from the overnight culture until OD600 reached 0.6-1.0, 0.75 mM IPTG was added in the culture. From then on, the culture was incubated in 28 ℃ for 4 h. To prepare protein samples, 50µl sonication buffer (300mM NaCl, 50mM NaPi, pH 8) was used for resuspending the pellets after centrifugation. The pellets were sonicated until the solution became clear. The samples were spun down for 2 min at 13000 rpm. Then the supernatant was transferred into new tubes and 1 µl PMSF (10mg/ml) was added. For the pellets, 30 µl lipsick urea buffer was used for resuspending the pellets. After checking with SDS-PAGE, the solubility of the protein could be determined.

5.5.2 Protein purification

Proteins with 6xHis tag in the study were purified using HisTrap™ HP 1 ml (GE Healthcare) as described in the operation manual.

5.5.3 Electrophoresis techniques

The protein samples were mixed with Roti®-Load 1 (Roth) for SDS-PAGE followed by heating at 100 ℃ for 5 min. Both forms of electrophoresis were performed on a 4.5% stacking gel following a 12% separating gel.

Proteins were electro-blotted onto Immobilon-P PVDF membrane (Sigma-Aldrich) after SDS-PAGE. The expressed proteins were studied by Western blotting with anti-

His antibody and detected using SuperSignal West Dura Extended Duration Substrate for HRP (Thermo Fisher Scientific) after incubation with anti-mouse IgG secondary antibody (Bio-Rad).

5.6 Protein-Nucleic Acid Interactions

5.6.1 Dual-luciferase-assay

The Dual Luciferase Assay protocol was modified based on Czemmel et al., (2009). Cells were bombarded with 1.6 µm gold particles (Bio-Rad) using the Model PDS-1000/He Biolistic Particle Delivery System from Bio-Rad with 4481 kPa helium pressure, a vacuum of 86 kPa and a distance of 9.5 cm. After bombardment, the cells and leaf discs were incubated on GC medium containing agar platesfor 48 h dark at 22 °C, then they were harvested and ground on ice in 200 µl of 2x Passive Lysis Buffer (PLB, Promega). Followed by a centrifugation step of the lysates for 1 min at 10000 rpm, measurement of the luciferase activities was performed according to the Dual-Luciferase Reporter assay system (DLR, Promega). Therefore, 10 µl lysate supernatant were mixed with 25 µl LARII (containing Firefly substrate beetle luciferin) and Stop & Glo® (containing Renilla substrate coelenterazine). The light emission, generated by active luciferases in the lysate, was measured with a Lumat LB 9507 Luminometer (Berthold Technologies). Relative luciferase activities were calculated as ratios between firefly and renilla (control) luciferase. The unbombared, grounded cell luminescence was used as background and was substracted from all measured values. All transfection experiments were repeated at least three times.

5.6.2 Electrophoretic Mobility Shift Assay

Electrophoretic Mobility Shift Assays (EMSAs) were performed using fluorescent probes produced by annealing complementary pairs of CY5 5′-labelled oligonucleotides (Eurofins) in annealing buffer (10 mM Tris-HCl pH 8, 1 mM EDTA,

50 mM NaCl). For binding reactions, protein was incubated with 200 fmol CY5 probe in binding buffer (final reaction conditions: 10 mM Tris-HCl pH 7.4, 75 mM NaCl, 1 mM DTT) for 20 min at room temperature. Samples were mixed with 6× Orange G loading dye and applied to a 6% native polyacrylamide gel in 0.5× TBE buffer. Electrophoresis was conducted at 150 V for 20 min and the gel shifts were subsequently recorded using the ImageQuant LAS 4000. Competitor assays were performed with a pre-incubation with non-labelled oligonucleotides as competitors prior to adding the CY5 5′-labelled probe.

5.7 Metabolites analysis

5.7.1 Methanolic extraction of flavonols

Leaf tissue samples are ground in liquid nitrogen and distributed in 50 mg aliquots in 2 ml reaction tubes with an addition of 0.4 ml 80% methanol and have to be vortexed for 20 sec. From here samples are kept dark (cover with aluminum foil) and cold (on ice). Afterwards sonicate samples for 20 min on ice (without aluminum foil). Incubate for 15 min on 70°C and centrifuge for 45 min 13.000 rpm on 4°C. Transfer the supernatant into a fresh 1.5 ml tube without disrupting the pellet. If pellet rests are transferred include a second short centrifugation for 1 min 13.000 rpm and repeat transfer into a fresh 1.5 ml tube. Evaporate 80% MeOH in the SpeedVac on 60°C for ~1h. Resolve the pellet in 1 µl 50 % MeOH/ mg starting material.

5.7.2 Thin layer chromatography with subsequent DPBA staining

Thin layer chromatography (TLC) and subsequent DPBA staining will be performed according to Stracke et al., (2007). 4 µl of methanolic extracts are spotted on 10 cm x10 cm silica-60 HPTLC-plates (Merck) used as the stationary phase. Adsorption chromatography will be carried out using a system of ethyl acetate, formic acid and water (8:1:1) as the mobile phase (20 ml) in a closed glass tank. Separated

phenylpropanoid compounds are stained afterwards by spraying a 1% DPBA(w/v) (also known as Naturstoffreagenz A; Roth 9920-1) solution in methanol (10 ml), followed by spraying 5% ethanolic polyethylene glycol(PEG) 4000 (w/v) solution (10 ml). The stained chromatograms are examined under UV light (346 nm) and photographed.

5.7.3 Cell wall components determination

Mature *Arabidopsis* stems (~8-week-old) and Brachypodium shoots were harvested, dried at 50 °C, ground into powder through 40 mesh screen and stored in sealed dry container until use. All the components determination experiments were conducted in Biomass and Bioenergy Research Centre, Huazhong Agricultural University, Wuhan, China during the external exchange period.

5.7.3.1 Plant wall polymer extraction and determination

The plant cell wall fraction method was used to extract hemicelluloses and cellulose as described by Peng et al. (2000) and Wu et al. (2013). Total hemicellulose was calculated based on total hexoses and pentoses determined in the hemicellulose fraction, and total hexoses were measured as cellulose in the cellulose fraction. All experiments were carried out in biological triplicate.

An UV–vis spectrometer (V-1100D, Shanghai MAPADA Instruments Co., Ltd.) was used to determine the hexoses and pentoses according to Sun et al. (2017)

5.7.3.2 Total lignin and monolignol assay

The total lignin content was measured by the two-step acid hydrolysis method according to Laboratory Analytical Procedure of the National Renewable Energy Laboratory with minor modifications as described by Sun et al. (2017) The three monomers of lignin were determined by HPLC as described by Jin et al. (2016) and Sun et al. (2017)

Table 5-3 Primers used in the study-1 (For cloning)

Primer	Sequence
gg_MsMYB31_fwd	aacaggtctcaggctATGGGGAGATCTCCGTGCT
gg_MsMYB31_rev	aacaggtctctctgaTTTCATCTCGAGGCTTCTG
gg_MsMYB42_fwd	aacaggtctcaggctATGGGGCGGTCGCCGTGC
gg_MsMYB42_rev	aacaggtctctctgaCTTCATCTCAAGGCCTCT
MsMYB31-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcATGGG GAGATCTCCGTGCT
MsMYB31_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTtTTTCATC TCGAGGCTTCTG
MsMYB42_attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcATGGG GCGGTCGCCGTGC
MsMYB42_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTtTCACTT CATCTCAAGGCCTCT
MYB31-R2R3-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtTCACCT CCGGATGTGCGTGTT
MYB42-R2R3-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtTCACCG GATGTGCGTGTTCCAG
gg_pAtMYB4_fwd	aacaggtctcaacctCGTCCAGTTAGAGTTCAACATTATT
gg_pAtMYB4_rev	aacaggtctcttgttCTTGTGGTTTAGATCTTATTTCGTC
proMsC4H1_fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcTGTGCT CAAGGCATCCG
proMsC4H1_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtGCTGGA CGCTTGGAGCC
proMsCAD2_fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcGGCAC ATCACTTTCCTC
proMsCAD2_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtGATCGA CACCGCTACGG
pMsCHS-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcACCAC TTGGAGCAGCCTATC
pMsCHS-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTtCGTCTC CGTCCGCTCTTT
proMsCCR_fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcTTGGG ACAGAGTGAGTAGATTGC
proMsCCR_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtCTTGGC GCGACGATGATG

Table 5-4 Primers used in the study-2 (For sequencing)

Primer	Sequence
seq_MsMYB31_F	ATGGGGAGATCTCCGTGCT
seq_MsMYB31_R	TTTCATCTCGAGGCTTCT
seqMsMYB42_fwd	ATGGGGCGGTCGCCGTGCTG
seqMsMYB42_rev	CTTCATCTCAAGGCCTCTGA
seq_pAtMYB4_rev	ACAATTCGAGTATTCGACTGTGTAC
seq_pAtMYB4_fwd	GTCAATAAATGTACACAGTCGAATA
LUC_F	CTAACATACGCTCTCCATCA
LUC_R	GGATAGAATGGCGCCGG
pGGA/C000_fwd	cgcaacgcaattaatgtgag
pGGA/C000_rev	cagattgtactgagagtgcacc
pGG-B-dummy_F	gtattcagtcgactggtaccaac
pGG-B-dummy_R	ttggtaccagtcgactgaatac
pGG-D-dummy_F	gtggatcctagataacctttac
pGG-D-dummy_R	acagggaatgaaggtaaagg
seq_Z003-A	GATCGCACCAGGTACCACCT

Table 5-5 Primers used in the study-3 (For qRT-PCR)

Primer	Sequence
qMsMYB31-fwd	AACCGTGACCGTGAGGAGA
qMsMYB31-rev	CGGGGTGGTTCTGATGATG
qMsMYB42_new_F	CGACCTCAACCTCGACCTCT
qMsMYB42_new_R	GTCCAGCTCCTCGTCTTCCT
qAtCHS-fwd	CTTGCCTTCTATCTGCCTACCTAC
qAtCHS-rev	TCCAGCACATATCACATATCACATC
qAtCHI-fwd	GGAGGCGGTTCTGGAATCTATC
qAtCHI-rev	TTCGTCCTTGTTCTTCATCATTAGC
qAtFLS1-fwd	CAAGGATTACAGTTACCGCAAGC
qAtFLS1-rev	CCACAACCACAAATTATTCTTCTCG
qAtDFR-fwd	TTGTTCGTGCCACCGTTCG
qAtDFR-rev	TCCTTCCTCAGATAAATCAGCCTTC
qAtF3H-fwd	GACCCTGGAACCATTACCTTG
qAtF3H-rev	ATCAGCATTCTTGAACCTCCC
qAtPAL1_F	AGCGCAACGTACCCGTTGAT
qAtPAL1_R	CGTAGGCTGCTCTTGCTGCT
qAtC4H_F	AACTGGCTTCAAGTCGGAGA
qAtC4H_R	GACCCATACGGAGGAGGAAG
qAtC3H1_F	AGGAGCGGTTGCGTTCAACA
qAtC3H1_R	AGCCCTTGCTCGTCCACAAC
qAtCCR1_F	CGCGTGGTCATCACCTCCTC
qAtCCR1_R	CAGCCTCAGGGTCACGGTTC
qAtF5H1_F	CCATAGGACGCGACCCAACC
qAtF5H1_R	AAATCCGGTACGCCCGGTTC
qAtCOMT_F	GAAGCTGCCCTCTTCGCCAT
qAtCOMT_R	ACGGAGGATACGGTCGAGCA
qAtHCT_F	ACACGAGACCAGCTTGTTGCT
qAtHCT_R	CTCGCGCCTTTCCCACTGAT
qAtCAD6_F	GGCTGCCAGAGACCCATCTG
qAtCAD6_R	CCACTTCATGCCCAGGAACCA
qAt4CL_F	CTAATGCCAAACTCGGTCAGG
qAt4CL_R	AGCTCCTGACTTAACCGGAAA
qAtCCoAOMT1_F	CTCACAAGATCGACTTCAGGG
qAtCCoAOMT1_R	ACGCTTGTGGTAGTTGATGTAG
Clath_F	TCGATTGCTTGGTTTGGAAGAT
Clath_R	GCACTTAGCGTGGACTCTGTTTGC
PP2A_F	GCTAGCTCCTGTCATGGGTC

PP2A_R	TCATGTTCGGAACCCTGTCC
UBC_F	CTGAACCAGACAGCCCACTT
UBC_R	CTCTGATATCACCCGACCGC
Exp_F	GAGCTGAAGTGGCTTCCATGAC
Exp_R	GGTCCGACATACCCATGATCC

Table 5-6 Vectors used in the study

Vector	Resistance	Description
pDONR201	Kanamycin	Donor vector Gateway
pDONR221	Kanamycin	Donor vector Gateway
pART7_GW	Ampicillin	35S overexpression destination vector
pLUC_GW	Ampicillin	Firefly luciferase destination vector
pRLUC	Ampicillin	Renilla luciferase 35S overexpression
pETG10A	Ampicillin	6xHis tag destination vector
pETG60A	Ampicillin	Nus-A+6xHis tag destination vector
pGGA000	Ampicillin	Empty promoter entry
pGGA004	Ampicillin	35S-promotor
pGGB003	Ampicillin	B-dummy
pGGC000	Ampicillin	Empty CDS entry
pGGD001	Ampicillin	GFP
pGGD002	Ampicillin	D-dummy
pGGE009	Ampicillin	UBQ10 terminator
pGGF001	Ampicillin	pMAS::BastaR::tMAS
pGGZ003	Spectinomycin	destination vector

List of Abbreviations

%	percent
4CL	4-coumarate:CoA ligase
A. tumefaciens	Agrobacterium tumefaciens
AMV	Avian Myeloblastosis Virus
ASL	acid-soluble lignin
BAP	6-Benzylaminopurine
bHLHs	basic Helix-Loop-Helix factors
C3H	p-coumarate 3-hydroxylase
C4H	cinnamic acid 4-hydroxylase
CAD	cinnamyl alcohol dehydrogenase
CaMV 35S	Cauliflower mosaic virus promoter
CCoAOMT	caffeoyl-CoA O-methyltransferase
CCR	cinnamoyl-CoA reductase
CDS	coding sequence
CHI	chalcone isomerase
CHS	chalcone synthase
Col-0	Columbia-0
COMT	caffeic acid o-methyltransferase
CTAB	Cetyl trimethylammonium bromide
DAPI	4′,6-diamidino-2-phenylindole
DEX	dexamethasone
DLA	dual luciferase assay
DMSO	dimethyl sulfoxide
E. Coli	Escherichia coli
EBGs	early flavonoid biosynthesis related genes
EMSA	Electrophoretic Mobility Shift Assay
F3H	flavanone 3-hydroxylase
F3'H	Flavonoid 3′-hydroxylase
F5H	ferulate-5-hydroxylase
FLS	flavonol synthase
GFP	green fluorescent protein

GR	glucocorticoid receptor
HCT	p-hydroxycinnamoyl transferase
HPTLC	High-performance thin-layer chromatography
IL	insoluble lignin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kilodalton
Kin	kinetin
LAC	laccase
LBGs	late flavonoid biosynthetic genes
Ler-0	Landsberg erecta
LUC	firefly luciferase
MES	2-(N-morpholino)ethanesulfonic acid
MS medium	Murashige and Skoog medium
mut	mutated
N. benthamiana	Nicotiana benthamiana
NAA	Naphthaleneacetic acid
NLS	nuclear localization signal
NusA	N-utilization substance
OD	optical density
ORF	open reading frame
PA	proanthocyanidin
PAL	phenylalanine ammonia-lyase
PAR	parabolic aluminized reflector
PBCs	Perennial biomass crops
PCR	polymerase chain reaction
PEG	Polyethylene glycol
pH	negative log10 of the hydrogen ion concentration expressed in mol/L
PIPES	piperazine-N,N′-bis(2-ethanesulfonic acid)
PMSF	phenylmethanesulfonylfluoride or phenylmethylsulfonyl fluorid
POX	peroxidase
qRT-PCR	real-time reverse transcription-PCR
REN	renilla luciferase

S/G	ratio of syringyl to guaiacyl
SCM	secondary cell wall MYBs
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELEX	systematic evolution of ligands by exponential enrichment
SMRE	secondary wall MYB-responsive element
SNBE	secondary wall NAC-binding element
SND	secondary wall–associated NAC domain protein
TE	tris-EDTA
TF	transcription factor
UV	Ultraviolet
VNS	VND, NST/SND, and SMB
WT	wild-type

References

AGARWAL, T., GROTEWOLD, E., DOSEFF, A. I. & GRAY, J. 2016. MYB31/MYB42 Syntelogs Exhibit Divergent Regulation of Phenylpropanoid Genes in Maize, Sorghum and Rice. *Sci Rep,* 6**,** 28502.

AGRAWAL, A. 2011. Pharmacological activities of flavonoids: a review. *International journal of pharmaceutical sciences and nanotechnology,* 4**,** 1394-1398.

ALVIRA, P., TOMÁS-PEJÓ, E., BALLESTEROS, M. & NEGRO, M. 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. *Bioresource technology,* 101**,** 4851-4861.

AMBAWAT, S., SHARMA, P., YADAV, N. R. & YADAV, R. C. 2013. MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants,* 19**,** 307-321.

APPELHAGEN, I., LU, G. H., HUEP, G., SCHMELZER, E., WEISSHAAR, B. & SAGASSER, M. 2011. TRANSPARENT TESTA1 interacts with R2R3-MYB factors and affects early and late steps of flavonoid biosynthesis in the endothelium of Arabidopsis thaliana seeds. *The Plant Journal,* 67**,** 406-419.

AZUMA, A., YAKUSHIJI, H., KOSHITA, Y. & KOBAYASHI, S. 2012. Flavonoid biosynthesis-related genes in grape skin are differentially regulated by temperature and light conditions. *Planta,* 236**,** 1067-1080.

BAXTER, H. L., MAZAREI, M., LABBE, N., KLINE, L. M., CHENG, Q., WINDHAM, M. T., MANN, D. G., FU, C., ZIEBELL, A. & SYKES, R. W. 2014. Two-year field analysis of reduced recalcitrance transgenic switchgrass. *Plant Biotechnology Journal,* 12**,** 914-924.

BESSEAU, S., HOFFMANN, L., GEOFFROY, P., LAPIERRE, C., POLLET, B. & LEGRAND, M. 2007. Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. *The Plant Cell,* 19**,** 148-162.

BHALLA, A., BANSAL, N., KUMAR, S., BISCHOFF, K. M. & SANI, R. K. 2013. Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresource technology,* 128**,** 751-759.

BONAWITZ, N. D. & CHAPPLE, C. 2010. The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annual review of genetics,* 44**,** 337-363.

BRODIN, M., VALLEJOS, M., OPEDAL, M. T., AREA, M. C. & CHINGA-CARRASCO, G. 2017. Lignocellulosics as sustainable resources for production of bioplastics–A review. *Journal of Cleaner Production,* 162**,** 646-664.

BROSSE, N., DUFOUR, A., MENG, X., SUN, Q. & RAGAUSKAS, A. 2012. *Miscanthus*: a fast-growing crop for biofuels and chemicals production. *Biofuels, Bioproducts and Biorefining,* 6**,** 580-598.

BROWN, D. E., RASHOTTE, A. M., MURPHY, A. S., NORMANLY, J., TAGUE, B. W., PEER, W. A., TAIZ, L. & MUDAY, G. K. 2001. Flavonoids act as negative regulators of auxin transport in vivo in Arabidopsis. *Plant physiology,* 126**,** 524-535.

CASS, C. L., PERALDI, A., DOWD, P. F., MOTTIAR, Y., SANTORO, N., KARLEN, S. D., BUKHMAN, Y. V., FOSTER, C. E., THROWER, N. & BRUNO, L. C. 2015. Effects of PHENYLALANINE AMMONIA LYASE (PAL) knockdown on cell wall composition, biomass digestibility, and biotic and abiotic stress responses in Brachypodium. *Journal of Experimental Botany,* 66**,** 4317-4335.

CAVALLINI, E., MATUS, J. T., FINEZZO, L., ZENONI, S., LOYOLA, R., GUZZO, F., SCHLECHTER, R., AGEORGES, A., ARCE-JOHNSON, P. & TORNIELLI, G. B. 2015. The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. *Plant Physiology,* 167**,** 1448-1470.

CHAW, S.-M., CHANG, C.-C., CHEN, H.-L. & LI, W.-H. 2004. Dating the monocot–dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of molecular evolution,* 58**,** 424-441.

CHEN, F. & DIXON, R. A. 2007. Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol,* 25**,** 759-61.

CHEN, L., AUH, C. K., DOWLING, P., BELL, J., CHEN, F., HOPKINS, A., DIXON, R. A. & WANG, Z. Y. 2003. Improved forage digestibility of tall fescue (Festuca arundinacea) by transgenic down-regulation of cinnamyl alcohol dehydrogenase. *Plant Biotechnology Journal,* 1**,** 437-449.

CHERUBINI, F. 2010. The biorefinery concept: using biomass instead of oil for producing energy and chemicals. *Energy conversion and management,* 51**,** 1412-1421.

CHUNG, J.-H. & KIM, D.-S. 2012. *Miscanthus* as a potential bioenergy crop in East Asia. *Journal of Crop Science and Biotechnology,* 15**,** 65-77.

DAHMEN, N., HENRICH, E. & HENRICH, T. 2017. Synthesis gas biorefinery. *Biorefineries.* Springer.

DAHMEN, N., LEWANDOWSKI, I., ZIBEK, S. & WEIDTMANN, A. 2018. Integrated lignocellulosic value chains in a growing bioeconomy: Status quo and perspectives. *GCB Bioenergy*.

DEMIRBAŞ, A. 2001. Biomass resource facilities and biomass conversion processing for fuels and chemicals. *Energy conversion and Management,* 42**,** 1357-1378.

DIXON, R. A., ACHNINE, L., KOTA, P., LIU, C. J., REDDY, M. S. & WANG, L. 2002. The phenylpropanoid pathway and plant defence—a genomics perspective. *Molecular plant pathology,* 3**,** 371-390.

DOHERTY, W. O., MOUSAVIOUN, P. & FELLOWS, C. M. 2011. Value-adding to cellulosic ethanol: Lignin polymers. *Industrial Crops and products,* 33**,** 259-276.

DOHLEMAN, F., HEATON, E., LEAKEY, A. & LONG, S. P. 2009. Does greater leaf-level photosynthesis explain the larger solar energy conversion efficiency of *Miscanthus* relative to switchgrass? *Plant, cell & environment,* 32**,** 1525-1537.

DOUGLAS, C. J. 1996. Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends in Plant Science,* 1**,** 171-178.

DU, H., FENG, B.-R., YANG, S.-S., HUANG, Y.-B. & TANG, Y.-X. 2012. The R2R3-MYB transcription factor gene family in maize. *PloS one,* 7**,** e37463.

DUBOS, C., STRACKE, R., GROTEWOLD, E., WEISSHAAR, B., MARTIN, C. & LEPINIEC, L. 2010. MYB transcription factors in Arabidopsis. *Trends Plant Sci,* 15**,** 573-81.

EUDES, A., LIANG, Y., MITRA, P. & LOQUE, D. 2014. Lignin bioengineering. *Curr Opin Biotechnol,* 26**,** 189-98.

FALCONE FERREYRA, M. L., RIUS, S. & CASATI, P. 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Frontiers in plant science,* 3**,** 222.

FARAGE, P. K., BLOWERS, D., LONG, S. P. & BAKER, N. R. 2006. Low growth temperatures modify the efficiency of light use by photosystem II for CO2 assimilation in leaves of two chilling-tolerant C4 species, Cyperus longus L. and *Miscanthus*× giganteus. *Plant, Cell & Environment,* 29**,** 720-728.

FISCHER, G., PRIELER, S. & VAN VELTHUIZEN, H. 2005. Biomass potentials of *Miscanthus*, willow and poplar: results and policy implications for Eastern Europe, Northern and Central Asia. *Biomass and Bioenergy,* 28**,** 119-132.

FORNALÉ, S., LOPEZ, E., SALAZAR-HENAO, J. E., FERNÁNDEZ-NOHALES, P., RIGAU, J. & CAPARROS-RUIZ, D. 2014. AtMYB7, a new player in the regulation of UV-sunscreens in Arabidopsis thaliana. *Plant and Cell Physiology,* 55**,** 507-516.

FORNALE, S., SHI, X., CHAI, C., ENCINA, A., IRAR, S., CAPELLADES, M., FUGUET, E., TORRES, J. L., ROVIRA, P., PUIGDOMENECH, P., RIGAU, J., GROTEWOLD, E., GRAY, J. & CAPARROS-RUIZ, D. 2010. ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J,* 64**,** 633-44.

FORNALE, S., SONBOL, F. M., MAES, T., CAPELLADES, M., PUIGDOMENECH, P., RIGAU, J. & CAPARROS-RUIZ, D. 2006. Down-regulation of the maize and Arabidopsis thaliana caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Mol Biol,* 62**,** 809-23.

FOSTER, C. E., MARTIN, T. M. & PAULY, M. 2010. Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) part I: lignin. *JoVE (Journal of Visualized Experiments)***,** e1745.

FU, C., XIAO, X., XI, Y., GE, Y., CHEN, F., BOUTON, J., DIXON, R. A. & WANG, Z.-Y. 2011. Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *BioEnergy Research,* 4**,** 153-164.

GOLFIER, P., VOLKERT, C., HE, F., RAUSCH, T. & WOLF, S. 2017. Regulation of secondary cell wall biosynthesis by a NAC transcription factor from *Miscanthus*. *Plant Direct,* 1**,** e00024.

GOLFIER, P. M. 2018. *Transcriptional regulation of secondary cell wall biosynthesis in Miscanthus sinensis.*

GOPAL, G. J. & KUMAR, A. 2013. Strategies for the production of recombinant protein in Escherichia coli. *The protein journal,* 32**,** 419-425.

GRAY, J., CAPARROS-RUIZ, D. & GROTEWOLD, E. 2012. Grass phenylpropanoids: regulate before using! *Plant Sci,* 184**,** 112-20.

GRAY, W. M., ÖSTIN, A., SANDBERG, G., ROMANO, C. P. & ESTELLE, M. 1998. High temperature promotes auxin-mediated hypocotyl elongation in Arabidopsis. *Proceedings of the National Academy of Sciences,* 95**,** 7197-7202.

HARMSEN, P. & HACKMANN, M. M. 2013. *Green building blocks for biobased plastics: biobased processes and market development*, Wageningen UR-Food & Biobased Research.

HARVEY, J. 2007. A versatile solution-growing *Miscanthus* for bioenergy. *Renewable Energy World,* 10**,** 86.

HATTON, D., SABLOWSKI, R., YUNG, M. H., SMITH, C., SCHUCH, W. & BEVAN, M. 1995. Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. *The Plant Journal,* 7**,** 859-876.

HE, F., MACHEMER-NOONAN, K., GOLFIER, P., UNDA, F., DECHERT, J., ZHANG, W., HOFFMANN, N., SAMUELS, L., MANSFIELD, S. D., RAUSCH, T. & WOLF, S. 2019. The in vivo impact of MsLAC1, a *Miscanthus* laccase isoform, on lignification and lignin composition contrasts with its in vitro substrate preference. *BMC Plant Biology,* 19**,** 552.

HE, X., HALL, M. B., GALLO-MEAGHER, M. & SMITH, R. L. 2003. Improvement of forage quality by downregulation of maize O-methyltransferase. *Crop Science,* 43**,** 2240-2251.

HEATON, E., VOIGT, T. & LONG, S. P. 2004. A quantitative review comparing the yields of two candidate C4 perennial biomass crops in relation to nitrogen, temperature and water. *Biomass and bioenergy,* 27**,** 21-30.

HEATON, E. A., DOHLEMAN, F. G. & LONG, S. P. 2008. Meeting US biofuel goals with less land: the potential of *Miscanthus*. *Global change biology,* 14**,** 2000-2014.

HEATON, E. A., DOHLEMAN, F. G., MIGUEZ, A. F., JUVIK, J. A., LOZOVAYA, V., WIDHOLM, J., ZABOTINA, O. A., MCISAAC, G. F., DAVID, M. B. & VOIGT, T. B. 2010. *Miscanthus*: a promising biomass crop. *Advances in botanical research.* Elsevier.

HELLER, W. & FORKMANN, G. 2017. Biosynthesis of flavonoids. *The Flavonoids Advances in Research since 1986.* Routledge.

HENDRIKS, A. & ZEEMAN, G. 2009. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresource technology,* 100**,** 10-18.

HIRANO, K., KONDO, M., AYA, K., MIYAO, A., SATO, Y., ANTONIO, B. A., NAMIKI, N., NAGAMURA, Y. & MATSUOKA, M. 2013. Identification of transcription factors involved in rice secondary cell wall formation. *Plant Cell Physiol,* 54**,** 1791-802.

HÖLL, J., LINDNER, S., WALTER, H., JOSHI, D., POSCHET, G., PFLEGER, S., ZIEGLER, T., HELL, R., BOGS, J. & RAUSCH, T. 2019. Impact of pulsed UV-B stress exposure on plant performance: How recovery periods stimulate secondary metabolism while reducing adaptive growth attenuation. *Plant, cell & environment,* 42**,** 801-814.

HÖLL, J., VANNOZZI, A., CZEMMEL, S., D'ONOFRIO, C., WALKER, A. R., RAUSCH, T., LUCCHIN, M., BOSS, P. K., DRY, I. B. & BOGS, J. 2013. The R2R3-MYB transcription factors MYB14 and MYB15 regulate stilbene biosynthesis in Vitis vinifera. *The Plant Cell,* 25**,** 4135-4149.

HOLLÓSY, F. 2002. Effects of ultraviolet radiation on plant cells. *Micron,* 33**,** 179-197.

I. LEWANDOWSKIA, J. C. C.-B., J.M.O. SCURLOCKC; W. HUISMAND 2000. *Miscanthus*-European experience with a novel energy crop. *Biomass and Bioenergy*.

JAMES, A. M., MA, D., MELLWAY, R., GESELL, A., YOSHIDA, K., WALKER, V., TRAN, L., STEWART, D., REICHELT, M. & SUVANTO, J. 2017. Poplar MYB115 and MYB134 transcription factors regulate proanthocyanidin synthesis and structure. *Plant physiology,* 174**,** 154-171.

JIN, H., COMINELLI, E., BAILEY, P., PARR, A., MEHRTENS, F., JONES, J., TONELLI, C., WEISSHAAR, B. & MARTIN, C. 2000. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *The EMBO journal,* 19**,** 6150-6161.

JIN, H. & MARTIN, C. 1999. Multifunctionality and diversity within the plant MYB-gene family. *Plant molecular biology,* 41**,** 577-585.

JUNG, J. H., VERMERRIS, W., GALLO, M., FEDENKO, J. R., ERICKSON, J. E. & ALTPETER, F. 2013. RNA interference suppression of lignin biosynthesis increases fermentable sugar yields for biofuel production from field-grown sugarcane. *Plant biotechnology journal,* 11**,** 709-716.

KAGALE, S., LINKS, M. G. & ROZWADOWSKI, K. 2010. Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in Arabidopsis. *Plant physiology,* 152**,** 1109-1134.

KEATING, J. D., PANGANIBAN, C. & MANSFIELD, S. D. 2006. Tolerance and adaptation of ethanologenic yeasts to lignocellulosic inhibitory compounds. *Biotechnology and bioengineering,* 93**,** 1196-1206.

KHANNA, M., DHUNGANA, B. & CLIFTON-BROWN, J. 2008. Costs of producing *Miscanthus* and switchgrass for bioenergy in Illinois. *Biomass and Bioenergy,* 32**,** 482-493.

KIM, W. C., KIM, J. Y., KO, J. H., KANG, H. & HAN, K. H. 2014. Identification of direct targets of transcription factor MYB46 provides insights into the transcriptional regulation of secondary wall biosynthesis. *Plant Mol Biol,* 85**,** 589-99.

KIM, W. C., KO, J. H., KIM, J. Y., KIM, J., BAE, H. J. & HAN, K. H. 2013. MYB46 directly regulates the gene expression of secondary wall-associated cellulose synthases in Arabidopsis. *Plant J,* 73**,** 26-36.

KUBASEK, W. L., SHIRLEY, B. W., MCKILLOP, A., GOODMAN, H. M., BRIGGS, W. & AUSUBEL, F. M. 1992. Regulation of flavonoid biosynthetic genes in germinating Arabidopsis seedlings. *The Plant Cell,* 4**,** 1229-1236.

KUBO, M., UDAGAWA, M., NISHIKUBO, N., HORIGUCHI, G., YAMAGUCHI, M., ITO, J., MIMURA, T., FUKUDA, H. & DEMURA, T. 2005. Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev,* 19**,** 1855-60.

LASK, J., WAGNER, M., TRINDADE, L. M. & LEWANDOWSKI, I. 2019. Life cycle assessment of ethanol production from *Miscanthus*: A comparison of production pathways at two European sites. *GCB Bioenergy,* 11**,** 269-288.

LEWANDOWSKI, I., SCURLOCK, J. M. O., LINDVALL, E. & CHRISTOU, M. 2003. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass and Bioenergy,* 25**,** 335-361.

LI, J., OU-LEE, T.-M., RABA, R., AMUNDSON, R. G. & LAST, R. L. 1993. Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. *The Plant Cell,* 5**,** 171-179.

LOIS, R., DIETRICH, A., HAHLBROCK, K. & SCHULZ, W. 1989. A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. *The EMBO journal,* 8**,** 1641-1648.

LYGIN, A. V., UPTON, J., DOHLEMAN, F. G., JUVIK, J., ZABOTINA, O. A., WIDHOLM, J. M. & LOZOVAYA, V. V. 2011. Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop–*Miscanthus*. *Gcb Bioenergy,* 3**,** 333-345.

MA, D. & CONSTABEL, C. P. 2019. MYB repressors as regulators of phenylpropanoid metabolism in plants. *Trends in plant science*.

MA, D., REICHELT, M., YOSHIDA, K., GERSHENZON, J. & CONSTABEL, C. P. 2018. Two R2R3-MYB proteins are broad repressors of flavonoid and phenylpropanoid metabolism in poplar. *The Plant Journal,* 96**,** 949-965.

MCCARTHY, R. L., ZHONG, R. & YE, Z. H. 2009. MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell Physiol,* 50**,** 1950-64.

MEHRTENS, F., KRANZ, H., BEDNAREK, P. & WEISSHAAR, B. 2005. The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant physiology,* 138**,** 1083-1096.

MELLWAY, R. D., TRAN, L. T., PROUSE, M. B., CAMPBELL, M. M. & CONSTABEL, C. P. 2009. The wound-, pathogen-, and ultraviolet B-responsive MYB134 gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in poplar. *Plant Physiology,* 150**,** 924-941.

MENON, V. & RAO, M. 2012. Trends in bioconversion of lignocellulose: biofuels, platform chemicals & biorefinery concept. *Progress in energy and combustion science,* 38**,** 522-550.

MITRA, M., AGARWAL, P., KUNDU, A., BANERJEE, V. & ROY, S. 2019. Investigation of the effect of UV-B light on Arabidopsis MYB4 (AtMYB4) transcription factor stability and detection of a putative MYB4-binding motif in the promoter proximal region of AtMYB4. *PloS one,* 14.

NAKANO, Y., YAMAGUCHI, M., ENDO, H., REJAB, N. A. & OHTANI, M. 2015. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front Plant Sci,* 6**,** 288.

NARAYANA, K. R., REDDY, M. S., CHALUVADI, M. & KRISHNA, D. 2001. Bioflavonoids classification, pharmacological, biochemical effects and therapeutic potential. *Indian journal of pharmacology,* 33**,** 2-16.

NESI, N., DEBEAUJON, I., JOND, C., STEWART, A. J., JENKINS, G. I., CABOCHE, M. & LEPINIEC, L. 2002. The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *The Plant Cell,* 14**,** 2463-2479.

NEUTELINGS, G. 2011. Lignin variability in plant cell walls: contribution of new models. *Plant Science,* 181**,** 379-386.

NURUZZAMAN, M., MANIMEKALAI, R., SHARONI, A. M., SATOH, K., KONDOH, H., OOKA, H. & KIKUCHI, S. 2010. Genome-wide analysis of NAC transcription factor family in rice. *Gene,* 465**,** 30-44.

OHTANI, M. & DEMURA, T. 2019. The quest for transcriptional hubs of lignin biosynthesis: beyond the NAC-MYB-gene regulatory network model. *Current opinion in biotechnology,* 56**,** 82-87.

OLSEN, A. N., ERNST, H. A., LEGGIO, L. L. & SKRIVER, K. 2005. NAC transcription factors: structurally distinct, functionally diverse. *Trends in plant science,* 10**,** 79-87.

OLSON, S. N., RITTER, K., ROONEY, W., KEMANIAN, A., MCCARL, B. A., ZHANG, Y., HALL, S., PACKER, D. & MULLET, J. 2012. High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. *Biofuels, Bioproducts and Biorefining,* 6**,** 640-655.

PASHKOVSKIY, P. P., KARTASHOV, A. V., ZLOBIN, I. E., POGOSYAN, S. I. & KUZNETSOV, V. V. 2016. Blue light alters miR167 expression and microRNA-targeted auxin response factor genes in Arabidopsis thaliana plants. *Plant Physiology and Biochemistry,* 104**,** 146-154.

PAUL, S. & DUTTA, A. 2018. Challenges and opportunities of lignocellulosic biomass for anaerobic digestion. *Resources, Conservation and Recycling,* 130**,** 164-174.

PEER, W. A. & MURPHY, A. S. 2007. Flavonoids and auxin transport: modulators or regulators? *Trends in plant science,* 12**,** 556-563.

PROUSE, M. B. & CAMPBELL, M. M. 2012. The interaction between MYB proteins and their target DNA binding sites. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1819**,** 67-77.

RABINOWICZ, P. D., BRAUN, E. L., WOLFE, A. D., BOWEN, B. & GROTEWOLD, E. 1999. Maize R2R3 Myb genes: sequence analysis reveals amplification in the higher plants. *Genetics,* 153**,** 427-444.

RAES, J., ROHDE, A., CHRISTENSEN, J. H., VAN DE PEER, Y. & BOERJAN, W. 2003. Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiology,* 133**,** 1051-1071.

RAGAUSKAS, A. J., BECKHAM, G. T., BIDDY, M. J., CHANDRA, R., CHEN, F., DAVIS, M. F., DAVISON, B. H., DIXON, R. A., GILNA, P. & KELLER, M. 2014. Lignin valorization: improving lignin processing in the biorefinery. *science,* 344**,** 1246843.

REDDY, M. S., CHEN, F., SHADLE, G., JACKSON, L., ALJOE, H. & DIXON, R. A. 2005. Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.). *Proceedings of the National Academy of Sciences,* 102**,** 16573-16578.

RIVAS, S., VILA, C., ALONSO, J., SANTOS, V., PARAJÓ, J. & LEAHY, J. 2019. Biorefinery processes for the valorization of *Miscanthus* polysaccharides: from constituent sugars to platform chemicals. *Industrial Crops and Products,* 134**,** 309-317.

SHADLE, G., CHEN, F., SRINIVASA REDDY, M. S., JACKSON, L., NAKASHIMA, J. & DIXON, R. A. 2007. Down-regulation of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry,* 68**,** 1521-9.

SHEN, H., HE, X., POOVAIAH, C. R., WUDDINEH, W. A., MA, J., MANN, D. G., WANG, H., JACKSON, L., TANG, Y., STEWART, C. N., JR., CHEN, F. & DIXON, R. A. 2012. Functional characterization of the switchgrass (Panicum virgatum) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol,* 193**,** 121-36.

SHEN, H., YIN, Y., CHEN, F., XU, Y. & DIXON, R. A. 2009. A Bioinformatic Analysis of NAC Genes for Plant Cell Wall Development in Relation to Lignocellulosic Bioenergy Production. *BioEnergy Research,* 2**,** 217-232.

SHINNERS, K. J., BOETTCHER, G., MUCK, R., WEIMER, P. & CASLER, M. 2010. Harvest and storage of two perennial grasses as biomass feedstocks. *Transactions of the ASABE,* 53**,** 359-370.

SONBOL, F. M., FORNALE, S., CAPELLADES, M., ENCINA, A., TOURINO, S., TORRES, J. L., ROVIRA, P., RUEL, K., PUIGDOMENECH, P., RIGAU, J. & CAPARROS-RUIZ, D. 2009. The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in Arabidopsis thaliana. *Plant Mol Biol,* 70**,** 283-96.

STRACKE, R., ISHIHARA, H., HUEP, G., BARSCH, A., MEHRTENS, F., NIEHAUS, K. & WEISSHAAR, B. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. *The Plant Journal,* 50**,** 660-677.

STRACKE, R., WERBER, M. & WEISSHAAR, B. 2001. The R2R3-MYB gene family in Arabidopsis thaliana. *Current opinion in plant biology,* 4**,** 447-456.

TANG, X., ZHUANG, Y., QI, G., WANG, D., LIU, H., WANG, K., CHAI, G. & ZHOU, G. 2015. Poplar PdMYB221 is involved in the direct and indirect regulation of secondary wall biosynthesis during wood formation. *Sci Rep,* 5**,** 12240.

TAYLOR, L. P. & GROTEWOLD, E. 2005. Flavonoids as developmental regulators. *Current opinion in plant biology,* 8**,** 317-323.

TREUTTER, D. 2006. Significance of flavonoids in plant resistance: a review. *Environmental Chemistry Letters,* 4**,** 147.

VAN ACKER, R., LEPLÉ, J.-C., AERTS, D., STORME, V., GOEMINNE, G., IVENS, B., LÉGÉE, F., LAPIERRE, C., PIENS, K. & VAN MONTAGU, M. C. 2014. Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proceedings of the National Academy of Sciences,* 111**,** 845-850.

VOELKER, S. L., LACHENBRUCH, B., MEINZER, F. C., JOURDES, M., KI, C., PATTEN, A. M., DAVIN, L. B., LEWIS, N. G., TUSKAN, G. A. & GUNTER, L. 2010. Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiology,* 154**,** 874-886.

VOGT, T. 2010. Phenylpropanoid biosynthesis. *Molecular plant,* 3**,** 2-20.

VORWERK, S., SOMERVILLE, S. & SOMERVILLE, C. 2004. The role of plant cell wall polysaccharide composition in disease resistance. *Trends in plant science,* 9**,** 203-209.

WANG, H. Z. & DIXON, R. A. 2012. On-off switches for secondary cell wall biosynthesis. *Mol Plant,* 5**,** 297-303.

WANG, L., RAN, L., HOU, Y., TIAN, Q., LI, C., LIU, R., FAN, D. & LUO, K. 2017. The transcription factor MYB115 contributes to the regulation of proanthocyanidin biosynthesis and enhances fungal resistance in poplar. *New Phytologist,* 215**,** 351-367.

WANG, Y., BOUCHABKE-COUSSA, O., LEBRIS, P., ANTELME, S., SOULHAT, C., GINEAU, E., DALMAIS, M., BENDAHMANE, A., MORIN, H. & MOUILLE, G. 2015. LACCASE5 is required for lignification of the Brachypodium distachyon Culm. *Plant physiology,* 168**,** 192-204.

WILKINS, O., NAHAL, H., FOONG, J., PROVART, N. J. & CAMPBELL, M. M. 2009. Expansion and diversification of the Populus R2R3-MYB family of transcription factors. *Plant Physiology,* 149**,** 981-993.

WINKEL-SHIRLEY, B. 2002. Biosynthesis of flavonoids and effects of stress. *Current opinion in plant biology,* 5**,** 218-223.

XU, B., ESCAMILLA-TREVIÑO, L. L., SATHITSUKSANOH, N., SHEN, Z., SHEN, H., PERCIVAL ZHANG, Y. H., DIXON, R. A. & ZHAO, B. 2011. Silencing of 4-coumarate: coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New phytologist,* 192**,** 611-625.

XUE, S., LEWANDOWSKI, I., WANG, X. & YI, Z. 2016. Assessment of the production potentials of *Miscanthus* on marginal land in China. *Renewable and Sustainable Energy Reviews,* 54**,** 932-943.

YAMAGUCHI, M., GOUE, N., IGARASHI, H., OHTANI, M., NAKANO, Y., MORTIMER, J. C., NISHIKUBO, N., KUBO, M., KATAYAMA, Y., KAKEGAWA, K., DUPREE, P. & DEMURA, T. 2010. VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol,* 153**,** 906-14.

YAN, J., CHEN, W., LUO, F., MA, H., MENG, A., LI, X., ZHU, M., LI, S., ZHOU, H. & ZHU, W. 2012. Variability and adaptability of *Miscanthus* species evaluated for energy crop domestication. *Gcb Bioenergy,* 4**,** 49-60.

YANG, F., MITRA, P., ZHANG, L., PRAK, L., VERHERTBRUGGEN, Y., KIM, J. S., SUN, L., ZHENG, K., TANG, K., AUER, M., SCHELLER, H. V. & LOQUE, D. 2013. Engineering secondary cell wall deposition in plants. *Plant Biotechnol J,* 11**,** 325-35.

YANG, L., ZHAO, X., RAN, L., LI, C., FAN, D. & LUO, K. 2017a. PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Sci Rep,* 7**,** 41209.

YANG, Z., LIU, B., SU, J., LIAO, J., LIN, C. & OKA, Y. 2017b. Cryptochromes orchestrate transcription regulation of diverse blue light responses in plants. *Photochemistry and photobiology,* 93**,** 112-127.

YOSHIDA, K., MA, D. & CONSTABEL, C. P. 2015. The MYB182 protein down-regulates proanthocyanidin and anthocyanin biosynthesis in poplar by repressing both structural and regulatory flavonoid genes. *Plant Physiol,* 167**,** 693-710.

ZHANG, L., WANG, Y., SUN, M., WANG, J., KAWABATA, S. & LI, Y. 2014. BrMYB4, a suppressor of genes for phenylpropanoid and anthocyanin biosynthesis, is down-regulated by UV-B but not by pigment-inducing sunlight in turnip cv. Tsuda. *Plant and Cell Physiology,* 55**,** 2092-2101.

ZHAO, J., ZHANG, W., ZHAO, Y., GONG, X., GUO, L., ZHU, G., WANG, X., GONG, Z., SCHUMAKER, K. S. & GUO, Y. 2007. SAD2, an importin -like

protein, is required for UV-B response in Arabidopsis by mediating MYB4 nuclear trafficking. *Plant Cell,* 19**,** 3805-18.

ZHAO, Q. & DIXON, R. A. 2011. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends Plant Sci,* 16**,** 227-33.

ZHONG, R., LEE, C., MCCARTHY, R. L., REEVES, C. K., JONES, E. G. & YE, Z. H. 2011. Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant Cell Physiol,* 52**,** 1856-71.

ZHONG, R., LEE, C. & YE, Z. H. 2010a. Functional characterization of poplar wood-associated NAC domain transcription factors. *Plant Physiol,* 152**,** 1044-55.

ZHONG, R., LEE, C. & YE, Z. H. 2010b. Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. *Mol Plant,* 3**,** 1087-103.

ZHONG, R., MCCARTHY, R. L., HAGHIGHAT, M. & YE, Z. H. 2013. The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS One,* 8**,** e69219.

ZHONG, R., RICHARDSON, E. A. & YE, Z. H. 2007a. The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell,* 19**,** 2776-92.

ZHONG, R., RICHARDSON, E. A. & YE, Z. H. 2007b. Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta,* 225**,** 1603-11.

ZHONG, R. & YE, Z.-H. 2011. MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant and Cell Physiology,* 53**,** 368-380.

ZHONG, R. & YE, Z. H. 2012. MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol,* 53**,** 368-80.

ZHONG, R. & YE, Z. H. 2014. Complexity of the transcriptional network controlling secondary wall biosynthesis. *Plant Sci,* 229**,** 193-207.

ZHONG, R. & YE, Z. H. 2015. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol,* 56**,** 195-214.

ZHONG, R., YUAN, Y., SPIEKERMAN, J. J., GULEY, J. T., EGBOSIUBA, J. C. & YE, Z. H. 2015. Functional Characterization of NAC and MYB Transcription Factors Involved in Regulation of Biomass Production in Switchgrass (Panicum virgatum). *PLoS One,* 10**,** e0134611.

ZHOU, H., LIN-WANG, K., WANG, F., ESPLEY, R. V., REN, F., ZHAO, J., OGUTU, C., HE, H., JIANG, Q., ALLAN, A. C. & HAN, Y. 2019. Activator-type R2R3-MYB genes induce a repressor-type R2R3-MYB gene to balance anthocyanin and proanthocyanidin accumulation. *New Phytol,* 221**,** 1919-1934.

ZHOU, J., LEE, C., ZHONG, R. & YE, Z. H. 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell,* 21**,** 248-66.

ZHUANG, Q., QIN, Z. & CHEN, M. 2013. Biofuel, land and water: maize, switchgrass or *Miscanthus*? *Environmental Research Letters,* 8**,** 015020.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Prof. Dr. Thomas Rausch, for his continuous support of my Ph.D study. I would not have made it without his patience, motivation and guidance. I also feel really grateful for the support from China Scholarship Council (CSC), because of the cooperation between the graduate program and CSC, I seized the opportunity to be a member of the project.

Besides, my sincere thanks go to my second supervisor Dr. Sebastian Wolf, and TAC meeting member Prof. Dr. Michael Wink and Dr. Emmanuel Gaquerel, for their insightful comments and encouragement which incented me to widen my research from various perspectives

I would like to thank all my lab mates for their continued support in COS, especially Dr. Katja Machemer-Noonan and Dr. Roland Gromes for their selfless help to correct my thesis. Thanks to Heike Steininger, Cornelia Walter, and Angelika Wunderlich for all the technical and management assistance. I am also thankful to Dr. Tanja Peskan-Berghoefer, Hanni Truong and my group buddies Dr. Philippe Golfier and Dr. Feng He in the BBW-Forwerts graduate program. And thanks to Jiaxin Zhong, for being a not only a workmate but also a true friend of mine. Special thanks also go to AG Reissig in COS and Prof. Dr. Liangcai Peng's group in HZAU, for the pleasant cooperation and happy memories between us.

I truly appreciate all the care and support from my friends, the most memorable thing is the warmth you give me.

Last but not least, I would like to express my appreciation to my family: my parents and my husband for supporting me spiritually throughout the studying period and all my life with their true love.