Universitat de Lleida

Dissection and modulation of isoprenoid biosynthesis in higher plants

Xin Huang

Universitat de Lleida

TESI DOCTORAL

Dissection and modulation of isoprenoid biosynthesis in higher plants

Xin Huang

Memòria presentada per optar al grau de Doctor per la Universitat de Lleida
Programa de Doctorat en Ciència i Tecnologia Agrària i Alimentària

Directors
Paul Christou
Teresa Capell

Tutor/a
Paul Christou

2023

Front cover paintings by Teresa Capell
“Birthday cake” oli sobre tela 4 teles de 22x22 (2013)
Back cover painting by Teresa Capell
“Grape Expectations!” oli sobre tela 27x43 (2015)

Supervisors: **Paul Christou**

Teresa Capell

Department of Plant Production and Forestry Science

School of Agricultural and Forestry Engineering

University of Lleida

2023

Signatures P. Christou T. Capell

DEDICATION

To my parents

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor Dr. Paul Christou for his unwavering support and guidance throughout my academic journey. It has been an absolute pleasure to have your guidance in all the time of research and writing of this thesis. A very special thanks to my co-supervisor Dr. Teresa Capell for her enthusiasm, patience, and her excellent advocacy capacity in disseminating our scientific work worldwide. Without your mentorship and encouragement, this PhD would not have been possible.

I extend my sincerest appreciation to Changfu Zhu for his constructive suggestions and useful advice during my research. I would also like to thank Dr. Lourdes Gomez Gomez (University of Castilla-La Mancha, Spain), for making facilities for carotenoid analysis available and Dr. Oussama Ahrazem for his kind help with HPLC data analysis. I would also like to thank Dr. Stefan Schillberg (Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Germany) for accepting me to his laboratory for a short visit and Dr. Andreas Schiermeyer (also at FhIME) for his guidance on CRISPR/Cas9 analysis.

I gratefully acknowledge the University of Lleida, especially the "Departament de Producció Vegetal i Ciència Forestal" for offering me the facilities for my PhD research. I would like to thank the "Agencia de Gestión de Ayudas Universitarias y de Investigación" (AGAUR, Spain) for the PhD fellowship.

I would like to thank Dr. Ludovic Bassie and Dr. Victoria Armario-Najera for their help in all the molecular techniques. A special thank you to Nuria Gabernet for assisting me with many official documents. I would like to thank Jaume Capell for looking after my maize plants. I would like to thank my friend and colleague, Wenhshu, for helping me get started in the lab and supporting me with her technical knowledge. I would like to thank Andrea for making me a member of her great family. I would like to thank Pedro and Erkan for helping me with all the experiments. To Ashwin and Guillermo for being incomparable lab colleagues for insightful discussions. To Hugo and Sheila for helping me with tissue culture. To all the lab colleagues from then and now: Xin Jin, Can, Amaya, Maria, Jaume, Jose, Derry, Alex, and Jessica, thank you for all the assistance and encouragement throughout my studies. So many moments are valuable and memorable, and I will always remember those great times.

I want to express my deep appreciation to my family, mother and father. Thank you for all your love and support. Their unwavering support has been a source of strength for me during challenging times. All my love and thanks for everything. Words are powerless to express my gratitude.

SUMMARY

My thesis focuses on identifying and characterizing the regulatory mechanisms of isoprenoid biosynthesis in higher plants by multigene engineering and gene editing. The major aim of my thesis was to better understand the isoprenoid biosynthesis pathway's bottlenecks by analyzing genes involved in isoprenoid accumulation and their functionality for future use in metabolic engineering applications. The overall objectives were to: a) enhance the production of crocins in two Nicotiana species and compare the accumulation of crocins in these species; b) investigate the regulation of the transcription factor OsBZ8 on isoprenoid biosynthetic pathway genes in rice embryo and endosperm; c) explore the potential of CRISPR/Cas9 to generate a series of maize mutants with altered strigolactone contents.

I focused on the biosynthetic enzymes of two crocin-producing plants, *(Crocus sativus*) *Cs*CCD2L and (*Buddleja davidii*) *Bd*CCD4.1 and transferred them to two different Nicotiana species. I created transgenic lines expressing different combinations of transgenes and investigated the resulting crocin profiles to determine how expression of the integrated transgene complement influenced the accumulation of crocin. Engineered *N. glauca* plants expressing C*sCCD2L* alone accumulated the highest levels of crocins. My results indicated that *N. glauca* is a better host for crocin production than *N. tabacum*.

I demonstrated that the transcription factor OsBZ8 and five associated genes, *AACT3*, *HMGS1*, *HMGR1*, *DXS2*, *IPPI1*, which are involved in the rice MVA and MEP pathways, exhibited similar expression patterns in rice endosperm and embryo. Moreover, I confirmed that OsBZ8 specifically binds to a G-box or a hybrid G/C-box in the promoters of *HMGR1*, *DXS2* and *IPPI1*.

I used CRISPR/Cas9 technology to modulate the expression of *ZmCCD7* and *ZmCCD8*, which are involved in Strigolactones (SL) biosynthesis in maize. I was able to recover SpCas9-induced mutations in two independent plant lines in the coding region of *ZmCCD8*. In depth analysis of the *ZmCCD8* mutants I recovered will contribute towards a better understanding of the molecular basis controlling structural traits in maize. Future genome sequencing of subsequent generations as well as metabolomic analysis is needed to assess fully the impact of these mutations. Collectively my thesis provides novel insights of the regulation of isoprenoid biosynthesis in higher plants at different levels, which will make it easier to predict the effect of metabolic engineering in plants for nutritional improvement or the production of valuable metabolites.

RESUMEN

Mi tesis se centra en identificar y caracterizar mediante ingeniería genética, introduciendo genes o editando su genoma, los mecanismos reguladores de la biosíntesis de isoprenoides en plantas superiores. El principal objetivo de mi tesis ha sido intentar superar las limitaciones de la biosíntesis de isoprenoides analizando los genes implicados en su acumulación y cuáles son sus funciones, para poder ser utilizados en nuevas aplicaciones de ingeniería metabólica. Los objetivos generales fueron: a) mejorar la producción de crocinas en dos especies de Nicotiana y comparar su acumulación; b) investigar la regulación del factor de transcripción OsBZ8 en genes de la vía biosintética de isoprenoides en embrión y endospermo de arroz; c) explorar el potencial de CRISPR/Cas9 para generar una serie de mutantes de maíz con contenido alterado de estrigolactona.

Enfoqué mis estudios en las enzimas biosintéticas de dos plantas productoras de crocinas, CsCCD2L (*Crocus sativus*) y BdCCD4.1 (*Buddleja davidii*) y transferí los genes que codifican para estas enzimas a dos especies diferentes de Nicotiana. Creé líneas transgénicas que expresaban diferentes combinaciones de genes e investigué los perfiles de crocinas resultantes para determinar cómo la expresión del complemento transgénico integrado influía en la acumulación de las crocinas. Las plantas de *N. glauca* transgénicas que solo expresaban *CsCCD2L* acumularon los niveles más altos de crocinas. Mis resultados indicaron que *N. glauca* es una mejor plataforma para la producción de crocinas que *N. tabacum*.

También demostré que el factor de transcripción OsBZ8 y cinco genes asociados, *AACT3*, *HMGS1*, *HMGR1*, *DXS2*, *IPPI1*, que están involucrados en las vías MVA y MEP del arroz, exhibían patrones de expresión similares en el endospermo y en el embrión del arroz. Además, confirmé que OsBZ8 se une específicamente a una G-box o una G/C-box híbrida en los promotores de *HMGR1*, *DXS2* e *IPPI1*.

Usé la tecnología CRISPR/Cas9 para modular la expresión de los genes *ZmCCD7* y *ZmCCD8*, que están involucrados en la biosíntesis de estrigolactonas (SL) en el maíz. Induje mutaciones por SpCas9 en la región codificante de *ZmCCD8*, en dos líneas independientes de plantas. El análisis de los mutantes *ZmCCD8* contribuirá a una mejor comprensión de la base molecular que controla las características estructurales del maíz. Sera necesario secuenciar el genoma de las siguientes generaciones, así como el análisis metabolómico, para evaluar el impacto de estas mutaciones. En conjunto, mi tesis proporciona nuevos conocimientos sobre la regulación de la biosíntesis de isoprenoides en plantas superiores a diferentes niveles, lo que facilitará la predicción del efecto de la ingeniería metabólica en las plantas para la mejora nutricional o la producción de metabolitos importantes.

RESUM

La meva tesi es centra en identificar i caracteritzar mitjançant enginyeria genètica, (introduint gens o editant el genoma) els mecanismes reguladors de la biosíntesi d'isoprenoides en plantes superiors. El principal objectiu de la meva tesi ha estat intentar superar les limitacions en la biosíntesi d'isoprenoides, analitzant els gens implicats en la seva producció i quines son les seves funcions, per tal de poder ser utilitzats en noves aplicacions d'enginyeria metabòlica. Els objectius generals van ser: a) millorar la producció de crocines en dues espècies de Nicotiana i comparar-ne l'acumulació; b) investigar la regulació del factor de transcripció *Os*BZ8 en gens de la via biosintètica d'isoprenoides en embrió i endosperma d'arròs; c) explorar el potencial del sistema CRISPR/Cas9 per generar una sèrie de mutants de blat de moro amb contingut alterat d'estrigolactona.

Vaig enfocar els meus estudis en els enzims biosintètics de dues plantes productores de crocines, *Cs*CCD2L (*Crocus sativus*) i *Bd*CCD4.1 (*Buddleja davidii*) i vaig transferir els gens que les codifiquen a dues espècies diferents de Nicotiana. Vaig crear línies transgèniques que expressaven diferents combinacions dels gens i vaig investigar els perfils de crocines resultants per determinar com l'expressió del complement transgènic integrat influïa en l'acumulació de les crocines. Les plantes de *N. glauca* transgèniques que només expressaven *CsCCD2L* van acumular els nivells més alts de crocines. Els resultats obtinguts van indicar que *N. glauca* és una millor plataforma per a la producció de crocines que *N. tabacum*.

També vaig demostrar que el factor de transcripció *Os*BZ8 i cinc gens associats, *AACT3*, *HMGS1*, *HMGR1*, *DXS2*, *IPPI1*, que estan involucrats a les vies MVA i MEP de l'arròs, exhibien patrons d'expressió similars a l'endosperm i a l'embrió de l'arròs. A més, vaig confirmar que *Os*BZ8 s'uneix específicament a una Gbox o una G/Cbox híbrida en els promotors de *HMGR1*, *DXS2* i *IPPI1*.

Vaig fer servir la tecnologia CRISPR/Cas9 per modular l'expressió dels gens *ZmCCD7* i *ZmCCD8*, que estan involucrats en la biosíntesi d'estrigolactones al blat de moro. Induí mutacions per SpCas9 a la regió codificant de *ZmCCD8*, en dues línies independents de plantes. L'anàlisi dels mutants *ZmCCD8* contribuirà a millorar els coneixements de la base molecular que controla les característiques fisiològiques estructurals del blat de moro. Caldrà seqüenciar el genoma de les següents generacions així com l'anàlisi metabolòmica per avaluar l'impacte d'aquestes mutacions. En conjunt, la meva tesi proporciona nous coneixements sobre la regulació de la biosíntesi d'isoprenoides en plantes superiors a diferents nivells, cosa que facilitarà la predicció de l'efecte de l'enginyeria metabòlica a les plantes per a la millora nutricional o la producció de metabòlits importants.

TABLE OF CONTENTS

INDEX OF FIGURES

Generation Introduction

Chapter I

Chapter II

Chapter III

Chapter IV

INDEX OF TABLES

ABBREVIATIONS

C_5H_8	isoprene units
$C_{10}H_{16}$	monoterpenes
$C_{15}H_{24}$	sesquiterpenes
$C_{20}H_{32}$	diterpenes
$C_{30}H_{48}$	triterpenes
$C_{40}H_{64}$	tetraterpenes
IPP	isopentenyl diphosphate
MVA	mevalonate
MEP	2-C-methyl-D-erythritol 4-phosphate
CAGR	compound annual growth rate
ABA	abscisic acid
BR	brassinosteroids
CK	cytokinins
GA	gibberellins
SL	strigolactones
DXS	deoxyxylulose phosphate synthase
VAD	vitamin A deficiency
CRISPR/Cas9	Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9
TF	Transcription factors
CsCCD2L	*Crocus sativus* carotenoid cleavage dioxygenase 2L
BrCrtZ	*Brevundimonas* sp. β-carotene hydroxylase
$AtOr^{Mut}$	*Arabidopsis thaliana* ORANGE mutant
CCDs	carotenoid cleavage dioxygenases
HTCC	4-hydroxy-2,6,6-trimethyl-1-cyclohexene-1-carboxaldehyde
ALDH	aldehyde dehydrogenase
UGTs	glucosyltransferases
PSY	phytoene synthase
PDS	phytoene desaturase
LCYB	β-lycopene cyclase;
LCYE	ε-lycopene cyclase
BCH	β-carotene hydroxylase
HPLC-DAD	high-performance liquid chromatography-diode array detector
qRT-PCR	quantitative real-time polymerase chain reaction
CaMV	*cauliflower mosaic virus*
DW	dry weight

CrtZ	β-carotene hydroxylase
*Bd*CCD4.1	*Buddleja* carotenoid cleavage dioxygenase 4.1
WT	Wild type
ORF	open reading frame
hph	hygromycin phosphotransferase
DMAPP	dimethylallyl diphosphate
AACT	acetoacetyl-CoA thiolase
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzymeA
*Os*HMGS	*Oryza sativa* 3-hydroxy-3-methylglutaryl coenzymeA synthase
OsHMGR	*Oryza sativa* 3-hydroxy-3-methylglutaryl reductase
*Os*DXS2	*Oryza sativa* deoxyxylulose phosphate coenzymeA synthase 2
*Os*IPPI	*Oryza sativa* isopentenyl diphosphate isomerase
*Os*BZ8	*Oryza sativa* the bZIP class Abscisic acid Responsive Element binding factor
MVAP	Mevalonate phosphate
FPP	farnesyl diphosphate
GPP	geranyl diphosphate
GGPP	geranylgeranyl diphosphate
ABRE	Abscisic acid Responsive Element
TSS	transcription start sites
PF	promoter fragment
CDS	coding sequence
AD	activation domain
AbA	aureobasidin A
Em	embryo
En	endosperm
Y1H	yeast one-hybrid
bHLH	basic helix-loop-helix
ABRE	abscisic acid-responsive element
ABRE-CE	ABRE coupling element
bZIP	basic region/leucine zipper motifs
N	Nitrogen
P	Phosphorus
AM	arbuscular mycorrhizal
D27	DWARF 27
ZmCCD7	*Zea mays* carotenoid cleavage dioxygenase 7
ZmCCD8	*Zea mays* carotenoid cleavage dioxygenase 8
MAX1	More Axillary Growth 1
4 DO	4-deoxyorobanchol

gRNA	guide RNA
PAM	protospacer adjacent motif
*Sp*Cas9	*Streptococcus pyogenes* Cas9
*Sa*Cas9	*Staphylococcus aureus* Cas9
*St*Cas9	*Streptococcus thermophilus* Cas9
*Nm*Cas9	*Neisseria meningitides* Cas9
ZmUbi1	*Zea mays* ubiquitin 1
sgRNA	single guide RNA
PPT	phosphinothricin
IZEs	immature zygotic embryos
MS	Murashige & Skoog basal salt mixture
BAP	6-Benzylaminopurine
DSB	double strand breaks
NLS	nuclear localization sequences
RNAi	RNA interference

GENERAL INTRODUCTION

1. Definition and structural classification of isoprenoids

Isoprenoids, also known as terpenoids, are a large and very diverse class of natural products (>70,000 compounds) with many industrial and medicinal applications (Vranova et al. 2013). Despite their critical physiological functions, isoprenoids are synthesized in very small amounts in living organisms, including eukaryotes (plants, animals, fungi) and prokaryotes (bacteria and yeast) (Thulasiram et al. 2007; Moser and Pichler 2019). Multi-cellular organisms coordinate their growth, development and responses to environmental stimuli through signaling molecules for intercellular communication, many of them are isoprenoids (Van Norman et al. 2011). Most isoprenoids were first discovered in plants. Extraction from plant material remains a major method of production for many of them. However, it is increasingly challenging to meet the growing demand for many isoprenoids from plant sources due to the slow growth rate and low isoprenoid content of plants (Ajikumar et al. 2010).

According to the isoprene rule developed by Wallach and Rutzicka in the late nineteenth and mid-twentieth centuries, isoprenoids are broadly classified on the basis of the number of isoprene units (C_5H_8) in the molecule (Ruzicka 1953). Thus, all natural terpenes consist of the basic molecular formula $(C_5H_8)_n$, where 'n' is the number of isoprene units linked together and a prefix in their name denotes the number of terpene units in a particular isoprenoid molecule (e.g., 2, monoterpene; 3, sesquiterpene; 4, diterpene, etc.) (Kumari et al. 2013). These linear molecules are cyclized, oxidized and rearranged to create structural and functional diversity through coordinated intracellular modification (Carruthers and Lee 2021). Two units of the simplest hemiterpene C_5H_8 are condensed to form monoterpenes ($C_{10}H_{16}$), sesquiterpenes ($C_{15}H_{24}$), diterpenes ($C_{20}H_{32}$), triterpenes ($C_{30}H_{48}$), and tetraterpenes ($C_{40}H_{64}$) by adding successive C_5H_8 units. High molecular weight molecules are also found in nature and are known as 'polyterpenes'. An example of such polyterpene is natural rubber, which contains ca: 4,000 isoprene units (**Figure 1**).

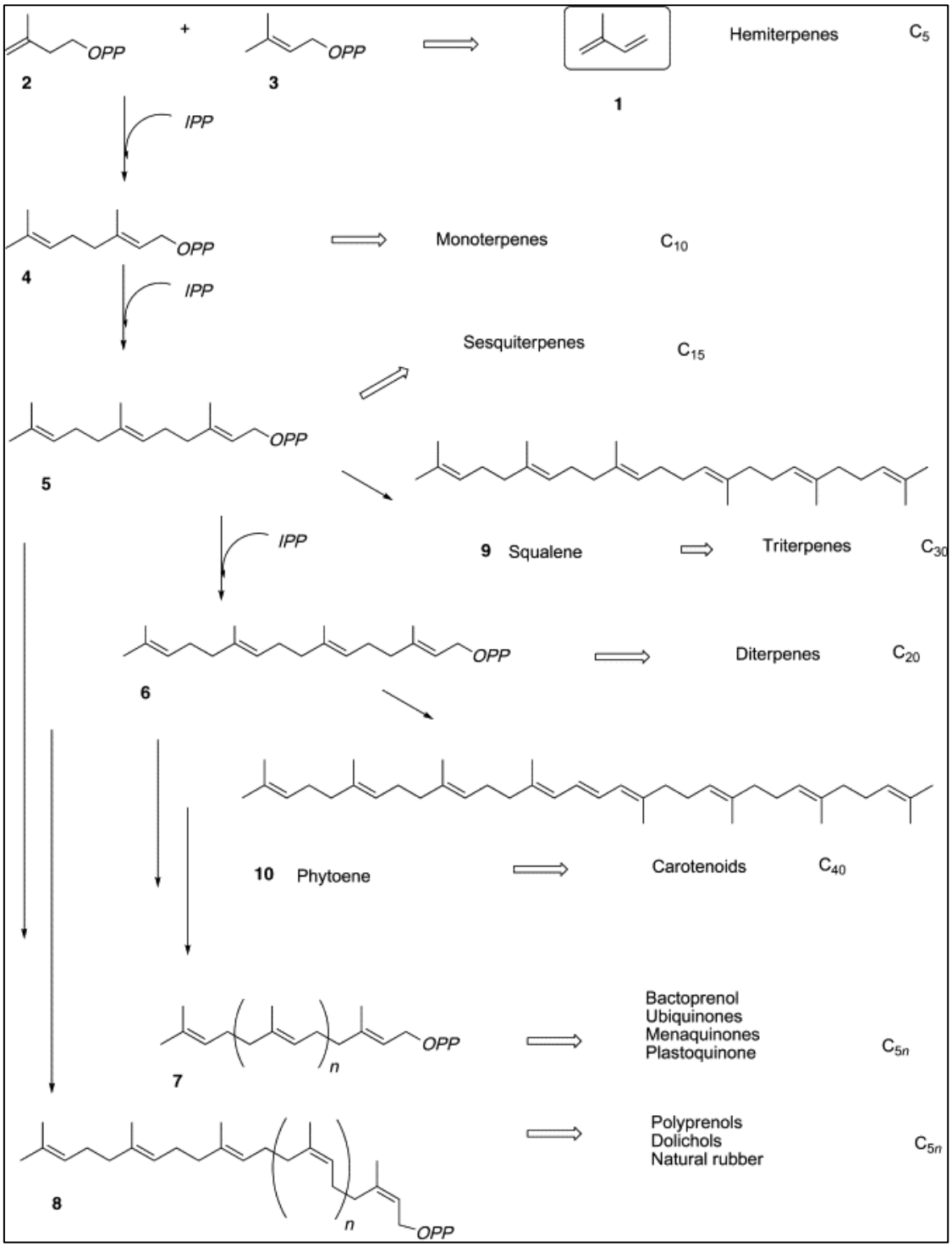


Figure 1. Biogenesis of the major isoprenoid classes. (**1**) Isoprene; (**2**) isopentenyl diphosphate; (**3**) dimethylallyl diphosphate; (**4**) geranyl diphosphate; (**5**) farnesyl diphosphate; (**6**) geranylgeranyl diphosphate; (**7**) all-*trans* polyprenyl diphosphates; (**8**) *cis*-polyprenyl diphosphates; (**9**) squalene; (**10**) phytoene (Rohmer et al. 2008)

2. Isoprenoid biosynthesis and function in higher plants

Isoprenoids are derived from the C-5 precursors isopentenyl diphosphate (IPP) and its isomer, dimethylallyl diphosphate which can be synthesized via two pathways: the mevalonate (MVA) pathway and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (Zhu et al. 2008; Farré et al. 2010). Higher plants and some algae are unique because both pathways operate in parallel in these systems. Other organisms such as yeast (MVA), animals (MVA) and green algae (MEP) use only one or the other pathway (Carruthers and Lee 2021). The precursors for sesquiterpenes and triterpenes, such as phytosterols, dolichols and the side chain of ubiquinone, are generally derived from the MVA pathway, which is localized in the cytosol and mitochondria. Precursors for isoprene, monoterpenes, carotenoids, chlorophyll and isoprenoid hormones are produced by the MEP pathway, which is located in the plastids (**Figure 2**) (Kirby and Keasling 2009). Spatial compartmentalization of the MVA and MEP pathways is advantageous because it helps plants to develop and defend themselves, by separating isoprenoid-derived primary and secondary metabolism (Vranová et al. 2013).

In plants, isoprenoids carry out a variety of biochemical functions, from fundamental primary metabolism to specialized secondary metabolism and ecological interactions with the environment (Kumari et al. 2013). Primary isoprenoid metabolites are crucial for survival and reproduction, for example, quinones as electron carriers, sterols as major membrane components, pigment carotenoids as photoprotectors and antioxidants (Pulido et al. 2012). Secondary isoprenoid metabolites, such as gibberellins, strigolactones, abscisic acid for hormonal signaling, are not considered to be essential for development. However, many secondary metabolites have attracted attention due to their high economic value as pharmaceuticals (artemisinin and pacilitaxel), nutraceuticals, biofuels (isoprenol and prenol) (Ro et al. 2006; Zheng et al. 2013; Biggs et al. 2016).

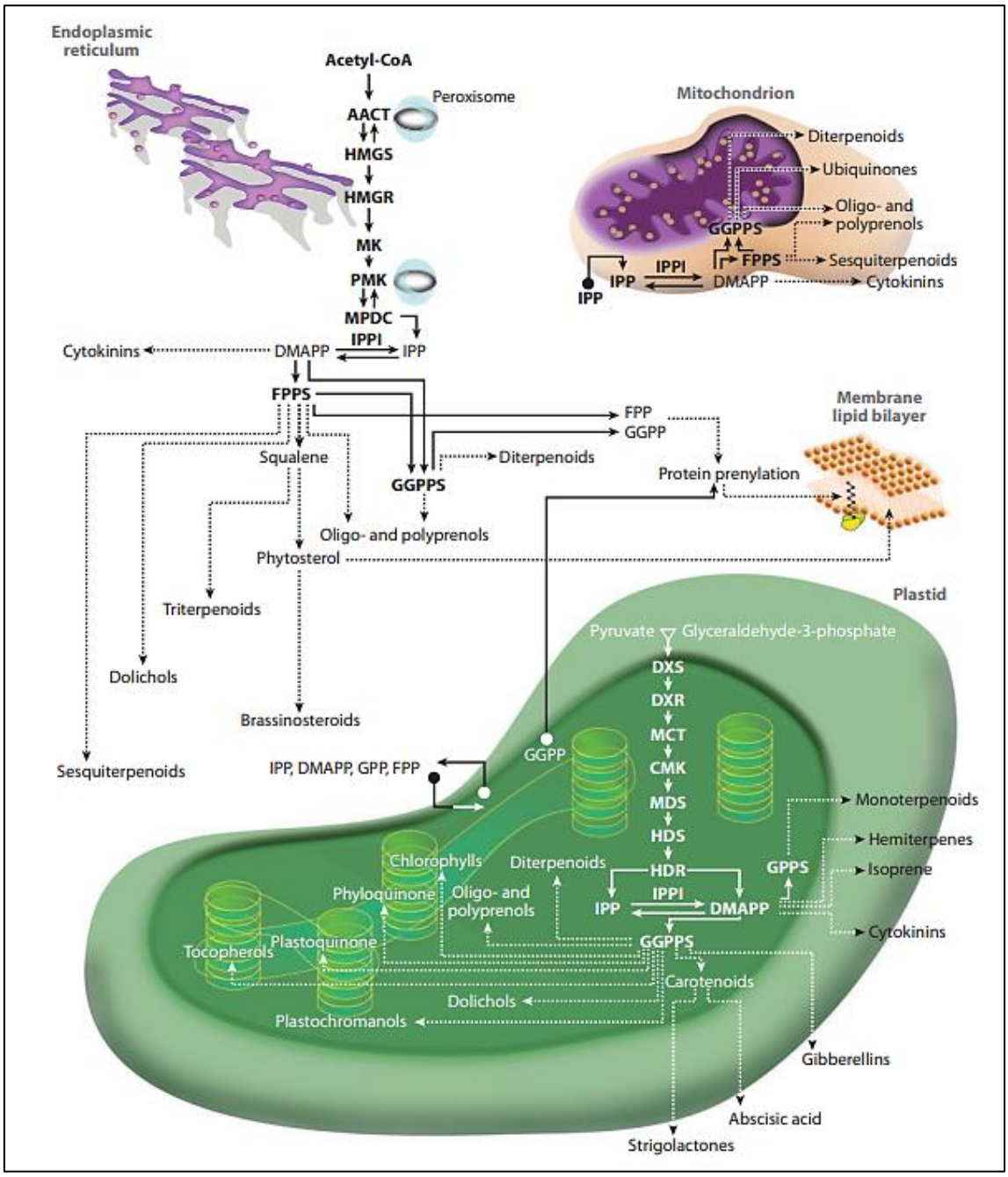


Figure 2. Subcellular compartmentalization of the MVA and MEP pathways in plant cells. Solid-line arrows indicate a single enzymatic step, dashed-line arrows indicate more than one enzymatic step, and circle-ended arrows indicate cross-membrane transport. White arrows indicate enzymatic reactions occurring in the plastids. Potential exchange of intermediate molecules derived from the MVA or MEP pathway is shown (Vranová et al. 2013).

3. Nutritional importance and applications of isoprenoids

Many isoprenoids have important uses in human health and nutrition (Gershenzon and Dudareva 2007). Due to their potential use as commercial products and their impact on human health, some of these natural products have been very attractive targets for bioengineering. For example, carotenoids have several uses in the food, animal feed, cosmetics, and health supplement sectors (Kirby and Keasling 2009). Examples include β-carotene, astaxanthin, and tocopherols. The global carotenoids market had been projected to grow at a compound annual growth rate (CAGR) of 5.7% from US $2.0 billion in 2022 to US $2.7 billion by 2027, due to increasing consumer awareness of the health advantages of carotenoids as antioxidants and such as the use of lutein in particular to prevent age-associated macular degeneration disease and related problems (https://www.bccresearch.com/market-research/food-and-beverage/the-global-market-for-carotenoids.html).

Animal metabolism depends heavily on a number of isoprenoids. Carotenoids are an indispensable group of lipid-soluble pigments which humans cannot synthesize, thus acquiring them through the diet (Ahrazem et al. 2022). Many important carotenoids are tetraterpenoids, produced from 8 isoprene molecules, thus containing 40 carbon atoms (Sun et al. 2022). Vitamin A, which is essential for good vision, and proper growth, reproduction, and neurological development, is derived from tetraterpene carotenoid pigments. Vitamin K, required for blood clotting, and vitamin E, essential for reproduction, are vitamins that are entirely or in part derived from isoprenoid biosynthesis. Isoprenoids are also involved in the biosynthesis of ubiquinones (coenzyme Q), which are essential for the oxidation of food to produce energy. Fish and other animal liver is particularly abundant in oils, the majority of which are acyclic triterpenoid hydrocarbons, mainly squalene. Certain isoprenoid compounds impact the development and mating behavior of insects, signal danger or deter predators, or indicate the route between the nest and food sources (Rizvi et al. 2021).

Ancient Egyptians and Greeks were aware that a dietary element affected visual acuity, but it was not until the 1930s that the significance of carotenoids and their metabolites was recognized, and β-carotene and retinol were thoroughly defined. Vitamin A activity is now recognized in the metabolites retinol, retinal, and retinoic acid, as well as several provitamin A carotenoids, most notably β-carotene. In addition, retinol, retinoic acid, and their metabolites are now understood to have a plethora of additional roles in human metabolism, from embryogenesis through adulthood, including growth and development,

reproduction, cancer, and infection resistance. They are major natural antioxidants that have positive health effects.

Crocins are the most important active ingredient found in *Crocus sativus*. They are water-soluble apocarotenoid pigments with high economic value used in the food and, to a lesser degree, in the pharmaceutical industries. Apocarotenoids are a subclass of isoprenoids that are produced in plants by the oxidative cleavage of carotenoids (Auldridge et al. 2006). *Crocus* is a genus in the Iridaceae Juss family, which contains ca: 85 flowering plant species (Gohari et al. 2013). Saffron stigmas have been shown to have a wide range of biological actions, including antioxidant, cytotoxic, antibacterial, anti-depressive, hypolipidemic, and antiparasitic properties. *Crocus* spp. are commonly cultivated for their valuable crocin content (Mykhailenko et al. 2019). The autumn-flowering *C. sativus* is the most expensive natural source of crocins, which accumulate at relatively high levels during the development of the stigma and confer a dark red color (Moraga et al. 2009). It takes between 120,000 and 200,000 flowers to produce 1 kg of dried saffron stigma (Ahrazem et al. 2022). Studies also have shown that *C. sativus* stigmas are useful in the treatment of mental disorders, neurological illnesses, learning and memory dysfunctions, cardiovascular diseases, atherosclerosis, hyperlipidemia, type 2 diabetes, hypertension, ulcer, fatty liver disease, epilepsy, and convulsions (Mykhailenko et al. 2019). In addition to *Crocus* spp., *Gardenia jasminoides* fruits and *Buddleja davidii* petals are additional sources of crocins, but at a much lower scale compared with the stigma of *C. sativus* (Guijarro-Diez et al. 2017; Diretto et al. 2021).

4. Isoprenoids as plant hormones

In plants, isoprenoids play multiple functions in photosynthesis, photoprotection, pigmentation, and signaling (Sun et al. 2022). Several phytohormones are important isoprenoid metabolites. Five of the eight major hormones commonly recognized in plants, are produced exclusively or largely from isoprenoids. These include abscisic acid (ABA), brassinosteroids (BR), cytokinins (CK), gibberellins (GA), and strigolactones (SL).

SL are structurally related plant signaling molecules which were identified as inducers of root parasitic plant seed germination of the genera *Striga*. They are involved in a number of different functions as exogenous signals that increase hyphal branching in arbuscular mycorrhizal fungi (**Figure 3**) (Akiyama et al. 2005). As internal signals, SL inhibit shoot branching, lateral and adventitious roots formation, and boost the elongation of the primary root which benefits plants by improving phosphate and nitrate uptake (Gomez-

Roldan et al. 2008; Rasmussen et al. 2012). Therefore, they must be delicately regulated to promote nutrient uptake efficiently and improve water and fertilizer acquisition. Otherwise, manipulation of SL metabolism can induce parasitic weed germination. Other processes that involve SL signaling functions include stem elongation, leaf expansion and senescence, and responses to drought and salinity (Brewer et al. 2013; Screpanti et al. 2016).

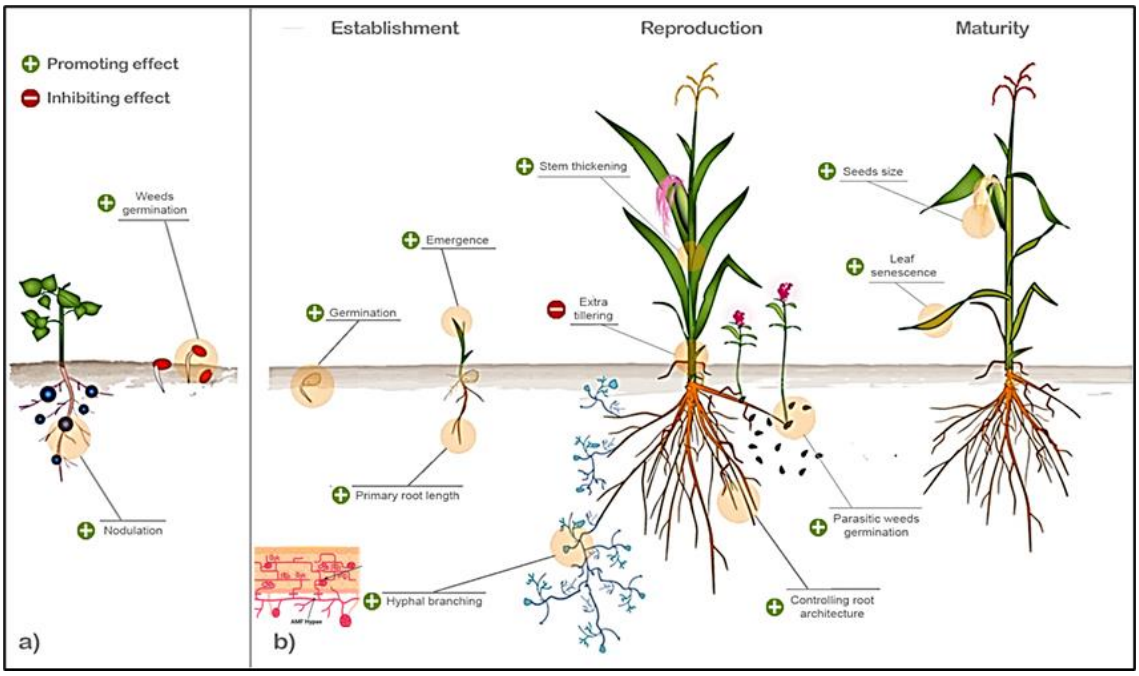


Figure 3. SL functions in the rhizosphere (a) and in plant development (b). (Modified from Screpanti et al. 2016).

5. Metabolic engineering strategies to modulate isoprenoid levels and composition in plants

Studies on the isoprenoid precursor pathways have dominated plant genetic engineering efforts to enhance isoprenoid biosynthesis. Overexpression of deoxyxylulose phosphate synthase (DXS), which initiates the MEP pathway, has been linked to an increase in plant isoprenoid levels in a number of studies (Farré et al. 2011). Sesquiterpenes are typically produced in the cytosol, whereas monoterpenes are mostly produced in the plastid. Therefore, the engineering of isoprenoid products depends on the flux of precursors provided by the core isoprene biosynthetic pathways. Introduction of heterologous genes, modification of existing native pathways, or eliminating competing pathways are some of the strategies to increase the pool of essential precursors and/or target products. The most direct approach to increase the levels of target isoprenoids in any given organism, is the overexpression of genes encoding enzymes in the biosynthetic pathway (Bailey et al. 2002). Previous studies in our laboratory resulted in transgenic maize plants that

simultaneously engineer three different vitamin metabolic pathways. Transgenic kernels contained 169-fold more β-carotene, 6-fold more ascorbate, and double the amount of folate as compared to wild-type plants (Naqvi et al. 2009). Similarly, other research groups have generated transgenic tomatoes with the aroma/flavor compounds linalool and geraniol in the fruit (Lewinsohn et al. 2001; Davidovich-Rikanati et al. 2007). With the intention of preventing vitamin A deficiency (VAD) in developing nations, conventional breeding and genetic engineering have been used to produce crops with increased carotene levels. For example, "Golden rice" and "Golden banana" have been developed to increase their β-carotenoid to combat VAD. The isoprenoid biosynthetic pathway is well understood and has been widely explored in many plant species. In recent years, attention has turned to the regulatory control of carotenoid metabolism especially with the discovery of the CRISPR-Cas system (Kumar et al. 2022). Transcription factors (TF) can also influence isoprenoid production in plants by modulating pathway gene expression. Single or multiple pathway gene expression can be altered using TF which bind cis elements in the promoter regions of the genes they regulate. TF such as WRKY, bZIP, MYC2, bHLH families are involved in the regulation of MEP and MVA pathways in response to stresses (Okada et al. 2009; Yamamura et al. 2015; Fu et al. 2018). These TF regulate the biosynthesis of phytoalexins, carotenoid and steroidal glycoalkaloids by directly targeting the MEP and MVA pathways. However, the mechanisms regulating the efficient cooperation of both pathways are not fully understood. It is, therefore, very important to examine TF that coordinate the metabolic flux through both pathways.

6. References

Ahrazem O, Zhu C, Huang X, Rubio-Moraga A, Capell T, Christou P, Gómez-Gómez L. (2022). Metabolic Engineering of Crocin Biosynthesis in Nicotiana Species. *Front. Plant Sci* **13**:861140.

Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* **330**: 70-74.

Akiyama K, Matsuzaki K-i, Hayashi H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435**: 824-827.

Auldridge ME, McCarty DR, Klee HJ. (2006). Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Curr Opin Plant Biol* **9**: 315-321.

Bailey JE, Sburlati A, Hatzimanikatis V, Lee K, Renner WA, Tsai PS. (2002). Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnol Bioeng* **79**: 568-579.

Bick JA, Lange BM. (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch Biochem Biophys* **415**: 146-154.

Biggs BW, Lim CG, Sagliani K, Shankar S, Stephanopoulos G, De Mey M, Ajikumar PK. (2016). Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **113**: 3209-3214.

Brewer PB, Koltai H, Beveridge CA. (2013). Diverse roles of strigolactones in plant development. *Mol Plant* **6**: 18-28.

Carruthers DN, Lee TS. (2021). Diversifying Isoprenoid Platforms via Atypical Carbon Substrates and Non-model Microorganisms. *Front Microbiol* **12**: 791089.

Davidovich-Rikanati R, Sitrit Y, Tadmor Y, Iijima Y, Bilenko N, Bar E, Carmona B, Fallik E, Dudai N, Simon JE. (2007). Enrichment of tomato flavor by diversion of the early plastidial terpenoid pathway. *Nat. Biotechnol.* **25**: 899-901.

Diretto G, López-Jiménez AJ, Ahrazem O, Frusciante S, Song J, Rubio-Moraga Á, Gómez-Gómez L. (2021). Identification and characterization of apocarotenoid modifiers and carotenogenic enzymes for biosynthesis of crocins in Buddleja davidii flowers. *J. Exp. Bot.* **72**: 3200-3218.

Farré G, Bai C, Twyman RM, Capell T, Christou P, Zhu C. (2011). Nutritious crops producing multiple carotenoids--a metabolic balancing act. *Trends Plant Sci* **16**: 532-540.

Farré G, Sanahuja G, Naqvi S, Bai C, Capell T, Zhu C, Christou P. (2010). Travel advice on the road to carotenoids in plants. *Plant Sci.* **179**: 28-48.

Fu J, Liu Q, Wang C, Liang J, Liu L, Wang Q. (2018). ZmWRKY79 positively regulates maize phytoalexin biosynthetic gene expression and is involved in stress response. *J. Exp. Bot.* **69**: 497-510.

Gershenzon J, Dudareva N. (2007). The function of terpene natural products in the natural world. *Nat Chem Biol* **3**: 408-414.

Gohari AR, Saeidnia S, Mahmoodabadi MK. (2013). An overview on saffron, phytochemicals, and medicinal properties. *Pharmacogn Rev.* **7**: 61.

Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pagès V, Dun EA, Pillot J-P, Letisse F, Matusova R, Danoun S, Portais J-C. (2008). Strigolactone inhibition of shoot branching. *Nature* **455**: 189-194.

Guijarro-Diez M, Castro-Puyana M, Crego AL, Marina ML. (2017). Detection of saffron adulteration with gardenia extracts through the determination of geniposide by liquid chromatography–mass spectrometry. *J Food Compost Anal* **55**: 30-37.

Kirby J, Keasling JD. (2009). Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu Rev Plant Biol* **60**: 335-355.

Kumar D, Yadav A, Ahmad R, Dwivedi UN, Yadav K. (2022). CRISPR-Based Genome Editing for Nutrient Enrichment in Crops: A Promising Approach Toward Global Food Security. *Front Genet* **13**: 932859.

Kumari S, Priya P, Misra G, Yadav G. (2013). Structural and biochemical perspectives in plant isoprenoid biosynthesis. *Phytochem Rev.* **12**: 255-291.

Lütke-Brinkhaus F, Liedvogel B, Kleinig H. (1984). On the biosynthesis of ubiquinones in plant mitochondria. *Eur J Biochem* **141**: 537-541.

Lewinsohn E, Schalechet F, Wilkinson J, Matsui K, Tadmor Y, Nam K-H, Amar O, Lastochkin E, Larkov O, Ravid U, Hiatt W, Gepstein S, Pichersky E. (2001). Enhanced Levels of the Aroma and Flavor Compound S-Linalool by Metabolic Engineering of the Terpenoid Pathway in Tomato Fruits. *Plant Physiol.***127**: 1256-1265.

Moraga ÁR, Rambla JL, Ahrazem O, Granell A, Gómez-Gómez L. (2009). Metabolite and target transcript analyses during Crocus sativus stigma development. *Phytochemistry* **70**: 1009-1016.

Moser S, Pichler H. (2019). Identifying and engineering the ideal microbial terpenoid production host. *Applied Microbiology and Biotechnology* **103**: 5501-5516.

Mykhailenko O, Kovalyov V, Goryacha O, Ivanauskas L, Georgiyants V. (2019). Biologically active compounds and pharmacological activities of species of the genus Crocus: A review. *Phytochemistry* **162**: 56-89.

Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Perez Conesa D, Ros G, Sandmann G, Capell T, Christou P. (2009). Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 7762-7767.

Okada A, Okada K, Miyamoto K, Koga J, Shibuya N, Nojiri H, Yamane H. (2009). OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice. *J. Biol. Chem.* **284**: 26510-26518.

Pulido P, Perello C, Rodriguez-Concepcion M. (2012). New insights into plant isoprenoid metabolism. *Mol Plant* **5**: 964-967.

Rasmussen A, Beveridge CA, Geelen D. (2012). Inhibition of strigolactones promotes adventitious root formation. *Plant signaling & behavior* **7**: 694-697.

Rizvi SAH, George J, Reddy GV, Zeng X, Guerrero A. (2021). Latest developments in insect sex pheromone research and its application in agricultural pest management. *Insects* **12**: 484.

Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940-943.

Ruzicka L. (1953). The isoprene rule and the biogenesis of terpenic compounds. *Experientia* **9**: 357-367.

Screpanti C, Fonné-Pfister R, Lumbroso A, Rendine S, Lachia M, De Mesmaeker A. (2016). Strigolactone derivatives for potential crop enhancement applications. *Bioorganic & medicinal chemistry letters* **26**: 2392-2400.

Sun T, Rao S, Zhou X, Li L. (2022). Plant carotenoids: Recent advances and future perspectives. *Molecular Horticulture* **2**: 3.

Thulasiram HV, Erickson HK, Poulter CD. (2007). Chimeras of two isoprenoid synthases catalyze all four coupling reactions in isoprenoid biosynthesis. *Science* **316**: 73-76.

Van Norman JM, Breakfield NW, Benfey PN. (2011). Intercellular communication during plant development. *Plant Cell* **23**: 855-864.

Vranová E, Coman D, Gruissem W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol* **64**: 665-700.

Vranova E, Coman D, Gruissem W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol* **64**: 665-700.

Yamamura C, Mizutani E, Okada K, Nakagawa H, Fukushima S, Tanaka A, Maeda S, Kamakura T, Yamane H, Takatsuji H. (2015). Diterpenoid phytoalexin factor, a bHLH transcription factor, plays a central role in the biosynthesis of diterpenoid phytoalexins in rice. *Plant J.* **84**: 1100-1113.

Zheng Y, Liu Q, Li L, Qin W, Yang J, Zhang H, Jiang X, Cheng T, Liu W, Xu X. (2013). Metabolic engineering of Escherichia coli for high-specificity production of isoprenol and prenol as next generation of biofuels. *Biotechnology for biofuels* **6**: 1-13.

Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T. (2008). Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proc Natl Acad Sci U S A* **105**: 18232-18237.

AIMS AND OBJECTIVES

Aims

The major aim of my thesis was to better understand the isoprenoid biosynthesis pathway's bottlenecks by analyzing genes involved in isoprenoid accumulation and their functionality for future use in metabolic engineering applications.

To reach these aims, the following specific objectives were addressed:

- Transform *CsCCD2L* into *N. glauca* and *N. tabacum*; Co-transform *CsCCD2L*, *BrCrtZ*, *AtOr*^{Mut} into *N. tabacum* and analyze the expression of transgenes and the endogenous *PSY1*, *PSY2*, *LCYB*, and *BCH* carotenoid biosynthetic genes; determine accumulation of crocins and carotenoids in transgenic *N. glauca* and *N. tabacum;* determine the stability of these metabolites during storage in leaf and petal tissues.
- Analyze the promoter regions of particular MVA and MEP pathway genes to determine whether the promoters include specific elements that can be recognized and bound by the OsBZ8 transcription factor; determine the expression patterns of *OsBZ8* in different rice tissues; compare the expression pattern of *OsBZ8* with that of selected genes (*AACT3, HMGS1, HMGR1, DXS2, IPPI1*) in rice endosperm and embryo.
- Induce targeted mutations in the coding regions and/or corresponding promoter regions of *ZmCCD7* and *ZmCCD8* to generate maize mutants with altered strigolactone contents and compositions in the elite M37W maize inbred line.

Chapter I

Metabolic engineering of crocin biosynthesis in *Nicotiana* species

1.0. Abstract

Crocins are high-value soluble pigments that are used as colorants and supplements. Their presence in nature is extremely limited, and consequently, the high cost of these metabolites hinders their use by other sectors, such as the pharmaceutical and cosmetic industries. The carotenoid cleavage dioxygenase 2L (*CsCCD2L*) is the rate-limiting enzyme in the biosynthetic pathway of crocins in *Crocus sativus*. *CsCCD2L* was introduced into *N. tabacum* and *N. glauca* in order to further understand the biosynthetic pathway of these molecules. In addition, a chimeric construct containing the *Brevundimonas* sp. β-carotene hydroxylase gene (*BrCrtZ*), the *Arabidopsis thaliana ORANGE* mutant gene ($AtOr^{Mut}$), and *CsCCD2L* was also introduced into *N. tabacum*. Quantitative and qualitative studies on carotenoids and apocarotenoids in the transgenic plants expressing *CsCCD2L* alone showed higher crocin level accumulation in transgenic *N. glauca* plants, reaching ca: 400 µg/g DW in leaves, while in *N. tabacum* accumulation was ca: 36 µg/g DW. In contrast, *N. tabacum* plants co-expressing *CsCCD2L*, *BrCrtZ*, and $AtOr^{Mut}$ accumulated 3.5-fold higher crocin levels compared to *N. tabacum* plants only expressing *CsCCD2L*. Crocins with three and four sugar molecules were the main molecular species in both host systems. These results demonstrate that the production of saffron apocarotenoids is feasible in engineered *Nicotiana* species and establishes a basis to develop strategies that may ultimately lead to the commercial exploitation of these valuable pigments for multiple applications.

1.1. Introduction

Carotenoids are an important group of natural and lipid-soluble pigments that are widely used as food colorants, nutraceuticals, animal feed additives, cosmetic ingredients, and health supplements (Fraser and Bramley 2004; Zhu et al. 2007; Farre et al. 2014). They are synthesized by plants, algae, fungi, and bacteria, and exhibit a yellow to red coloration, depending on the type of carotenoids and the concentration they reach in the different cells. In general, animals do not synthesize carotenoids *de novo*, so those found in animals are either directly accumulated from food or partly modified through metabolic reactions (Maoka 2011). Carotenoids play multiple physiological and nutritional functions as biological antioxidants, especially in relation to human health in the prevention of human diseases and maintaining good health by enhancing our immune system (Fiedor and Burda 2014). In all living organisms, carotenoids act as substrates to produce apocarotenoids (Ahrazem et al. 2016a). These secondary metabolites are not simply carotenoid breakdown products. Rather, they provide color and aroma to flowers and fruits, but also act as signaling molecules and hormones (Walter et al. 2010). Apocarotenoids are present as volatile soluble compounds. They include crocins, which are glycosylated derivatives of the apocarotenoid crocetin present in a limited number of plants (Ahrazem et al. 2017). *Crocus sativus* L. is the main natural source of crocins, which accumulate at high levels during the development of the stigma and confer a dark red coloration (Moraga et al. 2009). Together with the monoterpene aldehydes picrocrocin (β-D-glucopyranoside of hydroxyl-β-cyclocitral) and its deglycosylated derivative, safranal (2,6,6-trimethyl-1,3-cyclohexadiene-1 carboxaldehyde) are responsible for its color, bitter flavor, and aroma of the saffron spice, respectively. In addition, they exhibit a strong coloring potential, and crocins have been extensively used in medical applications for their capacity to relieve different diseases in humans, mainly due to their action as free radical quenchers that can modulate the redox status of the cells which are under oxidative stress (Bukhari et al. 2018). Crocins are associated with a broad range of health benefits because of their analgesic and sedative properties (Amin and Hosseinzadeh 2012), neurological protection, and anticancer activities (Finley and Gao 2017). Furthermore, clinical trials indicate that crocins have a positive effect on the treatment of depression and dementia (Lopresti and Drummond 2014). The health-promoting properties of crocins, along with their ability to act as a natural water-soluble colorant, have led to intense interest in elucidating the biosynthetic pathway in order to develop new and renewable sources producing these valuable metabolites (Diretto et al. 2021).

Commercial utilization of crocins from saffron is limited due to elevated prices, a consequence of the high labor costs in harvesting and processing flower material (Ahrazem et al. 2015). One kg of saffron spice can cost over $10,000. Saffron yield is greatly affected by cultivation methods and environmental conditions. In Castile-La Mancha (Spain), approximately 173,250 saffron flowers weighing over 68 kg are needed to produce 1 kg of saffron spice (Serrano-Díaz et al. 2012). Studies have shown that the yield of saffron, which is obtained from the dry stigmas of Crocus flowers, varies greatly depending on the region and cultivation methods used. In Morocco, the yield of saffron can reach 6 kg/ha (Lage and Cantrell 2009), in India, 3 kg/ha and in Italy, 10 kg/ha. Greece and Iran have reported yields ranging from 4 to 7 kg/ha (Kothari et al. 2021; Kumar et al. 2022). In addition to *Crocus*, gardenia (*Gardenia jasminoides*) fruits are also a commercial source of crocins, and transcriptome analyses have identified the genes that encode enzymes responsible for crocin production in this plant. However, gardenia does not accumulate picrocrocin (Pfister et al. 1996; Moras et al. 2018; Xu et al. 2020). Among other plants that produce crocins, but which are not commercially exploited due to the low quantities they accumulate, are *Buddleja* species. In *B. davidii*, crocins accumulate in the basal part of the flower corolla (Ahrazem et al. 2017). Therefore, only a few plant species can synthesize crocins, and this is because the carotenoid cleavage dioxygenases (CCDs) responsible for their production are not usually present in plants. Plant CCDs are divided into five groups, CCD1, CCD2, CCD4, CCD7 and CCD8, differing in their substrate specificities and cleavage sites (Walter et al. 2010).

In *Crocus*, *Gardenia* and *Buddleja* species, zeaxanthin is the precursor of crocetin (**Figure 1.1**). Cleavage of the zeaxanthin molecule at the 7,8 and 7′,8′ double bonds generates one molecule of crocetin dialdehyde and two molecules of 4-hydroxy-2,6,6-trimethyl-1-cyclohexene-1-carboxaldehyde (HTCC) (Frusciante et al. 2014; Ahrazem et al. 2016a; Ahrazem et al. 2017). Crocetin dialdehyde is further converted to crocetin by aldehyde dehydrogenase (ALDH) enzymes, and crocetin is the substrate of glucosyltransferases (UGTs) that catalyze the formation of crocins, catalyzing the transfer of glucose molecules to both ends of the crocetin molecule and produce crocins with different numbers of attached glucose molecules (Moraga et al. 2004; Nagatoshi et al. 2012). HTCC is also recognized by UGTs in the stigma, resulting in the formation of picrocrocin, which is further metabolized to produce the volatile compound safranal (**Figure 1.1**) (López-Jimenez et al. 2021). In *C. sativus* CCD2 catalyzes the sequential cleavage of zeaxanthin at the 7,8;7′,8′ double bonds (Frusciante et al. 2014). *Cs*CCD2L

is a plastidic enzyme that belongs to a new CCD subfamily only isolated from *Crocus* species (Ahrazem et al. 2016a), which is closely related to the CCD1 subfamily (Ahrazem et al. 2016a; Ahrazem et al. 2016b). In *Buddleja davidii* and *Gardenia jasminoides*, the CCD enzymes that catalyze the same reaction belong to the CCD4 subfamily (Ahrazem et al. 2016a; Xu et al. 2020). The *Buddleja* enzymes, *Bd*CCD4.1 and *Bd*CCD4.3 are also plastidic enzymes expressed in flowers (Ahrazem et al. 2017).

The carotenoid biosynthetic pathway has been modified by genetic engineering in higher plants to increase the amount of carotenoids and/or to produce specific carotenoids (Farre et al. 2014). In earlier studies, *CsCCD2L* (a longer CCD2 version) was transiently expressed on its own and was sufficient to allow significant crocin accumulation in *Nicotiana benthamiana* (Marti et al. 2020). In this chapter, I described experiments in *Nicotiana* species to express *CsCCD2L* alone or in combination with genes involved in the carotenoid biosynthesis and accumulation, such as *BrCrtZ* and *AtOr*^{Mut}, in the latter case to increase the availability of substrates for *Cs*CCD2L activity. Metabolic analyses showed the accumulation of these nutritional substances reaching almost 400 µg/g DW in *N. glauca* which indicates that *N. glauca* is a better system than *N. tabacum* to produce crocins because the transgenic *N. glauca* plants expressing *Cs*CCD2L alone could accumulate much higher amounts of crocins as compared with the transgenic *N. tabacum* plants co-expressing *CsCCD2L*, *BrCrtZ*, and *AtOr*^{Mut}.

1.2. Aims and objectives

The aim of this chapter was to enhance the production of crocins in *N. glauca* and *N. tabacum* leaves and to compare the accumulation of crocins in the two species.

To reach these aims, the following specific objectives were pursued:

- transform *CsCCD2L* into *N. glauca* and *N. tabacum*.
- Co-transform *CsCCD2L*, *BrCrtZ*, *AtOr*^{Mut} into *N. tabacum* and analyze the expression of transgenes and the endogenous *PSY1*, *PSY2*, *LCYB,* and *BCH* carotenoid biosynthetic genes.
- determine accumulation of crocins and carotenoids in transgenic *N. glauca* and *N. tabacum.*

1.3. Materials and methods

1.3.1. Vector construction

Two plasmids were created to engineer the saffron apocarotenoid pathway (**Figure 1.1**) in *N. glauca* and *N. tabacum* (**Figure 1.2**). The tobacco Ubi.4U, *Arabidopsis At*UBQ10 and CaMV35S promoters were used to control the expression of *BrCrtZ, AtOr^{Mut} and CsCCD2L*, respectively. The GoldenBraid strategy was followed to construct the vectors (Sarrion-Perdigones et al. 2013; Sarrion-Perdigones et al. 2014). Briefly, the complete open reading frames of *CsCCD2L*, *BrCrtZ* and *AtOr^{Mut}* were domesticated by removing the BsmBI and BsaI restriction sites in the original sequence using the primers listed in **Table 1.1**. The products were cloned in the level 0 vector pUPD2 of the GoldenBraid modular cloning system. The resulting plasmids pUPD2-CsCCD2L, pUPD2-BrCrtZ and pUPD2-AtOr^{Mut} were then used to construct two recombinant binary vectors as follows: pDGB3Ω1[p35S:CCD2L: T35S-pNos: Hyg:T35S], and pDGB3α1[p35S:CCD2L:T35S-pNos:Hyg:T35S-pTUBI4U:BrCrtZ:T35S pAtUBQ10:AtOrMut:T35S]. All vectors were constructed by Dr. O. Ahrazem (Departamento de Ciencia y Tecnología Agroforestal y Genética, Instituto Botánico, Universidad de Castilla-La Mancha, Campus Universitario, Albacete, Spain).

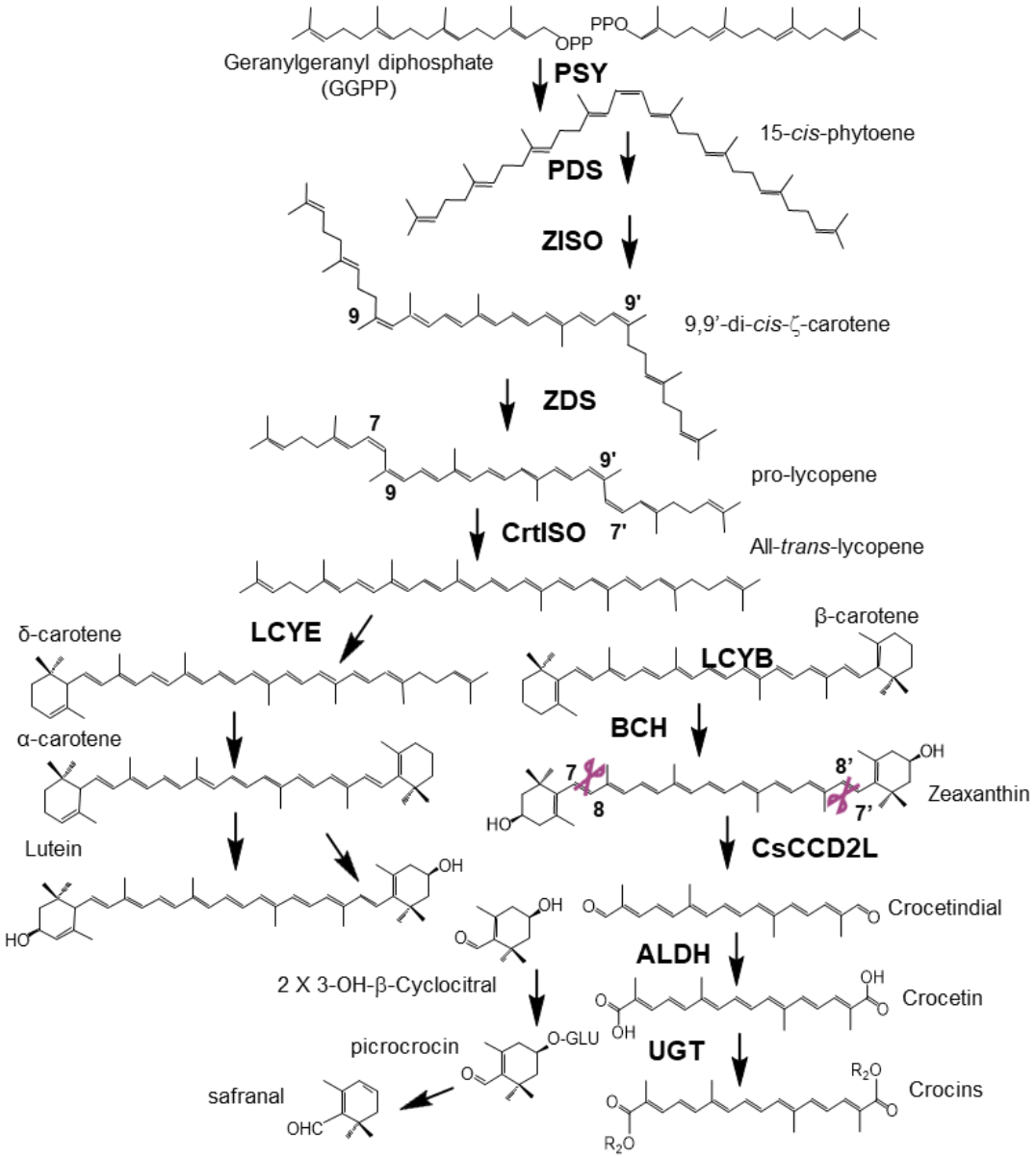


Figure 1.1. Biosynthetic pathway of crocins in saffron stigma. PSY, phytoene synthase; PDS, phytoene desaturase; Z-ISO, ζ-carotene isomerase; ZDS, ζ-carotene desaturase; CrtISO, carotene isomerase; LCYB, β-lycopene cyclase; LCYE, ε-lycopene cyclase; BCH, β-carotene hydroxylase; ALDH, aldehyde dehydrogenase. 3-OH-β-cyclocitral also known as HTCC (hydroxy-2,6,6-trimethyl-1-cyclohexen-1-carboxaldehyde). Cleavage positions recognized by *Cs*CCD2L are shown in purple.

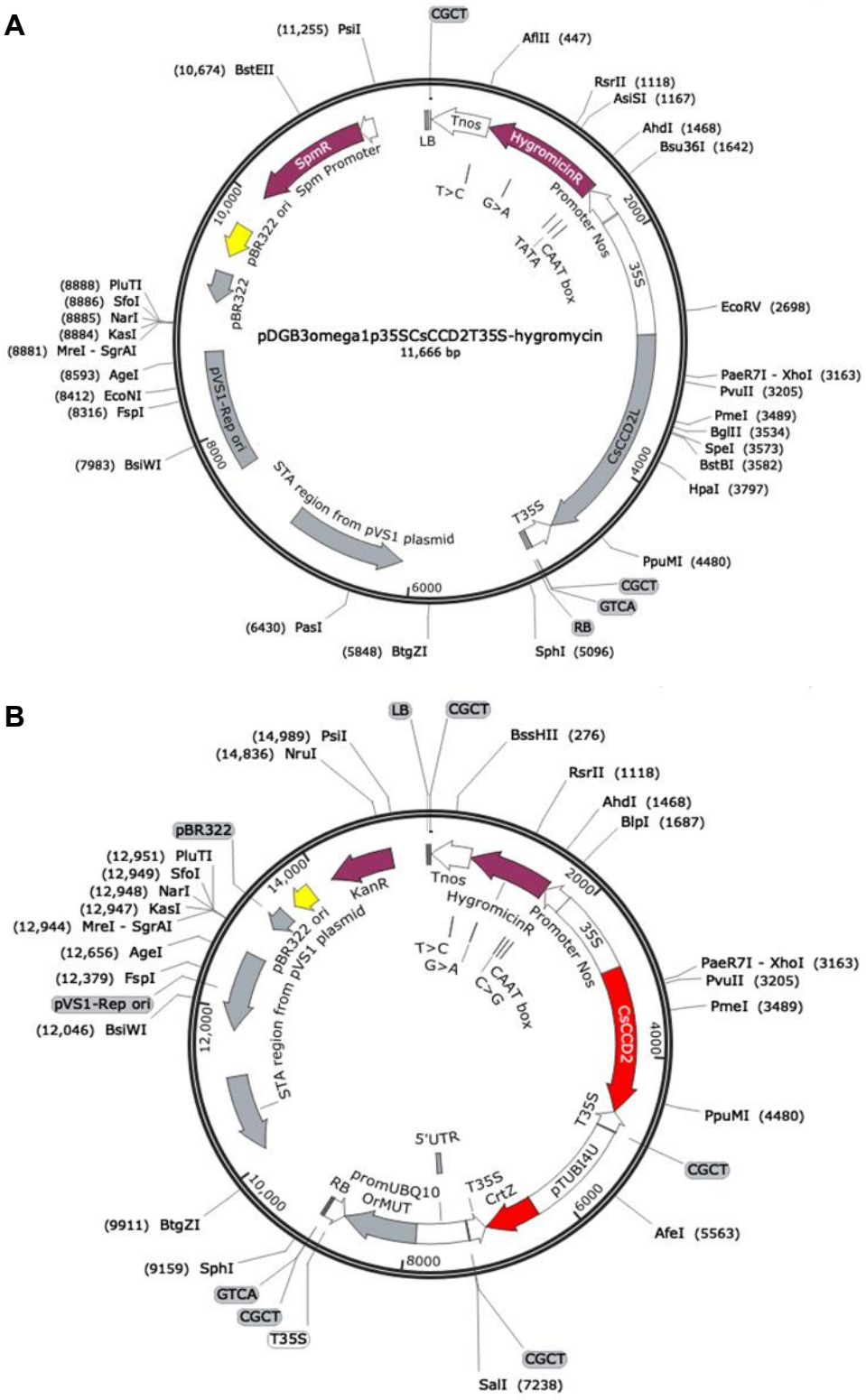


Figure 1.2. Schematic representation of the constructs used to transform **(A)** *N. glauca* and **(B)** *N. tabacum* plants.

Table 1.1. Oligonucleotides used for gene expression analyses and for *CsCCD2L*, *BrCrtZ* and *AtOr*^{Mut} domestication.

Gene	Primer sequence (5′-3′)	Reference
*CsCCD2L*_F	ACATGTCGCCTTGAGAGTCC	This work
*CsCCD2L*_R	TCAGATTTGATGCCAGGTTG	
*BrCrtZ*_F	TGGCTTACTTGGATCGCTCT	This work
*BrCrtZ*_R	TCCAACAGCAACCATAACGA	
AtOr^{Mut}_F	GTTGGTGTGATTTCGGCTTT	This work
AtOr^{Mut}_R	GTTTTGGGCGGTGATAGAGA	
*PSY1*_F	GAAGCCGGAGATCCCTCTCC	(Wang et al. 2021)
*PSY1*_R	TTGCCCAAATAGCCCTTCTTC	
*PSY2*_F	GGCAGCTGAGATCTACCGATG	(Wang et al. 2021)
*PSY2*_R	TGCACATACTTCGCCACAACG	
*LCYB*_F	GACAATACAACTAAAGATCTTGATAG	(Wang et al. 2021)
*LCYB*_R	CATAAGCTACTTGATATCCAGGAT	
*BCH*_F	ATGGCCGCCAGCAGAATTTC	(Shi et al. 2014)
*BCH*_R	CTCAATTTTCATTTCAATCTCCTCTGTC	
*Actin*_F	GTATGTCGCCATTCAAGCCGTTCT	(Moreno et al. 2016)
*Actin*_R	ACGGAGGATAGCATGTGGCAAAGCAT	
*BrCrtZ domestication*_F	GCGCCGTCTCGCTCGAATGGCTTCTATGATATCCTCTTC	This work
*BrCrtZ domestication*_R	GCGCCGTCTCGCTCAAAGCTCAAGCTCCGCTAGAAGAAG	
AtOr^{Mut} *domestication*_F	GCGCCGTCTCGCTCGAATGTCATCTTTGGGTAGGATTTTG	This work
AtOr^{Mut} *domestication*_R	GCGCCGTCTCGCTCAAAGCTCAATCGAAAGGGTCGATACG	
*CsCCD2L-1 domestication*_F	GCGCCGTCTCGCTCGAATGGAATCTCCTGCTACTAAATTA	This work
*CsCCD2L-1 domestication*_R	GCGCCGTCTCGTTGTCTCTGCCTCCTCCTTA	
*CsCCD2L-2 domestication*_F	GCGCCGTCTCGACAAGTAAGAAGAAGCCCAAAC	This work
*CsCCD2L-2 domestication*_R	GCGCCGTCTCGCTCAAAGCTCATGTCTCTGCTTGGTGCT	

1.3.2. Genetic transformation of *N. glauca* and *N. tabacum*

Agrobacterium tumefaciens strain LBA4404 was used for the above binary plasmid transformation. This strain was grown on YEB medium supplemented with 50 μg/mL rifampicin and 50 μg/mL gentamicin final concentrations for 2 days at 28°C. The plasmids pDGB3 Ω1 [p35S:CCD2L:T35S-pNos:Hyg:T35S], and pDGB3 α1 [p35S:CCD2L:T35S-pNos:Hyg:T35S-pTUBI4U:BrCrtZ:T35S-pAtUBQ10:AtOrMut: T35S] were introduced into the bacteria by electroporation. Clones were selected on YEB agar plates containing 100 μg/mL rifampicin, 50 μg/mL spectinomycin or 50 μg/mL kanamycin, and 25 μg/mL gentamicin. *Agrobacterium*-mediated transformation of *N. glauca* and *N. tabacum* was performed according to the leaf disc method of Horsch (Horsch et al., 1985). For the selection of the transformants, the antibiotic hygromycin B at a final concentration of 50 μg/mL was used.

Wild-Type (WT) tobacco (*N. glauca* and *N. tabacum* cv. SR1) and transgenic plants were grown in pots with soil (Traysubstrat, Klasmann-Deilmann GmbH, Postfach 1250, 49741 Geeste, Germany) in a controlled growth chamber with a 25/20 °C day/night temperature regime, a 12-h photoperiod (mean irradiance 100 μmol m^{-2} s^{-1}) and 60–90% relative humidity. The fully expanded mature leaves (the 5th and 6th leaves) were collected from five Wt and five transgenic plants for each line. All materials were frozen in liquid nitrogen and stored at –80 °C. After leaf sample collection, Wt and transgenic *N. glauca* plants were grown in a controlled growth chamber with a 25/20 °C day/night temperature regime, a 12-h photoperiod (mean irradiance 400 μmol m^{-2} s^{-1}) and 60–90% relative humidity. WT and transgenic tobacco plants were self-pollinated to obtain seeds.

1.3.3. Carotenoid and apocarotenoid extraction and quantification

Metabolite extraction and analysis differed depending on the nature of the metabolites. Polar and apolar metabolites were extracted from 50 to 5 mg of lyophilized leaves, respectively. For the analysis of polar metabolites, leaves were extracted in cold 50% methanol. The soluble fractions were analyzed by high-performance liquid chromatography-diode array detector (HPLC-DAD) as previously described (Marti et al. 2020). The insoluble fractions (carotenoids) were extracted with 1:2 cold extraction solvents (50:50 methanol: $CHCl_3$) and analyzed by HPLC-DAD as previously described (Marti et al. 2020). Pigments were quantified by integrating peak areas that were converted to concentrations by comparison with the standards and as reported before (Diretto et al. 2019). All the samples were analyzed in triplicate. Metabolite extraction

and metabolomic analysis were performed by Dr. O. Ahrazem (Departamento de Ciencia y Tecnología Agroforestal y Genética, Instituto Botánico, Universidad de Castilla-La Mancha, Campus Universitario, Albacete, Spain)

1.3.4. Gene expression analysis by quantitative real-time PCR

Total RNA was isolated from leaves using the RNeasy Plant Mini Kit (Qiagen, Germany), after DNase treatment, first-strand cDNA synthesis was performed using the SuperScript First-Strand Synthesis System (Takara, Otsu, Japan) primed with oligo (dT) 18 following the manufacturer's instructions. qRT-PCR was performed with a Fluorescent Quantitative PCR Detector (Apply Biosystems). SYBR Green real-time PCR Master Mix (Promega) was used. The actin gene was used as a reference gene. qRT-PCR products were assessed by melting curve to ensure the specificity of the amplification in the reactions. Three technical replicates were carried out for each biological sample. Conditions for qRT-PCR cycling were 95°C for 3 min, 95°C for 20 s, 60°C for 20 s, 40 cycles. The relative expression level of each gene was calculated using the $2^{-\Delta\Delta Cq}$ method. Primers used in the qRT-PCR analysis are listed in **Table 1.1**.

1.4. Results

1.4.1. Introduction of *CsCCD2L* in the *N. glauca* and *N. tabacum* genomes

We set out to produce crocins in two different species *N. tabacum* and *N. glauca*. Two binary vectors were constructed using the Goldenbraid strategy (Sarrion-Perdigones et al. 2013; Sarrion-Perdigones et al. 2014). The first construct (**Figure 1.2A**) contains only *CsCCD2L* under the control of the *cauliflower mosaic virus* (CaMV) 35S promoter together with the hygromycin phosphotransferase selectable marker (*hpt*) gene. The second construct (**Figure 1.2B**) carried three genes: *CsCCD2L*, *BrCrtZ* and $AtOr^{Mut}$ under the control of CaMV35S promoter, the *tobacco* polyubiquitin Ubi.U4 and the *Arabidopsis* AtUBQ10 promoter, respectively, together with the hygromycin gene as a selection marker.

To produce crocins in transgenic plants, two species (*N. tabacum* and *N. glauca*) were transformed with *CsCCD2L* under the control of the CaMV35S promoter (**Figure 1.2A**). Seventy-four (74) putative transgenic *N. tabacum* plants were obtained. Four out of these seventy-four plants showed a bleaching phenotype (**Figure 1.3A**). All plants were fertile and were self-pollinated to obtain T1 seeds. *N. tabacum* T1 seeds were plated on hygromycin B (at a final concentration of 50 μg/ml) selection MS medium. All lines

showed resistance/susceptible segregation. T1 and T2 transgenic plants were grown in the greenhouse as described in section **1.3 Materials and methods** (mean irradiance 100 μmol m^{-2} s^{-1}), exhibited a normal phenotype (**Figure 1.3B**) and were fertile. Based on segregation patterns of hygromycin-resistant and sensitive seedlings, twenty-nine out of the fifty-five lines were shown to be single locus insertions for the introduced transgene.

Ten putative transgenic *N. glauca* plants (T0) were obtained and screened by HPLC-DAD analysis. All lines produced crocins in leaves. All plants were fertile and were self-pollinated to obtain T1 seeds, which took approximately one and half years. The transgenic T1 seeds were plated on hygromycin B similar to the *N tabacum* plants above. All lines showed resistance/susceptible phenotypes under hygromycin selection. T1 transgenic plants were grown in a greenhouse as described in section **1.3 Materials and methods** (mean irradiance 100 μmol m^{-2} s^{-1}) and exhibited a faint yellow leaf phenotype (**Figure 1.3C**). Based on segregation of hygromycin-resistant and sensitive seedlings, seven out of the ten lines were single locus transgene insertion lines.

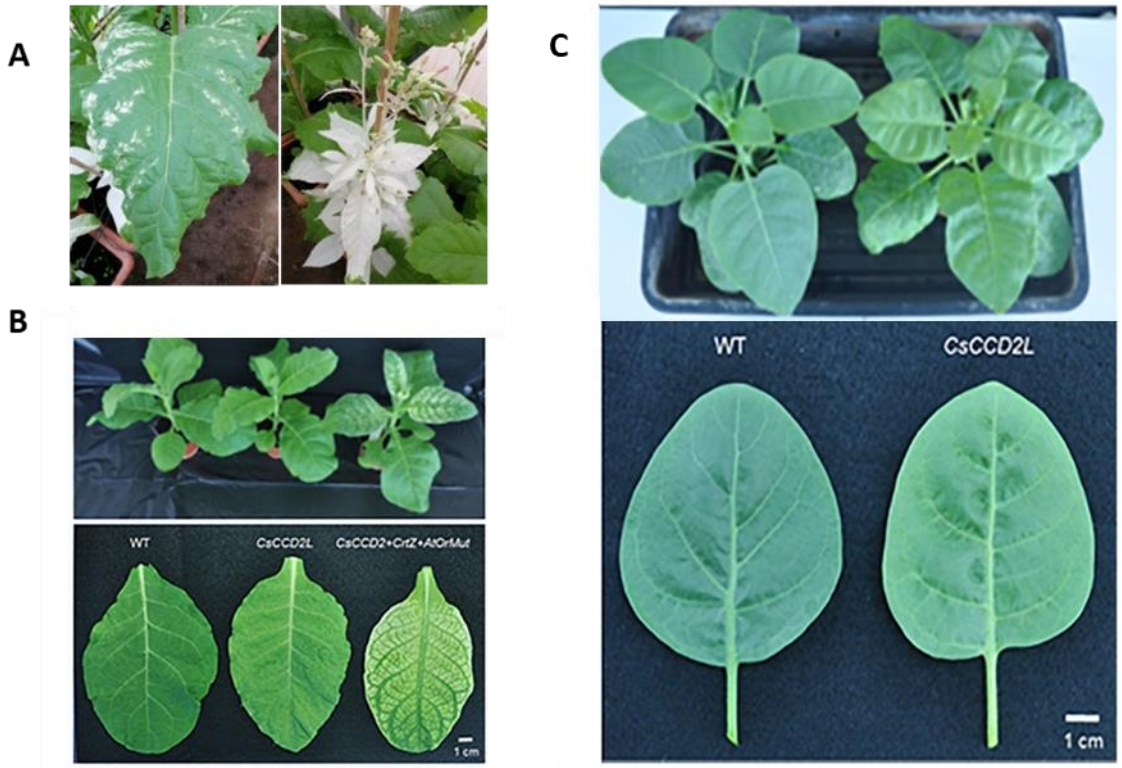


Figure 1.3. Tobacco plant phenotypes. **(A)** Bleaching phenotype of some T0 *N. tabacum* plants overexpressing *CsCCD2L*. **(B)** Plants and sixth leaves from *N. tabacum* WT control and T2 homozygous lines expressing *CsCCD2L* or co-expressing *CsCCD2L* + *BrCrtZ* + *AtOr*^{Mut}. **(C)** Plants and sixth leaf from *N. glauca* WT control and T1 hygromycin resistant plants expressing *CsCCD2L*.

1.4.2. Introduction of *CsCCD2L, BrCrtZ*, and *AtOr^{Mut}* in the *N. tabacum* genome

The flower formation of *N. glauca* depends very much on the season, even in a greenhouse. The plant prefers long-day conditions and high intensity of sunlight. Thus, it may take up to 1 year after seedling transfer to soil before the first flowers set seeds (Gerjets et al. 2007). Therefore, we have chosen the faster-growing *N. tabacum* as a model to improve crocin biosynthesis in leaves by transformation with one binary vector carrying *CsCCD2L, BrCrtZ* and *AtOr^{Mut}* under the control of three different promoters (CaMV35S promoter, *N. tabacum* Ubi.U4 promoter, and *Arabidopsis* ubiquitin 10 promoter, respectively) (**Figure 1.2B**). Seventy-five putative transgenic *N. tabacum* T0 plants were generated and screened by crocin analysis in leaves. Thirty-two transgenic plants accumulated crocins in leaves. All the crocin accumulating transgenic *N. tabacum* T0 plants were fertile and were self-pollinated to obtain T1 seeds. T1 seeds were plated on selection MS medium containing hygromycin B. All 32 lines showed resistance/susceptible segregation under hygromycin selection. T1 and T2 transgenic plants were grown in a greenhouse (mean irradiance 100 μmol m^{-2} s^{-1}) and exhibited a yellow leaf phenotype (**Figure 1.3B**). Based on the segregation of hygromycin-resistant and sensitive seedlings, 15 out of the 32 lines were deduced to be single locus transgene insertion lines.

1.4.3. Crocin accumulation in transgenic *N. glauca* and *N. tabacum* plants

When compared with the chromatogram of the extract of the WT *N. glauca* (**Figure 1.4**) and *N. tabacum* (**Figure 1.5**) the chromatograms of the transgenic lines showed several peaks with maximum absorbance from 433 to 440 nm. Further analyses of the retentions times and the spectra data compared with the standards for crocins led to the identification of crocins with different degrees of glycosylation, ranging from one to four glucose molecules (**Figures 1.4B, 1.5B**). Such apocarotenoids were absent in the WT extracts. *N. glauca* lines accumulated a 10-fold higher amount of crocins than *N. tabacum* lines. Line Ng5 accumulated ca: 400 μg/g DW of crocins (**Figures 1.4C**). To evaluate whether the introduction of *AtOr^{Mut}* and *BrCrtZ* in addition to *CsCCD2L* may boost the level of crocins in *N. tabacum*, polar extracts from seven T2 homozygous lines were analyzed to determine the amount of crocins (**Figure 1.6C**). All the lines accumulated higher levels of crocins as compared to those lines from *N. tabacum* expressing *CsCCD2L* alone. Line Nt24 accumulated ca: 136 μg/g DW of crocins (**Figure 1.6C**). Interestingly, *N. glauca* transgenic plants accumulated far higher levels of crocins.

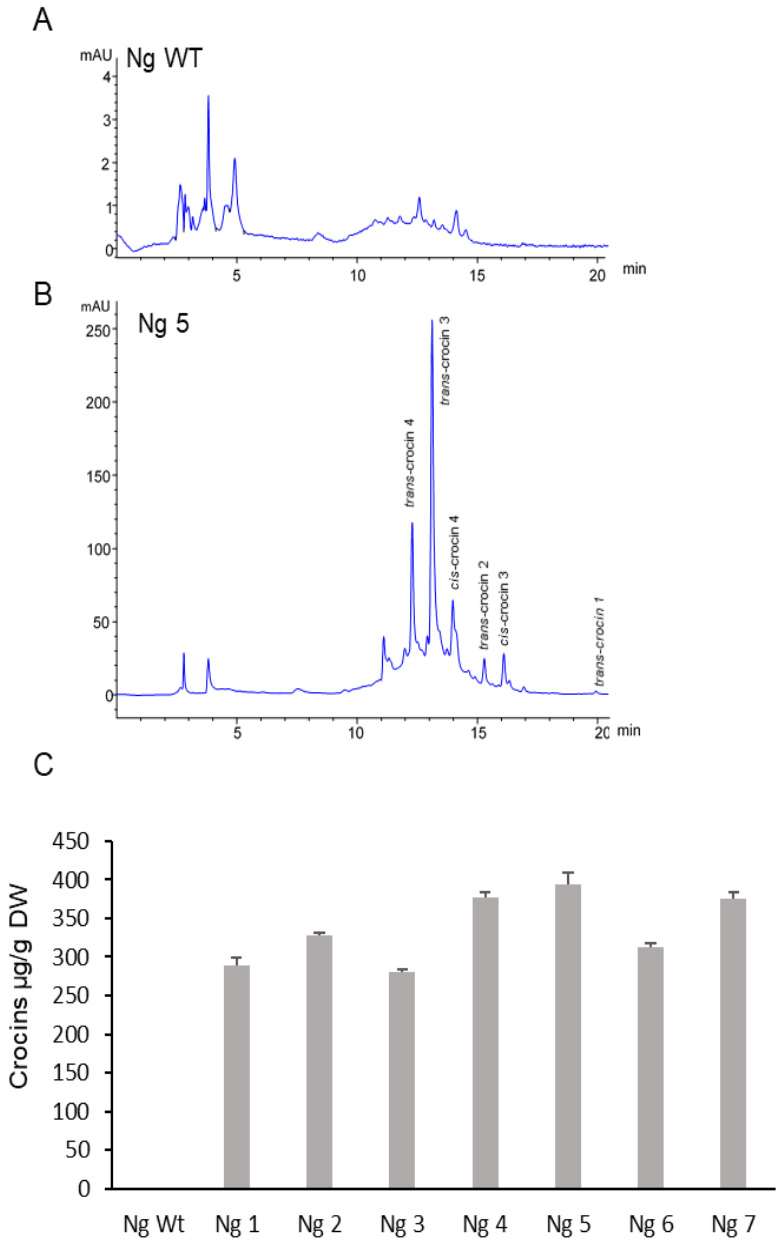


Figure 1.4. Accumulation of apocarotenoids in WT and transgenic T1 lines of *N. glauca* expressing *CsCCD2L*. **(A)** High-performance liquid chromatography-diode array detector (HPLC-DAD) analysis of polar extracts of WT *N. glauca* leaves at 440 nm. **(B)** HPLC-DAD analysis of polar extracts of transgenic *N. glauca* leaves at 440 nm. Note that peaks for the abundant crocins in the transgenic lines (trans-crocin 4, trans-crocin 3, cis-crocin 4, trans-crocin 2, cis-crocin 3, and trans-crocin 1) are completely absent from Wt plants. mAU, milli-absorbance units. **(C)** Apocarotenoid accumulation in leaves of WT and transgenic lines. Analyses were done in triplicate. Error bars represent the SD. DW, dry weight.

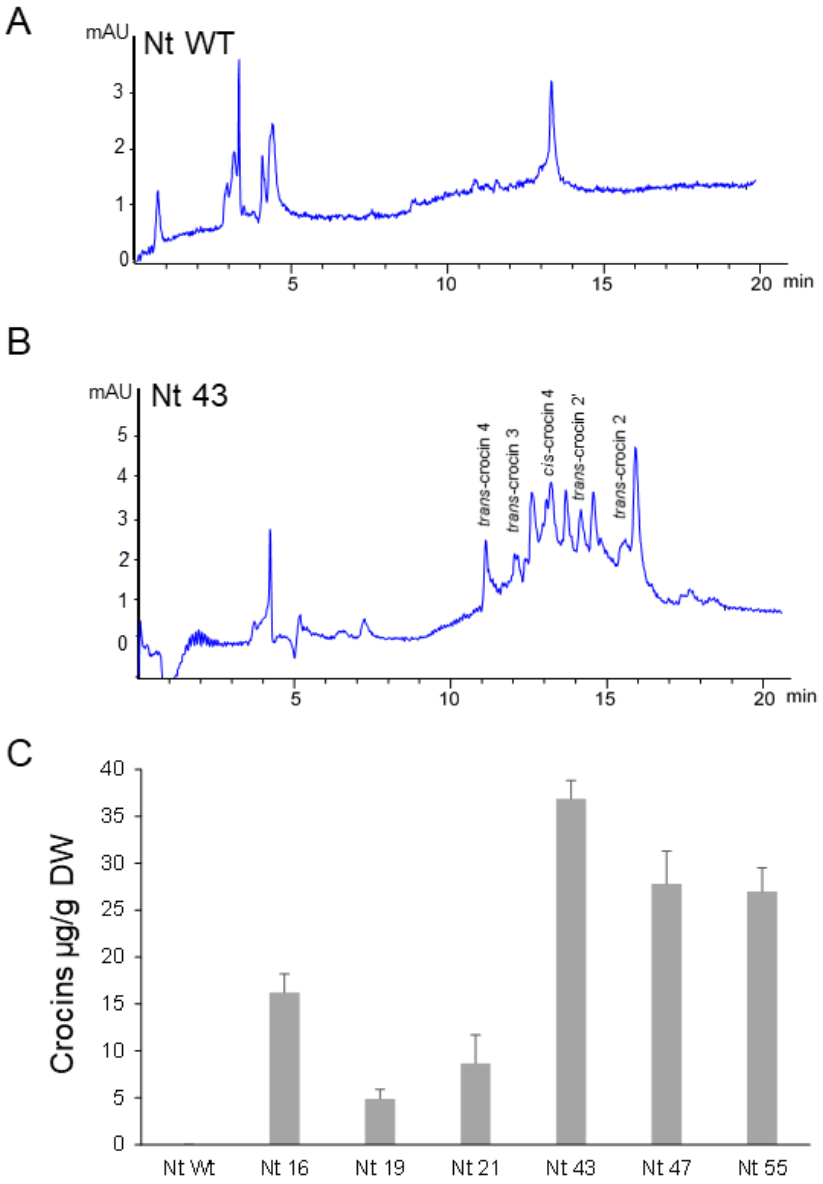


Figure 1.5. Accumulation of apocarotenoids in wild-type (WT) and transgenic T2 homozygous lines of *N. tabacum* expressing *CsCCD2L*. **(A)** HPLC-DAD analysis of polar extracts of WT leaves at 440 nm. **(B)** HPLC-DAD analysis of polar extracts of transgenic leaves at 440 nm. Peaks for the abundant crocins in the transgenic lines (trans-crocin 4, trans-crocin 3, cis-crocin 4, trans-crocin 2', and trans-crocin 2) are completely absent from Wt plants. mAU, milli-absorbance units. (**C)** Apocarotenoid accumulation in leaves of WT and transgenic lines. Analyses were done in triplicate. Error bars represent the SD. DW, dry weight.

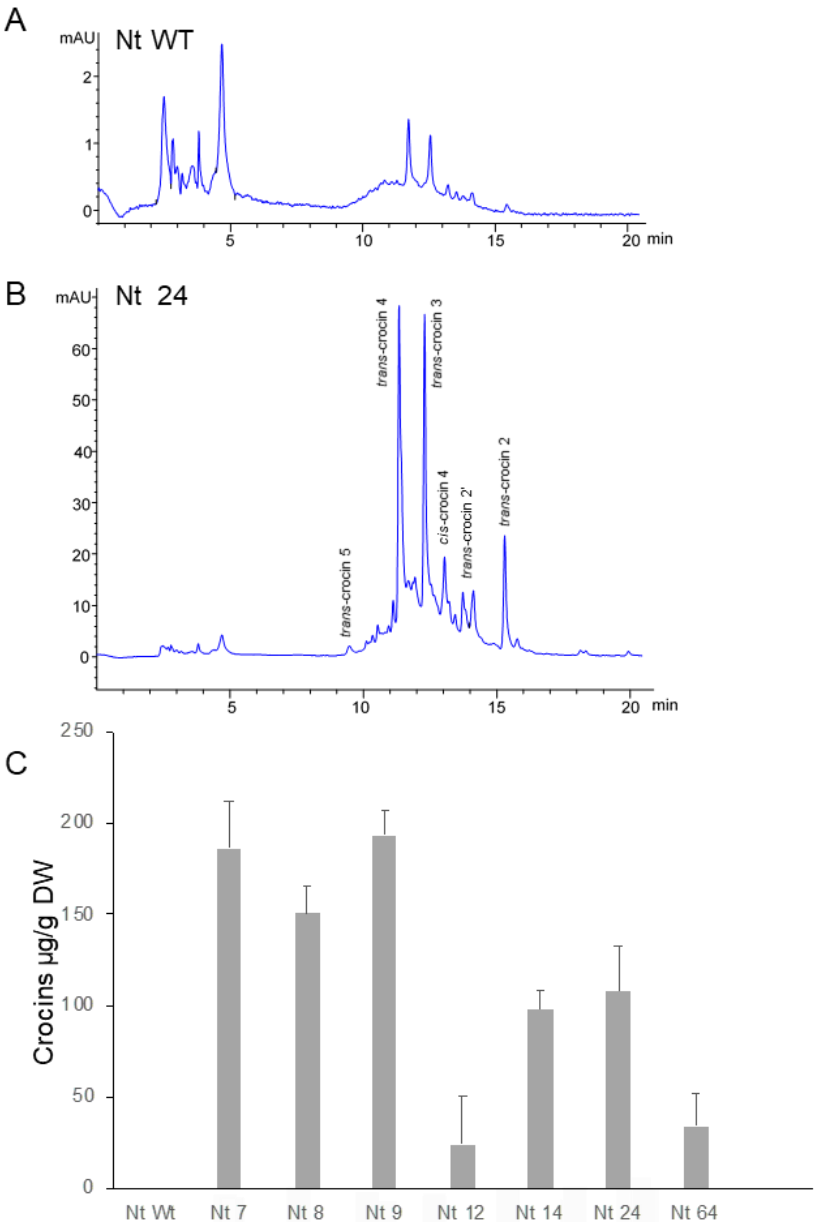


Figure 1.6. Accumulation of apocarotenoids in wild-type (WT) and transgenic T2 homozygous lines of *N. tabacum* co-expressing *CsCCD2L+BrCrtZ+AtOr*^{Mut}. (**A)** HPLC-DAD analysis of polar extracts of Wt leaves at 440 nm. **(B)** HPLC-DAD analysis of polar extracts of transgenic leaves at 440 nm. Peaks for the abundant crocins in the transgenic lines (trans-crocin 4, trans-crocin 3, cis-crocin 4, trans-crocin 2´, and trans-crocin 2) are completely absent from Wt plants. mAU, milli-absorbance units. **(C)** Apocarotenoid accumulation in leaves of WT and transgenic lines. Analyses were done in triplicate. Error bars represent the SD. DW, dry weight.

1.4.4. Carotenoid accumulation in transgenic *N. glauca* and *N. tabacum* plants

Lines accumulating higher levels of crocins were selected for more in-depth analyses. Carotenoids were extracted from four *N. tabacum* lines expressing *CsCCD2L*, *BrCrtZ*, and *AtOr*^{Mut}, three *N. glauca* lines expressing *CsCCD2L*, and WT controls and analyzed by HPLC-DAD (**Figures 1.7, 1.8 and Tables 1.2, 1.3**). In general, the carotenoid content of all transgenic lines was much lower compared to their corresponding WT controls. The most prominent differences in the carotenoid composition in leaves were for neoxanthin, β-carotene, and lutein. Neoxanthin, β-carotene, and lutein levels were substantially decreased in the leaves of transgenic plants compared with WT (**Figures 1.7C, 1.8C**). Changes in the levels of chlorophyll were more pronounced in transgenic plants of *N. glauca* than *N. tabacum* (**Figures 1.7B, 1.8B**).

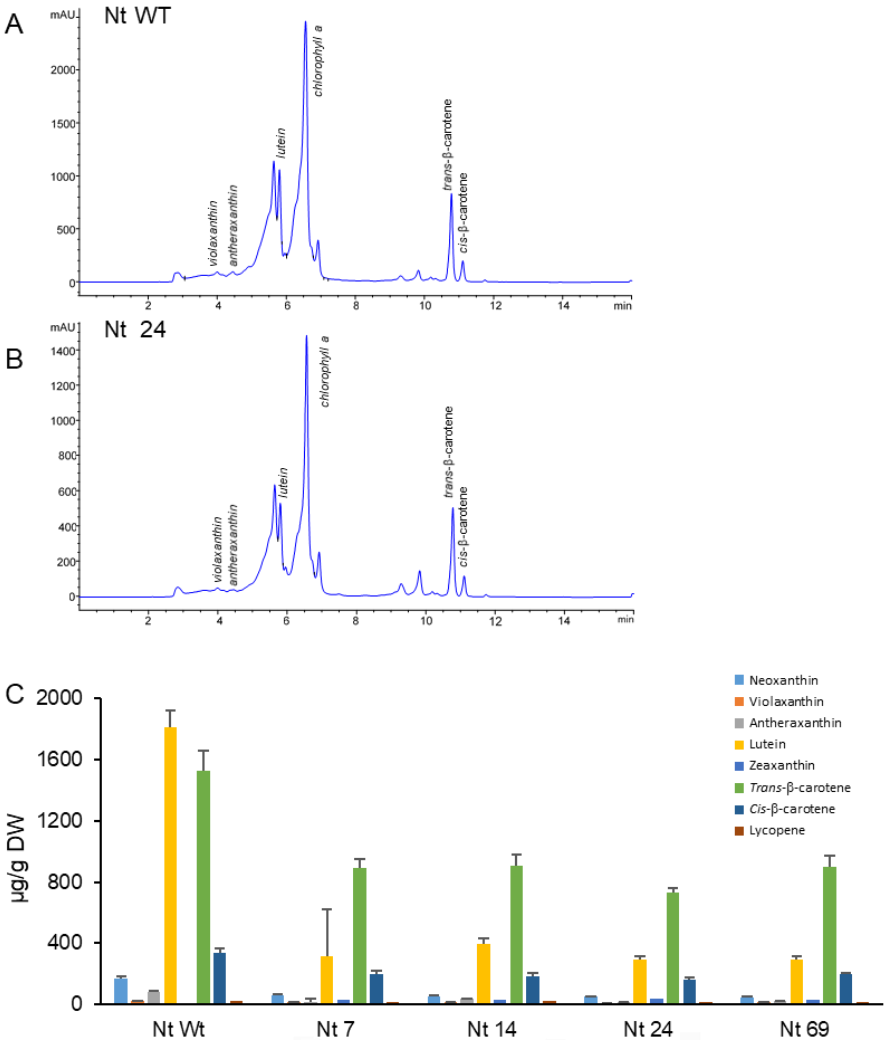


Figure 1.7. Accumulation of carotenoids in WT and transgenic T2 lines of *N. tabacum* expressing *CsCCD2L+BrCrtZ+AtOr*^{Mut}. **(A)** HPLC-DAD analysis of apolar extracts of WT leaves at 450 nm. **(B)** HPLC-DAD analysis of apolar extracts of transgenic leaves at 450 nm. mAU, milli-absorbance units. **(C)** Carotenoid levels in leaves of WT and transgenic lines. Analyses were done in triplicate. Error bars represent the SD. DW, dry weight.

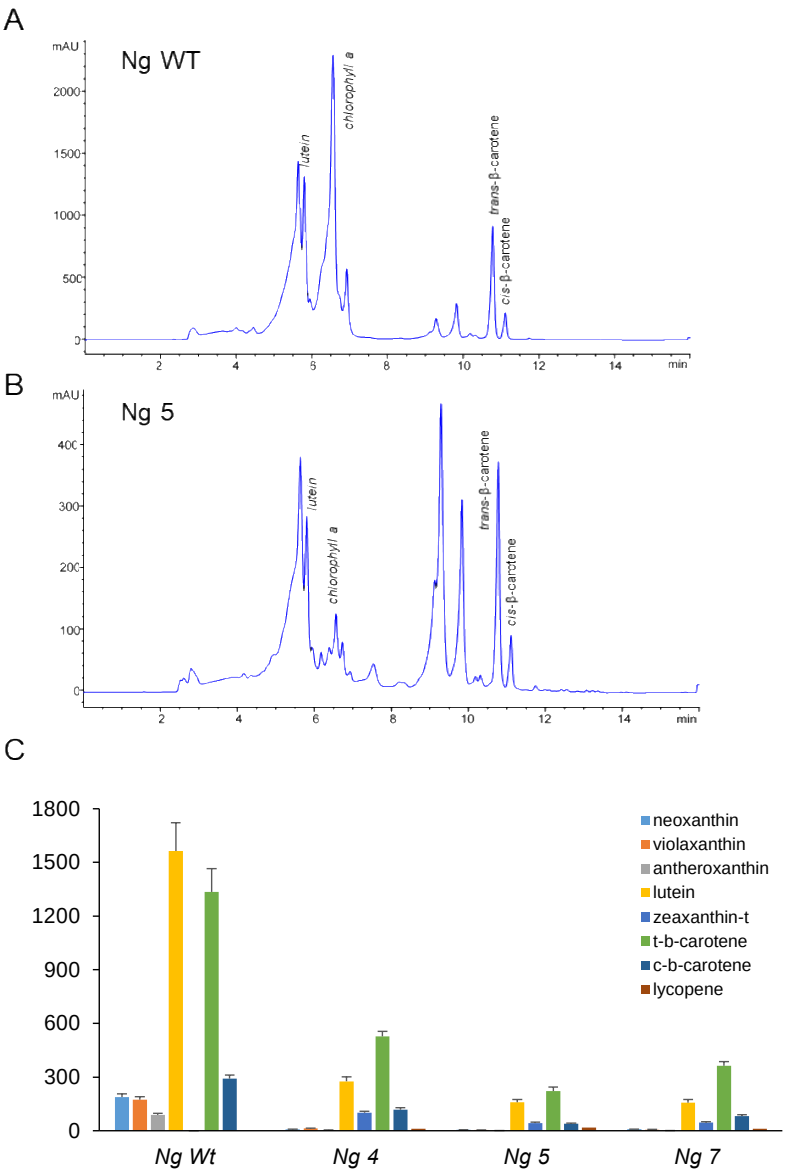


Figure 1.8. Accumulation of carotenoids in WT and transgenic T1 lines of *N. glauca* expressing *CsCCD2L*. **(A)** HPLC-DAD analysis of apolar extracts of WT leaves at 450 nm. **(B)** HPLC-DAD analysis of apolar extracts of transgenic leaves at 450 nm. mAU, milli-absorbance units. **(C)** Carotenoid levels in leaves of Wt and transgenic lines. Analyses were done in triplicate. Error bars represent the SD. DW, dry weight.

Table 1.2. Carotenoid content (µg/g DW) of leaf tissues of *N. tabacum* WT and transformants co-expressing *CsCCD2L*, *BrCrtZ* and *AtOr*^{Mut}.

Carotenoids	**Nt WT**	**Nt 7**	**Nt 14**	**Nt 24**	**Nt 69**
Neoxanthin	172.4±12.5	63.60±4.10	56.30±4.80	52.68±3.99	46.10±3.90
Violaxanthin	22.9±2.90	15.02±1.06	14.23±1.23	9.77±0.90	17.50±1.70
Antheroxanthin	84.34±9.21	18.80±2.65	36.90±2.90	13.60±1.41	23.10±2.10
Lutein	1811.77±112.23	312.16±22.66	397.60±38.20	293.21±25.25	294.11±25.22
Zeaxanthin	0	30.25±3.48	29.35±3.40	35.21±6.02	32.97±3.02
Trans-β-carotene	1528.43±130.01	891.50±61.40	907.30±71.60	731.42±33.02	895.70±75.60
Cis-β-carotene	336.05±34.22	197.58±19.58	186.90±17.10	160.74±15.87	195.61±13.45
Lycopene	26,3±1.28	16.80±1.70	19.90±1.50	18.80±1.78	19.70±1.80
Total carotenoids	3955.89	1545.71	1648.48	1315.43	1524.79

Table 1.3. Carotenoid content (µg/g DW) of leaf tissues of *N. glauca* WT and transformants expressing *CsCCD2L*.

Carotenoids	**Ng WT**	**Ng 4**	**Ng 5**	**Ng 7**
Neoxanthin	187.80±17.93	7.68±0.71	5.45±0.56	8.17±0.70
Violaxanthin	173.24±18.31	13.11±1.11	4.31±0.41	6.20±0.72
Antheroxanthin	88.23±8.28	3.42±0.72	1.35±0.40	1.53±0.58
Lutein	1564.23±160.02	275.33±25.33	159.15±15.05	157.56±17.35
Zeaxanthin	0	101.23±7.66	42.85±5.82	45.08±4.37
Trans-β-carotene	1335.11±129.18	527.84±31.04	221.65±20.69	363.32±23.32
Cis-β-carotene	291.33±20.27	118.01±12.01	39.55±2.55	81.97±6.97
Lycopene	17.9±1.34	11.62±1.74	18.25±1.31	12.13±1.15
Total carotenoids	3639.94	1058.24	492.56	675.96

1.4.5. Transgene expression and its impact on the expression of selected endogenous carotenogenic genes

The expression levels of transgenes and selected endogenous carotenoid biosynthetic genes (*PSY1*, *PSY2*, *LCYB*, and *BCH*) in leaves were monitored by RT-qPCR (**Figures 1.8A, B**). The expression of *CsCCD2L* in transgenic *N. tabacum* plants was much higher than those of *BrCrtZ* and *AtOr*^{Mut}, implying that the activity of CaMV35S promoter for *CsCCD2L* is much stronger than those of tobacco polyubiquitin Ubi.4U and the Arabidopsis AtUBQ10 promoters for *BrCrtZ* and *AtOr*^{Mut}, respectively. Endogenous carotenoid biosynthetic genes, including phytoene synthase (*PSY1* and *PSY2*) genes,

lycopene β-cyclase (*LCYB*) and β-carotene hydroxylase (*BCH*) genes, were downregulated in all the transgenic plants as compared with the Wt control.

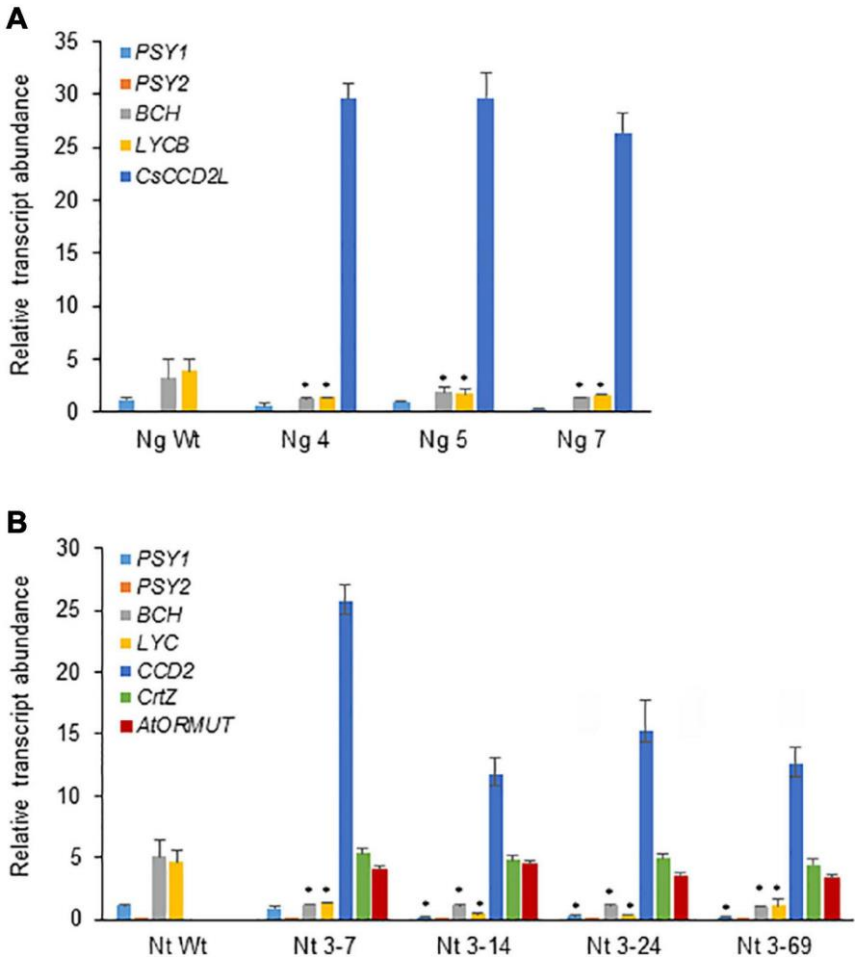


Figure 1.9. Relative abundance of different transcripts in *N. glauca* **(A)** and *N. tabacum* **(B)** wild-type (Wt) and selected transgenic lines (Ng lines expressing *CsCCD2L*, Nt lines co-expressing *CsCCD2L*, *BrCrtZ*, and *AtOr*^{Mut}). RT-qPCR was used to quantify gene expression levels in three technical replicates per sample Error bars represent the SD. The asterisk above the bars indicates P-values (Student's t-test) statistically significant ($P < 0.05$).

1.5. Discussion

Crocins are mainly known for being responsible for the color of the saffron spice. In addition to their pigment capacity, crocins show bioactive properties with several therapeutic and pharmacological applications (Mykhailenko et al. 2019). The saffron apocarotenoids had a global market size valued at 374.6 million USD in 2020 and is expected to reach USD 721.5 million by 2028, being the medical application segment one projected to be the fastest-growing segment. Between 120,000 and 200,000 flowers are needed to produce One kg of dried saffron stigma threads, which equates to 370-470 h of work. Consequently, the process is very labor-intensive and risky since it is highly dependent on environmental conditions, leading to high costs.

To overcome these limitations for obtaining saffron apocarotenoids while considering the advances of metabolic engineering, its progression into synthetic biology, and the elucidation of the saffron pathway, we were able to engineer plants from both *N. glauca* and *N. tabacum* expressing *CsCCD2L*. Several efforts have been made by the scientific community to transfer the synthesis of apocarotenoids from saffron to other hosts, looking for the low-cost production of these rare metabolites. It is known that the use of the saffron *Cs*CCD2L enzyme resulted in a maximum accumulation of 1.22, 15.70, and 4.42 mg/l of crocetin in *Saccharomyces cerevisiae* (Chai et al. 2017) and *Escherichia coli* (Wang et al. 2019). In a subsequent study aimed to confirm, in *planta*, the role of a novel UGTs in picrocrocin biosynthesis (Diretto et al. 2019), *N. benthamiana* leaves were transiently transformed, via *A. tumefaciens*, with *CsCCD2L*, alone or in combination with *UGT709G1*, which led to the production of 30.5 µg/g DW of crocins, with a glycosylation degree ranging from 1 to 4 (crocin 1–4). In addition, Marti et al. using a virus-driven expression of *CsCCD2L* in adult *N. benthamiana* plants demonstrated that *CsCCD2L* expression alone is sufficient for significant crocin accumulation in transiently transformed tobacco (up to 2.18 mg/g DW) (Marti et al. 2020),.

The expression of *CsCCD2L* in *N. glauca* and *N. tabacum* allowed the accumulation of notable amounts of up to 400 and 36 µg/g DW of crocins, respectively. The underlying reason for the accumulated crocin variation observed among the analyzed *N. glauca* and *N. tabacum* lines remains unknown. The amount of crocins obtained is much higher in lines from *N. glauca* than in those from *N. tabacum* and the ones previously described using an *A. tumefaciens*-mediated transient expression of the *CsCCD2L* in *N. benthamiana* leaves, alone or in combination with *UGT709G1* (30.5 µg/g DW of crocins), this accumulation is lower than the one reported in leaves of adult *N. benthamiana* using a recombinant virus that expressed *CsCCD2L* (2.18 mg/g DW of crocins) (Marti et al. 2020). However, crocin production in *N. glauca* provides more effective compared to the data obtained for the production of other carotenoids in this plant; expression of the β-carotene ketolase (CrtO) gene led to a total ketocarotenoid concentration in leaves of 136.6 (young) or 156.1 (older) µg/g dry weight and in petals of 165 µg/g dry weight (Zhu et al. 2007; Mortimer et al. 2017).

The profiles of crocins were different from those reported in saffron stigma, where *trans*-crocin 4, followed by *trans*-crocin 3 were the major crocins detected. In contrast, in our transgenic plants, *trans*-crocin 3 was more abundant than *trans*-crocin 4. In *N. benthamiana* leaves expressing transiently only *CsCCD2L*, the major crocin species were

the *trans*-crocin 3, and *trans*-crocin with two glucose molecules (**Figure 1.10**). These data indicated that different endogenous UGTs belonging to *Nicotiana* species recognize crocetin and crocins and are able to transfer different glucose molecules by changing the qualitative pattern. However, other factors, such as the presence of hydrolase activities that affect the stability of the different crocins synthesized in tobacco cells, and the transport of these crocins to the vacuole and their stability therein, could be influencing the different profiles observed.

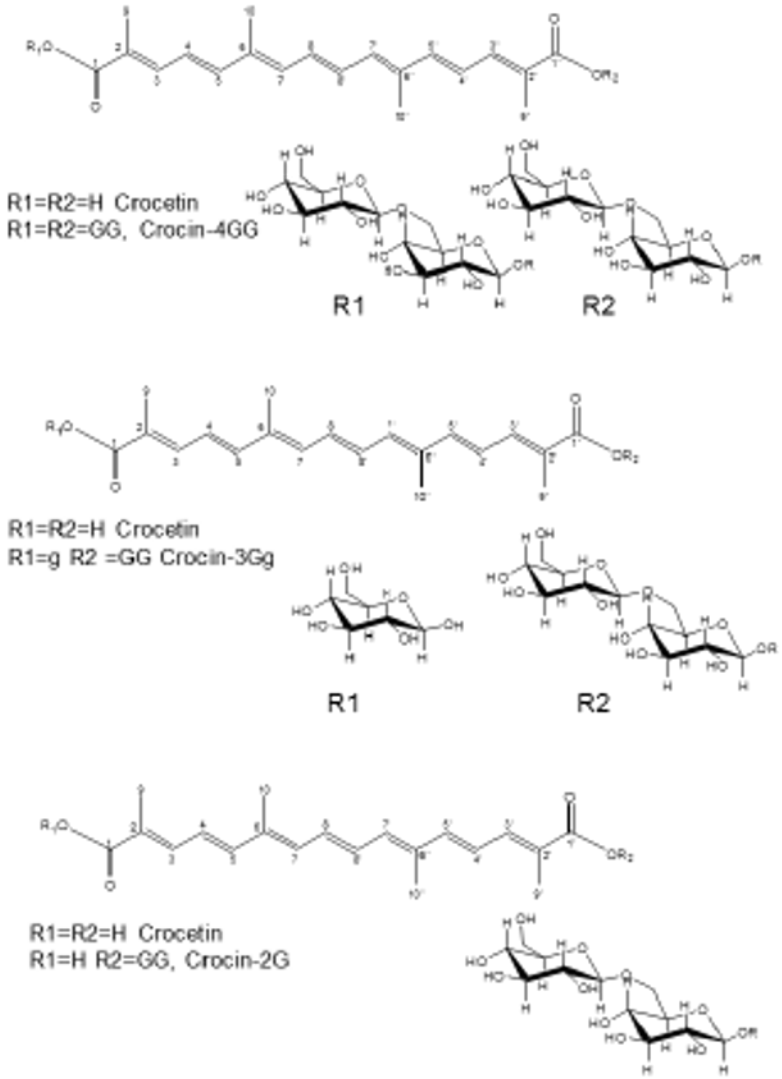


Figure 1.10. Structure of glucosylated crocins.

It is well-known that *BCH* overexpression resulted in increased zeaxanthin and xanthophyll contents, both in microbes and plants (Lagarde et al. 2000; Du et al. 2010; Arango et al. 2014). Therefore, overexpression of *BCH* could enhance the levels of zeaxanthin as the preferred substrate of *Cs*CCD2L (Frusciante et al. 2014). On the other hand, the *Orange* gene has proven to play an important role in carotenoid accumulation by activating chromoplast differentiation in non-green tissues (Lu et al. 2006), although its expression in green tissues exerts no effect on carotenoid levels, we decided to determine whether its introduction could increase the carotenoid content by promoting the formation of carotenoid-sequestering structures. In this context, *BrCrtZ* and *AtOr*^{Mut,} were introduced to boost the crocin content in *N. tabacum*, since the accumulation of these metabolites was 10-fold lower than *N. glauca*. The introduction of the two transgenes together with *CsCCD2L* has most likely led to an increase in the carotenoid pool used as a substrate by the enzyme, consequently, in a 3.5-fold increase in crocins compared to the sole expression of *CsCCD2L*.

By introducing *CsCCD2L* alone or in combination with *BrCrtZ* and *AtOr*^{Mut}, a decreased level of the major and other minor carotenoids normally present in *N. glauca* and *N. tabacum* was detected. Among all the carotenoids detected, lutein was one of the carotenoids showing the higher decrease, likely due to the ability of *CsCCD2L* to act over the β-ring of this molecule as reported in previous *in vitro* study where lutein can acts as substrate of this cleavage activity (Frusciante et al. 2014). Zeaxanthin, the crocin precursor, is not a major carotenoid in *Nicotiana* plant leaves, although the production of the crocins in *N. glauca* and *N. tabacum* suggests that the expression of *CsCCD2L* can drag the leaf metabolic flux toward the production of these apocarotenoids. The level of β-carotene was also lower among the *N. glauca* transgenic plants, the same phenomenon has been shown in *N. tabacum* plants transformed with the three genes suggesting a hydroxylation of β-carotene to zeaxanthin under the influence of *CsCCD2L* in *N. glauca* and the combination of the exogenous *BrCrtZ* and *CsCCD2L* in *N. tabacum*. Gotz et al. previously described the increase of zeaxanthin in tobacco plants (*Nicotiana tabacum* L. cv. Samsun) by transformation with a heterologous carotenoid gene encoding β-carotene hydroxylase (CrtZ) from *Erwinia uredovora* under constitutive promoter control, the increase of zeaxanthin in transgenic plants content is due to catalytic activity of the additional β-carotene hydroxylase rather than being primarily caused by an elevation in the conversion of violaxanthin to zeaxanthin in high light (Gotz et al. 2002).

It has been reported that the expression of an appropriate *CsCCD2L* in *N. benthamiana* is sufficient to activate the apocarotenoid pathway in this plant (Diretto et al., 2019). It seems that both *N. glauca* and *N. tabacum* followed the same pattern as *N. benthamiana*. The crocins produced in this report, using metabolically engineered *N. glauca* (400 μg/g DW) and *N. tabacum* (136 μg/g DW), are lower than those obtained in *N. benthamiana* (2 mg/g DW) using a virus-driven system expressing transient *CsCCD2L*. However, the stable expression of *CsCCD2L* can compensate for the low apocarotenoid productivities by reducing costs and labor. Further increase of *N. glauca* and *N. tabacum* content in these apocarotenoids can also be achieved by combining different types of promoters and the expression of other carotenogenic genes to increase the content of the precursors of crocins, zeaxanthin, and lutein.

1.6. Conclusions

In this chapter, I described the transfer of *CsCCD2L* in *N. glauca* and *N. tabacum* plants or *CsCCD2L* in combination with *BrCrtZ* and *AtOr*^{Mut} *in N. tabacum* plants leading to crocin production in both cases. Focused metabolic engineering intended toward the production of crocins will establish *N. glauca* as a potential industrial host to produce these apocarotenoids. *N. glauca* is not only a better system for crocin biosynthesis and accumulation than *N. tabacum* but also a nicotine-free species and a perennial plant from which the leafy biomass can be harvested repeatedly for crocin production.

1.7. References

Ahrazem O, Diretto G, Argandona J, Rubio-Moraga A, Julve JM, Orzaez D, Granell A, Gomez-Gomez L. (2017). Evolutionarily distinct carotenoid cleavage dioxygenases are responsible for crocetin production in Buddleja davidii. *J. Exp. Bot.* **68**: 4663-4677.

Ahrazem O, Gomez-Gomez L, Rodrigo MJ, Avalos J, Limon MC. (2016a). Carotenoid cleavage oxygenases from microbes and photosynthetic organisms: features and functions. *Int. J. Mol. Sci.* **17**:1781.

Ahrazem O, Rubio-Moraga A, Argandona-Picazo J, Castillo R, Gomez-Gomez L. (2016b). Intron retention and rhythmic diel pattern regulation of carotenoid cleavage dioxygenase 2 during crocetin biosynthesis in saffron. *Plant molecular biology* **91**: 355-374.

Ahrazem O, Rubio-Moraga A, Nebauer SG, Molina RV, Gomez-Gomez L. (2015). Saffron: its phytochemistry, developmental processes, and biotechnological prospects. *J. Agric. Food Chem.* **63**: 8751-8764.

Amin B, Hosseinzadeh H. (2012). Evaluation of aqueous and ethanolic extracts of saffron, Crocus sativus L., and its constituents, safranal and crocin in allodynia and hyperalgesia induced by chronic constriction injury model of neuropathic pain in rats. *Fitoterapia* **83**: 888-895.

Arango J, Jourdan M, Geoffriau E, Beyer P, Welsch R. (2014). Carotene hydroxylase activity determines the levels of both alpha-carotene and total carotenoids in orange carrots. *The Plant cell* **26**: 2223-2233.

Bukhari SI, Manzoor M, Dhar MK. (2018). A comprehensive review of the pharmacological potential of Crocus sativus and its bioactive apocarotenoids. *Biomedicine & pharmacotherapy* **98**: 733-745.

Chai F, Wang Y, Mei X, Yao M, Chen Y, Liu H, Xiao W, Yuan Y. (2017). Heterologous biosynthesis and manipulation of crocetin in Saccharomyces cerevisiae. *Microb. Cell Factories* **16**: 54.

Diretto G, Ahrazem O, Rubio-Moraga A, Fiore A, Sevi F, Argandona J, Gomez-Gomez L. (2019). UGT709G1: a novel uridine diphosphate glycosyltransferase involved in the biosynthesis of picrocrocin, the precursor of safranal in saffron (Crocus sativus). *New Phytol.* **224**: 725-740.

Diretto G, López-Jimenez AJ, Ahrazem O, Frusciante S, Song J, Rubio-Moraga A, Gomez-Gomez L. (2021). Identification and characterization of apocarotenoid modifiers and carotenogenic enzymes for biosynthesis of crocins in Buddleja davidii flowers. *J. Exp. Bot.* **72**: 3200-3218.

Du H, Wang N, Cui F, Li X, Xiao J, Xiong L. (2010). Characterization of the beta-carotene hydroxylase gene DSM2 conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice. *Plant Physiol*.**154**: 1304-1318.

Farre G, Blancquaert D, Capell T, Van Der Straeten D, Christou P, Zhu C. (2014). Engineering complex metabolic pathways in plants. *Annu. Rev. Plant Biol.* **65**: 187-223.

Fiedor J, Burda K. (2014). Potential role of carotenoids as antioxidants in human health and disease. *Nutrients* **6**: 466-488.

Finley JW, Gao S. (2017). A perspective on Crocus sativus L. (saffron) constituent crocin: a potent water-soluble antioxidant and potential therapy for Alzheimer's Disease. *J. Agric. Food Chem.* **65**: 1005-1020.

Fraser PD, Bramley PM. (2004). The biosynthesis and nutritional uses of carotenoids. *Progress in lipid research* **43**: 228-265.

Frusciante S, Diretto G, Bruno M, Ferrante P, Pietrella M, Prado-Cabrero A, Rubio-Moraga A, Beyer P, Gomez-Gomez L, Al-Babili S, Giuliano G. (2014). Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *The Proc. Natl. Acad. Sci. U.S.A.* **111**: 12246-12251.

Gerjets T, Sandmann M, Zhu C, Sandmann G. (2007). Metabolic engineering of ketocarotenoid biosynthesis in leaves and flowers of tobacco species. *Biotechnology journal* **2**: 1263-1269.

Gotz T, Sandmann G, Romer S. (2002). Expression of a bacterial carotene hydroxylase gene (crtZ) enhances UV tolerance in tobacco. *Plant molecular biology* **50**: 129-142.

Kothari D, Thakur M, Joshi R, Kumar A, Kumar R. (2021). Agro-climatic suitability evaluation for saffron production in areas of western Himalaya. *Front. Plant Sci* **12**: 657819.

Kumar A, Devi M, Kumar R, Kumar S. (2022). Introduction of high-value Crocus sativus (saffron) cultivation in non-traditional regions of India through ecological modelling. *Scientific reports* **12**: 11925.

Lagarde D, Beuf L, Vermaas W. (2000). Increased production of zeaxanthin and other pigments by application of genetic engineering techniques to Synechocystis sp. strain PCC 6803. *Applied and Environmental Microbiology* **66**: 64-72.

Lage M, Cantrell CL. (2009). Quantification of saffron (Crocus sativus L.) metabolites crocins, picrocrocin and safranal for quality determination of the spice grown under different environmental Moroccan conditions. *Scientia Horticulturae* **121**: 366-373.

López-Jimenez AJ, Frusciante S, Niza E, Ahrazem O, Rubio-Moraga A, Diretto G, Gomez-Gomez L. (2021). A new glycosyltransferase enzyme from family 91, UGT91P3, is responsible for the final glucosylation step of crocins in saffron (Crocus sativus L.). *International journal of molecular sciences* **22**: 8815.

Lopresti AL, Drummond PD. (2014). Saffron (Crocus sativus) for depression: a systematic review of clinical studies and examination of underlying antidepressant mechanisms of action. *Human psychopharmacology* **29**: 517-527.

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J, Kochian LV, Kupper H, Earle ED, Cao J, Li L. (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. *The Plant cell* **18**: 3594-3605.

Maoka T. (2011). Carotenoids in marine animals. *Marine drugs* **9**: 278-293.

Marti M, Diretto G, Aragones V, Frusciante S, Ahrazem O, Gomez-Gomez L, Daros JA. (2020). Efficient production of saffron crocins and picrocrocin in Nicotiana benthamiana using a virus-driven system. *Metab. Eng*. **61**: 238-250.

Moraga AR, Nohales PF, Perez JA, Gomez-Gomez L. (2004). Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from Crocus sativus stigmas. *Planta* **219**: 955-966.

Moraga AR, Rambla JL, Ahrazem O, Granell A, Gomez-Gomez L. (2009). Metabolite and target transcript analyses during Crocus sativus stigma development. *Phytochemistry* **70**: 1009-1016.

Moras B, Loffredo L, Rey S. (2018). Quality assessment of saffron (Crocus sativus L.) extracts via UHPLC-DAD-MS analysis and detection of adulteration using gardenia fruit extract (Gardenia jasminoides Ellis). *Food chemistry* **257**: 325-332.

Moreno JC, Cerda A, Simpson K, Lopez-Diaz I, Carrera E, Handford M, Stange C. (2016). Increased Nicotiana tabacum fitness through positive regulation of carotenoid, gibberellin and chlorophyll pathways promoted by Daucus carota lycopene β-cyclase (Dclcyb1) expression. *J. Exp. Bot.* **67**: 2325-2338.

Mortimer CL, Misawa N, Perez-Fons L, Robertson FP, Harada H, Bramley PM, Fraser PD. (2017). The formation and sequestration of nonendogenous ketocarotenoids in transgenic Nicotiana glauca. *Plant Physiol*.**173**: 1617-1635.

Mykhailenko O, Kovalyov V, Goryacha O, Ivanauskas L, Georgiyants V. (2019). Biologically active compounds and pharmacological activities of species of the genus Crocus: A review. *Phytochemistry* **162**: 56-89.

Nagatoshi M, Terasaka K, Owaki M, Sota M, Inukai T, Nagatsu A, Mizukami H. (2012). UGT75L6 and UGT94E5 mediate sequential glucosylation of crocetin to crocin in Gardenia jasminoides. *FEBS letters* **586**: 1055-1061.

Pfister S, Meyer P, Steck A, Pfander H. (1996). Isolation and structure elucidation of carotenoid-glycosyl esters in gardenia fruits (Gardenia jasminoides Ellis) and saffron (Crocus sativus Linne). *J. Agric. Food Chem.* **44**: 2612-2615.

Sarrion-Perdigones A, Palaci J, Granell A, Orzaez D. (2014). Design and construction of multigenic constructs for plant biotechnology using the GoldenBraid cloning strategy. *Methods Mol. Biol.* **1116**: 133-151.

Sarrion-Perdigones A, Vazquez-Vilar M, Palaci J, Castelijns B, Forment J, Ziarsolo P, Blanca J, Granell A, Orzaez D. (2013). GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol*.**162**: 1618-1631.

Serrano-Díaz J, Sánchez AM, Maggi L, Martínez-Tomé M, García-Diz L, Murcia MA, Alonso GL. (2012). Increasing the applications of Crocus sativus flowers as natural antioxidants. *Journal of Food Science* **77**: C1162-C1168.

Shi Y, Wang R, Luo Z, Jin L, Liu P, Chen Q, Li Z, Li F, Wei C, Wu M, Wei P, Xie H, Qu L, Lin F, Yang J. (2014). Molecular cloning and functional characterization of the lycopene epsilon-cyclase gene via virus-induced gene silencing and its expression pattern in Nicotiana tabacum. *International journal of molecular sciences* **15**: 14766-14785.

Walter MH, Floss DS, Strack D. (2010). Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles. *Planta* **232**: 1-17.

Wang W, He P, Zhao D, Ye L, Dai L, Zhang X, Sun Y, Zheng J, Bi C. (2019). Construction of Escherichia coli cell factories for crocin biosynthesis. *Microb. Cell Factories* **18**: 1-11.

Wang Z, Zhang L, Dong C, Guo J, Jin L, Wei P, Li F, Zhang X, Wang R. (2021). Characterization and functional analysis of phytoene synthase gene family in tobacco. *BMC Plant Biology* **21**: 1-18.

Xu Z, Pu X, Gao R, Demurtas OC, Fleck SJ, Richter M, He C, Ji A, Sun W, Kong J, Hu K, Ren F, Song J, Wang Z, Gao T, Xiong C, Yu H, Xin T, Albert VA, Giuliano G, Chen S, Song J. (2020). Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC biology* **18**: 1-14.

Zhu C, Gerjets T, Sandmann G. (2007). Nicotiana glauca engineered for the production of ketocarotenoids in flowers and leaves by expressing the cyanobacterial crtO ketolase gene. *Transgenic Res.* **16**: 813-821

Chapter II

The biosynthesis of non-endogenous apocarotenoids in transgenic *Nicotiana glauca*

2.0. Abstract

Crocins are high-value compounds with industrial and food applications. Enzymes involved in the production of these compounds have been identified in saffron, Buddleja, and gardenia. In this study, the *Buddleja* carotenoid cleavage dioxygenase 4.1 (*Bd*CCD4.1) was constitutively expressed in *Nicotiana glauca*, a *Nicotiana* species with carotenoid-pigmented petals. The transgenic lines produced significant levels of crocins in their leaves and petals. However, the accumulation of crocins was, in general, higher in the leaves, reaching ca: 302 µg/g DW. The production of crocins was associated with decreased levels of endogenous carotenoids, mainly β-carotene. The stability of crocins in leaf and petal tissues was evaluated after three years of storage at room temperature in the dark, showing an average reduction of 58% in the petals, and 78% in the leaves.

2.1. Introduction

The carotenoid backbone can be truncated by the enzymatic or non-enzymatic removal of fragments, generating apocarotenoids (Walter and Strack 2011). Apocarotenoids are terpenoid compounds with an extended conjugated double-bond system. A sequence of hydroxylation and epoxidation reactions leads to the production of the xanthophylls, including crocins, lutein, zeaxanthin, antheraxanthin, and violaxanthin (Bai et al. 2014). Based on the number of carbons and double bonds in the apocarotenoid skeleton, these compounds exhibit distinctive yellow, orange, and red colors, while these properties also determine their degree of volatility (Ahrazem et al. 2016a). In non-photosynthetic tissues, such as petals and fruits, xanthophylls are present as unconjugated molecules or as mono- or diesters of fatty acids (Mariutti and Mercadante 2018), with implications for xanthophyll storage, retention and cleavage. In addition, crocins are highly unsaturated and subject to off-flavors and tissue changes during processing and storage due to oxidative and hydrolytic reactions (Ordoudi et al. 2015).

Crocins are high-value apocarotenoids used in the food and pharmaceutical industries, and to a lesser extend gardenia, which accumulates much lower levels (Chen et al. 2010). New natural sources are required to satisfy this substantial increase in the demand for crocins (Ahrazem et al. 2017). An alternative is the development of more cost-effective production strategies in heterologous systems by the over-expression of rate-limiting enzymes in crocin biosynthesis. *C. sativus* stigma, gardernia fruits and *Buddleja* spp. petals share the common feature of expression of carotenoid cleavage dioxygenase (CDD) activities, such as CCD2L or CCD4 subfamily, which have been elucidated in the crocin biosynthetic pathway (Ahrazem et al. 2017; Diretto et al. 2021; Zheng et al. 2021). Heterologous systems have been exploited with varying degrees of success in bacteria (Wang et al. 2019), yeast (Chai et al. 2017), and in plants by using different species and transgenic approaches to produce saffron and its derivatives (Ahrazem et al. 2016b; Martí et al. 2020; Ahrazem et al. 2022). Among natural sources of these compounds, *Buddleja davidii* accumulates crocins in petals, but at lower levels (ca:10-fold lower) compared to the stigma of saffron (Diretto et al. 2021). Although these levels are much higher than those obtained using other heterologous systems, *B. davidii* is not readily amenable to agricultural production. Even though Buddleja may not be a suitable production host, its native biosynthetic pathway for crocin biosynthesis has been elucidated, and the rate-limiting CCD enzyme, *Bd*CCD4 has been isolated (Ahrazem et al. 2017; Diretto et al. 2021) (**Figure 2.1**). The CCD4 family is the largest family of plant CCDs and is generally

divided into two groups, based on their substrate specificity and the position of the scissile double bond (Zheng et al. 2021). The main clade contains Arabidopsis and potato CCD4 enzymes that show a 9,10 (9′,10′) double bond cleavage activity (Gonzalez-Jorge et al. 2013; Bruno et al. 2016), and the second clade, BdCCD4.1 from *Buddleja davidii* and GjCCD4a from *Gardenia jasminoides* which cleave the 7,8 (7′,8′) site of zeaxanthin to produce crocetin dialdehyde, the precursor of crocin. In addition, *Gj*CCD4a cleaves at the same positions (7,8;7′,8′) in the lycopene and β-carotene backbone, also forms crocetin diadehyde (Zheng et al. 2021). The activity of *Bd*CCD4.1 is similar to that of *Cs*CCD2L in that it utilizes only zeaxanthin as substrate. However, its ability to act on other plant-based carotenoids has not yet been explored (Ahrazem et al. 2017; Martí et al. 2020).

N. glauca has been used previously as a platform to produce ketocarotenoids (Zhu et al. 2007), taking advantage of its highly pigmented petals that mainly accumulate lutein. In addition, *N. glauca* grows in arid environments and has immense potential as a biofuel stock due to its exceptionally high content of hydrocarbons in the leaves (González et al. 2012; Mortimer et al. 2012; Usade et al. 2018).

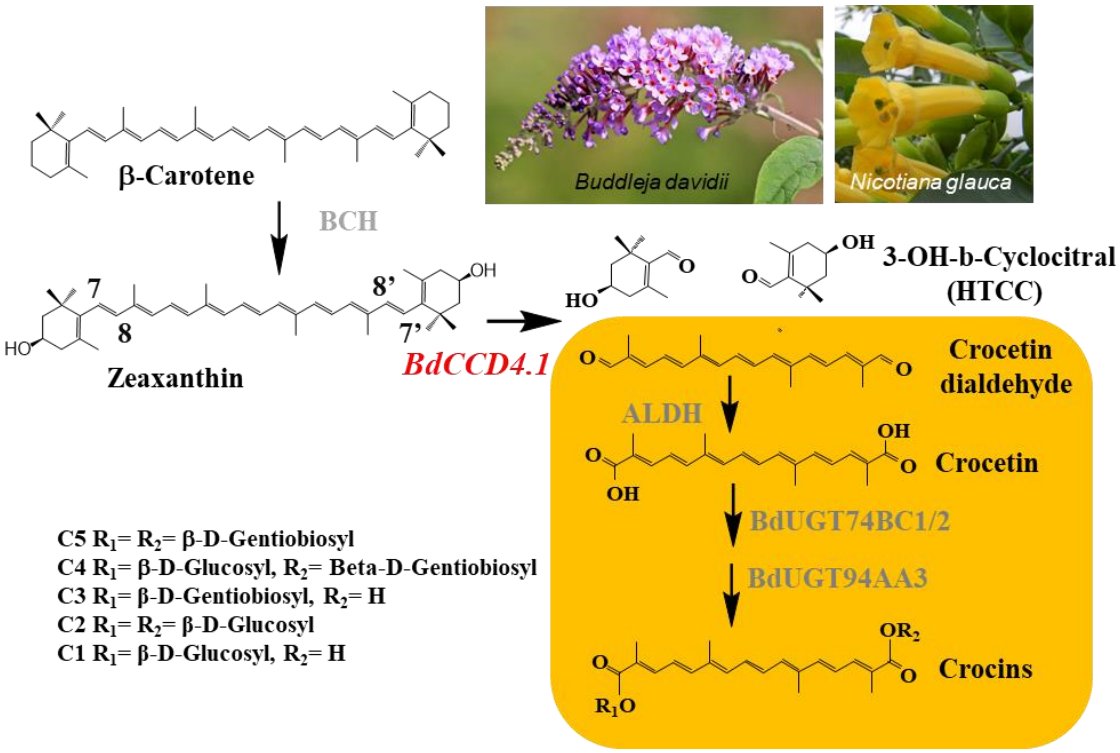


Figure 2.1. Biosynthetic pathway of crocins from zeaxanthin in *Buddleja davidii* petals. Pictures of *B. davidii* and *N. glauca* petals are shown in the upper right. The recombinant enzyme introduced in *N. glauca* under the control of the CaMV35S promoter is shown in red. The yellow background shows the stages of transformation of crocetindial into the different crocins present in the petals of *B. davidii*. The arrows show each of the conversion steps catalyzed by the corresponding enzyme in grey. The different substitutions for the crocins present in the petals of *B. davidii* are shown in the lower left.

2.2. Aims and objectives

The major aim of this chapter was to compare the performance of *BdCCD4.1* and *CsCCD2L* for the production of crocins in *N. glauca* and to determine the stability of these metabolites during storage in leaf and petal tissues.

To reach these aims, the following specific objectives were pursued:

- determine accumulation of crocins and carotenoids in transgenic *N. glauca.*
- quantification of apocarotenoids in leaves and petals of transgenic *N. glauca* after 3-year period of storage under room temperature in the dark.

2.3. Materials and methods

2.3.1. Vector construction

The plasmid for transformation was created with the BdCCD4.1 gene. The CaMV35S promotor was used to control the expression of the transgene. The Goldenbraid strategy was followed to construct the vectors as previously described (**section 1.3.1**). Briefly, the complete open reading frame (ORF) of *Bd*CCD4.1 was domesticated by removing BsmBI and BsaI in the original sequence using the primers. The product was cloned in the vector pUPD2 of the Goldenbraid modular cloning system. The resulting plasmid pUPD2-BdCCD4.1 was used to construct the recombinant binary vector: pDGB3Ω1[p35S:BdCCD4.1:T35S-pNos:Hyg:T35S]. The vector was constructed by Dr. O. Ahrazem (Departamento de Ciencia y Tecnología Agroforestal y Genética, Instituto Botánico, Universidad de Castilla-La Mancha, Campus Universitario, Albacete, Spain).

2.3.2. Genetic transformation of *N. glauca*

Agrobacterium tumefaciens strain LBA4404 was used for transformation by electroporation with the construct pDGB3Ω1[p35S:BdCCD4.1:T35S-pNos:Hyg:T35S]. Transformants were selected on YEB agar plates containing 100 μg/ml rifampicin, 50 μg/ml spectinomycin, and 25 μg/ml gentamicin. Agrobacterium-mediated transformation of *N. glauca* was performed according to the leaf disc method (Horsch et al. 1985). For selection of transformants, hygromycin B at a concentration of 50 μg/ml was used.

Wild type (WT) *N. glauca* and transgenic plants were grown under controlled growth conditions with a 25/20°C day/night temperature cycle, a 12-h photoperiod (mean irradiance 100 μmol m−2 s−1) and 60-90% relative humidity. Mature leaves (the 5th and 6th leaves) were collected from 5 WT and 5 transgenic plants for each line, were frozen in liquid nitrogen and stored at -80 °C until used.

2.3.3. DNA analysis

A PCR reaction with primers specific to *BdCCD4.1* was conducted to verify the presence of the transgene. Genomic DNA was extracted from *N. glauca* leaves using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The amplification conditions were as follows: 95°C for 2 min, followed by 35 cycles at 95°C for 20 s, 55°C for 20 s, and 72°C for 1 min per kilobase, followed by a final extension at 72°C for 5 min. PCR products were separated and visualized on 0.8% agarose gel stained with ethidium bromide using a UV transilluminator.

2.3.4. Carotenoid and apocarotenoid extraction and quantification

HPLC-DAD analysis as described in chapter I **section 1.3.3**. The experiment carried out by Dr. O. Ahrazem (Departamento de Ciencia y Tecnología Agroforestal y Genética, Instituto Botánico, Universidad de Castilla-La Mancha, Campus Universitario, Albacete, Spain).

2.4. Results

2.4.1. Generation of *N. glauca* plants expressing *BdCCD4.1*

Agrobacterium-mediated transformation was used to introduce *BdCCD4.1* into *N. glauca*. The transformation vector contained the *BdCCD4.1* coding sequence (KX374547) under the control of the CaMV35S promoter and the *Hph* gene for selection with the antibiotic hygromycin. Following antibiotic selection, five independent transformed lines were selected and grown in the greenhouse. Two non-transgenic (PCR-negative) plants were used as controls.

2.4.2. Levels of crocins in leaves and petals of transgenic *N. glauca*

In order to quantify apocarotenoids in the plants, a previously described HPLC system (Martí et al. 2020; Ahrazem et al. 2022) was used for the analyses of polar extracts of leaves and petals. Typical chromatograms of the apocarotenoids formed in leaves and petals of transgenic *N. glauca* plants expressing *BdCCD4.1* are shown in **Figure 2.2 A,B**. The profiles of crocins in leaves and petals clearly differ between these tissues, suggesting the presence of different tissue-specific endogenous glycosyltransferases. The predominant crocin moieties contained three and two glucose molecules. The amounts of crocins in petals and leaves in the transgenic lines are shown in **Figure. 2.2C**. In three transgenic lines, the levels of crocins were higher in the leaves than in the petals (lines #5, #7 and #9). In two transgenic lines, crocin levels were similar in the leaves and petals (lines #4 and #8). The levels of crocins in the five lines we analyzed ranged from 137 to

302 µg/g. Previously, we analyzed the levels of crocins in *N. glauca* leaves from plants transformed with *CsCCD2L* **(section 1.4.3, Figure 1.4)**. The levels of crocins in the leaves of these plants were within the same order of magnitude, but higher than those obtained with *BdCCD4.1* (1.5-fold difference on average). We also analyzed the crocin content of the petals of lines we had generated previously expressing *CsCCD2L* **(section 1.3.2)**. Interestingly, the levels of crocins in the petals from transgenic *CsCCD2L* were lower, with a range of 150-200 µg/g **(Figure 2.3)**, than in the leaves, with a range of 280-400 µg/g, suggesting a lower availability of substrates for *Cs*CCD2L in the petals. In fact, zeaxanthin was found at higher levels in leaves than in petals in *N. glauca*, where lutein predominates (Mortimer et al. 2017).

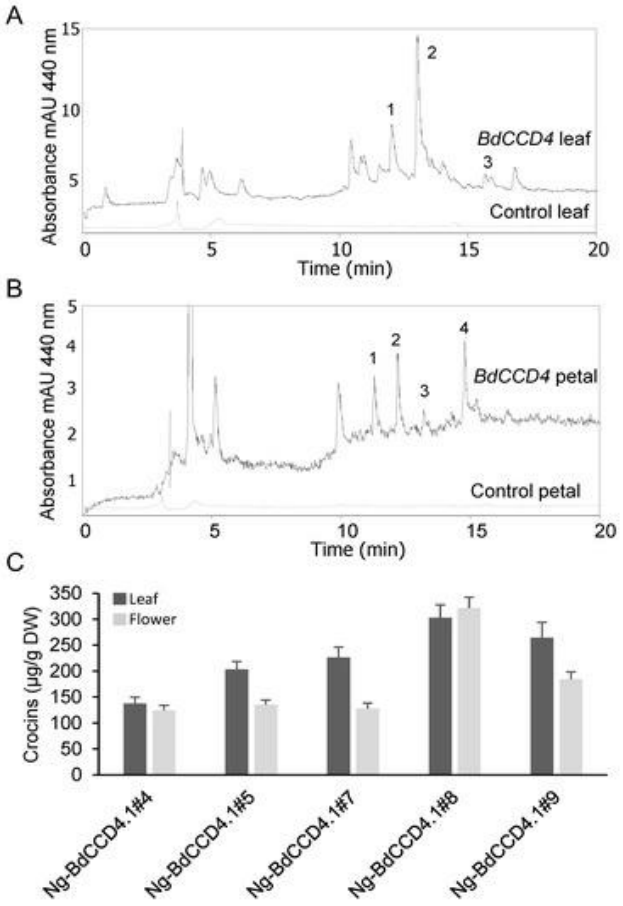


Figure 2.2. Accumulation of crocins in Wt and transgenic T1 lines of *N. glauca* constitutively expressing *BdCCD4.1*. **(A)** HPLC-DAD analysis of polar extracts of Wt and *N. glauca* leaves at 440 nm. **(B)** HPLC-DAD analysis of polar extracts of transgenic and Wt N. glauca petals at 440 nm. Peaks for the abundant crocins in the transgenic lines are denoted by numbers, in leaves (1= trans-crocin3; 2= trans-crocin-2; 3= cis-crocin 2) and petals (1= trans-crocin 4; 2= trans-crocin 3; 3= trans-crocin 2; 4= cis-crocin 2); these crocins were completely absent from Wt plants. mAU, milli-absorbance units. **(C)** Apocarotenoid accumulation in the leaves and petals of transgenic lines. Analyses were conducted in triplicate. Error bars represent the SD. DW, dry weight.

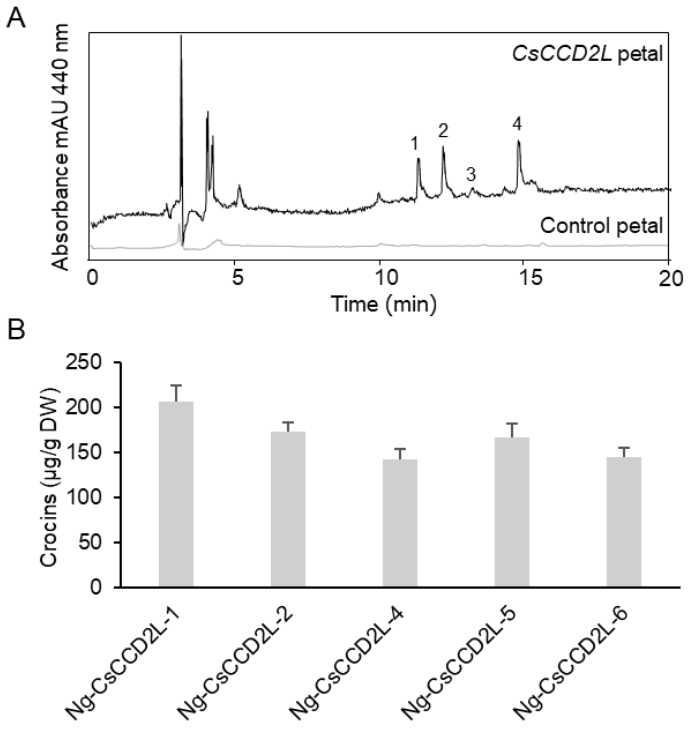


Figure 2.3. Accumulation of apocarotenoids in Wt and transgenic T1 lines of *N. glauca* expressing *CsCCD2L*. **(A)** HPLC-DAD analysis of polar extracts of Wt *N. glauca* petals at 440 nm. The peaks for the abundant crocins in the transgenic lines are as follows: 1= trans-crocin 4; 2= trans-crocin 3, 3= trans-crocin 2, 4= cis-crocin 2, and crocins were completely absent from Wt plants. mAU, milli-absorbance units. **(B)** Crocin accumulation in the petals of transgenic lines. Analyses were conducted in triplicate. Error bars represent the SD. DW, dry weight.

2.4.3. Levels of endogenous carotenoids in leaves and petals of transgenic *N. glauca* expressing *BdCCD4.1*

The levels of endogenous carotenoids were analyzed in the leaves and petals of transgenic *BdCCD4.1* plants **(Figure 2.4)**. Five major carotenoids were detected in the extracts of the control petals and leaves: violaxanthin, lutein, zeaxanthin, all-*trans*-β-carotene, and its geometric isomer 9-*cis*-β-carotene. In all of the transgenic lines, there was a reduction in the content of these carotenoids in the leaves as well as in the petals **(Figures 2.4A, B)**. This suggests that the production of crocins occurred at the expense of the native carotenoids present in these tissues. In the leaves, zeaxanthin was not detected in the transgenic plants, and the lutein concentration was reduced by 74-92%. For β-carotene, the fluctuations were greater between the different lines, with 23-70% reduction in its concentration with respect to the WT plants. Similar fluctuations have been previously observed in *N. glauca* plants engineered to produce ketocarotenoids (Mortimer et al. 2017). Interestingly, the lower reduction in the β-carotene levels was observed in those lines accumulating higher levels of crocins (lines #7 and #9, with a 23.6% and 27.4% reduction, respectively). In the petals, a reduction in concentration was also observed for lutein and β-carotene, whose levels fluctuated within broad ranges of 44-92% and 45-71%

for lutein and β-carotene, respectively. As observed in the leaves, a lower reduction was observed for lines #7 and #9, which showed higher levels of crocins in the petals **(Figure 2.2)**.

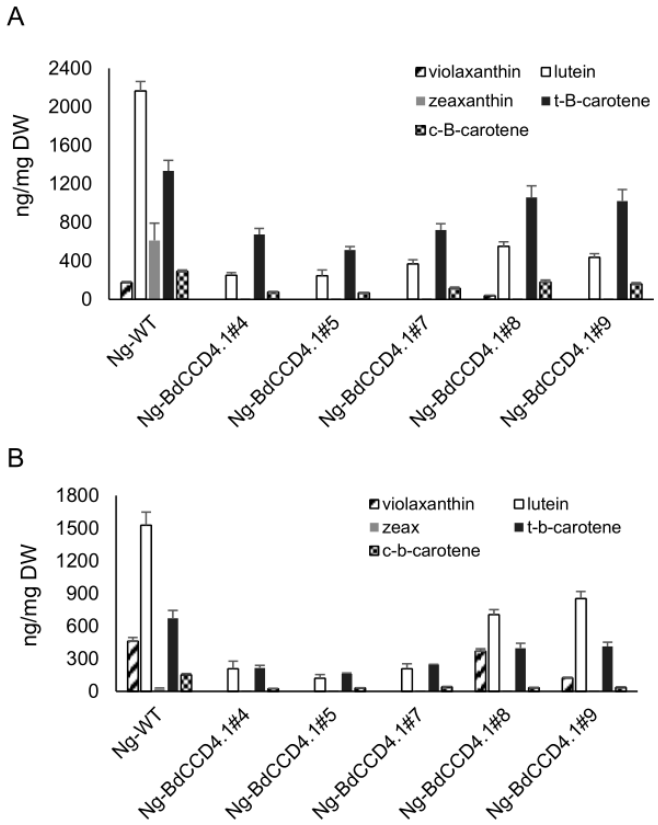


Figure 2.4. Accumulation of carotenoids in WT and transgenic T1 lines of *N. glauca* transformed with *BdCCD4.1*. **(A)** Carotenoid levels in non-polar extracts of leaves from the WT and transgenic lines of *N. glauca*. **(B)** Carotenoid levels in non-polar extracts of petals from the WT and transgenic *N. glauca*. Analyses were conducted in triplicate. Error bars represent the SD. DW, dry weight.

2.4.4. Stability of endogenous apocarotenoids in leaves and petals of transgenic *N. glauca*

No information is available on the stability of apocarotenoids during the storage of tissues from transgenic plants. Similar to other carotenoids, crocins, being highly unsaturated, are prone to degradation during storage. However, in contrast to the lipophilic carotenoids, the stability of crocins is largely affected by water (Tsimidou and Biliaderis 1997). We determined the stability of crocins in lyophilized samples stored at room temperature after a 3-year period. The general chromatographic profile at 440 nm of polar extracts from the leaves and petals of the transgenic samples analyzed after 3 years of storage was clearly different (**Figure 2.5**). In general, in the petals and leaves, there was a clear reduction in the content of crocins, and crocins with two and one glucose molecules predominated (**Figure 2.5A, B**). In addition, crocin degradation in petals was less pronounced than in leaves, at 54-60% in petals and 71-83% in leaves.

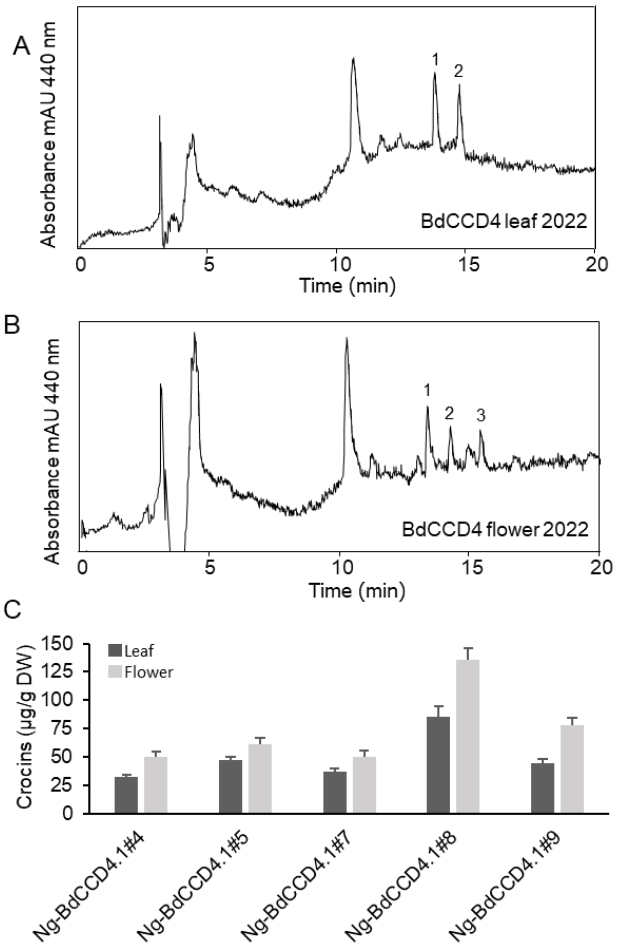


Figure 2.5. Accumulation of crocins in the WT and transgenic T1 lines of *N. glauca* constitutively expressing *BdCCD4.1*. **(A)** HPLC-DAD analysis of polar extracts from WT and *N. glauca* leaves at 440 nm. **(B)** HPLC-DAD analysis of polar extracts from transgenic and WT *N. glauca* petals at 440 nm. Peaks for the abundant crocins in the transgenic lines are denoted by numbers in leaves (1 = trans-crocin 2; 2 = trans-crocin 1) and in petals (1 = trans-crocin 2; 2 = trans-crocin 1; 3 = cis-crocin 2). Crocins were completely absent from the Wt plants. mAU, milli-absorbance units. **(C)** Apocarotenoid accumulation in leaves and petals of transgenic lines. Analyses were done in triplicate. Error bars represent the SD. DW, dry weight.

2.5. Discussion

The Production of high-value saffron apocarotenoids in heterologous systems has been reported in bacteria, yeast, and plants with different degrees of success, including tomato, *N. benthamiana*, and *N. glauca* (Liu et al. 2020; Martí et al. 2020; Ahrazem et al. 2022). Among the different heterologous systems, plants are an attractive production platform due to the presence of native enzymes for the proper production of carotenoid substrates for the biosynthesis of crocins. Previously, the *BdCCD4.1* was transiently expressed in *N. benthamiana* plants using a viral vector , resulting in the accumulation of up to 79±8.6 µg/g DW of crocins in leaves. In this chapter, I investigated whether the stable expression of *BdCCD4.1* might result in an increased level of crocins when expressed in *N. glauca* plants. Recently, *N. glauca* and *N. benthamiana* transgenic lines expressing *CsCCD2L* were evaluated for their capacity to produce crocins , with transgenic *N. glauca* plants

shown to be more efficient, suggesting that, in *N. benthamiana*, other factors may limit crocin accumulation.

We evaluated the levels of crocins in the leaves and petals of control and transgenic lines of *N. glauca* plants. Crocins were only detected in transgenic lines, and in three out of the five lines I analyzed, a higher accumulation of crocins was observed in leaves, while two lines showed similar levels of crocins in petals and leaves. However, the accumulated crocins were qualitatively different in terms of their glucosylation patterns. The addition of sugar moieties to the crocetin molecule is mediated in *N. glauca* by unspecific UDP-glycosyltransferases, which increases the solubility of crocins, allowing their transport to the vacuole (Bowles et al. 2006). Similarly, in *N. glauca* plants modified to produce ketocarotenoids, the levels of ketocarotenoids were higher in petals than in the photosynthetic tissues, while the profile of the different ketocarotenoids was also different (Mortimer et al. 2017). Interestingly, in transgenic lines expressing *CsCCD2L*, the levels of crocins in the petals were lower than those in the leaves. The higher accumulation in leaves might have been due to the presence of higher levels of the substrate zeaxanthin (Zhu et al. 2007). Nevertheless, this possible tissue-specific activity of endogenous CCDs in leaves and petals, acting to reduce the carotenoid pool, should also be considered, as well as the availability of free, non-esterified zeaxanthin for crocin biosynthesis.

The stability of bioactive compounds with putative health benefits under different storage conditions is crucial to guarantee their properties for a long period, and this is especially relevant for high-value metabolites such as crocins (Jafari et al. 2018). Similar to other carotenoids, crocins have a highly unsaturated chemical structure and are consequently susceptible to degradation during storage (Selim et al. 2000). They are especially sensitive to heat and light, and their stability is also affected by water (Ordoudi et al. 2015). Earlier studies with dry saffron stigma stored at room temperature showed a 20% loss of crocins in a 12-week period (Shahidi et al., 2008), and for standardized extracts kept at 4 °C for 22 months, a 30% loss was reported (Suchareau et al., 2021). Therefore, it is necessary to test the stability of crocins in crocin-rich *N. glauca* plants in order to facilitate future applications. Freeze-drying is considered as the most effective method for preserving and maintaining bioactive chemical compounds in plant tissues (Ishwarya et al. 2015) and is especially recommended for heat-sensitive materials and biotechnological products. Freeze-dried commercial saffron contained higher amounts of safranal and crocin compared with different traditional methods for saffron preservation (Acar et al. 2015; Chen et al. 2020). In this study, crocins were extracted from freeze-dried leaves and petals,

and the tissues were stored at room temperature in the dark for a long period of time. After three years of storage, the levels of crocins were reduced by an average of 58 % in the petals, and 78% in the leaves. The higher stability of crocins in the petals might have been due to the different crocin moieties that accumulated in each tissue. The petals contained crocins with a more complex glucosylation pattern, with crocins with three and four glucose molecules, while crocins with two glucose molecules were predominant in the leaves. Although studies on saffron stability showed similar kinetics of degradation for crocins with different degrees of glucosylation, the presence of a greater number of glucose molecules suggested a slower rate of degradation of crocins in the samples (Morimoto et al. 1994; Sánchez et al. 2008). In the petals, the crocins with two glucose molecules were in a cis- configuration. Studies on carotenoid isomers have shown that some cis-isomers are more stable than trans-isomers (Khoo et al. 2011), and this could be the case for crocins as well. Previous studies on the stability of β-carotene from Golden Rice during storage indicated losses after six months ranging from 68% to 80% depending on the processing, temperature, and packing conditions (Bollinedi et al. 2019). Similarly, 50 to 65% carotenoid losses have been reported for biofortified maize after 4 to 6 months under conventional storage conditions (Burt et al. 2010; Ortiz et al. 2016).

2.6. Conclusions

The results of this study demonstrate that the stable expression of *BdCCD4.1* in *N. glauca* plants resulted in the accumulation of up to 321.6±21.3 µg/g DW and 302.7±25.6 µg/g DW of crocins in petals and leaves, respectively. However, these levels were lower compared to plants expressing *CsCCD2L*. Crocin levels in *N.glauca* (even after 3 years of storage) were still higher than those previously reported in transient expression experiments in *N.benthamiana* plants (Martí et al. 2020), or in *N. tabacum* by chloroplast transformation expressing a *CCD4* from *Bixa Orellana* (Frusciante et al. 2021), demonstrating that stable expression of *BdCCD4.1* in *N. glauca* not only affords high levels of crocins, but also provides important information about the stability of these compounds in different tissues.

2.7. References

Acar B, Sadikoglu H, Doymaz I. (2015). Freeze-drying kinetics and diffusion modeling of saffron (Crocus sativus L.). *Journal of Food Processing and Preservation* **39**: 142-149.

Ahrazem O, Diretto G, Argandona J, Rubio-Moraga A, Julve JM, Orzaez D, Granell A, Gomez-Gomez L. (2017). Evolutionarily distinct carotenoid cleavage dioxygenases are responsible for crocetin production in Buddleja davidii. *J. Exp. Bot.* **68**: 4663-4677.

Ahrazem O, Gomez-Gomez L, Rodrigo MJ, Avalos J, Limon MC. (2016a). Carotenoid cleavage oxygenases from microbes and photosynthetic organisms: features and functions. *Int. J. Mol. Sci.* **17**: 1781.

Ahrazem O, Rubio-Moraga A, Berman J, Capell T, Christou P, Zhu C, Gomez-Gomez L. (2016b). The carotenoid cleavage dioxygenase CCD2 catalysing the synthesis of crocetin in spring crocuses and saffron is a plastidial enzyme. *New Phytol.* **209**: 650-663.

Ahrazem O, Zhu C, Huang X, Rubio-Moraga A, Capell T, Christou P, Gómez-Gómez L. (2022). Metabolic engineering of crocin biosynthesis in Nicotiana species. *Front. Plant Sci* **13**: 861140.

Bai C, Rivera SM, Medina V, Alves R, Vilaprinyo E, Sorribas A, Canela R, Capell T, Sandmann G, Christou P, Zhu C. (2014). An in vitro system for the rapid functional characterization of genes involved in carotenoid biosynthesis and accumulation. *Plant J.* **77**: 464-475.

Bollinedi H, Dhakane-Lad J, Gopala Krishnan S, Bhowmick PK, Prabhu KV, Singh NK, Singh AK. (2019). Kinetics of β-carotene degradation under different storage conditions in transgenic Golden Rice® lines. *Food Chemistry* **278**: 773-779.

Bowles D, Lim EK, Poppenberger B, Vaistij FE. (2006). Glycosyltransferases of lipophilic small molecules. *Annu. Rev. Plant Biol.* **57**: 567-597.

Bruno M, Koschmieder J, Wuest F, Schaub P, Fehling-Kaschek M, Timmer J, Beyer P, Al-Babili S. (2016). Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites. *J. Exp. Bot.* **67**: 5993-6005.

Burt AJ, Grainger CM, Young JC, Shelp BJ, Lee EA. (2010). Impact of postharvest handling on carotenoid concentration and composition in high-carotenoid maize (Zea mays L.) kernels. *J. Agric. Food Chem.* **58**: 8286-8292.

Chai F, Wang Y, Mei X, Yao M, Chen Y, Liu H, Xiao W, Yuan Y. (2017). Heterologous biosynthesis and manipulation of crocetin in Saccharomyces cerevisiae. *Microb. Cell Factories* **16**: 1-14.

Chen D, Xing B, Yi H, Li Y, Zheng B, Wang Y, Shao Q. (2020). Effects of different drying methods on appearance, microstructure, bioactive compounds and aroma compounds of saffron (Crocus sativus L.). *LWT* **120**: 108913.

Diretto G, López-Jiménez AJ, Ahrazem O, Frusciante S, Song J, Rubio-Moraga Á, Gómez-Gómez L. (2021). Identification and characterization of apocarotenoid modifiers and carotenogenic enzymes for biosynthesis of crocins in Buddleja davidii flowers. *J. Exp. Bot.* **72**: 3200-3218.

Frusciante S, Demurtas OC, Sulli M, Mini P, Aprea G, Diretto G, Karcher D, Bock R, Giuliano G. (2021). Heterologous expression of Bixa orellana cleavage dioxygenase 4–3 drives crocin but not bixin biosynthesis. *Plant Physiol*.**188**: 1469-1482.

Gonzalez-Jorge S, Ha S-H, Magallanes-Lundback M, Gilliland LU, Zhou A, Lipka AE, Nguyen Y-N, Angelovici R, Lin H, Cepela J, Little H, Buell CR, Gore MA, DellaPenna D. (2013). CAROTENOID CLEAVAGE DIOXYGENASE4 Is a Negative Regulator of β-Carotene Content in Arabidopsis Seeds. *The Plant Cell* **25**: 4812-4826.

González A, Tezara W, Rengifo E, Herrera A. (2012). Ecophysiological responses to drought and salinity in the cosmopolitan invader Nicotiana glauca. *Brazilian Journal of Plant Physiology* **24**: 213-222.

Horsch RB, Rogers SG, Fraley RT. (1985). Transgenic plants. *Cold Spring Harbor Symposia on Quantitative Biology* **50**: 433-437.

Ishwarya SP, Anandharamakrishnan C, Stapley AGF. (2015). Spray-freeze-drying: A novel process for the drying of foods and bioproducts. *Trends in Food Science & Technology* **41**: 161-181.

Jafari SM, Bahrami I, Dehnad D, Shahidi SA. (2018). The influence of nanocellulose coating on saffron quality during storage. *Carbohydrate Polymers* **181**: 536-542.

Khoo H-E, Prasad KN, Kong K-W, Jiang Y, Ismail A. (2011). Carotenoids and their isomers: color pigments in fruits and vegetables. *Molecules* **16**: 1710-1738.

Liu T, Yu S, Xu Z, Tan J, Wang B, Liu Y-G, Zhu Q. (2020). Prospects and progress on crocin biosynthetic pathway and metabolic engineering. *Computational and Structural Biotechnology Journal* **18**: 3278-3286.

Mariutti LRB, Mercadante AZ. (2018). Carotenoid esters analysis and occurrence: What do we know so far? *Archives of Biochemistry and Biophysics* **648**: 36-43.

Martí M, Diretto G, Aragonés V, Frusciante S, Ahrazem O, Gómez-Gómez L, Daròs J-A. (2020). Efficient production of saffron crocins and picrocrocin in Nicotiana benthamiana using a virus-driven system. *Metab. Eng.***61**: 238-250.

Morimoto S, Umezaki Y, Shoyama Y, Saito H, Nishi K, Irino N. (1994). Post-harvest degradation of carotenoid glucose esters in saffron. *Planta medica* **60**: 438-440.

Mortimer CL, Bramley PM, Fraser PD. (2012). The identification and rapid extraction of hydrocarbons from Nicotiana glauca: A potential advanced renewable biofuel source. *Phytochemistry Letters* **5**: 455-458.

Mortimer CL, Misawa N, Perez-Fons L, Robertson FP, Harada H, Bramley PM, Fraser PD. (2017). The Formation and Sequestration of Nonendogenous Ketocarotenoids in Transgenic Nicotiana glauca. *Plant Physiol*.**173**: 1617-1635.

Ordoudi SA, Cagliani LR, Lalou S, Naziri E, Tsimidou MZ, Consonni R. (2015). 1H NMR-based metabolomics of saffron reveals markers for its quality deterioration. *Food Research International* **70**: 1-6.

Ortiz D, Rocheford T, Ferruzzi MG. (2016). Influence of temperature and humidity on the stability of carotenoids in biofortified maize (Zea mays L.) genotypes during controlled postharvest storage. *J. Agric. Food Chem.* **64**: 2727-2736.

Sánchez AM, Carmona M, A. Ordoudi S, Z. Tsimidou M, Alonso GL. (2008). Kinetics of Individual Crocetin Ester Degradation in Aqueous Extracts of Saffron (Crocus sativus L.) upon Thermal Treatment in the Dark. *J. Agric. Food Chem.* **56**: 1627-1637.

Selim K, Tsimidou M, Biliaderis CG. (2000). Kinetic studies of degradation of saffron carotenoids encapsulated in amorphous polymer matrices. *Food Chemistry* **71**: 199-206.

Tsimidou M, Biliaderis CG. (1997). Kinetic Studies of Saffron (Crocus sativus L.) Quality Deterioration. *J. Agric. Food Chem.* **45**: 2890-2898.

Usade B, Tohge T, Scossa F, Sierro N, Schmidt M, Vogel A, Bolger A, Kozlo A, Enfissi EMA, Morrel K, Regenauer M, Hallab A, Ruprecht C, Gundlach H, Spannagl M, Koram Y, Mayer KFX, Boerjan W, Fraser P, Persson S, Ivanov N, Fernie A. (2018). The genome and metabolome of the tobacco tree, Nicotiana glauca : a potential renewable feedstock for the bioeconomy. bioRxiv, 351429.

Walter MH, Strack D. (2011). Carotenoids and their cleavage products: Biosynthesis and functions. *Natural Product Reports* **28**: 663-692.

Wang W, He P, Zhao D, Ye L, Dai L, Zhang X, Sun Y, Zheng J, Bi C. (2019). Construction of Escherichia coli cell factories for crocin biosynthesis. *Microb. Cell Factories* **18**: 1-11.

Zheng X, Yang Y, Al-Babili S. (2021). Exploring the Diversity and Regulation of Apocarotenoid Metabolic Pathways in Plants. *Front. Plant Sci* **12**: 787049.

Zhu C, Gerjets T, Sandmann G. (2007). Nicotiana glauca engineered for the production of ketocarotenoids in flowers and leaves by expressing the cyanobacterial crtO ketolase gene. *Transgenic Res.* **16**: 813-821.

Chapter III

Transcriptional regulation of particular MEP and MVA pathway genes in rice seed

3.0. Abstract

Isoprenoids are natural products derived from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). In plants, these precursors can be synthesized via two alternative pathways: the mevalonate (MVA) pathway and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. Therefore, the regulation of these pathways and the function of the corresponding genes need to be elucidated in detail to develop effective strategies for isoprenoid metabolic engineering in plants. In this chapter, I analyzed the promoter regions of five genes (*AACT3, HMGS1, HMGR1, DXS2, IPPI1*) involved in the rice MEP and MVA pathways. I found that the 300 bp promoter regions upstream of the transcription start sites of all these genes contain a G-box or a hybrid G/C-box that is specifically bound by the *Os*BZ8 transcription factor. I have determined the expression of *OsBZ8* in different rice tissues and compared it with the expression patterns of *AACT3, HMGS1, HMGR1, DXS2,* and *IPPI1* in rice endosperm and embryo to understand how the OsBZ8 regulates the expression of these genes. I found that *OsBZ8* and these five genes exhibited similar expression patterns in rice endosperm and embryo, suggesting that OsBZ8 regulates the expression of these genes in rice seed. Moreover, I have confirmed that the OsBZ8 specifically binds to a G-box or a hybrid G/C-box in the promoters of *HMGR1, DXS2, and IPPI1* (rate-limiting enzymes). These findings point towards important aspects of the regulatory mechanisms of the genes and provide new insights into rice seed quality breeding using genetic engineering.

3.1. Introduction

3.1.1. Isoprenoids and Isoprenoids biosynthesis

Isoprenoids, also called terpenoids, are a large group of natural products with important medical and industrial properties. They comprise tens of thousands of chemicals with diverse functional groups. They include essential metabolites such as sterols, acting as membrane stabilizers or as precursors of steroid hormones, carotenoids of photosynthesizing organisms as well as a multitude of secondary metabolites with less obvious roles (Krivoruchko and Nielsen 2015). Many isoprenoids are used as pharmaceuticals, nutrients, fragrances, flavoring substances, fine chemicals and fuels (Ajikumar et al. 2008). Most of these compounds were originally discovered in plants and extraction from plant tissues remains a major production route (Ajikumar et al. 2010). Although they possess diverse structures and biological functions, they are all derived from two precursors IPP and DMAPP. Isoprenoid biosynthesis begins with the formation of IPP and its isomer DMAPP, by the MVA pathway in the cytosol and the MEP pathway in plastids (Vranova et al. 2013) (**Figure 3.1**)**.** Most organisms have only one of these pathways, for example, yeast and animals only possess the MVA pathway, and green algae only possess the MEP pathway (Disch and Rohmer 1998; Disch et al. 1998; Kovacs et al. 2002). In plants, IPP and DMAPP can be synthesized via both pathways. Although the two pathways are compartmentalized in plants, cross-flow of intermediates has been demonstrated and it depends on the type of supplied precursor, the type of final product, and cultivation conditions (Schuhr et al. 2003). The enzyme deoxyxylulose phosphate synthase (DXS) catalyzes the initial step in the MEP pathway (**Figure 3.1**). It is encoded by a multigene family in plants (Vranova et al. 2013). This multigene family consists of three paralogs. The first, DXS1, plays an essential role in the biosynthesis of photosynthetically derived terpenoids such as chlorophylls and carotenoids in the leaves of monocots and dicots. DXS2 plays a role in the production of secondary terpenoid metabolites such as non-provitamin A carotenoids in the kernels of yellow maize. DXS3 has not been functionally characterized and it is only present in the Poaceae family and a few dicots and monocots among Angiosperms (Miyamoto et al. 2015). In rice and maize, DXS3 expression was detectable in several tissues at low levels, in particular in the maize endosperm where it correlates with carotenoid accumulation (Miyamoto et al. 2015). DXS enzymes have been reported to function differently in different plant species. For example, in Arabidopsis and tobacco, DXS2 is a rate-limiting enzyme for leaf carotenoids and chlorophyll (You et al. 2020). Studies in ripening tomato fruits and potato tubers have

shown that overexpression of *Escherichia coli DXS* resulted in a significant increase in the accumulation of isoprenoids, specifically phytoene and β-carotene (Enfissi et al. 2005; Morris et al. 2006; Bai et al. 2016). The MVA pathway starts in the cytoplasm with the conversion of three molecules of acetyl-CoA to 3-hydroxy-3-methylglutaryl (HMG-CoA) by acetoacetyl-CoA thiolase and HMG-CoA synthase (AACT and HMGS, respectively). In the rate-limiting step of the pathway, HMG-CoA reductase (HMGR), converts HMG-CoA to mevalonate. Mevalonate phosphate (MVAP), after phosphorylation of mevalonate, enters the peroxisome, where it is converted to the five-carbon building block isopentenyl pyrophosphate (IPP) (Henry et al. 2018). Isopentenyl diphosphate isomerase (IPPI) is involved in the final step of the MVA/MEP pathways to convert IPP to DMAPP. The sequential addition of IPP units to DMAPP generates different prenyl diphosphates such as geranyl diphosphate (GPP, C10), farnesyl diphosphate (FPP, C15, and C30), and geranylgeranyl diphosphate (GGPP, C20, and C40)—the precursors of monoterpenes, sesquiterpenes, and diterpenes, respectively (Vranova et al. 2013; Vavitsas et al. 2018).

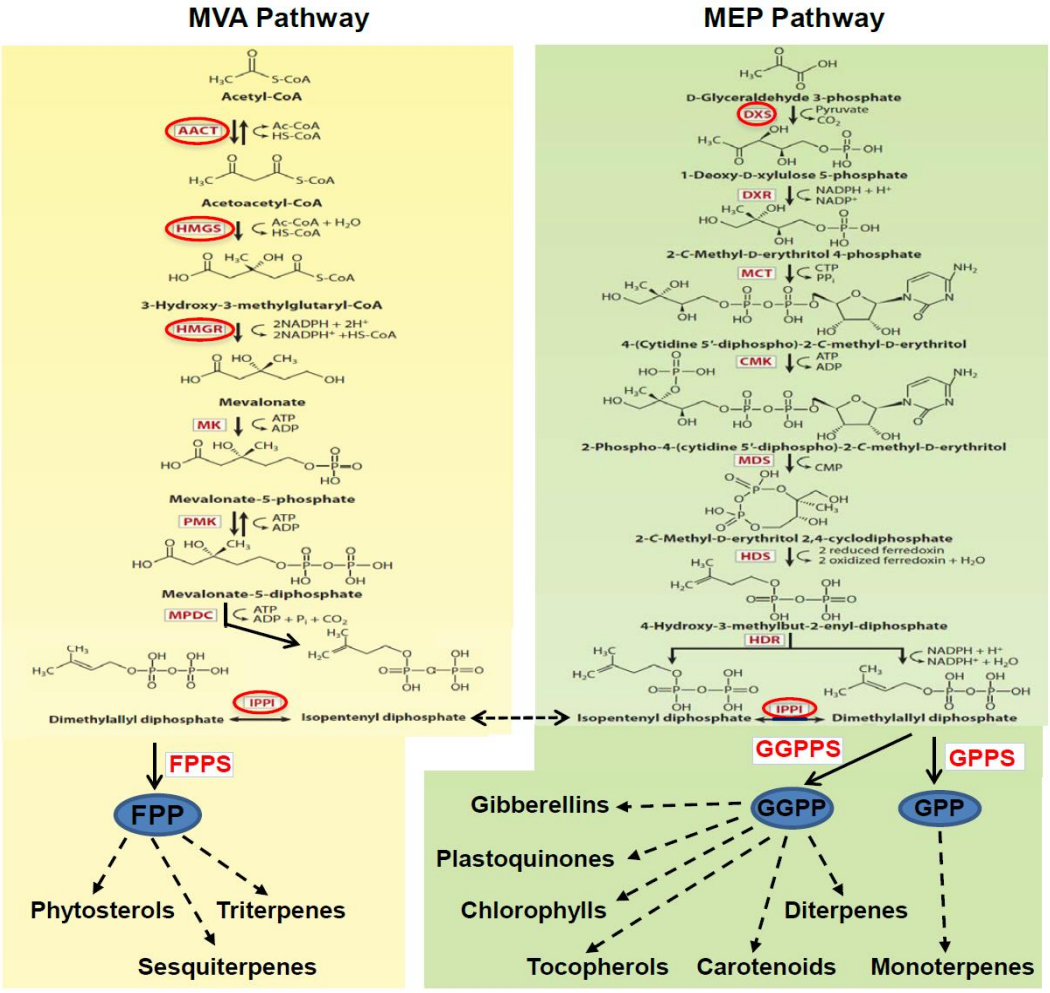


Figure 3.1. Mevalonate (MVA) and 2-C-methyl-D-erythritol-4-phosphate (MEP) pathways for the synthesis of prenyl diphosphates in higher plants. The metabolic pathway in yellow is the MVA pathway; the metabolic pathway in green is the MEP pathway (Modified from Vranova et al. 2013).

3.1.2. Transcriptional regulation of isoprenoid biosynthesis in plants

Regulation of gene expression extends through all stages of plant growth and development. In addition to the enzymes involved in the isoprenoid biosynthetic pathway, transcription factors (TF) also control the formation of intermediates and final products (Okada et al. 2009). TF play a central role in the transcription process by recruiting, interacting, or adjusting components of the transcriptional machinery, activating or inhibiting the transcription of target genes (Okada et al. 2009). They are binding proteins located in the nucleus that specifically interact with *cis*-acting elements in the promoter regions of genes to regulate the expression of the genes (Sahu et al. 2022).

Previous studies in our laboratory identified TF based on screening of *cis*-acting elements in the promoter regions of genes of the MEP pathway (Jin et al. 2019). The bZIP class Abscisic acid Responsive Element (ABRE)-binding factor OsBZ8 cloned from rice embryos specifically binds to the G-box or hybrid G/C-box within the 300 bp upstream of the target gene transcription start sites (TSS) and activates abscisic acid (ABA)-responsive gene expressions such as *Osem* and *Rab16A* (Nakagawa et al. 1996; RoyChoudhury et al. 2008). In this chapter, retrieval of the promoter sequences corresponding to 300 bp upstream of the TSS of *OsAACT3*, *OsHMGS1*, *OsHMGR1*, *OsDXS2*, and *OsIPPI1*, confirmed that promoter regions of these genes contain a G-box or a hybrid G/C-box where TF OsBZ8 recognizes and specifically binds. I confirmed that *OsBZ8* activates the expression of *OsHMGR1*, *OsDXS2*, and *OsIPPI1*, suggesting a likely regulatory mechanism controlling the expression of these genes.

3.2. Aims and objectives

The aim of this chapter was to investigate the transcriptional regulation of the transcription factor *Os*BZ8 on particular MEP and MVA pathway genes in rice embryo and endosperm.

Specific objectives were addressed:

- analyze the promoter regions of particular MVA and MEP pathway genes to determine whether the promoters include specific elements that can be recognized and bound by the *Os*BZ8 transcription factor.
- determine the expression patterns of *OsBZ8* in different rice tissues.
- compare the expression pattern of *OsBZ8* with that of selected genes (*AACT3, HMGS1, HMGR1, DXS2, IPPI1*) in rice endosperm and embryo.
- determine the role of *Os*BZ8 on the expression of the selected genes in the MVA and MEP pathways.

3.3. Materials and methods

3.3.1. Gene and promoter sequences and identification of *cis*-regulatory elements

Sequences of *OsBZ8, OsAACT, OsHMGS, OsHMGR, OsDXS, OsIPPI* and corresponding promoters (2 kb upstream of TSS) of rice (*Oryza. sativa* L. cv Nipponbare) involved in MEP and MVA pathways under consideration were retrieved from NCBI (*National Center for Biotechnology Information*) (https://www.ncbi.nlm.nih.gov). The tools PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) (Lescot et al. 2002) and PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE) (Higo et al. 1999) were used to identify *cis*-regulatory elements in the corresponding promoters.

3.3.2. Plant materials

Rice (*Oryza sativa* L. cv. EYI105) plants were grown in a controlled growth chamber with a 28/20°C day/night temperature regime, a 10h photoperiod and 60–90% relative humidity. Flag leaf, endosperm, and embryo tissues were collected 25 days after pollination. Root, stem, and leaf tissues were collected at 7 and 14 days after germination. Immature panicle, ovary, inner glume, and outer glume tissues were collected 20 days after pollination. All plant materials were rapidly frozen in liquid nitrogen and stored at –80°C.

3.3.3. Gene expression analysis

Total RNA was isolated from the different tissues using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and DNA was removed with DNase I (RNase-Free DNase Set, Qiagen) (Jin et al. 2021). First-strand cDNA was synthesized from 1 µg total RNA using Ominiscript Reverse Transcriptase (Qiagen) in a 20 µL total reaction, and qRT-PCR was performed on a CFX96 system (Bio-Rad Laboratories, Hercules, CA, USA) using the primers listed in **Table 3.1**. The amplified DNA fragments for each gene were confirmed by sequencing. The expression levels of the genes in different samples were normalized against rice actin mRNA (*OsActin*). Three biological replicates each comprising three technical replicates were used for each sample.

Table 3.1. Primers used for qRT-PCR analysis.

Gene	Forward primer (5′- 3′)	Reverse primer (5′- 3′)
OsBZ8	CCCGCACGATTAACATCGAA	ATATGAACAACGCGAGCCAG
OsAACT1	TATAAGTGACGGTGCTGCCG	ACACGAGAAGACTCCAAGCC
OsAACT2	ACAGTGAACCAAGTTCGCCT	TTCCTGTGGGACTCAACAGC
OsAACT3	GAATGAAATGAAGCGCCGGG	TGGCGCACGTTTGAGAATTG
OsDXS1	CTCAAGGGAGGGAAGAACAA	ACACCTGCTTGTTGTCGTTG
OsDXS2	TGTTGTGGAGCTCGCTATTG	TCCTCCCACCTAGATCCCTT
OsDXS3	ACCTCCTCGGGAAGAAGAAG	GAGGGACACCTGCTTGTTGT
OsHMGS1	GCCTACGCCTTCCTCCCAAT	ATGCCCACGTCCTTCCTCTC
OsHMGS2	GGGATGGACGCTACGGTCTT	TAGCAGCAGCACCACCTGTT
OsHMGS3	TTGTTGCCTCCTGGGACGTT	GATCTCCTCGTCGGCCTTCC
OsHMGR1	CTCAAGGGAGGGAAGAACAA	ACACCTGCTTGTTGTCGTTG
OsHMGR2	TGTTGTGGAGCTCGCTATTG	TCCTCCCACCTAGATCCCTT
OsHMGR3	ACCTCCTCGGGAAGAAGAAC	GAGGGACACCTGCTTGTTGT
OsIPPI1	GGTGGATGAACAAGACAATG	ATCGTTGCTGGAGTAGGAGT
OsIPPI2	TTTAGTGGACGAACAGGACA	GCAGACCTTTGCTGAAGTAAC
OsActin	GACTCTGGTGATGGTGTCAGC	GCTTCTCCTTTATGTCTCTGAC

OsBZ8, rice (*Oryza sativa*) the bZIP class Abscisic acid Responsive Element (ABRE)-binding factor; *OsAACT*, rice acetoacetyl-CoA thiolase; *OsDXS1/2/3*, rice 1-deoxy-D-xylulose-5-phosphate synthase 1/2/3 genes; *OsHMGS1/2/3*, rice HMG-CoA synthase 1/2/3 genes; *OsHMGR1/2/3*, rice HMG-CoA reductase 1/2/3 genes; *OsIPPI1/2*, rice isopentenyl diphosphate isomerase 1/2 genes; *OsActin*, rice actin gene.

3.3.4. Yeast one-hybrid system

The yeast one-hybrid system involves two components: ‘bait DNA’ and ‘prey protein’. The bait DNA sequence was cloned into pAbAi and the prey protein coding sequence (CDS) was cloned into pGADT7 AD. Resistance to aureobasidin A (AbA) was conferred by the *AUR-1C* on the pAbAi vector. The *Ura* and *Leu* nutrient marker genes were on the pAbAi and pGADT7 vectors, respectively. In addition, the pGADT7 AD vector contains the yeast GAL4 transcription activation domain (AD), which activates the expression of the AbA reporter gene when the prey protein binds to the bait DNA sequence, allowing cells to grow on media containing the AbA antibiotic.

The 300 bp wild-type promoter fragment (abbreviated as PF in the vector names) upstream of the TSS of *OsHMGR1*, *OsDXS2*, and *OsIPPI1* from rice leaf was amplified using primers containing additional sequences providing appropriate restriction sites

(**Table 3.2**). Controls were synthetic promoter fragments OsHMGR1PF-mG/C, OsDXS2PF-mG/C, OsIPPI1PF-mG containing the corresponding mutated hybrid G/C-box (mG/C, 5′-GACGTG-3′ to 5′-GgCcTG-3′) or mutated G-box (mG, 5′-CACGTG-3′ to 5′-CgCcTG-3´). The wild-type and synthetic promoter fragments were independently inserted into the pAbAi vector at the HindIII and XhoI sites to form the bait-pAbAi vector, which were then introduced into yeast strain Y1HGold using Yeastmaker Yeast Transformation System 2 (Clontech) grown on medium lacking uracil. To determine the minimal inhibitory concentration of AbA, the bait strains were separately screened on synthetic defined medium lacking uracil and containing different concentrations of AbA (50, 100, 200, 300, or 400 ng/ml). CDS of *OsBZ8* from rice embryo was introduced into pGADT7 AD vector at the EcoRI and BamHI sites to generate the prey vector pGADT7-OsBZ8. The prey vector was transformed into the bait strains and grown on synthetic defined medium lacking leucine and containing 200ng ml^{-1} AbA (minimal inhibitory concentration) for 72h at 30°C. The integrity of all intermediate and final constructs was confirmed by sequencing.

Table 3.2. Primers used for the construction of yeast one-hybrid vectors.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
OsHMGR1	AAGCTTAGCGAGTTCATCACATCAACCAC	CTCGAGCTAAGCGGCAAGCCGCAAGCAGA
OsDXS2	AAGCTTGAGGGGAGAACGCGGAGAGCGGA	CTCGAGGCGCTCTTGACATGACAGATTAG
OsIPPI1	AAGCTTGACATAAAGAGACAAGAGAACAC	CTCGAGGGAGGCGGGGTGATAAATAACGG
OsBZ8	GAATTCGGAGAATAATGATGATGACGA	GGATCCTTGCGGGGAAGGGGATGAGT

OsHMGR1, rice (*Oryza sativa*) HMG-CoA reductase 1 gene; *OsDXS2*, rice 1-deoxy-D-xylulose-5-phosphate synthase 2 gene; *OsIPPI1*, rice isopentenyl diphosphate isomerase 1 gene; *OsBZ8*, rice (*Oryza sativa*) the bZIP class Abscisic acid Responsive Element (ABRE)-binding factor; Hind*III* (5′-AAGCTT-3′) and XhoI (5′-CTCGAG-3′) sites (underlined) were introduced at the 5' ends of each primer of the bait sequences. EcoRI (5′-GAATTC-3′) and BamHI (5′-GGATC-3′) sites (underlined) were introduced at the 5' ends of each primer of the prey sequence.

3.3.5. Statistical analysis

Significant differences in quantitative dependent variables for any two alternative treatments were calculated using Student's *t*-test and more than two treatments using Duncan's test, with a *p*-value of 0.05 considered significant.

3.4. Results

3.4.1. In silico promoter analysis of genes involved in the MEP and MVA pathways

The full-length rice (*O. sativa* cultivar: Nipponbare) CDS of particular genes involved in the MVA and MEP pathways was used to identify their corresponding promoter regions [up to 300 bp upstream of the transcription start sites (TSS)] (shown in **Table 3.4**).

Table 3.3. In silico analysis of cis-acting regulatory elements in the promoters of selected MVA and MEP pathway genes.

Gene	*Cis*-acting regulatory elements	GenBank accession number	Chromosome number
OsAACT1	---	XM_015795681 (mRNA); AP014965	9
OsAACT2	---	XM_015765825 (mRNA); AP014957	1
OsAACT3	hybrid G/C-box (GACGTG, -144)	XM_015769949 (mRNA); AP014958	2
OsAACT4	----	XM_015759346 (mRNA); AP014966	10
OsHMGS1	G-box (CACGTG, -89)	XM_015757147 (mRNA); AP014965	9
OsHMGS2	---	NM_001403686(mRNA); AP014964	8
OsHMGS3	---	XM_015776782 (mRNA); AP014959	3
OsHMGR1	hybrid G/C-box (GACGTG, -107)	XM_015768351 (mRNA); AP014958	2
OsHMGR2	---	XM_015756838 (mRNA); AP014965	9
OsHMGR3	---	XM_015792764 (mRNA); AP014964	8
OsDXS1	---	NM_001402527 (mRNA); AP014961	5
OsDXS2	hybrid G/C-box (GACGTG, -69)	XM 015787004 (mRNA); AP014962	6
OsDXS3	--	AK_100909 (mRNA); AP014963	7
OsIPPI1	G-box (CACGTG, -130)	XM_015791312 (mRNA); AP014963	7
OsIPPI2	---	XM_015782225 (mRNA); AP014961	5

The positions of the *cis*-acting regulatory elements are relative to the first nucleotide of the cDNA of the corresponding genes. G-box or hybrid G/C-box 300 bp upstream of the TSS is shown in red. Abbreviations: *OsAACT1/2/3*, rice acetoacetyl-CoA thiolase 1/2/3 genes; *OsHMGS1/2/3*, rice HMG-CoA synthase 1/2/3 genes; *OsHMGR1/2/3*, rice HMG-CoA reductase 1/2/3 genes; *OsDXS1/2/3*, rice 1-deoxy-D-xylulose-5-phosphate synthase 1/2/3 genes; *OsIPPI1/2*, rice isopentenyl diphosphate isomerase 1/2 genes.

3.4.2. mRNA accumulation patterns of *OsBZ8* in different tissues

OsBZ8 mRNA accumulation in different rice tissues was determined by qRT-PCR. *OsBZ8* mRNA was abundant in all tissues except the stem. The highest *OsBZ8* mRNA accumulation was obtained 14 days after germination in leaves. *OsBZ8* mRNA was 1.8-fold more abundant compared to 7-day-old leaves. *OsBZ8* mRNA also accumulated in roots with no obvious difference in abundance between 7 and 14 days after germination. The levels of *OsBZ8* mRNA accumulation significantly increased (1.7-fold) in the embryo compared to the endosperm, 25 days after pollination (**Figure 3.2**).

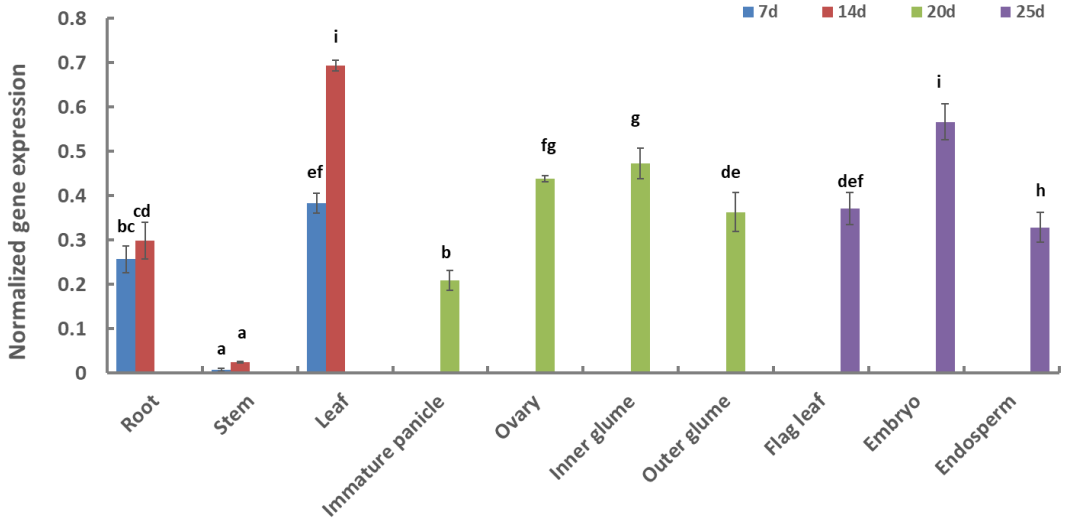


Figure 3.2. mRNA accumulation of *OsBZ8* in different tissues. Root, stem, and leaf tissues were collected at 7 and 14 days after germination, whereas immature panicle, ovary, inner glume, and outer glume tissues were collected 20 days after pollination. Flag leaf, embryo and endosperm tissues were collected 25 days after pollination. Values represent the mean of three biological replicates and bars represent standard deviations. Expression levels were normalized to the rice actin gene. Different letters indicate significant differences among groups at $P< 0.05$ by Duncan's multiple range test.

3.4.3. mRNA accumulation of *OsBZ8, OsAACT, OsHMGS, OsHMGR, OsDXS* and *OsIPPI* in rice endosperm and embryo

There are three isoforms of *AACT*, *HMGS*, *HMGR*, and *DXS*, and two *IPPI* isoforms in rice. *OsBZ8*, *OsAACT2/3*, *OsHMGS1/2*, *OsHMGR1/2/3*, *OsDXS2*, and *OsIPPI1* mRNA accumulation was detectable in both embryo and endosperm tissues, and the mRNA of all the above genes was more abundant in the embryo than in the endosperm. *OsAACT2* and *OsAACT3* mRNA accumulation were 1.7-fold and 1.9-fold higher in the embryo compared to the endosperm. *OsHMGS1* and *OsHMGS2* mRNA were 2.6-fold and 1.6-fold higher in the embryo than in the endosperm. *OsHMGR1* mRNA was minimally accumulated in both tissues. *OsHMGR2* and *OsHMGR3* mRNA were strongly upregulated in the embryo, ca: 3.4-fold and 3.6-fold higher than in the endosperm,

respectively. *OsDXS2* was the only mRNA isoform detectable in both tissues and the mRNA accumulation was 1.5-fold higher in the embryo compared to the endosperm. *OsIPP1* mRNA was 3.9-fold higher in the embryo than in the endosperm. *OsAACT1*, *OsHMGS3,* and *OsIPPI2* mRNA were detectable only in the embryo **(Figure 3.3)**.

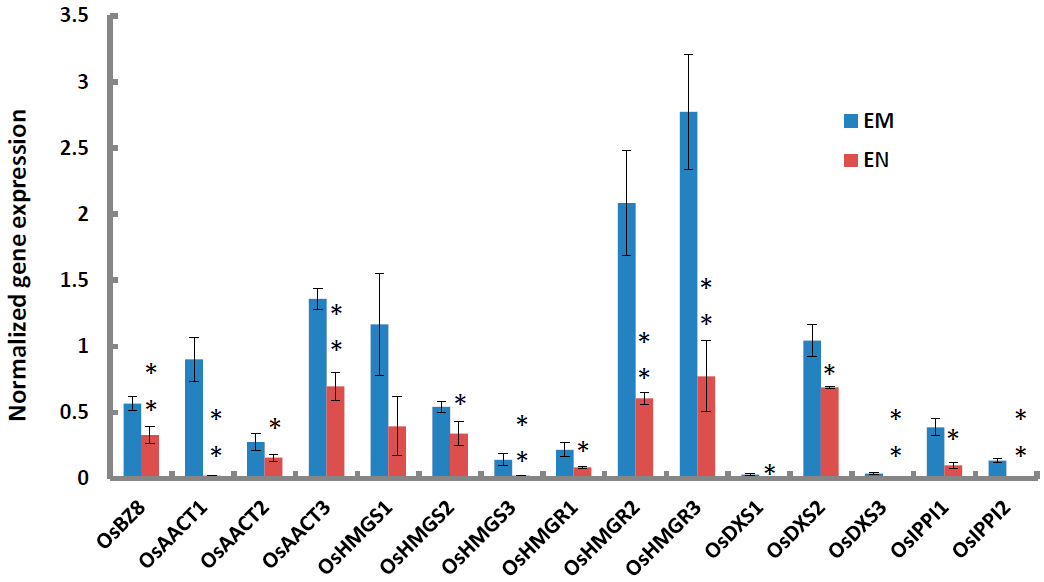


Figure 3.3. *OsBZ8, OsAACT1/2/3, OsHMGS1/2/3, OsHMGR1/2/3, OsDXS1/2/3,* and *OsIPPI1/2* mRNA accumulation in rice embryo (Em) and endosperm (En) 25 days after pollination. Values represent the mean of three biological replicates and bars represent SD. Expression levels were normalized to the rice actin gene as determined by qRT-PCR. Asterisks indicate statistically significant differences (*, P <0.05; **, P <0.01; Student's t-test).

3.4.4. Interactions of *Os*BZ8 with the G-box and hybrid G/C-box promoter motifs

HMGR and DXS are rate-limiting enzymes in the MVA and MEP pathways, respectively. IPPI is shared by both pathways. Consequently, we focused on these three genes for a more in-depth analysis. I used the yeast one-hybrid (Y1H) system to determine whether OsBZ8 interacts with the G-box and hybrid G/C-box in the promoters of *OsHMGR1*, *OsDXS2* and *IPPI1*. Wild-type bait plasmids pAbAi-OsHMGR1PF, pAbAi-OsDXS2PF, pAbAi-OsIPPI1PF, mutant bait plasmids pAbAi-OsHMGR1PF-mG/C, pAbAi-OsDXS2PF-mG/C, pAbAi-OsIPPI1PF-mG and pAbAi void bait plasmids (**3.3.4 Materials and methods**) were separately introduced into Y1HGold yeast cells to form bait strains. They can grow on medium lacking uracil but not on medium containing 200ng ml^{-1} AbA. I used 200 ng ml^{-1} as the minimal inhibitory concentration of AbA for the bait plasmids. The prey plasmid pGADT7-OsBZ8 was transformed into each of the bait strains. Only strains transformed with the corresponding wild-type bait plasmids (A3/B3/C3) containing either an intact G-box or a hybrid G/C-box were able to grow on leucine-deficient medium containing 200ng ml^{-1} aureobasidin A (AbA), whereas strains transformed with pAbAi void bait plasmids (A1/B1/C1) and the corresponding mutant

bait plasmids (A2/B2/C2) hardly grew, indicating that *Os*BZ8 can specifically bind to the hybrid G/C-box or G-box motifs in the *OsHMGR1*, *OsDXS2* and *IPPI1* promoters in yeast. Thus, the Y1H system demonstrated the interactions between TF *Os*BZ8 with *OsHMGR1*, *OsDXS2* and *IPPI1* via binding promoter elements (**Figure 3.4**).

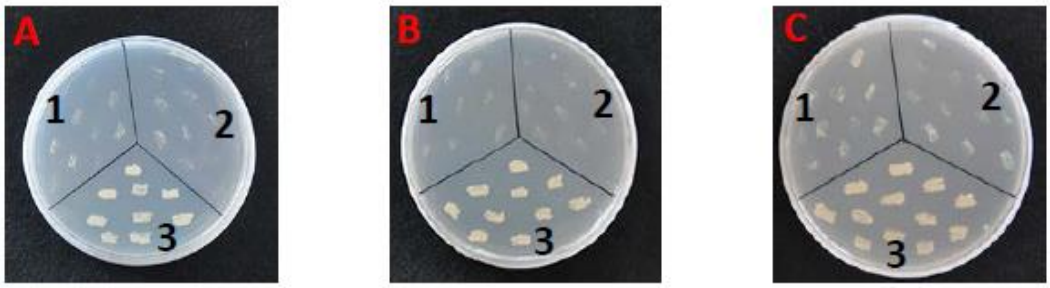


Figure 3.4. Identification and characterization of the TF OsBZ8. *Os*BZ8 recognizes and specifically binds to G-box or hybrid G/C-box within the promoter region 300 bp upstream of TSS of *OsHMGR1* **(A)**, *OsDXS2* **(B)**, and *OsIPPI1* **(C)**. (1) pAbAi void bait plasmid; (2) mutant bait plasmid; (3) wild-type bait plasmid.

3.5. Discussion

The metabolism and function of isoprenoids have been studied extensively at the biochemical, structural and genetic levels (Christianson 2017; Stanley and Yuan 2019; Perez et al. 2022). Transcriptional regulation of isoprenoid biosynthetic genes in plants is a key intervention point to gain insights into isoprenoid metabolism and develop robust strategies for metabolic engineering. Very little is known about the regulation of these genes in plants. A number of TF have been shown to influence carotenogenic genes in dicots. These include RIN and BBX20 in tomato, RAP2.2 and PIF1 in *Arabidopsis*, maize C1 in *Salvia miltiorrhiza*, etc. (Welsch et al. 2007; Toledo-Ortiz et al. 2010; Martel et al. 2011; Xiong et al. 2019). The MADS-box protein RIPENING INHIBITOR (RIN) and SlBBX20, a B-box (BBX) zinc-finger TF, activated the expression of phytoene synthase (*PSY*), which resulted in higher levels of carotenoids in tomato (Martel et al. 2011; Xiong et al. 2019). Experiments in Arabidopsis showed that *At*RAP2.2, a member of the AP2 TF family and phytochrome-interacting factor 1 (PIF1) can bind to the ATCTA motif in the *AtPSY* promoter region, resulting in a reduced accumulation of carotenoids, by negatively regulating PSY enzyme activity (Welsch et al. 2007; Toledo-Ortiz et al. 2010). The heterologous expression of maize *C1* (TF R2R3-MYB) significantly increased the accumulation of diterpenoid tanshinones (bioactive nor-diterpenoid constituents) through up-regulation of the MVA pathway genes *SmAACT, SmHMGS, SmHMGR, SmIPPI* and the MEP pathway genes *SmDXS, SmIPPI* in *Salvia miltiorrhiza* hairy roots (Zhao et al. 2015). These TF promoted or inhibited the expressions of single or multiple genes

involved in isoprenoid biosynthesis in dicots. The regulation of isoprenoid metabolism in monocots is not as well understood, even though carotenoid-deficient cereals are a major target for metabolic engineering (Zhu et al. 2008; Bai et al. 2016; Kempinski et al. 2019). In rice, the basic helix-loop-helix (bHLH) TF MYC2, a global regulator of JA signaling, upregulated the MEP pathway genes *DXR*, *CMK*, *MDS* and *HDS* and induced genes encoding terpene synthases, including *linalool synthase* and *sesquiterpene synthase* (Ogawa et al. 2017). Earlier work from our group unraveled the effects of putative *cis*-regulatory elements (P-box and AACA motif) in the *ZmBCH2* promoter by in silico promoter analysis (Jin et al. 2019). It was shown that TFs ZmPBF and ZmGAMYB independently transactivated the *Zea mays* β-carotene hydroxylase 2 gene (*ZmBCH2*) promoter in maize (Jin et al. 2019). More recently, our group also elucidated the mechanistic basis of coordinated light-upregulated gene expression during the de-etiolation of etiolated rice leaves. Light-responsive *cis*-regulatory elements were identified in all the promoter regions of each light-upregulated gene, providing an important link between the observed phenotype during de-etiolation and the molecular machinery controlling expression of MVA and MEP pathways genes (Jin et al. 2020). However, there is no information on whether a single transcription factor can simultaneously regulate the co-expression of multiple genes in the two pathways in rice.

I used in silico analysis to identify *cis*-acting regulatory elements in the promoters of five structural gene families in the MEP and MVA pathways in rice. I determined that the 300 bp promoter region upstream of the TSS of *OsAACT3*, *OsHMGS1*, *OsHMGR1*, *OsDXS2* and *OsIPPI1* contains a G-box (5´-CACGTG-3´) or a hybrid G/C-box motif (5´-GACGTG-3´). The sequences of these boxes are identical to two abscisic acid (ABA) responsive elements, the typical ABA-responsive element (ABRE, CACGTG[T/C/G]) and the ABRE coupling element (ABRE-CE, [C/A] ACGCG[T/C/A]) (RoyChoudhury et al. 2008). The bZIP TF family (basic region/leucine zipper motifs) are widely involved in plant ABA/abiotic stress responses, such as improving the ability of plants to adapt to drought, salinity, and cold (Yu et al. 2020). bZIP preferentially binds to DNA sequence motifs containing a 5´-ACGT-3´ core, such as G-box (CACGTG), C-box (GACGTC), A-box, (TACGTA), ABRE (ACGTGGC) (Foster et al. 1994). It was shown that OsBZ8, a rice bZIP protein, is the master regulatory trans-acting factor that interacts with ABREs as well as ABRE-CE in the vegetative tissues of indica rice cultivars to enhance ABA/abiotic stress-inducible gene expression (RoyChoudhury et al. 2008). Treatment of

mature rice embryos, seedlings, and suspension-cultured cells with ABA, induced the accumulation of *OsBZ8* mRNA (RoyChoudhury et al. 2008).

I used yeast one-hybrid analysis to elucidate the mechanism of coordinated expression of the selected isoprenoid biosynthetic pathway genes and their interactions with *Os*BZ8. I chose three specific genes for this analysis (*OsHMGR1*, *OsDXS2* and *OsIPPI1*). HMGR and DXS are the main rate-limiting steps in the MEP and MVA pathways (Vranova et al. 2013). IPPI is the only enzyme in both MVA and MEP pathways and our earlier results demonstrated that the MVA pathway is not functional (blocked) without IPPI expression (Jin et al. 2020). Yeast one-hybrid analysis confirmed that the *Os*BZ8 can recognize and specifically bind to the G-box or hybrid G/C-box within 300 bp upstream of TSS of rice *OsHMGR1*, *OsDXS2* and *OsIPPI1*. These results suggest a possible mechanism by which *Os*BZ8 regulates the co-expression of these specific genes in rice.

My results show that *Os*BZ8 was preferentially transcribed in rice embryo with a significant 1.7-fold increase compared to endosperm, and exhibits similar expression patterns to *OsAACT3*, *OsHMGS1*, *OsHMGR1*, *OsDXS2* and *OsIPPI1*. Both OsBZ8 and the above genes were highly expressed in rice embryo, whereas expression in the endosperm was low. Earlier work also reported that the OsBZ8 family members OsTGAP1 and OsbZIP79 are involved in the transcriptional regulation of isoprenoid biosynthesis in rice. Overexpression of *OsTGAP1,* positively regulated diterpenoid phytoalexin biosynthetic gene (*OsDXS3* and *OsKSL4*) transcription by binding to the TGACGT motif in the promoter region, resulting in the formation of momilactones and phytocassanes. These metabolites are not present in wild-type roots. These results agree with my findings that *Os*BZ8 can regulate *OsDXS2* transcription by binding to the hybrid G/C-box (GACGTG). In contrast, the OsTGAP1 interacting protein, OsbZIP79, was shown to act as a negative regulator of diterpenoid phytoalexins. The overexpression of *OsbZIP79* in suspension-culture rice cells resulted in a 50% reduction in the mRNA levels of all MEP pathway genes (including *OsDXS3*) and an 80% reduction in the levels of both momilactones and phytocassanes (Miyamoto et al. 2015). These data collectively suggest multiple TF may act competitively or cooperatively to modulate the expression of isoprenoid biosynthetic genes.

3.6. Conclusions

I have demonstrated that *Os*BZ8 is involved in the transcriptional regulation of a number of MVA and MEP pathway genes. However, overexpression and down-regulation of *OsBZ8* in rice endosperm are required to elucidate the precise contribution of *Os*BZ8 to isoprenoid biosynthesis in rice endosperm. This is important to better understand broader implications of perturbing isoprenoid metabolism in rice endosperm and its impact on the whole plant, which will make it easier to predict the effect of metabolic engineering in cereals for nutritional improvement or the production of valuable metabolites, abiotic stress tolerance, and other biotechnological targets.

3.7. References

Ajikumar PK, Tyo K, Carlsen S, Mucha O, Phon TH, Stephanopoulos G. (2008). Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol Pharm* **5**: 167-190.

Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* **330**: 70-74.

Bai C, Capell T, Berman J, Medina V, Sandmann G, Christou P, Zhu C. (2016). Bottlenecks in carotenoid biosynthesis and accumulation in rice endosperm are influenced by the precursor-product balance. *Plant Biotechnol J* **14**: 195-205.

Christianson DW. (2017). Structural and Chemical Biology of Terpenoid Cyclases. *Chemical Reviews* **117**: 11570-11648.

Disch A, Rohmer M. (1998). On the absence of the glyceraldehyde 3-phosphate/pyruvate pathway for isoprenoid biosynthesis in fungi and yeasts. *FEMS Microbiology Letters* **168**: 201-208.

Disch A, Schwender J, MÜller C, Lichtenthaler HK, Rohmer M. (1998). Distribution of the mevalonate and glyceraldehyde phosphate/pyruvate pathways for isoprenoid biosynthesis in unicellular algae and the cyanobacterium Synechocystis PCC 6714. *Biochemical Journal* **333**: 381-388.

Enfissi EM, Fraser PD, Lois LM, Boronat A, Schuch W, Bramley PM. (2005). Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotechnol J* **3**: 17-27.

Foster R, Izawa T, Chua N-H. (1994). Plant bZIP proteins gather at ACGT elements. *The FASEB Journal* **8**: 192-200.

Henry LK, Thomas ST, Widhalm JR, Lynch JH, Davis TC, Kessler SA, Bohlmann J, Noel JP, Dudareva N. (2018). Contribution of isopentenyl phosphate to plant terpenoid metabolism. *Nat Plants* **4**: 721-729.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27**: 297-300.

Jin X, Bai C, Bassie L, Nogareda C, Romagosa I, Twyman R, Christou P, Zhu C. (2019). ZmPBF and ZmGAMYB transcription factors independently transactivate the promoter of the maize (Zea mays) β-carotene hydroxylase 2 gene. *New Phytol.* **222**: 793-804.

Jin X, Baysal C, Drapal M, Sheng Y, Huang X, He W, Shi L, Capell T, Fraser PD, Christou P, Zhu C. (2021). The Coordinated Upregulated Expression of Genes Involved in MEP, Chlorophyll, Carotenoid and Tocopherol Pathways, Mirrored the Corresponding Metabolite Contents in Rice Leaves during De-Etiolation. *Plants (Basel)* **10**: 1456.

Jin X, Baysal C, Gao L, Medina V, Drapal M, Ni X, Sheng Y, Shi L, Capell T, Fraser PD. (2020). The subcellular localization of two isopentenyl diphosphate isomerases in rice suggests a role for the endoplasmic reticulum in isoprenoid biosynthesis. *Plant Cell Reports* **39**: 119-133.

Kempinski C, Jiang Z, Zinck G, Sato SJ, Ge Z, Clemente TE, Chappell J. (2019). Engineering linear, branched-chain triterpene metabolism in monocots. *Plant biotechnology journal* **17**: 373-385.

Kovacs W, Olivier L, Krisans S. (2002). Central role of peroxisomes in isoprenoid biosynthesis. *Progress in Lipid Research* **41**: 369-391.

Krivoruchko A, Nielsen JB. (2015). Production of natural products through metabolic engineering of Saccharomyces cerevisiae. *Current opinion in biotechnology* **35**: 7-15.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* **30**: 325-327.

Martel C, Vrebalov J, Tafelmeyer P, Giovannoni JJ. (2011). The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner. *Plant Physiol* **157**: 1568-1579.

Miyamoto K, Nishizawa Y, Minami E, Nojiri H, Yamane H, Okada K. (2015). Overexpression of the bZIP transcription factor OsbZIP79 suppresses the production of diterpenoid phytoalexin in rice cells. *Plant Physiol* **173**: 19-27.

Morris WL, Ducreux LJ, Hedden P, Millam S, Taylor MA. (2006). Overexpression of a bacterial 1-deoxy-D-xylulose 5-phosphate synthase gene in potato tubers perturbs the isoprenoid metabolic network: implications for the control of the tuber life cycle. *J Exp Bot* **57**: 3007-3018.

Nakagawa H, Ohmiya K, Hattori T. (1996). A rice bZIP protein, designated OSBZ8, is rapidly induced by abscisic acid. *Plant J.* **9**: 217-227.

Ogawa S, Kawahara-Miki R, Miyamoto K, Yamane H, Nojiri H, Tsujii Y, Okada K. (2017). OsMYC2 mediates numerous defence-related transcriptional changes via jasmonic acid signalling in rice. *Biochemical and Biophysical Research Communications* **486**: 796-803.

Okada A, Okada K, Miyamoto K, Koga J, Shibuya N, Nojiri H, Yamane H. (2009). OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice. *J Biol Chem* **284**: 26510-26518.

Perez L, Alves R, Perez-Fons L, Albacete A, Farre G, Soto E, Vilaprinyo E, Martinez-Andujar C, Basallo O, Fraser PD, Medina V, Zhu C, Capell T, Christou P. (2022). Multilevel interactions between native and ectopic isoprenoid pathways affect global metabolism in rice. *Transgenic Res* **31**: 249-268.

RoyChoudhury A, Gupta B, Sengupta DN. (2008). Trans-acting factor designated OSBZ8 interacts with both typical abscisic acid responsive elements as well as abscisic acid responsive element-like sequences in the vegetative tissues of indica rice cultivars. *Plant Cell Reports* **27**: 779-794.

Sahu B, Hartonen T, Pihlajamaa P, Wei B, Dave K, Zhu F, Kaasinen E, Lidschreiber K, Lidschreiber M, Daub CO, Cramer P, Kivioja T, Taipale J. (2022). Sequence determinants of human gene regulatory elements. *Nature Genetics* **54**: 283-294.

Schuhr CA, Radykewicz T, Sagner S, Latzel C, Zenk MH, Arigoni D, Bacher A, Rohdich F, Eisenreich W. (2003). Quantitative assessment of crosstalk between the two isoprenoid biosynthesis pathways in plants by NMR spectroscopy. *Phytochemistry Reviews* **2**: 3-16.

Stanley L, Yuan Y-W. (2019). Transcriptional Regulation of Carotenoid Biosynthesis in Plants: So Many Regulators, So Little Consensus. *Front Plant Sci* **10**: 1017.

Toledo-Ortiz G, Huq E, Rodríguez-Concepción M. (2010). Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 11626-11631.

Vavitsas K, Fabris M, Vickers CE. (2018). Terpenoid Metabolic Engineering in Photosynthetic Microorganisms. *Genes* **9**: 520.

Vranova E, Coman D, Gruissem W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol* **64**: 665-700.

Welsch R, Maass D, Voegel T, Dellapenna D, Beyer P. (2007). Transcription factor RAP2.2 and its interacting partner SINAT2: stable elements in the carotenogenesis of Arabidopsis leaves. *Plant Physiol* **145**: 1073-1085.

Xiong C, Luo D, Lin A, Zhang C, Shan L, He P, Li B, Zhang Q, Hua B, Yuan Z. (2019). A tomato B-box protein Sl BBX 20 modulates carotenoid biosynthesis by directly activating PHYTOENE SYNTHASE 1, and is targeted for 26S proteasome-mediated degradation. *New Phytol.* **221**: 279-294.

You MK, Lee YJ, Kim JK, Baek SA, Jeon YA, Lim SH, Ha SH. (2020). The organ-specific differential roles of rice DXS and DXR, the first two enzymes of the MEP pathway, in carotenoid metabolism in Oryza sativa leaves and seeds. *BMC Plant Biology* **20**: 1-16.

Yu Y, Qian Y, Jiang M, Xu J, Yang J, Zhang T, Gou L, Pi E. (2020). Regulation Mechanisms of Plant Basic Leucine Zippers to Various Abiotic Stresses. *Front Plant Sci* **11**: 1258.

Zhao S, Zhang J, Tan R, Yang L, Zheng X. (2015). Enhancing diterpenoid concentration in Salvia miltiorrhiza hairy roots through pathway engineering with maize C1 transcription factor. *J Exp Bot* **66**: 7211-7226.

Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T. (2008). Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proc. Natl. Acad. Sci. U.S.A.* **105**: 18232-18237.

Chapter IV

Knocking out strigolactone biosynthetic genes in maize and its impact on broad metabolism and phenotype

4.0. Abstract

Strigolactones (SL) are hormones with diverse functions in plants. In maize, they are produced by Carotenoid Cleavage Dioxygenase *7* (*Zm*CCD7) and Carotenoid Cleavage Dioxygenase *8* (*Zm*CCD8), which catalyze the cleavage of 9-cis-β-carotene and 9-cis-β-apo-10′-carotene, respectively. Their functions include cues for parasitic weed seed germination, modulation of root architecture, promotion of tillering, increased drought stress tolerance and proliferation of symbiotic Arbuscular Mycorrhiza (AM) fungi which help the plant to have an extended root system to improve water and nutrient acquisition. Maize production is intimately associated with two constraints: water and mineral availability, specifically N and P. Precise genome editing using CRISPR/Cas9 technology to create novel maize (corn) varieties with higher productivity while simultaneously minimizing environmental impact directly provides this opportunity. I used CRISPR/Cas9 gene editing technology to modulate the expression of *ZmCCD7* and *ZmCCD8* to assess their function in elite M37W inbred maize lines. I confirmed the presence of SpCas9, SaCas9, and sgRNA vectors in more than 250 independent transgenic plants. I was able to recover SpCas9-induced mutations in two independent plant lines in the coding region of *ZmCCD8*. I was not able to recover any mutations of *ZmCCD7*. SaCas9 accumulation was not detectable by immunoblot analysis even in the lines where SaCas9 expression was detectable at the mRNA level. Nonetheless, in-depth analysis of the *ZmCCD8* mutants that I recovered will contribute toward a better understanding of the molecular basis controlling structural traits in maize. Future genome sequencing of subsequent generations as well as metabolomic analysis is needed to assess fully the impact of these mutations.

4.1. Introduction

4.1.1. Strigolactones and their role in plant growth and development

SL are carotenoid-derived secondary metabolites that play essential roles in the regulation of plant growth and development (Al-Babili and Bouwmeester 2015). Since the first natural Strigolactone, strigol, was isolated from root exudates of cotton (*Gossypium hirsutum*) and identified as a germination stimulant for *Striga lutea*, more than 35 natural strigolactones have been identified in exudates of different plant species (Cook et al. 1966; Li et al. 2023). *Striga hermonthica* among diverse *Striga* species is the main biotic constraint limiting crop yield (Yoneyama 2019). SL also serve as host recognition signals for symbiotic arbuscular mycorrhizal (AM) fungi in the rhizosphere(Akiyama et al. 2005). AM form symbiotic associations with the roots of more than 80% of terrestrial plants (Bahadur et al. 2019). Fixed organic carbon from the host plant is exchanged with water and minerals (nitrate and phosphate) absorbed by the fungi through an extensive network of hyphae (Bonfante and Genre 2010). Such a mutual relationship is particularly important for nutrients with low mobility in the soil. More recently, SL were shown to be involved in abiotic stress responses in plants, not only indirectly through AM symbiosis but also as chemical signals directly coordinating plant architecture in response to nutrient availability (Sun et al. 2016; Chai and Schachtman 2022). Plants in nutrient (N or P)-deficient environments significantly promote SL production, resulting in reduced shoot branching, increased primary/adventitious root elongation, along with extensive growth of lateral roots and root hairs for better adaptation to P or N deficiencies (**Figure 4.1**) (Saeed et al. 2017). SL are also involved in crosstalk with other hormones, such as auxin, abscisic acid, cytokinin, and gibberellins. Thus SL are involved in different developmental processes in plants (Bhoi et al. 2021).

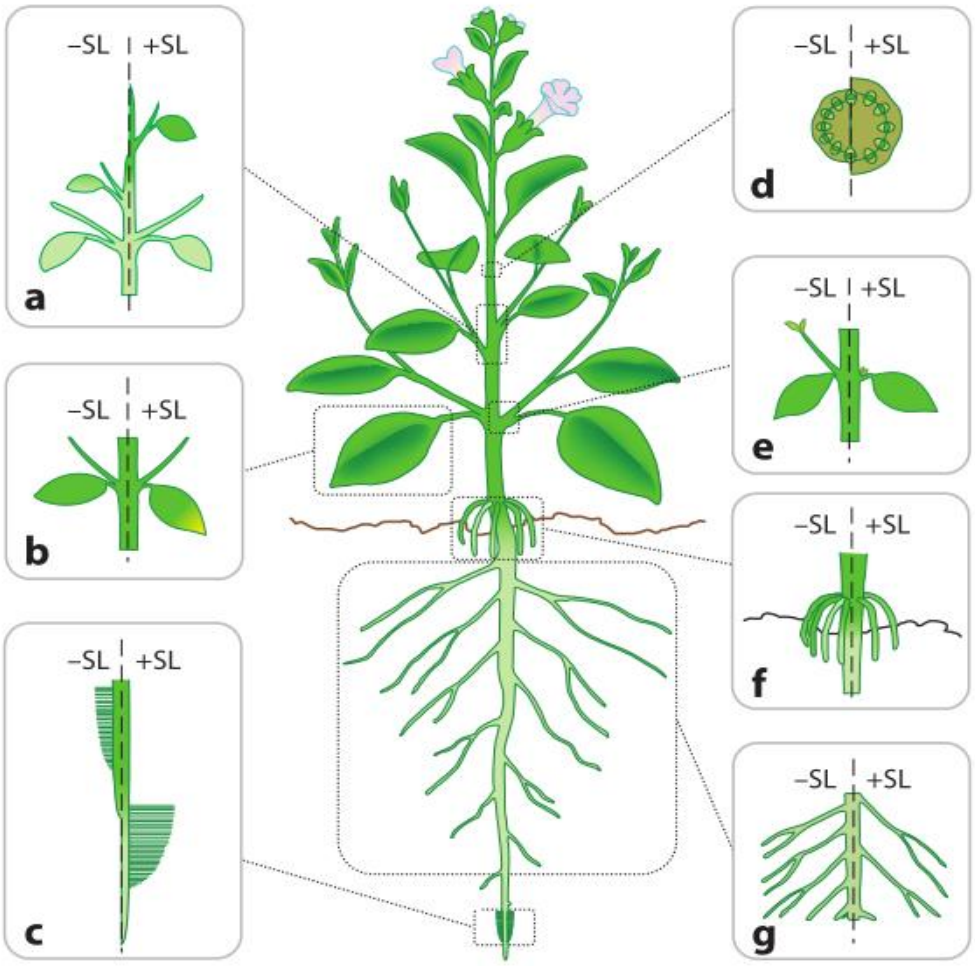


Figure 4.1. Roles of strigolactones in plant development (modified from Al-Babili and Bouwmeester 2015).

4.1.2. Chemistry of strigolactones

In 1995, Butler first described the collective term "strigolactone" to refer to structurally related strigol compounds, containing a common butenolide D-ring, essential for biological activity, linked to a less conserved tricyclic lactone ABC ring by an enol-ether bridge (Butler 1995). Based on the core ABC ring, SL can be classified into canonical forms which contain the ABC ring system as in strigol and non-canonical forms which lack the A, B or C ring, as in carlactone, which has been identified as an endogenous strigolactone precursor in rice (Ito et al. 2022). In addition, the canonical SL can be further divided into strigol and orobanchol types according to the stereochemistry of the C-ring (β- and α-orientations, respectively) (**Figure 4.2**) (Aliche et al. 2020). In general, non-canonical SL are less stable (Yoneyama et al. 2018). Even though canonical SL are chemically unstable and decompose rapidly in the soil, AM fungi and root parasitic weeds utilize this property of SL as cues to identify the presence of host roots (Xie et al. 2010).

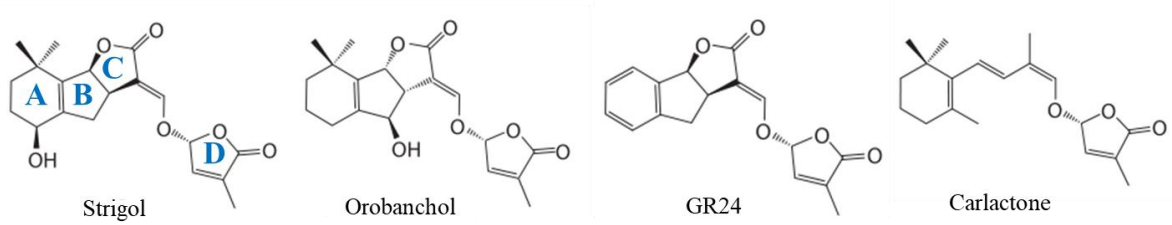


Figure 4.2. Structures of natural and synthetic SL. Canonical SL: Strigol (strigol-type) and Orobanchol (orobanchol-type); non-canonical SL: carlactone (a SL precursor lacking the BC-ring); synthetic SL minic GR24.

4.1.3. Biosynthesis of strigolactones

SL are plant apocarotenoids, a class of naturally occurring organic compounds derived from oxidative cleavage of carotenoids (Al-Babili and Bouwmeester 2015). Although the biosynthesis of SL has not been fully elucidated, many SL biosynthetic enzymes have been identified in Arabidopsis (Sorefan et al. 2003), petunia (Snowden et al. 2005), pea (Hayward et al. 2009), rice (Arite et al. 2007), and maize (Li et al. 2023) (Figure 4.3.). The core pathway of strigolactone biosynthesis starts in the plastids from all-trans-β-carotene which is isomerized from a trans to a cis configuration by β-carotene isomerase DWARF 27 (D27). The conversion of cis-β-carotene to carlactone, the first SL-like carbon skeleton, proceeds via two sequential reactions involving carotenoid cleavage dioxygenase 7 (CCD7) and 8 (CCD8) (Alder et al. 2012; Bruno and Al-Babili 2016). Considering the structural differences between carlactone and strigolactones, CCD7 and CCD8 are not sufficient to produce strigolactones. Genetic evidence indicates that carlactone is modified by diverse enzymes to give rise to various SL derivatives (Seto et al. 2014). For example, in Arabidopsis, carlactone is oxidized by a cytochrome P450, encoded by the More Axillary Growth 1 (MAX1) homolog *At*Max1, to produce carlactoneacid, while in rice, MAX1 homologs *Os*900 and *Os*1400 convert carlactone to 4-deoxyorobanchol (4 DO) and orobanchol, respectively (Ito et al. 2022). Both CCDs have been identified in rice and Arabidopsis by applying synthetic strigolactone analog (GR24) on ccd7 or ccd8 mutant plants. These experiments showed that not only the shoot branching phenotype could be rescued, but also the ability of the root for symbiosis with AM could be restored (Umehara et al. 2008; Beveridge and Kyozuka 2010; Ruyter-Spira et al. 2011). The SL content in roots of mutant plants was lower and the plants had shorter primary roots and a higher density of lateral roots, compared with WT plants (Ruyter-Spira et al. 2011; Arite et al. 2012).

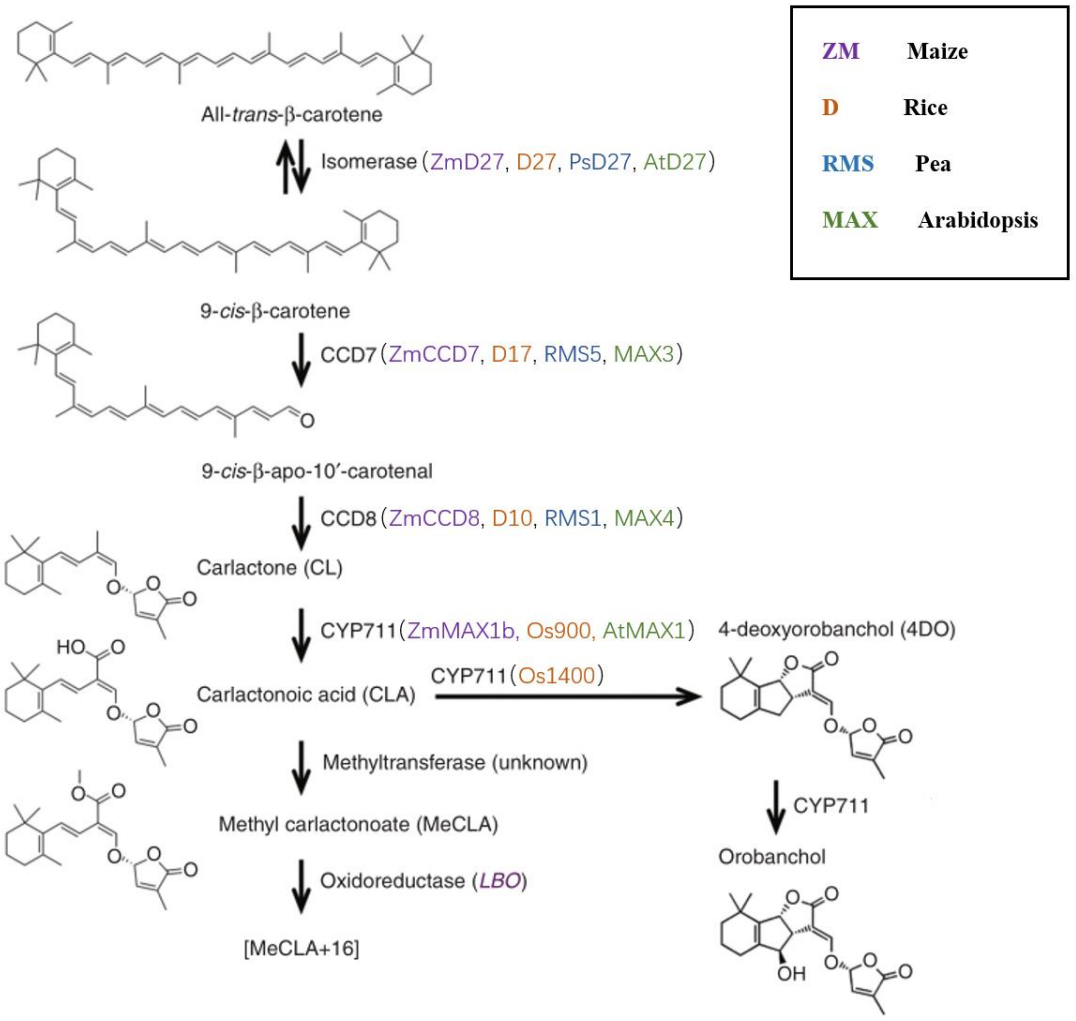


Figure 4.3. Strigolactone biosynthesis.

4.1.4. Modification of maize root architecture by CRISPR/Cas9

Maize (*Zea mays* L.) is a major food security crop. It provides nutrients for humans and animals and is an essential raw material for starch, oil, protein, alcoholic beverages, food sweeteners, and fuel. High grain quality and productivity, require high nitrogen and phosphorus fertilizer inputs. Root system architecture is important for the absorption of nutrients and water, as well as influencing hormone biosynthesis, particularly under abiotic stress conditions (Coudert et al. 2010). The Clustered Regularly Interspaced Short Palindromic Repeats CRISPR/Cas system has emerged as the predominant genome editing tool, in various species, including bacteria, mammals, and plants, because of ease of use, simple cloning procedures, and higher efficiency compared to other genome editing technologies such as ZFNs and TALENs (Bortesi and Fischer 2015; Baysal et al. 2016; Zhu et al. 2017). The CRISPR/Cas system is a microbial adaptive immune system that defends bacteria from repeat viral infections (Rath et al. 2015). Three CRISPR/Cas systems (types I, II, and III) with different mechanisms of action have been identified in

bacteria and archaea. The Type II system is the simplest and has two main components: a CRISPR-associated protein Cas9 and guide RNA (gRNA). The platform allows a gRNA ca: 20 nucleotides long, which is complementary to the target DNA sequence, to direct the Cas9 nuclease with combined helicase and nuclease activity to the target DNA sequence(s) (Zhu et al. 2017). The CRISPR system utilizes a protospacer adjacent motif (PAM), which is a short G-rich oligonucleotide sequence between 3-8 nt long, downstream of the target DNA sequence (Mojica et al. 2009). The *Streptococcus pyogenes* (*Sp*Cas9) is the predominant nuclease in use. *Staphylococcus aureus* Cas9 (*Sa*Cas9) has also been reported in rice and Arabidopsis as an alternative nuclease (Ran et al. 2015; Kaya et al. 2016). The protospacer adjacent motif (PAM) sequence for SaCas9 (5′-NNGRRT-3′) differs from that of SpCas9 (5′-NGG-3′) and this alternative Cas9 nuclease increases selectivity at potential cleavage target sites (Kaya et al. 2016). In addition, the less common *Streptococcus thermophilus* Cas9 (*St*Cas9) and the *Neisseria meningitides* cas9 (*Nm*Cas9) have also been used for genome engineering of human cells (Horvath et al. 2008; Lee et al. 2016).

CRISPR/Cas9-mediated genome editing has been used to modulate different traits related to maize plant architecture, including leaf angle, number of branches and grains, and plant height, among others (Svitashev et al. 2015; Li et al. 2017; Shi et al. 2017; Wang et al. 2019), However, reports focusing on root architecture are rare. Recent studies have shown that knocking out *lrt1* generated maize plants with fewer lateral roots and shorter primary roots compared to wild-type plants (Baer et al. 2023). Additionally, knocking out the transcription factors *SBP20, SBP25,* and *SBP27* which play important roles in regulating many growth and development processes, resulted in an increased density of crown roots. Crown roots are essential for water and nutrient uptake and also enhance the plant's lodging resistance (Wang et al. 2022). These studies highlight the importance of understanding factors that determine maize root architecture and provide a theoretical basis and technical strategy for improving the performance and resilience of the maize root system.

4.2. Aims and objectives

The aim of this chapter was to explore the potential of CRISPR/Cas9 to modulate strigolactone levels in maize. The main objective was to induce targeted mutagenesis in the coding regions and/or corresponding promoter regions of *ZmCCD7* and *ZmCCD8* to generate maize mutants with altered strigolactone content and compositions.

4.3. Materials and methods

4.3.1. Cloning of *ZmCCD7* and *ZmCCD8* fragments and selection of CRISPR/Cas9 target sites

The MaizeGDB [Maize Genetics and Genomic Database (http://www.maizegdb.org)] and the National Center for Biotechnology Information (NCBI) GenBank databases (https://www.ncbi.nlm.nih.gov/genbank) were searched for sequences homologous to *ZmCCD7* (LOC100502521,chr2) and *ZmCCD8* (LOC100502522,chr3) in inbred maize line B73 RefGen_v4 (MGSC) using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Based on the B73 maize homologous genomic DNA sequence, I designed cloning primers to amplify the *ZmCCD7* and *ZmCCD8* fragments using genomic DNA from maize leaves **(Table 4.1)**. Genomic DNA extraction and PCR amplification were performed as described in section **4.3.4**. Purified PCR products (Geneclean II kit, MP Biomedicals, Solon, OH, USA) were cloned into the pGEM-T Easy vector (Promega) and sequenced. CHOPCHOP Tool (http://chopchop.cbu.uib.no) and CRISPR RGEN Tools (http://www.rgenome.net) were used to design gRNAs. Activity scores for gRNAs were predicted using sgRNA Scorer 2.0 (https://crispr.med.harvard.edu).

Table 4.1. Primers used for cloning the M37W *ZmCCD7* and *ZmCCD8*

Genes/Promoter	**Primer sequence (**5′-3′**)**
*ZmCCD7*_promoter_F	GAATAAAATGCATGTGTGTGTAGGGCGTGTAA
*ZmCCD7*_promoter_R	AGACTTGCTAACTGTTAGAGTGGACACTAACC
*ZmCCD7*_Exon1_F	ACCGAGACCGGCACAACACAAACAGAAAC
*ZmCCD7*_Exon1_R	AGAGGTGGTTGACTGCTAGTGGTGGCAGTGACG
*ZmCCD8*_promoter_F	AAGATTATAGACATACATGGGAGGAGCACTTTG
*ZmCCD8*_promoter_R	AAATACGCACAACGACGAAGCCATAGTG
*ZmCCD8*_Exon1_F	ACTACTGCCCTCTCCTTTCCAG
*ZmCCD8*_Exon1_R	CGATCAACTCTTCCCGAAAACA

ZmCCD7, maize (*Zea mays*) carotenoid cleavage dioxygenase 7. *ZmCCD8*, maize carotenoid cleavage dioxygenase 8.

4.3.2. Vector construction

Three expression vectors were designed with different Cas9 gene coding sequences under the control of different promoters: (1) the SaCas9 expression vector **(**p*Zm*Ubi1-SaCas9, **Figure 4.4A)** was constructed as follows: the coding sequence of SaCas9 was codon-optimized for maize using the tool offered by Integrated DNA Technologies Inc. (IDT: https://eu.idtdna.com/CodonOpt), the synthesized *SaCas9* fragment was purchased from IDT (Coralville, Iowa, USA) and then cloned into the pAL76 (Christensen and Quail 1996) containing the maize ubiquitin-1 promoter and first intron (*ZmUbi1*+1st); (2) the SpCas9 expression vector driven by the *2X35S* promoter **(**pJIT163-2NLSCas9, **Figure 4.4B)** was obtained from Dr. C. Gao (Chinese Academy of Sciences, Beijing, China). The vector was already codon-optimized for rice (Shan et al. 2013); (3) the SpCas9 expression vector driven by the *ZmUbi1* promoter **(**pZmUbi1_SpCas9, **Figure 4.4C)** was constructed as follows: the *2X35S* promoter in the original pJIT163-2NLSCas9 was removed by KpnI and BamHI sites and replaced with p*ZmUbi1*+1st from pAL76. The Cas9 gene coding sequences in all constructs included nuclear localization signals (NLS) at both ends.

Three gRNA expression vectors were designed with different gRNA scaffolds. (1) the single gRNA (sgRNA) expression vectors **(**sgRNA_Sa and sgRNA_Sp, **Figure 4.4 D, E)** were constructed as follows: SaCas9 gRNA scaffold or SpCas9 gRNA scaffold with the maize *U3* promoter, respectively, were synthesized (Sigma-Aldrich, St Louis, MO, USA) and then cloned into pUC57 to generate gRNA intermedia vectors pZmU3_Sa and pZmU3_Sp. Each target site was synthesized as a pair of reverse complementary oligonucleotides and sub-cloned into pZmU3_Sa or pZmU3_Sp at the BbsI site; (2) the multiplexed gRNA expression vector **(**4XgRNAs_Sp, **Figure 4.4F)** used to deliver multiple gRNAs simultaneously, was constructed as follows: a synthetic fragment containing four independent gRNAs (T5, T6, T11, T12), SpCas9 scaffolds and tRNA linkers (GeneArt, Thermo Fisher Scientific, USA), were cloned into gRNA intermediate vector pZmU3_SpCas9 at the EcoRI and BamHI sites (Baer et al. 2023). The selectable marker gene controlled by the *ZmUbi1* promoter **(**pTRAux-bar, **Figure 4.4G)** was described earlier (Naqvi et al. 2009) (Zhu et al. 2008). Transformation construct fidelity was verified by sequencing all junctions in the constructs (Stab Vida, Caparica, Portugal).

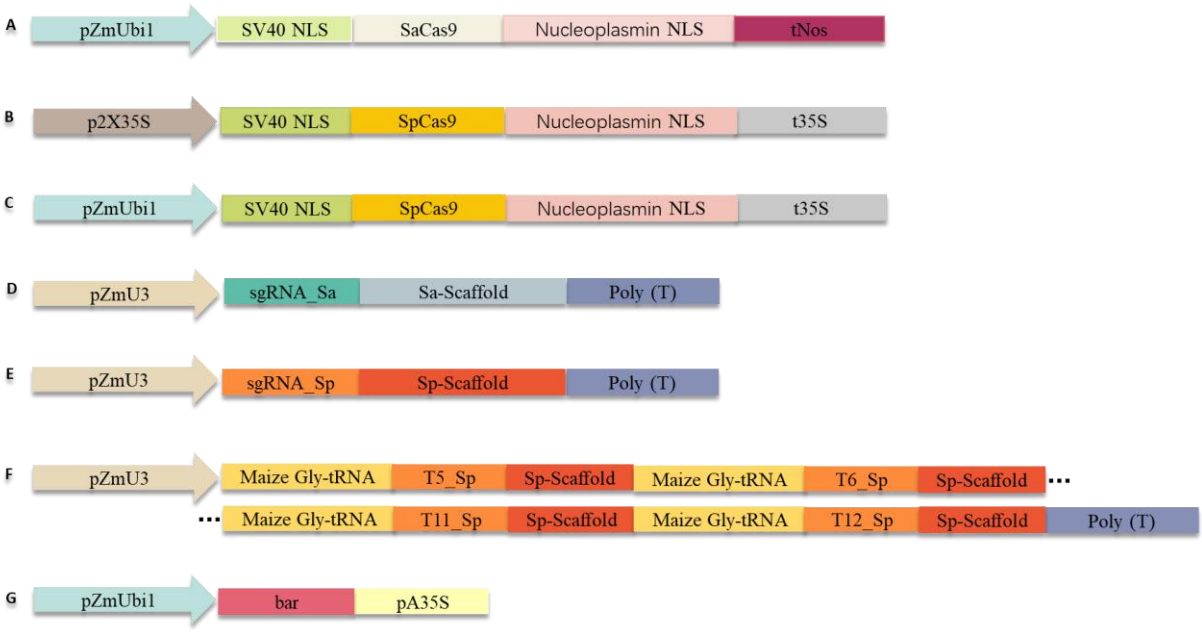


Figure 4.4. Constructs for maize transformation. (A) pZmUbi1_SaCas9. (B) pJIT163-2NLSCas9. (C) pZmUbi1_SpCas9. (D) sgRNA_Sa. (E) sgRNA_Sp. (F) 4XgRNAs_Sp. (G) pTRAux-bar.

4.3.3. Maize transformation, selection, and regeneration of transgenic plants

M37W immature zygotic embryos (*Zea mays* L., cv. M37W, IZEs) were excised 10-14 days after pollination and cultured on N6 (Duchefa Biochemie BV, Haarlem, Netherlands) medium in the dark. After 5 days, type I callus were transferred to osmoticum (Osm) medium for 3 hours prior to bombardment. Embryos were bombarded with 10 mg of DNA-coated gold particles (Christou et al. 1991). The Cas9 vector, sgRNA_*ZmCCD7* vector, sgRNA_*ZmCCD8* vector and the selectable marker vector were used at a 3:3:3:1 DNA ratio (calculated by molecular weight) **(Table 4.2)**. The selectable marker gene *bar* confers resistance to phosphinothricin (PPT), which was added to the culture medium to facilitate selection. IZEs were transferred to N6 medium 16 hours after bombardment and incubated for three days in the dark before being transferred to selection medium containing 3 mg/L PPT and sub-cultured after two weeks as described previously (Ramessar et al. 2008). The PPT-resistant embryogenic type I callus were transferred to regeneration medium (Reg-I) containing 3 mg/L PPT for shoot formation in the light, and two weeks later transferred to regeneration medium (Reg-II) with 1 mg/L PPT. After 2 to 4 weeks, well-developed plantlets were transferred to the rooting medium (R-M) without PPT to develop roots. Multiple plants with well-formed shoots (>3 cm) and roots (>2 cm) were regenerated from each herbicide-resistant callus and hardened off in the soil. All experiments were performed at 25 °C. Plants were either self- or cross-pollinated depending on pollen availability. Transgenic plants were grown in the growth chamber

and greenhouse at 28/20°C day/night temperature with a 10-h photoperiod and 60–90% relative humidity. The media composition is listed in **Table 4.3**.

Table 4.2. Combinations of transgene constructs for maize transformation.

Experiment	gRNA_*ZmCCD7*	gRNA_*ZmCCD8*	Cas9
I	T1_Sa	T7_Sa	pZmUbi1_SaCas9
II	T2_Sa	T8_Sa	pZmUbi1_SaCas9
III	T3_Sa	T9_Sa	pZmUbi1_SaCas9
IV	T4_Sp	T10_Sp	pJIT163-2NLSCas9
V	T5_Sp	T11_Sp	pJIT163-2NLSCas9
VI	T6_Sp	T12_Sp	pJIT163-2NLSCas9
VII	4XgRNAs_Sp		pZmUbi1_SpCas9

I, II, III: two sgRNAs (one targeting *ZmCCD7*, the other targeting *ZmCCD8*) transformed together with pZmUbi1_SaCas9. IV, V, VI: two sgRNAs (one targeting *ZmCCD7*, the other targeting *ZmCCD8*) transformed together with pJIT163-2NLSCas9. VII: four gRNAs (two targeting *ZmCCD7*, the other two targeting *ZmCCD8*) transformed together with pZmUbi1_SpCas9.

Table 4.3. Media composition (for 1 L)

Item	N6	Osm	S-M	Reg-I	Reg-II	Reg-III	R-M
N6 vitamins	4 g	4 g	4 g	/	/	/	/
MS	/	/	/	4.33 g	4.33 g	4.33 g	4.33 g
Casein hydrolysate	0.1 g	0.1 g	0.1 g	/	/	/	/
L-proline	2.88 g	2.88 g	2.88 g	/	/	/	/
Sucrose	20 g	20 g	20 g	30 g	30 g	30 g	10 g
D-Mannitol	/	36.4 g	/	/	/	/	/
D-sorbitol	/	36.4 g	/	/	/	/	/
Myo-inositol	0.1 g	0.1 g	0.1 g	/	/	/	/
2.4 D (5 mg/ml)	200 ul	200 ul	200 ul	50 ul	50 ul	50 ul	/
$CuSO_4$ 5H2O (0.5 mg/ml)	50 ul	50 ul	50 ul	/	/	/	/
Na_2MoO_4 $2H_2O$ (1 mg/ml)	250 ul	250ul	250 ul	/	/	/	/
Adjust pH to 5.8 with KOH							
Gelrite	4 g	4 g	4 g	4 g	4 g	4 g	4 g
Autoclave at 121 °C for 16 min							
BAP (1 mg/ml)	/	/	/	10 ml	10 ml	10 ml	/
Vitamin 1000X	/	/	/	1 ml	1 ml	1 ml	1 ml
PPT (10 mg/ml)	/	/	300 ul	300 ul	100 ul	/	/
$AgNO_3$ (10 mg/ml)	1 ml	1 ml	1 ml	85 ul	85 ul	85 ul	85 ul
IBA (1 mg/ml)	/	/	/	/	/	/	500 ul

MS: Murashige & Skoog basal salt mixture, BAP: 6-Benzylaminopurine.

4.3.4. DNA analysis

Genomic DNA was isolated from frozen leaves (0.1–0.2 g) finely ground in liquid nitrogen, by phenol extraction and ethanol precipitation (Bassie et al. 2008). Three pairs of primers were used to identify transgenic plants containing the transgene fragments (SaCas9, SpCas9, sgRNA) by PCR (primers are listed in **Table 4.4)**. I used 100 ng genomic DNA and GoTaq DNA Polymerase with the reagents and reaction conditions recommended by the manufacturer (Promega, Fitchburg, WI, USA).

Table 4.4. Primers used for PCR analysis and RT-PCR analysis.

Genes	**Primers used for PCR analysis**
SaCas9_F	5′-ATGGCGGCTTTACCTCCTTC-3′
SaCas9_R	5′-GGGTAGTCGTCCGTGATGTC-3′
SpCas9_F	5′-TCCCACTATCCTTCGCAAGACCC-3′
SpCas9_R	5′-GGTCCGCCTTATCAGTCGAGTCC-3′
sgRNA_F	5′-TGTGAGTTGTGACCGATGCT-3′
sgRNA_R	5′-GACTCGGTGCCACTTTTTCAA-3′
Genes	**Primers used for RT PCR analysis**
SaCas9_RT_F	5′-ATTCTCACAATCTATCAATCGTCG-3′
SaCas9_RT_R	5′-TCGCTATTTGGTTGTCATTCGT-3′
SpCas9_RT_F	5′-TCCCACTATCCTTCGCAAGACCC-3′
SpCas9_RT_R	5′-GGTCCGCCTTATCAGTCGAGTCC-3′

SaCas9, *Staphylococcus aureus* Cas9 gene. SpCas9, *Streptococcus pyogenes* Cas9 gene.

4.3.5. Gene expression analysis

Primers are listed in **Table 4.4**. Gene expression analysis as described in **section 3.3.4**.

4.3.6. Protein extraction and immunoblot analysis

Leaf tissue ca. 0.1–0.2 g was ground in liquid nitrogen, thawed in an equal volume of extraction buffer (20 mM Tris-HCl pH 7.5, 5 mM ethylenediaminetetraacetic acid (EDTA) 0.1% Tween-20, 0.1% sodium dodecylsulfate (SDS), 2 mM phenylmethanesulfonylfluoride (PMSF) and vortexed for 1 h at 4 °C. Cell debris was removed by centrifugation at 15,300 g for 20 min at 4 °C, and the supernatant was collected and stored at –80 °C. The protein concentration in the supernatants was determined using the Bradford method (AppliChem, Darmstadt, Germany). I fractionated

80 µg of total maize protein by denaturing SDS-PAGE in polyacrylamide gels containing 10% SDS at 200 V for 60 min and then electro-transferred the protein to an Immobilon FL polyvinylidene difluoride (PVDF) membrane (Merck, Darmstadt, Germany) using a semidry transfer apparatus (Bio-Rad, Hercules, CA, USA) at 20 V for 45 min. The membrane was immersed in 5% non-fat milk in Tris-buffered saline with Tween-20 (TBST) solution (0.2 M Tris-HCl pH 7.6, 1.37 M NaCl, 0.1% Tween-20) for 1 h at room temperature. Membranes were incubated with anti-Cas9 antibody in 5% non-fat milk in TBST overnight at 4 °C. Anti-SaCas9 polyclonal antibody (AB356480, Sigma-Aldrich), diluted 1:1000 and anti-SpCas9 polyclonal antibody (ABS2202, Sigma-Aldrich) diluted 1:500. After three rinses in TBST for 10 min each, the membranes were subsequently incubated with a goat anti-rabbit secondary antibody conjugated to alkaline phosphatase (Sigma-Aldrich) diluted 1:5000 in 2% non-fat milk in TBST for 1h at room temperature. Membranes were washed three times for 10 min each with TBST buffer and the signals were detected using SIGMAFAST BCIP/NBT tablets (Sigma-Aldrich).

4.3.7. Analysis of mutations

Target regions were amplified using the genomic DNA from transgenic plants (section 4.1.1) using sequencing primers for preliminary evaluation of the amplicons by Sanger sequencing **(Table 4.5)** (Stab Vida, Caparica, Portugal). Purified PCR products (Geneclean II kit, MP Biomedicals, Solon, OH, USA) from potential mutant lines that showed multiple peaks in the chromatograms, were cloned into the pGEM-T Easy vector (Promega) and 25 clones from each line were sequenced to characterize the structure of each mutation.

Table 4.5. Primers used for sequence analysis.

Genes/Promoter	Primer sequence
*ZmCCD7*_S_promoter_F	5′-GAATAAAATGCATGTGTGTGTAGGGCGTGTAA-3′
*ZmCCD7*_S_promoter_R	5′-AGACTTGCTAACTGTTAGAGTGGACACACTAACC-3′
*ZmCCD8*_S_promoter_F	5′-AAGCATGTAACTCATGAGTAGAAGACAAGTCGAG-3′
*ZmCCD8*_S_promoter_R	5′-GCTGCAAATACGCACAACGACGAAGCCATAGTG-3′
*ZmCCD7*_S_Exon1_F	5′-ACCGAGACCGGCACAACACAAACAGAAAC-3′
*ZmCCD7*_S_Exon1_R	5′-AGAGGTGGTTGACTGCTAGTGGTGGCAGTGACG-3′
*ZmCCD8*_S_Exon1_F	5′-ACTACTGCCCTCTCCTTTCCAGGCGACTTTGCA-3′
*ZmCCD8*_S_Exon1_R	5′-CTTGTACGCCTCCGACTCGATCTGCCGGTGC-3′

4.4. Results

4.4.1. Cloning of the *ZmCCD7* and *ZmCCD8* promoters and gene sequences

I amplified and cloned 1 kb upstream and downstream of the start codon of *ZmCCD7* and *ZmCCD8* genes of the maize varieties M37W and B73 to determine sequence differences in these regions to design the most efficient gRNAs and avoid mismatches. Sequencing and alignment analysis showed that there are 1, 2 and 6 bp deletions and 1 bp insertion in the promoter region (1 kb upstream of the start codon) of the M37W *ZmCCD7* sequence compared to that of B73, in addition to a 3 bp insertion in the first exon of *ZmCCD7* in M37W **(Table 4.5)**. I also demonstrated that there is a 3 bp deletion and a 4 bp insertion in the promoter region (1 kb upstream of the start codon) of the M37W *ZmCCD8* sequence compared to that of B73. The sequence of the first exon of *ZmCCD8* was identical between the two genotypes **Figure 4.5.**

(A)

```
>ZmCCD7_Promoter
B73_Promoter      ACTAGACTCGAACAAAATGCATGTAGTTACAGTCACGAGTAACCTAGCCGTTGGCTGAAT  464
M37W_Promoter     ACTAGACTCGAACAAAATGCATGTAGTTACAGTCACGAGTAACCTAGCCGTTGGCTGAAT  418
B73_Promoter      AAAATGCATGTGTGTGTAGGGCGTGTAATTTGCTAAGCAGCAACATATCTTACTGCATGT  524
M37W_Promoter     AAAATGCATGTGTGTGTAGGGCGTGTAATTTGCTAAGCAGCAAC-AATCTTACTGCATGT  477
B73_Promoter      GCTTGATCACGTTTTTCGTTCCGAAACCCAAAACCAACGCGCCCCTCTCGTGCGGTACCT  584
M37W_Promoter     GCTTGGTCACGTTTTTCGTTCCGAAACCCAAAACCAACGCGCCCCTCTCGTGCGGTACCT  537
B73_Promoter      TCAATGTAATGACAATTTGGTTGAGCAGCGGCCATGGGCAAATCAGTGGCTCTCGAATAG  644
M37W_Promoter     CCAATGTAATGACAATATGGTTGAGCAGCGGCCATGGGCAAATCAGTGGCTCTCGATTAG  597
B73_Promoter      CGACGCGGTCGGCGACAGCGAGCGACGAGCGCTTCCTGGGGAGAGTCTTGTCGCAGCCTT  704
M37W_Promoter     CGACGCGGTCGGCGACAGCGAGCGACGAGCGCTTCCTGGGGAGAGTCTTGTCGCAGCCTC  657
                                                  T1(+)
B73_Promoter      ATTCGAGATTTGTATAGGCACAAGCACAACTAGCTAATCCAGTTACCCCTTTTTTTAATC  764
M37W_Promoter     ATTCGAGATTTGTATAGGC------ACAACTAGCTAATCCAGTTACCCCTTTTTTTAATC  711
B73_Promoter      GAGAAAGGAAGCGCCGCAGAAGCATCACGATTATCATGCCACAACAAACAAGACACCCTG  824
M37W_Promoter     GAGAAAGGAAGCGCCGCAGAAGCATCAAGATTATCATGCCACAACAAACAAGCCACCCTG  771
                                                        T2(+)
B73_Promoter      GGATTCCACACCCAAAATGCCACCGATACCCTGGCCAAAACCCTCCAAAAAGCATCCCCC  884
M37W_Promoter     GGATTCCACACCCAAAATGCCACCGATTCCATGGCCAAAACGCTCCAAAAAGCA-CCCCC  830
B73_Promoter      CGGAGATACCATCCACACCACCTAGCTATTCTTCTTCCACGCAACAAGGTGTACCGTTAA  944
M37W_Promoter     CGGAGATACCATCCACATCACCTAGCTATTCTTCTTCCACGCAACAAGGTGTACCGTTAT  890
                                                     T3(-)
B73_Promoter      ATTTGATTAAACGCTCCCATCATGAAATCCAACCAGTTCTGTGTAGATCGATTCAAATTG  1004
M37W_Promoter     ATTTGATTAAACGCTCCCATCATGAAATCCAACCAGTTCTGTGTAGATCGATTCAAATTG  950
B73_Promoter      ACACCATATACTAGAAATCAAGTTTTTT-TTTCTTTTTTACTGCTTATCTCATGTCACAA  1063
M37W_Promoter     ACACCATATGCTAGAAATCAAGTTTTCTCTCTCTTTTTTACTCATTATCTAATGTCATAA  1010
B73_Promoter      GAAAGCACTCTTAGTTTAATGATTTTTGTAAGACGGACTAATTTTTAGCTCTGAGATTTA  1123
M37W_Promoter     GAAAGCACTCTTAGTTTAATGATTTTTGTAAGACGAACTAATTTTTAGCTCTGAGATCCA  1070
B73_Promoter      GACATATCCTCTCGATGTTGAAGTATATAAGTGGATTGACTATATTTTTCCATTAACTTA  1183
M37W_Promoter     GACATA--TTCTCGGTGTTGAAGTATATAAGTGGATTAACTATCTTTTTCTATTAGCTTA  1128
B73_Promoter      ATTTTTTTGGGTTAAACTGATTGGTGCGTCCACTCTAATATGGTATCAAAGTTAGATGTCT  1243
M37W_Promoter     ATCTTTTTGGGTTGAACTGATTAGTGCGTCCACTCTAATATGGTATCAAAGC-CGATGTCT  1187
B73_Promoter      CGAGTTCAAATCTTGATAGAGATTTTATTTATGCATCCACCAATTTATTCCCACATTTGC  1303
M37W_Promoter     CGAATTCGAATTCTGGCAGAGATTTTATTTATGCATCCACCAATTTATTCCCATGTTTGC  1247
B73_Promoter      GCCTTTCTCTCTCTGACTGCACGTGAGTGTTGAAATATATAAGTGGATTGACCTTCTTCA  1363
M37W_Promoter     GCTTTTCTCACTCTGACTGCACGTGAGTGTTGAAGTATATAAGTGGATTGACCTTCTTCT  1307
B73_Promoter      AACAACTTAATCTTTTTGGTTGAACTGGTTAGTGTGTCCACTCTAACAATTAGCAAGTCT  1423
M37W_Promoter     ATCAATTTAATCTTTTAGGTTAAACTGGTTAGTGTGTCCACTCTAACAGTTAGCAAGTCT  1367
```

(B)

>ZmCCD7_Exon1

```
B73_Exon1      ACCGAGACCGGCACAACACAAACAGAAACATGTATCAACACACCATGGCCGCCATGCACG   60
M37W_Exon1     ACCGAGACCGGCACAACACAAACAGAAACATGTATCAACACACCATGCCCGCCATGCACG   60
B73_Exon1      CCATCGTGCACCACCGCGCACCGCCGCCTGGCCGCTGCTGTCGCGGCCACGGGCGCAGCG  120
M37W_Exon1     CCGTCGTGCACCACCGCGCGCCGCCGCCTGGCCGCTGCTGTCGCGGCCACGGGCGCAGCG  120
B73_Exon1      TCGTCGTCCGCGCGTCGGCCGCCACCGTCACCACCAGCATCCCCGGGTCCGCGGCGACAG  180
M37W_Exon1     TCGTCGTCCGCGCGTCGGCCGCCACCGTCACCACCAGCATCCCCGGGTCCGCGGCGACAG  180
B73_Exon1      TGCCGGACTCGCCGTCCGCGGCGTTCTGGGACTACAACCTCCTATTCCGGTCGCAGCGCG  240
M37W_Exon1     TGCCGGACTCGCCGTCCGCGGCGTTCTGGGACTACAACCTCCTATTCCGGTCGCAGCGCG  240
                                  T4(-)
B73_Exon1      CCGAGTCCCCCGACCCCGTGGTGCTCCGCGTCACGGAGGGCGCGATCCCGCCGGACTTCC  300
M37W_Exon1     CCGAGTCCCCCGACCCCGTCTCGCTCCGCGTCACGGAGGGCGCGATCCCGCCGGACTTCC  300
B73_Exon1      CGGCGGGCACCTACTACCTCGCCGGTCCCGGGATGTTCACCGACGACCACGGGTCCACGG  360
M37W_Exon1     CGGCGGGCACCTACTACCTCGCCGGTCCCGGGATGTTCACCGACGACCACGGGTCCACGG  360
                                        T5(-)
B73_Exon1      TCCACCCGCTCGACGGCCACGGCTACCTCCGCTCGTTCCGCTTCGACGCCAGCGGCGCGG  420
M37W_Exon1     TCCACCCGCTCGACGGCCACGGCTACCTCCGCTCGTTCCGCTTCGACGCCAGCGGCGCGG  420
B73_Exon1      CGCACTACTCCGCGCGGTACGTGGAGACGGCGGCGAAGCGGGAGGAGCACGACGCGGGCG  480
M37W_Exon1     CGCACTACTCCGCGCGGTACGTGGAGACGGCGGCGAAGCGGGAGGAGCACGACGCGGGCG  480
B73_Exon1      GCGCGTCGTGGCGGTTCACGCACCGGGGCCCCTTCTCGGTGCTGCAGGGTGGGAGCCGGG  540
M37W_Exon1     GCGCGTCGTGGCGGTTCACGCACCGGGGCCCCTTCTCGGTGCTGCAGGGCGGGAGCCGGG  540
B73_Exon1      TGGGCAACGTGAAGGTGATGAAGAACGTGGCCAACACCAGCGTGCTGCGCTGGGGCGGCC  600
M37W_Exon1     TGGGCAACGTGAAGGTGATGAAGAACGTGGCCAACACCAGCGTGCTGCGCTGGGGCGGCC  600
B73_Exon1      GCGTGCTCTGCCTCTGGGAAGGCGGGGAGCCGTACGAGCTGGACCCGCGGACGCTGGAGA  660
M37W_Exon1     GCGTGCTCTGCCTCTGGGAAGGCGGGGAGCCGTACGAGCTGGACCCGCGGACGCTGGAGA  660
                                              T6(+)
B73_Exon1      CCATCGGCCCGTTCGACATCCTCGGCAGCCTCTGCGGCGGTACCGACG---AAGTGGCAC  717
M37W_Exon1     CCATCGGCCCGTTCGACATCCTCGGCAGCCTCGGCGGCGGTACCGACGACGAAGCGGCAC  720
B73_Exon1      GAGACGCCAGCGGCGAGGCTGCGCATCACGGGCGCCGGCAGCCGTGGCTGCAGGAGGCAG  777
M37W_Exon1     GAGACGCCAGCAGCGAGGCTGCGCGTCACGGGCGGCGGCAGCCGTGGCTGCAGGAGGCAG  780
B73_Exon1      GGATCGACGTGGCCGCGCGCCTGCTGCGACCGGTCCTCAGTGGTCTGTACCGTCACTGCC  837
M37W_Exon1     GGATCGACGTGGCCGCGCGCCTGCTGCGACCGGTCCTCAGTGGTCTGTACCGTCACTGCC  840
B73_Exon1      ACCACTAGCAGTCAACCACCTCT        860
M37W_Exon1     ACCACTAGCAGTCAACCACCTCT        863
```

(C)

```
>CCD8_Promoter
B73_Promoter        AAGATTATAGACATACATGGGAGGAGCACTTTGTTGTGGTTGTCGTTGTAATGGAAGGAG          60
M37W_Promoter       AAGATTATAGACATACATGGGAGGAGCACTTTGTTGTGGTTGTCGTTGTAATGGAAAGAG          60
B73_Promoter        ACCTTAGTGATGCTCAGGCAGCAAACGCTCACCACCAGGCATTAATCGTCCCTAGACCCC         120
M37W_Promoter       ACCTTAGTGATGCTCAGGCAGCAAACGCTCACCACCAGGCATTAATCGTCCCTAGACCCC         120
B73_Promoter        TGCATCCGTTGGGTCCCCAGCTCCCCACGGAAGTCGCACGCCTCCCTGCTCCTTGACATC         180
M37W_Promoter       TGCATCCGTTGCGTCCCCAGCTCCCCACGGAAGTCGCGCGCCTCCCTGCTCCCTGACATT         180
                          T7 (-)
B73_Promoter        GTAGGCCCCTGACTTCTTCATGGGGTGTGGCCTCATGTTCCTGGCTCCCTGGTTATCGTG         240
M37W_Promoter       GTAGGCCCTTGGCTTCTTCATGGGGTGTGGCCTCATGTTCCTGGCTCCCTGGTTATCGTG         240
                          T8 (+)
B73_Promoter        CGATCCACATCAACGACCTAGGTACACGGCCTCACGGGATGCACTATCATCGTCTACCGA         300
M37W_Promoter       CGATCCACATCGACGACCTAGGTACACGGCCTCACGGGATGCACTATCGTCATCTACCGA         300
B73_Promoter        GGATTCAAAGAAGAGATAGCTATTGTTGCGGTCCTCAGTCTACGTAAGTGCGGAGCACGC         360
M37W_Promoter       GGATTCGATGAAGAGATAGCTATGGTTGCGGTCCTCGGTCTACGGAAGTGCGGAGCACGC         360
B73_Promoter        TAGAGGGAGCGCGGGGTGGCGATGGGAAGGGAAGTGGATCGATATTAGAATAGGATTAGA         420
M37W_Promoter       TAGAGGGAGGCGGGGGTGGCGATGGGAAGGGAAGTGGATCGGTATTAGAATAGGATTGGA         420
                                                          T9 (+)
B73_Promoter        GTGTGGAGTACTTGCGCCTATGTGGTGCGGCTAGGGTTAGGAGCGGGAGGCCAAAGGCGG         480
M37W_Promoter       GTGTGGAGTACTTGCGCCTATGTGGTGCGGATAGGGTTAGGAGCGGAAGGCCAAAGACGG         480
B73_Promoter        AGTTCTATGCACTATGCAGCCTCGCAGCAGCTTGGGCTTCAATGCTTGTTGTTCGGTAAC         540
M37W_Promoter       AGGTCTATGCACTATGCAGCCTCGCACCAGCTTGGGCTTCAATGCTTGTTGTTCGGTAAC         540
B73_Promoter        TTGAGTGTTCGGTTCCGTTTTAACTAGAAACTGAAGCTGTAACCGAACATCGAATTTTGC         600
M37W_Promoter       TTGAGAGTTTGGCTCCGTTTTAACTAGAAACTGAAGCTGTATCCGAACATCGAATTTTGC         600
B73_Promoter        GGATTTACGAACCGAAATCGAACCAGAACTCTAAAACCCCAAACTTTCGGTTCAGTTTGG         660
M37W_Promoter       GGATTTACAAACCGAAATCGAACCGGTACTTTAAAAACCCAGACTTTCGGTTCAGTTTGG         660
B73_Promoter        TTCTGTTTTTGGTTTTGGGTGAGAATATGCACAGCCCTAGGCTAGATCATCGGTTCGCCG         720
M37W_Promoter       TTCTGTTTTTGGTTTTCGGTGAGAATATGCACAGCCCTAGGCTAGATCATCGGTTCGCCG         720
B73_Promoter        GGCCAGTGCCCAACTGCCCACCAACCGGAGGAAACCAGGGCAGCGACAGCCAGCAACGTG         780
M37W_Promoter       GGCCAGTGCCCAACTGCCCACCAACCGGAGGAAACCAGGGCAGCGACAGCCAGCAACGTG         780
B73_Promoter        ACAACGTAATGCAAACAAAGGGAGCGCGTCATTTTGCTCAAAGAATTTGCCCAAGGAAAG         840
M37W_Promoter       ACAACGTAATGCAAACAAAGGGAGCGCGTCATTTTGCTCAGAGAATTTGCCCAAGGAAAG         840
B73_Promoter        ATTCCACCCCGTTGCCTCCCTACGCCGCGGCTATATATATGGCCTTGACCAGGGCACGAA         900
M37W_Promoter       ATTCCACCCCGTTGCCTCCCTACGCCGCGGCTATATATATGGCCTTGACCCGGGCACGAA         900
B73_Promoter        CTCGTCCCAAACCAAGCCAGATCATACTAGTACACACCACACTCCCCCGGCTAGCACCGT         960
M37W_Promoter       CTCGTCCCAAACCAAGCCAGATCATACCAGTACACACCACACTCCCCCGGCTAGCACCGT         960
B73_Promoter        GCTCTCTAGCTAGCTAGCTAAGCTACTACTGCCCTCTCCTTTCCAGGCGACTTTGCAACC        1020
M37W_Promoter       GCTCTCT--ATAGCTAGCTAAGCTACTACTGCCCTCTCCTTTCCAGGCGACTCTGCAACC        1018
B73_Promoter        CTAGCAAG----AGCATTTTCGCCATGTCTCCCACTATGGCTTCGTCGTTGTGCGTATTT        1076
M37W_Promoter       CTAGCTAGCAAGAGCATTTTCGCCATGTCTCCCACTATGGCTTCGTCGTTGTGCGTATTT        1078
B73_Promoter        GCAGC       1081
M37W_Promoter       GCAGC       1083
```

(D)

```
>ZmCCD8_Exon1
M37W_Exon1   ACTACTGCCCTCTCCTTTCCAGGCGACTTTGCAACCCTAGCAAGAGCATTTTCGCCATGT      60
M37W_Exon1   CTCCCACTATGGCTTCGTCGTTGTGCGTATTTGCAGCGATGTCTGGCGCCAGCGGCAGGC     120
                                                  T10(+)
M37W_Exon1   CGTCGGCCGGTGGCTCGGCGGTACCGGGCCGTCTGTCCAGCAGCACACAGGGGGGCAAGG     180
                                                        T11(+)
M37W_Exon1   GAAAGCGGGCCGTGGTGCAGCCGCTCGCGGCTAGCGTGGTGACGGAGACGCCAGCGCCGG     240
M37W_Exon1   CCGTAGCTCCGGCTCGGCCCGTCGTCGACGCCCCGCCGCGCCGCCGTGGGGGCCGCGGCA     300
M37W_Exon1   CCGTCGAGCACGCGGCGTGGAAGAGCGTCCGGCAGGAGAGGTGGGAGGGGGCGCTGGAGC     360
                                    T12(+)
M37W_Exon1   TGGAGGGAGAGCTGCCGCTCTGGCTGGTGGGTACCACTATTCATCTGGCTATTGGTTTGA     420
M37W_Exon1   TCATCGATCTCTGCTGCCTTTTTGTTTTCGGGAAGAGTTGATCG          480
```

Figure 4.5. Alignment of ZmCCD7/ZmCCD8 genomic DNA sequences of B73 and M37W. **(A)** Sequence of *ZmCCD7* promoter 1kb upstream of the start codon. **(B)** Sequence of *ZmCCD7* Exon1 (first exon). **(C)** Sequence of *ZmCCD8* promoter 1kb upstream of the start codon. **(D)** Sequence of *ZmCCD8* Exon1. Mismatches are shown in blue; the target sites are shown in red, and the PAM is in green.

4.4.2. gRNA design

Six gRNAs were designed to target the promoter region and the first exon of each gene (without considering any predicted cis-regulatory elements). For targeting the promoter region of *ZmCCD7*, three sgRNA sequences **(**T1, T2, T3, **Figure 4.6A)** were individually cloned into vector pZmU3_SaCas9 to generate sgRNA vectors, T1_Sa, T2_Sa, and T3_Sa. For targeting the first exon of *ZmCCD7*, three sgRNA sequences **(**T4, T5, T6, **Figure 4.6A)** were individually cloned into vector pZmU3_SpCas9 to generate sgRNA vectors, T4_Sp, T5_Sp, and T6_Sp. The same strategy was followed for targeting the promoter region and first exon of *ZmCCD8* **(Figure 4.6B)**. In addition, 4 gRNA (T5, T6, T11, T12) were cloned into the pZmU3_SpCas9 vector. The target sites are shown in **Figure 4.6.**

Figure 4.6. Schematic representation of the gRNA target sites. (A) Target sites in *ZmCCD7*. (B) Target sites in *ZmCCD8*. Introns are represented by lines and exons by boxes. The expanded sequence corresponds to the target sites in red and the PAM in blue. PAM Type/SaCas9: 5'-NNGRRT-3' (R can be A or G), PAM Type/SpCas9: 5'-NGG-3' (N is any nucleotide).

4.4.3. Recovery and molecular characterization of transgenic plants

The expected fragments of Cas9 gene and gRNA were amplified by genomic PCR to identify transgenic lines **(Table 4.6)**. The number of SaCas9 positive plants (at DNA level) was 153. The sgRNA_SaCas9 was also detected in 133 of these plants **(Figure 4.6, A)**. The number of SpCas9 positive plants was 121, with 98 plants also containing sgRNA_SpCas9 **(Figure 4.6, B)**. Eight putative positive lines (based on genomic DNA amplification of *Cas9* and gRNA) were selected to confirm the integrity of the Cas9 transgene by Sanger sequencing **(Figure 4.7, C, D)**. Detailed genomic PCR analysis and Sanger sequencing of amplified fragments are shown in **Figure 4.7.**

Table 4.6. Cas9 and gRNA positive transgenic plants (T0)

Experiment	Cas9 positive	gRNA positive
I	62	55
II	51	49
III	40	29
IV	26	18
V	47	37
VI	48	43

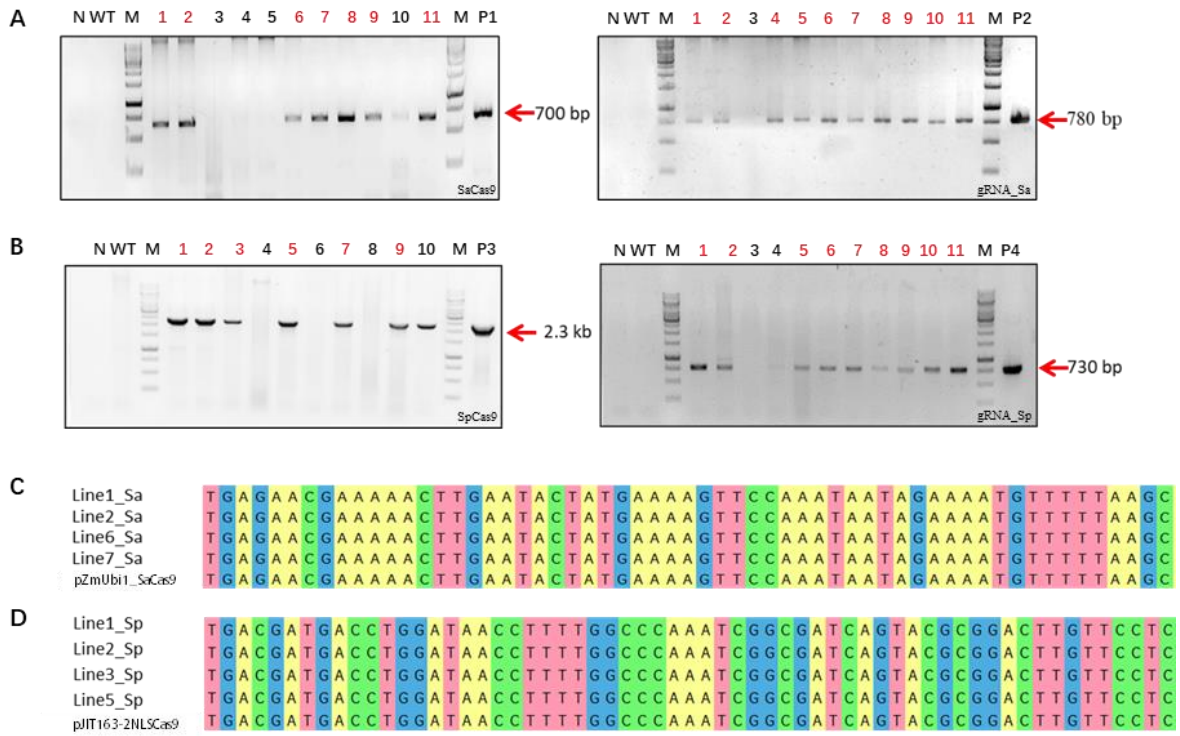


Figure 4.7. Confirmation of transgene integrity of T0 maize plants by genomic PCR and Sanger sequencing. **(A).** PCR amplification of SaCas9 (left) and sgRNA (right), Lanes 1-11: representative independent transformed lines obtained in experiments I, II, III. **(B).** PCR amplification of SpCas9 (left) and sgRNA (right), Lanes 1-11: representative independent transformed lines obtained in experiments IV, V, VI. N: negative PCR control with no DNA template; WT: wild-type M37W; M: 1kb maker; P1, P2, P3, P4: positive control of the corresponding plasmids pZmUbi1_SaCas9, pZmU3_sgRNA_SaCas9, pJIT163-2NLSCas9 and pZmU3_sgRNA_SpCas9. respectively, red lane numbers and arrows indicate positive lines and expected molecular weight, respectively. **(C).** Lines 1, 2, 6, 7 are representative SaCas9-PCR positive plants. **(D).** Lines 1, 2, 3, 5 are representative SpCas9-PCR positive plants.

4.4.4. Expression of *Cas9* in transgenic plants

Based on genomic PCR, 44 independent plant lines were selected for further analyses by RT-PCR and western blot to determine *Cas9* gene mRNA and protein accumulation. I confirmed *SaCas9* and *SpCas9* mRNA accumulation consistent with their expected size in 4 and 8 plant lines, respectively (**Figures 4.8 and 4.9)**. However, western blot analysis showed that there was no protein accumulation at the expected molecular weight for SaCas9 (124 kDa) or SpCas9 (160 kDa). An unexpected fragment (40 kDa) (most likely degraded SaCas9 fragment) was identified in 5 plant lines (**Figure 4.8B**, lanes 1 and 2, representative examples).

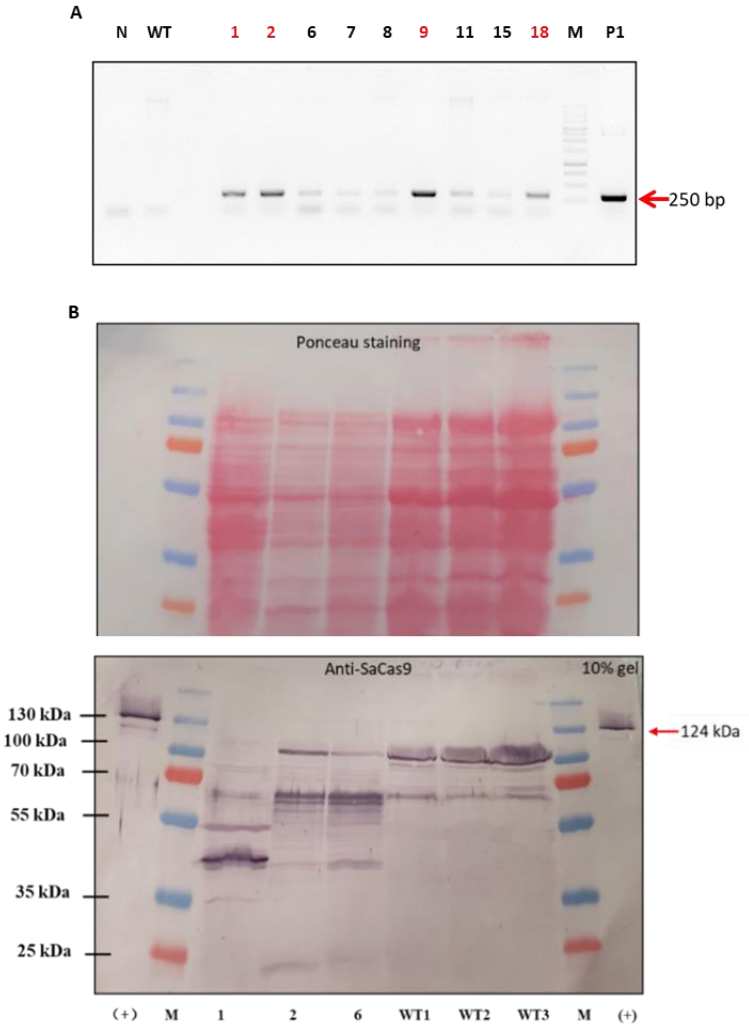


Figure 4.8. RT-PCR and western blot analyses of maize plants for SaCas9. **(A).** RT-PCR analysis. Lane order: lane N-no template negative control; lane WT-wild type M37W plants as negative control; lanes 1-18-independent transgenic lines; lane M-1kb DNA marker; lane P1-pZmUbi1_SaCas9 plasmid (positive control). **(B).** Ponceau staining solution (upper image) was used to detect proteins on the membrane post transfer for the purpose of protein transfer and loading verification. Western blot (lower image). Lane (+): SaCas9 pure protein as positive control; lanes 1, 2, 6-three independent transgenic lines; lane WT 1, 2, 3-three independent wild-type M37W lines as negative control; lane M-protein marker in kDa. Red lane numbers and arrows indicate positive lines and expected molecular weight, respectively.

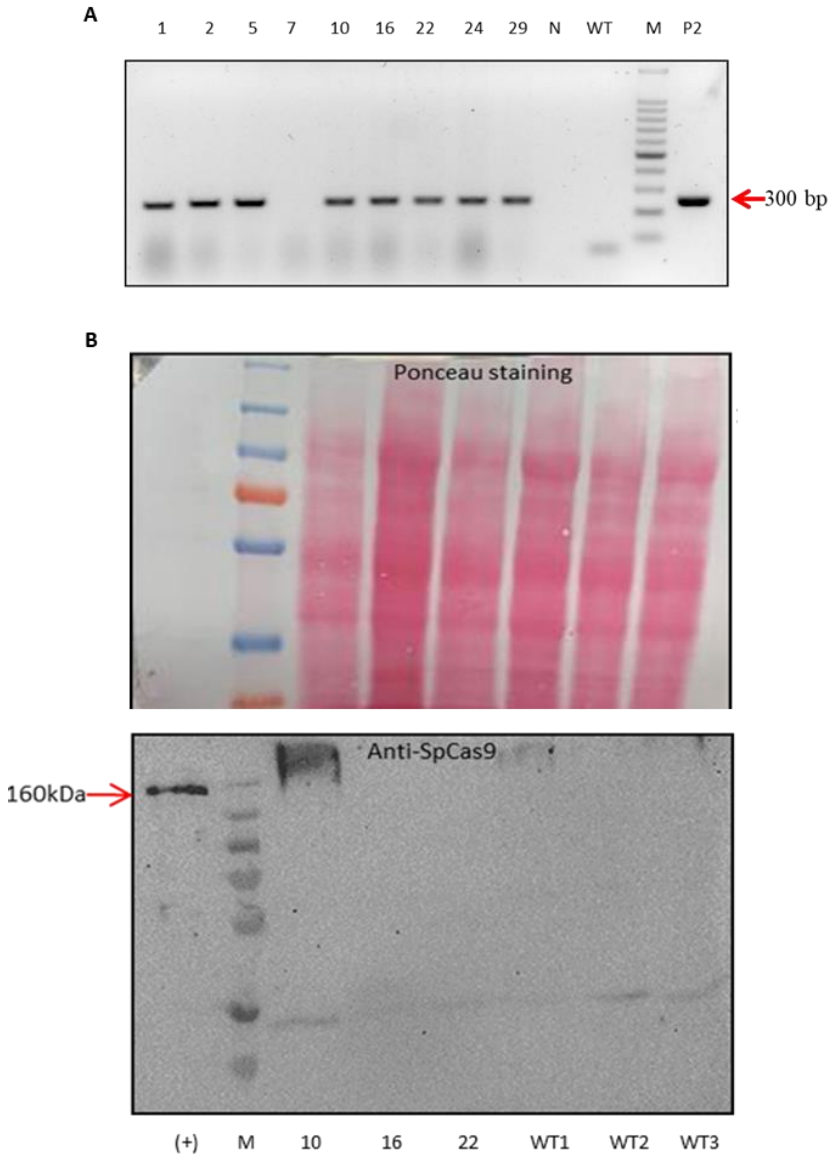
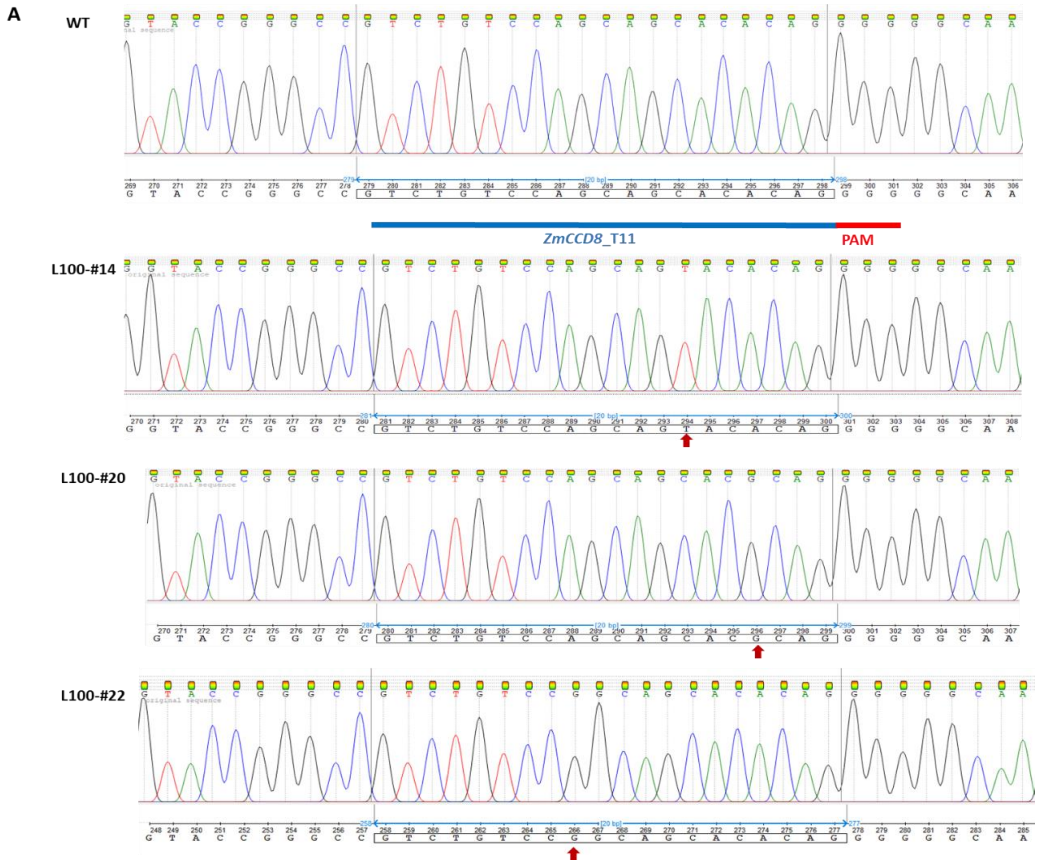


Figure 4.9. RT-PCR and western blot analyses of maize plants for SpCas9. **(A)** RT-PCR analysis. Lane order: lanes 1-29-independent transgenic lines; lane N-no template negative control; lane WT-wild type M37W plants as negative control; lane M-1kb DNA marker; lane P2-pZmUbi1_SpCas9 plasmid (positive control). **(B)** Ponceau staining solution (upper image) was used to detect proteins on the membrane post transfer for the purpose of protein transfer and loading verification. Western blot (lower image). Lane (+): SpCas9 pure protein as positive control; lanes 10-16-22-three independent transgenic lines; lane WT 1, 2, 3-three independent wild-type M37W lines as negative controls; lane M-protein marker in kDa. Red lane numbers and arrows indicate positive lines and expected molecular weight, respectively.

4.4.5. Detection of mutations

Sanger sequencing and mutation analysis confirmed that SpCas9 controlled by the pZmUbi1 generated chimeric mutations in two independent plant lines (L100 and L101). These plants were derived from experiment VII (please see **Table 4.2**) in which I used 4 gRNAs targeting the *ZmCCD7* and *ZmCCD8* first exon sequences simultaneously. Preliminary sequencing showed PCR-amplified fragment with multiple peaks in the target region (the first exon of *ZmCCD8*). I cloned the PCR-amplified fragments covering the target regions into the pGEM-T Easy vector to characterize the mutations. Line L100 contained different substitutions on the target site T11 (L100-#14, L100-#20 and L100-#22, **Figure 4.10**), indicating that each regenerated plant was a chimera of mutant and non-mutant cells. Similar results were obtained for Line L101 on the T12 sequence in its clone L101-#15, G replaced with A at the 5th base. Mutation analysis and recovery of additional transgenic lines are ongoing.



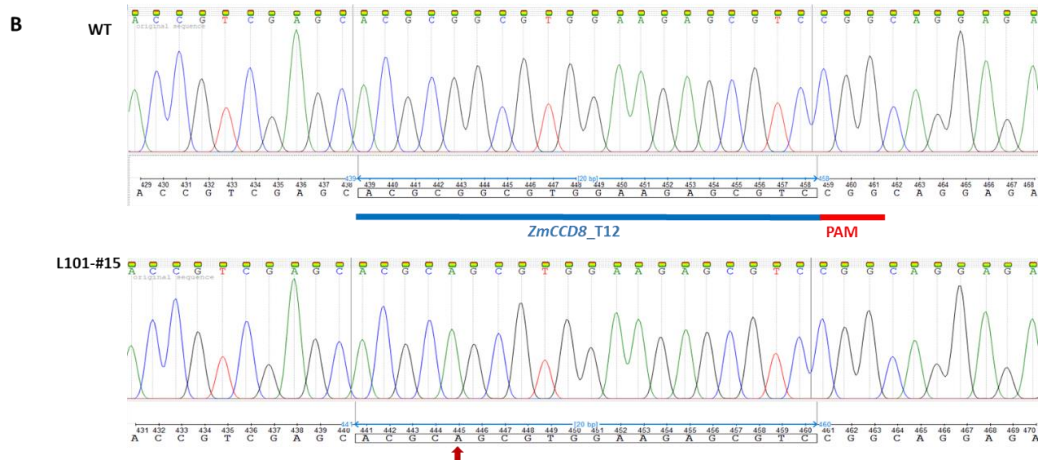


Figure 4.10. Detection of mutations at the ZmCCD8 T11 and T12 sites. **(A).** Sanger sequencing chromatograms of the three mutant clones (#14, #20, and #22) of L100. **(B).** Sanger sequencing chromatogram of the mutant clone (#15) of Line 101. WT: the wild-type sequence is shown for comparison, with the PAM highlighted in red and the target sequence in blue. Mutated bases are indicated with red arrows.

4.5. Discussion

The Green Revolution resulted in an extraordinary increase in staple food productivity over the past 70 years. The production of cereal crops, mainly maize, rice, and wheat tripled during this period with only a 30% increase in cultivated land area (Pingali 2012). However, these increases in productivity were directly dependent on chemical inputs, one of which was the excessive application of chemical fertilizers, in particular nitrogen and phosphorus, which not only resulted in serious contamination of soil and groundwater but also land degradation (Wang et al. 2019). A global meta-analysis of field studies on the impact of AM fungi, indicated that AM were directly responsible for increased grain yields by 16% for a number of cereal crops including wheat, maize, rice, and sorghum, because of increased P and N uptake (Zhang et al. 2019).

Strigolactones are important signaling molecules that promote symbiotic interactions between roots and AM fungi (Al-Babili and Bouwmeester 2015). Their biological roles were illustrated by the use of SL-deficient mutants of pea, Arabidopsis, maize, and rice. AM colonization of roots of the mutant plants were much lower compared to wild-type plants (Gomez-Roldan et al. 2008; Umehara et al. 2008). SL also induce germination of parasitic weeds. SL-deficient mutants of maize and tomato accumulated lower levels of SL, resulting in limited germination of the parasitic weeds *Striga hermonthica*, Phelipanche, and Orobanche (Dor et al. 2011; Stanley et al. 2021). In maize, infestation by parasitic *Striga hermonthica* caused yield losses ranging from 30% to 100% in sub-Saharan regions (Mutsvanga et al. 2022). Similarly, the production of tomato is highly susceptible to infestation by *Orobanche ramosa* and *Orobanche aegyptiaca*, causing severe yield losses of up to 75% (Dor et al. 2011). Therefore, SL also play a role in

limiting germination of parasitic weeds and presents a promising approach to reduce crop losses caused by these parasitic weeds. Reciprocal grafting experiments on these mutants demonstrated that SL also regulate important agricultural traits related to plant architecture, such as the number of tillers and the structure and development of primary and lateral roots (Koltai 2011; Sun et al. 2022). These findings highlight the potential of SL in improving crop yields and quality.

The modification of plant genomes has been practiced initially through selective breeding, then by mutagenesis and transgenesis, and in the last decade by genome editing. Traditional breeding methods can produce diverse germplasm. By crossing different varieties, breeders can create new combinations of genes, resulting in a wider range of traits. While this can be useful for developing crops that are adapted to different environmental conditions, exhibit improved disease resistance, or increase yields, it is a time-consuming process that may not always result in the desired combination of traits, especially of the gene pool of a species is narrow. Traditionally, plant genes are mutagenized randomly using chemical, physical, or biological agents, such as ethyl methane sulfonate and transfer DNA, and extensive screening has to be carried out to select mutants with the desired mutations without causing negative effects on plant growth and development. More recently, different genome editing techniques have been applied in plants: Meganucleases, Zinc finger nucleases (ZFNs), Transcription activator-like effector nucleases (TALENs) and CRISPR/Cas9 offered several benefits. These nucleases allowed the creation of new plant genotypes by introducing precisely targeted double-strand breaks that are resolved by endogenous repair pathways. Unlike traditional breeding methods, genome editing using engineered nucleases is not dependent on random recombination or integration events, which made it more efficient, predictable and less laborious (Townsend et al., 2009; Zhang et al., 2013, Zhu et al 2016).

The SL biosynthetic genes *CCD7* and *CCD8* and their orthologs RAMOSUS5 (RMS5) and RMS1 in pea, MORE AXILLARY BRANCHING3 (MAX3) and MAX4 in Arabidopsis have been shown to regulate plant architecture and parasitic weed germination. In these mutants, shoot branching is increased compared to wild-type (Sorefan et al. 2003; Booker et al. 2004; Johnson et al. 2006; Arite et al. 2007; Yang et al. 2017). In rice, Cas9 generated transgenic plants with homozygous or bi-allelic mutations of the *OsCCD7* gene were severely dwarfed with more tillers (Yang et al 2017). In tomato (*Solanum lycopersicum*), knocking down *SICCD7* or *SICCD8* by siRNA or knocking out *SICCD8* by CRISPR/Cas9 resulted in the expected morphological changes with increased

shoot branching, excessive adventitious roots and plant dwarfism and a decrease in strigolactone content (Vogel et al. 2010; Kohlen et al. 2012). In maize, there are no reports on SL-deficient mutants generated by engineered nucleases. Several SL-deficient mutants were generated using chemical mutagenesis and transposon mutagenesis (Vollbrecht et al 2010). The first report on functional characterization of the *ZmCCD8* involved *ZmCCD8* knock out plants with a dissociation transposon insertion (*zmccd8::Ds*) (Guan et al 2012). Analysis of the *zmccd8* mutant revealed a modest increase in branching, reduction in stem diameter, and delay in the development of adventitious roots.

I aimed to modulate the expression of *ZmCCD7* and *ZmCCD8* simultaneously by creating a series of mutations in the promoter or first exon regions of these genes using CRISPR/Cas9 gene editing. To determine the efficiency of targeted mutagenesis, I carried out different experiments where I tested two Cas9 endonucleases, SaCas9 driven by p*ZmUbi1* and SpCas9 driven by p*2X35S* or p*ZmUbi1*. I used single and multiple gRNA expression cassettes with the combinations of the above constructs and the bar selectable marker gene in transformation experiments. After the preliminary mutation analysis of 153 transgenic maize plant lines containing SaCas9 gene and 133 lines containing SpCas9 gene, potential mutations were identified in the target regions of *ZmCCD8*. To confirm the mutations, the target regions were cloned from the putative mutant plants and 25 clones were sequenced for each line. Two mutant lines, L100 and L101, were identified with 1-bp substitutions. Line L100 had three clones with different mutations, even though they were derived from the same plant. L101 had one clone with a substitution mutation. These results suggest that both lines were chimeras of mutated and non-mutated cells, indicating that the somatic mutations may occur after the embryogenic cell division. Similar observations were previously reported in maize, tomato, rice, and Arabidopsis (Brooks et al. 2014; Feng et al. 2014; Zhang et al. 2014; Svitashev et al. 2015). In tomato, 6 of the 8 T0 transgenic lines transformed with the Cas9 construct carried mutations, and 4 of them were chimeric. PCR genotyping showed that two plants carried hetero and homozygous mutations (Brooks et al. 2014). In rice, analysis of the genotypes and frequency of edited genes in T0 plants showed that the CRISPR/Cas9 system was highly efficient, with target genes edited in nearly half of the transformed embryogenic cells before their first division. However, 23 of the 57 (40.4 %) plants genotyped were chimeric (Zhang et al 2014). In soybean (*Glycine max*), Lie et al, transformed embryonic callus using particle bombardment. The target alleles of mutations often contained different sequences, such that two of the three clones exhibited different sequences in the callus

lines (Li et al. 2015).

In my experiments, I did not recover a sufficient number of independent mutants to allow me to determine whether targeting a specific gene (in my case *ZmCCD7* or *ZmCCD8*) or a particular region in the gene/promoter sequence, might impact the efficiency of recovering targeted mutations. In *Vitis vinifera,* Ren et al. simultaneously knocked out the *VvCCD7* and *VvCCD8* using CRISPR/Cas9. Interestingly, targeted mutations were only achieved in *CCD8*. Mutant plants exhibited more extensive branching than the corresponding wild-type plants (Ren et al. 2020). In rice targeting the second exon of *OsCCD7* by CRISPR/Cas9 resulted in mutant plants which produced an extraordinary number of tillers, ca: 145 tillers per plant, with a significantly lower seed setting compared to targeting the first or seventh exons which also resulted in mutant plants with ca: 101 or 61 tillers, respectively (Yang et al. 2017; Butt et al. 2018). Comparing results between *Vitis* and rice suggests that targeting efficiencies may be influenced by the nature of the target gene(s) in particular when genes are associated with growth and development. Taken together, careful gRNA selection is critical for effective genome editing and mutations at multiple target sites may occur independently. GC content, target strand and target location are all factors that will affect the stability of sgRNA (Wang et al. 2014). The other critical factor for increased mutation efficiency is the features of the Cas9 nuclease construct. For example, the Cas9 nuclease itself, codon usage and promoters. To optimize the efficiency of the Cas9 nuclease in maize, I investigated two endonucleases with different codon usage and promoters. The first nuclease, maize codon optimized SaCas9 under the control of the p*ZmUbi1*, failed to generate targeted mutations in 133 transgenic maize plants in successive independent experiments, even though I was able to confirm the integrity of the SaCas9 and sgRNA cassettes in the transgenic maize plants that I recovered. Currently, there are no reports on editing the maize genome using SaCas9. However, Cheng et al., attempted to knock out the maize *MS8* gene (*Male Sterility 8*), which mainly affects the meiotic stage of anther development by CRISPR/Cas9 and reported that T0 transgenic plants (containing SpCas9) were not mutated. Mutations were recovered in T1 and T2 progenies (Chen et al. 2018). On this basis I crossed or selfed transgenic maize plants (containing SaCas9 and gRNA) to generate T1 plants. Subsequently, I will sequence and analyze six target sites of *ZmCCD7* (T1, T2, T3) and *ZmCCD8* (T7, T8, T9) in T1 plants to determine whether late mutations will take place in progeny. The other nuclease, I used was the rice codon-optimized SpCas9, I evaluated a construct encoding SpCas9 controlled by the p*ZmUbi1*, and

compared that to a construct controlled by the p*2X35S*, which successfully edited *OsBEIIb*, *Waxy*, and *APL2* genes in rice in our previous studies (Baysal et al. 2016; Pérez et al. 2018; Pérez et al. 2019). I was not able to recover any mutations in maize when I used SpCas9 driven by the *2X35S* promoter even in the lines that I confirmed SpCas9 gene expression at mRNA level. In contrast SpCas9 under the control of the p*ZmUbi1* resulted in mutation in the maize plants I generated. Furthermore, the number of nuclear localization signals (NLSs), and the presence or absence of introns are also known to influence mutation efficiencies in plants. Grützner et al., failed to detect Cas9 protein (without introns) in Arabidopsis. After the introduction of 13 introns to the maize codon optimized SpCas9 coding sequence, not only the protein could be detected, but also the mutation rate was improved by 50%. In *N. benthamiana*, the accumulation of Cas9 protein with two NLS was lower than that of Cas9 proteins with one NLS (Grützner et al. 2021). Therefore, species-to-species differences may also have an impact on inducing mutations.

4.6. Conclusions and future work

In summary, I was able to recover targeted mutations in the coding region of *ZmCCD8* by CRISPR/Cas9 mediated genome editing. However, it is unclear at present why similar mutations in *ZmCCD7* could not be recovered. In order to address this, further experiments need to be carried out. Future experiments should focus on designing and testing different gRNAs targeting different coding regions of *ZmCCD7* and recovering more plant lines. It will also be important to test SpCas9 with introns and determine its editing efficiency. Nonetheless, in depth analysis of the *ZmCCD8* mutants I recovered will contribute towards a better understanding of the molecular basis controlling structural traits in maize. Future genome sequencing of subsequent generations as well as metabolomic analysis is needed to assess fully the impact of these mutations.

4.7. References

Akiyama K, Matsuzaki K, Hayashi H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435**: 824-827.

Al-Babili S, Bouwmeester HJ. (2015). Strigolactones, a novel carotenoid-derived plant hormone. *Annu. Rev. Plant Biol.* **66**: 161-186.

Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, Ghisla S, Bouwmeester H, Beyer P, Al-Babili S. (2012). The path from β-carotene to carlactone, a strigolactone-like plant hormone. *Science* **335**: 1348-1351.

Aliche EB, Screpanti C, De Mesmaeker A, Munnik T, Bouwmeester HJ. (2020). Science and application of strigolactones. *New Phytol.* **227**: 1001-1011.

Arite T, Iwata H, Ohshima K, Maekawa M, Nakajima M, Kojima M, Sakakibara H, Kyozuka J. (2007). DWARF10, an RMS1/MAX4/DAD1 ortholog, controls lateral bud outgrowth in rice. *Plant J.* **51**: 1019-1029.

Arite T, Kameoka H, Kyozuka J. (2012). Strigolactone positively controls crown root elongation in rice. *Journal of Plant Growth Regulation* **31**: 165-172.

Baer M, Taramino G, Multani D, Sakai H, Jiao S, Fengler K, Hochholdinger F. (2023). Maize lateral rootless 1 encodes a homolog of the DCAF protein subunit of the CUL4-based E3 ubiquitin ligase complex. *New Phytol.* **237**: 1204-1214.

Bahadur A, Batool A, Nasir F, Jiang S, Mingsen Q, Zhang Q, Pan J, Liu Y, Feng H. (2019). Mechanistic Insights into Arbuscular Mycorrhizal Fungi-Mediated Drought Stress Tolerance in Plants. *International Journal of Molecular Sciences* **20**: 4199.

Bassie L, Zhu C, Romagosa I, Christou P, Capell T. (2008). Transgenic wheat plants expressing an oat arginine decarboxylase cDNA exhibit increases in polyamine content in vegetative tissue and seeds. *Molecular Breeding* **22**: 39-50.

Baysal C, Bortesi L, Zhu C, Farré G, Schillberg S, Christou P. (2016). CRISPR/Cas9 activity in the rice OsBEIIb gene does not induce off-target effects in the closely related paralog OsBEIIa. *Molecular Breeding* **36**: 1-11.

Beveridge CA, Kyozuka J. (2010). New genes in the strigolactone-related shoot branching pathway. *Current opinion in plant biology* **13**: 34-39.

Bhoi A, Yadu B, Chandra J, Keshavkant S. (2021). Contribution of strigolactone in plant physiology, hormonal interaction and abiotic stresses. *Planta* **254**: 1-21.

Bonfante P, Genre A. (2010). Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis. *Nat Commun* **1**: 48.

Booker J, Auldridge M, Wills S, McCarty D, Klee H, Leyser O. (2004). MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. *Current biology* **14**: 1232-1238.

Bortesi L, Fischer R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol Adv* **33**: 41-52.

Brooks C, Nekrasov V, Lippman ZB, Van Eck J. (2014). Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System. *Plant Physiol*.**166**: 1292-1297.

Bruno M, Al-Babili S. (2016). On the substrate specificity of the rice strigolactone biosynthesis enzyme DWARF27. *Planta* **243**: 1429-1440.

Butt H, Jamil M, Wang JY, Al-Babili S, Mahfouz M. (2018). Engineering plant architecture via CRISPR/Cas9-mediated alteration of strigolactone biosynthesis. *BMC plant biology* **18**: 1-9.

Chai YN, Schachtman DP. (2022). Root exudates impact plant performance under abiotic stress. *Trends in Plant Sci.* **27**: 80-91.

Chen R, Xu Q, Liu Y, Zhang J, Ren D, Wang G, Liu Y. (2018). Generation of transgene-free maize male sterile lines using the CRISPR/Cas9 system. *Front Plant Sci* **9**: 1180.

Christensen AH, Quail PH. (1996). Ubiquitin promoter-based vectors for high-level expression of

selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res.* **5**: 213-218.

Christou P, Ford TL, Kofron MJ. (1991). Production of Transgenic Rice (Oryza Sativa L.) Plants from Agronomically Important Indica and Japonica Varieties via Electric Discharge Particle Acceleration of Exogenous DNA into Immature Zygotic Embryos. *Bio/Technology* **9**: 957-962.

Cook CE, Whichard LP, Turner B, Wall ME, Egley GH. (1966). Germination of Witchweed (Striga lutea Lour.): Isolation and Properties of a Potent Stimulant. *Science* **154**: 1189-1190.

Coudert Y, Périn C, Courtois B, Khong NG, Gantet P. (2010). Genetic control of root development in rice, the model cereal. *Trends Plant Sci* **15**: 219-226.

Dor E, Yoneyama K, Wininger S, Kapulnik Y, Yoneyama K, Koltai H, Xie X, Hershenhorn J. (2011). Strigolactone Deficiency Confers Resistance in Tomato Line SL-ORT1 to the Parasitic Weeds Phelipanche and Orobanche spp. *Phytopathology* **101**: 213-222.

Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang D-L, Wang Z, Zhang Z, Zheng R, Yang L. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* **111**: 4632-4637.

Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pagès V, Dun EA, Pillot JP, Letisse F, Matusova R, Danoun S, Portais JC, Bouwmeester H, Bécard G, Beveridge CA, Rameau C, Rochange SF. (2008). Strigolactone inhibition of shoot branching. *Nature* **455**: 189-194.

Grützner R, Martin P, Horn C, Mortensen S, Cram EJ, Lee-Parsons CW, Stuttmann J, Marillonnet S. (2021). High-efficiency genome editing in plants mediated by a Cas9 gene containing multiple introns. *Plant Communications* **2**: 100135.

Hahn F, Mantegazza O, Greiner A, Hegemann P, Eisenhut M, Weber AP. (2017). An efficient visual screen for CRISPR/Cas9 activity in Arabidopsis thaliana. *Front Plant Sci* **8**: 39.

Hayward A, Stirnberg P, Beveridge C, Leyser O. (2009). Interactions between Auxin and Strigolactone in Shoot Branching Control. *Plant Physiol*.**151**: 400-412.

Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R. (2008). Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *J Bacteriol* **190**: 1401-1412.

Ito S, Braguy J, Wang JY, Yoda A, Fiorilli V, Takahashi I, Jamil M, Felemban A, Miyazaki S, Mazzarella T, Chen G-TE, Shinozawa A, Balakrishna A, Berqdar L, Rajan C, Ali S, Haider I, Sasaki Y, Yajima S, Akiyama K, Lanfranco L, Zurbriggen MD, Nomura T, Asami T, Al-Babili S. (2022). Canonical strigolactones are not the major determinant of tillering but important rhizospheric signals in rice. *Science Advances* **8**: eadd1278.

Johnson X, Brcich T, Dun EA, Goussot M, Haurogné K, Beveridge CA, Rameau C. (2006). Branching genes are conserved across species. Genes controlling a novel signal in pea are coregulated by other long-distance signals. *Plant Physiol.* **142**: 1014-1026.

Kaya H, Mikami M, Endo A, Endo M, Toki S. (2016). Highly specific targeted mutagenesis in plants using Staphylococcus aureus Cas9. *Sci Rep* **6**: 26871.

Kohlen W, Charnikhova T, Lammers M, Pollina T, Tóth P, Haider I, Pozo MJ, de Maagd RA, Ruyter-Spira C, Bouwmeester HJ. (2012). The tomato CAROTENOID CLEAVAGE DIOXYGENASE 8 (S l CCD 8) regulates rhizosphere signaling, plant architecture and affects reproductive development through strigolactone biosynthesis. *New Phytol.* **196**: 535-547.

Lee CM, Cradick TJ, Bao G. (2016). The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells. *Molecular Therapy* **24**: 645-654.

Li C, Dong L, Durairaj J, Guan J-C, Yoshimura M, Quinodoz P, Horber R, Gaus K, Li J, Setotaw Y. (2023). Maize resistance to witchweed through changes in strigolactone biosynthesis. *Science* **379**: 94-99.

Li C, Liu C, Qi X, Wu Y, Fei X, Mao L, Cheng B, Li X, Xie C. (2017). RNA-guided Cas9 as an in vivo desired-target mutator in maize. *Plant Biotechnol J* **15**: 1566-1576.

Li Z, Liu Z-B, Xing A, Moon BP, Koellhoffer JP, Huang L, Ward RT, Clifton E, Falco SC, Cigan AM. (2015). Cas9-Guide RNA Directed Genome Editing in Soybean. *Plant Physiol.* **169**: 960-970.

Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading)* **155**: 733-740.

Mutsvanga S, Gasura E, Setimela PS, Nyakurwa CS, Mabasa S. (2022). Nutritional management and maize variety combination effectively control Striga asiatica in southern Africa. *CABI Agriculture and Bioscience* **3**: 1-14.

Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Perez Conesa D, Ros G, Sandmann G, Capell T. (2009). Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 7762-7767.

Pérez L, Soto E, Farré G, Juanos J, Villorbina G, Bassie L, Medina V, Serrato AJ, Sahrawy M, Rojas JA. (2019). CRISPR/Cas9 mutations in the rice Waxy/GBSSI gene induce allele-specific and zygosity-dependent feedback effects on endosperm starch biosynthesis. *Plant Cell Rep* **38**: 417-433.

Pérez L, Soto E, Villorbina G, Bassie L, Medina V, Muñoz P, Capell T, Zhu C, Christou P, Farré G. (2018). CRISPR/Cas9-induced monoallelic mutations in the cytosolic AGPase large subunit gene APL2 induce the ectopic expression of APL2 and the corresponding small subunit gene APS2b in rice leaves. *Transgenic Res.* **27**: 423-439.

Pingali PL. (2012). Green Revolution: Impacts, limits, and the path ahead. *Proc. Natl. Acad. Sci. U.S.A.* **109**: 12302-12308.

Ramessar K, Rademacher T, Sack M, Stadlmann J, Platis D, Stiegler G, Labrou N, Altmann F, Ma J, Stöger E, Capell T, Christou P. (2008). Cost-effective production of a vaginal protein microbicide to prevent HIV transmission. *Proc. Natl. Acad. Sci. U.S.A.* **105**: 3727-3732.

Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F. (2015). In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**: 186-191.

Rath D, Amlinger L, Rath A, Lundgren M. (2015). The CRISPR-Cas immune system: biology, mechanisms and applications. *Biochimie* **117**: 119-128.

Ren C, Guo Y, Kong J, Lecourieux F, Dai Z, Li S, Liang Z. (2020). Knockout of VvCCD8 gene in grapevine affects shoot branching. *BMC plant biology* **20**: 1-8.

Ruyter-Spira C, Kohlen W, Charnikhova T, van Zeijl A, van Bezouwen L, de Ruijter N, Cardoso C, Lopez-Raez JA, Matusova R, Bours R. (2011). Physiological effects of the synthetic strigolactone analog GR24 on root system architecture in Arabidopsis: another belowground role for strigolactones? *Plant Physiol.* **155**: 721-734.

Saeed W, Naseem S, Ali Z. (2017). Strigolactones Biosynthesis and Their Role in Abiotic Stress Resilience in Plants: A Critical Review. *Front Plant Sci* **8**: 1487.

Seto Y, Sado A, Asami K, Hanada A, Umehara M, Akiyama K, Yamaguchi S. (2014). Carlactone is an endogenous biosynthetic precursor for strigolactones. *Proc. Natl. Acad. Sci. U.S.A.* **111**: 1640-1645.

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu J-L, Gao C. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **31**: 686-688.

Shi J, Gao H, Wang H, Lafitte HR, Archibald RL, Yang M, Hakimi SM, Mo H, Habben JE. (2017). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol J* **15**: 207-216.

Snowden KC, Simkin AJ, Janssen BJ, Templeton KR, Loucas HM, Simons JL, Karunairetnam S, Gleave AP, Clark DG, Klee HJ. (2005). The Decreased apical dominance1/Petunia hybrida CAROTENOID CLEAVAGE DIOXYGENASE8 gene affects branch production and plays a role in leaf senescence, root growth, and flower development. *Plant Cell* **17**: 746-759.

Sorefan K, Booker J, Haurogné K, Goussot M, Bainbridge K, Foo E, Chatfield S, Ward S, Beveridge C, Rameau C. (2003). MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea. *Genes Dev* **17**: 1469-1474.

Sun H, Tao J, Gu P, Xu G, Zhang Y. (2016). The role of strigolactones in root development. *Plant Signal Behav* **11**: e1110662.

Svitashev S, Young JK, Schwartz C, Gao H, Falco SC, Cigan AM. (2015). Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol.* **169**: 931-945.

Umehara M, Hanada A, Yoshida S, Akiyama K, Arite T, Takeda-Kamiya N, Magome H, Kamiya Y, Shirasu K, Yoneyama K. (2008). Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455**: 195-200.

Vogel JT, Walter MH, Giavalisco P, Lytovchenko A, Kohlen W, Charnikhova T, Simkin AJ, Goulet C, Strack D, Bouwmeester HJ. (2010). SlCCD7 controls strigolactone biosynthesis, shoot branching and mycorrhiza-induced apocarotenoid formation in tomato. *Plant J.* **61**: 300-311.

Wang B, Zhu L, Zhao B, Zhao Y, Xie Y, Zheng Z, Li Y, Sun J, Wang H. (2019). Development of a Haploid-Inducer Mediated Genome Editing System for Accelerating Maize Breeding. *Mol Plant* **12**: 597-602.

Wang T, Wei JJ, Sabatini DM, Lander ES. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**: 80-84.

Wang Y, Tang Q, Kang Y, Wang X, Zhang H, Li X. (2022). Analysis of the Utilization and Prospects of CRISPR-Cas Technology in the Annotation of Gene Function and Creation New Germplasm in Maize Based on Patent Data. *Cells* **11**: 3471.

Xie X, Yoneyama K, Yoneyama K. (2010). The strigolactone story. *Annual review of phytopathology* **48**: 93-117.

Yang X, Chen L, He J, Yu W. (2017). Knocking out of carotenoid catabolic genes in rice fails to boost carotenoid accumulation, but reveals a mutation in strigolactone biosynthesis. *Plant Cell Rep* **36**: 1533-1545.

Yoneyama K, Xie X, Yoneyama K, Kisugi T, Nomura T, Nakatani Y, Akiyama K, McErlean CS. (2018). Which are the major players, canonical or non-canonical strigolactones? *J. Exp. Bot.* **69**: 2231-2239.

Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, Mao Y, Yang L, Zhang H, Xu N. (2014). The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J* **12**: 797-807.

Zhang S, Lehmann A, Zheng W, You Z, Rillig MC. (2019). Arbuscular mycorrhizal fungi increase grain yields: a meta-analysis. *New Phytol.* **222**: 543-555.

Zhu C, Bortesi L, Baysal C, Twyman RM, Fischer R, Capell T, Schillberg S, Christou P. (2017). Characteristics of Genome Editing Mutations in Cereal Crops. *Trends Plant Sci* **22**: 38-52.

Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T. (2008). Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proc. Natl. Acad. Sci. U.S.A.* **105**: 18232-18237.

GENERAL DISCUSSION

General discussion

The biosynthesis of secondary metabolites and their regulation have long been a focus of research because of their beneficial effects on human health and industrial applications. The production of high-value isoprenoids in heterologous systems has been reported in bacteria, yeast, and plants. Plants are an attractive production platform due to their inherent photosynthetic and carbon utilization machinery as well as the presence of native enzymes for the production of substrates for the biosynthesis of isoprenoids. For example, tomato (*Solanum lycopersicum*), rice (*Oryza sativa*), tobacco (*Nicotiana tabacum*), maize (*Zea mays*), potato (*Solanum tuberosum*), carrot (*Daucus carota*) have all been shown to be potential platforms for the production of useful isoprenoids (Fraser et al. 2002; Gerjets and Sandmann 2006; Jayaraj et al. 2008; Zhu et al. 2008; Huang et al. 2013; Li et al. 2019; Liu et al. 2020; Ahrazem et al. 2022; Zhu et al. 2022). Metabolic engineering is a promising strategy for identifying and developing valuable secondary metabolites in plant production systems, from lead optimization to industrial-scale production. Genetic engineering can be used to better understand the biosynthesis and regulatory mechanisms of specialized isoprenoids produced only in a few plant species by introducing several genes or transcription factors.

The most expensive spice pigment, crocin(s) and the phytohormones strigolactones are apocarotenoids produced by carotenoid degradation by plant carotenoid cleavage dioxygenases (CCDs). These include many subfamilies CCD1, CCC2, CCD4, CCD7 and CCD8, which generate a huge diversity of apocarotenoids (Gonzalez-Jorge et al. 2013; Wang et al. 2019; Dutta et al. 2021). In this study, I focused on biosynthetic enzymes of two crocin-producing plants, (*Crocus sativus*) *CsCCD2L* and (*Buddleja davidii*) *BdCCD4.1*. I was able to engineer *N. glauca* and *N. tabacum* plants expressing *CsCCD2L and BdCCD4.1*. In the leaves of transgenic *N. glauca* plants expressing *CsCCD2L* alone, the levels of crocins reached 400 ug/g DW, nearly 10-fold higher than those obtained in *N. tabacum* and even 3-fold higher than in transgenic *N. tabacum* plants co-expressing *CsCCD2L, BrCrtZ,* and *AtOrMut*. This indicated that *N. glauca* is a better host for crocin production than *N. tabacum*.

Further studies in *N. glauca* revealed that *CsCCD2* and *BdCCD4.1* had contrasting effects on crocin formation, in different tissues. Leaves of *N. glauca* expressing *CsCCD2L* accumulated 1.5-fold higher amounts of crocins compared to those expressing *BdCCD4.1*; the accumulation of crocin in the petals of *N. glauca* expressing *CsCCD2L* ranged from

56-86 µg/g DW, which was 3-fold lower than in plants expressing *BdCCD4.1*. In *N. glauca*, zeaxanthin accumulated at higher levels in the leaves than in the highly pigmented petals, where lutein predominates (Zhu et al. 2007). This indicated that the availability of substrates for *CsCCD2L* is higher in leaves than in petals. However, the availability of substrate for *BdCCD4.1* was the same in both tissues. This tissue-specific activity of endogenous CCDs in leaves and petals also controls the metabolic flux to the production of crocins.

Another approach to optimize metabolite content and composition is to identify and explore the transcription factors (TF) that control metabolism. In plants, TF have been reported to simultaneously activate or repress multiple genes in the isoprenoid pathway, providing a convenient shortcut to multi-gene interventions. Examples have been reported mostly in dicotyledons, and include TF such as WRKY, ERF, NAC and MADS (Jensen et al. 2010; Wang et al. 2010). The heterologous expression of maize C1 (TF R2R3-MYB) significantly increased the accumulation of diterpenoid tanshinones (bioactive nor-diterpenoid molecules) through up-regulation of the MVA and the MEP pathway genes in *Salvia miltiorrhiza* hairy roots (Zhao et al. 2015). However, the mechanisms are largely unknown in monocotyledons, although carotenoid-deficient grains are a major target for metabolic engineering. There is no information on whether a single transcription factor can simultaneously regulate the co-expression of multiple genes in the two pathways in rice. In this thesis, retrieval of promoter sequences of *OsAACT3, OsHMGS1, OsHMGR1, OsDXS2, OsIPPI1* in the MVA or the MEP pathway, revealed common *cis*-acting elements such as a G-box or a hybrid G/C-box. These may represent binding sites for OsBZ8, an ABA-inducible rice bZIP TF. I confirmed that OsBZ8 and these five genes were preferentially transcribed in rice endosperm in a similar manner and OsBZ8 can activate the expression of *OsHMGR1, OsDXS2* and *OsIPPI1* in yeast. Similar expression profiles may be caused by OsBZ8, thus coordinating the expression of the above five genes simultaneously in the plants. My results provide insights into the regulatory basis of the coordinated upregulation of carotenogenic gene expression in rice endosperm.

Modulation of gene or TF expression can produce advantageous effects, but also unwanted traits or undesirable phenotypes which need to be identified and removed to ensure proper growth, development, and yield. For example, strigolactones are important plant hormones involved in nitrogen, phosphorus and water use efficiency and acquisition,

through promoting symbiosis between beneficial fungi (AMF) and plant root systems, as well as by modulating root architecture, shoot branching, leaf senescence, parasitic weed germination and plant defense (Akiyama et al. 2005; Koltai 2011; Al-Babili and Bouwmeester 2015; Waters et al. 2017). Therefore, in order to elucidate the precise role of SL, not only upregulation but also downregulation of the related genes (*ZmCCD7* and *ZmCCD8*) is required. This can be achieved conveniently by a CRISPR/Cas9 strategy. Recent work utilized the power of CRISPR/Cas9 to engineer rice plant architecture through genome editing of *OsCCD7*, resulting in a significantly higher tiller phenotype with an average of 60-140 tillers per plant combined with a 10-20% reduction in height (Yang et al. 2017; Butt et al. 2018). Similarly, another study demonstrated that *SlCCD8* knock out tomato mutants also exhibited a 2-3 fold increase in branching and a decrease in height by ca 50% (Bari et al. 2019).

Utilizing similar CRISPR/Cas9 genome-editing strategies, I attempted to modulate the expression of *ZmCCD7* and *ZmCCD8* in maize, a major global food security staple and feed crop, requiring excessive nitrogen and phosphorus fertilizer inputs, which are major contributors to greenhouse gas emissions, directly impacting climate change and environmental degradation world-wide. I targeted *ZmCCD7* and *ZmCCD8* simultaneously using SpCas9 (*Streptococcus pyogenes* Cas9) and SaCas9 (*Staphylococcus aureus* Cas9) on multiple target sequences in the promoter and coding regions.

My results demonstrated that SpCas9 targeted the coding regions of *ZmCCD8* effectively and induced targeted mutations in the gene. Among plants with mutations, the number of sequencing reads for each allele in the same plant was different, indicating plants were chimeric. This has not been reported before and further work will unravel important developmental and genetic components involved in chimerism. SaCas9 was not effective in inducing mutations, even though the protospacer adjacent motif (PAM) sequence of SaCas9 (5′-NNGRRT-3′) is 2 nt longer than that of SpCas9 (5′-NGG-3′) with higher sequence recognition capacity. Up to now, SaCas9 was shown to be effective in only a few plant species, including Arabidopsis, tobacco, rice, and soybean (Qin et al. 2019; Beying et al. 2020; Huang et al. 2021; Zhang et al. 2022). Other studies in maize indicate that CRISPR/Cas9-derived mutations in T0 plants were stably transmitted to subsequent generations (Zhu et al. 2016; Li et al. 2017). I anticipate that homozygous and transgene-free mutant lines will be obtained in the T1 and subsequent generations. My results in

maize demonstrate the power of genome editing to dissect biological mechanisms and ultimately improve crop performance.

References

Ahrazem O, Zhu C, Huang X, Rubio-Moraga A, Capell T, Christou P, Gómez-Gómez L. (2022). Metabolic Engineering of Crocin Biosynthesis in Nicotiana Species. *Front. Plant Sci* **13**: 861140.

Akiyama K, Matsuzaki K-i, Hayashi H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435**: 824-827.

Al-Babili S, Bouwmeester HJ. (2015). Strigolactones, a novel carotenoid-derived plant hormone. *Annu. Rev. Plant Biol.* **66**: 161-186.

Bari VK, Nassar JA, Kheredin SM, Gal-On A, Ron M, Britt A, Steele D, Yoder J, Aly R. (2019). CRISPR/Cas9-mediated mutagenesis of CAROTENOID CLEAVAGE DIOXYGENASE 8 in tomato provides resistance against the parasitic weed Phelipanche aegyptiaca. *Scientific reports* **9**: 11438. (

Beying N, Schmidt C, Pacher M, Houben A, Puchta H. (2020). CRISPR–Cas9-mediated induction of heritable chromosomal translocations in Arabidopsis. *Nature plants* **6**: 638-645.

Butt H, Jamil M, Wang JY, Al-Babili S, Mahfouz M. (2018). Engineering plant architecture via CRISPR/Cas9-mediated alteration of strigolactone biosynthesis. *BMC plant biology* **18**: 1-9.

Dutta S, Muthusamy V, Chhabra R, Baveja A, Zunjare RU, Mondal TK, Yadava DK, Hossain F. (2021). Low expression of carotenoids cleavage dioxygenase 1 (ccd1) gene improves the retention of provitamin-A in maize grains during storage. *Molecular Genetics and Genomics* **296**: 141-153.

Fraser PD, Romer S, Shipton CA, Mills PB, Kiano JW, Misawa N, Drake RG, Schuch W, Bramley PM. (2002). Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 1092-1097.

Gerjets T, Sandmann G. (2006). Ketocarotenoid formation in transgenic potato. *J. Exp. Bot.* **57**: 3639-3645.

Gonzalez-Jorge S, Ha S-H, Magallanes-Lundback M, Gilliland LU, Zhou A, Lipka AE, Nguyen Y-N, Angelovici R, Lin H, Cepela J. (2013). Carotenoid cleavage dioxygenase4 is a negative regulator of β-carotene content in Arabidopsis seeds. *The Plant Cell* **25**: 4812-4826.

Huang J-C, Zhong Y-J, Liu J, Sandmann G, Chen F. (2013). Metabolic engineering of tomato for high-yield production of astaxanthin. *Metab. Eng.***17**: 59-67.

Huang TK, Armstrong B, Schindele P, Puchta H. (2021). Efficient gene targeting in Nicotiana tabacum using CRISPR/SaCas9 and temperature tolerant LbCas12a. *Plant biotechnology journal* **19**: 1314-1324.

Jayaraj J, Devlin R, Punja Z. (2008). Metabolic engineering of novel ketocarotenoid production in carrot plants. *Transgenic Res.* **17**: 489-501.

Jensen MK, Kjaersgaard T, Nielsen MM, Galberg P, Petersen K, O'shea C, Skriver K. (2010). The Arabidopsis thaliana NAC transcription factor family: structure–function relationships and determinants of ANAC019 stress signalling. *Biochemical Journal* **426**: 183-196.

Koltai H. (2011). Strigolactones are regulators of root development. *New Phytol.* **190**: 545-549.

Li C, Liu C, Qi X, Wu Y, Fei X, Mao L, Cheng B, Li X, Xie C. (2017). RNA-guided Cas9 as an in vivo desired-target mutator in maize. *Plant Biotechnol J* **15**: 1566-1576.

Li J, Mutanda I, Wang K, Yang L, Wang J, Wang Y. (2019). Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in Nicotiana benthamiana. *Nat. Commun.* **10**: 4850.

Liu T, Yu S, Xu Z, Tan J, Wang B, Liu YG, Zhu Q. (2020). Prospects and progress on crocin biosynthetic pathway and metabolic engineering. *Computational and structural biotechnology journal* **18**: 3278-3286.

Qin R, Li J, Li H, Zhang Y, Liu X, Miao Y, Zhang X, Wei P. (2019). Developing a highly efficient and wildly adaptive CRISPR-SaCas9 toolset for plant genome editing. **17**: 706-708.

Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 22338-22343.

Wang JY, Haider I, Jamil M, Fiorilli V, Saito Y, Mi J, Baz L, Kountche BA, Jia K-P, Guo X. (2019). The apocarotenoid metabolite zaxinone regulates growth and strigolactone biosynthesis in rice. *Nat. Commun.* **10**: 810.

Waters MT, Gutjahr C, Bennett T, Nelson DC. (2017). Strigolactone signaling and evolution. *Annu. Rev. Plant Biol.* **68**: 291-322.

Yang X, Chen L, He J, Yu W. (2017). Knocking out of carotenoid catabolic genes in rice fails to boost carotenoid accumulation, but reveals a mutation in strigolactone biosynthesis. *Plant Cell Reports* **36**: 1533-1545.

Zhang Y, Cai Y, Sun S, Han T, Chen L, Hou W. (2022). Using Staphylococcus aureus Cas9 to Expand the Scope of Potential Gene Targets for Genome Editing in Soybean. *International Journal of Molecular Sciences* **23**: 12789.

Zhao S, Zhang J, Tan R, Yang L, Zheng X. (2015). Enhancing diterpenoid concentration in Salvia miltiorrhiza hairy roots through pathway engineering with maize C1 transcription factor. *J. Exp. Bot.* **66**: 7211-7226.

Zhu C, Bai C, Gomez-Gomez L, Sandmann G, Baysal C, Capell T, Christou P. (2022). Rice callus as a high-throughput platform for synthetic biology and metabolic engineering of carotenoids. *Methods in enzymology* **671**: 511-526.

Zhu C, Gerjets T, Sandmann G. (2007). Nicotiana glauca engineered for the production of ketocarotenoids in flowers and leaves by expressing the cyanobacterial crtO ketolase gene. *Transgenic Res.* **16**: 813-821.

Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T. (2008). Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proc. Natl. Acad. Sci. U.S.A.* **105**: 18232-18237.

Zhu J, Song N, Sun S, Yang W, Zhao H, Song W, Lai J. (2016). Efficiency and Inheritance of Targeted Mutagenesis in Maize Using CRISPR-Cas9. *Journal of genetics and genomics* **43**: 25-36.

GENERAL CONCLUSIONS

General conclusions

- Transformation of *CsCCD2L* in *N. glauca* and *N. tabacum* plants or *CsCCD2L* in combination with *BrCrtZ* and *AtOr^{Mut} in N. tabacum* plants led to the production of crocin.
- Transgenic *N. glauca* plants expressing *CsCCD2L* accumulated up to 400 ug/g DW crocin in leaves, approximately 10 times that of transgenic *N. tabacum* plants.
- Transgenic *N. glauca* plants expressing *BdCCD4.1* accumulated similar levels in leaves and petals, up to 302.7 and 321.6 µg/g DW, respectively.
- The production of crocin was associated with decreased levels of endogenous carotenoids, including β-carotene, in leaves as well as in the petals.
- Crocin levels in *N. glauca* (even after 3 years of storage) were still higher than those previously reported in transient expression experiments in *N. benthamiana* plants, or in *N. tabacum* by chloroplast transformation expressing a *CCD4* from *Bixa Orellana*.
- *Os*BZ8 is involved in the transcriptional regulation of MVA and MEP pathway genes and exhibited similar expression patterns in rice endosperm and embryo, suggesting that this transcription factor regulates the expression of these genes.
- A 300 bp promoter region upstream of the transcription start site of the rate limiting enzymes *HMGR1, DXS2 and IPPI1* in the early part of the carotenoid pathway contains a G-box or a hybrid G/C-box that is specifically bound by the *Os*BZ8 transcription factor in rice.
- SpCas9-induced mutations were recovered in two independent plant lines in the coding region of *ZmCCD8*.
- In depth characterization and metabolomic analysis of *ZmCCD8* mutants will contribute toward a better understanding of the molecular basis controlling structural traits in maize.