Mechanistic Investigations to Enhance Carotenoid Content and Composition in Rice Endosperm

Derry Alvarez

Universitat de Lleida

TESI DOCTORAL

Mechanistic Investigations to Enhance Carotenoid Content and Composition in Rice Endosperm

Derry Alvarez

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
Victoria Armario-Najera

Supervisors: **Paul Christou**

Victoria Armario-Najera

Department of Plant Production and Forestry Science
School of Agricultural and Forestry Engineering
University of Lleida

July 11th of 2022

Signatures P. Christou V. Armario-Najera

DEDICATION

To my family and friends

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Dr. Paul Christou and co-supervisor Dr. Victoria Armario-Najera for supporting me over the years with great enthusiasm and patience. 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 Dr. Teresa Capell for making me a member of this great family, her constant feedback, her limitless motivation and energy, and for selfless supporting throughout all these years. I would like to thanks Dr. Changfu Zhu for his kindness and wisdom in all the molecular and carotenoid topic. Without all their guidance and encouragement this PhD would not have been achievable.

I would like to thank Dr. Paul Fraser (The Royal Holloway, University of London), for accepting me to his laboratory and making available facilities for metabolomic analysis and Dr. Gerrish for his kindness and contribution to my research. I would also like to thank Dr. Lourdes Gomez-Gomez (University of Castilla-La Mancha, Spain) for making available GC-MS facilities for metabolic analysis and GC-MS data analysis.

I am grateful to the University of Lleida, specially to the "Departament de Producció Vegetal i Ciència Forestal", for offering me the platform for my PhD study. I would like to thank Dr. Ludovic Bassie for helping in techniques and for the coffee endless conversations, Nuria Gabernet for be an invaluable person that help organizing all the paperwork and Jaume Capell for taking care of all my plants and their wise counsels. To Pedro Cerda and Guillermo Sobrino for helping me with all the experiments, all the moments we shared and for being more than friends and becoming family. To all the lab colleagues from then and now: Alexandre, Can, Jose, Amaya, Maria, Jaume, Marco, Andrea, Xin, Wenshu, Ashwin thank you for all of the assistance and encouragement.

I would like to thank my entire family, Mom and Dad, my sister and niece, Pucho, Juan and Bebo, thank you for all your love and support. Also, I would like to thanks to my love Ana, for understanding and supporting me when I needed it most. All my love and thanks for everything. Words are powerless to express my gratitude.

Summary

My research project focused on investigations to better understand the mechanism of carotenoid accumulation in rice callus and endosperm. A major aim is to better understand the bottleneck of carotenoid biosynthetic pathway, by analyzing genes involved in carotenoid accumulation and their functionality for further use in metabolic engineering applications.

The overall aims were to: 1) reconstitute a native carotenoid pathway in the rice endosperm by defining the minimum number of genes necessary to accumulate β-carotene; 2) restoring *OsPSY1* expression in the endosperm by re-activation of the promoter; 3) test if chloroplast division factor genes in combination with carotenogenic genes enhance carotenoid accumulation in rice endosperm.

PSY1 commits the first step of the carotenoid biosynthetic pathway. In rice endosperm, the *PSY1* promoter is not active so the first part of my work has been to define the minimum number of necessary genes to accumulate β-carotene in rice endosperm. In early experiments, rice plants were transformed with *OsPSY1* and *OsPDS*. The results indicated that the combination of these two genes leads to the accumulation of β-carotene in rice callus. However, co-transformation of the subsequent genes (*OsZDS*, *OsZISO* and *OsCRTISO*) will be important for the full understanding of each genes and the function in the accumulation of carotenoid in rice endosperm. Future work needs to focus on generating a population of transgenic rice plants with the combination of all the above genes in order to dissect the function of each gene in the pathway.

The second part of my project was focus on the reactivation of the *OsPSY1* gene promoter specifically in rice endosperm. To achieve this aim, we identified six potential endosperm specific *cis*-regulatory elements in the promoter region of *OsPSY1*: three prolamin p-boxes, one AACA motif, one GCN4-like motif and an Opaque 2 box (O2 box). To test our hypothesis, we constructed two synthetic *OsPSY1* promoters with 6-motif and 4-motif corrected driving *GFP*. The wild type version of *OsPSY1* promoter under the control of *GFP* was used as a negative control. The results confirmed that *GFP* was highly expressed and the GFP protein was accumulated and detected in the transgenic rice callus lines containing the corrected 6- and 4-motif *OsPSY1* promoter. Taken together, these results confirm that the corrected *OsPSY1* promoter activated gene expression. The *cis*-

regulatory elements have the potential to regulate *OsPSY1*, and when present they can activate the *OsPSY1* promoter, strongly supporting their role in the expression of *OsPSY1* in rice endosperm.

Moreover, we describe a third strategy to understand the role of the plastid division factors *AtPDV1* and *AtARC3* in the accumulation of carotenoids in rice endosperm. One of our strategies was to combine the plastid division genes (*AtARC3* and *AtPDV1*) with carotenogenic genes (*ZmPSY1* and *PaCRTI*). This strategy did not result in any significant differences in carotenoid accumulation and chromoplast quantity. In the second strategy we combined the plastid division factors (*AtARC3* and *AtPDV1*) with $AtOr^{His}$ in order to enhance carotenoid accumulation in the rice endosperm. We demonstrate co-transformation of rice with the three genes (*AtARC3*, *AtPDV1* and $AtOr^{His}$). Also we confirmed mRNA accumulation for the three genes in rice callus. However, we are far from identifying all the necessary factors controlling chromoplast division and number. Further experiments will be required to identify and characterize proteins that associate with these factors in order to determine their precise role in the accumulation of carotenoids in the rice endosperm.

Resumen

Mi proyecto de investigación se centra en las investigaciones para comprender mejor el mecanismo de acumulación de carotenoides en callo y endospermo de arroz. Uno de los principales objetivos es comprender mejor el cuello de botella de la vía biosintética de los carotenoides, analizando los genes implicados en la acumulación de carotenoides y su funcionalidad para su posterior uso en aplicaciones de ingeniería metabólica.

Los objetivos generales eran: 1) reconstituir una vía nativa de carotenoides en el endospermo de arroz definiendo el número mínimo de genes necesarios para acumular β-caroteno; 2) restaurar la expresión de *OsPSY1* en el endospermo mediante la reactivación de su promotor; 3) probar si los genes del factor de división del cloroplasto en combinación con los genes carotenogénicos mejoran la acumulación de carotenoides en el endospermo de arroz.

El PSY1 realiza el primer paso de la vía biosintética de los carotenoides. En el endospermo de arroz, el promotor de *PSY1* no está activo, por lo que la primera parte de mi trabajo ha sido definir el número mínimo de genes necesarios para acumular β-caroteno en el endospermo de arroz. En los primeros experimentos, se transformaron plantas de arroz con *OsPSY1* y *OsPDS*. Los resultados indicaron que la combinación de estos dos genes conduce a la acumulación de β-caroteno en el callo de arroz. Sin embargo, la co-transformación de los genes subsiguientes (*OsZDS*, *OsZISO* y *OsCRTISO*) será importante para la plena comprensión de cada uno de los genes y la función en la acumulación de carotenoides en el endospermo del arroz. El trabajo futuro debe centrarse en la generación de una población de plantas de arroz transgénicas con la combinación de todos los genes mencionados para diseccionar la función de cada gen en la ruta metabólica.

La segunda parte de mi proyecto se centró en la reactivación del promotor del *OsPSY1* específicamente en el endospermo del arroz. Para lograr este objetivo, identificamos seis posibles elementos *cis*-reguladores específicos del endospermo en la región promotora de *OsPSY1*: tres cajas p de prolamina, un motivo AACA, un motivo similar a GCN4 y una caja opaca 2 (caja O2). Para probar nuestra hipótesis, construimos dos promotores sintéticos de *OsPSY1* con 6 motivos y 4 motivos corregidos para la expresión de GFP. La versión de tipo salvaje del promotor *OsPSY1* controlando la expresión de *GFP* se utilizó como control negativo. Los resultados confirmaron que *GFP* se expresaba en gran medida

y que la proteína GFP se acumulaba y se detectaba en las líneas de callos de arroz transgénico que contenían el promotor *OsPSY1* corregido con 6 y 4 motivos. En conjunto, estos resultados confirman que el promotor corregido de *OsPSY1* activó la expresión del gen. Los elementos reguladores *cis* tienen el potencial de regular *OsPSY1*, y cuando están presentes pueden activar el promotor de *OsPSY1*, apoyando fuertemente su papel en la expresión de *OsPSY1* en el endospermo del arroz.

Además, describimos una tercera estrategia para entender el papel de los factores de división del cloroplasto *AtPDV1* y *AtARC3* en la acumulación de carotenoides en el endospermo del arroz. Una de nuestras estrategias fue combinar los genes de división del cloroplasto (*AtARC3* y *AtPDV1*) con genes carotenogénicos (*ZmPSY1* y *PaCRTI*). Esta estrategia no dio lugar a diferencias significativas en la acumulación de carotenoides y en la cantidad de cromoplastos. En la segunda estrategia combinamos los factores de división de cloroplastos (*AtARC3* y *AtPDV1*) con $AtOr^{His}$ para potenciar la acumulación de carotenoides en el endospermo del arroz. Demostramos la co-transformación del arroz con los tres genes (*AtARC3*, *AtPDV1* y $AtOr^{His}$). También confirmamos la acumulación de ARNm para los tres genes en el callo de arroz. Sin embargo, estamos lejos de identificar todos los factores necesarios que controlan la división y el número de cromoplastos. Serán necesarios más experimentos para identificar y caracterizar las proteínas que se asocian a estos factores con el fin de determinar su papel preciso en la acumulación de carotenoides en el endospermo del arroz.

Resum

El meu projecte de recerca es centra en les investigacions per comprendre millor el mecanisme d'acumulació de carotenoides en call i endosperma d'arròs. Un dels principals objectius és comprendre millor el coll d'ampolla de la via biosintètica dels carotenoides, analitzant els gens implicats en l'acumulació de carotenoides i la seva funcionalitat per utilitzar-los posteriorment en aplicacions d'enginyeria metabòlica.

Els objectius generals eren: 1) reconstituir una via nativa de carotenoides a l'endosperma d'arròs definint el nombre mínim de gens necessaris per acumular β-carotè; 2) restaurar l'expressió d'*OsPSY1* a l'endosperma mitjançant la reactivació del seu promotor; 3) provar si els gens del factor de divisió del cloroplast en combinació amb els gens carotenogènics milloren l'acumulació de carotenoides a l'endosperma d'arròs.

El PSY1 fa el primer pas de la via biosintètica dels carotenoides. A l'endosperma d'arròs, el promotor de *PSY1* no està actiu, per la qual cosa la primera part del meu treball ha estat definir el nombre mínim de gens necessaris per acumular β-carotè a l'endosperma d'arròs. Als primers experiments, es van transformar plantes d'arròs amb *OsPSY1* i *OsPDS*. Els resultats van indicar que la combinació d'aquests dos gens condueix a l'acumulació de β-carotè al call d'arròs. Tanmateix, la co-transformació dels gens subsegüents (*OsZDS*, *OsZISO* i *OsCRTISO*) serà important per a la plena comprensió de cadascun dels gens i la funció en l'acumulació de carotenoides a l'endosperma de l'arròs. El treball futur ha de centrar-se en generar una població de plantes d'arròs transgèniques amb la combinació de tots els gens esmentats per diseccionar la funció de cada gen a la ruta metabòlica.

La segona part del meu projecte es va centrar en la reactivació del promotor de l'*OSPSY1* específicament a l'endosperma de l'arròs. Per assolir aquest objectiu, identifiquem sis possibles elements *cis*-reguladors específics de l'endosperma a la regió promotora d'*OsPSY1*: tres caixes p de prolamina, un motiu AACA, un motiu similar a GCN4 i una caixa opaca 2 (caixa O2). Per provar la nostra hipòtesi, construïm dos promotors sintètics de *OsPSY1* amb 6 motius i 4 motius corregits per a l'expressió de *GFP*. La versió de tipus salvatge del promotor *OsPSY1* controlant l'expressió de *GFP* es va utilitzar com a control negatiu. Els resultats van confirmar que *GFP* s'expressava en gran mesura i que la proteïna GFP s'acumulava i es detectava a les línies de tripa d'arròs transgènic que contenien el promotor *OsPSY1* corregit amb 6 i 4 motius. Tot plegat, aquests resultats confirmen que el promotor corregit d'*OsPSY1* va activar l'expressió del gen. Els elements

reguladors *cis* tenen el potencial de regular *OsPSY1*, i quan són presents poden activar el promotor d'*OsPSY1*, recolzant fortament el seu paper en l'expressió d'*OsPSY1* a l'endosperma de l'arròs.

A més, descrivim una tercera estratègia per entendre el paper dels factors de divisió del cloroplast *AtPDV1* i *AtARC3* a l'acumulació de carotenoides a l'endosperma de l'arròs. Una de les nostres estratègies va ser combinar els gens de divisió del cloroplast (*AtARC3* i *AtPDV1*) amb gens carotenogènics (*ZmPSY1* i *PaCRTI*). Aquesta estratègia no va donar lloc a diferències significatives en l'acumulació de carotenoides i la quantitat de cromoplasts. A la segona estratègia combinem els factors de divisió de cloroplasts (*AtARC3* i *AtPDV1*) amb $AtOr^{His}$ per potenciar l'acumulació de carotenoides a l'endosperm de l'arròs. Demostrem la co-transformació de l'arròs amb els tres gens (*AtARC3*, *AtPDV1* i $AtOr^{His}$). També confirmem l'acumulació d'ARNm per als tres gens al call d'arròs. Tot i així, estem lluny d'identificar tots els factors necessaris que controlen la divisió i el nombre de cromoplasts. Seran necessaris més experiments per identificar i caracteritzar les proteïnes que s'associen a aquests factors per tal de determinar-ne el paper precís en l'acumulació de carotenoides a l'endosperma de l'arròs.

Table of Contents

Acknowledgements ..... **I**

Summary ..... **II**

Resumen ..... **III**

Resum ..... **IV**

Table of Contents ..... **V**

Index of Figures ..... **VII**

Index of Tables ..... **VIII**

List of Abbreviations ..... **IX**

Chapter I. GENERAL INTRODUCTION 1

1.0. General Introduction 3

1.1. Importance of pro-vitamin A carotenoid in human health and nutrition 3

1.2. Carotenoid biosynthetic pathway in plants 4

1.3. Strategies for enhancing β-carotene accumulation in plants 6

AIMS AND OBJECTIVES 21

Chapter II. Dissecting the genetic components required to reconstitute a native pathway to β-carotene in rice endosperm 25

2.0. Abstract 27

2.1. Introduction 29

2.1.1. Engineered rice with enhanced levels of pro-vitamin A 29

2.1.2 Regulation of the early carotenoid pathway in rice endosperm 32

2.3.1. Cloning and vector construction 36

2.3.2. Rice transformation 37

2.3.3. Nucleic acid isolation and cDNA synthesis 38

2.3.4. Gene expression analysis by quantitative real-time PCR 38

2.3.5. Carotenoid Analysis 39

2.4.1. Cloning and characterization of *OsPSY1, OsPDS, OsZISO, OsZDS* and *OsCRTISO* genes. 40

2.4.3. Expression analysis of the rice *PSY1, PDS* genes. 42

2.4.4. Identification of minimum set of genes to reconstitute the rice carotenoid pathway in callus. 43

2.4.5. Carotenoid analysis in transgenic callus 44

Chapter III. Identification and repair of cis-acting elements in the endogenous rice PSY1 gene promoter to restore functionality in the endosperm 53

3.0. Abstract 55

3.1. Introduction 57

3.1.1. Cis-acting elements and their importance in modulating gene expression. 57

3.1.2. Transcription factors in endosperm-specific promoters in cereals. 58
3.1.3. Rice *PSY1* gene promoter function and its regulation in different tissues 60
3.3.6. Gene expression analysis by quantitative real-time PCR 64
3.4.2. Expression patterns of corrected synthetic 4- and 6-motif *OsPSY1* promoter using green fluorescence protein (GFP). 68
3.5. Discussion 71
Chapter IV. **The role of plastid division factors (*AtPDV1* and *AtARC3*) and Or gene for enhancing carotenoid accumulation in rice endosperm** 85
4.0. Abstract 87
4.0. Introduction 89
4.1. Plastid division genes and *Orange* (*Or*) gene 89
4.2. Aims and Objectives 92
4.3. Material and Methods 92
4.3.3. Rice transformation 93
4.3.4. Gene expression analysis 93
4.4. Results 94
4.4.1. Co-expression of *ZmPSY1* and *PaCRTI* with *AtARC3* and *AtPDV1* in rice callus. Carotenoid analysis and chromoplast number in transgenic callus. 94
4.4.2. Cloning and functional characterization of *AtPDV1*, *AtARC3* and *AtOr^{His}* genes 96
4.4.3. Co-expression of *AtOr^{His}* with *AtARC3* and *AtPDV1* in rice callus and endosperm. 97
4.5. Discussion 97
4.6. Conclusions 100
4.7. References 101
Chapter V. **GENERAL DISCUSSION** 105
6.1. General discussion 107
6.2. References 112
GENERAL CONCLUSIONS 115
General Conclusions 117

INDEX OF FIGURES

Chapter I.

Figure 1. 1. Vitamin A in humans consists of a group of molecules with a common retinyl group 9

Chapter II

Figure 2. 1. Isoprenoid biosynthetic pathways in the plant cell. The MVA pathway enzymes are shown in blue and the MEP pathway enzymes in green 34
Figure 2. 2. Carotenoid biosynthesis pathway in plants and equivalent steps in bacteria 35
Figure 2. 3. Schematic representation of expression vectors. 40
Figure 2. 4. Restriction enzyme analysis of the transformation vectors..... 40
Figure 2. 5. Transformation strategies to elucidate the necessary genes required to accumulate carotenoids in rice endosperm. 42
Figure 2. 6. PCR analysis confirming presence of input transgenes. 42
Figure 2. 7. mRNA accumulation of OsPSY1 and OsPDS in rice callus..... 43
Figure 2. 8. Line #31 change in color and relative expression over a three-week period. .. **¡Error! Marcador no definido.**
Figure 2. 9. Phenotype of transgenic rice callus expressing carotenogenic genes OsPSY1 and OsPDS.. 44
Figure 2. 10. Carotenoid levels in lines #22, #31, #34 expressing OsPSY1 and OsPDS. 45

Chapter III.

Figure 3. 1. Sequence of the rice (O. sativa Japonica cvv. Nipponbare) Phytoene synthase 1 gene (OsPSY1) promoter and part of the cDNA region (GenBank AP014962.1)..... 66
Figure 3. 2. Schematic representation of expression vectors. 67
Figure 3. 3. Restriction enzyme digestion of transformation vectors..... 68
Figure 3. 4. Confirmation of 4MOspsy1, 6MOspsy1, wtOspsy1 and LMWG cassette insertion by PCR. 68
Figure 3. 5A and B. GFP transgene expression and western blot analysis in different rice callus lines with (A) 6MOsPSY1-GFP-TNos and (B) wtOsPSY1-GFP-TNos 69
Figure 3. 6A and B. Analysis of GFP gene expression and western blot analysis in different rice callus lines with (A) 4MOsPSY1-GFP-TNos and (B) LMWG-GFP-TNos..... 70
Figure 3. 7A and B. Western blot analysis of GFP in different rice seeds lines with (A) 4MOsPSY1-GFP-TNos, (B) 4MOsPSY1-GFP-TNos and wtOsPSY1-GFP-TNos 70

Chapter IV.

Figure 4. 1A, B. mRNA accumulation and carotenoid level in rice callus.. 95
Figure 4. 2A, B. Chromoplast number and protoplast isolation from rice callus..... 95
Figure 4. 3. Schematic representation of expression vectors 96
Figure 4. 4. Restriction enzyme analysis of the transformation vectors..... 96
Figure 4. 5. mRNA accumulation of Or^{His}, AtARC3 and AtPDV1 in rice callus 97

INDEX OF TABLES

Chapter II.

Table 2. 1. Primers used to amplify the OsPSY1 and OsPDS. 37
Table 2. 2. Primers used for qRT-PCR analysis. 38
Table 2. 3. PCR primers and gene accession numbers. Restriction enzyme sequence for cloning marked in red. 41

Chapter III.

Table 3. 1. Cis-acting elements identified in the promoters of LMW glutenin gene types 62
Table 3. 2. Primers used to amplify the OsPSY1 gene promoter region. 63
Table 3. 3. Primers used for qRT-PCR analysis. 64
Table 3. 5. Target sequences to be corrected for the 4 motif OsPSY1 gene promoter. 67
Table 3. 4. Target sequences to be corrected for the 6 motif OsPSY1 gene promoter. 67

Chapter IV.

Table 4. 1. Primers used to amplify AtPDV1, AtARC3, ZmPSY1, CRTI and Or^{His} genes. 93
Table 4. 2. Primers used for qRT-PCR analysis. ..... **¡Error! Marcador no definido.**

ABBREVIATIONS

ABA	abscisic acid
ARC3	accumulation and replication of chloroplast 3
*At*DXS	*Arabidopsis thaliana* 1-deoxy-d-xylulose-5-phosphate synthase
*At*DXS	*Arabidopsis thaliana* 1-deoxy-D-xylulose-5-phosphate synthase
AtOr	*Arabidopsis thaliana* orange gene
BCH	beta-carotene hydroxylase
BCH	β-carotene hydroxylase
bZIP	basic leucine zipper
CCDs	carotenoid cleavage dioxygenase
CHY	non-heme carotene hydroxylases
CRTB	bacterial phytoene synthase
CRTB	bacterial phytoene synthase
CRTI	bacterial desaturase
CRTISO	prolycopene isomerase
CRTW	bacterial β-carotene ketolase
DET1	de-etiolated1
DMAPP	dimethylallyl diphosphate
DRI	dietary reference intake
DRP5B	dynamin-related protein 5b
EDTA	ethylenediaminetetraacetic acid
FtsZ	filamenting temperature sensitive mutant Z protein
GBFs	G-box-binding factors
GFP	green fluorescence protein
GGPP	geranylgeranyl pyrophosphate
GGPS	geranylgeranyl pyrophosphate synthase
GR1	golden rice 1
GR2	golden rice 2
HMGR	3-hydroxy-3-methylglutaryl coenzyme A reductase
IDI	isopentenyl diphosphate isomerase
IPP	isopentenyl pyrophosphate
LCY-β	lycopene β-cyclase
LCY-ε	lycopene ε-cyclase
LMWG	low molecular weight glutenin
MEP	2-C-methyl-D-erythritol 4-phosphate
MVA	mevalonate
NCEDs	9-cis-epoxycarotenoid dioxygenases
O2	maize opaque 2
*Os*DXR	deoxyxylulose 5-phosphate reductoisomerase
*Os*DXS2	deoxyxylulose 5-phosphate synthase
PARC6	paralog of accumulation and replication of chloroplast 6
P-box	prolamin box
PDS	phytoene desaturase
PDV1	plastid division 1
PIP5K	phosphatidylinositol-4-phosphate 5-kinase
PMSF	phenylmethanesulfonylfluoride
PSY	phytoene synthase
PVDF	polyvinylidene difluoride

qRT-PCR	quantitative real-time PCR
RAE	retinol activity equivalents
RNAi	RNA interference
RuBisCO	ribulose 1,5- bisphosphate carboxylase/oxygenase
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
Skn-1	skinhead 1
SNP	single nucleotide polymorphism
TBST	Tris-buffered saline with Tween-20
VAD	vitamin A deficiency
VDE	violaxanthin de-epoxidase
ZDS	z-carotene desaturase
ZEP	zeaxanthin epoxidase
ZISO	z-carotene isomerase
*Zm*PSY1	*Zea mays* phytoene synthase 1
Z-ring	ring-like structure

Chapter I. GENERAL INTRODUCTION

1.0. General Introduction

1.1. Importance of pro-vitamin A carotenoid in human health and nutrition

Global food security is one of the main targets of world leaders, principally combating the problem of hidden hunger. Present estimations are that nearly 690 million people are hungry, up by 10 million people in one year and by nearly 60 million in five years (FAO, 2020). Hidden hunger or micronutrient malnutrition are a public health concern worldwide. One of the micronutrient deficiencies is vitamin A deficiency (VAD). Retinoids, including vitamin A are essential nutrients for the human (and animal) diet. They contain a common retinyl group, comprising an unsubstituted β-ionone ring and an isoprenoid side chain. Retinal or retinaldehyde is necessary for the production of rhodopsin and for the maintenance of epithelial and immune cells, whereas the acidic form (retinoic acid) is a morphogen that plays important roles in cell growth, differentiation, and organogenesis (Kirschfeld et al 1977) (**Figure 1. 1**). Retinal is derived from retinol, which is obtained from the diet (meat and dairy products) in the form of esters such as retinyl esters. Vitamin A deficiency causes many health problems including the deterioration of light-sensitive rod cells essential for low-light vision, and in some cases can lead to an irreversible form of blindness called xerophthalmia (Sommer 2008).

Humans (and animals) possess the enzyme β-carotene 15,150- monooxygenase, which also allows the direct synthesis of retinal from pro-vitamin A carotenoids. Carotenoids are usually found in fruits and vegetables, such as oranges, broccoli, spinach, carrots, squash, sweet potatoes and pumpkins (Harrison 2005). In populations which lack access to animal-derived food, plants are therefore an important dietary source of vitamin A precursors (Farré et al 2011). Retinol Activity Equivalents (RAE) is the recommended unit for vitamin A intake considering bioavailability. One RAE is equivalent to 1 µg of retinol, 2 µg of supplemental β-carotene, 12 µg of dietary β-carotene, 24 µg dietary alpha-carotene or beta-cryptoxanthin (IOM, 2001). The dietary reference intake (DRI) for vitamin A is 900 RAE for males equivalent to 900 µg or 3000 IU of retinol, 1800 µg of β-carotene supplement or 1.08 mg of β-carotene in food (IOM 2001), 700 RAE for females and 400-500 RAE for children.

The majority of the population in the industrialized world has a diverse diet ensuring the DRI for vitamin A. However, the situation in developing countries is totally different, with many living on a cereal-based diet with low levels of essential nutrients, including

β-carotene. Low vitamin A intake is strongly associated with health issues during periods of high nutritional demand, such as in infancy, childhood, pregnancy, and lactation. More than 190 million children and 19.1 million pregnant women around the world have a serum retinol concentration below 0.70 μm/L (WHO 2009). Another effect of vitamin A deficiency in pregnant and lactating women include increased maternal and infant mortality, anemia risk, and slower infant growth and development.

1.2. Carotenoid biosynthetic pathway in plants

Plants produce a diversity of primary and secondary metabolites. Primary metabolites, such as amino acids, enzymes, sugars and organic acids, are mainly involved in plant development, reproduction and are also crucial for plant survival. Secondary metabolites, including isoprenoids and flavonoids, are mainly involved in the adaptation of plants to biotic and abiotic stress responses. Plant secondary metabolites also have a wide range of uses in the biotechnology, agricultural, agrochemical and pharmaceutical industries (Verpoorte et al 2002; Vasconsuelo and Boland 2007).

Isoprenoids (also called terpenoids) are an important group of secondary metabolites, forming one of the largest families of natural products found in all organisms (Capell and Christou 2004). Isoprenoids are derived from two independent and very conserved pathways, the mevalonate (MVA) and 2-C-methyl-D-erythritol 4-phosphate (MEP). The MVA pathway is localized in the cytosol of all eukaryotes, archaea, and some bacteria; and the MEP pathway is active in most bacteria and in the plastids of plants (Vranová et al 2013). Nonetheless, some bacteria and plants have the ability to use both pathways (Vranová et al 2013). The isoprenoids are derived from the same precursors, the five-carbon unit isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP), which are interconverted by the enzyme isopentenyl diphosphate isomerase (IDI) (Giuliano 2017a).

Carotenoids form a very important group of isoprenoids. They are tetraterpenoids, i.e. they comprise eight condensed C5 isoprenoid precursors generating a C40 linear backbone. In plants, this condensation reaction involves the isomeric precursors IPP and DMAPP and occurs in the plastids (Chappell 1995). IPP and DMAPP are derived predominantly from the plastidial MEP pathway (Rodríguez-Concepción and Boronat 2002). The same precursors are formed by the cytosolic MVA pathway, and there is

evidence of intermediate interchange between the two pathways (Rodríguez-Concepción 2006).

In the MEP pathway carotenoid biosynthesis starts with the C5 prenyl phosphates, IPP and DMAPP. They are interconverted by IDI and then the two alternative forms condense to form geranylgeranyl pyrophosphate (GGPP), the precursor of a series of isoprenoid molecules, including gibberellins, quinones, the isoprenoid moieties of chlorophylls and tocopherols and carotenoids, by geranylgeranyl pyrophosphate synthase (GGPS). The condensation of two GGPP molecules by phytoene synthase (PSY) to 15-cis-phytoene, is the first committed step in carotenoid biosynthesis (Misawa et al 1994). 15-Cis-phytoene is converted into all-trans-lycopene by a poly-cis pathway involving two desaturases (phytoene desaturase or PDS and z-carotene desaturase or ZDS) and two isomerases (z-carotene isomerase or ZISO and prolycopene isomerase or CRTISO) (Li et al 2007; Chen et al 2010). A bacterial desaturase/isomerase (CRTI) converts directly 15-cis phytoene into all trans-lycopene, substituting for the four plant enzymes, and for this reason is often used in metabolic engineering. Lycopene is the substrate for two competing cyclases: lycopene β-cyclase (LCY-β) and lycopene ε-cyclase (LCY-ε), introducing, respectively β-rings (double bond at the 5,6 position) and ε-rings (double bond at the 4,5 position). Carotenoids containing non-substituted β-rings are vitamin A precursors. The two cyclases start two competing branches: the α-branch proceeds through α-carotene (ε-β-carotene) to its 3,3´ diol, lutein, while the β-branch proceeds through β-carotene (β-β-carotene) to its 3,3´ diol, zeaxanthin. In most plants, two non-heme carotene hydroxylases (CHY1 and CHY2) and two heme hydroxylases (CYP97A and CYP97C) introduce hydroxyl groups at the 3 and 3´ positions of the rings. In the β-branch, epoxidation by zeaxanthin epoxidase (ZEP) and de-epoxidation by violaxanthin de-epoxidase (VDE) generate the xanthophyll cycle in leaves.

The carotenoid biosynthetic pathway in plants produces four pro-vitamin A carotenoids, having at least one retinyl group. Two of these molecules (α-carotene and β-carotene) accumulate in significant amounts whereas the others (γ-carotene and β-cryptoxanthin) are intermediates and are converted rapidly into downstream products (Farré et al 2010a; Zhu et al 2010). In humans (and animals), several carotenoids derived from plant sources act as antioxidants and protect against diseases, while other carotenoids are precursors of vitamin A and retinoid compounds.

1.3. Strategies for enhancing β-carotene accumulation in plants

Developing countries require a solid strategy to address VAD and increase access to foods rich in pro-vitamin A. Metabolic engineering may have a role in increasing the levels of these compounds since pro-vitamin A carotenoids are synthesized de novo by plants. The focus needs to be on β-carotene because it is the most important and potent of the four pro-vitamin A carotenoids. As well as enhancing the synthesis of β-carotene, additional approaches include the inhibition of post-β-carotene steps to prevent conversion to zeaxanthin and other derivatives, and increasing the ability of plant cells to store β-carotene, thereby providing a metabolic sink and preventing feedback inhibition.

The ¨push strategy¨ is the most common approach to enhance β-carotene content in plants, using a combination of genes to recreate the complete pathway to accumulate carotenoids. The use of the *Zea mays* phytoene synthase 1 (*Zm*PSY1) and bacterial desaturase (*CRTI*) are the most common components in the push strategy¨ in order to increase the accumulation of carotenoids in the rice endosperm (Ye et al. 2000; Paine et al. 2005; Baisakh et al. 2006; Aluru et al. 2008; Bai et al. 2016). Other approaches include the *Arabidopsis thaliana* Orange gene (*AtOr*) in combination with *ZmPSY1* and *CRTI* to promote the differentiation of chromoplasts thus generating a metabolic sink that promotes the accumulation of carotenoids in the endosperm (Lu et al 2006; Bai et al 2016). Another strategy to generate carotenoid precursors combines *Arabidopsis thaliana* 1-deoxy-d-xylulose-5-phosphate synthase (*AtDXS*) with *ZmPSY1* and *CRTI*. In maize and wheat, the majority of the strategies were similar to those used in rice to recreate the biosynthetic pathway and accumulate carotenoids in the endosperm (**Table 1. 1**).

A strategy combining the ¨push strategy¨ with a ¨block strategy¨ was used in wheat (*Triticum aestivum* L. cv. Bobwhite) using bacterial phytoene synthase (*CRTB*) and RNA interference (RNAi) to block the endogenous carotenoid hydroxylase gene (*TaHYD*) (Zeng et al 2015a). In canola (*Brassica napus*) the ¨block strategy¨ was used to downregulate the expression of lycopene ε-cyclase (ε-CYC) using RNAi (Yu et al 2008). RNAi repressed the ε-CYC and led to increased levels of β-carotene, lutein, zeaxanthin and violaxanthin in *B. napus* seeds (**Table 1.1**). In tomato (*Solanum lycopersicum*) several publications exemplify different combination of genes used to enhance carotenoid content but the majority of the examples use a ¨push strategy¨ similarly to rice. A ¨block strategy¨ was used in tomato to silence the DE-ETIOLATED1 (*DET1*) gene by RNAi

(Davuluri et al 2005). This cause increases in the contents of β-carotene and lycopene in mature fruits (**Table 1. 1**).

1.4. Engineered plants with enhanced levels of pro-vitamin A carotenoid

Several metabolic engineering approaches have been made to increase the levels of nutritionally relevant carotenoids in staple cereals. Strengthening the rate-limiting enzyme with a high flux control coefficient is the main target for metabolic engineering (Zhu et al 2008a; Bai et al 2016; Zhai et al 2016). Alternatively, it may be desirable to change the carotenoid composition or enhance the endogenous carotenoid pathway in the tissues of interest. One of the first examples of engineering carotenoid accumulation used the daffodil (*Narcissus pseudonarcissus*) *PSY* gene in rice endosperm. However, this resulted in the accumulation of phytoene but not desaturated products (Burkhardt et al 1997). Golden Rice is the first example of carotenoid accumulation in the endosperm. Golden Rice was achieved with the expression of daffodil *PSY* and *CRTI*. The endosperm carotenoid content in Golden Rice 1 was 1.6 μg/d dry weight, demonstrating the need for a desaturase along with *PSY* to achieve carotenoid accumulation in the endosperm (Ye et al. 2000; Paine et al. 2005). Golden Rice 2 utilized the *Zea mays PSY* along with *CRTI* (Paine et al 2005). The endosperm carotenoid content of Golden Rice 2 was 37 μg/d dry weight (23- fold increase compared to Golden Rice 1) (Paine et al 2005).

Furthermore, overexpression in potato tubers of the cauliflower *Or* regulator, a molecular chaperone able to induce formation of chromoplasts in heterologous systems (Giuliano and Diretto 2007), results in increased PSY stability (Li et al 2012). The tomato Stay Green SlSGR1 protein regulates fruit lycopene and β-carotene accumulation through direct interaction with the PSY enzyme (Luo et al 2013).

The endosperm of cereals are deficient in carotenoid and in many minerals (Zhu et al 2007, 2008a; Harjes et al 2008; Cong et al 2009). Introducing the bacterial phytoene synthase (*CRTB*) and *CRTI* genes under the control of an enhanced λ-zein promoter to provide strong endosperm-specific expression, increased the total carotenoid content to 33.6 g/g dry weight versus the control (1.01 g/g dry weight; Aluru et al. 2008). Beta-carotene hydroxylase (*BCH*) and *LCY-e* are important genes linked with β-carotene and total carotenoid accumulation in maize. Mutations at the *LCY-e* locus redirect the flux towards the β-branch, increasing β-carotene content in maize endosperm (Harjes et al. 2008). A significant advance was achieved with the development of a combinatorial

nuclear transformation system designed to dissect and modify the carotenoid biosynthetic pathway in maize, using the white-endosperm elite South African inbred M37W (Zhu et al. 2008).

Combinatorial nuclear transformation involves co-transformation with multiple genes involved in carotenoid biosynthesis (on individual plasmids) and then screening a library of resulting transformants for plants with appropriate metabolic profiles. Early experiments involved the introduction of *ZmPSY1*, *Gentiana lutea LYCB*, *BCH*, *CRTI* and bacterial β-carotene ketolase (*CRTW*) under the control of endosperm-specific promoters. This recreates the entire pathway from GGPP to zeaxanthin and also adds the enzyme CRTW that converts β-carotene into downstream ketocarotenoids. The M37W genetic background provided a blank template because the endosperm in this variety has no PSY activity and therefore lacks carotenoids. The plants carrying random combinations of genes resulted in a metabolically diverse library comprising plants with a range of carotenoid profiles and endosperm colors ranging from yellow to scarlet. The grains contained high levels of β-carotene, lycopene, zeaxanthin and lutein, as well as further commercially-relevant ketocarotenoids such as astaxanthin and adonixanthin (Zhu et al 2008a). In a subsequent study also in maize, plants transformed with multiple genes enabled the simultaneous modulation of three metabolic pathways in the endosperm, increasing the levels of the three target nutrients: ascorbate, β-carotene, and folate (Naqvi et al 2009). The same M37W line was used as the basis for these experiments because, in addition to the absence of carotenoids, it also has very low ascorbate and folate levels. Kernels from the transgenic plants contained six times the normal level of ascorbate (110 µg/g dry weight), 169-fold more β-carotene than normal (60 µg/g dry weight) and twice the normal level of folate (1.94 µg/g dry weight) (Naqvi et al 2009).

Carotenoid levels in elite wheat lines (*Triticum aestivum* EM12) have also been improved by metabolic engineering. Transgenic wheat expressing maize *PSY1* under the control of the endosperm-specific 1Dx5 promoter in combination with constitutively expressed *CRTI* produced yellow grains containing 10.8-fold the carotenoid levels of wild type EM12 plants (Cong et al 2009). Hexaploid tritordeum accumulate higher levels of carotenoids than their respective wheat parents or hybrids derived from crosses between wild diploid barley and durum wheat (Alvarez et al 1999). A wheat population, which was previously characterized for endosperm color was used to map the *PSY1* and *PSY2*

genes against four QTLs affecting endosperm color, with one showing strong linkage (Clarke et al 2006; Pozniak et al 2007).

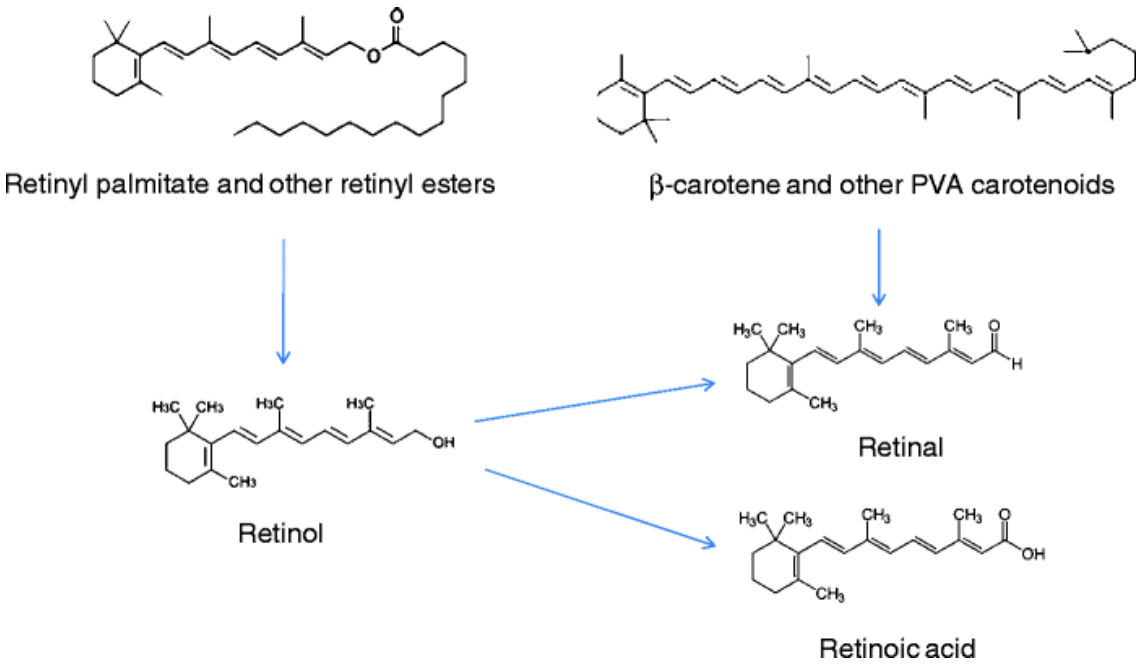


Figure 1. 1. Vitamin A in humans consists of a group of molecules with a common retinyl group. The biologically active molecules are retinal and retinoic acid (Bai et al 2011).

Table 1. 1. *Total carotenoids levels and b-carotene levels in wild type and transgenic plants.*

(DW = dry weight, FW = fresh weight, ND = not determined)

[a] Different carotenoid levels cited for each species (wild type) reflect the different varieties used in each investigation, [b] We converted dry weight to fresh weight assuming the water content of tomato fruit is 90%.

Species	Genes (origin)	Total carotenoid levels in wild type [a]	Total carotenoid levels (increase relative to wild type) in transgenic plants	β-Carotene levels in wild type [a]	β-Carotene levels (increase relative to wild type) in transgenic plants	References
Rice	*PSY1* (*Narcissus pseudonarcissus*; daffodil) and *CRTI* (*P. ananatis*)	ND	1.6 µg/g DW	ND	1.4 µg/g DW	(Ye et al. 2000)
	PSY1 (*Zea mays*; maize) and *CRTI* (*P. ananatis*)	ND	37 µg/g DW	ND	31 µg/g DW	(Paine et al 2005)
	PSY1 (*Zea mays*, maize) and *CRTI* (*P. ananatis*)	ND	1.06 µg/g DW	ND	0.41 µg/g DW	(Baisakh et al 2006)
	CRTB and *CRTI* (*P. ananatis*)	0.99 µg/g DW	33.6 µg/g DW (34)	0.98 µg/g DW	9.8 µg/g DW	(Aluru et al 2008)
	PSY1 (*Zea mays*, maize), *CRTI* (*P. ananatis*) and *AtOr* (*A. thaliana* ORANGE)	ND	25.83 µg/g DW	ND	10.52 µg/g DW	(Bai et al 2016)

	PSY1 (*Zea mays*, maize), *CRTI* (*P. ananatis*) and *AtDXS* (*A. thaliana* 1-deoxy-d-xylulose-5-phosphate synthase)	ND	31.78 µg/g DW	ND	16.61 µg/g DW	(Bai et al 2016)
Maize	*PSY1* (*Zea mays*; maize) *CRTI* (*P. antoeaananatis*) *CRTW* (*P. aracoccus* spp) *LYCB* (*Gentiana lutea*)	1.10 µg/g DW	146.7 µg/g DW (133)	0.14 µg/g DW	57.35 µg/g DW (410)	(Zhu et al 2008a)
	PSY1 (maize) *CRTI* (*Pananatis*)	1.45 µg/g DW	163.2 µg/g DW (112)	0.35 µg/g DW	59.32 µg/g DW (169)	(Naqvi et al 2009)
Wheat	*PSY1* (maize) *CRTI* (*P. ananatis*)	0.46 µg/g DW	4.96 µg/g DW (10.8)	ND	ND	(Cong et al 2009)
	CRTB and CRTI (bacterial)	0.58 µg/g DW	4.06 µg/g DW	0.05 µg/g DW	3.21 µg/g DW	(Wang et al 2014)
	CRTB (bacteria) and *HYDi* (wheat)	1.22 µg/g DW	9.31 µg/g DW	0.16 µg/g DW	5.55 µg/g DW	(Zeng et al 2015b)

Canola	*CRTB* (*P.ananatis*)	36 µg/g FW	1055 µg/g FW (29.3)	5 µg/g FW	401 µg/g FW (80.2)	(Shewmaker et al 1999)
	CRTB (*P. ananatis*)	ND	1341 µg/g FW	ND	739 µg/g FW	(Ravanello et al 2003)
	CRTE and *CRTB* (*P. ananatis*)		1023 µg/g FW		488 µg/g FW	
	CRTB (*P. ananatis*) *CRTI* *(P. ananatis)*		1412 µg/g FW		857 µg/g FW	
	CRTB and *CRTY* (*P. ananatis*)		935 µg/g FW		459 µg/g FW	
	CRTB (*P. ananatis*) and *LYCB* (*Brassica napus;* canola*)*		985 µg/g FW		488 µg/g FW	
	CRTB, CRTI and *CRTY* (*P. ananatis*)		1229 µg/g FW		846 µg/g FW	
	lycopene ε-cyclase (*Arabidopsis*) RNAi to 5' end	5.34 µg/g FW	227.78 µg/g FW (42.5)	0.49 µg/g FW	90.76 µg/g FW (185.2)	(Yu et al 2008)
	lycopene ε-cyclase (*Arabidopsis*) RNAi to 3' end		94.09 µg/g FW (17.6)		27.02 µg/g FW (55)	
	idi, *CRTE*, *CRTB*, *CRTI* and *CRTY* (*P. ananatis*)	21.7 µg/g FW	656.7 µg/g FW (30)	0.2 µg/g FW	214.2 µg/g FW (1070)	(Fujisawa et al 2009)

	CRTZ, CRTW (*Brevundimonas* sp)					
	microRNA miR156b (*Arabidopsis*)	3 µg/g FW (10% water content)	6.9 µg/g FW (2.45) (10% water content)	0.08 µg/g FW (10% water content)	0.38 µg/g FW (4.5) (10% water content)	(Wei et al 2010)
Tomato [b]	*CRTI* (*P. ananatis*)	285 µg/g FW	137.2 µg/g FW (0.5)	27.1 µg/g FW	52 µg/g FW (1.9)	(Römer et al 2000)
	S/BBX20	87.31 µg/g FW	161.70 µg/g FW	54.17 µg/g FW	96.34 µg/g FW	(Xiong et al 2019)
	LYCB (*Solanum lycopersicum*; tomato)	66 µg/g FW	109 µg/g FW (1.7)	7 µg/g FW	57 µg/g FW (7.1)	(Rosati et al 2000)
	AtOR^{his}	71.12 µg/g FW	175.26 µg/g FW	9.03 µg/g FW	28.04 µg/g FW	(Yazdani, Sun, Yuan, Zeng, Theodore W. Thannhauser, et al 2019)
	LYCB (*Arabidopsis*) *BCH* (*Capsicum annuum*; pepper)	66.3 µg/g FW	100.7 µg/g FW (1.5)	5 µg/g FW	63 µg/g FW (12)	(Dharmapuri et al 2002)
	CRTB (*P. ananatis*)	285.7 µg/g FW	591.8 µg/g FW (2.1)	33 µg/g FW	82.5 µg/g FW (2.5)	(Fraser et al 2002)
	LYCB (tomato)	94.5 µg/g FW	215.2 µg/g FW (2.3)	4.4 µg/g FW	205.0 µg/g FW (46.6)	(D'Ambrosio et al 2004)
	DXS (*Escherichia*	460 µg/g FW	720 µg/g FW (1.6)	50 µg/g FW	70.0 µg/g FW (1.4)	(Enfissi et al 2005)

	coli)					
	DET1 (tomato)	128.2 µg/g FW	417.9 µg/g FW	35.1 µg/g FW	145.5 µg/g FW	(Enfissi et al 2010)
	Antisense *DET-1* (tomato)	36.4 µg/g FW	83.8 µg/g FW (2.3)	1.63 µg/g FW	13 µg/g FW (8)	(Davuluri et al 2005)
	CRY2 (tomato)	87.6 µg/g FW in ripe fruit pericarps	149 µg/g FW in ripe fruit (1.7) pericarps (17)	7.8 µg/g FW in ripe fruit pericarps	10.1 µg/g FW in ripe fruit (1.3) pericarps (13)	(Giliberto et al 2005)
	PSY1 (tomato)	181.20 µg/g FW	227.67 µg/g FW (1.25)	58.62 µg/g FW	81.93 µg/g FW (1.4)	(Fraser et al 2007)
	fibrillin (pepper)	325 µg/g FW	650 µg/g FW (2.0)	90 µg/g FW	150 µg/g FW (1.6)	(Simkin et al 2007)
	CRTY (*Erwinia herbicola*)	372.66 µg/g FW	323.71 µg/g FW (09)	6.91 µg/g FW	28.61 µg/g FW (4)	(Wurbs et al 2007)
	LYCB (*Narcissus pseudonarcissus*; daffodil)	76.67 µg/g FW	115 µg/g FW (1.5)	19 µg/g FW	95 µg/g FW (5)	(Apel and Bock 2009)

1.5 References

Aluru M, Xu Y, Guo R, Wang Z, Li S, White W, Wang K, Rodermel S (2008). Generation of transgenic maize with enhanced provitamin A content. *J Exp Bot*., **59**:3551–3562

Alvarez JB, Martín LM, Martín A (1999). Genetic variation for carotenoid pigment content in the amphiploid Hordeum chilense × Triticum turgidum conv. durum. *Plant Breed*., **118**:187–189

Apel W, Bock R (2009). Enhancement of Carotenoid Biosynthesis in Transplastomic Tomatoes by Induced Lycopene-to-Provitamin A Conversion. *Plant Physiol*., **151**:59–66

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

Bai C, Twyman RM, Farré G, Sanahuja G, Christou P, Capell T, Zhu C (2011). A golden era—pro-vitamin A enhancement in diverse crops. *Vitr Cell Dev Biol - Plant*., **47**:205–221

Baisakh N, Rehana S, Rai M, Oliva N, Tan J, Mackill DJ, Khush GS, Datta K, Datta SK (2006). Marker-free transgenic (MFT) near-isogenic introgression lines (NIILs) of 'golden' indica rice (cv. IR64) with accumulation of provitamin A in the endosperm tissue. *Plant Biotechnol J*., **4**:467–475

Burkhardt PK, Beyer P, Wünn J, Klöti A, Armstrong GA, Schledz M, von Lintig J, Potrykus I (1997). Transgenic rice (Oryza sativa) endosperm expressing daffodil (Narcissus pseudonarcissus) phytoene synthase accumulates phytoene, a key intermediate of provitamin A biosynthesis. *Plant J*., **11**:1071–1078

Capell T, Christou P (2004). Progress in plant metabolic engineering. *Curr Opin Biotechnol*., **15**:148–154

Chappell J (1995). Biochemistry and Molecular Biology of the Isoprenoid Biosynthetic Pathway in Plants. *Annu Rev Plant Physiol Plant Mol Biol*., **46**:521–547

Chen Y, Li F, Wurtzel ET (2010). Isolation and Characterization of the Z-ISO Gene Encoding a Missing Component of Carotenoid Biosynthesis in Plants . *Plant Physiol*., **153**:66–79

Clarke FR, Clarke JM, McCaig TN, Knox RE, DePauw RM (2006). Inheritance of yellow pigment concentration in seven durum wheat crosses. *Can J Plant Sci*., **86**:133–141

Cong L, Wang C, Chen L, Liu H, Yang G, He G (2009). Expression of phytoene synthase1 and Carotene Desaturase crtI Genes Result in an Increase in the Total Carotenoids Content in Transgenic Elite Wheat (Triticum aestivum L.). *J Agric Food Chem*., **57**:8652–8660

D'Ambrosio C, Giorio G, Marino I, Merendino A, Petrozza A, Salfi L, Stigliani AL, Cellini F (2004). Virtually complete conversion of lycopene into β-carotene in fruits of tomato plants transformed with the tomato lycopene β-cyclase (tlcy-b) cDNA.

Plant Sci., **166**:207–214

Davuluri GR, van Tuinen A, Fraser PD, Manfredonia A, Newman R, Burgess D, Brummell DA, King SR, Palys J, Uhlig J, Bramley PM, Pennings HMJ, Bowler C (2005). Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nat Biotechnol*., **23**:890–895

Dharmapuri S, Rosati C, Pallara P, Aquilani R, Bouvier F, Camara B, Giuliano G (2002). Metabolic engineering of xanthophyll content in tomato fruits. *FEBS Lett*., **519**:30–34

Enfissi EMA, Barneche F, Ahmed I, Lichtlé C, Gerrish C, McQuinn RP, Giovannoni JJ, Lopez-Juez E, Bowler C, Bramley PM, Fraser PD (2010). Integrative transcript and metabolite analysis of nutritionally enhanced DE-ETIOLATED1 downregulated tomato fruit. *Plant Cell*., **22**:1190–1215

Enfissi EMA, Fraser PD, Lois L-M, 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

FAO, IFAD, UNICEF W and W (2020). The State of Food Security and Nutrition in the World 2020

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

Fraser PD, Enfissi EMA, Halket JM, Truesdale MR, Yu D, Gerrish C, Bramley PM (2007). Manipulation of Phytoene Levels in Tomato Fruit: Effects on Isoprenoids, Plastids, and Intermediary Metabolism. *Plant Cell*., **19**:3194–3211

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*., **99**:1092 LP – 1097

Fujisawa M, Takita E, Harada H, Sakurai N, Suzuki H, Ohyama K, Shibata D, Misawa N (2009). Pathway engineering of Brassica napus seeds using multiple key enzyme genes involved in ketocarotenoid formation. *J Exp Bot*., **60**:1319–1332

Giliberto L, Perrotta G, Pallara P, Weller JL, Fraser PD, Bramley PM, Fiore A, Tavazza M, Giuliano G (2005). Manipulation of the Blue Light Photoreceptor Cryptochrome 2 in Tomato Affects Vegetative Development, Flowering Time, and Fruit Antioxidant Content. *Plant Physiol*., **137**:199–208

Giuliano G (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr Opin Biotechnol*., **44**:169–180

Giuliano G, Diretto G (2007). Of chromoplasts and chaperones. *Trends Plant Sci*., **12**:529–531

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008). Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science (80-).*, **319**:330–333

Harrison EH (2005). Mechanisms of digestion and absorption of dietary vitamin A. *Annu Rev Nutr*., **25**:87–103

Institute of Medicine (2001). Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc. The National Academies Press, Washington, DC

Kirschfeld K, Franceschini N, Minke B (1977). Evidence for a sensitising pigment in fly photoreceptors. *Nature*., **269**:386–390

Li F, Murillo C, Wurtzel ET (2007). Maize Y9 Encodes a Product Essential for 15-cis-ζ-Carotene Isomerization. *Plant Physiol*., **144**:1181–1189

Li L, Yang Y, Xu Q, Owsiany K, Welsch R, Chitchumroonchokchai C, Lu S, Van Eck J, Deng X-X, Failla M, Thannhauser TW (2012). The Or Gene Enhances Carotenoid Accumulation and Stability During Post-Harvest Storage of Potato Tubers. *Mol Plant*., **5**:339–352

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*., **18**:3594–3605

Luo Z, Zhang J, Li J, Yang C, Wang T, Ouyang B, Li H, Giovannoni J, Ye Z (2013). A STAY-GREEN protein SlSGR1 regulates lycopene and β-carotene accumulation by interacting directly with SlPSY1 during ripening processes in tomato. *New Phytol*., **198**:442–452

Misawa N, Truesdale MR, Sandmann G, Fraser PD, Bird C, Schuch W, Bramley PM (1994). Expression of a Tomato cDNA Coding for Phytoene Synthase in *Escherichia coli*, Phytoene Formation *In Vivo* and *In Vitro*, and Functional Analysis of the Various Truncated Gene Products. *J Biochem*., **116**:980–985

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*., **106**:7762–7767

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL, Drake R (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat Biotechnol*., **23**:482–487

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor Appl Genet*., **114**:525–537

Ravanello MP, Ke D, Alvarez J, Huang B, Shewmaker CK (2003). Coordinate expression of multiple bacterial carotenoid genes in canola leading to altered carotenoid production. *Metab Eng*., **5**:255–263

Rodríguez-Concepción M (2006). Early Steps in Isoprenoid Biosynthesis: Multilevel Regulation of the Supply of Common Precursors in Plant Cells. *Phytochem Rev*., **5**:1–15

Rodríguez-Concepción M, Boronat A (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol*., **130**:1079–1089

Römer S, Fraser PD, Kiano JW, Shipton CA, Misawa N, Schuch W, Bramley PM (2000). Elevation of the provitamin A content of transgenic tomato plants. *Nat Biotechnol*., **18**:666–669

Rosati C, Aquilani R, Dharmapuri S, Pallara P, Marusic C, Tavazza R, Bouvier F, Camara B, Giuliano G (2000). Metabolic engineering of beta-carotene and lycopene content in tomato fruit. *Plant J*., **24**:413–420

Shewmaker CK, Sheehy JA, Daley M, Colburn S, Ke DY (1999). Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. *Plant J*., **20**:401–412

Simkin AJ, Gaffé J, Alcaraz J-P, Carde J-P, Bramley PM, Fraser PD, Kuntz M (2007). Fibrillin influence on plastid ultrastructure and pigment content in tomato fruit. *Phytochemistry*., **68**:1545–1556

Sommer A (2008). Vitamin a deficiency and clinical disease: an historical overview. *J Nutr*., **138**:1835–1839

Vasconsuelo A, Boland R (2007). Molecular aspects of the early stages of elicitation of secondary metabolites in plants. *Plant Sci*., **172**:861–875

Verpoorte R, Contin A, Memelink J (2002). Biotechnology for the production of plant secondary metabolites. *Phytochem Rev*., **1**:13–25

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

Wang C, Zeng J, Li Y, Hu W, Chen L, Miao Y, Deng P, Yuan C, Ma C, Chen X, Zang M, Wang Q, Li K, Chang J, Wang Y, Yang G, He G (2014). Enrichment of provitamin A content in wheat (Triticum aestivum L.) by introduction of the bacterial carotenoid biosynthetic genes CrtB and CrtI. *J Exp Bot*., **65**:2545–2556

Wei S, Yu B, Gruber MY, Khachatourians GG, Hegedus DD, Hannoufa A (2010). Enhanced Seed Carotenoid Levels and Branching in Transgenic Brassica napus Expressing the Arabidopsis miR156b Gene. *J Agric Food Chem*., **58**:9572–9578

World Health Organization (2009). Global prevalence of vitamin A deficiency in populations at risk 1995-2005 : WHO global database on vitamin A deficiency. *WHO Iris*., 55

Wurbs D, Ruf S, Bock R (2007). Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. *Plant J*., **49**:276–288

Xiong C, Luo D, Lin A, Zhang C, Shan L, He P, Li B, Zhang Q, Hua B, Yuan Z, Li H, Zhang J, Yang C, Lu Y, Ye Z, Wang T (2019). A tomato B-box protein SlBBX20 modulates carotenoid biosynthesis by directly activating PHYTOENE SYNTHASE 1, and is targeted for 26S proteasome-mediated degradation. *New Phytol*., **221**:279–294

Xudong Y, Salim A-B, Andreas K, Jing Z, Paola L, Peter B, Ingo P (2000). Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science (80-)*., **287**:303–305

Yazdani M, Sun Z, Yuan H, Zeng S, Thannhauser TW, Vrebalov J, Ma Q, Xu Y, Fei Z, Van Eck J, Tian S, Tadmor Y, Giovannoni JJ, Li L (2019). Ectopic expression of ORANGE promotes carotenoid accumulation and fruit development in tomato. *Plant Biotechnol J*., **17**:33–49

Yu B, Lydiate DJ, Young LW, Schäfer UA, Hannoufa A (2008). Enhancing the carotenoid content of Brassica napus seeds by downregulating lycopene epsilon cyclase. *Transgenic Res*., **17**:573–585

Zeng J, Wang X, Miao Y, Wang C, Zang M, Chen X, Li M, Li X, Wang Q, Li K, Chang J, Wang Y, Yang G, He G (2015a). Metabolic Engineering of Wheat Provitamin A by Simultaneously Overexpressing CrtB and Silencing Carotenoid Hydroxylase (TaHYD). *J Agric Food Chem*., **63**:9083–9092

Zeng J, Wang X, Miao Y, Wang C, Zang M, Chen X, Li M, Li X, Wang Q, Li K, Chang J, Wang Y, Yang G, He G (2015b). Metabolic Engineering of Wheat Provitamin A by Simultaneously Overexpressing CrtB and Silencing Carotenoid Hydroxylase (TaHYD). *J Agric Food Chem*., **63**:9083–9092

Zhai S, Xia X, He Z (2016). Carotenoids in Staple Cereals: Metabolism, Regulation, and Genetic Manipulation. *Front Plant Sci*., **7**:1197

Zhu C, Bai C, Sanahuja G, Yuan D, Farré G, Naqvi S, Shi L, Capell T, Christou P (2010). The regulation of carotenoid pigmentation in flowers. *Arch Biochem Biophys*., **504**:132–141

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 C, Naqvi S, Gomez-Galera S, Pelacho AM, Capell T, Christou P (2007). Transgenic strategies for the nutritional enhancement of plants. *Trends Plant Sci*., **12**:548–555

AIMS AND OBJECTIVES

Aims

My research project focused on investigations to better understand the mechanism of carotenoid accumulation in rice callus and endosperm. A major aim was to develop a more in depth understanding of the regulation of the carotenoid biosynthetic pathway in rice. This also encompassed the development of a system to characterize functionally candidate genes implicated in the pathway for further use in metabolic engineering applications.

Objectives

In order to reach these aims, the following objectives were address:

- Reconstitute a plant-only pathway to β-carotene in rice by defining the precise nature and number of rice (phytoene) desaturases and/or (zeta-carotene) isomerases.
- Construct a rice *PSY1* gene synthetic promoter with 4 and 6 *cis*-elements to activate the promoter region of *OsPSY1* gene.
- Test if the plastid division factors genes (*AtPDV1* and *AtARC3*) enhance the carotenoid accumulation in rice callus and endosperm.

Chapter II

Dissecting the genetic components required to reconstitute a native pathway to β-carotene in rice endosperm

2.0.Abstract

The rice endosperm carotenoid biosynthetic pathway (similarly to other plants) is characterized by multiple branches, complex regulatory controls, and competition for intermediate precursor molecules. These are some of the hurdles that jointly limit the synthesis of carotenoids. The major focus in carotenoid synthesis is to accumulate β-carotene. Therefore, when the limitation of carotenoid accumulation in rice endosperm is somehow surpassed, other questions were raised. The golden rice results raise questions about the regulation of the carotenoid pathway. This chapter focused in the dissection and understand the limitations in endogenous rice endosperm carotenoid biosynthesis. We overexpress the *OsPSY1* and *OsPDS* genes in order to define the minimal number of genes required to reconstitute the pathway to β-carotene. Using the established rice callus platform, we demonstrate the accumulation of carotenoids specifically β-carotene. This revealed that in order to accumulate carotenoids in rice endosperm the *OsPSY1* and *OsPDS* genes need to be active in the endosperm. This results provide an exciting new data for the understanding of the rice endosperm carotenoid pathway and also new strategies for carotenoid accumulation in engineered rice plants.

2.1. Introduction

2.1.1. Engineered rice with enhanced levels of pro-vitamin A

The endosperm of cereals is a major food staple worldwide, but it is scarce in many vitamins and minerals, including carotenoids (Zhu et al 2007, 2008a). Earlier approaches to enhance the levels of pro-vitamin A in rice endosperm followed alternative strategies. Carotenoid content in rice endosperm can be enhanced by the expression of carotenogenic enzymes, if isoprenoid precursors are available. In rice endosperm, carotenoid biosynthesis is blocked at the first committed enzymatic step because phytoene synthase 1 (PSY1), which converts GGPP into phytoene is not active (**Figure 2. 1**). In addition, there is a limited flux in the subsequent desaturation and isomerization reactions leading to lycopene.

Strategies to enhance carotenoid production in rice endosperm are exemplified by the creation of Golden Rice 1 (GR1). GR1 was engineered by expressing the *N. pseudonarcissus* PSY1 and bacterial CRTI enzymes in rice endosperm (**Figure 2. 2**). GR1 accumulated 1.6 µg/g DW of β-carotene, with a low accumulation of lutein and zeaxanthin. The accumulation of lutein and zeaxanthin suggest that the lycopene α(ε), β-cyclases and the hydroxylase are either constitutively expressed in normal rice endosperm or induced upon lycopene formation. A number of transgenic Indica rice varieties were generated subsequently using the same strategy to increase β-carotene content in the endosperm in elite varieties (Datta et al. 2003; Parkhi et al. 2005; Baisakh et al. 2006). All transgenic indica rice lines had very similar carotenoid levels in the range of 1.05 µg/g to 2.00 µg/g DW. The amount of β-carotene in the endosperm of all GR1 varieties was very low and insufficient to meet the RDI for vitamin A in the diet (IOM 2001) .

A systematic comparison of different plant PSY1 enzymes in combination with CRTI demonstrated that the origin of the enzyme had a major impact on the levels of β-carotene in rice endosperm (Paine et al 2005). *PSY1* from different carotenoid-rich sources (daffodil, pepper, tomato, rice, and maize) was stably transformed into maize callus. The origin of *PSY* influenced the amounts of carotenoid in the transgenic callus, confirmed by the colored carotenoid composition (calculated in percent of total carotenoids). The highest β-carotene composition was maize (51.3%) and rice (35.5%). The tomato (15.8%), daffodil (15.9%) and pepper (10.1%) resulted in less β-carotene composition. Based on the callus results they use the *PSY* from the maize, rice, tomato, pepper and daffodil to transform rice. The highest β-carotene content in the T1 seed was achieved

using either the maize (12.9 µg/g DW) or rice (15.6 µg/g DW) PSY. The pepper (3.66 µg/g DW) or tomato (0.86 µg/g DW) *PSY1* resulted in intermediate β-carotene content, whereas daffodil *PSY* gave the lowest levels (0.82 µg/g DW). In order to avoid silencing through co-suppression, the maize *PSY* rather than its rice homolog was used for subsequent experiments. Golden Rice 2 (GR2) was thus developed using maize *PSY1* and bacterial *CRTI*, with an accumulation of up to 37 µg/g total carotenoids (a 23-fold increase compared to the use of a daffodil *PSY*) with β-carotene being by far the predominant carotenoid (31 µg/g DW) (Paine et al 2005). α-Carotene (15%), β-cryptoxanthin (2%), zeaxanthin and lutein (less than one percent1% each) were very minor components of the total carotenoid content in the endosperm. The GR2 cassette (*ZmPSY1* and *PaCRTI*) was introduced into Kitaake rice (*japonica*) using targeted gene insertion (Dong et al 2020). The use of SSU-*CRTI* (fusion of the DNA encoding the chloroplast transit peptide from the pea Rubisco small subunit with *CRTI*) and *ZmPSY* by CRISPR-Cas9 was another technique used to deliver carotenogenic genes in order to accumulate carotenoids in rice endosperm. Broader applicability of these results, however, are constrained by a number of experimental limitations. There was no significant improvement in β-carotene content (7.90 µg/g DW) compare to GR2 and also a fragment of the CRISPR vector was integrated into the plant intergenic region causing extra steps to eliminate the non-necessary elements inserted (Dong et al 2020).

Isoprenoids in the cytosol and mitochondria are derived from the MVA pathway, whereas those in plastids are derived from the MEP pathway. The MVA pathway is present in archaebacterial, some gram-positive bacteria, yeasts and animals, while most gram-negative bacteria, cyanobacteria and green algae use the MEP pathway for the biosynthesis of terpenoids (Disch et al. 1998; Vranová et al. 2013). Plants use both pathways to synthesize isoprenoids and to regulate their biosynthesis. The MVA and MEP pathways produce the C5 prenyl phosphates, IPP, and DMAPP, the two precursors for carotenoid biosynthesis (Rodríguez-Concepción 2010; Pulido et al 2012; Giuliano 2017a).

The roles of the MEP pathway enzymes deoxyxylulose 5-phosphate synthase (*OsDXS2*) and deoxyxylulose 5-phosphate reductoisomerase (*OsDXR*) were investigated in carotenoid metabolism in rice endosperm (You et al. 2020). (Ha et al 2010). Plants co-expressing *OsDXS2* with *PSY1* and *CRTI* accumulated modest levels of β-carotene in the endosperm (9.22 µg/g DW), whereas β-carotene levels in plants in which *OsDXR*

replaced *OsDXS* were similar to those in plants expressing only *PSY1* and *CRTI* (1.09 µg/g DW) (You et al 2020). These experiments used rice codon optimized *PSY1* and *CRTI,* from Capsicum and *P. annatis*, respectively. It was thus demonstrated that the *OsDXS2* is a rate-limiting enzyme in the biosynthesis of IPP/DMAPP.

Different approaches were used to explore and overcome bottlenecks in the early carotenoid pathway in rice endosperm. The first approach uses the *Arabidopsis thaliana* 1-deoxy-D-xylulose-5-phosphate synthase (*At*DXS) that is a rate-limiting enzyme in production of plastid-derived isopentenyl diphosphate. Analysis of several transgenic lines showed that plants overexpressing *DXS* had increased levels of isoprenoids (Estévez et al 2001). The second approach uses *AtOr* that triggers the differentiation of proplastids and/or other noncolored plastids into chromoplasts, thus enhancing storage sink capacity for carotenoid biosynthesis accumulation (Lu et al 2006). Or proteins were shown to interact directly and post-transcriptionally regulate *PSY* in controlling carotenoid biosynthesis (Zhou et al 2015; Chayut et al 2017). The two strategies were in combination of carotenogenic genes *ZmPSY1* and *CRTI* (Bai et al. 2016). It was found that *AtDXS* combined with *ZmPSY1* and *CRTI* significantly enhanced the accumulation of carotenoids in rice endosperm with a β-carotene content of 16.6 µg/g DW, confirming that the supply of isoprenoid precursors such as GGPP is a rate-limiting step. The combined expression of *ZmPSY1*, *CRTI* and *AtOR* also boosted carotenoid accumulation (10.5 µg/g DW β-carotene content) through the creation of a metabolic sink.

The 3-Hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) is a key regulatory enzyme and a major rate-limiting enzyme in the MVA pathway (Enfissi et al 2005)(**Figure 2. 1).** In *S. cerevisiae* overexpression of the catalytic domain of the *HMGR* (truncated HMG-CoA reductase gene or *tHMG1*) boosts isoprenoid biosynthesis (Li et al 2013). tHMG1 in combination with maize *PSY1* and bacterial *CRTI* was used to elucidate the role of *HMGR* in carotenoid accumulation rice endosperm (Tian et al. 2019). The strategy was developed using the combination of carotenogenic genes *ZmPSY1* and *CRTI* with *tHMG1*. The study concluded that *tHMG1* enhanced the total carotenoid (14.45 µg/g DW) and β-carotene (10.51 µg/g DW) accumulation in rice endosperm compared with the control construct (*ZmPSY1*/*CRTI*) that accumulates 8.05 µg/g DW in total carotenoids and 5.15 µg/g DW in β-carotene.

2.1.2 Regulation of the early carotenoid pathway in rice endosperm

The rice endosperm carotenoid biosynthetic pathway (similarly to other plants) is characterized by multiple branches, complex regulatory controls, and competition for intermediate precursor molecules. These are some of the hurdles that jointly limit the synthesis of carotenoids. IPP synthesized through the MVA pathway can be transported to mitochondria for the biosynthesis of mitochondrial isoprenoids such as ubiquinones (Pulido et al 2012)(**Figure 2. 1**).

The plastidial MEP pathway synthesizes as well IPP and DMAPP for the biosynthesis of carotenoids (Zhu et al 2007; Pulido et al 2012; Giuliano 2017a). Early experiment trying to elucidate the importance of the different carotenoid biosynthesis genes started with the expression of *PSY1*. The expression of daffodil *PSY* gene demonstrates the importance of the accumulation of phytoene. Therefore, the expression of *PSY* alone does not result in colored carotenoid accumulation. The results suggest that in order to accumulate carotenoids, desaturases and/or isomerases enzymes need to be expressed in combination with phytoene synthase 1 gene (Burkhardt et al 1997). The use of bacterial CRTI catalyze the reaction by substituting the plant desaturase and isomerase steps in the plant carotenoid pathway. Interestingly the expression of *CRTI* alone cannot produce any carotenoid, because of the low levels of the precursor phytoene (Schaub et al 2005)(**Figure 2. 2**).

The PSY1 is a very tightly regulated enzyme affected by light (Jin et al 2021) (**Figure 2. 1**). The interaction of MEP pathway with PSY1 was observed with the appearance of plastidial isoprenoids in de-etiolated rice leaves. This interaction correlates with the upregulation of *Os*DXS1 and the increase of chlorophyll, carotenoid, and tocopherol (Rodríguez-Concepción 2010; Jin et al 2021). In etiolated seedlings, a de-repression of *PSY1* and *DXS* gene expression contributes to a burst in the production of carotenoids in coordination with chlorophyll biosynthesis and chloroplast development for an optimal transition to photosynthetic metabolism during de-etiolation (Jin et al 2021). It is likely that such interaction contributes to ensuring that the prenyl diphosphate precursors required for carotenoid biosynthesis will be supplied when necessary, as presumed from the phenotypes resulting when flux through only one of the two pathways (MVA and MEP) is up-regulated (Rodríguez-Concepción 2010). The upregulation of key genes in MVA and MEP pathways are the different strategies used for the accumulation of carotenoids in rice endosperm.

In the MVA pathway, the expression of *HMGR* with the combination of *PSY1* and *CRTI* genes significantly enhanced the accumulation of carotenoids in rice endosperm (Tian et al 2019). The study demonstrated that a metabolic sink towards carotenoids can be generated by boosting the flux through the MVA pathway. The study of the genes involved in the carotenoid pathway concludes that the pathway was with a low amount of a key metabolite (phytoene) (Ye et al. 2000; Paine et al. 2005; Schaub et al. 2005). The results suggest that the precursor genes in the MEP pathways are also limited. The overexpression of *DXS* in combination with carotenogenic genes (*ZmPSY1/CRTI*) demonstrates the capacity of carotenoid accumulation in the rice endosperm (You et al 2020). Also, establish the need of phytoene precursor in order to avoid the bottleneck of the carotenoid pathway in the endosperm. Moreover, the use of *Or* gene in combination with carotenogenic genes (*ZmPSY1/CRTI*) demonstrates the capacity to create a sink and accumulate carotenoids (Tzuri et al 2015; Bai et al 2016; Yazdani, Sun, Yuan, Zeng, Theodore W. Thannhauser, et al 2019). In addition, overexpression of key precursor genes (*Or*, *DXS* with *PSY1*, and *CRTI*) demonstrates the ability to accumulate carotenoid in rice endosperm.

Another hurdle in carotenoid pathway regulation is downstream of the pathway. The major focus in carotenoid synthesis is to accumulate β-carotene. Therefore, when the limitation of carotenoid accumulation in rice endosperm is somehow surpassed, other questions were raised. The GR results raise questions about the regulation of the carotenoid pathway. The yellow endosperm in GR instead of a red endosperm suggested an activation of endogenous enzymes downstream of the carotenoid pathway(Schaub et al 2005). The studies concluded that the expression of endogenous lycopene cyclases (*LCYs*) can metabolize lycopene to α-carotene and β-carotene. (Schaub et al 2005; Harjes et al 2008; Giuliano 2017a). Moreover, downstream of the carotenoid pathway, the accumulation of other carotenoids (lutein and zeaxanthin) suggests that α- and β-carotene hydroxylases (BCH) enzymes are active and metabolize the subsequent steps in the carotenoid pathway (Schaub et al 2005; Farré et al 2010a; Rodríguez-Concepción 2010; Giuliano 2017a). Finally, the carotenoid pathway can end in degradation, through the carotenoid cleavage dioxygenase (*CCDs*) gene family that is responsible for abscisic acid (ABA), apocarotenoids and strigolactones (Rodríguez-Concepción 2010; Ruiz-Sola and Rodríguez-Concepción 2012; Giuliano 2017a; Moreno et al 2021) (**Figure 2. 1**).

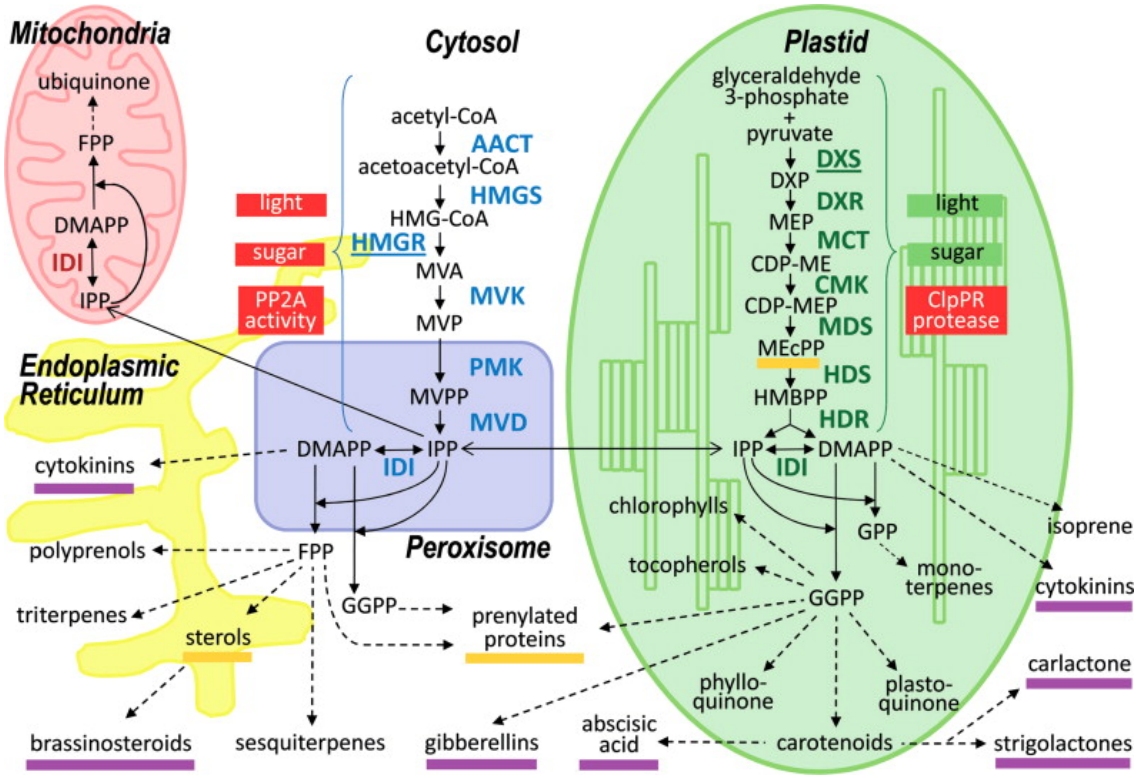


Figure 2. 1. Isoprenoid biosynthetic pathways in the plant cell. The MVA pathway enzymes are shown in blue and the MEP pathway enzymes in green (Pulido et al 2012).

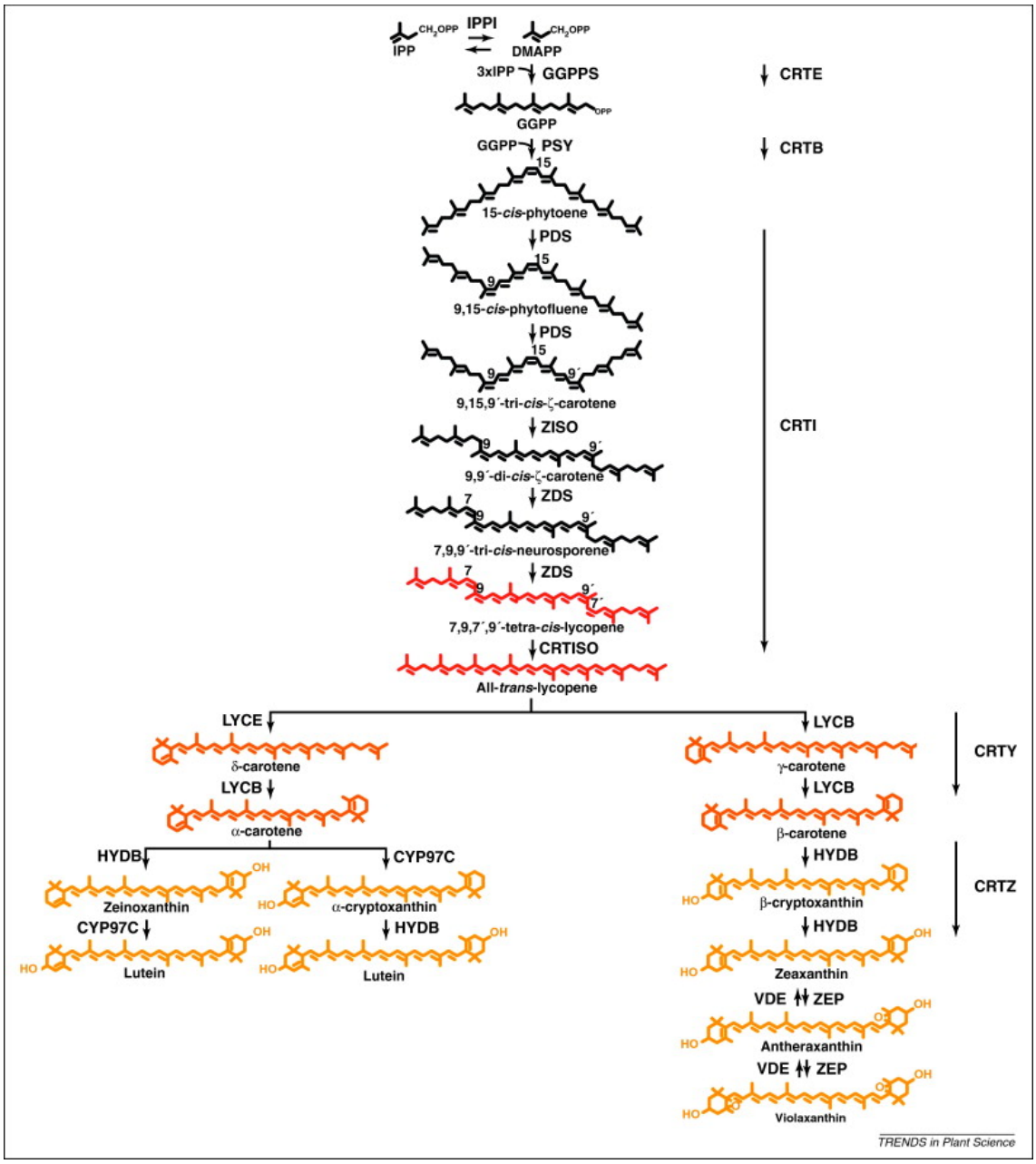


Figure 2. 2. Carotenoid biosynthesis pathway in plants and equivalent steps in bacteria (Farré et al 2011).

2.2. Aims and objectives

The aims of this chapter were to dissect and understand the limitations in endogenous rice endosperm carotenoid biosynthesis.

Objectives:

- Define the precise nature and number of rice (phytoene) desaturases and/or (zeta-carotene) isomerases required to reconstitute the pathway to β-carotene.
- Introduce the minimum set of genes to reconstitute the pathway to β-carotene.
- Use transcriptomic and metabolomics analyses to understand the impact of the above genetic interventions in rice endosperm.

2.3. Materials and Methods

2.3.1. Cloning and vector construction

The *OsPSY1* and *OsPDS*, genes were cloned directly from rice leaf mRNA by RT-PCR using *OsPSY1* and *OsPDS* forward and reverse primers (Table 2. 1). The primers were designed to flank regions of each gene based on sequences in GenBank (accession numbers FJ971175 and XM_015777614). The RT-PCR products were transferred to vector pGEM-Teasy (Promega, Madison, WI, USA) to generate pGEM-*OsPSY1* and pGEM-*OsPDS* for sequencing. The *OsZISO*, *OsZDS* and *OsCRTISO* genes were synthesized by Synbio-technologies Inc. (NJ, USA) using the vector pUC-57 based on the sequences in GenBank (accession numbers XM_015762760.1, XM_015791038.1, XM_015761536.1 respectively.

Vectors pGEM-*OsPSY1*, pUC57-*OsZISO*, pUC57-*OsZDS*, pUC57-*OsCRTISO* were digested with XbaI and EcoRI, and pGEM-*OsPDS* was digested with XbaI and SacI, then were individually transferred to pHorp-P (Sorensen et al., 1996), containing the endosperm specific barley D-hordein promoter and the rice ADPGPP (ADP-Glucose pyrophosphorylase) terminator, to generate pHord-*OsPSY1*-ADPGPP, pHord-*OsPDS*-ADPGPP, pHord-*OsZISO*-ADPGPP, pHord-*OsZDS*-ADPGPP and, pHord-*OsCRTISO*-ADPGPP respectively. The integrity of all intermediate and final constructs was verified by sequencing.

Table 2. 1. Primers used to amplify the *OsPSY1* and *OsPDS*.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
OsPSY1	TCTAGAATGGCGGCCATCACGCTCCTA	GAATTCCTACTTCTGGCTATTTCTCAGTGA
OsPDS	TCTAGAATGGATACTGGCTGCCTGTCATC	GAGCTCCTAGGAGGCAACAGGAACTTCA

OsPSY1, rice (*Oryza sativa* ssp. *japonica*) phytoene synthase 1 gene; XbaI (5´-TCTAGA-3´) and EcoRI (5´-GAATTC-3´) sites (underlined and red) were introduced into each primer. *OsPDS*, phytoene desaturase gene; XbaI (5´-TCTAGA-3´) and SacI (5´-GAGCTC-3´) sites (underlined and red) were introduced into each primer.

2.3.2. Rice transformation

Rice seeds from Oryza sativa. cv. Bomba (provided by https://www.illaderiu.com/en/home-3/) were dehusked and surface-sterilized in 70% ethanol with continuous shaking for 5 minutes, placed in a 50% sodium hypochlorite solution with three drops of Tween for 30 minutes with agitation and rinsed five times in sterile distilled water. The seeds were dried on sterile filter paper and placed on Murashige-Skoog (Murashige and Skoog 1962) Proliferation Medium (4.4 g L^{-1} MSP including Gamborg B5 Vitamins, Duchefa Biochemie) supplemented with casein hydrolysate 300 mg L^{-1}, proline 500 mg L^{-1}, sucrose 30 g L^{-1}, and 500 μL L^{-1} 2,4 D (from 5 mg mL^{-1} stock in ethanol), solidified with 3.5 g L^{-1} phytagel (0.35%), and incubated in the dark at 28°C. The pH of the medium was adjusted to 5.8 using KOH before autoclaving at 120ºC for 20 min. Seeds were germinated for six or seven. Mature zygotic embryos were then separated from the endosperm and transferred to MS Osmoticum medium (MSO) (MS 4.4 g L^{-1} including Gamborg B5 Vitamins, casein hydrolysate 300 mg L^{-1}, proline 500 mg L^{-1}, sucrose 30 g L^{-1}, and 500 μl L^{-1} 2,4 D (from 5 mg mL^{-1} stock in ethanol) and 72.8 g L^{-1} mannitol solidified with 3.5 g L^{-1} phytagel). Twenty mature embryos were put in each MSO petri dish for 4 hours before and 16 hours after bombardment. The mature embryo scutellum was orientated in such a way as to be in the direct path of the accelerated gold particles (Sudhakar et al 1998; Valdez et al 1998; Saba-Mayoral et al 2022).

2.3.3. Nucleic acid isolation and cDNA synthesis

Genomic DNA was extracted from 5g of frozen leaf tissue as described by Sambrook et al. (1989). Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, and USA) and DNA was removed with DNase I (RNase-Free DNase Set, Qiagen).

Total RNA was quantified using a Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and 1µg of total RNA was used as the template for first-strand cDNA synthesis with Ominiscript Reverse Transcriptase in a 20µl total reaction volume according to the manufacturer's recommendations (Qiagen).

2.3.4. Gene expression analysis by quantitative real-time PCR

Quantitative real-time PCR (qRT-PCR) was performed on a CFX96 system (Bio-Rad Laboratories, Hercules, CA, USA) using a 25µl mixture containing 10 ng cDNA, 1x iQ SYBR Green Supermix (Bio-Rad) and 0.2µM of each primer. To calculate relative expression levels, serial dilutions (0.2–125ng) were used to produce standard curves for each gene. Triplicate PCRs in 96-well optical reaction plates were carried out with the following profile: a heating step for 3 min at 95°C was followed by 40 cycles of 95°C for 10s, 59°C for 30s, and 72°C for 20s. Amplification specificity was confirmed by melt curve analysis of the final PCR products in the temperature range 50–90°C with fluorescence acquired after each 0.5°C increment. The fluorescence threshold value and gene expression data were calculated using CFX96 system software. Primer combinations are listed in supporting information (Table 2. 2).

Table 2. 2. Primers used for qRT-PCR analysis.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
OsPSY1	TATGCTCATGACGGAGGACC	GCATCAAGCATGTCGTAGGG
OsPDS	TTGCGGGACAACTTCCTACT	AACAACCTGTAGAGCACCGA
OsZISO	AGCTATGGCAAGTTCAGGGT	TTGTCAACAGCTGCCACTTC
OsZDS	AGCATCACAAGGATGTGGGA	TGAAACCTACCACCCCTGTC
OsCRTISO	GCCTGCTGATACTGATTGCC	CCTGGACAGACCTTCCCAAT
OsACTIN	GACTCTGGTGATGGTGTCAGC	TCATGTCCCTCACAATTTCC

OsPSY1, rice (*Oryza sativa* ssp. *japonica*) phytoene synthase 1 gene; *OsPDS*, rice phytoene desaturase gene; *OsZISO*, rice zeta-carotene isomerase gene; *OsZDS*, rice zeta-carotene desaturase gene; *OsCRTISO*, rice prolycopene isomerase 1 gene; *OsACTIN*, rice actin gene.

2.3.5. Carotenoid Analysis

Chromatographic and mass spectrometry (MS) analyses were performed in Dr. Paul Fraser laboratory at Royal Holloway University. Callus samples were freeze-dried before analysis, and 10 mg of powdered tissue was extracted with 500µl water: methanol (50:50) and vortexed for 30 min at room temperature. I then added 500µl chloroform and vortexed the mixture before centrifuging at 15,300 g for 3min at room temperature. The upper phase was collected, evaporated, and re-dissolved in HPLC-grade ethyl acetate for UPLC analysis using an Acquity system (Waters, Watford, UK) with an Ethylene Bridged Hybrid (BEH C18) column (2.1×100mm, 1.7mm) with a BEH C18 VanGuard pre-column (2.1×50 mm, 1.7mm). Mobile phase A was methanol: water (50:50), and mobile phase B was acetonitrile: ethyl acetate (75:25) and the flow rate was 0.5ml min−1. All solvents were passed through a 0.2mm filter before use. The gradient started at 30% A for 0.5 min and was then stepped to 0.1% A for 5.5min and then to 30% A for the last 2min. The column temperature was maintained at 30°C and the sample temperature at 8°C. Continuous online scanning was performed across the UV/visible range from 250 to 600nm, using a Waters extended wavelength photo-diode array detector (Nogueira et al 2013).

2.4. Results

2.4.1. Cloning and characterization of *OsPSY1*, *OsPDS*, *OsZISO*, *OsZDS* and *OsCRTISO* genes.

The phytoene synthase 1 (*Os*PSY1) and phytoene desaturase (*Os*PDS) cDNAs encoding the corresponding full-length enzymes were amplified by RT-PCR from *O. sativa* ssp. Japonica cv. Niponnbare cDNA. The 15-cis-zeta-carotene isomerase (*Os*ZISO), zeta carotene desaturase (*Os*ZDS) and prolycopene isomerase (*Os*CRTISO) were synthesized by Synbio Technologies Inc. (https://www.synbio-tech.com/) (Table 1). All vectors were confirmed by restriction digestion and sequencing (**Table 2. 3**, **Figure 2. 3** and **Figure 2. 4**).

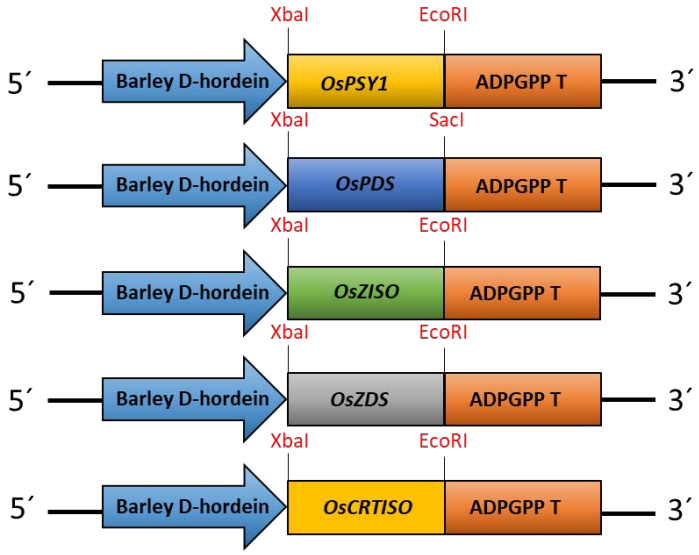


Figure 2. 3. Schematic representation of expression vectors p-Hord-OsPSY1, p-Hord-OsPDS, p-Hord-OsZISO, p-Hord-OsZDS and p-Hord-OsCRTISO with the restriction enzymes (marked in red) used for cloning.

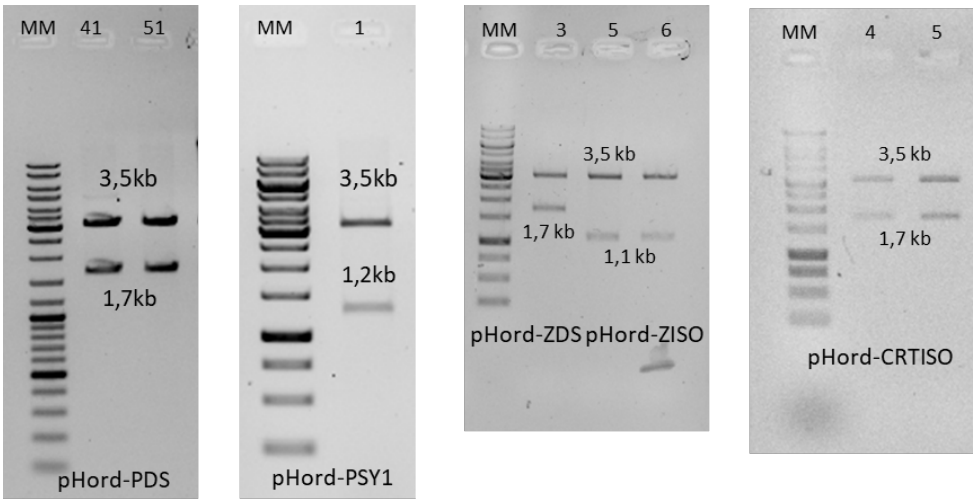


Figure 2. 4. Restriction enzyme analysis of the transformation vectors.

Table 2. 3. PCR primers and gene accession numbers. Restriction enzyme sequence for cloning marked in red.

Target site (Accession Numbers)	Primer	Primer sequence (5'→3')
OsPSY1 (1263bp) (FJ971175)	OsPSY1 Forward	TCTAGAATG GCGGCCATCACGCTCCTA
	OsPSY1 Reverse	GAATTCCTACTTCTGGCTATTTCTCAGTGA
OsPDS (1755bp) (tXM_015777614)	OsPDS Forward	TCTAGAATGGATACTGGCTGCCTGTCATC
	OsPDS Reverse	GAGCTCCTAGGAGGCAACAGGAACTTCA
Synthetic genes		
OsZISO (XM_015762760.1)	1016bp	Genes synthetized by Synbio Technologies, Inc. with the specific restriction enzyme sites (XbaI and EcoRI) for cloning.
OsZDS (XM_015791038.1)	1749bp	
OsCRTISO (XM_015761536.1)	1773bp	

2.4.2. Rice transformation

I co-transformed 7-day-old mature zygotic rice embryos with the above five constructs containing individual transgenes and the hpt-selectable marker gene. The *hpt* gene was expressed constitutively, and expression of the five carotenogenic genes was driven by the barley d-hordein endosperm-specific promoter. The transformation of rice was carried out following different strategies (**Figure 2. 5**), to achieve the combination of the minimum genes needed to reconstitute the carotenoid pathway in the endosperm. Transformed rice callus were analyzed by PCR to ascertain the genetic make-up of each independent line (**Figure 2. 6**).

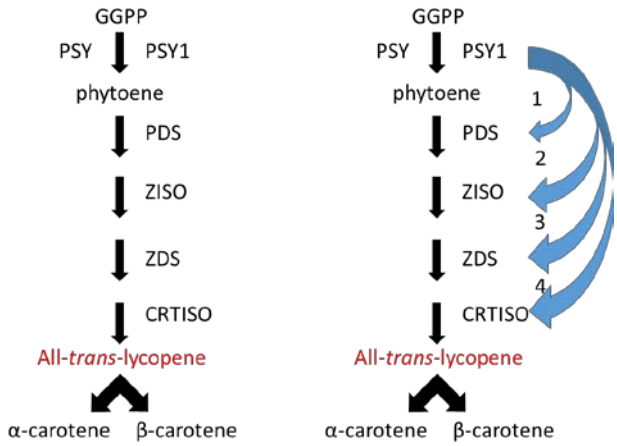


Figure 2. 5. Transformation strategies to elucidate the necessary genes required to accumulate carotenoids in rice endosperm.

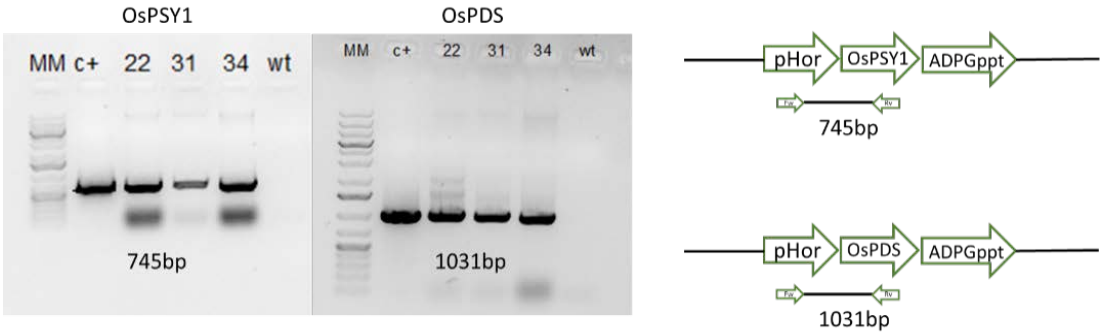
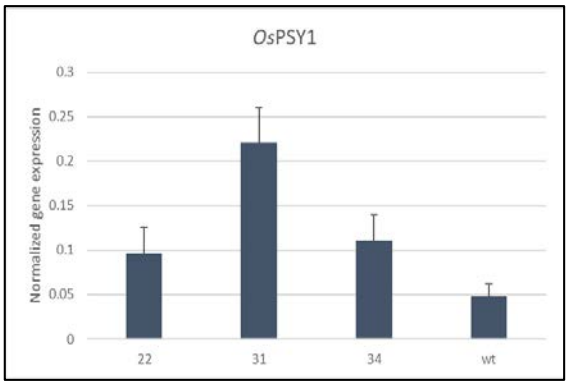


Figure 2. 6. PCR analysis confirming presence of input transgenes.

2.4.3. Expression analysis of the rice *PSY1, PDS* genes.

The expression of the *OsPSY1* and *OsPDS* genes was analyzed by quantitative real-time PCR as shown in **Figure 2. 7**. As expected, *Os*PSY1 relative expression was higher in all the callus lines compared with wild-type (wt) (*O. sativa* ssp. Japponica cv. Nipponbare). The relative expression of *Os*PSY1 in the highest-expressing line (line 31) was 3 fold higher than wt. In the case of *OsPDS* the relative expression was similar in lines 22, 34 and wt. The *OsPDS* relative expression in line 31 was 8 fold higher than wt.



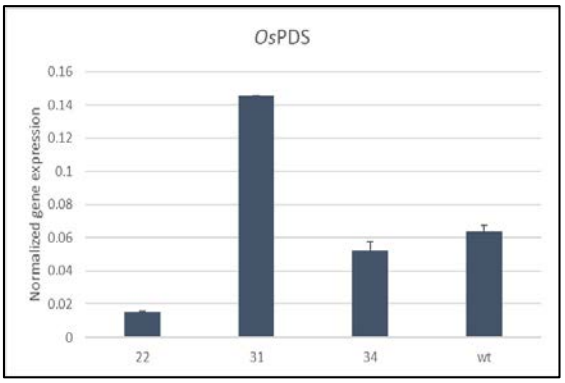


Figure 2. 7. mRNA accumulation of OsPSY1 and OsPDS in rice callus. Expression levels were normalized to the rice actin gene, as determined by qRT-PCR. Each value represents the average of three experiments and the error bars indicates the standard deviation.

2.4.4. Identification of minimum set of genes to reconstitute the rice carotenoid pathway in callus.

During selection the transgenic rice callus ranged in color from white through various shades of yellow to orange, presumably reflecting the expression of *OsPSY1* and *OsPDS* (**Figure 2. 8**). I found a precise correlation between the phenotypes and expressed transgenes (at the mRNA level) after three weeks in culture (**Figure 2. 8**). The WT callus did not have any changes in relative expression or color. Moreover, I analyze the carotenoid content to correlate the phenotype, transgene expression with carotenoid accumulation. The results show an increase in carotenoid levels in lines expressing *OsPSY1* and *OsPDS*. Most of the increase in carotenoid content was attributed to increases in the levels of β-carotene (**Figure 2. 10**).

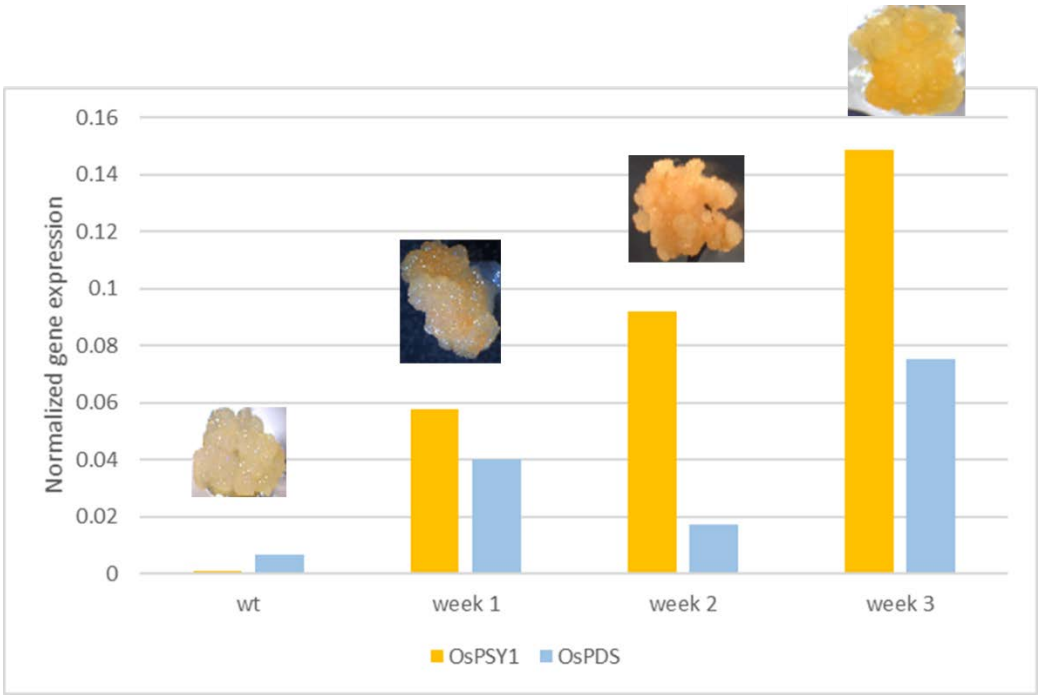


Figure 2. 8. Line #31 change in color and relative expression over a three-week period.

2.4.5. Carotenoid analysis in transgenic callus

Embryo-derived rice callus is white in color and accumulates only minimal levels of carotenoids (**Figure 2. 9**). The β-carotene content of lines #22, #31, #34 and wt were 156.7, 331.3, 318.7 and 72.02 µg/g Dry Weight (DW), respectively (**Figure 2. 10**). The highest-accumulating line (#31) produced more than 3 times β-carotene than the wt.

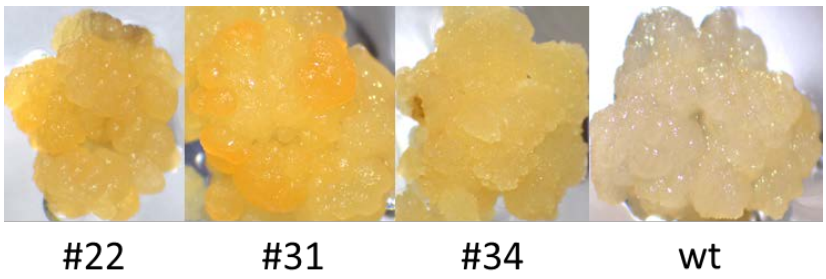


Figure 2. 9. Phenotype of transgenic rice callus expressing carotenogenic genes OsPSY1 and OsPDS. Lines #22, #31 and #34 are callus co-expressing OsPSY1 and OsPDS. wt, wild type callus.

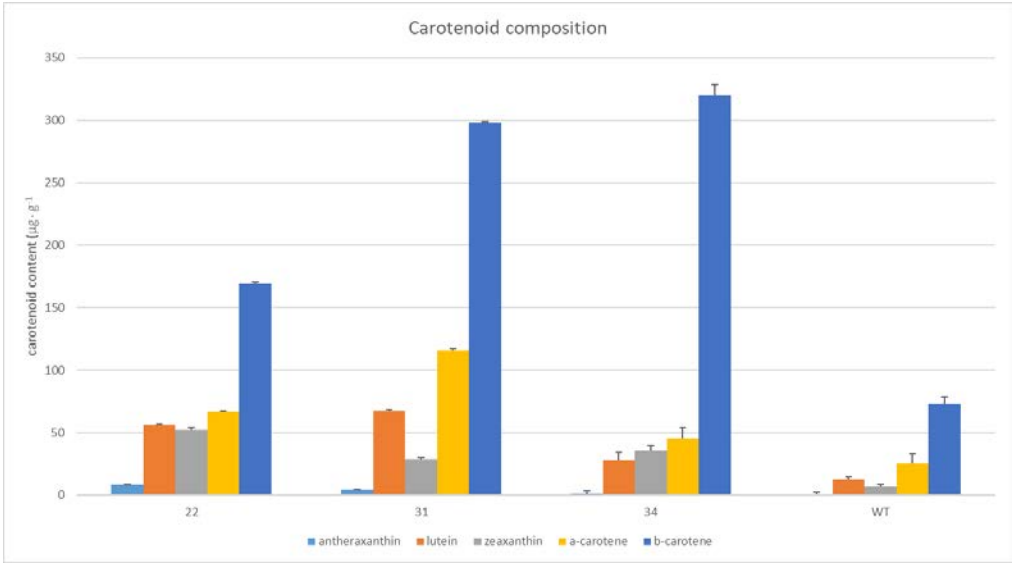


Figure 2. 10. Carotenoid levels in lines #22, #31, #34 expressing OsPSY1 and OsPDS.

2.5. Discussion

Isoprenoids are derived from two independent, conserved pathways, the mevalonate (MVA) and 2-C-methyl-D-erythritol 4-phosphate (MEP) pathways. Carotenoids are an important groups of isoprenoids. The first committed step in carotenoid biosynthesis is the condensation of two geranylgeranyl diphosphate (GGPP) molecules by phytoene synthase (PSY) to 15-cis-phytoene (Misawa et al 1994). 15-Cis-phytoene is converted into all-trans-lycopene by a poly-cis pathway involving two desaturases (phytoene desaturase or PDS and z-carotene desaturase or ZDS) and two isomerases (z-carotene isomerase or ZISO and prolycopene isomerase or CRTISO) (Li et al 2007; Chen et al 2010). Earlier literature reported that there was no expression of *PSY1* gene in rice endosperm (Schaub et al 2005). *PSY1* is the first committed step in plant carotenoid biosynthesis and is blocked in rice endosperm (Misawa et al 1994; Al-Babili et al 2006; Bai et al 2014, 2016). During rice seed development, the subsequent genes (*PDS*, *ZDS* and *CRTISO*) are expressed in the mature seed stage which indicates that some regulatory mechanism (at post-transcriptional/translational level or during localization in different subcellular compartments) affects enzyme synthesis/activity resulting in a carotenoid free endosperm (Schaub et al 2005; Chaudhary et al 2010). *OsPDS* was shown to play an

important role in rice carotenoid biosynthesis (Hable et al 1998; Schaub et al 2005; Ruiz-Sola and Rodríguez-Concepción 2012; Koschmieder et al 2017). The inhibition of PDS resulted in the production of bleached tissues, eventually followed by necrosis (Koschmieder et al 2017). There are several herbicides that inhibit this critical enzyme including norflurazon and diflufenican (Ruiz-Sola and Rodríguez-Concepción 2012). Moreover, OsPDS is needed to convert the products generated from OsPSY1 (Hable et al 1998; Schaub et al 2005). OsPDS catalyzes the introduction of two double bonds into 15-*cis*-phytoene, yielding 9,15-di-*cis*-phytofluene as intermediate and 9,15,9'-tri-*cis*-ζ-carotene as the end product (Ruiz-Sola and Rodríguez-Concepción 2012; Koschmieder et al 2017).

Our lab had previously demonstrated that combinatorial transformation can unravel bottlenecks in metabolic pathways in plants (Bai et al 2014, 2016). Our aim has been to define the minimum set of genes necessary to reconstitute the carotenoid pathway in rice endosperm. In order to achieve the aim, we co-transformed all the rice carotenoid biosynthetic genes (*OsPSY1*, *OsPDS*, *OsZISO*, *OsZDS*, *OsCRTISO*) driven by endosperm specific promoters in order to characterize the carotenoid biosynthesis pathway and generate a library with all the possible combination of carotenoid biosynthesis genes (**Figure 2. 5**).

In pilot experiments we co-transformed rice with *OsPSY1* and *OsPDS* in order to start with the minimum number of genes in the reconstitution of the pathway. The screening was established using PCR and a color phenotype in callus. In earlier way from our lab we established the correlation of phenotype with metabolite accumulation in rice callus (Bai et al 2014) (**Figure 2. 8** and **Figure 2. 9**). Metabolic engineering has been a trial and error approach where the elimination of one bottleneck often merely serves to reveal the next bottleneck. In contrast, combinatorial transformation generates a library of metabolic variants allowing the best strategy to be deduced in a single experiment if all necessary combinations of transgenes are represented in the population (Zhu et al 2008b; Li et al 2010; Bai et al 2014). We previously described a combinatorial gene platform based on white maize (M37W) endosperm, rice callus and endosperm for functional analysis of combinations of metabolic genes (Zhu et al 2018; Bai et al 2014). Similarly, to white maize endosperm, rice callus is also white and does not accumulate significant levels of carotenoids. The endosperm tissue can also be studied in more detail by quantitative analysis of all carotenoids, and this can be extended to any other metabolic pathway that

is also missing from endosperm tissue. The usefulness of the maize and rice platform in the functional characterization of metabolic genes is constrained by the time needed to regenerate transgenic plants carrying combinations of transgenes. Researchers have reported the use of callus tissue from rice, maize, *A. thaliana*, to characterize gene function and expression levels (Paine et al 2005; Bai et al 2014; Sun et al 2020).

The clear differences in color of white wild type rice callus versus the yellow-orange lines expressing *OsPSY1* and *OsPDS* suggested carotenoid accumulation (**Figure 2.9**). In other investigations callus color were used and demonstrated as an indicator for carotenoid accumulation (Bai et al 2014; Sun et al 2020). Our results are consistent with the experiments in rice callus expressing *ZmPSY1* and *PaCRTI*. Such callus were yellow-orange in color (Bai et al 2014). The results show that the relative expression of *OsPSY1* was more than two fold higher than wild type *OsPSY1* in rice callus. The *OsPDS* relative expression in rice callus was not significantly higher than in the wild type, similarly to other published reports (Paine et al 2005). In order to demonstrate that the combination of *OsPSY1* and *OsPDS* are the minimum genes required to reconstitute the native pathway to β-carotene, we performed a carotenoid content analysis. The results showed an increase in total carotenoids in the transgenic callus lines up contained up to 514 µg/g DW total carotenoids (**Figure 2. 10**) compared to wild type which accumulated 116 µg/g DW. Also β-carotene levels in rice callus increased two fold with accumulation up to 156 µg/g DW (**Figure 2. 10**) in transgenic lines compare to wild type callus (72.02 µg/g DW).

Our results suggest that the expression of *OsPSY1* and *OsPDS* are the minimum genes necessary to reconstitute the carotenoid pathway in rice callus. However, these preliminary results need to be confirmed in seeds and with the co-transformation of all the pathway genes. The carotenoid biosynthesis pathway is a highly regulated pathway (Rodríguez-Concepción 2010; Jin et al 2021). Although carotenoid biosynthesis in plants has been well investigated, extensive studies on its regulation are relatively limited. The pathway is not only highly regulated, but it is also characterized by multiple branches and competition for intermediate precursor molecules. These are some of the difficulties found in the synthesis of carotenoids. In the MVA pathway, *HMGR* had been demonstrated to be an important regulatory gene to maximize accumulation of carotenoids (Tian et al 2019). In the MEP pathway the overexpression of *DXS* demonstrated the capacity of carotenoid accumulation (You et al 2020). All the experiments were carried out in combination with well characterized genes such as *PSY1*

and *CRTI* (Paine et al 2005). In addition, there is a limited flux in the subsequent desaturation and isomerization reactions leading to lycopene. We used *OsPDS* to carry out the desaturation reaction and corroborate whether subsequent reactions lead to carotenoid accumulation. Our results suggest that the expression of *OsPSY1* and *OsPDS* are the minimum genes necessary to reconstitute the carotenoid pathway in rice callus. However, these preliminary results need to be confirmed in seeds and with the co-transformation of all the endogenous rice carotenogenic genes.

2.6. Conclusions

In summary, using a previously developed rice callus platform that enables functional characterization of candidate genes, we elucidated the minimum genes necessary to reconstitute the native pathway to β-carotene in rice callus. The overexpression of *OsPSY1* and *OsPDS* results in a yellow to orange callus phenotype. Carotenoid analysis demonstrated an increased accumulation of carotenoids, specifically a high accumulation of β-carotene. Our preliminary results suggest that the minimum genes necessary to reconstitute the carotenoid pathway in rice callus are *OsPSY1* and *OsPDS*. Co-transformation of the subsequent genes (*OsZDS*, *OsZISO* and *OsCRTISO*) will be important for the full understanding of the pathway and to have insights of gene regulation. Future work needs to focus on generating a population of transgenic rice plants with the combination of all the above genes. Once the population is in hand gene expression analysis and metabolite composition at the endosperm level can be carried out to better describe the pathway in rice endosperm.

2.7. References

Al-Babili S, Hoa TTC, Schaub P (2006). Exploring the potential of the bacterial carotene desaturase CrtI to increase the β-carotene content in Golden Rice. *J Exp Bot*., **57**:1007–1014

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

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

Baisakh N, Rehana S, Rai M, Oliva N, Tan J, Mackill DJ, Khush GS, Datta K, Datta SK (2006). Marker-free transgenic (MFT) near-isogenic introgression lines (NIILs) of 'golden' indica rice (cv. IR64) with accumulation of provitamin A in the endosperm tissue. *Plant Biotechnol J*., **4**:467–475

Burkhardt PK, Beyer P, Wünn J, Klöti A, Armstrong GA, Schledz M, von Lintig J, Potrykus I (1997). Transgenic rice (Oryza sativa) endosperm expressing daffodil (Narcissus pseudonarcissus) phytoene synthase accumulates phytoene, a key intermediate of provitamin A biosynthesis. *Plant J*., **11**:1071–1078

Chaudhary N, Nijhawan A, Khurana JP, Khurana P (2010). Carotenoid biosynthesis genes in rice: structural analysis, genome-wide expression profiling and phylogenetic analysis. *Mol Genet Genomics*., **283**:13–33

Chayut N, Yuan H, Ohali S, Meir A, Sa'ar U, Tzuri G, Zheng Y, Mazourek M, Gepstein S, Zhou X, Portnoy V, Lewinsohn E, Schaffer AA, Katzir N, Fei Z, Welsch R, Li L, Burger J, Tadmor Y (2017). Distinct Mechanisms of the ORANGE Protein in Controlling Carotenoid Flux. *Plant Physiol*., **173**:376–389

Chen Y, Li F, Wurtzel ET (2010). Isolation and Characterization of the Z-ISO Gene Encoding a Missing Component of Carotenoid Biosynthesis in Plants . *Plant Physiol*., **153**:66–79

Datta K, Baisakh N, Oliva N, Torrizo L, Abrigo E, Tan J, Rai M, Rehana S, Al-Babili S, Beyer P, Potrykus I, Datta SK (2003). Bioengineered 'golden' indica rice cultivars with β-carotene metabolism in the endosperm with hygromycin and mannose selection systems. *Plant Biotechnol J*., **1**:81–90

Dong OX, Yu S, Jain R, Zhang N, Duong PQ, Butler C, Li Y, Lipzen A, Martin JA, Barry KW, Schmutz J, Tian L, Ronald PC (2020). Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat Commun*., **11**:1–10

Enfissi EMA, Fraser PD, Lois L-M, 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

Estévez JM, Cantero A, Reindl A, Reichler S, León P (2001). 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in

plants. *J Biol Chem*., **276**:22901–22909

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

Giuliano G (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr Opin Biotechnol*., **44**:169–180

Ha S-H, Liang YS, Jung H, Ahn M-J, Suh S-C, Kweon S-J, Kim D-H, Kim Y-M, Kim J-K (2010). Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. *Plant Biotechnol J.*, **8**:928–938

Hable WE, Oishi KK, Schumaker KS (1998). Viviparous-5 encodes phytoene desaturase, an enzyme essential for abscisic acid (ABA) accumulation and seed development in maize. *Mol Gen Genet MGG*., **257**:167–176

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008). Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science (80-).*, **319**:330–333

Institute of Medicine (2001). Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc. The National Academies Press, Washington, DC

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, Switzerland).*, **10**:1456

Koschmieder J, Fehling-Kaschek M, Schaub P, Ghisla S, Brausemann A, Timmer J, Beyer P (2017). Plant-type phytoene desaturase: Functional evaluation of structural implications. *PLoS One*., **12**:e0187628

Li F, Murillo C, Wurtzel ET (2007). Maize Y9 Encodes a Product Essential for 15-cis-ζ-Carotene Isomerization. *Plant Physiol*., **144**:1181–1189

Li Q, Farre G, Naqvi S, Breitenbach J, Sanahuja G, Bai C, Sandmann G, Capell T, Christou P, Zhu C (2010). Cloning and functional characterization of the maize carotenoid isomerase and β-carotene hydroxylase genes and their regulation during endosperm maturation. *Transgenic Res*., **19**:1053–1068

Li Q, Sun Z, Li J, Zhang Y (2013). Enhancing beta-carotene production in Saccharomyces cerevisiae by metabolic engineering. *FEMS Microbiol Lett*., **345**:94–101

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*., **18**:3594–3605

Misawa N, Truesdale MR, Sandmann G, Fraser PD, Bird C, Schuch W, Bramley PM (1994). Expression of a Tomato cDNA Coding for Phytoene Synthase in *Escherichia coli*, Phytoene Formation *In Vivo* and *In Vitro*, and Functional Analysis of the Various Truncated Gene Products. *J Biochem*., **116**:980–985

MK You, 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 Biol*., **20**:167

Moreno JC, Mi J, Alagoz Y, Al-Babili S (2021). Plant apocarotenoids: from retrograde signaling to interspecific communication. *Plant J*., **105**:351–375

Nogueira M, Mora L, Enfissi EMA, Bramley PM, Fraser PD (2013). Subchromoplast sequestration of carotenoids affects regulatory mechanisms in tomato lines expressing different carotenoid gene combinations. *Plant Cell*., **25**:4560–4579

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL, Drake R (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat Biotechnol*., **23**:482–487

Parkhi V, Rai M, Tan J, Oliva N, Rehana S, Bandyopadhyay A, Torrizo L, Ghole V, Datta K, Datta SK (2005). Molecular characterization of marker-free transgenic lines of indica rice that accumulate carotenoids in seed endosperm. *Mol Genet Genomics*., **274**:325–336

Pulido P, Perello C, Rodriguez-Concepcion M (2012). New Insights into Plant Isoprenoid Metabolism. *Mol Plant*., **5**:964–967

Rodríguez-Concepción M (2010). Supply of precursors for carotenoid biosynthesis in plants. *Arch Biochem Biophys*., **504**:118–122

Ruiz-Sola MÁ, Rodríguez-Concepción M (2012). Carotenoid biosynthesis in Arabidopsis: a colorful pathway. *Arab B*., **10**:e0158–e0158

Saba-Mayoral A, Bassie L, Christou P, Capell T (2022). Development of a facile genetic transformation system for the Spanish elite rice paella genotype Bomba. Transgenic Res. https://doi.org/10.1007/s11248-022-00303-z

Schaub P, Al-Babili S, Drake R, Beyer P (2005). Why Is Golden Rice Golden (Yellow) Instead of Red? *Plant Physiol*., **138**:441–450

Sudhakar D, Fu X, Stoger E, Williams S, Spence J, Brown DP, Bharathi M, Gatehouse JA, Christou P (1998). Expression and immunolocalisation of the snowdrop lectin, GNA in transgenic rice plants. *Transgenic Res*., **7**:371–378

Sun T, Yuan H, Chen C, Kadirjan-Kalbach DK, Mazourek M, Osteryoung KW, Li L (2020). OR(His), a Natural Variant of OR, Specifically Interacts with Plastid Division Factor ARC3 to Regulate Chromoplast Number and Carotenoid Accumulation. *Mol Plant*., **13**:864–878

Tian Y-S, Wang B, Peng R-H, Xu J, Li T, Fu X-Y, Xiong A-S, Gao J-J, Yao Q-H (2019). Enhancing carotenoid biosynthesis in rice endosperm by metabolic engineering. *Plant Biotechnol J*., **17**:849–851

Tzuri G, Zhou X, Chayut N, Yuan H, Portnoy V, Meir A, Sa'ar U, Baumkoler F, Mazourek M, Lewinsohn E (2015). A 'golden'SNP in CmOr governs the fruit flesh color of melon (C ucumis melo). *Plant J.*, **82**:267–279

Valdez M, Cabrera-Ponce JL, Sudhakar D, Herrera-Estrella L, Christou P (1998). Transgenic Central American, West African and Asian elite rice varieties resulting from particle bombardment of foreign DNA into mature seed-derived explants utilizing three different bombardment devices. *Ann Bot.*, **82**:795–801

Xudong Y, Salim A-B, Andreas K, Jing Z, Paola L, Peter B, Ingo P (2000). Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science (80-).*, **287**:303–305

Yazdani M, Sun Z, Yuan H, Zeng S, Thannhauser TW, Vrebalov J, Ma Q, Xu Y, Fei Z, Van Eck J, Tian S, Tadmor Y, Giovannoni JJ, Li L (2019). Ectopic expression of ORANGE promotes carotenoid accumulation and fruit development in tomato. *Plant Biotechnol J.*, **17**:33–49

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 Biol.*, **20**:167

Zhou X, Welsch R, Yang Y, Álvarez D, Riediger M, Yuan H, Fish T, Liu J, Thannhauser TW, Li L (2015). Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc Natl Acad Sci.*, **112**:3558–3563

Zhu C, Farré G, Zanga D, Lloveras J, Michelena A, Ferrio JP, Voltas J, Slafer G, Savin R, Albajes R, Eizaguirre M, Lopez C, Cantero-Martínez C, Díaz-Gómez J, Nogareda C, Moreno JA, Angulo E, Estany J, Pena RN, et al (2018). High-carotenoid maize: development of plant biotechnology prototypes for human and animal health and nutrition. *Phytochem Rev.*, **17**:195–209

Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T (2008a). 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 C, Naqvi S, Gomez-Galera S, Pelacho AM, Capell T, Christou P (2007). Transgenic strategies for the nutritional enhancement of plants. *Trends Plant Sci.*, **12**:548–555

Chapter III

Identification and repair of *cis*-acting elements in the endogenous rice *PSY1* promoter to restore functionality in the endosperm

3.0.Abstract

There have been exhaustive efforts to characterize the molecular structure of genomic elements involved in transcriptional regulation of gene expression. *Cis*-acting elements are DNA sequences that regulate the transcription of one or more genes. *Trans*-acting factors are proteins that recognize and bind directly or indirectly *cis*-acting elements and regulate their activity. *Cis*-regulatory elements in promoters provide specific binding sites for corresponding transcription factors. The identification of promoter sequences and their specific elements are critical for the fine regulation of introduced transgenes. Seeds are often a preferred target for transgene expression, particularly if the goal is to accumulate a heterologous product that might interfere with vegetative growth at high concentrations, or to improve the nutritional quality of seeds used as staple foods. In rice endosperm the biosynthesis of carotenoids is blocked at the first enzymatic step (PSY1) which converts geranylgeranyl diphosphate (GGPP) into phytoene. The aim of this chapter was to identify and repair *cis*-acting elements in the rice *PSY1* gene promoter in order to restore gene functionality in the endosperm. We identified six potential *cis*-regulatory elements in the promoter region of *PSY1*: three prolamin p-boxes, one AACA motif, one GCN4-like motif and an Opaque 2 box (O2 box). Moreover, we constructed two synthetic promoters using corrected 6- and 4-motif *OsPSY1* gene promoters. The first promoter contained the 6-motif before mentioned and the 4-motif promoter contained three p-box and one AACA box. Also, we generated a wild type synthetic promoter using the full length flanking region. We used *GFP* as a reporter for all the constructs. We confirmed that *GFP* was highly expressed and the GFP protein was detected in the transgenic rice callus lines containing the corrected 4- and 6-motif *OsPSY1* promoter. These results confirm that the corrected *OsPSY1* promoter activated gene expression. Our results suggest that the use of *cis*-regulatory elements provide an opportunity to regulate carotenogenic gene expression in metabolically engineered plants.

3.1. Introduction

3.1.1. Cis-acting elements and their importance in modulating gene expression.

There have been exhaustive efforts to characterize the molecular structure of genomic elements involved in transcriptional regulation of gene expression. Transcription involves the initial interaction of RNA polymerase II with several elements within the genomic DNA. The transcription factor binding regions are located upstream of the coding region of a gene, and the DNA sequence that includes these regions is generally referred to as the gene promoter (Potenza et al 2004). Promoters are very important in the control of the overall spatiotemporal expression of a gene, either driving or repress the transcription. Over the years, many promoters have been isolated from a wide variety of organisms and used in plant genetic engineering (Venter 2007; Peremarti et al 2010; Mithra et al 2017).

Cis-acting elements are sequences on the DNA that regulate the transcription of one or more genes (Hernandez-Garcia and Finer 2014). *Trans*-acting factors are proteins that recognize and bind directly or indirectly *cis*-acting elements and regulate their activity. *Cis*-regulatory elements in promoters provide specific binding sites for corresponding transcription factors (Hernandez-Garcia and Finer 2014). One example of a c*is*-regulatory element is the G-box, an abundant, well-characterized, multifunctional plant promoter regulatory element (Zhang et al 2019). The G-box, was identified in the promoter region of a light-regulated gene encoding the small subunit of ribulose 1,5- bisphosphate carboxylase/oxygenase (RuBisCO) (Giuliano et al 1988). A G-box core element consists of a 6-bp DNA sequence 'CACGTG' with variations in the first and last nucleotides, which is minimally needed to recruit G-box-binding factors (GBFs) such as basic leucine zipper (bZIP) proteins or basic helix-loop-helix (bHLH) proteins (Heim et al 2003).

Tissue specific *cis*-elements were characterized and identified, which included synthetic constitutive promoter elements, inducible promoter elements, and tissue-specific promoter elements (Ali and Kim 2019). Endosperm specific *cis*-elements consist of prolamin box (P-box) and the GCN4-motif (Hammond-Kosack et al 1993; Juhász et al 2011) (Table 3. 1). Most of the cereal seed storage proteins gene promoters contain one GCN4 motif, two P-box, one skinhead 1 (Skn-1) and one GCAA motifs (Fauteux and Strömvik 2009; Juhász et al 2011). The Skn-1 motif (GTCAT) is a *cis*-regulatory element for endosperm-specific expression, similar to a DNA binding domain in *C. elegans* (Rupert et al 1998). The GCAA motif is a highly conserved DNA sequence in the promoters of storage protein genes found in rice and maize, suggesting the possibility of

their regulatory role in transcription (So and Larkins 1991; Takaiwa et al 1996). Alongside this element, the AACA/TA, ACGT and CCAAT motifs were also suggested to be involved in the regulation of prolamin gene expression (Hammond-Kosack et al 1993; Takaiwa et al 1996).

With the increases of plant genomic sequences available (Schnable et al. 2009; Schmutz et al. 2010; Brenchley et al. 2012) and the development and broad availability of bioinformatics tools (Lescot et al 2002; Bailey et al 2009; Lichtenberg et al 2010) the approach to characterize different regulatory elements are easier and faster. Different aspects, such as specificity, spacing, and copy number, of *cis*-regulatory elements with their specific transcription factors have been studied (Ali and Kim 2019). For example, in rice are already described different elements specific for leaves (ACE element), endosperm (G-box), and seeds (GA-responsive element) (Ding et al 2021). Minimal promoters containing different combinations of elements (DRE/CRT, ABRE, G-box, MYB, MYC, P-box) from different inducible promoters were also used to study the nature of the elements. Synthetic promoters have been crucial for basic studies of signaling and transcriptional activation. The specificity of synthetic promoters is supported by the interactions among *cis*-motifs and their corresponding *trans*-elements for enhanced transcription (Mehrotra et al 2011; Dey et al 2015). Corroboration of those promoters and their regulatory elements using transgene expression assays and nucleotide mutagenesis is important to understand their function (Hernandez-Garcia and Finer 2014). Several synthetic promoters have been tested for combination of specific *cis*-regulatory elements with an essential promoter or hybrids with different regulatory sequences (Ali and Kim 2019).

3.1.2. Transcription factors in endosperm-specific promoters in cereals.

Gene promoters are essential to understanding the overall regulation of gene expression in plants. The identification of promoter sequences and their specific elements are critical for the fine regulation of introduced transgenes (Hernandez-Garcia and Finer 2014). The majority of the promoters used in biotechnological applications are: (1) constitutive (unregulated and active *in vivo* throughout development in all tissues); (2) spatiotemporal (specific to tissue or a development phase); (3) inducible (regulated by the application of an abiotic or biotic stress); and (4) synthetic (contain defined regulatory elements specifically located upstream to or within promoter sequences)(Peremarti et al 2010; Hernandez-Garcia and Finer 2014; Shimada et al 2014). Spatiotemporal promoters

in seed, are seed-specific promoters isolated from genes with restricted expression during seed development. Seeds are often a preferred target for transgene expression, particularly if the goal is to accumulate a heterologous product that might interfere with vegetative growth at high concentrations or to improve the nutritional quality of seeds used as staple foods (Peremarti et al 2010). Cereal storage proteins are expressed only in the starchy endosperm during the mid- and late developmental stage of the grain. Several endosperm-specific promoters from rice have been isolated and characterized, most of which are glutelin promoters (Qu and Takaiwa 2004; Kawakatsu et al 2008; Qu et al 2008; Urriola and Rathore 2014), with 10 of the 15 glutelin promoters been examined in stably transformed transgenic rice plants (Kawakatsu and Takaiwa 2010). From promoters that confer strong expression in the aleurone and subaleurone layers (GluA and GluB) to promoters that confer direct expression throughout the whole endosperm (GluC and GluD).

Endosperm-specific expression of seed storage protein genes is regulated by interactions of multiple *cis*-acting elements in their promoters (Onodera et al 2001). Three main transcription factor families are involved in the endosperm-specific activation of the prolamin genes, the bZIP, the DOF and the MYB transcription factors (Juhász et al 2011). Basic leucine zipper (bZIP) transcription factors have wide binding specificity, binding the ACGT core sequences, GCN4 core (TGAGTCA) in wheat prolamins or the O2 recognition site in maize (TCCACGTAGA) (Izawa et al 1993) (Table 3. 1). Differences in expression patterns among glutelin genes have been shown to be caused by the arrangement or diversity of *cis*-elements determining endosperm specificity (Kawakatsu et al 2008). The first identified sequence which was considered to be necessary for endosperm-specific expression was located around 300 bp upstream of the transcription start site (Kreis et al 1985). The prolamin box (P-box) or endosperm element is located -300-bp in the promoter of seed storage proteins (Kawakatsu et al 2008; Qu et al 2008). The P-box is well conserved in seed storage protein gene promoters (Hartings et al 1990) and consists of two independent *cis*-elements, the P-box (P box; TG(T⁄C⁄A)AAAG) and the GCN4 motif (TGA(G⁄ C)TCA) (Hammond-Kosack et al 1993). Research demonstrates the importance of precise *cis*-elements (P-box, GCN4 motif, AACA motif) for the expression in endosperm (Chuan-Yin Wu et al 1998; Chuan-Yin Wu et al 1998; Wu et al 2000; Kawakatsu et al 2008; Jin et al 2019). Interestingly, mutations in the P box and ACGT motif did not change tissue specificity but caused 10-fold and 4-fold

reductions in expression, respectively, but when the AACA motif is mutated results in a complete loss in promoter activity (Wu et al 2000).

Maize Opaque 2 (O2) is a transcriptional regulator of seed specific protein gene expression (Hartings et al 1990). O2 was originally isolated from opaque maize kernels that show a reduction in mRNA expression levels and in zein proteins (Kodrzycki et al 1989). In barley (*Hordeum vulgare*), the regulatory mechanisms of seed specific protein genes consists of five *cis*-elements able to recognize eight transcription factors belonging to four families (bZIP of the Opaque-2 family, and the B3, DOF, and MYB proteins), which are all reported to be regulators of seed specific protein genes (Ravel et al 2014) (**Table 3. 1**).

3.1.3. Rice *PSY1* gene promoter function and its regulation in different tissues

In the past years, attention has turned to the regulation of carotenoid accumulation at multiple levels: transcriptional, post-transcriptional, post-translational, storage/degradation, and feedback regulation (Stanley and Yuan 2019). This has led to the discovery of numerous carotenoid regulatory mechanisms such as the post-translational regulation of *PSY* by *Or* (Lu et al 2006); Zhou et al 2015), the catabolism of carotenoids by CCDs and 9-cis-epoxycarotenoid dioxygenases (NCEDs) (Auldridge et al 2006; Ohmiya et al 2006; Vallabhaneni et al 2010).

In most plant taxa, *PSY* constitutes a small gene family of generally three members with differential transcriptional regulation (Walter et al 2015). In most plants, phytoene synthase is encoded by three paralog genes: a) *PSY1* associated with carotenoid accumulation in endosperm, and essential for maintaining leaf carotenoid content; b) *PSY2* plays an important role in controlling leaf carotenogenesis during greening;, and c) *PSY3* associated mostly with abiotic stress responses and also involved in root carotenogenesis (Gallagher et al 2004; Welsch et al 2008; Li, Vallabhaneni and Wurtzel 2008). *PSY* duplication has provided an opportunity for sub-functionalization whereby gene family members vary in tissue specificity. For example, *PSY1* and *PSY2* are regulated by light during photomorphogenesis and increased carotenoid biosynthesis during greening (Welsch et al 2000; Rodríguez-Villalón et al 2009). Also induced during carotenoid accumulation in flowers and fruits (Fraser et al 1999). In addition, induced *PSY3* expression was shown to provide ABA formation in roots (Welsch et al 2008; Li, Vallabhaneni and Wurtzel 2008). *PSY1* catalyzes the highly regulated,

frequently rate-limiting synthesis of the carotenoid biosynthetic pathway controlling the carbon flux (Álvarez et al 2016). Nonetheless, *PSY1* is also subjected to posttranscriptional control mechanisms.

In rice endosperm the biosynthesis of carotenoids is blocked at the first enzymatic step (PSY) which converts GGPP into phytoene (Bai et al 2016). There is also limited flux in the subsequent desaturation reaction, which generates lycopene. Researchers demonstrate differences in levels of mRNA of carotenoid biosynthetic genes in rice endosperm with *PSY1* mRNA the least abundant, 130-fold lower than in leaves, followed by mRNAs for *LCY* (25-fold), *HYD* (22-fold), *PDS* (10-fold), *ZDS* (8-fold), and *CRTISO* (4-fold). (Schaub et al. 2005**)**.

Table 3. 1. Cis-acting elements identified in the promoters of LMW glutenin gene types (Juhász et al. 2011).

Name	**DNA binding motif (5'-3')**	**Transcription factor**	**References**
(CA)n element	CAAACAC	bZIP	(Stålberg et al 1996)
CCAAT box	CCAAT	CBF	(Albani and Robert 1995)
G box like element	TGACGT	bZIP	(Menkens et al 1995)
GCAA motif1	GCAAAAGTG	bZIP	(Takaiwa et al 1996)
GCAA motif2	GCAAAAGTA	bZIP	(Juhász et al 2011)
GCN4 like motif 1	GTGAGTCAT	O2 bZIP	(Albani et al 1997)
GCN4 like motif 2	ATGAGTCAT	O2 bZIP	(Müller and Knudsen 1993)
GCN4 like motif 3	GTGTGACAT	O2 bZIP	(Albani et al 1997)
AACA/TA motif 1	TAACAA	R2R3 MYB	(Diaz et al 2002)
AACA/TA motif 2	AACAAA	R2R3 MYB	(Diaz et al 2002)
MYB1AT core	AAACCA	R2R3 MYB	(Abe et al 1997)
P-box 1	TGTAAAGT	PBF DOF	(Kreis et al 1985)
P-box 2	TGCAAAG	PBF DOF	(Sugiyama et al 1985)
P-box 3	TGTAAAG	PBF DOF	(Colot et al 1987)
P-box4	TGCAAAC	PBF DOF	(Norre et al 2002)
P-box5	TGCAAAAG	PBF DOF	(Norre et al 2002)
P-box7	TGTAAAAGT	PBF DOF	(Shirsat et al 1989)
RY core site	CATGCA	ABI3/VP1	(Suzuki et al 1997)
Skn-1 like motif	GTCAT	bZIP	(Blackwell et al 1994)
SPA bZIP	GATGACGTGTC	O2 bZIP	(Mena et al 1998)
TATA-box 1	CTATAAATA	TBP	(Bernard et al 2010)
TATA-box 2	GTATAAAG	TBP	(Bernard et al 2010)
TATA like motif	TATAA	TBP	(Bernard et al 2010)

DNA binding motif = recognition sites, Transcription factors = bZIP- basic leuzin zipper, CBF = CCAAT binding factor, O2-bZIP = TF similar to Opaque 2 transcription factor, R2R3 MYB = MYB transcription factor with two or three binding motif, PBF DOF = prolamin box binding factors of the DNA-binding with one finger domain transcription factor family, ABI3/VP1 = Abscisic acid insensitive3 or viviparous1 transcription factors, TBP = TATA binding protein.

3.2. Aims and Objectives

Aim: The aim of this chapter was to identify and repair cis-acting elements in the rice *PSY1* gene promoter in order to restore gene functionality in the endosperm.

Objectives:

- Construct a complete synthetic *OsPSY1* gene promoter by the *de novo* combination of six or 4 *cis*-elements driving *GFP*.
- Use the wheat low-molecular-weight glutenin (LMWG) promoter, which contains two P-boxes, two O2-boxes and a CAAT box as a positive control.
- Use the wild type *OsPSY1* promoter driving *GFP* as a negative control.

3.3. Materials and Methods

3.3.1. Cloning and sequencing the *OsPSY1* gene promoter.

The *OsPSY1* gene promoter sequence was cloned directly from rice leaf cDNA by PCR using *OsPSY1* forward and reverse primers (Table 3. 2). The primers were designed to flank the *OsPSY1* promoter region based on sequences in GenBank (accession number AP014962.1). The PCR product was transferred to vector pGEM-Teasy (Promega, Madison, WI, USA) to generate pGEM-*OsPSY1p* for sequencing.

Table 3. 2. Primers used to amplify the OsPSY1 gene promoter region.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
OsPSY1p	TGCAAGTACTCTAATAAGCTTTAGCTACCACTGCT	GGATCCTATTATACTGGATGGGATCAATCGCTGGCCGGAGA

OsPSY1p, rice (*Oryza sativa* ssp. *japonica*)) phytoene synthase 1 gene promoter; HindIII (5´-AAGCTTT-3´), BamHI (5´-GGATCC-3´) sites (underlined and red) were introduced into each primer, respectively.

3.3.2. Constructing and sequencing the corrected synthetic 4- and 6-motif *OsPSY1* promoter.

The corrected synthetic 4- and 6-motif *OsPSY1* promoter were synthesize by Synbio-technologies Inc. (NJ, USA) and cloned into pUC-57 vector.

3.3.3. Vector containing the low molecular weight glutenin (LMWG) promoter.

The p326 vector containing the LWM glutenin promoter and nos terminator was used as a positive control in subsequent experiments (Stoger et al., 1999).

3.3.4. Vector construction

Vectors pGEM-Wt*OsPSY1*-TNos, pGEM-6M*OsPSY1*-TNos, pGEM-4M*OsPSY*1-TNos, and p326 (LMWG-nosT) were digested with BamHII and HindIII, then were individually transferred to 35S-*GFP*-TNos (Addgene plasmid # 80127; http://n2t.net/addgene:80127; RRID: Addgene_80127), to generate wt*OsPSY1-GFP*-TNos, 6M*OsPSY1-GFP*-TNos, 4M*OsPSY1-GFP*-TNos and LMWG-*GFP*-TNos respectively. The integrity of all intermediate and final constructs was verified by sequencing.

3.3.5. Rice transformation

Rice transformation was performed as described in chapter 2.3.2

3.3.6. Gene expression analysis by quantitative real-time PCR

Gene expression analysis was performed as described in chapter 2.4.4. Primer combinations are listed in supporting information Table 3. 3.

Table 3. 3. Primers used for qRT-PCR analysis.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
GFP	GAAGCAGCACGACTTCTTCAA	TATAGACGTTGTGGCTGTTGTAGT
OsACTIN	GACTCTGGTGATGGTGTCAGC	TCATGTCCCTCACAATTTCC

GFP, enhance green fluorescence protein gene; *OsACTIN*, rice actin gene.

3.3.7. Protein extraction and western blot analysis

Total rice protein extracts were prepared by grinding 0.1–0.2 g callus or leaf tissue in liquid nitrogen and thawing the powder in 0.2–0.4 mL 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). The mixture was vortexed for 1 h at 4 C°. Cell debris were removed by centrifugation at 15,000 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). Total rice protein (80mg) were fractionated by denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) in

polyacrylamide gels containing 10% SDS at 200 V for 60 min, and then electro-transferred 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-GFP polyclonal antibody SAB4301138 (Sigma-Aldrich) diluted 1:2000 in 5% non-fat milk in TBST overnight at 4 C°, then rinsed three times for 10 min in TBST. The membranes were subsequently incubated with an alkaline phosphatase-conjugated goat anti-rabbit secondary antibody (Sigma-Aldrich) (diluted 1:5000 in 2% non-fat milk in TBS-T) for 1 h at room temperature followed by three 10 min rinses in TBST. Signals were detected using SIGMAFAST BCIP/NBT tablets (Sigma-Aldrich).

3.4. Results

3.4.1. Cloning and bioinformatics analysis of the wild-type *OsPSY1* promoter, 4- and 6-motif corrected *OsPSY1* promoter and LMWG promoter.

The *OsPSY1* gene promoter sequence was searched in the NCBI blast query (BLAST: Basic Local Alignment Search Tool (nih.gov)) using the sequence database of the *OsPSY1* gene located on chromosome 6. The 2.5-kb upstream 5′-flanking region (GenBank AP014962.1) was cloned from rice leaf genomic DNA (*O. sativa* Japonica cvv Nipponbare) (**Figure 3. 1**). The *OsPSY1* gene is not transcribed in the endosperm (Bai et al 2014), therefore, I searched the *OsPSY1* promoter region and identified *cis*-regulatory elements potentially responsible for the gene repression . I identified six potential *cis*-regulatory elements: three prolamin boxes (p-box), one AACA motif, one GCN4-like motif and an Opaque 2 box (O2 box) (Table 3. 5). The p-box, AACA motif, GNC4 motif, and O2 box motifs are known to be responsible for endosperm-specific expression in monocots (Fauteux and Strömvik 2009) and are likely to be involved in the endosperm-specific expression of *OsPSY1*. The 4- and 6-motif corrected *OsPSY1* gene promoter were developed using the target sequences shows in Table 3. 5Table 3. 4. The LMWG promoter was cloned into the expression vector containing the green fluorescence protein (*GFP*) and the Nos terminator as positive control. To test the synthetic promoters (4- and 6-motif

corrected) we used the reporter gene *GFP* and the Nos terminator (**Figure 3. 2**). All vectors were confirmed by restriction enzyme digestions and sequencing (**Figure 3. 3**).

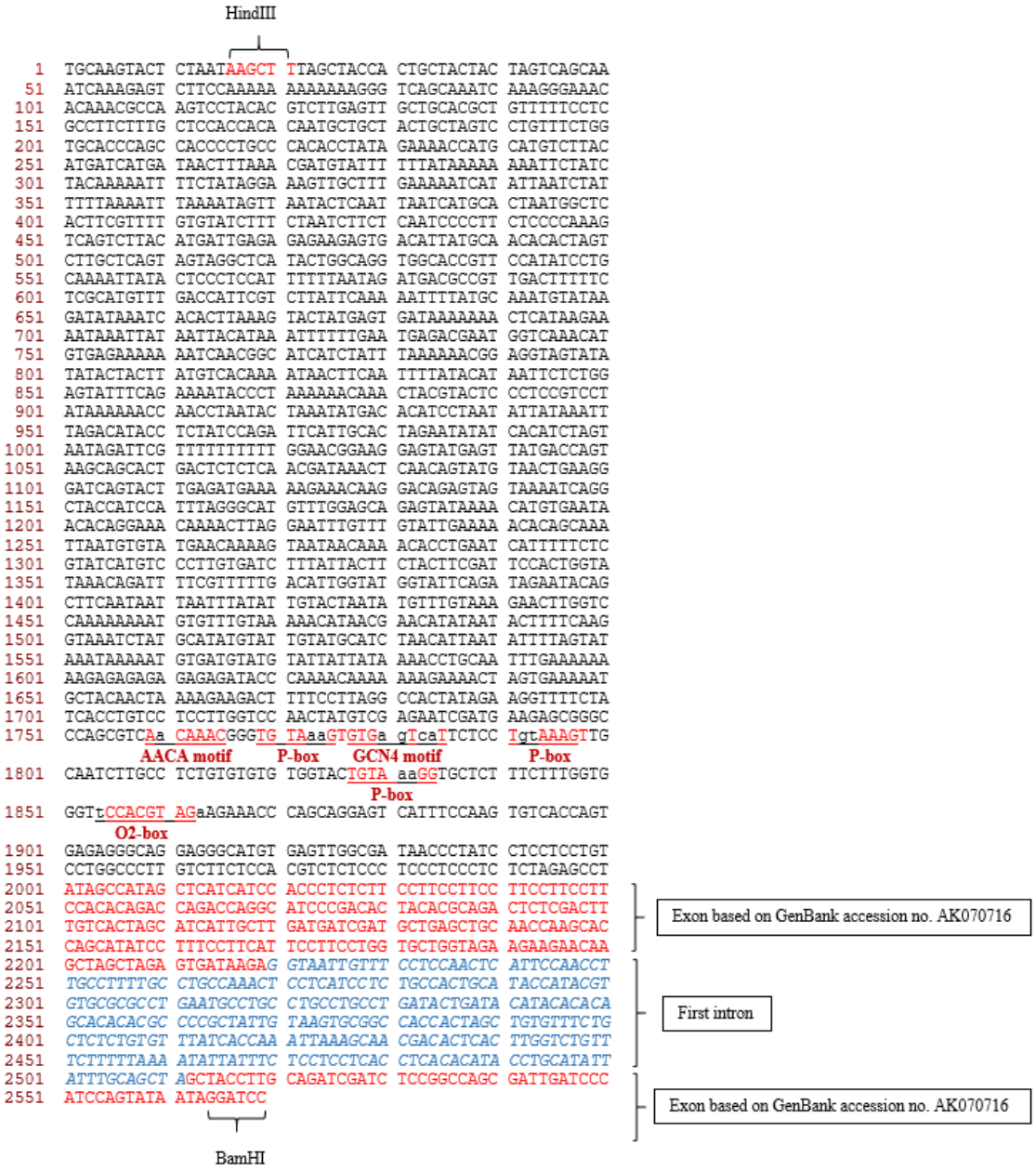


Figure 3. 1. Sequence of the rice (*O. sativa* Japonica cvv. Nipponbare) Phytoene synthase 1 gene (*OsPSY1*) promoter and part of the cDNA region (GenBank AP014962.1). The first intron (positions 2219– 2511) is in blue italics. The AACA motif, P-box, GCN4 motif and O2-box are in red and underlined.

Table 3. 5. Target sequences to be corrected for the 6 motif OsPSY1 gene promoter.

Target site	Sequence name	Original sequence (5'→3')	Corrected sequence (5'→3')
Target 1	AACA motif	ACCAAAC	AACAAAC
Target 2	(P-box) (-300 ELEMENT)	TGTATGGT	TGTAAAGT
Target 3	GCN4-like motif (GZM)	GTGTATTCT	GTGAGTCAT
Target 4	(P-box) (-300 ELEMENT)	TCCAAAGT	TGTAAAGT
Target 5	(P-box) (-300 ELEMENT)	TGTAGCGG	TGTAAAGG
Target 6	(O2-box)	GCCACGTAGG	TCCACGTAGA

Table 3. 4. Target sequences to be corrected for the 4 motif OsPSY1 gene promoter.

Target site	Sequence name	Original sequence (5'→3')	Corrected sequence (5'→3')
Target 1	AACA motif	ACCAAAC	AACAAAC
Target 2	(P-box) (-300 ELEMENT)	TGTATGGT	TGTAAAGT
Target 4	(P-box) (-300 ELEMENT)	TCCAAAGT	TGTAAAGT
Target 5	(P-box) (-300 ELEMENT)	TGTAGCGG	TGTAAAGG

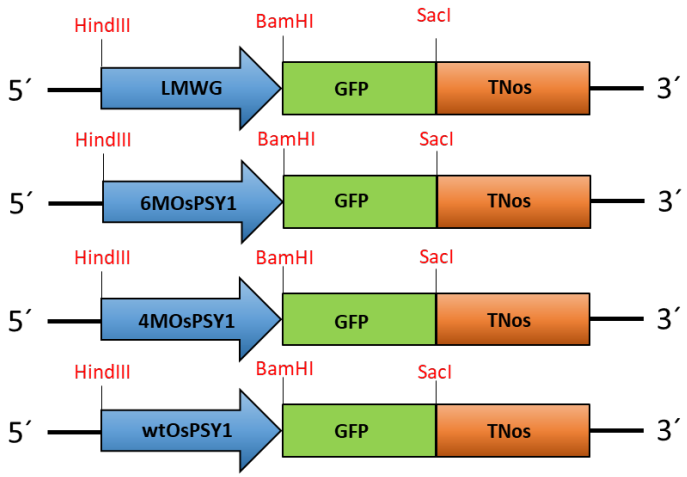


Figure 3. 2. Schematic representation of expression vectors LMWG-GFP-TNos, 6MOsPSY1-GFP-TNos, 4MOsPSY1-GFP-TNos and wtOsPSY1-GFP-TNos with the restriction enzymes used for cloning (marked in red).

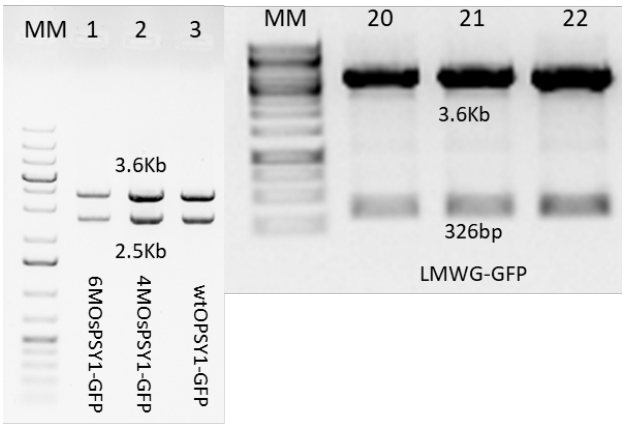


Figure 3. 3. Restriction enzyme digestion of transformation vectors (6MOsPSY1-GFP-TNos, 4MOsPSY1-GFP-TNos, wtOsPSY1-GFP-TNos, LMWG-GFP-TNos).

3.4.1. Molecular characterization of the synthetic wt, 4- and 6-motif *OsPSY1* promoter and LMWG promoter.

The constructs (6M*OsPSY1*-*GFP*-TNos, 4M*OsPSY1*-*GFP*-TNos, wt*OsPSY1*-*GFP*-TNos, LMWG-*GFP*-TNos) were introduced into rice each one independently by particle bombardment, and several independent transformants were analyzed by PCR. Callus lines were analyzed by PCR using primers overlapping the promoter region and the *GFP* transgene (**Figure 3. 4**).

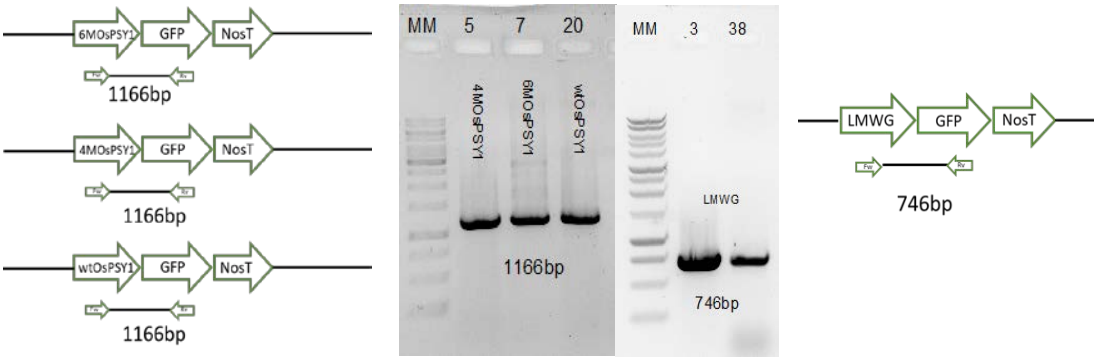


Figure 3. 4. Confirmation of 4MOspsy1, 6MOspsy1, wtOspsy1 and LMWG cassette insertion by PCR. Schematic representation showing PCR primer landing sites and amplicon sizes. PCR analysis on callus lines using primer pairs demonstrating gene presence.

3.4.2. Expression patterns of corrected synthetic 4- and 6-motif *OsPSY1* promoter using green fluorescence protein (GFP).

The expression patterns of the constructs (4M*OsPSY1*-*GFP*-TNos, 6M*OsPSY1*-*GFP*-TNos, wt*OsPSY1*-*GFP*-TNos and LMWG-*GFP*-TNos) were analyzed by quantitative real-time PCR (**Figure 3. 5** and **3. 6**). As expected, the synthetic 4- and 6-motif promoters

increased the relative expression of *GFP* compared with the wild-type promoter (wt*OsPSY1-GFP*-TNos). The GFP protein was detected using an anti-GFP antibody revealing protein bands with the expected molecular mass of ~27 kDa (**Figure 3. 5, 3. 6** and **3. 7**).

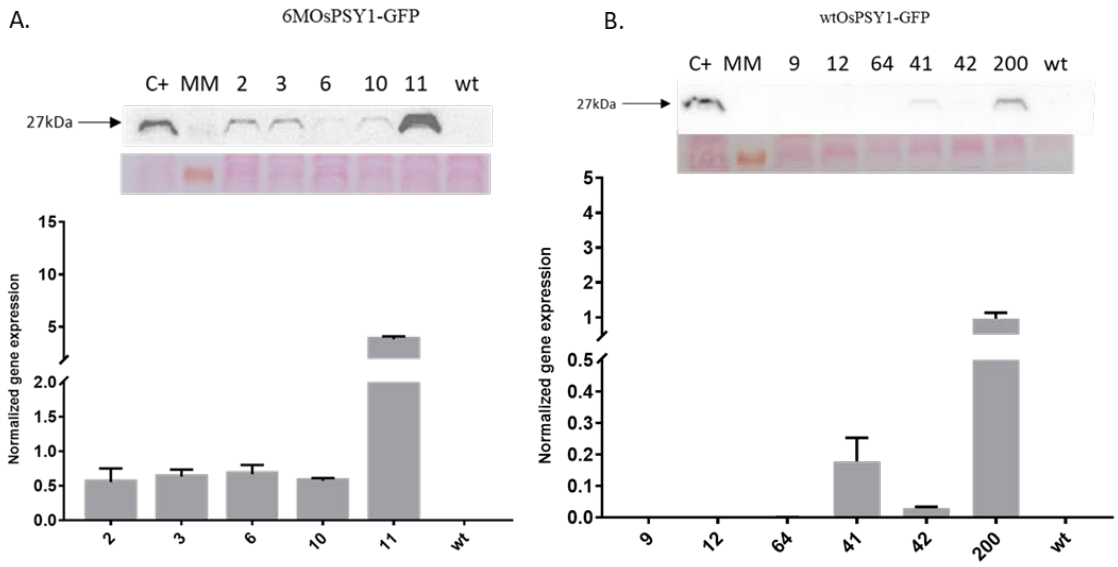


Figure 3. 5A and B. GFP transgene expression and western blot analysis in different rice callus lines with (A) 6MOsPSY1-GFP-TNos and (B) wtOsPSY1-GFP-TNos. Western blot analysis of rice callus using an anti-GFP antibody. (A) Lanes 2, 3. 6, 10, 11, 12 and 81 represent independent lines of 6MOsPSY1-GFP-TNos. (B) Lanes 9, 12. 64, 41, 42 and 200 represent independent lines of wtOsPSY1-GFP-TNos. Line wt = wild type callus rice (negative control). The expression of GFP was measured by qRT-PCR in rice callus. Expression levels were normalized against the rice actin gene. Each value represents the mean of three different biological replicates and error bars represent standard deviations.

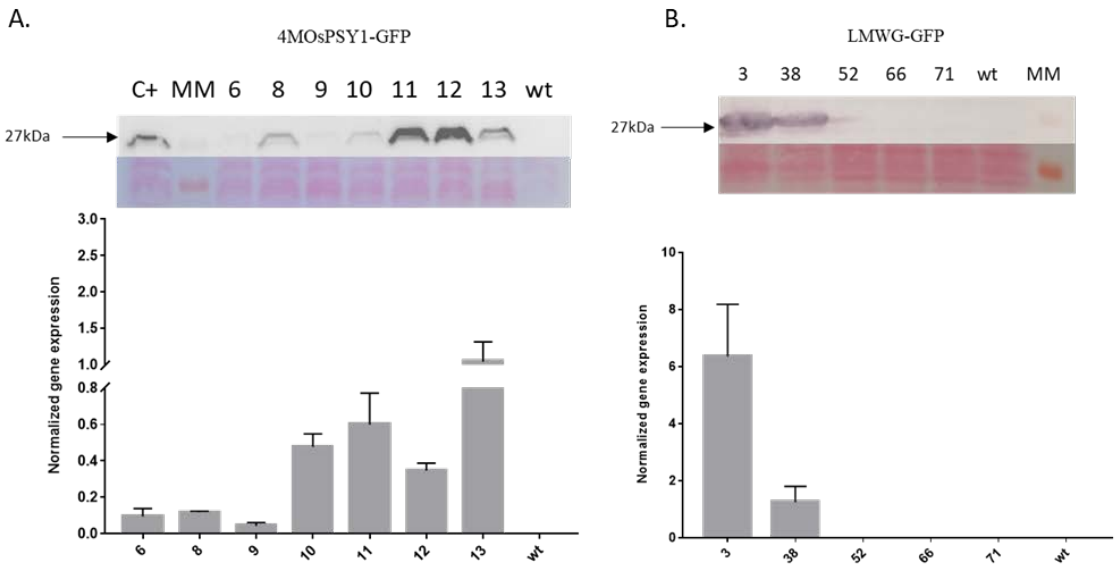


Figure 3. 6A and B. Analysis of GFP gene expression and western blot analysis in different rice callus lines with (A) 4MOsPSY1-GFP-TNos and (B) LMWG-GFP-TNos. Western blot analysis of rice callus using an anti-GFP antibody. (A) Lanes 6, 8. 9, 10, 11, 12 and 13 represent independent lines of 4MOsPSY1-GFP-TNos. (B) Lanes 3, 38. 52, 66 and 71 represent independent lines of LMWG-GFP-TNos. Line wt = wild type callus rice (negative control). The expression of GFP was measured by qRT-PCR in rice callus. Expression levels were normalized against the rice actin gene. Each value represents the means of three different biological replicates and error bars represent standard deviations.

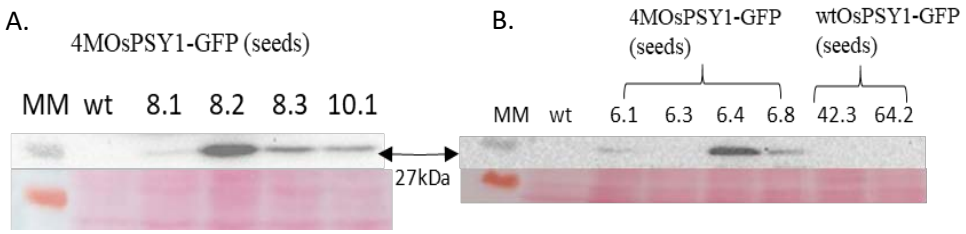


Figure 3. 7A and B. Western blot analysis of GFP in different rice seeds lines with (A) 4MOsPSY1-GFP-TNos, (B) 4MOsPSY1-GFP-TNos and wtOsPSY1-GFP-TNos. Western blot analysis of rice seeds using an anti-GFP antibody. (A) Lanes 8.1, 8.2, 8.3 and 10.1 represent independent lines of 4MOsPSY1-GFP-TNos. (B) Lanes 6.1, 6.3, 6.4 and 6.8 represent independent lines of 4MOsPSY1-GFP-TNos. (B) Lanes 42.3 and 64.2 represent independent lines of wtOsPSY1-GFP-TNos. Line wt = wild type callus rice (negative control).

3.5. Discussion

Accumulation of carotenoids occurs in most plant tissues; however, the types and amounts of carotenoid accumulation in green tissues are highly conserved. In contrast, carotenoid accumulation levels in non-green tissues such as flowers, fruits, and seeds, can be affected by many factors, such as developmental stage, abiotic and biotic factors etc. (Howitt and Pogson 2006; Zhu et al 2010; Giuliano 2017). Furthermore, a high accumulation of carotenoids is achieved based on the biosynthesis rate, cell storage capacity, and their catabolism and degradation rates (Farré et al 2010; Ruiz-Sola and Rodríguez-Concepción 2012; Giuliano 2017) (**Figure 2.1**). Phytoene synthase (PSY) is one of the most important enzyme in the carotenoid biosynthetic pathway (Rodríguez-Concepción 2010; Pulido et al 2012; Bai et al 2014; Giuliano 2017; Yazdani et al 2019). Different *PSY* paralogs are differentially expressed in different tissues based on abiotic and biotic stress conditions, such as drought, excessive light, salt, temperature, insects, fungi and post-transcriptional feedback control (Hable et al 1998; Ruiz-Sola and Rodríguez-Concepción 2012; Jin et al 2019). For example Arabidopsis contains only a single copy of *PSY*; however, there are two different splicing variants in Arabidopsis with different 5′ UTRs (Sun et al 2018a). The 5′ UTRs exhibit different translation efficiency to regulate the enzymatic activity of PSY (Álvarez et al 2016). Most plant species, such as tomato, rice, and maize, have three copies of *PSY*, with different expression patterns (Welsch et al 2008; Li, Vallabhaneni, Yu, et al 2008; Chaudhary et al 2010). Studies in Arabidopsis demonstrated that the expression level of *PSY* transcripts after salt or ABA treatment was higher in roots, whereas there was no significant expression in shoots (Ruiz-Sola and Rodríguez-Concepción 2012). In addition, during de-etiolation of Arabidopsis seedlings, *PSY* expression is up-regulated by the transcription factor phytochrome-interacting factor 1 (PIF1), and other members of this transcription factor family, such as bHLH (basic helix-loop-helix) (Toledo-Ortiz et al 2010; Nisar et al 2015; Sun et al 2018b).

With the increasing number of available genes and gene assembly technologies, there is a need to develop an efficient toolbox of promoters and *cis*-acting elements to control more complex agronomic traits, such as yield improvement and metabolic engineering. Considering that plants contain thousands of genes, it would not be surprising that many promoters and regulatory elements remain to be discovered (Hernandez-Garcia and Finer 2014). Transcription factor binding regions are located upstream of the coding region of

a gene, and the DNA sequence that includes these regions is generally referred to as the gene promoter (Potenza et al 2004). Promoters are very important in the control of the overall spatiotemporal expression of a gene. Over the years, many promoters have been isolated from a wide variety of organisms and have been used in plant genetic engineering (Venter 2007; Peremarti et al 2010; Mithra et al 2017). Some promoters are induced by a stimulus or a regulator, but the mechanism through which transcription is initiated is not fully understood. In other cases, mechanisms that affect transcription have been proposed, such as direct binding to a gene promoter by a transcription factor (Liu and Stewart 2016).

The transcriptional regulation of carotenoid biosynthetic genes in plants is a key intervention point to gain insights into metabolism. Almost every tissue type in every species studied to date appears to utilize a different group of transcriptional regulators for carotenoid biosynthesis (Giuliano 2017). Nonetheless, carotenoid biosynthesis is differentially regulated in various tissue types, as carotenoids are responsible for very different functions in different organs, e.g., in leaves as essential components of the photosynthetic apparatus vs. in flowers and fruits for coloration to attract pollinators (Giuliano et al 2008; Zhu et al 2010; Diretto et al 2019). Little is known about the transcriptional regulation of carotenoid biosynthesis in cereals or indeed in other plants.

A few transcription factors have been shown to influence carotenogenic gene expression in dicots. For example, transcription factors RAP2.2 [encodes a member of the ERF (ethylene response factor) subfamily B-2 of the plant-specific ERF/AP2 transcription factor family] and PIF1, specifically bind the *PSY* promoter and represses the expression of *PSY* and the accumulation of carotenoids in Arabidopsis seedlings (Welsch et al 2007; Toledo-Ortiz et al 2010). Nonetheless, HY5 (long hypocotyl 5), a PIF antagonist, directly binds a *cis*-element (G-box) in the promoter of *PSY* and promotes the accumulation of carotenoids associated with photosynthesis (Toledo-Ortiz et al 2014). Tomato RIN (ripening inhibitor) and BBX20, a MADS-box and B-box (BBX) zinc-finger transcription factors, promotes the accumulation of lycopene by directly regulating the expression of *PSY* (Martel et al 2011; Xiong et al 2019). Moreover, papaya CpNAC1 was shown to bind directly the NAC-binding site in CpPDS2/4 (phytoene desaturase) promoters (Fu et al 2017).

C*is*-regulatory elements in promoters provide specific binding sites for corresponding transcription factors (Hernandez-Garcia and Finer 2014). One example of a *cis*-regulatory

element is the G-box, an abundant, well-characterized, multifunctional plant promoter regulatory element (Zhang et al 2019). The G-box, was identified in the promoter region of a light-regulated gene encoding the small subunit of ribulose 1,5- bisphosphate carboxylase/oxygenase (RuBisCO) (Giuliano et al 1988). A G-box core element consists of a 6-bp DNA sequence 'CACGTG' with variations in the first and last nucleotides, which is minimally needed to recruit G-box-binding factors (GBFs) such as basic leucine zipper (bZIP) proteins or bHLH proteins (Heim et al 2003). In order to recruit the transcription factors for tissue-specific expression of the rice *PSY1* promoter we searched for *cis*-acting elements known to be present in the promoter region of proteins expressed in the endosperm.

Tissue specific *cis*-elements which include synthetic constitutive promoter elements, inducible promoter elements, and tissue-specific promoter elements have been identified and characterized (Ali and Kim 2019). Endosperm specific *cis*-elements include the prolamin box (P-box) and the GCN4-motif (Hammond-Kosack et al 1993; Juhász et al 2011) (Table 3.1). Most of the cereal seed specific promoters contain one GCN4 motif, two P-box, one skinhead 1 (Skn-1) and one GCAA motif (Fauteux and Strömvik 2009; Juhász et al 2011). The Skn-1 motif (GTCAT) is a *cis*-regulatory element for endosperm-specific expression, similar to a DNA binding domain in *C. elegans* (Rupert et al 1998). The GCAA motif is a highly conserved DNA sequence in the promoters of storage protein genes found in rice and maize, suggesting the possibility of their regulatory role in transcription (So and Larkins 1991; Takaiwa et al 1996). Alongside this element, the AACA/TA, ACGT and CCAAT motifs were also suggested to be involved in the regulation of prolamin gene expression (Hammond-Kosack et al 1993; Takaiwa et al 1996). The P-box is a highly conserved 7-bp sequence [5′-TG(T/A/C)AAA(G/A)-3′] found in the promoters of many cereal seed storage protein genes (Diaz et al 2002; Biłas et al 2016).

We focused on the regulation of *OsPSY1* because this gene is the first committed step in the rice biosynthetic pathway (Rodríguez-Concepción 2010; Pulido et al 2012; Bai et al 2014; Giuliano 2017; Yazdani et al 2019). The *cis*-regulatory elements and corresponding transcription factors which control the endosperm-preferred expression of *OsPSY1* have not yet been identified. We hypothesized that important regulatory elements would be located in the upstream promoter region. Moreover, we searched the *OsPSY1* promoter region and identified the classical *cis*-regulatory elements with differences in a number

of nucleotides of the classical sequence. For example, the P-box classical sequence is TGTAAAGT and using bioinformatics tools we identified a variant of the P-box (TGTATGGT) in the promoter region of *PSY1*. This finding suggests that the mutations in the P-box sequence might be responsible for the lack of expression of *PSY1* in the endosperm. We identified six potential *cis*-regulatory elements in the promoter region of *PSY1*: three prolamin p-boxes, one AACA motif, one GCN4-like motif and an Opaque 2 box (O2 box) (**Table 3. 4**). Literature reports suggest that the insertion or changes of *cis*-regulatory elements in the promoter sequence can regulate gene expression (Rodríguez-Leal et al 2017; Jin et al 2019; Zhang et al 2019).

CRISPR/Cas9 was used to recreate a known fruit size quantitative trait locis (QTL) in tomato by targeting *cis*-regulatory motifs in the 15-bp repressor element downstream of tomato WUS (SlWUS), a conserved homeobox gene that promotes stem cell proliferation (Rodríguez-Leal et al 2017). In maize, nested deletion analysis revealed that two *cis*-regulatory elements (AACA motif and P-box) determine the activity of the *ZmBCH2* promoter in seeds (Jin et al 2019). In soybean manipulation of the upstream 2- to 4-bp flanking sequence of G-boxes in a relatively weak soybean glycinin promoter (GmScream3) led to significant increases in the activity of the promoter when the first was converted into a classical G-box (CACGTG) (Zhang et al 2019).

We constructed two synthetic promoters using 4- and 6-motif corrected *OsPSY1* gene promoters (**Tables 3. 4 and 3. 5**). We generated a wild type synthetic promoter using the full length flanking region spanning from position –2500 in the *OsPSY1* upstream promoter as negative control (wt*OsPSY1*-*GFP*-TNos), which includes the first intron (**Figure 3.1**). As positive control we used an endosperm specific LMWG promoter which has three different variants of a GCN4 like motif and seven different p-box variants of (Hammond-Kosack et al 1993; Juhász et al 2011). We used *GFP* as a reporter for all the constructs (**Figure 3.2**). The use of GFP as a reporter has already been demonstrated for the analysis of *cis*-regulatory elements in gene regulation (Zhang et al 2019). Among the four *cis*-regulatory elements found in the 4-motif corrected *OsPSY1* promoter, three P-boxes and a single AACA motif were located within the first 300 bp upstream of the transcription start site, often the key regulatory region in endosperm-specific genes.

Having developed and introduced into rice the constructs for the activation of the *OsPSY1* promoter, we analyzed gene expression and protein accumulation. We

confirmed that the reporter gene (*GFP*) was highly expressed and the GFP protein was detected in the transgenic rice callus lines containing the corrected *OsPSY1* promoter 4- and 6-motif (**Figure 3.5A, 3.6A, B**). Taken together, these results confirm that the corrected *OsPSY1* promoter activated gene expression compared to the wild type *OsPSY1* promoter. The *cis*-regulatory elements have the potential to regulate *OsPSY1*, and when present they can activate the *OsPSY1* promoter, strongly supporting their role in the regulation of *OsPSY1* in rice endosperm. As expected, the construct containing the wild type *OsPSY1* gene promoter did not significantly expresses *GFP* in rice callus and endosperm (**Figure 3.5B, 3.7B**). It was demonstrated that *OsPSY1* was not expressed in rice endosperm (Schaub et al 2005) suggesting that the P-box and AACA motif are required for *OsPSY1* promoter activity in rice endosperm. We also investigated the regulatory potential of a synthetic 4-motif *OsPSY1* promoter in transgenic rice endosperm expressing *GFP*. Rice transformed with the synthetic 4-motif *OsPSY1* promoter accumulate GFP in the endosperm (**Figure 3.7A, B**), suggesting that the *cis*-regulatory elements are essential and exhibit tissue specific activation as demonstrated in other reports (Juhász et al 2011; Ali and Kim 2019; Jin et al 2019; Zhang et al 2019). Studies in rice, maize, wheat and soybean demonstrated the capacity of the *cis*-elements to regulate gene expression (Juhász et al 2011; Liu et al 2018; Jin et al 2019; Zhang et al 2019). In maize six *cis*-elements (RY repeats, GCN4, the p-box, Skn-1, and the ACGT and AACA motifs) were fused to the *ZmBD1* promoter to generate a putative bidirectional promoter (Liu et al 2018), resulting in endosperm- specific expression and high accumulation of anthocyanin in the endosperm. In rice, experiments investigated tissue specificity of the *cis*-elements P-box and AACA (Jin et al 2019). By constructing a series of 5′ deletions in the ZmBCH2 promoter region, investigators identified the minimum region needed for endosperm expression. Four cis-regulatory elements were found in the minimum promoter. A single P-box and a single AACA motif were located within the first 300 bp upstream of (Jin et al 2019). Tissue specificity was monitored by using a reporter gene (*gusA*).

In soybean, a synthetic promoter was developed using different repeats of G-box motifs (Zhang et al 2019). Using a relatively weak soybean glycinin promoter (GmScream3) and manipulating the sequence to obtain different repeats of G-box motif led to a significant increase in GFP expression. The results demonstrated the capacity of the G-box to activate gene expression. Our results using a synthetic *OsPSY1* promoter for

tissue-specific expression of *GFP* are similar to those obtained in rice and soybean by these investigators in terms of the tissue specific expression and generation of a synthetic promoter with endosperm specific *cis*-elements.

In the absence of the *cis*-transcription factors, GFP accumulation in the transgenic rice lines expressing the wild type *OsPSY1* promoter (wt*OsPSY1*-*GFP*-TNos) was low but detectable in some callus lines (**Figure 3.5B**), probably reflecting the insertion of the cassette near a strong enhancer that might affect the promoter in the same way as an enhancer trap (Springer 2000). However, the same rice lines with wild type *OsPSY1* promoter (wt*OsPSY1*-*GFP*-TNos) did nor accumulated *GFP* in rice endosperm (**Figure 3.7B**), as expected because *OsPSY1* is not expressed in rice endosperm (Schaub et al 2005; Bai et al 2014).

3.5. Conclusions

We demonstrate experimentally that rice phytoene synthase 1 (*OsPSY1*) is regulated by a combination of *cis*-acting elements in the upstream promoter region of the gene. Our results suggest that the expression of *OsPSY1* in the endosperm is dependent on the tissue-specificity of *cis*-regulatory elements. We provide experimental evidence that tissue specific *cis*-elements can regulate *GFP* expression in rice endosperm (**Figure 3.7A, B**). The combination of AACA, P-box, GCN4-like and O2 motifs in the synthetic *OsPSY1* promoter (4-and 6-motif corrected synthetic promoter versions) determined the activity of the *OsPSY1* promoter. These results are consistent with the results obtained in rice and soybean using *cis*-regulatory elements to regulate the expression of reporter genes (Jin et al 2019; Zhang et al 2019).

We identified the sequence of candidate *cis*-regulatory elements based on known regulatory mechanisms in the promoters of other cereal genes and confirmed that the corresponding *cis*-elements (P-box and AACA motif) are required for rice endosperm-specific expression. In transgenic rice plants expressing the corrected 4-motif *OsPSY1* promoter driving *GFP* the recombinant protein accumulated in the endosperm. In contrast, no GFP accumulation was seen when the wild type *OsPSY1* (synthetic) promoter was used. Our results suggest that the use of *cis*-regulatory elements provide an opportunity for the regulation of carotenogenic gene expression in metabolically engineered plants.

3.6. References

Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997). Role of Arabidopsis MYC and MYB homologs in drought-and abscisic acid-regulated gene expression. *Plant Cell*., **9**:1859–1868

Albani D, Hammond-Kosack MC, Smith C, Conlan S, Colot V, Holdsworth M, Bevan MW (1997). The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell*., **9**:171–184

Albani D, Robert LS (1995). Cloning and characterization of a Brassica napus gene encoding a homologue of the B subunit of a heteromeric CCAAT-binding factor. *Gene*., **167**:209–213

Ali S, Kim W-C (2019). A Fruitful Decade Using Synthetic Promoters in the Improvement of Transgenic Plants . Front. Plant Sci. 10

Álvarez D, Voß B, Maass D, Wüst F, Schaub P, Beyer P, Welsch R (2016). Carotenogenesis Is Regulated by 5'UTR-Mediated Translation of Phytoene Synthase Splice Variants. *Plant Physiol*., **172**:2314–2326

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

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

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

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009). MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res*., **37**:W202–W208

Bernard V, Brunaud V, Lecharny A (2010). TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genomics*., **11**:1–15

Biłas R, Szafran K, Hnatuszko-Konka K, Kononowicz AK (2016). Cis-regulatory elements used to control gene expression in plants. *Plant Cell, Tissue Organ Cult*., **127**:269–287

Blackwell TK, Bowerman B, Priess JR, Weintraub H (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. *Science (80-)*., **266**:621–628

Brenchley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, Allen AM, McKenzie

N, Kramer M, Kerhornou A, Bolser D, Kay S, Waite D, Trick M, Bancroft I, Gu Y, Huo N, Luo M-C, Sehgal S, Gill B, et al (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*., **491**:705–710

Chaudhary N, Nijhawan A, Khurana JP, Khurana P (2010). Carotenoid biosynthesis genes in rice: structural analysis, genome-wide expression profiling and phylogenetic analysis. *Mol Genet Genomics*., **283**:13–33

Colot V, Robert LS, Kavanagh TA, Bevan MW, Thompson RD (1987). Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J*., **6**:3559–3564

Dey N, Sarkar S, Acharya S, Maiti IB (2015). Synthetic promoters in planta. *Planta*., **242**:1077–1094

Diaz I, Vicente-Carbajosa J, Abraham Z, Martínez M, Isabel-La Moneda I, Carbonero P (2002). The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. *Plant J*., **29**:453–464

Ding Y, Zhu J, Zhao D, Liu Q, Yang Q, Zhang T (2021). Targeting Cis-Regulatory Elements for Rice Grain Quality Improvement . Front. Plant Sci. 12

Diretto G, Jin X, Capell T, Zhu C, Gomez-Gomez L (2019). Differential accumulation of pelargonidin glycosides in petals at three different developmental stages of the orange-flowered gentian (Gentiana lutea L. var. aurantiaca). *PLoS One*., **14**:e0212062

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

Fauteux F, Strömvik M V (2009). Seed storage protein gene promoters contain conserved DNA motifs in Brassicaceae, Fabaceae and Poaceae. *BMC Plant Biol*., **9**:1–11

Fraser PD, Kiano JW, Truesdale MR, Schuch W, Bramley PM (1999). Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid synthesis in ripening fruit. *Plant Mol Biol*., **40**:687–698

Fu C-C, Han Y-C, Kuang J-F, Chen J-Y, Lu W-J (2017). Papaya CpEIN3a and CpNAC2 Co-operatively Regulate Carotenoid Biosynthesis-Related Genes CpPDS2/4, CpLCY-e and CpCHY-b During Fruit Ripening. *Plant Cell Physiol*., **58**:2155–2165

Gallagher CE, Matthews PD, Li F, Wurtzel ET (2004). Gene Duplication in the Carotenoid Biosynthetic Pathway Preceded Evolution of the Grasses. *Plant Physiol*., **135**:1776–1783

Giuliano G (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr Opin Biotechnol*., **44**:169–180

Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR (1988). An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc Natl Acad Sci*., **85**:7089 LP – 7093

Giuliano G, Tavazza R, Diretto G, Beyer P, Taylor MA (2008). Metabolic engineering of carotenoid biosynthesis in plants. *Trends Biotechnol*., **26**:139–145

Hable WE, Oishi KK, Schumaker KS (1998). Viviparous-5 encodes phytoene desaturase, an enzyme essential for abscisic acid (ABA) accumulation and seed development in maize. *Mol Gen Genet MGG*., **257**:167–176

Hammond-Kosack MC, Holdsworth MJ, Bevan MW (1993). In vivo footprinting of a low molecular weight glutenin gene (LMWG-1D1) in wheat endosperm. *EMBO J*., **12**:545–554

Hartings H, Lazzaroni N, Marsan PA, Aragay A, Thompson R, Salamini F, Di Fonzo N, Palau J, Motto M (1990). The b-32 protein from maize endosperm: characterization of genomic sequences encoding two alternative central domains. *Plant Mol Biol*., **14**:1031–1040

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC (2003). The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol Biol Evol*., **20**:735–747

Hernandez-Garcia CM, Finer JJ (2014). Identification and validation of promoters and cis-acting regulatory elements. *Plant Sci*., **217**–**218**:109–119

Howitt CA, Pogson BJ (2006). Carotenoid accumulation and function in seeds and non-green tissues. *Plant Cell Environ*., **29**:435–445

Izawa T, Foster R, Chua N-H (1993). Plant bZIP protein DNA binding specificity. *J Mol Biol*., **230**:1131–1144

Jin X, Bai C, Bassie L, Nogareda C, Romagosa I, Twyman RM, 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

Juhász A, Makai S, Sebestyén E, Tamás L, Balázs E (2011). Role of Conserved Non-Coding Regulatory Elements in LMW Glutenin Gene Expression. *PLoS One*., **6**:e29501

Kawakatsu T, Takaiwa F (2010). Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol J*., **8**:939–953

Kawakatsu T, Yamamoto MP, Hirose S, Yano M, Takaiwa F (2008). Characterization of a new rice glutelin gene GluD-1 expressed in the starchy endosperm. *J Exp Bot*., **59**:4233–4245

Kodrzycki R, Boston RS, Larkins BA (1989). The opaque-2 mutation of maize differentially reduces zein gene transcription. *Plant Cell*., **1**:105–114

Kreis M, Forde BG, Rahman S, Miflin BJ, Shewry PR (1985). Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol*., **183**:499–502

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 Res*., **30**:325–327

Li F, Vallabhaneni R, Wurtzel ET (2008). PSY3, a new member of the phytoene synthase gene family conserved in the Poaceae and regulator of abiotic stress-induced root carotenogenesis. *Plant Physiol*., **146**:1333–1345

Li F, Vallabhaneni R, Yu J, Rocheford T, Wurtzel ET (2008). The maize phytoene synthase gene family: overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. *Plant Physiol*., **147**:1334–1346

Lichtenberg J, Kurz K, Liang X, Al-ouran R, Neiman L, Nau LJ, Welch JD, Jacox E, Bitterman T, Ecker K, Elnitski L, Drews F, Lee SS, Welch LR (2010). WordSeeker: concurrent bioinformatics software for discovering genome-wide patterns and word-based genomic signatures. *BMC Bioinformatics*., **11**:S6

Liu W, Stewart CN (2016). Plant synthetic promoters and transcription factors. *Curr Opin Biotechnol*., **37**:36–44

Liu X, Li S, Yang W, Mu B, Jiao Y, Zhou X, Zhang C, Fan Y, Chen R (2018). Synthesis of Seed-Specific Bidirectional Promoters for Metabolic Engineering of Anthocyanin-Rich Maize. *Plant Cell Physiol*., **59**:1942–1955

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*., **18**:3594–3605

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

Mehrotra R, Gupta G, Sethi R, Bhalothia P, Kumar N, Mehrotra S (2011). Designer promoter: an artwork of cis engineering. *Plant Mol Biol*., **75**:527–536

Mena M, Vicente-Carbajosa J, Schmidt RJ, Carbonero P (1998). An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *Plant J*., **16**:53–62

Menkens AE, Schindler U, Cashmore AR (1995). The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends Biochem Sci*., **20**:506–510

Mithra SVA, Kulkarni K, Srinivasan R (2017). Plant Promoters: Characterization and Applications in Transgenic Technology BT - Plant Biotechnology: Principles and Applications. In: Abdin MZ, Kiran U, Kamaluddin, Ali A (eds). Springer Singapore, Singapore, pp 117–172

Müller M, Knudsen S (1993). The nitrogen response of a barley C-hordein promoter is

controlled by positive and negative regulation of the GCN4 and endosperm box. *Plant J.*, **4**:343–355

Nisar N, Li L, Lu S, Khin NC, Pogson BJ (2015). Carotenoid metabolism in plants. *Mol Plant.*, **8**:68–82

Norre F, Peyrot C, Garcia C, Rancé I, Drevet J, Theisen M, Gruber V (2002). Powerful effect of an atypical bifactorial endosperm box from wheat HMWG-Dx5 promoter in maize endosperm. *Plant Mol Biol.*, **50**:699–712

Ohmiya A, Kishimoto S, Aida R, Yoshioka S, Sumitomo K (2006). Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals. *Plant Physiol.*, **142**:1193–1201

Onodera Y, Suzuki A, Wu C-Y, Washida H, Takaiwa F (2001). A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. *J Biol Chem.*, **276**:14139–14152

Peremarti A, Twyman RM, Gómez-Galera S, Naqvi S, Farré G, Sabalza M, Miralpeix B, Dashevskaya S, Yuan D, Ramessar K, Christou P, Zhu C, Bassie L, Capell T (2010). Promoter diversity in multigene transformation. Plant Mol. Biol. 73:363–378

Potenza C, Aleman L, Sengupta-Gopalan C (2004). Targeting transgene expression in research, agricultural, and environmental applications: Promoters used in plant transformation. *Vitr Cell Dev Biol - Plant.*, **40**:1–22

Pulido P, Perello C, Rodriguez-Concepcion M (2012). New Insights into Plant Isoprenoid Metabolism. *Mol Plant.*, **5**:964–967

Qu LQ, Takaiwa F (2004). Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice. *Plant Biotechnol J.*, **2**:113–125

Qu LQ, Xing YP, Liu WX, Xu XP, Song YR (2008). Expression pattern and activity of six glutelin gene promoters in transgenic rice. *J Exp Bot.*, **59**:2417–2424

Ravel C, Fiquet S, Boudet J, Dardevet M, Vincent J, Merlino M, Michard R, Martre P (2014). Conserved cis-regulatory modules in promoters of genes encoding wheat high-molecular-weight glutenin subunits. *Front Plant Sci.*, **5**:

Rodríguez-Concepción M (2010). Supply of precursors for carotenoid biosynthesis in plants. *Arch Biochem Biophys.*, **504**:118–122

Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB (2017). Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell.*, **171**:470–480

Rodríguez-Villalón A, Gas E, Rodríguez-Concepción M (2009). Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown Arabidopsis seedlings. *Plant J.*, **60**:424–435

Ruiz-Sola MÁ, Rodríguez-Concepción M (2012). Carotenoid biosynthesis in Arabidopsis: a colorful pathway. *Arab B.*, **10**:e0158–e0158

Rupert PB, Daughdrill GW, Bowerman B, Matthews BW (1998). A new DNA-binding motif in the Skn-1 binding domain–DNA complex. *Nat Struct Biol*., **5**:484–491

Schnable P, Doreen W, S. FR, C. SJ, Fusheng W, Shiran P, Chengzhi L, Jianwei Z, Lucinda F, A. GT, Patrick M, Denise RA, Laura C, S. KS, Chad T, Cindy S, Kim D, Catrina F, Bill C, et al (2009). The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science (80-).*, **326**:1112–1115

Schaub P, Al-Babili S, Drake R, Beyer P (2005). Why Is Golden Rice Golden (Yellow) Instead of Red? *Plant Physiol*., **138**:441–450

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, et al (2010). Genome sequence of the palaeopolyploid soybean. *Nature*., **463**:178–183

Shimada T, Yamazaki Y, Tanaka K, Ishihama A (2014). The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of Escherichia coli. *PLoS One*., **9**:e90447

Shirsat A, Wilford N, Croy R, Boulter D (1989). Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. *Mol Gen Genet MGG*., **215**:326–331

So J-S, Larkins BA (1991). Binding of an endosperm-specific nuclear protein to a maize beta-zein gene correlates with zein transcriptional activity. *Plant Mol Biol*., **17**:309–319

Springer PS (2000). Gene traps: tools for plant development and genomics. *Plant Cell*., **12**:1007–1020

Stålberg K, Ellerström M, Ezcurra I, Ablov S, Rask L (1996). Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic Brassica napus seeds. *Planta*., **199**:515–519

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

Sugiyama T, Rafalski A, Peterson D, Söll D (1985). A wheat HMW glutenin subunit gene reveals a highly repeated structure. *Nucleic Acids Res*., **13**:8729–8737

Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L (2018a). Carotenoid metabolism in plants: the role of plastids. *Mol Plant*., **11**:58–74

Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L (2018b). Carotenoid Metabolism in Plants: The Role of Plastids. *Mol Plant*., **11**:58–74

Suzuki M, Kao CY, McCarty DR (1997). The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *Plant Cell*., **9**:799–807

Takaiwa F, Yamanouchi U, Yoshihara T, Washida H, Tanabe F, Kato A, Yamada K (1996). Characterization of common cis-regulatory elements responsible for the

endosperm-specific expression of members of the rice glutelin multigene family. *Plant Mol Biol*., **30**:1207–1221

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*., **107**:11626–11631

Toledo-Ortiz G, Johansson H, Lee KP, Bou-Torrent J, Stewart K, Steel G, Rodríguez-Concepción M, Halliday KJ (2014). The HY5-PIF regulatory module coordinates light and temperature control of photosynthetic gene transcription. *PLoS Genet*., **10**:e1004416

Urriola J, Rathore KS (2014). Temporal and spatial activities of a rice glutelin promoter in transgenic sorghum. *Plant Cell, Tissue Organ Cult*., **116**:227–234

Vallabhaneni R, Bradbury LMT, Wurtzel ET (2010). The carotenoid dioxygenase gene family in maize, sorghum, and rice. *Arch Biochem Biophys*., **504**:104–111

Venter M (2007). Synthetic promoters: genetic control through cis engineering. *Trends Plant Sci*., **12**:118–124

Walter MH, Stauder R, Tissier A (2015). Evolution of root-specific carotenoid precursor pathways for apocarotenoid signal biogenesis. *Plant Sci*., **233**:1–10

Welsch R, Beyer P, Hugueney P, Kleinig H, von Lintig J (2000). Regulation and activation of phytoene synthase, a key enzyme in carotenoid biosynthesis, during photomorphogenesis. *Planta*., **211**:846–854

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

Welsch R, Wust F, Bar C, Al-Babili S, Beyer P (2008). A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. *Plant Physiol*., **147**:367–380

Wu Chuan-Yin, Adach T, Hatano T, Washida H, Suzuki A, Takaiwa F (1998). Promoters of rice seed storage protein genes direct endosperm-specific gene expression in transgenic rice. *Plant cell Physiol*., **39**:885–889

Wu Chuan-Yin, Suzuki A, Washida H, Takaiwa F (1998). The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. *Plant J*., **14**:673–683

Wu C, Washida H, Onodera Y, Harada K, Takaiwa F (2000). Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. *Plant J*., **23**:415–421

Xiong C, Luo D, Lin A, Zhang C, Shan L, He P, Li B, Zhang Q, Hua B, Yuan Z, Li H, Zhang J, Yang C, Lu Y, Ye Z, Wang T (2019). A tomato B-box protein SlBBX20 modulates carotenoid biosynthesis by directly activating PHYTOENE

SYNTHASE 1, and is targeted for 26S proteasome-mediated degradation. *New Phytol*., **221**:279–294

Yazdani M, Sun Z, Yuan H, Zeng S, Thannhauser TW, Vrebalov J, Ma Q, Xu Y, Fei Z, Van Eck J (2019). Ectopic expression of ORANGE promotes carotenoid accumulation and fruit development in tomato. *Plant Biotechnol J*., **17**:33–49

Zhang N, McHale LK, Finer JJ (2019). Changes to the core and flanking sequences of G-box elements lead to increases and decreases in gene expression in both native and synthetic soybean promoters. *Plant Biotechnol J*., **17**:724–735

Zhou X, Welsch R, Yang Y, Álvarez D, Riediger M, Yuan H, Fish T, Liu J, Thannhauser TW, Li L (2015). Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc Natl Acad Sci*., **112**:3558–3563

Zhu C, Bai C, Sanahuja G, Yuan D, Farré G, Naqvi S, Shi L, Capell T, Christou P (2010). The regulation of carotenoid pigmentation in flowers. *Arch Biochem Biophys*., **504**:132–141

Chapter IV

The role of plastid division factors (*AtPDV1* and *AtARC3*) and *Or* gene for enhancing carotenoid accumulation in rice endosperm

4.0. Abstract

Plastids are essential sites of photosynthesis and many other important cellular metabolic processes in plants. Plastids exist universally in plants and include various types such as etioplasts, amyloplasts, chloroplasts, and chromoplasts. Chloroplasts are the photosynthetic plastids in green tissues, whereas chromoplasts are important organelles for carotenoid biosynthesis and accumulation. Chromoplast number and size define and strongly influence final carotenoid levels; however, little is known about the machinery that controls chromoplast number and size. One of the chloroplast division factor involved in the initiation of chloroplast division is Accumulation and Replication of Chloroplast 3 (*ARC3*). Another chloroplast division factor is the Plastid Division 1 (*PDV1*) which is recruited to the division site through direct interactions with other division factors such as: ARC6, PARC6 and PDV2 in order to complete chloroplast division. Here we describe preliminary experiments aiming to understand the role of the plastid division factors *AtPDV1* and *AtARC3* in the accumulation of carotenoids in rice endosperm. Our first strategy was to combine the plastid division genes (*AtARC3* and *AtPDV1*) with carotenogenic genes (*ZmPSY1* and *PaCRTI*). This strategy did not result in any significant differences in carotenoid accumulation and chromoplast quantity. In order to develop an alternative strategy to enhance carotenoid accumulation in rice endosperm we combined an Orange gene(Or^{His}) that was previously shown to encode a chaperone-like protein that induces the formation of chromoplasts, creating a metabolic sink for carotenoids, with plastid division genes. Therefore, in our second strategy we combined *AtARC3* and *AtPDV1* with $AtOr^{His}$. We demonstrate co-transformation of rice with the three genes (*AtARC3*, *AtPDV1* and $AtOr^{His}$). We also confirmed mRNA accumulation for the three genes in rice callus. Further analyses will focus on determining whether the strategy alters chromoplast number and size and in turn carotenoid composition in rice endosperm.

4.0.Introduction

4.1.Plastid division genes and *Orange* (*Or*) gene

Plastids are essential sites of photosynthesis and many other important cellular metabolic activities in plants. Plastids exist universally in plants and include various types like etioplasts, amyloplasts, chloroplasts, and chromoplasts (Lopez-Juez and Pyke 2004). Chloroplasts are the photosynthetic plastids in plant green tissues. Carotenoids are localized in chloroplast thylakoid membranes for photosynthesis and photoprotection. Chloroplasts evolved from a cyanobacterial endosymbiont and multiply by binary fission. The Accumulation and Replication of Chloroplast 3 (*ARC3*) gene is a chloroplast division factor involved in the initiation of chloroplast division. ARC3 is a chimera of the prokaryotic filamenting temperature sensitive mutant Z protein (FtsZ) (essential for cell division) and part of the eukaryotic phosphatidylinositol-4-phosphate 5-kinase (PIP5K) (Vitha et al 2003). The proteins above are located on the outer surface of the chloroplast in a ring-like (Z-ring) structure at the early stage of chloroplast division (McAndrew et al 2008). After the formation of the Z-ring, the outer envelope-spanning proteins Plastid Division 1 (PDV1) and PDV2 are recruited to the division site through direct interactions with Paralog of ARC6 (PARC6) and ARC6 (Glynn et al 2008, 2009). FtsZ and ARC6 are derived from the cell division machinery of the cyanobacterial endosymbiont (Vitha et al 2003), and PARC6 is derived from ARC6 by gene duplication and differentiation (Glynn et al 2009). Finally, Dynamin-Related Protein5b (DRP5B; also known as ARC5), a member of the dynamin family of self-assembling proteins, is recruited by PDV1 and PDV2 to chloroplasts to complete chloroplast division (Miyagishima et al 2006). DRP5B is specific to plants and algae and is thought to have evolved from a dynamin-related protein involved in eukaryotic cytokinesis (Miyagishima et al 2008; Holtsmark et al 2013). PDV1 and PDV2 are specific to land plants (Miyagishima et al 2006), in which the levels of PDV proteins determine the rate of chloroplast division (Okazaki et al 2009).

Chromoplasts are plastids that synthesize and accumulate carotenoids (Egea et al 2010). Chromoplast number and size define the capacity for carotenoid accumulation. The single-locus *Or* mutation in cauliflower (*Brassica oleracea*) confers high levels of β-carotene accumulation in tissues where the accumulation of carotenoids is normally repressed (Lu et al 2006). The *Or* gene appears to be plant specific with homologs present in all plant species examined, including algae. *Or* encodes a plastidic DnaJ cysteine-rich domain-containing protein (Lu et al 2006; Tzuri et al 2015).

Overexpression of *Or* leads to carotenoid accumulation in different species. In cereals the expression of *Or* gene in rice and maize were achieved. In rice the expression of *Or* in combination with *ZmPSY1* and *CRTI* enhanced the carotenoid accumulation in endosperm (Bai et al 2016). Increase from no detectable carotenoids in wild-type rice endosperm to an accumulation of 25.83μg/g DW in total carotenoids and 10.52μg/g DW in β-carotene were reported. Other carotenoids detected were zeaxanthin and lutein but less than 2 μg/g DW. The expression of *Or* in maize (M37W) endosperm also enhance the carotenoids in endosperm (Berman et al 2017). The total carotenoid content increased from 0.7μg/g DW in wild-type to 25μg/g DW in the transgenic lines. The levels and composition of carotenoids differed compared to rice. Maize endosperm accumulate mainly: zeaxanthin (12μg/g DW), antheraxanthin (7μg/g DW), lutein (5μg/g DW) and violaxanthin (2μg/g DW). Other carotenoid enhanced in rice like β-carotene were not detected in maize lines.

When the *Or* gene was expressed in potato tubers the total carotenoid increase was from 5μg/g DW in wild-type to 31μg/g DW in transgenic lines (Lopez et al 2008). Other carotenoids detected were: lutein (9μg/g DW), phytoene (10μg/g DW) and β-carotene (5μg/g DW). Total carotenoid in storage roots of purple-fleshed sweet potato increased from less than 1μg/g DW in wild-type to 8μg/g DW in *Or*-expressing lines (Park et al 2015). Additionally, were detected: zeaxanthin (2.4μg/g DW) and 9z-β-carotene (1μg/g DW). The sweet potato expression of *Or* were analyzed in callus. The total carotenoid accumulation were more than 20 folds (40μg/g DW) compare to wilt-type (less than 5μg/g DW). The composition of enhanced carotenoids was: α-carotene (1.5 μg/g DW), β-carotene (10 μg/g DW), lutein (10 μg/g DW), zeaxanthin (6 μg/g DW) and β-cryptoxanthin (14 μg/g DW).

Investigation with *Cucumis melo* (melon) identified a single nucleotide polymorphism (SNP) between *Cm*Or alleles in orange versus green/white-flesh fruit (Tzuri et al 2015). The single (SNP) causes a change of an arginine to histidine in the *Cm*Or protein. The role of the amino acid substitution was defined by site-directed mutagenesis changing Or^{Arg} to Or^{His}. Functional analysis of the *Or* gene polymorphism in *Arabidopsis thaliana* callus confirmed the role of Or^{His} to induce carotenoid accumulation (Tzuri et al 2015). Additional investigations using the Or^{His} variant of Arabidopsis *Or*, which functions as the melon Or^{His} allele, induces chromoplast formation (Yuan et al 2015) These results suggest an additional role for Or^{His} as a chromoplast number regulator in addition to its

functions in carotenoid biosynthesis and chromoplast biogenesis. Moreover, the *AtOr*^{wt} and *AtOr*^{His} were expressed in tomato (Yazdani, Sun, Yuan, Zeng, Theodore W Thannhauser, et al 2019). The results show that the tomato expressing *Or*^{His} accumulate more total carotenoids (160µg/g FW) than the *Or*^{wt} (92µg/g FW) at red fruit stage. The carotenoid composition also shows differences in accumulation between *AtOr*^{His} and *AtOr*^{wt}: phytoene (18 to 6.2 µg/g FW), phytofluene (12 to 6.1 µg/g FW), lycopene (77 to 64 µg/g FW) and b-carotene (27 to 14 µg/g FW). While the variants of *Or* are able to regulate phytoene synthase (PSY) post-translationally, the orange variants exhibit an additional function in promoting chromoplast biogenesis (Lu et al 2006; Tzuri et al 2015; Wang et al 2015; Yuan et al 2015; Chayut et al 2017; Welsch et al 2018).

Report demonstrate that *Or*^{His} is associated with division sites in non-green plastids (Sun et al 2020). *Or*^{His} specifically interacts with ARC3 and reduce the interaction of PARC6–ARC3. Also the expression of *Or*^{His} with *ARC3* gene in Arabidopsis enhances carotenoid accumulation by 85% compared to the lines expressing only *Or*^{His} (Sun et al 2020). Earlier reports showed that upregulation of *PDV1* and *PDV2* expression significantly increases the number of chloroplasts in Arabidopsis (Okazaki et al 2009). Moreover, the combination of *Or*^{His} with *PDV1* also increases more than two-fold the total carotenoid accumulation in Arabidopsis callus (Sun et al 2020). The use of plastid division genes presents a novel strategy for carotenoid accumulation in crops.

4.2. Aims and Objectives

Aim: The aim of this chapter was to understand the role of the plastid division factors (*AtPDV1* and *AtARC3*) in the accumulation of carotenoids in rice endosperm.
Objectives:

- Co-transform AtPDV1 and AtARC3 with Or^{his} gene to enhance the accumulation of carotenoids in rice endosperm.
- Use *ZmPSY1* and *CRTI* driven by two endosperm specific promoters (low molecular weight glutenin and d-hordein, respectively) as a positive control for carotenoid accumulation.
- Use transcriptomic and metabolomics analysis to elucidate the impact of the plastid division factors on carotenoid accumulation in rice endosperm.

4.3. Material and Methods

4.3.1. Cloning and sequencing the *AtPDV1, AtARC3, ZmPSY1, CRTI* and Or^{his} cDNAs

The *AtPDV1*, *AtARC3* and Or^{his} genes were kindly provided by Dr. L. Li (Cornell University, Ithaca, NY, USA). *AtPDV1*, *AtARC3* and Or^{his} genes are based on sequences in GenBank (accession numbers At5g53280, At1g75010 and At5G61670). *ZmPSY1* and *CRTI* were cloned as described in Bai et al., 2014 based on sequences in Genbank (accession numbers AY324431 and D90087.2).

4.3.2. Vector construction

AtPDV1 and *AtARC3* were amplify by PCR from vector provided by Dr. L. Li (Cornell University, Ithaca, NY, USA) and restriction sites were added. The PCR products were digested with the corresponding restriction enzymes (Table 4. 1) and ligated into pAL76 vector, containing the maize ubiquitin-1 promoter plus first intron and Nos transcriptional terminator.

The maize *PSY1* cDNA was cloned from maize inbred line B73 by RT–PCR using forward and reverse primer (Table 4.1) based on the *PSY1* sequence (GenBank accession number AY324431). The product was transferred to pGEM-T Easy (Promega) for

sequencing, and then to the p326 vector containing the LWM glutenin promoter and nos terminator (Stoger et al., 1999). The *Pantoea ananatis* (formerly *Erwinia uredovora*) CRTI gene was fused in-frame with the transit peptide sequence from the P. vulgaris small subunit of ribulose bisphosphate carboxylase (Schreier et al., 1985) in plasmid pYPIET4 (Misawa et al., 1993), and amplified by PCR using forward and reverse primers (Table 4. 1). The product was transferred to pGEM-T Easy for sequencing, and then to pHorp–P (Sorensen et al., 1996) containing the barley D–hordein promoter and the rice ADPGPP terminator. All transformation constructs were verified by sequencing.

Table 4. 1. Primers used to amplify AtPDV1, AtARC3, ZmPSY1, CRTI and Or^{His} genes.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
AtPDV1	GGATCCATGGGAGAAATGGAGATC GAAGAAATCG	CCCGGGTTAACCACGAGCCATCAT TACGTC
AtARC3	GGATCCATGCCGATTTCTATGGAA CTTCCAGT	AAGCTTTCAATCTCCGGCGTCCACTTG TTTCC
AtOr^{his}	GCGGCCGCATGTCATCTTTGGGTA GGATTTT	GAATTCTCAATCGAAAGGGTCGATA CGA
ZmPSY1	AGGATCCATGGCCATCATACTCGT ACGAG	AGAATTCTAGGTCTGGCCATTTCTC AATG
PaCRTI	ATCTAGAATGGCTTCTATGATATC CTCTTC	AGAATTCTCAAATCAGATCCTCCA GCATCA

AtPDV1, Arabidopsis thaliana plastid division factor 1; BamHI (5´-GGATCC-3´) and XmaI (5´-CCCGGG-3´) sites (underlined and red) were introduced into each primer. *AtARC3*, Arabidopsis thaliana Accumulation and Replication of Chloroplast 3; BamHI (5´-GGATCC-3´) and HindIII (5´-AAGCTT-3´) sites (underlined and red) were introduced into each primer. *$AtOr^{his}$*, Arabidopsis thaliana orange gene mutant; NotI (5´-GCGGCCGC-3´) and EcoRI (5´-GAATTC-3´) sites (underlined and red) were introduced into each primer. *ZmPSY1*, Maize (*Zea mays*) phytoene synthase 1 gene; BamHI (5´-GGATCC-3´) and EcoRI (5´-GAATTC-3´) sites (underlined and red) were introduced into each primer. *PaCRTI*, bacterial (*Pantoea ananatis*) phytoene desaturase; XbaI (5´-TCTAGA-3´) and EcoRI (5´-GAATTC-3´) sites (underlined and red) were introduced into each primer.

4.3.3. Rice transformation

Rice transformation was performed as described in chapter 2.3.2

4.3.4. Gene expression analysis

Gene expression analysis was performed as described in chapter 2.3.4. Primer combinations are listed in supporting information

Table 4. 2. Primers used for qRT-PCR analysis.

Gene	Forward sequence (5′-3′)	Reverse sequence (5′-3′)
AtPDV1	ACCGCTCTTGAGAACCTTGA	TCCATAGTTCTTGCCGTGGT
AtARC3	GAGGGGGAACCATCCAGAAA	TCTCCAGAATCAGCTTGCCA
Orhis	TTCGAAGCCACCGAAACAAG	AGGAGCTAGTAGACCACCGA
CrtI	TCAGTTTCTGGACTATTCACGCGCGGT	CTGGCAACCTTACTGTAAACGCTTCTCCATG
ZmPSY1	GTCCGAGCAGAAGGTCTACG	CTGTCCTCCTACACCACACA
OsACTIN	GACTCTGGTGATGGTGTCAGC	TCATGTCCCTCACAATTTCC

AtPDV1, Arabidopsis thaliana plastid division factor 1. *AtARC3*, Arabidopsis thaliana Accumulation and Replication of Chloroplast 3. *AtOr*^{his}, Arabidopsis thaliana orange gene mutant. *ZmPSY1*, Maize (*Zea mays*) phytoene synthase 1 gene. *PaCRTI*, bacterial (*Pantoea ananatis*) phytoene desaturase. *OsACTIN*, rice actin gene.

4.4.Results

4.4.1. Co-expression of *ZmPSY1* and *PaCRTI* with *AtARC3* and *AtPDV1* in rice callus. Carotenoid analysis and chromoplast number in transgenic callus.

The expression patterns of all constructs (LMWG-*ZmPSY1*-TNos, pHord-*PaCRTI*-TNos, Ubi-*AtARC3*-TNos, Ubi-*AtPDV1*-TNos) were analyzed by quantitative real-time PCR (**Figure 1A**). *AtARC3* and *AtPDV1* relative mRNA accumulation was higher compared with *ZmPSY1* and *PsCRTI* in callus. Differences in expression levels might be attributed to the promoters used in the different vectors. *AtARC3* and *AtPDV1* were controlled by a strong constitutive promoter (maize ubiquitin 1), *ZmPSY1* and *PsCRTI* by an endosperm specific promoter (Low molecular weight glutenin and d-hordein promoter respectively). The β-carotene content of line #91 and wild type were very low and not a significant difference was seen (**Figure 4. 1B**). The chromoplast number in the isolated protoplasts from wild type was similar to line 91 (**Figure 4. 2A, B**).

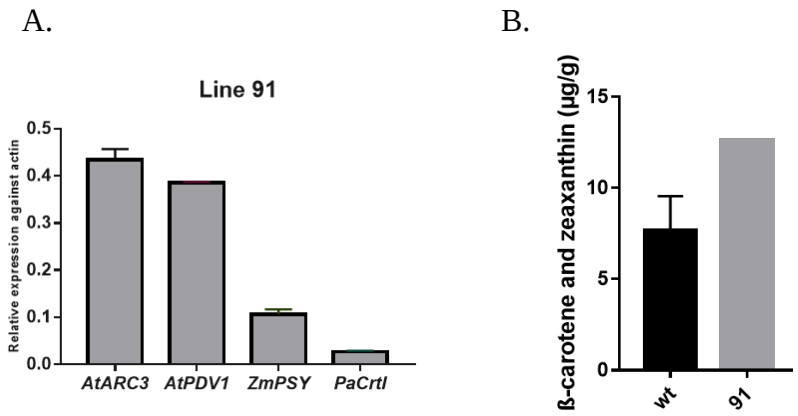


Figure 4. 1A, B. mRNA accumulation and carotenoid level in rice callus. (A) mRNA accumulation of ZmPSY1, PaCRTI, AtARC3 and AtPDV1 in rice callus. Expression levels were normalized to the rice actin gene, as determined by qRT-PCR. (B) β-carotene and zeaxanthin levels in lines #91 and wild type (wt).

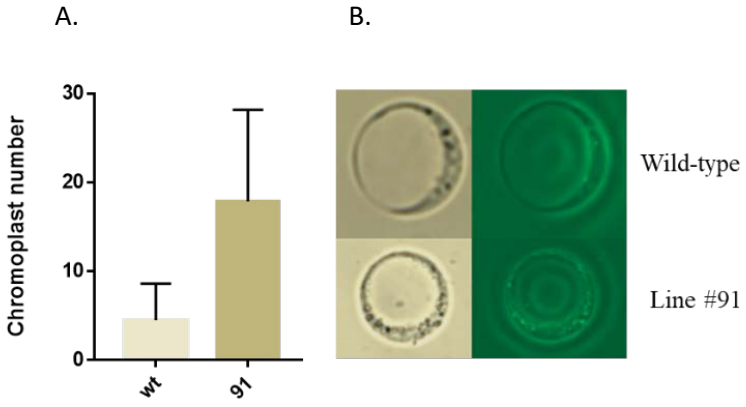


Figure 4. 2A, B. Chromoplast number and protoplast isolation from rice callus. (A) Chromoplast number in wild type callus and line 91. (B) Protoplast isolated from rice callus of wild type and line 91.

4.4.2. Cloning and functional characterization of *AtPDV1*, *AtARC3* and *AtOr^{His}* genes.

The *A. thaliana* plastid division 1 (*AtPDV1*), *A. thaliana* Accumulation and Replication of Chloroplast 3 (*AtARC3*) and *A. thaliana* Orange-histidine (*AtOr^{His}*) cDNAs encoding the corresponding full-length genes were kindly provided by Dr. L. Li (Cornell University, NY, USA). *AtARC3* and *AtPDV1* were cloned in a vector containing the maize ubiquitin promoter (with first intron) and the Nos terminator. *AtOr^{His}* was cloned in a vector containing the barley d-hordein promoter and the rice ADPGPP terminator (**Figure 4. 3**). All vectors were confirmed by restriction digestion enzymes and sequencing (**Figure 4. 4**).

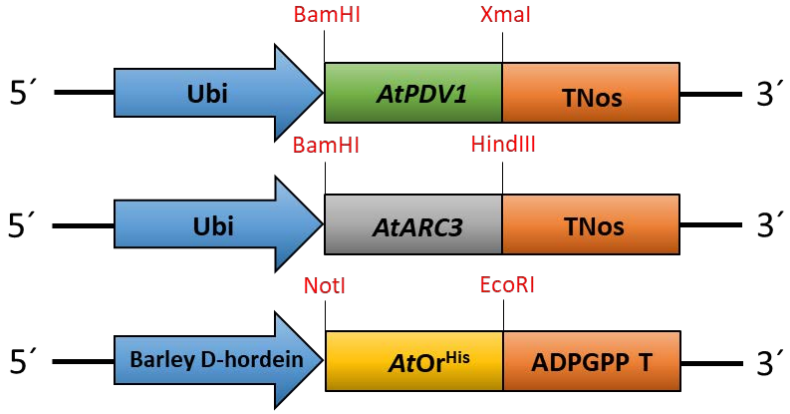


Figure 4. 3. Schematic representation of expression vectors Ubi-AtPDV1-TNos, Ubi-AtARC3-TNos and Ubi-AtOr^{His}-TADP.

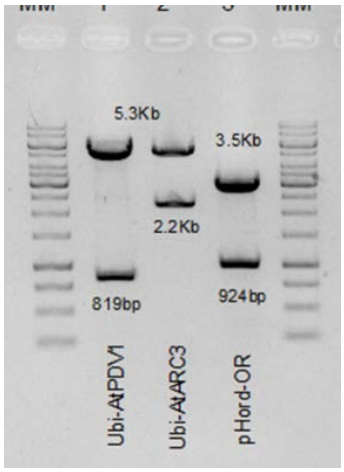


Figure 4. 4. Restriction enzyme analysis of the transformation vectors.

4.4.3. Co-expression of *AtOr^{His}* with *AtARC3* and *AtPDV1* in rice callus and endosperm.

The expression patterns of all constructs (pHord-*Or^{His}*-ADPT, Ubi-*AtARC3*-TNos, Ubi-*AtPDV1*-TNos) were analyzed by quantitative real-time PCR (**Figure 4. 5**). *AtARC3* and *AtPDV1* relative mRNA accumulation was higher compared with *Or^{His}* in callus lines. Differences in expression levels might be attributed to the promoters used in the different vectors. *AtARC3* and *AtPDV1* were controlled by a strong constitutive promoter (maize ubiquitin 1) and the *Or^{His}* gene by an endosperm specific promoter (d-hordein promoter).

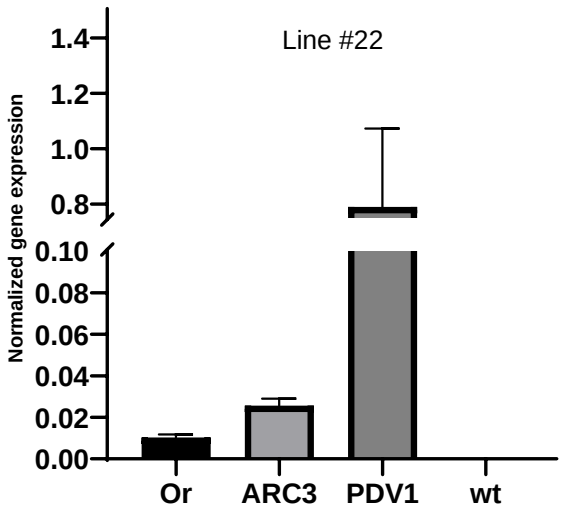


Figure 4. 5. mRNA accumulation of Or^{His}, AtARC3 and AtPDV1 in rice callus. Expression levels were normalized to the rice actin gene, as determined by qRT-PCR. Each value represents the average of three experiments and the error bars indicates the standard deviation.

4.5. Discussion

In plants, carotenoids accumulate in specialized pigment-bearing structures known as plastoglobuli, within chromoplasts (Bréhélin et al 2007). Chromoplasts are the key organelles for carotenoid biosynthesis and accumulation. Chromoplast number and size define and strongly influence the final carotenoid level (Sun et al 2018); however, little is known about the machinery that controls chromoplast number and size. One of the

chloroplast division factor involved in the initiation of chloroplast division is Accumulation and Replication of Chloroplast 3 (*ARC3*) gene. *ARC3* is involved in the structure at the early stage of chloroplast division (McAndrew et al 2008). Another chloroplast division factor is the Plastid Division 1 (*PDV1*) which is recruited to the division site through direct interactions with other division factors such as: ARC6, PARC6 and PDV2 in order to complete the chloroplast division (Miyagishima et al 2006). The levels of PDV protein determines the rate of chloroplast division (Okazaki et al 2009).

Here we describe preliminary experiments aiming to understand the role of the plastid division factors *AtPDV1* and *AtARC3* in the accumulation of carotenoids in rice endosperm. Our first strategy was to combine the plastid division genes (*AtARC3* and *AtPDV1*) with carotenogenic genes (*ZmPSY1* and *PaCRTI*). *ZmPSY1* and bacterial *CRTI* are the most common genes used in order to increase the accumulation of carotenoids in rice endosperm (Ye et al. 2000; Paine et al. 2005; Baisakh et al. 2006; Aluru et al. 2008; Bai et al. 2016). Unfortunately, this strategy did not result in any significant differences in carotenoid accumulation and chromoplast quantity (**Figure 4.1A, B** and **4.2A, B**). Our results were similar to studies that use *ZmPSY1* and *CRTI*. Our best line accumulates around 13 µg/g β-carotene and zeaxanthin (**Figure 4.1B**) being in the average of studies in different varieties of rice that range from 1.0 to 37 µg/g carotenoid accumulation (Paine et al 2005; Parkhi et al 2005; Dong et al 2020). Also the chromoplast number did not show significant differences compared with the wild type rice callus (**Figure 4.2A**). A possible explanation of the similarity of chromoplast number between wild type and the lines we generated here is that the plastid division genes are specific for chloroplast rather than chromoplast division (Chen et al 2018).

Other strategies for carotenoid accumulation include the *AtOr* gene in combination with *ZmPSY1* and *CRTI* to promote the differentiation of chromoplasts thus generating a metabolic sink that promotes the accumulation of carotenoids in the endosperm (Lu et al 2006; Bai et al 2016).The cauliflower Orange (*Or*) gene was discovered by the analysis of an orange curd cauliflower mutant, and was shown to encode a chaperone-like protein that induces the formation of chromoplasts, creating a metabolic sink for carotenoids (Lu et al 2006). Further investigations in *Cucumis melo* (melon) identified a single nucleotide polymorphism (SNP) between *Cm*Or alleles in orange versus green/white-flesh fruit (Tzuri et al 2015). The SNP causes a change from an arginine to histidine in the *Cm*Or

protein. Functional analysis of the *Or* gene polymorphism in Arabidopsis callus confirmed the role of Or gene with the histidine change (Or^{His}) to induce carotenoid accumulation (Tzuri et al 2015). In recent research it was demonstrated that the overexpression of $AtOr^{His}$ in combination with *ZmPSY1* and *CRTI* increased total carotenoids in rice callus from near zero in wild-type rice endosperm to 25.83μg/g DW in the transgenic plants expressing the three genes; β-carotene levels were 10.52μg/g DW (Bai et al 2016).

In order to develop new strategies to enhance carotenoid accumulation in rice endosperm we combined carotenogenic genes with plastid division genes. Our second strategy was a combination of *AtARC3* and *AtPDV1* with $AtOr^{His}$ (**Figure 4.5**). It was already demonstrated that $AtOr^{His}$ specifically interacts with *AtARC3* to regulate chromoplast number and enhances carotenoid accumulation by 85% compared to the lines expressing only Or^{His} (Sun et al 2020). Earlier reports demonstrated that upregulation of *AtPDV1* and *AtPDV2* expression significantly increases the number of chloroplasts in Arabidopsis (Okazaki et al 2009). Moreover, upregulation of *PDV1* was shown to dramatically enhance carotenoid levels in $AtOr^{His}$ callus (Sun et al 2020). In the same study, *AtARC3* and *AtPDV1* have been implicated in $AtOr^{His}$ regulated chromoplast duplication, which provides a foundation for dissecting the binary fission mechanism in regulating chromoplast numbers. Plastid division factors affect carotenoid levels and chromoplast number (Sun et al 2020). $AtOr^{His}$ has been shown to enhance carotenoid accumulation in various plants like melon, tomato, Arabidopsis and rice (Bai et al 2014; Tzuri et al 2015; Sun et al 2020). Chromoplast number and size determine total carotenoid accumulation (Sun et al 2018). Indeed, increased chromoplast size and number is accompanied by enhanced carotenoid levels in tomato high pigment mutants (Galpaz et al 2008). *AtARC3* is known to affect chloroplast size and may have a similar effect on chromoplast size (Glynn et al 2007).

Co-expression of *AtARC3* or *AtPDV1* with $AtOr^{His}$ in Arabidopsis callus was shown to enhance significantly carotenoid accumulation (Sun et al 2020). We combined both plastid division factors (*AtARC3* and *AtPDV1*) with $AtOr^{His}$ in order to enhance carotenoid accumulation in the rice endosperm. Our results demonstrate co-transformation of rice with the three genes (*AtARC3*, *AtPDV1* and $AtOr^{His}$). Also we confirmed mRNA accumulation for the three genes in rice callus. Earlier studies showed that the callus system can be used to predict the effect of genes in metabolic pathways

(Bai et al 2014). Nonetheless the callus system can be used even with endosperm specific promoters that are active in callus and have restricted activity in mature plants (Bai et al 2014).

Further carotenoid and chromoplast analysis will determine whether these lines exhibit enhanced carotenoid content in rice endosperm. Nevertheless, the demonstration that co-expression of *AtARC3*, *AtOr^{His}* and *AtPDV1* affects positively carotenoid levels suggests the potential of plastid division factors to serve as new genetic tools for carotenoid enrichment in plants. However, we are far from identifying all the necessary factors controlling chromoplast division and number. Further experiments will be required to identify and characterize proteins that associate with these factors in order to determine their precise role in the accumulation of carotenoids in the rice endosperm.

4.6. Conclusions

We explored strategies for carotenoid accumulation in rice endosperm, focusing specifically on two strategies. The first strategy co-expressing the carotenogenic genes *ZmPSY1* and *PaCRTI* in combination with chloroplast division factors (*AtARC3* and *AtPDV1*) did not achieve an increase in carotenoid levels and chromoplast numbers compare to wild type rice callus (**Figure 4.1A, B** and **4.2A, B**).

In the second strategy we utilized *AtOr^{His}* in combination with both chloroplast division factors. We generated rice callus lines containing and expressing the input genes at the mRNA level. Further analyses will focus on determining whether the strategy alters chromoplast number and size and in turn carotenoid composition.

4.7. References

Aluru M, Xu Y, Guo R, Wang Z, Li S, White W, Wang K, Rodermel S (2008). Generation of transgenic maize with enhanced provitamin A content. *J Exp Bot*., **59**:3551–3562

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

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

Baisakh N, Rehana S, Rai M, Oliva N, Tan J, Mackill DJ, Khush GS, Datta K, Datta SK (2006). Marker-free transgenic (MFT) near-isogenic introgression lines (NIILs) of 'golden' indica rice (cv. IR64) with accumulation of provitamin A in the endosperm tissue. *Plant Biotechnol J*., **4**:467–475

Berman J, Zorrilla-López U, Medina V, Farré G, Sandmann G, Capell T, Christou P, Zhu C (2017). The Arabidopsis ORANGE (AtOR) gene promotes carotenoid accumulation in transgenic corn hybrids derived from parental lines with limited carotenoid pools. *Plant Cell Rep*., **36**:933–945

Bréhélin C, Kessler F, van Wijk KJ (2007). Plastoglobules: versatile lipoprotein particles in plastids. *Trends Plant Sci*., **12**:260–266

Chayut N, Yuan H, Ohali S, Meir A, Sa'ar U, Tzuri G, Zheng Y, Mazourek M, Gepstein S, Zhou X, Portnoy V, Lewinsohn E, Schaffer AA, Katzir N, Fei Z, Welsch R, Li L, Burger J, Tadmor Y (2017). Distinct Mechanisms of the ORANGE Protein in Controlling Carotenoid Flux. *Plant Physiol*., **173**:376–389

Chen C, MacCready JS, Ducat DC, Osteryoung KW (2018). The Molecular Machinery of Chloroplast Division. *Plant Physiol*., **176**:138–151

Dong OX, Yu S, Jain R, Zhang N, Duong PQ, Butler C, Li Y, Lipzen A, Martin JA, Barry KW, Schmutz J, Tian L, Ronald PC (2020). Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat Commun*., **11**:1–10

Egea I, Barsan C, Bian W, Purgatto E, Latché A, Chervin C, Bouzayen M, Pech J-C (2010). Chromoplast differentiation: current status and perspectives. *Plant cell Physiol*., **51**:1601–1611

Galpaz N, Wang Q, Menda N, Zamir D, Hirschberg J (2008). Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. *Plant J*., **53**:717–730

Glynn JM, Froehlich JE, Osteryoung KW (2008). Arabidopsis ARC6 coordinates the division machineries of the inner and outer chloroplast membranes through interaction with PDV2 in the intermembrane space. *Plant Cell*., **20**:2460–2470

Glynn JM, Miyagishima S, Yoder DW, Osteryoung KW, Vitha S (2007). Chloroplast division. *Traffic*., **8**:451–461

Glynn JM, Yang Y, Vitha S, Schmitz AJ, Hemmes M, Miyagishima S, Osteryoung KW (2009). PARC6, a novel chloroplast division factor, influences FtsZ assembly and is

required for recruitment of PDV1 during chloroplast division in Arabidopsis. *Plant J.*, **59**:700–711

Holtsmark I, Lee S, Lunde KA, Auestad K, Maple-Grødem J, Møller SG (2013). Plastid division control: the PDV proteins regulate DRP5B dynamin activity. *Plant Mol Biol*., **82**:255–266

Lopez-Juez E, Pyke KA (2004). Plastids unleashed: their development and their integration in plant development. *Int J Dev Biol*., **49**:557–577

Lopez AB, Van Eck J, Conlin BJ, Paolillo DJ, O'Neill J, Li L (2008). Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J Exp Bot*., **59**:213–223

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*., **18**:3594–3605

McAndrew RS, Olson BJSC, Kadirjan-Kalbach DK, Chi-Ham CL, Vitha S, Froehlich JE, Osteryoung KW (2008). In vivo quantitative relationship between plastid division proteins FtsZ1 and FtsZ2 and identification of ARC6 and ARC3 in a native FtsZ complex. *Biochem J*., **412**:367–378

Miyagishima S, Froehlich JE, Osteryoung KW (2006). PDV1 and PDV2 mediate recruitment of the dynamin-related protein ARC5 to the plastid division site. *Plant Cell*., **18**:2517–2530

Miyagishima S, Kuwayama H, Urushihara H, Nakanishi H (2008). Evolutionary linkage between eukaryotic cytokinesis and chloroplast division by dynamin proteins. *Proc Natl Acad Sci*., **105**:15202–15207

Okazaki K, Kabeya Y, Suzuki K, Mori T, Ichikawa T, Matsui M, Nakanishi H, Miyagishima S-Y (2009). The PLASTID DIVISION1 and 2 components of the chloroplast division machinery determine the rate of chloroplast division in land plant cell differentiation. *Plant Cell*., **21**:1769–1780

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL, Drake R (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat Biotechnol*., **23**:482–487

Park S-C, Kim SH, Park S, Lee H-U, Lee JS, Park WS, Ahn M-J, Kim Y-H, Jeong JC, Lee H-S (2015). Enhanced accumulation of carotenoids in sweetpotato plants overexpressing IbOr-Ins gene in purple-fleshed sweetpotato cultivar. *Plant Physiol Biochem*., **86**:82–90

Parkhi V, Rai M, Tan J, Oliva N, Rehana S, Bandyopadhyay A, Torrizo L, Ghole V, Datta K, Datta SK (2005). Molecular characterization of marker-free transgenic lines of indica rice that accumulate carotenoids in seed endosperm. *Mol Genet Genomics*., **274**:325–336

Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L (2018). Carotenoid Metabolism in Plants: The Role of Plastids. *Mol Plant*., **11**:58–74

Sun T, Yuan H, Chen C, Kadirjan-Kalbach DK, Mazourek M, Osteryoung KW, Li L

(2020). OR(His), a Natural Variant of OR, Specifically Interacts with Plastid Division Factor ARC3 to Regulate Chromoplast Number and Carotenoid Accumulation. *Mol Plant.*, **13**:864–878

Tzuri G, Zhou X, Chayut N, Yuan H, Portnoy V, Meir A, Sa'ar U, Baumkoler F, Mazourek M, Lewinsohn E (2015). A 'golden'SNP in CmOr governs the fruit flesh color of melon (C ucumis melo). *Plant J.*, **82**:267–279

Vitha S, Froehlich JE, Koksharova O, Pyke KA, van Erp H, Osteryoung KW (2003). ARC6 is a J-domain plastid division protein and an evolutionary descendant of the cyanobacterial cell division protein Ftn2. *Plant Cell.*, **15**:1918–1933

Wang Z, Ke Q, Kim MD, Kim SH, Ji CY, Jeong JC, Lee H-S, Park WS, Ahn M-J, Li H (2015). Transgenic alfalfa plants expressing the sweetpotato Orange gene exhibit enhanced abiotic stress tolerance. *PLoS One.*, **10**:e0126050

Welsch R, Zhou X, Yuan H, Álvarez D, Sun T, Schlossarek D, Yang Y, Shen G, Zhang H, Rodriguez-Concepcion M (2018). Clp protease and OR directly control the proteostasis of phytoene synthase, the crucial enzyme for carotenoid biosynthesis in Arabidopsis. *Mol Plant.*, **11**:149–162

Xudong Y, Salim A-B, Andreas K, Jing Z, Paola L, Peter B, Ingo P (2000). Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science (80-).*, **287**:303–305

Yazdani M, Sun Z, Yuan H, Zeng S, Thannhauser TW, Vrebalov J, Ma Q, Xu Y, Fei Z, Van Eck J (2019). Ectopic expression of ORANGE promotes carotenoid accumulation and fruit development in tomato. *Plant Biotechnol J.*, **17**:33–49

Yuan H, Owsiany K, Sheeja TE, Zhou X, Rodriguez C, Li Y, Welsch R, Chayut N, Yang Y, Thannhauser TW (2015). A single amino acid substitution in an ORANGE protein promotes carotenoid overaccumulation in Arabidopsis. *Plant Physiol.*, **169**:421–431

Chapter V

GENERAL DISCUSSION

6.1. General discussion

Global food security is one of the main targets of world leaders, aiming primarily to combat hidden hunger. Hidden hunger or micronutrient malnutrition is a public health concern worldwide. Hidden hunger is described often as a multiple micronutrient deficiency (Lowe 2021). Malnutrition affects a further two billion people who even though consume enough calories; they do not obtain the essential nutrients needed to maintain a healthy life (Berman et al 2013). In most developing countries, diet is based on a single staple crop, which invariably contains low amounts of nutrients (e.g. rice). One of the most widespread micronutrient deficiencies is vitamin A deficiency (VAD). An effective way of achieving food security is by using food biofortification, aiming to improve crop production as well as nutrient content. Currently, the main strategies to biofortify food crops are: conventional breeding and genetic engineering, or a combination of the two. Conventional breeding improves crops without using recombinant DNA technology, but using mutagenesis or marker assisted selection to introgress genes from distant relatives, which similarly would not occur naturally (Bai et al 2011). The relatively long lead time to have an impact and its dependence on a compatible gene pool makes this approach very limited. The long time frames to obtain nutritionally improved lines are one of the challenges in conventional breeding (Pérez-Massot et al 2013). On the other hand, genetic engineering can generate biofortified crops by transferring genes directly into breeding lines, obtaining transgenic plants with enhanced nutritional traits. Compared to conventional breeding, genetic engineering has the advantages of direct engineering of breeding lines, multiple simultaneous biofortification of different nutrients and unrestricted access to genetic diversity including recombinant genes (Zhu et al 2007; Alvarez et al 2021). An entry barrier of genetic engineering is the regulatory global framework which makes the technology prohibitively expensive, particularly for developing countries and public institutions, at least in the case of philanthropic, not-for-profit applications. Because of the beneficial effects of secondary metabolites on human health and industry, biosynthesis of secondary metabolites and their regulation have always been a prime focus of research.

I employed three different strategies to biofortify rice, a staple food crop, using metabolic engineering. The carotenoid biosynthetic pathway in rice endosperm is blocked in the first committed step, the condensation of two geranylgeranyl diphosphate (GGPP) molecules by phytoene synthase (PSY) to 15-cis-phytoene (Misawa et al 1994; Bai et al

2014). We had previously demonstrated that combinatorial transformation can unravel bottlenecks in metabolic pathways in plants (Bai et al 2014, 2016). Our aim has been to define the minimum set of genes necessary to reconstitute the carotenoid pathway in rice endosperm. Metabolic engineering has been a trial and error approach where the elimination of one bottleneck often merely serves to reveal the next bottleneck. In contrast, combinatorial transformation generates a library of metabolic variants allowing the best strategy to be deduced in a single experiment if all necessary combinations of transgenes are represented in the population (Zhu et al 2008; Li et al 2010; Bai et al 2014). In pilot experiments we co-transformed rice with *OsPSY1* and *OsPDS* in order to start with the minimum number of genes in the reconstitution of the pathway. The screening was established using PCR and a color phenotype in callus. In earlier published work from our lab we established the correlation of phenotype with metabolite accumulation in rice callus (Bai et al 2014). The clear differences in color of white wild type rice callus versus the yellow-orange lines expressing *OsPSY1* and *OsPDS* suggested carotenoid accumulation. In other investigations callus color were used and demonstrated as an indicator for carotenoid accumulation (Bai et al 2014; Sun et al 2020). Our results are consistent with the experiments in rice callus expressing *ZmPSY1* and *PaCRTI*. Such callus were yellow-orange in color (Bai et al 2014). Our results suggest that the expression of *OsPSY1* and *OsPDS* are the minimum genes necessary to reconstitute the carotenoid pathway in rice callus. However, these preliminary results need to be confirmed in seeds and with the co-transformation of all the pathway genes.

Many biosynthetic pathway genes and transcription factors have been characterized for their roles in regulating secondary metabolites biosynthesis (Patra et al 2013). The second strategy was focused on the regulation of *OsPSY1*, because this gene is the first committed step in the rice biosynthetic pathway (Rodríguez-Concepción 2010; Pulido et al 2012; Bai et al 2014; Giuliano 2017; Yazdani et al 2019). The *cis*-regulatory elements and corresponding transcription factors which control the endosperm-preferred expression of *OsPSY1* have not yet been identified. We hypothesized that important regulatory elements would be located in the upstream promoter region. We searched the *OsPSY1* promoter region and identified the classical *cis*-regulatory elements with differences in a number of nucleotides from the classical sequence. For example, the P-box classical sequence is TGTAAAGT and using bioinformatics tools we identified a variant of the P-box (TGTATGGT) in the promoter region of *PSY1*. This finding suggests that the mutations

in the P-box sequence might be responsible for the lack of expression of *PSY1* in the endosperm. We identified six potential *cis*-regulatory elements in the promoter region of *PSY1*: three prolamin p-boxes, one AACA motif, one GCN4-like motif and an Opaque 2 box (O2 box). Literature reports suggest that the insertion or changes of *cis*-regulatory elements in the promoter sequence can regulate gene expression (Rodríguez-Leal et al 2017; Jin et al 2019; Zhang et al 2019).

We constructed two synthetic promoters using 4- and 6-motif corrected *OsPSY1* gene promoters. We generated a wild type synthetic promoter using the full length flanking region spanning from position –2500 in the *OsPSY1* upstream promoter as negative control (wt*OsPSY1*-*GFP*-TNos), which includes the first intron. As positive control we used an endosperm specific LMWG promoter which has three different variants of a GCN4 like motif and seven different p-box variants of (Hammond-Kosack et al 1993; Juhász et al 2011). We used *GFP* as a reporter for all the constructs. The use of GFP as a reporter has already been demonstrated for the analysis of *cis*-regulatory elements in gene regulation (Zhang et al 2019). Among the four *cis*-regulatory elements found in the 4-motif corrected *OsPSY1* promoter, three P-boxes and a single AACA motif were located within the first 300 bp upstream of the transcription start site, often the key regulatory region in endosperm-specific genes.

Having developed and introduced into rice the constructs for the activation of the *OsPSY1* promoter, we analyzed gene expression and protein accumulation. We confirmed that the reporter gene (*GFP*) was highly expressed and the GFP protein was detected in the transgenic rice callus lines containing the corrected *OsPSY1* promoter 4- and 6-motif. Taken together, these results confirm that the corrected *OsPSY1* promoter activated gene expression compared to the wild type *OsPSY1* promoter. The *cis*-regulatory elements have the potential to regulate *OsPSY1*, and when present they can activate the *OsPSY1* promoter, strongly supporting their role in the regulation of *OsPSY1* in rice endosperm. As expected, the construct containing the wild type *OsPSY1* gene promoter did not significantly expresses *GFP* in rice callus and endosperm. It was demonstrated that *OsPSY1* was not expressed in rice endosperm (Schaub et al 2005) suggesting that the P-box and AACA motif are required for *OsPSY1* promoter activity in rice endosperm. We also investigated the regulatory potential of a synthetic 4-motif *OsPSY1* promoter in transgenic rice endosperm expressing *GFP*. Rice transformed with the synthetic 4-motif *OsPSY1* promoter accumulate GFP in the endosperm, suggesting that the *cis*-regulatory

elements are essential and exhibit tissue specific activation as demonstrated in other reports (Juhász et al 2011; Mithra et al 2017; Xiong et al 2019; Jin et al 2019).

In the absence of the *cis*-transcription factors, GFP accumulation in the transgenic rice lines expressing the wild type *OsPSY1* promoter (wt*OsPSY1*-*GFP*-TNos) was low but detectable in some callus lines, probably reflecting the insertion of the cassette near a strong enhancer that might affect the promoter in the same way as an enhancer trap (Springer 2000). However, the same rice lines with wild type *OsPSY1* promoter (wt*OsPSY1*-*GFP*-TNos) did nor accumulated *GFP* in rice endosperm, as expected because *OsPSY1* is not expressed in rice endosperm (Schaub et al 2005; Bai et al 2014).

Our third strategy for rice biofortification was the used of genes involved in chloroplast division and chromoplast biogenesis. Here we describe preliminary experiments aiming to understand the role of the plastid division factors *AtPDV1* and *AtARC3* in the accumulation of carotenoids in rice endosperm. Our first strategy was to combine the plastid division genes (*AtARC3* and *AtPDV1*) with carotenogenic genes (*ZmPSY1* and *PaCRTI*). *ZmPSY1* and bacterial *CRTI* are the most common genes used in order to increase the accumulation of carotenoids in rice endosperm (Ye et al. 2000; Paine et al. 2005; Baisakh et al. 2006; Aluru et al. 2008; Bai et al. 2016). Unfortunately, this strategy did not result in any significant differences in carotenoid accumulation and chromoplast quantity. Our results were similar to studies that use *ZmPSY1* and *CRTI*. Our best line accumulates around 13 µg/g β-carotene and zeaxanthin being in the average of studies in different varieties of rice that range from 1.0 to 37 µg/g carotenoid accumulation (Paine et al 2005; Parkhi et al 2005; Dong et al 2020). Also the chromoplast number did not show significant differences compared with the wild type rice callus. A possible explanation of the similarity of chromoplast number between wild type and the lines we generated here is that the plastid division genes are specific for chloroplast rather than chromoplast division (Chen et al 2018).

Other strategies for carotenoid accumulation include the *AtOr* gene in combination with *ZmPSY1* and *CRTI* to promote the differentiation of chromoplasts thus generating a metabolic sink that promotes the accumulation of carotenoids in the endosperm (Lu et al 2006; Bai et al 2016). In recent research it was demonstrated that the overexpression of $AtOr^{His}$ in combination with *ZmPSY1* and *CRTI* increased total carotenoids in rice callus from near zero in wild-type rice endosperm to 25.83µg/g DW in the transgenic plants expressing the three genes; β-carotene levels were 10.52µg/g DW (Bai et al 2016). In

order to develop new strategies to enhance carotenoid accumulation in rice endosperm we combined carotenogenic genes with plastid division genes. We combined both plastid division factors (*AtARC3* and *AtPDV1*) with *AtOr*^{His} in order to enhance carotenoid accumulation in the rice endosperm. Co-expression of *AtARC3* or *AtPDV1* with *AtOr*^{His} in Arabidopsis callus was shown to enhance significantly carotenoid accumulation (Sun et al 2020). Our results demonstrate co-transformation of rice with the three genes (*AtARC3*, *AtPDV1* and *AtOr*^{His}). Also we confirmed mRNA accumulation for the three genes in rice callus. Earlier studies showed that the callus system can be used to predict the effect of genes in metabolic pathways (Bai et al 2014). Nonetheless the callus system can be used even with endosperm specific promoters that are active in callus and have restricted activity in mature plants (Bai et al 2014). Further carotenoid and chromoplast analysis will determine whether these lines exhibit enhanced carotenoid content in rice endosperm. Nevertheless, the demonstration that co-expression of *AtARC3*, *AtOr*^{His} and *AtPDV1* affects positively carotenoid levels suggests the potential of plastid division factors to serve as new genetic tools for carotenoid enrichment in plants. However, we are far from identifying all the necessary factors controlling chromoplast division and number. Further experiments will be required to identify and characterize proteins that associate with these factors in order to determine their precise role in the accumulation of carotenoids in the rice endosperm.

6.2. References

Aluru M, Xu Y, Guo R, Wang Z, Li S, White W, Wang K, Rodermel S (2008). Generation of transgenic maize with enhanced provitamin A content. *J Exp Bot*., **59**:3551–3562

Alvarez D, Cerda-Bennasser P, Stowe E, Ramirez-Torres F, Capell T, Dhingra A, Christou P (2021). Fruit crops in the era of genome editing: closing the regulatory gap. *Plant Cell Rep*., **40**:915–930

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

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

Bai C, Twyman RM, Farré G, Sanahuja G, Christou P, Capell T, Zhu C (2011). A golden era—pro-vitamin A enhancement in diverse crops. *Vitr Cell Dev Biol - Plant*., **47**:205–221

Baisakh N, Rehana S, Rai M, Oliva N, Tan J, Mackill DJ, Khush GS, Datta K, Datta SK (2006). Marker-free transgenic (MFT) near-isogenic introgression lines (NIILs) of 'golden' indica rice (cv. IR64) with accumulation of provitamin A in the endosperm tissue. *Plant Biotechnol J*., **4**:467–475

Berman J, Zhu C, Pérez-Massot E, Arjó G, Zorrilla-López U, Masip G, Banakar R, Sanahuja G, Farré G, Miralpeix B, Bai C, Vamvaka E, Sabalza M, Twyman RM, Bassié L, Capell T, Christou P (2013). Can the world afford to ignore biotechnology solutions that address food insecurity? *Plant Mol Biol*., **83**:5–19

Chen C, MacCready JS, Ducat DC, Osteryoung KW (2018). The Molecular Machinery of Chloroplast Division. *Plant Physiol*., **176**:138–151

Dong OX, Yu S, Jain R, Zhang N, Duong PQ, Butler C, Li Y, Lipzen A, Martin JA, Barry KW, Schmutz J, Tian L, Ronald PC (2020). Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat Commun*., **11**:1–10

Giuliano G (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr Opin Biotechnol*., **44**:169–180

Hammond-Kosack MC, Holdsworth MJ, Bevan MW (1993). In vivo footprinting of a low molecular weight glutenin gene (LMWG-1D1) in wheat endosperm. *EMBO J*., **12**:545–554

Jin X, Bai C, Bassie L, Nogareda C, Romagosa I, Twyman RM, 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

Juhász A, Makai S, Sebestyén E, Tamás L, Balázs E (2011). Role of Conserved Non-Coding Regulatory Elements in LMW Glutenin Gene Expression. *PLoS One*., **6**:e29501

Li Q, Farre G, Naqvi S, Breitenbach J, Sanahuja G, Bai C, Sandmann G, Capell T, Christou P, Zhu C (2010). Cloning and functional characterization of the maize carotenoid isomerase and β-carotene hydroxylase genes and their regulation during endosperm maturation. *Transgenic Res*., **19**:1053–1068

Lowe NM (2021). The global challenge of hidden hunger: perspectives from the field. *Proc Nutr Soc*., **80**:283–289

Lu S, Van Eck J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of β-carotene accumulation. *Plant Cell*., **18**:3594–3605

Misawa N, Truesdale MR, Sandmann G, Fraser PD, Bird C, Schuch W, Bramley PM (1994). Expression of a Tomato cDNA Coding for Phytoene Synthase in *Escherichia coli*, Phytoene Formation *In Vivo* and *In Vitro*, and Functional Analysis of the Various Truncated Gene Products. *J Biochem*., **116**:980–985

Mithra SVA, Kulkarni K, Srinivasan R (2017). Plant Promoters: Characterization and Applications in Transgenic Technology BT - Plant Biotechnology: Principles and Applications. In: Abdin MZ, Kiran U, Kamaluddin, Ali A (eds). Springer Singapore, Singapore, pp 117–172

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL, Drake R (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat Biotechnol*., **23**:482–487

Parkhi V, Rai M, Tan J, Oliva N, Rehana S, Bandyopadhyay A, Torrizo L, Ghole V, Datta K, Datta SK (2005). Molecular characterization of marker-free transgenic lines of indica rice that accumulate carotenoids in seed endosperm. *Mol Genet Genomics*., **274**:325–336

Patra B, Schluttenhofer C, Wu Y, Pattanaik S, Yuan L (2013). Transcriptional regulation of secondary metabolite biosynthesis in plants. *Biochim Biophys Acta*., **1829**:1236–1247

Pérez-Massot E, Banakar R, Gómez-Galera S, Zorrilla-López U, Sanahuja G, Arjó G, Miralpeix B, Vamvaka E, Farré G, Rivera SM, Dashevskaya S, Berman J, Sabalza M, Yuan D, Bai C, Bassie L, Twyman RM, Capell T, Christou P, et al (2013). The contribution of transgenic plants to better health through improved nutrition: Opportunities and constraints. *Genes Nutr*., **8**:29–41

Pulido P, Perello C, Rodriguez-Concepcion M (2012). New Insights into Plant Isoprenoid Metabolism. *Mol Plant*., **5**:964–967

Rodríguez-Concepción M (2010). Supply of precursors for carotenoid biosynthesis in plants. *Arch Biochem Biophys*., **504**:118–122

Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB (2017). Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell*., **171**:470–480

Schaub P, Al-Babili S, Drake R, Beyer P (2005). Why Is Golden Rice Golden (Yellow) Instead of Red? *Plant Physiol*., **138**:441–450

Springer PS (2000). Gene traps: tools for plant development and genomics. *Plant Cell*., **12**:1007–1020

Sun T, Yuan H, Chen C, Kadirjan-Kalbach DK, Mazourek M, Osteryoung KW, Li L (2020). OR(His), a Natural Variant of OR, Specifically Interacts with Plastid Division Factor ARC3 to Regulate Chromoplast Number and Carotenoid Accumulation. *Mol Plant*., **13**:864–878

Xiong C, Luo D, Lin A, Zhang C, Shan L, He P, Li B, Zhang Q, Hua B, Yuan Z, Li H, Zhang J, Yang C, Lu Y, Ye Z, Wang T (2019). A tomato B-box protein SlBBX20 modulates carotenoid biosynthesis by directly activating PHYTOENE SYNTHASE 1, and is targeted for 26S proteasome-mediated degradation. *New Phytol*., **221**:279–294

Xudong Y, Salim A-B, Andreas K, Jing Z, Paola L, Peter B, Ingo P (2000). Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science (80-)*., **287**:303–305

Yazdani M, Sun Z, Yuan H, Zeng S, Thannhauser TW, Vrebalov J, Ma Q, Xu Y, Fei Z, Van Eck J (2019). Ectopic expression of ORANGE promotes carotenoid accumulation and fruit development in tomato. *Plant Biotechnol J*., **17**:33–49

Zhang N, McHale LK, Finer JJ (2019). Changes to the core and flanking sequences of G-box elements lead to increases and decreases in gene expression in both native and synthetic soybean promoters. *Plant Biotechnol J*., **17**:724–735

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*., **105**:18232–18237

Zhu C, Naqvi S, Gomez-Galera S, Pelacho AM, Capell T, Christou P (2007). Transgenic strategies for the nutritional enhancement of plants. *Trends Plant Sci*., **12**:548–555

GENERAL CONCLUSIONS

General Conclusions

1. The overexpression of *OsPSY1* and *OsPDS* in rice callus results in a yellow to orange phenotype. Carotenoid analysis in rice callus demonstrated an increased accumulation of carotenoids, specifically a high accumulation of β-carotene suggesting that the minimum genes required to accumulate carotenoids are *OsPSY1* and *OsPDS*
2. Co-transformation of the subsequent genes in the pathway (*OsZDS*, *OsZISO* and *OsCRTISO*) in conjunction with *OsPSY1* and *OsPDS* will be necessary to develop a full understanding of the carotenoid pathway in rice endosperm.
3. We demonstrated experimentally that rice phytoene synthase 1 (*OsPSY1*) is activated by a combination of *cis*-acting elements in the upstream promoter region of the gene.
4. The combination of AACA, P-box, GCN4-like and O2 motifs in the synthetic *OsPSY1* promoter (4-and 6-motif corrected synthetic promoter versions) activates the *OsPSY1* promoter in rice endosperm.
5. In transgenic rice endosperm *GFP* accumulated under the control of a corrected 4-motif *OsPSY1* promoter. In contrast, no GFP accumulation was seen when the wild type *OsPSY1* promoter was used.
6. The use of *cis*-regulatory elements in the promoter region provides an opportunity for the activation of carotenogenic gene expression in metabolically engineered plants.
7. The co-expression of the carotenogenic genes *ZmPSY1* and *PaCRTI* in combination with chloroplast division factors (*AtARC3* and *AtPDV1*) did not achieve an increase in carotenoid levels or chromoplast numbers compared to wild type rice callus.