



COLOUR-SPECIFIC GENES FROM BETALAIN PRODUCING PLANTS

Thèse de doctorat

Présentée à la
Faculté des Sciences de
l'Université de Lausanne

par

Maia ZAIKO

Diplômée en biotechnologie de l'Université
de Génie Chimique de Russie

Jury

Prof. Majed Chergui, Président
Prof. Jean-Pierre Zryd, Directeur de thèse
Dr. Ursula Hinz, Expert
Prof. Frederick Meins, Expert

LAUSANNE
2000



TABLE OF CONTENTS

Résumé	i
Abstract	ii
Abbreviations	iii
I. GENERAL INTRODUCTION	
1. Anthocyanins versus betalains	1
2. Betalain structure	3
3. Biosynthesis of betalains	4
4. Genetic model of betalain biosynthesis	6
5. What is the importance of research on betalain biosynthesis?	7
6. What do we know about enzymes involved in betalain metabolism?	9
7. Attempts to isolate a plant DOPA-dioxygenase	11
8. Aims of the present work and model systems	12
II. RESULTS AND DISCUSSION	
1. Subtractive hybridisation using magnetic beads	
Introduction	19
Results	22
Discussion	25
2. Differential display and RNA fingerprinting	
2.1 mRNA differential display	
Introduction	26
Results	29
Discussion	32
2.2 RNA fingerprinting	
Introduction	34
Results	35
Discussion	40
3. Construction of <i>Beta vulgaris</i> subtractive cDNA library	
Introduction	41
Results	44
Discussion	59
4. Construction of <i>P. grandiflora</i> subtractive cDNA library	
Introduction	62
Results	63
Discussion	90
III. GENERAL CONCLUSIONS	98
VI. MATERIALS AND METHODS	103
Annex	113
References	122

RESUME

Les bétalaïnes sont des pigments hydrosolubles qui s'accumulent dans les vacuoles des plantes de l'ordre des Caryophyllales ainsi que dans certains champignons. Ils représentent un bon système pour étudier le métabolisme secondaire. Les gènes codant pour les enzymes de la biosynthèse des bétalaïnes peuvent être utilisés comme gènes marqueurs grâce à la coloration des produits de la biosynthèse. Aucune enzyme végétale spécifique à la formation des bétalaïnes n'a été isolée à ce jour. Une approche moléculaire a été utilisée dans ce travail afin d'isoler des gènes liés à la présence des bétalaïnes. Elle est basée sur l'hypothèse que les gènes de la biosynthèse des bétalaïnes sont exprimés seulement dans des tissus colorés et que la régulation se fait au niveau du mRNA.

L'hybridation soustractive classique, le « mRNA differential display » et sa modification, le « RNA fingerprinting », n'ont pas permis d'isoler des gènes spécifique aux bétalaïnes. Un protocole modifié de soustraction utilisant la technique de PCR a été appliqué à deux plantes modèles, les cultures de cellules *Beta vulgaris* (*B. v.*) et les pétales de *Portulaca grandiflora* (*P. g.*). La soustraction entre les cDNA des lignées orange et verte de *B. v.* a donné 13 fragments spécifiques à la lignée orange. Leur spécificité a été confirmée par l'analyse Northern blot. Cependant, après le séquençage, tous les clones se sont avérés très similaires. Ils codent pour des protéines riches en glycine (GRP) et ont une grande similitude avec d'autres GRPs.

Le même protocole a été appliqué avec succès aux pétales immatures jaunes et blancs de *P. g.* Une banque de cDNA soustractive a été construite et plusieurs fragments différentiellement exprimés ont été isolés. L'analyse Northern blot a démontré que trois clones sont exprimés dans les pétales violets et jaunes, mais pas dans les blancs. Les cDNA complets ont été isolés par RACE. Le premier cDNA (L.13) code pour une protéine qui présente une grande similitude avec une polyamine oxydase de maïs, une monoamine oxydase d'*Aspergillus niger*, ainsi qu'avec d'autres flavoprotéines. Dans la partie N-terminale, elle contient un « $\beta\alpha\beta$ fold » (typique pour les flavoprotéines) qui lie le résidu ADP du FAD. Le pattern d'expression du L.13 est particulièrement intéressant, car il correspond à celui des gènes spécifiques aux bétalaïnes. Il est notamment exprimé dans les pétales et les tiges, mais pas dans les feuilles. Le second cDNA (V.33) n'a pas de similitude avec des séquences connues. Il est davantage exprimé dans les pétales violets immatures que dans les tiges. Le transcrit est également détectable dans les feuilles. Le troisième cDNA (L.6) présente une similitude avec plusieurs protéines et ORFs inconnus de différents organismes. Un transcrit de 1.3 kb est détecté exclusivement dans les pétales. Deux fragments spécifiques au phénotype jaune ont été isolés et partiellement caractérisés. Aucun des trois cDNAs spécifiques aux phénotypes colorés n'a complété la synthèse des bétalaïnes dans les pétales blancs de *P. g.* Afin d'élucider la fonction de ces gènes, il serait nécessaire de générer des transformants stables et d'entreprendre des études d'expression antisens.

ABSTRACT

Betalains are water-soluble pigments accumulated in the vacuoles of the plants from the order Caryophyllales and in certain mushrooms. They are a good system to study secondary metabolism. Expression of betalain genes gives coloured cells that can be scored visually; therefore, they can be used as marker genes for plant transformation. No enzymes specifically involved in betalain formation have been isolated from plants so far. Therefore, a molecular biology approach was used in the present work. It was assumed that betalain genes are expressed only in coloured tissues, regulation occurring at the mRNA level.

Classical subtractive hybridisation, mRNA differential display and its modification, RNA fingerprinting, failed to detect betalain-specific genes. A modified subtraction procedure taking advantage of the PCR technique was applied to two model systems, *Beta vulgaris* and *Portulaca grandiflora*. Subtraction between cDNA from orange and green cell lines of *B. vulgaris* yielded 13 orange-specific fragments. Their specificity was confirmed by Northern blot analysis. However, sequence analysis revealed that all the clones were related. They code for glycine-rich proteins and have similarity to other GRP.

The same subtraction protocol was successfully applied to *P. grandiflora* system using cDNA from yellow and white immature petals. The resulting subtractive cDNA library was differentially screened and several differentially expressed fragments were found. Northern blot analysis showed that three clones were expressed in violet and yellow petals, but not in white. The full-length cDNAs were isolated by RACE. The first cDNA (L.13) codes for a protein highly similar to a maize polyamine oxidase and a monoamine oxidase from *Aspergillus niger*, as well as to some other flavin-containing enzymes. At the N-terminal part it contains a $\beta\alpha\beta$ fold typical for flavoenzymes binding the ADP-moiety of the FAD. The expression pattern of L.13 gene is particularly interesting and corresponds to our prediction for a betalain-specific gene. It is expressed in colour-specific manner, i. E. in petals and stems, but not in leaves. The second cDNA (V.33) shows no similarity to known sequences. It is highly expressed in immature violet petals and less in stems. The transcript is also detectable in leaves. The third cDNA (L.6) has similarity to several unknown proteins and ORFs in different organisms. A 1.3 kb transcript is exclusively detected in petals. Two yellow-specific fragments were also isolated and partially characterised. None of the three colour-specific cDNAs complemented the betalain synthesis in white *P. g.* petals by ballistic transformation. However, generation of stable transformants and antisense expression studies must be done to elucidate the function of these genes.

ABBREVIATIONS

6-BAP	6-benzylaminopurine
BLAST	Basic Local Alignment Research Tool
bp	base pair
cDNA	complementary deoxyribonucleic acid
Ci	curie
Da	Dalton
DAO	diamine oxidase
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
Dod	DOPA-dioxygenase
DOPA	dihydroxyphenylalanine
ds	double stranded
DTT	dithiothreitol
EB	extraction buffer
EDTA	ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
FAD	flavin adenine dinucleotide
GRP	glycine-rich protein
HPLC	high-performance liquid chromatography
IPTG	isopropyl-1-thio- β -D-galactoside
kb	kilobase
LB	Luria-Bertani medium
MAO	monoamine oxidase
MOPS	3-(N-morpholino)propane sulfonic acid
mRNA	messenger ribonucleic acid
NAD	nicotinamide adenine dinucleotide
o/n	over night
ORF	open reading frame
PAO	polyamine oxidase
PCR	polymerase chain reaction
pfu	plaque forming units
PVP	polyvinylpyrrolidone
RACE	race amplification of cDNA ends
RNase	ribonuclease
rRNA	ribosomal ribonucleic acid
RT	reverse transcriptase
SDS	sodium dodecyl sulfate
SSC	sodium chloride/sodium citrate buffer
Tris HCl	tris (hydroxymethyl) aminomethane hydrochloride
UV	ultraviolet