

THESE

PRESENTEE A

L'UNIVERSITE BORDEAUX 1

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

PAR

Thilia FERRIER

POUR OBTENIR LE GRADE DE

DOCTEUR

SPECIALITE : BIOLOGIE VEGETALE

**Les facteurs de transcription MYB et la régulation
de la biosynthèse des flavonoïdes dans la baie de
raisin: analyse fonctionnelle et identification de
nouveaux candidats**

Soutenu le 14 novembre 2008

Après avis de :

Mr. David VENDEHENNE	Professeur de l'Université de Bourgogne	Rapporteur
Mme Frédérique PELS	Chargée de recherche INRA, Colmar	Rapporteur

Devant la commission d'examen formée de :

M. Serge DELROT	Professeur de l'Université de Bordeaux 2	Président
Mr. David VENDEHENNE	Professeur de l'Université de Bourgogne	Rapporteur
Mme Frédérique PELS	Chargée de recherche INRA, Colmar	Rapporteur
M. Jean Michel MERILLON	Professeur de l'Université de Bordeaux 2	Examineur

Membres invités:

Mr. François BARRIEU	MDC à l'Université de Bordeaux 1	Dir. de thèse
Mme. Muriel BARTHE	Directrice du Service Technique au CIVB	

Résumé

Les flavonoïdes (anthocyanes, flavonols et proanthocyanidines) sont des éléments clés de la qualité organoleptique des baies de raisin. Chez les végétaux, l'expression des gènes de la voie de biosynthèse de ces composés est contrôlée par des complexes protéiques organisés autour des facteurs de transcription de type MYB. Dans le cadre de cette thèse, une première approche s'est intéressée aux mécanismes de régulation de l'expression du gène *VvMyb5a* et de l'activité biologique de la protéine codée par ce gène. L'analyse du promoteur *VvMyb5a* a montré que son activité au cours du développement de la baie serait plutôt placée sous contrôle hormonal. Des expériences de double hybride ont révélé que la protéine VvMyb5a pouvait interagir avec une protéine kinase de type GAMYB et une protéine WD40. Une deuxième approche, basée sur l'analyse globale du transcriptome de mutants naturels de vigne affectés dans la biosynthèse des anthocyanes, a permis d'identifier deux nouveaux gènes *MYB* nommés *VvMybPA1* et *VvMyb24*. L'expression différentielle de ces gènes dans des baies de cépages rouges et blancs a été confirmée et leurs caractérisations fonctionnelles ont été engagées chez *Arabidopsis thaliana*.

Mots clés

Baie de raisin, qualité, flavonoïdes, facteur de transcription, MYB, mutants naturels de vigne

Abstract

Flavonoids, like anthocyanins, flavonols and condensed tannins, are key elements of the organoleptic quality of grape berries. In plants, expression of genes encoding enzymes of the flavonoid biosynthetic pathway is controlled by small protein complexes organised around MYB transcription factors. In the present work, we first focused on the regulatory mechanisms of *VvMyb5a* expression and on the biological activity of the corresponding protein. Promoter analysis indicated that *VvMyb5a* expression is probably mainly controlled by hormones. A yeast two-hybrid screen revealed that VvMyb5a can interact with a GAMYB type protein kinase and a WD40 protein. In a second time, global transcriptome analysis of grapevine natural mutants deficient in anthocyanin biosynthesis led to the identification of two new MYB genes, named *VvMybPA1* and *VvMyb24*. Differential expression of these two genes in red and white berry skins was confirmed by RT-PCR and their functional characterizations have been initiated in *Arabidopsis thaliana*.

Keywords

Grape berry, quality, flavonoids, development, transcription factor, MYB, grapevine natural mutants.

ABRÉVIATIONS

°C	Degré Celsius	dTTP	Déoxythymidine triphosphate
A	Absorbance	DTT	1,4-Dithiothréitol
aa	Acides aminés	EBG	“ Early Biosynthetic Genes ”
ABA	Acide abscissique	EDTA	Acide éthylènediaminetétraacétique
ACT	Acétyltransférases	EST	Expressed Sequence Tag
aa-dUTP	5-(3-aminoallyl)-2'deoxyuridine 5'triphosphate	F3H	Flavanone-3-hydroxylase
AD	“ Activation domain ”	F3'H	Flavonoïde 3'-hydroxylases
ADN	Acide désoxyribonucléique	F3'5'H	Flavonoïde 3'5'-hydroxylases
ADNc	ADN complémentaire	FLS	Flavonol synthase
ADNg	ADN génomique	FT	Facteurs de transcription
ADN-T	ADN transféré	GA	Acide gibbéréllique
AMPc	Adénosine 5'-monophosphate	GMPc	Guanoside 5'-monophosphate cyclique
ANR	Anthocyanidine réductase	GST	Glutathione S-transferase
ANS	Anthocyanidine synthase	GTP	Guanosine TriPhosphate
<i>Arabidopsis</i>	<i>Arabidopsis thaliana</i>	GUS	β-glucuronidase
ARN	Acide ribonucléique	h	Heure
ARNm	ARN messenger	HPLC	Chromatographie liquide à haute performance
ATP	Adénosine triphosphate	IAA	Alcool isoamylique
Aux	Auxine	IPTG	Isopropyl-β-thiogalactopyranoside
BB	Pellicule blanche du Béquignol mutant	Inr	Initiateur
BD	“ Binding domain ”	j	Jours
BET	Bromure d'éthidium	Kan	kanamycine
bHLH	Région basique/Hélice-boucle-Hélice	Kb	Kilobase
BR	Pellicule rouge du Béquignol mutant	kDa	Kilodalton
C4H	Cinnamate 4-hydroxylase	L	Litre
CaMV	“ cauliflower mosaic virus ”	LAR	Leucoanthocyanidine réductase
CHI	Chalcone isomérase	LB	Milieu de Luriani-Bertani
CHS	Chalcone synthase	LBG	“ Late Biosynthetic Gene ”
4CL	4-coumarate CoA ligase	LDOX	Leucoanthocyanidine dioxygénase
CTAB	Bromure d'hexadécyltriméthylammonium	min	Minutes
cv	Cultivar	mL	Millilitre
Cy3	Cyanine 3	mM	Millimolaire
Cy5	Cyanine 5	MS	Milieu Murashige et Skoog
Da	Dalton	MYB	Myéloblastosis

dATP	Déoxyadénosine triphosphate	NASC	Nottingham Arabidopsis Stock Center
DEPC	Diéthylpyrocarbonate	nm	Nanomètre
DFR	Dihydroflavonol réductase	NOS	Nopaline synthase
DNase	Désoxyribonucléase	NptII	Néomycine phosphotransférase II
dNTP	Désoxynucléoside 5'-triphosphate	OligodT	Oligodésoxyribonucléotide
DO	“ Dropout solution ”	OMT	O-méthyltransférase
dTTP	Déoxythymidine triphosphate	p/v	poids/volume
DTT	1,4-Dithiothréitol	pb	Paire de base
EBG	“ Early Biosynthetic Genes ”	PCR	Réaction de polymérisation en chaîne
EDTA	Acide éthylènediaminetétraacétique	Pfu	<i>Pyrococcus furiosus</i>
EST	Expressed Sequence Tag	pmol	Picomoles
F3H	Flavanone-3-hydroxylase	rpm	Rotation par minute
Rpm	Rotation par minute	TC	Tentative consensus
PA	Proanthocyanidines	TE	Tampon tris-EDTA
PAL	Phényalanine ammoniac-lyase	TIGR	The Institute for Genomic Research
PN	Pinot noir	Tm	“ Melting Temperature ”
PB	Pinot blanc	Tris	tri-(hydroxyméthyl) amino méthane
PVPP	Polyvinyl polypyrrolidone	TSS	Site d'initiation de la transcription
RE	Réticulum endoplasmique	U	Unités
RER	Face cytosolique du RE	UFGT	UDP Glucose Flavonoïd 3-O-Glucosyltransférase
RT-PCR	Transcription réverse suivie d'une PCR	µg	Microgramme
S	Seconde	µM	Micromolaire
SA	Acide salicylique	µL	Microlitre
SDS	Sodium dodécyle sulfate	U.V.	Ultraviolet
SSC	Standard sodium citrate	v/v	volume/volume
TAIR	The Arabidopsis Information Resource	WT	“ wild type ”
TAE	Tampon tris-acétate-EDTA	X-α-gal	5-bromo-4-chloro-3-indoïl α-D galactoside
Taq	<i>Thermophilus aquaticus</i>	X-gal	5-bromo-4-chloro-3-indoïl β-D galactoside

PLAN DETAILLE

AVANT-PROPOS.....1

Le Fruit de la vigne.....2
De la vigne au vin.....2
Le vin et la santé.....3
Le raisin, source de jouvence.....4
Objectifs du travail de thèse.....5

SYNTHESE BIBLIOGRAPHIQUE6

I- Présentation de la Vigne.....7

- 1) Systématique.....7
- 2) Notion de variétés, cépages, clones, cultivars8
- 3- Cycle végétatif de la vigne8
- 4) Morphologie et anatomie de la baie de raisin.....10
 - a- La pellicule10
 - b- La pulpe11
 - c- Les pépins11
 - d- Les faisceaux libéro-ligneux.....12
- 5) Physiologie de la baie de raisin12
 - a- Cycle de développement de la baie de raisin12
 - Phase I ou croissance herbacée*.....12
 - Phase II ou maturation*13
 - b- Aspects biochimiques du développement de la baie de raisin.....14
 - Acidité de la baie*14
 - Accumulation des sucres*.....15
 - Accumulation des composés aromatiques*16
 - c- Contrôle hormonal du développement de la baie17
 - Hormones intervenant dans la mise en place de la baie de raisin*18
 - Hormones intervenant au cours de la maturation de la baie de raisin*18

II- Métabolisme des Flavonoïdes dans la baie de raisin.....20

- 1- Flavonoïdes chez les végétaux21
- 2- Biosynthèse des flavonoïdes chez les végétaux23
 - Voie commune de biosynthèse des flavonoïdes*.....23
 - Biosynthèse des anthocyanes*24
 - Biosynthèse des proanthocyanidines*25
 - Organisation subcellulaire du métabolisme des flavonoïdes*26
 - Transport et compartimentation des flavonoïdes*27
- 3- Composés phénoliques identifiés dans la baie de raisin.....28
 - Les anthocyanes*28
 - Les proanthocyanidines*29
 - Les flavonols*30
- 4- Biosynthèse des flavonoïdes dans la baie de raisin31

III- Régulation transcriptionnelle du métabolisme des flavonoïdes.....33

- 1- Mécanisme général de la régulation des gènes chez les eucaryotes.....33
 - Les promoteurs et leurs éléments de régulation*33

<i>Les facteurs de transcription</i>	34
2- Aspects généraux sur la régulation de la voie de biosynthèse des flavonoïdes.....	35
a- Les éléments <i>cis</i> -régulateurs de la voie des flavonoïdes	35
b- Les facteurs de transcription régulateurs de la voie des flavonoïdes	37
i- Facteur de transcription MYB	37
<i>Identité des MYB</i>	37
<i>Les MYB chez les végétaux</i>	39
ii- Facteur de transcription bHLH.....	40
<i>Identité des bHLH</i>	40
<i>Les bHLH chez les végétaux</i>	40
c- Les WD40, une autre famille de régulateur de la voie des flavonoïdes	41
3-Gènes impliqués dans la régulation des gènes de biosynthèse des anthocyanes.....	43
<i>Régulation de la biosynthèse des anthocyanes chez les espèces modèles</i>	43
<i>Activateurs de la biosynthèse des anthocyanes</i>	43
<i>Répresseurs de la biosynthèse des anthocyanes</i>	44
<i>Régulation de la biosynthèse des anthocyanes dans la baie de raisin</i>	45
4-Gènes impliqués dans la régulation des gènes de biosynthèse des PA	46
<i>Régulation de la biosynthèse des PA chez Arabidopsis</i>	46
<i>Régulation de la biosynthèse des PA dans la baie de raisin</i>	47
5-Gènes impliqués dans la régulation des gènes de biosynthèse des flavonols.....	47

CHAPITRE 1:

Caractérisation des mécanismes régulateurs de l'expression du gène *VvMyb5a* et de l'activité de la protéine correspondante.....48

I- Introduction : Etat de l'art sur *VvMyb5a*49

II- Recherche de partenaires protéiques de la protéine *VvMyb5a*50

1) Analyse <i>in silico</i> de la séquence <i>VvMyb5a</i>	50
2) Recherche d'interacteurs protéiques du domaine GRD de <i>VvMyb5a</i>	52
a- Principe de la technique du double hybride chez la levure.....	52
b- Clonage du domaine GRD dans le plasmide proie et tests préliminaires.....	54
i- Clonage du domaine GRD.....	54
ii- Test préliminaires : toxicité et auto-activation	54
<i>Le peptide GRD est-il toxique pour les levures transformées ?</i>	55
<i>Le peptide GRD est-il un auto-activateur ?</i>	55
c-Criblage de la banque	55
d- Résultats du séquençage des clones positifs.....	56

III Etude fonctionnelle du promoteur de *VvMyb5a*58

1) Définition biologique du promoteur reconnu par l'ARN polymérase II	58
2) Clonage et analyse <i>in silico</i> du promoteur <i>VvMyb5a</i>	60
a- Clonage du promoteur <i>VvMyb5a</i>	60
b- Analyse <i>in silico</i> de la séquence promotrice de <i>VvMyb5a</i>	61
3) Recherche de motifs consensus du promoteur proximal de <i>VvMyb5a</i>	61
a- Recherche des motifs spécifiques du promoteur <i>VvMyb5a</i>	62
b- Boîtes de réponses aux hormones.....	62
<i>Motifs de régulation par l'ABA</i>	62
<i>Coopération de motifs cis-régulateurs pour une signalisation par les gibbérellines</i>	63
<i>Motif de régulation par d'autres hormones</i>	64

c- Boîtes de réponse aux stress environnementaux.....	65
<i>Motifs impliqués dans le signal lumière</i>	65
<i>Autres motifs impliqués dans les facteurs abiotiques</i>	66
d- Boîtes de réponse aux sucres	67
4- Dissection fonctionnelle du promoteur <i>VvMyb5a</i>	67
a- Clonage des régions promotrices dans le vecteur d'expression transitoire pAM35.....	69
b- Tests préliminaires.....	69
5- Identification d'éléments <i>trans</i> -régulateurs par la technique du Simple Hybride chez la levure.....	70
a- Principe de la technique du simple hybride chez la levure.....	70
b- Clonage des régions promotrices cibles dans le vecteur d'expression pHIS2	71
c- Tests préliminaires.....	71
d- Criblage de la banque	72
e- Résultats.....	73
IV- Discussion et perspectives	73
<i>Régulation transcriptionnelle du gène VvMyb5a</i>	73
<i>Les approches de criblage à grande échelle...</i>	76

CHAPITRE 2

Identification de nouveaux facteurs de transcription MYB par analyse de mutants naturels de la Vigne80

I- Approche gènes candidats sur le Pinot Noir et deux de ses mutants naturels, le Pinot blanc et le Pinot gris	81
1- Analyse de l'expression différentielle du facteur MYB R ₂ R ₃ , <i>VvMybPA1</i> , par la technique de la RT-PCR	82
2- Clonage et analyse <i>in silico</i> de la séquence d'ADNc du gène <i>VvMybPA1</i>	82
a- Clonage de la séquence ADNc du gène <i>VvMybPA1</i> par RACE-PCR.....	82
b- Analyse <i>in silico</i> de la séquence <i>VvMybPA1</i>	84
3- Caractérisation fonctionnelle du gène <i>VvMybPA1 in planta</i>	85
a- Isolement et caractérisation d'un mutant d'insertion d'ADN-T dans le gène <i>VvMybPA1</i>	85
i- Recherche de l'homologue de <i>VvMybPA1</i> chez <i>Arabidopsis thaliana</i>	85
ii- Obtention du mutant d'insertion.....	86
iii- Isolement des lignées mutantes homozygotes.....	87
iv- Vérification du site d'insertion de l'ADN-T dans les mutants sélectionnés	88
v- Confirmation de la perte d'expression d' <i>AtMyb82</i>	88
vi- Analyse préliminaire du phénotype du mutant KO62B	89
b- Obtention et caractérisation des plantes transgéniques présentant une surexpression du gène <i>VvMybPA1</i>	90
i- Obtention des transformants et expression du gène <i>VvMybPA1</i>	90
ii- Analyse des transformants T2	91
iii- Analyse phénotypique des transformants T3	92
II- Analyse globale du transcriptome des pellicules du cépage Béquignol mutant	93
1- Caractérisation du phénotype des baies du Béquignol mutant.....	93
a- Analyses phénotypiques et microscopiques	93
b- Analyses de la composition phénolique et de la teneur en sucre des pellicules.....	94
c- Analyses de l'expression de gènes <i>MYB</i>	95

2- Analyse globale comparée des pellicules du Béquignol mutant	95
3- Validation de l'expression différentielle de <i>CB913371</i> par RT-PCR semi-quantitative...	97
4- Identification et caractérisation fonctionnelle de <i>CB913371</i>	98
a- Clonage de la séquence d'ADNc du gène <i>CB913371</i>	98
b- Analyses <i>in silico</i> de la séquence codante de <i>CB913371</i> et recherche de séquences protéiques homologues	99
c- Profil d'expression de <i>VvMyb24</i> dans les différents organes de la vigne et au cours du développement de la baie de raisin	99

III- Discussion et Perspectives	100
<i>Du séquençage de la vigne à la génomique fonctionnelle...</i>	100
<i>Les mutants de couleurs, une ressource génétique utile pour identifier les régulateurs clés du métabolisme des anthocyanes.</i>	101
<i>VvMybPA1, gène MYB impliqué dans la synthèse des anthocyanes ou des tannins ?</i>	102
<i>VvMyb24, nouveau régulateur MYB de la voie des anthocyanes ?</i>	105

CONCLUSIONS GENERALES ET PERSPECTIVES.....108

MATERIELS ET METHODES112

I- Analyses bioinformatiques	113
1) Recherche de séquences	113
2) Profil d'expression <i>in silico</i>	113
3) Alignement des séquences et obtention des arbres phylogénétiques	113
4) Recherches bioinformatiques ciblées	114
a- Analyse des séquences promotrices.....	114
b- Recherche de domaines protéiques.....	114
II- Matériels	114
1) Amorces.....	114
2) Plasmides	115
a- Vecteur de clonage.....	115
b- Vecteurs d'expression eucaryote	115
i- Plasmides utilisés pour la recherche d'interactions protéiques par la technique du double hybride en levure.....	115
ii- Plasmides utilisés pour la recherche d'interaction ADN-protéine par la technique du simple hybride en levure.....	116
iii- Plasmide utilisé pour les tests d'activation du promoteur.....	117
iv- Plasmide utilisé pour la transgénèse végétale	117
3- Bactéries	117
a- Souches bactériennes	117
b- Milieux de culture.....	118
4- Levures	118
a- Souches de levures.....	118
b- Milieux de culture.....	118
5) Matériel végétal	119
a- Baies de raisin.....	119
i- Cépages.....	119
ii- Procédure de prélèvement	119

b- Graines d' <i>Arabidopsis thaliana</i> et conditions de culture	119
i- Ecotype	119
ii- Condition de culture in vitro	120
iii- Condition de culture en terre	120
III- Techniques de biologie moléculaire	120
1) Extraction des acides nucléiques	120
a- Extraction des ARN totaux	120
i Extraction d'ARN totaux de vigne	120
ii- Extraction d'ARN totaux d' <i>Arabidopsis thaliana</i>	121
iii- Traitement des ARN totaux à la DNase	121
b- Extraction d'ADN génomique d' <i>Arabidopsis thaliana</i>	121
2- Analyses des acides nucléiques	122
a- Quantification des acides nucléiques extraits	122
b- Electrophorèse des acides nucléiques.....	122
c- Réaction de transcription inverse (RT).....	122
d- Réaction de polymérisation en chaîne (PCR).....	122
3) Analyse du niveau d'expression des transcrits par RT-PCR.....	123
4) Clonage moléculaire	124
a- Préparation des fragments à cloner.....	124
i- Amplification et purification des fragments amplifiés par PCR	124
ii- Digestion de l'ADN par des enzymes de restriction	124
b- Ligation dans le vecteur d'intérêt	125
i- Principe de la réaction de ligation	125
ii- Ligation des produits amplifiés dans le vecteur pGEMT®easy.....	125
iii- Ligation des produits digérés dans les autres vecteurs.....	125
c- Transformation de bactéries thermocompétentes par choc thermique	126
i- Préparation des bactéries thermocompétentes.....	126
ii- Transformation de bactéries par choc thermique	126
d- Sélection des bactéries recombinantes	127
e- Vérification des clones positifs.....	127
i- Minipréparation d'ADN plasmidique.....	127
ii- Digestion de l'ADN plasmidique et séquençage des clones positifs	127
5) Clonage des extrémités 5' et 3' d'un gène par la technique de RACE-PCR.....	127
6) Techniques pour le criblage simple et double hybride chez la levure.....	128
a- Préparation et transformation des levures compétentes.....	128
b- Vérification de la toxicité de la protéine ou du peptide appât pour la technique du double hybride	129
c- Vérification de l'auto-activation de la protéine ou du peptide appât pour la technique du double hybride.....	129
d- Détermination de la quantité optimale de 3-AT pour la technique de simple hybride.....	129
e- Amplification de la banque ADNc double et simple hybride.....	130
f- Criblage des interactions simple et double hybride par co-transformation.....	130
g- Calcul de l'efficacité de co-transformation	131
h- Isolation des clones positifs en simple et double hybride chez la levure	131
i- Minipréparation d'ADN plasmidique des clones de levures.....	131
ii- Préparation et transformation des bactéries électrocompétentes.....	132
7) Techniques de microarray	132
a- Principe général de la technique de microarray.....	132
b- Principe d'analyses des lames de microarray	133
c- Type de lames utilisées pour l'analyse microarray.....	134
d- Synthèse des sondes marquées, co-hybridation et lavages des lames	134
i- Synthèse et marquages des sondes	134

ii- Dosage des sondes et détermination de l'efficacité d'incorporation des fluorochromes	135
iii- Co-hybridation et lavages des lames.....	136
<i>Prétraitement des lames</i>	136
<i>Préparation des sondes pour l'injection</i>	136
<i>Co-hybridation et lavages des lames</i>	136
e- Acquisition et analyse des images	137
i- Acquisition des images.....	137
ii- Analyse des images	137
f- Normalisation et statistiques	137
g- Ré-annotation des séquences " spottées " sur les lames	138
i- Mise à jour des annotations	138
ii- Catégories fonctionnelles	139

IV- Méthodes de transgénèse et d'analyses des plantes transgéniques.....	139
1) Préparation et transformation d'agrobactéries électrocompétentes.....	139
2) Transformation stable d' <i>Arabidopsis thaliana</i> et sélection des plantes	139
3) Méthodes d'analyses des plantes transgéniques	140

V- Méthodes d'analyses histologiques.....	140
1) Réalisation de coupes d'échantillons frais.....	140
2) Observation et acquisition des images.....	141

ANNEXES142

Annexe 1	143
Annexe 2	145
Annexe 3	147
Annexe 4	148
Annexe 5	149
Annexe 6	150
Annexe 7	151
Annexe 8	152
Annexe 9	154
Annexe 10	155
Annexe 11	156
Annexe 12	157
Annexe 13	158
Annexe 14	159

BIBLIOGRAPHIE.....164

PUBLICATIONS ET COMMUNICATIONS SCIENTIFIQUES191

Liste des figures

- Figure 1:** Classification selon Cronquist
- Figure 2:** Classification phylogénétique des Vitacées
- Figure 3:** Classification générale du genre *Vitis*
- Figure 4:** Stades repères phénologiques du cycle végétatif de la vigne
- Figure 5:** Organisation d'une baie de raisin
- Figure 6:** Structure détaillée d'une baie de raisin
- Figure 7:** Schéma représentant le développement de la baie de raisin
- Figure 8:** Répartition tissulaire des composés essentiels à l'élaboration de la qualité organoleptique de la baie de raisin
- Figure 9:** Régulation hormonal du développement de la baie de raisin
- Figure 10:** Structure chimique de base des flavonoïdes
- Figure 11:** Représentation schématique des principales classes de flavonoïdes
- Figure 12:** Structure chimique des six anthocyanes majeures de la baie de raisin
- Figure 13:** Structure chimique des tannins condensés ou proanthocyanidines
- Figure 14:** Structure chimique des principaux flavonols
- Figure 15:** Représentation schématique des voies de biosynthèse des principaux composés phénoliques présents chez les végétaux
- Figure 16:** Modèles d'organisation des enzymes du métabolisme des flavonoïdes en canal métabolique
- Figure 17:** Cinétique d'accumulation des composés phénoliques pendant le développement de la baie de raisin
- Figure 18:** Profil d'expression des gènes *VvLDOX*, *VvANR*, *VvLAR1* et *VvLAR2* au cours du développement de la baie dans les pépins (A) et la pellicule (B)
- Figure 19:** Profil d'expression des gènes de biosynthèse des anthocyanes au cours du développement de la baie dans la pellicule (A) et la pulpe (B)
- Figure 20:** Principaux domaines de liaison à l'ADN des facteurs de transcription
- Figure 21:** Structure tridimensionnelle d'une protéine MYB R₂R₃ interagissant avec la molécule d'ADN
- Figure 22:** Représentation des domaines fonctionnels des protéines de type MYB typiques chez les animaux et chez les végétaux
- Figure 23:** Structure tridimensionnelle d'une protéine bHLH MyoD interagissant avec la molécule d'ADN
- Figure 24:** Structure tridimensionnelle du domaine WD40 de la sous-unité beta de la protéine G hétérotrimérique
- Figure 25:** Représentation schématique du locus qui détermine la couleur des baies de Cabernet sauvignon
- Figure 26:** Différences génétiques à l'origine des cépages rouges et blancs
- Figure 27:** Représentation schématique de l'activité du complexe MYB-bHLH-WD40 impliqué dans la biosynthèse des proanthocyanidines dans les graines d'*Arabidopsis*
- Figure 28:** Activation des promoteurs des gènes codant les enzymes de la voie de biosynthèse des flavonoïdes chez la vigne par *VvMyb5a*
- Figure 29:** Implication des facteurs de transcriptions MYB R₂R₃ dans les mécanismes régulateurs de la biosynthèse des flavonoïdes au cours du développement de la baie de raisin
- Figure 30:** Analyse bioinformatique du locus *VvMyb5a*.
- Figure 31:** Analyse des similarités observées entre *VvMyb5a* et d'autres protéines MYB de plantes
- Figure 32:** Principe de la technique du double hybride chez la levure
- Figure 33:** Représentation schématique des domaines protéiques conservés dans la protéine *VvMyb5a* et du peptide cible utilisé pour l'approche double hybride.
- Figure 34:** Représentation schématique du crible effectué pour identifier les clones positifs en double hybride contre le domaine GRD de *VvMyb5a*.

- Figure 35:** Comparaison des séquences protéiques de GSVIVP0003204001 avec deux membres de la famille des protéines Mak kinase.
- Figure 36:** Représentation schématique de la structure des promoteurs de plantes.
- Figure 37:** Représentation schématique de la position des motifs consensus du promoteur proximal *VvMyb5a*
- Figure 38:** Représentation schématique de la position des éléments de réponse à l'acide abscissique identifiés dans le promoteur *VvMyb5a*.
- Figure 39:** Représentation schématique de la position des éléments de réponse aux gibbérellines identifiés dans le promoteur *VvMyb5a*.
- Figure 40:** Représentation schématique de la position des éléments de réponse aux hormones éthylène, auxine et acide salicylique identifiés dans le promoteur *VvMyb5a*.
- Figure 41:** Représentation schématique de la position des éléments de réponse à la lumière identifiés dans le promoteur *VvMyb5a*.
- Figure 42:** Représentation schématique de la position des éléments de réponse aux facteurs abiotiques dans le promoteur *VvMyb5a*.
- Figure 43:** Représentation schématique de la position des éléments de réponse aux sucres dans le promoteur *VvMyb5a*.
- Figure 44:** Représentation schématique des régions promotrices de *VvMyb5a* impliquées dans la réponse aux hormones, aux sucres et à la lumière.
- Figure 45:** Fusions transcriptionnelles entre divers fragments du promoteur du gène *VvMyb5a* délétés en 5' et le gène rapporteur de la β -glucuronidase (*GUS*).
- Figure 46:** Activité transcriptionnelle du promoteur *VvMyb5a* et des fragments délétés en 5' dans des protoplastes d'*Arabidopsis thaliana*.
- Figure 47:** Principe de la technique de criblage simple hybride chez la levure pour identifier des facteurs *trans*-régulateurs.
- Figure 48:** Représentation schématique du crible effectué pour identifier les clones positifs en simple hybride contre le fragment S₄ de *VvMyb5a*.
- Figure 49:** Analyse par RT-PCR semi-quantitative de l'expression de *VvMybPA1* au cours du développement des baies de Pinot noir, gris et blanc.
- Figure 50:** Analyse par électrophorèse en gel d'agarose des amplifications des extrémités 5' et 3' de l'ADNc de *VvMybPA1* par RACE-PCR.
- Figure 51:** Analyse par électrophorèse des amplifications de la séquence codante et génomique de *VvMybPA1*.
- Figure 52:** Analyse de la séquence génomique de *VvMybPA1*.
- Figure 53:** Comparaison des séquences protéiques de *VvMybPA1* avec deux membres de la famille des protéines MYB.
- Figure 54:** Analyse phylogénétique de *VvMYB24* et des 126 protéines MYB R2R3 d'*Arabidopsis thaliana*
- Figure 55:** Représentation schématique de l'ADN-T du vecteur pAC106 utilisé par GABI-kat.
- Figure 56:** Analyse par PCR du génotype des mutants d'insertion ADN-T pour le gène *AtMyb82* (*At5g52600*).
- Figure 57:** Identification du site d'insertion exact de l'ADN-T dans le mutant KO62B.
- Figure 58:** Expression du gène *AtMyb82* dans différents organes d'*Arabidopsis thaliana*
- Figure 59:** Vérification de l'absence de transcrits *AtMyb82* dans le mutant d'insertion KO62B par RT-PCR.
- Figure 60:** Analyse phénotypique préliminaire du mutant KO62B
- Figure 61:** Représentation schématique de l'ADN-T utilisé pour la surexpression de *VvMybPA1*
- Figure 62:** Phénotype observé sur les transformants 35S::*VvMybPA1*.
- Figure 63:** Analyse de l'expression du gène *VvMybPA1* dans les plantes transgéniques 35S::*VvMybPA1* de génération T3.
- Figure 64:** Photographies de grappes de baies du groupe variétal du cépage Béquignol
- Figure 65:** Observations des couches cellulaires de la pellicule de baie du Béquignol mutant, Béquignol et Béquignol blanc.
- Figure 66:** Analyse de la teneur en anthocyanes des pellicules rouges de Béquignol mutant et des pellicules de Béquignol rouge par HPLC.
- Figure 67:** Analyse de la teneur en flavonols des pellicules de Béquignol mutant, Béquignol rouge et Béquignol blanc par HPLC.
- Figure 68:** Analyse de la teneur en sucre des pulpes de Béquignol mutant, Béquignol rouge et Béquignol blanc.

- Figure 69:** Analyse par RT-PCR semi-quantitative des profils d'expression des gènes *VvMybA1*, *VvMybA3* et *VvMybPA1* dans les pellicules blanches et rouges du Béquignol mutant.
- Figure 69:** Validation de l'expression différentielle de *CB923371* par RT-PCR semi-quantitative dans les pellicules des baies de Pinot noir, de Pinot blanc et de Béquignol mutant.
- Figure 70:** Profil d'expression par RT-PCR semi-quantitative de *CB913371* dans les pellicules, les pulpes et les baies épépinées des cépages Pinot noir et blanc.
- Figure 71:** Analyse de la séquence génomique de *VvMyb24*.
- Figure 72:** Comparaison des séquences protéiques de *CB913371* et d'*AtMyb24*.
- Figure 73:** Analyse par RT-PCR semi-quantitative de l'expression de *VvMyb24* dans des organes de vigne et dans des baies de raisin à différents stades de développement.
- Figure 74:** Principe et différentes étapes de la technique microarray

Liste des tableaux

- Tableau I:** Protéines MYB, bHLH et WD40 impliquées dans le contrôle de la synthèse des anthocyanes chez les espèces modèles et dans la baie de raisin
- Tableau II:** Protéines MYB, bHLH et WD40 impliquées dans le contrôle de la biosynthèse des proanthocyanidines chez les espèces modèles et dans la baie de raisin.
- Tableau III:** Représentation schématique des résultats des tests d'autoactivation.
- Tableau IV:** Interacteurs protéiques possibles du domaine GRD de la protéine *VvMyb5a* identifiés par la technique du double hybride en levure
- Tableau V:** Analyse comparative des motifs *cis*-régulateurs identifiés dans le promoteur *VvMyb5a* par les outils de prédiction PLACE et MatInspector.
- Tableau VI:** Liste des éléments *cis*- et *trans*-régulateurs de réponse aux hormones identifiés dans le promoteur de *VvMyb5a* par les outils de prédiction PLACE et MatInspector.
- Tableau VII:** Liste des éléments *cis*- et *trans*-régulateurs impliqués dans le signal lumière identifiés par les outils de prédiction PLACE et MatInspector.
- Tableau VIII:** Liste des éléments *cis*- et *trans*-régulateurs impliqués dans différents stress abiotiques identifiés par les outils de prédiction PLACE et MatInspector.
- Tableau IX:** Liste des éléments *cis*- et *trans*-régulateurs impliqués dans la réponse aux sucres identifiés par les outils de prédiction PLACE et MatInspector.
- Tableau X:** Eléments *trans*-régulateurs interagissant avec le promoteur *VvMyb5a*, identifiés par la technique du simple hybride chez la levure.
- Tableau XI:** Souches de microorganismes utilisés.

AVANT-PROPOS

L'histoire de la vigne et du vin accompagne l'histoire de l'humanité depuis des millénaires. Comme l'homme, le cep est divers, changeant, souvent imprévisible. Le vin est fruit de la terre comme du travail des hommes, lié à ceux-ci par une complicité profonde. Un néophyte voit dans la vigne des plants bien ordonnés, alignés tels des militaires, tous identiques. Pourtant, chaque cep, résultat biologique de lentes mutations génétiques, est différent et vit son aventure individuelle tout en étant encadré afin d'éviter que la nature ne reprenne le dessus.

Le Fruit de la vigne...

Le raisin est un fruit exceptionnel : en tant qu'aliment énergétique, il est facilement digeste et assimilable par l'organisme. Ses atouts nutritifs sont ainsi reconnus depuis l'antiquité. Il est le 2^{ème} fruit dont parle la Bible, après la pomme. Né en Chine 2000 ans avant notre ère, il est l'un des fruits les plus anciennement connus. Riche en symbole, il représente la vie dans les tombeaux des pharaons et devient la figure du culte de Dionysos dans la Grèce antique. Jusqu'alors, la vigne était essentiellement destinée à la vinification et c'est seulement au XVI^{ème} siècle qu'elle sera considérée pour d'autres usages. C'est en outre François Ier qui, recevant en cadeau du Chasselas de la part de Soliman le Magnifique, fait entrer ces belles grappes à Fontainebleau. Le "raisin de table" acquiert ses lettres de noblesse et devient dessert du roi. A la fin du XIX^{ème}, le secteur viticole français traverse une crise. Sous l'égide des viticulteurs en quête de nouveaux débouchés, la production dédiée au raisin de table prend alors son plein essor. Il connaît au cours du XX^{ème} siècle un développement considérable, lié notamment aux évolutions des moyens de transport et de commercialisation.

De la vigne au vin...

Le vin est un produit unique. Tout au long des siècles, le vin reste un élément de fête, de culture ou une libation de choix pour préconiser une meilleure santé, avant même la découverte du *French Paradox* [1]. Boire du vin, toujours avec modération, c'est aussi rafraîchir sa mémoire culturelle. Il est, selon Colette "l'honneur des mets" ou, selon Alexander Fleming "ce qui rend les hommes heureux" (la pénicilline ne faisant que "guérir les humains"). Un pays moderne se doit de lui conserver son prestige et d'encourager tous les efforts qui tendent à en améliorer la qualité. Le vin est, et doit demeurer un produit de "

distinction » (cf. Pierre Bourdieu) dans un pays comme la France qui lui a toujours reconnu une place de choix.

L'usage du vin par les Gaulois a permis à nos ancêtres de s'intégrer dans la civilisation du vin et par là même de se distinguer culturellement des peuples de l'Europe du Nord buveurs de bière. Mais nulle part ailleurs, autant qu'en France, la liqueur de Bacchus n'a été élevée au rang de « boisson totem ». Dans un essai intitulé *Mythologies*, Roland Barthes écrit en 1957 que: « le vin est senti par la nation française comme un bien qui lui est propre, au même titre que ses 360 espèces de fromages et sa culture ». Trente ans plus tard à la question : « *Être Français, c'est selon vous d'abord... ?* », la réponse : « *Aimer le bon vin* » vient à la suite d'évidences comme « *Être né en France* » ou « *Parler français* » (selon un sondage réalisé en mai 1987 sur « Les Français et leur histoire » pour le n°100 du mensuel *L'Histoire*). Autrement dit, le vin apparaît à nos compatriotes comme un élément constitutif de la « francitude » voire comme un mythe fondateur de la nation française. Le flacon de vin est aussi un « lieu de mémoire ».

Car la France est le pays de référence de la civilisation du vin en même temps que la patrie de la gastronomie, laquelle culmine dans ses vins. Mais indépendamment de sa région natale ou du statut économique du consommateur, tous les vins sont attendus comme une expérience orgasmique [2]. Ainsi, la volonté des producteurs français est d'être irréprochable sur la qualité du produit et comme le dit Gérard Bertrand (leader qualitatif des vins premiums du sud de la France), « Le vin est grand, quand la main de l'homme a révélé le terroir qui l'a engendré ».

Le vin et la santé...

Les vertus thérapeutiques du vin ont été reconnues dès l'Antiquité. La plupart des médicaments que prescrivait Hippocrate étaient à base de vin, bien qu'à cette époque, aucun fondement scientifique ne le justifiait. A l'époque où l'eau était le véhicule de maladies et d'infections, Pasteur disait que le vin était la plus saine et la plus hygiénique des boissons. Depuis la découverte du « french paradox », de nombreuses études ont été menées sur l'alcool et ses effets sur la santé humaine. Des études épidémiologiques se sont multipliées, les plus significatives sont celles menées dans le Nord-est de la France sous l'égide du Pr. Serge Renaud, et celles menées au Danemark par Gronbaek et *al.* [3]. Ces études ont permis de montrer que les composés phénoliques, en dehors de leurs effets sensoriels et leur

contribution à la couleur du vin, ont des propriétés antimutagènes, anticarcinogènes, antiathérogènes et anticoagulantes. Leurs actions thérapeutiques interviendraient contre certaines pathologies chroniques comme l'athérosclérose, le diabète, l'hypertension et certains cancers. Des recherches sont encore nécessaires pour expliquer les effets observés lors des enquêtes épidémiologiques. Il est incontestable que l'abus de boissons alcoolisées a fait des ravages dans la société en étant à la fois source de problèmes de santé et de drames humains. Mais comme le recommande Arnaud de Villeneuve, « Buvez-en peu, mais qu'il soit bon, le bon vin sert de médecin, le mauvais vin est un poison ».

Le raisin, source de jeunesse...

On connaissait les vertus d'un verre de vin par jour, mais on ne soupçonnait pas toutes les richesses contenues dans un minuscule pépin de raisin. L'emploi du raisin en cosmétologie remonte au moins au XVII^e siècle, à la cour de Louis XIV, où il était à la mode de s'appliquer du vin vieilli sur le visage pour donner au teint de l'éclat. De même, son action éclaircissante était bien connue des vigneron français. Dans les années 90, le professeur Vercauteren ouvre la porte à un nouvel art de prendre soin de son corps et le petit grain rond recouvert de pruine, entra dans la Vinothérapie [4].

Le secteur viticole est l'un des secteurs les plus importants de la production agricole française. En 2005, la superficie du vignoble français était de 890 000 hectares la plaçant à la deuxième place derrière l'Espagne. Pourtant la France reste le premier producteur mondial du vin [5]. Si la part de l'Europe dans la production mondiale reste importante (74%), elle est en constante diminution depuis le milieu des années 90. D'une part, les vins des pays dits du Nouveau Monde (USA, Argentine, Chili, Australie et Afrique du Sud) sont arrivés sur le marché mondial. D'autre part, toutes les grandes zones viticoles mondiales connaissent une baisse de leur production face aux fluctuations climatiques. De fait, la volonté des producteurs français est d'être irréprochable sur la qualité du produit. Cela passe par l'expérimentation (choix du cépage), la conduite de la vigne, la qualité phytosanitaire des fruits, le respect de l'environnement, la rénovation variétale, une meilleure adaptabilité de la vigne aux contraintes environnementales. Pour pouvoir améliorer la culture de la vigne, il est impératif de connaître et comprendre les mécanismes moléculaires, biochimiques et physiologiques qui déterminent la qualité organoleptique des baies de raisin.

Objectifs du travail de thèse

Les flavonoïdes, et plus particulièrement les anthocyanes et les tannins condensés, sont des métabolites secondaires jouant un rôle important dans l'élaboration de la qualité organoleptique des baies de raisin et, *in fine*, des vins. La voie de biosynthèse de ces composés a été largement étudiée au cours de ces dernières années non seulement chez les espèces modèles et mais également chez la vigne. Cependant, les mécanismes moléculaires de contrôle de la voie de biosynthèse des flavonoïdes restent encore mal définis. A l'heure actuelle, des protéines de type MYB, bHLH, ou encore WD40 apparaissent impliquées dans la régulation de l'expression des gènes codant les enzymes de la voie de biosynthèse des flavonoïdes. Au début de ce travail de thèse, seuls trois facteurs de transcription MYB (*VvMyb5a*, *VvMyb5b* et *VvMybA1*) avaient été identifiés chez la vigne. Les gènes *VvMyb5a* et *VvMyb5b* avaient été isolés et caractérisés par L. Deluc lors d'un précédent travail de thèse dans le laboratoire.

Ainsi, dans un premier temps, nous avons poursuivi le travail engagé sur le gène *VvMyb5a*. Nous avons cherché à identifier des protéines régulatrices de l'expression de ce gène et de l'activité biologique de la protéine *VvMyb5a*. Cette approche a été réalisée en utilisant les techniques de simple et double hybride chez la levure. En parallèle, les mécanismes de contrôle de l'expression de *VvMyb5a* ont été étudiés grâce à une dissection fonctionnelle du son promoteur *via* une analyse *in silico* et le clonage de plusieurs régions du promoteur en amont du gène rapporteur *GUS* (β -glucuronidase).

Dans un second temps, nous avons recherché de nouveaux régulateurs du métabolisme des flavonoïdes dans les baies en utilisant des mutants naturels de vigne affectés dans la synthèse des anthocyanes. Le séquençage à grande échelle des EST (Expressed Sequence Tag) de vigne a permis d'identifier plusieurs régulateurs putatifs du métabolisme des flavonoïdes. L'analyse des profils d'expression de ces gènes par RT-PCR semi-quantitative chez le Pinot noir et deux de ses mutants naturels, le Pinot gris et le Pinot blanc, a mis en évidence l'expression différentielle d'un gène codant un facteur de transcription MYB nommé *VvMybPA1*. Une analyse globale du transcriptome de la pellicule des baies de Béquignol mutant a également été réalisée et a permis l'identification de *VvMyb24*, un autre gène codant une protéine de la famille MYB. La caractérisation fonctionnelle des gènes *VvMybPA1* et *VvMyb24* a alors été entreprise chez *Arabidopsis thaliana*.

SYNTHESE BIBLIOGRAPHIQUE

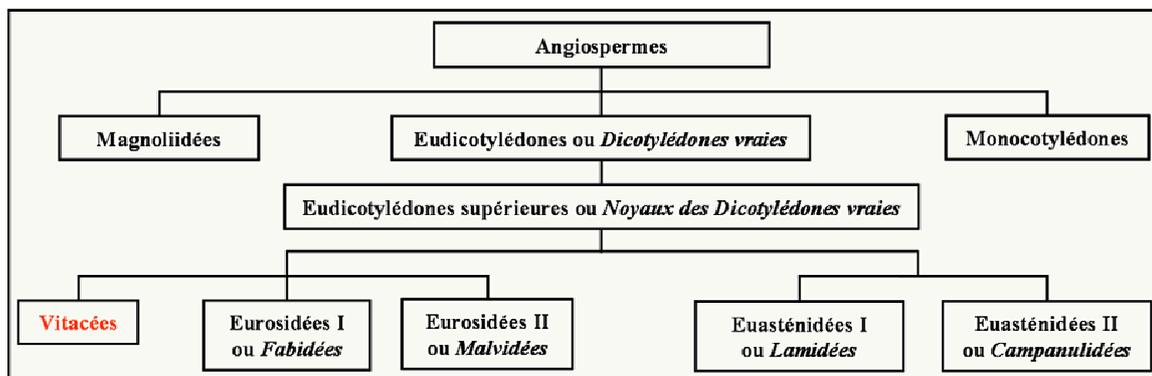


Figure 1. Classification selon Cronquist.

La classification de Cronquist est une classification des Angiospermes fondée essentiellement sur des critères morphologiques, anatomiques et chimiques [6].

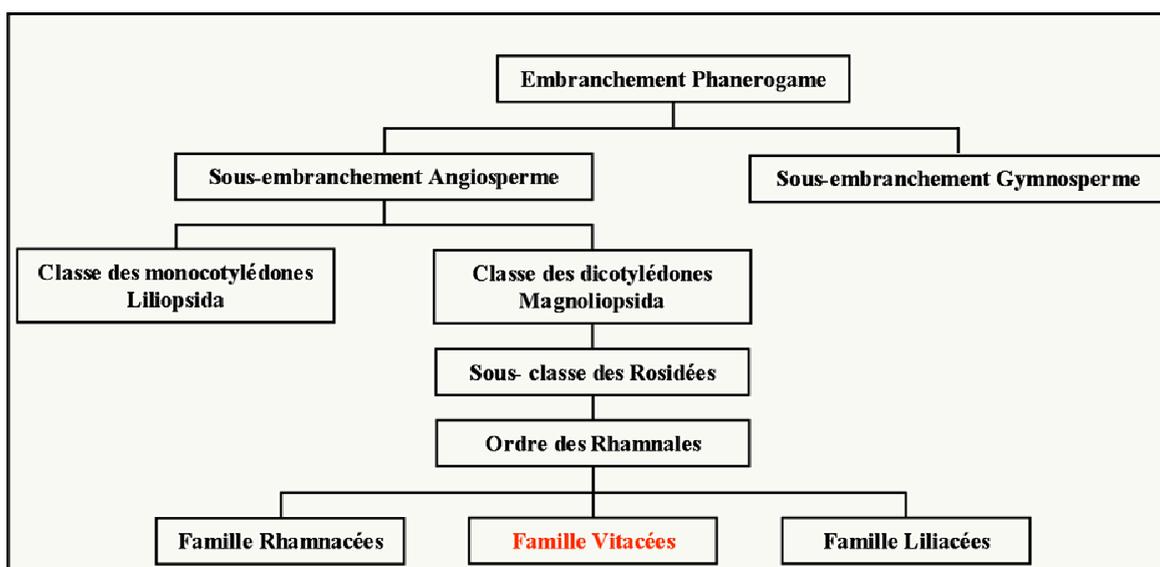


Figure 2. Classification phylogénétique des Vitacées.

Selon les travaux de l'Angiosperms Phylogeny Group (APG I en 1998 et APG II en 2003) et de Jansen et *al.* (2006), la famille des *Vitacées* est une famille basale de la famille des Rosidées [7-9].

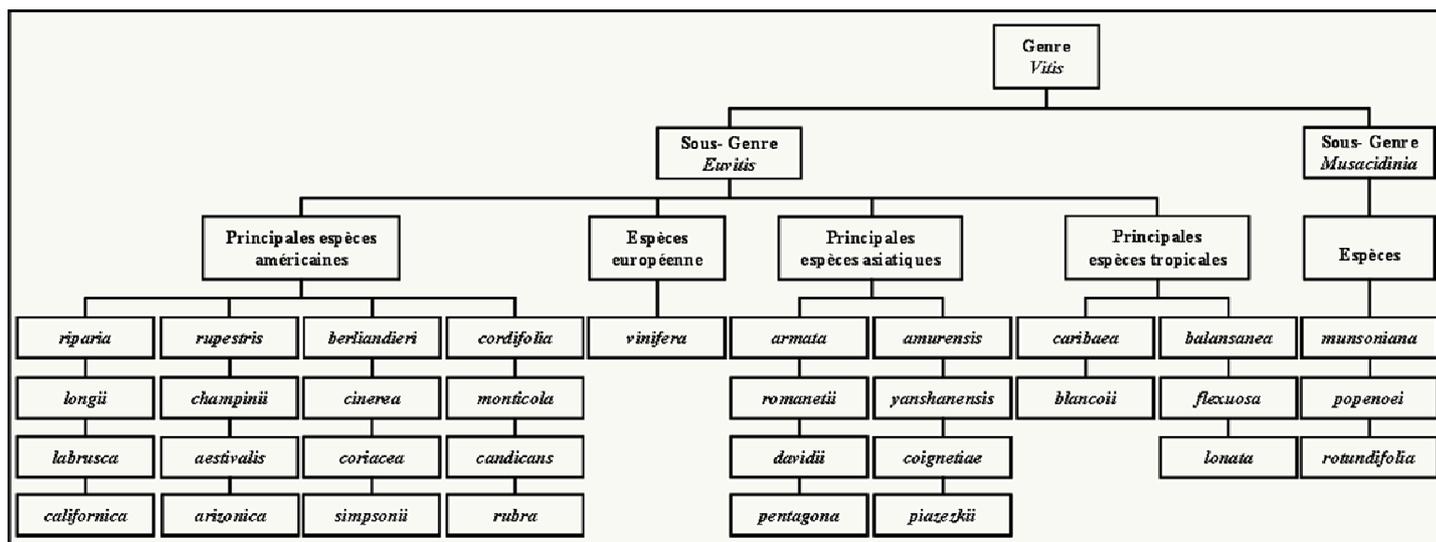


Figure 3. Classification générale du genre *Vitis* (d'après [10]).

Le genre *Vitis* se divise en deux sous-genres : *Muscadinia* et *Euvitis*. Le sous-genre *Muscadinia* comporte trois espèces originaires des Etats-Unis. Le sous-genre *Euvitis* comprend une soixantaine d'espèces divisées en quatre groupes, classés en fonction de leur origine géographique : américaine, européenne, asiatique et tropicale.

I- Présentation de la Vigne

1) Systématique

La vigne appartient à la famille des *Vitacées*, également appelée *Ampelidacées* (dans la littérature ancienne). Les plantes de cette famille sont des lianes ligneuses ou herbacées (avec une tige tubéreuse ou souterraine) qui possèdent des feuilles alternes, des vrilles et des inflorescences oppositifoliées, des étamines libres, un gynécée supère, et dont les fruits sont des baies.

Arthur Cronquist (1981) sur la base de critères morphologiques, anatomiques et chimiques, place cette dicotylédone dans la sous-classe des *Rosidées* dans l'ordre des *Rhamnales*; proche de la famille des *Rhamnacées* et des *Léécées* (figure 1) [6]. Mais aujourd'hui, les classifications phylogénétiques APG I [7], APG II [8] et les travaux de Jansen et *al.* présentent les *Vitacées* comme une famille ancestrale à celle des *Rosidées* (figure 2) [9].

La famille des *Vitacées* comprend 19 genres, dont deux sont fossiles. Seul le genre *Vitis* a une importance agricole. Il est divisé en deux sous-genres, *Euvtis* et *Muscadinia*, avec comme principale différence le nombre de chromosomes : $2n = 38$ pour le sous-genre *Euvtis* et $2n = 40$ pour le sous-genre *Muscadinia* (figure 3). Le sous-genre *Euvtis* est composé d'environ quarante espèces réparties sur les continents nord américain, européen et asiatique [11]. Les vignes d'origine américaines ont été introduites en Europe à la fin du XIX^{ème} siècle, suite à la dévastation des vignes européennes par le phylloxéra. Seul le greffage des cépages européens sur des pieds « porte-greffes » américains (*Vitis riparia*, *Vitis rupestris*) résistants au parasite a permis, dans la première moitié du XX^{ème} siècle, le sauvetage et la restauration des vignobles européens. Les vignes asiatiques (*Vitis amurensis*...) sont utilisées dans les programmes de croisements interspécifiques pour leur résistance au froid, particulièrement dans l'ex-URSS [10]. Enfin, il existe une seule vigne d'origine européenne, *Vitis vinifera*, qui permet de produire la quasi-totalité du vin consommé dans le monde. La viticulture française fait état de 500 cépages dont les représentants les plus connus se nomment Pinot, Cabernet sauvignon, Syrah, Gamay, Chardonnay....

Vitis vinifera est une espèce diploïde. Réparti sur 19 chromosomes, son génome est de petite taille, 475-500 Mb (approximativement 4 fois celui d'*Arabidopsis* mais 1/6^{ème} de celui du maïs) et 30 434 gènes. Le séquençage de la vigne a révélé que 41,4% de son génome était constituée de séquences répétées et de transposons [12]. Elle présente un polymorphisme remarquable en relation avec son caractère fortement hétérozygote (75%) [13]. Sous sa forme originale et sauvage, la vigne est dioïque. Cependant, la plupart des espèces modernes sont hermaphrodites [11].

2) Notion de variétés, cépages, clones, cultivars

La vigne cultivée *Vitis vinifera* comprend plus de 6 à 7000 variétés.

Un **clone**, ou **cultivar**, peut être défini comme le descendant par voie végétative d'une souche mère. Le **cépage**, unité taxonomique propre à *Vitis vinifera*, est le produit d'un semis ou d'un individu unique au départ, multiplié par voie végétative. Au cours des cycles de multiplication, des variations peuvent se produire et être fixées. Un cépage est donc composé d'un ensemble de clones suffisamment semblables entre eux pour être confondus sous un même nom, on parle de cépage-population [10]. Lorsque la variation touche un caractère évident et remarquable (apparition d'une forte densité de poils couchés) ou ayant des conséquences technologiques importantes (couleur de l'épiderme de la baie, couleur de la pulpe, particularité de la saveur), le clone concerné est alors considéré comme une nouvelle variété différenciée du cépage initial.

3) Cycle végétatif de la vigne

Dans les pays tempérés comme la France, le cycle végétatif de la vigne est caractérisé par des stades phénologiques bien décrits par Baggiolini [14] (figure 4). Ces derniers correspondent à la croissance des organes végétatifs (rameaux, feuilles, vrilles et racines).

Lorsque les températures diminuent, la vigne entre dans une période de **repos hivernal**. De novembre à février, la vigne est en dormance avec une activité interne et biochimique relativement importante. Elle stocke alors des réserves principalement sous forme d'amidon. Au début du printemps, lorsque la température du sol atteint 10 à 12°C, l'activité végétative débute. Elle se manifeste par des " pleurs ", correspondant à des remontées d'une forme de sève brute au niveau des plaies de taille [10]. Il n'est pas rare de constater que les pleurs durent près d'un mois, temps nécessaire à la plante pour cicatriser ses plaies de taille, ce qui peut occasionner l'humidification de jeunes bourgeons et ainsi accroître leur sensibilité au gel. L'apparition de ces pleurs précède l'étape de **débourrement**. La date de débourrement est fonction des cépages, de la température, de la latitude mais également de la vigueur du sarment, et du système de taille utilisé. Cette étape est caractérisée par une reprise d'activité endogène du bourgeon latent, il s'agit essentiellement de mitoses. Lorsque la somme des températures cumulées est suffisante, il y a croissance et développement du futur rameau (stades A à D de Baggiolini). Le bourgeon gonfle, écarte les écailles, laissant apparaître la villosité (bourre ou coton qui entoure et protège les organes primordiaux) et la pointe de la première feuille [10]. Le cycle végétatif se poursuit par une période de croissance,

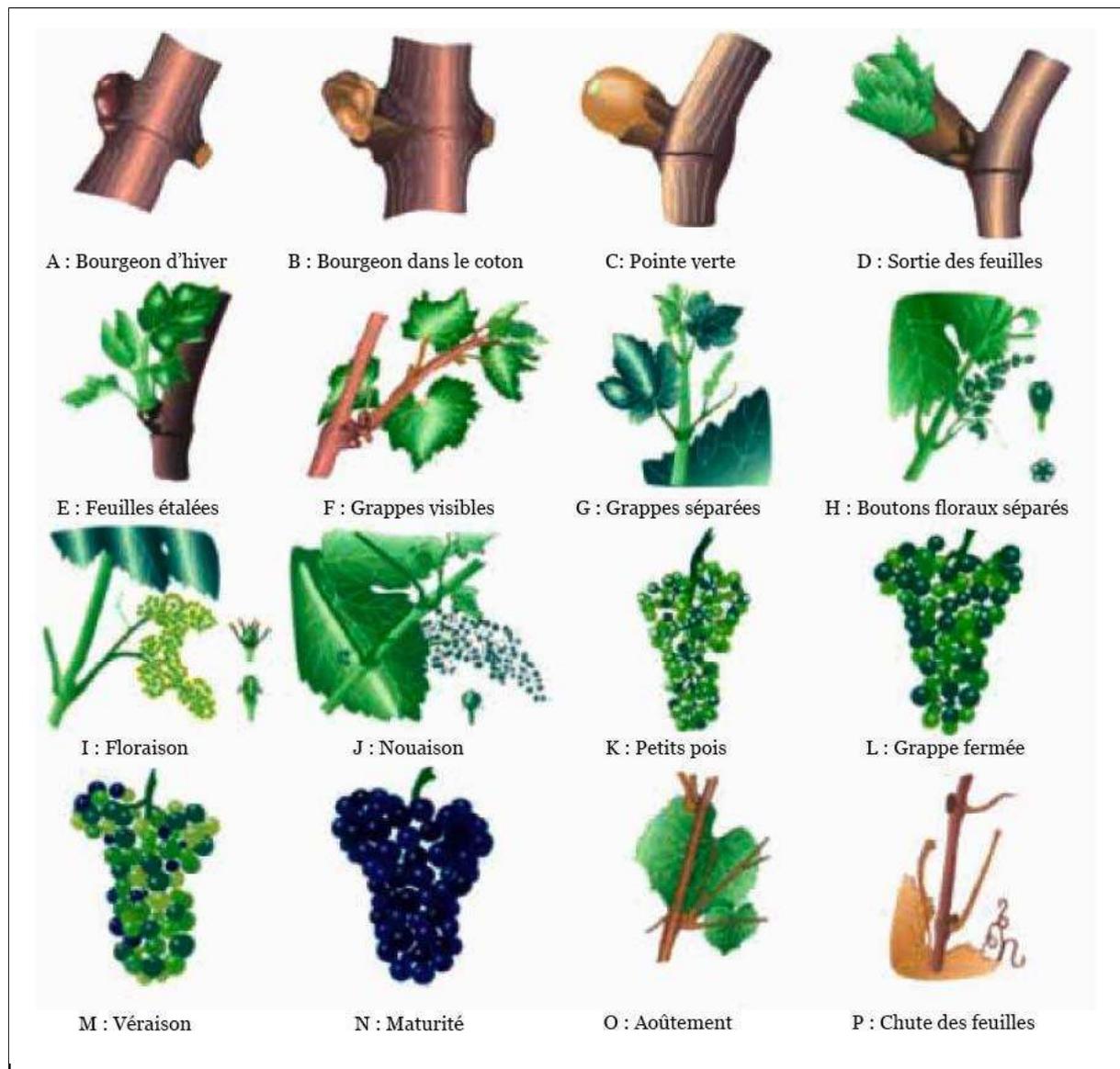


Figure 4. Stades repères phénologiques du cycle végétatif de la vigne [14].

- (A) Le repos hivernal.
- (B, C, D) Le débourrement.
- (E, F, G, H) La croissance.
- (I) La floraison.
- (J) La nouaison.
- (K et L) Fin de la phase herbacée.
- (M) La véraison.
- (N) La maturation suivie de l'aoûtement (O).
- (P) Fin du cycle de développement de la vigne.

caractérisée par l'allongement des rameaux issus des bourgeons latents, l'étalement et l'accroissement des jeunes feuilles puis la naissance de nouvelles feuilles.

Après le débourrement, les inflorescences apparaissent rapidement au sommet des pousses entre les premières feuilles. Au bout d'une à deux semaines, selon les conditions climatiques et la vitesse de croissance, les boutons floraux qui se présentent tout d'abord en masses compactes, se séparent et l'ensemble acquiert sa forme définitive. Les stades grappes visibles et grappes séparées constituent respectivement les stades F et G. La **floraison**, qui a lieu au mois de mai/juin, est l'une des phases les plus critiques du cycle végétatif de la vigne puisqu'elle va conditionner la récolte (stade I). En cas de conditions climatiques trop fraîches et/ou pluvieuses lors des périodes de floraison et d'après floraison, la fécondation ne sera que partielle, on parle alors de *coulure physiologique*. Certaines années, la fécondation est imparfaite et les ovaires mal fécondés produisent des baies de tailles réduites qui n'atteindront jamais la maturité : c'est le *millerandage* [10]. Après la fécondation, les ovules évoluent en graines (ou pépins) tandis que le reste de l'ovaire donne le fruit. Cette étape correspond à la nouaison [15] (stade J). Seuls les ovaires fécondés vont se développer en une baie au stade "petit pois" (stade K). Le feuillage, les rameaux et les racines continuent à s'étendre pendant toute la phase herbacée du cycle de développement des baies qui dure jusqu'à la fin juillet. Les baies sont vertes et se comportent à ce stade comme des organes chlorophylliens en croissance. Au mois d'août, la baie atteint le stade "véraison" qui marque la fin de la phase herbacée et le début de la maturation (stades L à N). A partir de la véraison débutent de nombreux changements physiologiques qui se poursuivent tout au long de la maturation (environ 45 j) jusqu'à ce que les baies soient mures et puissent être vendangées (fin septembre-début octobre). Après la phase de maturation, le raisin entre dans une phase de surmaturation, caractérisée par un flétrissement des baies, une augmentation de la concentration en sucres et une sensibilité accrue aux attaques fongiques et bactériennes.

Parallèlement à la véraison survient l'**aoûtement** (stade O). Il se caractérise par un brunissement de l'écorce des rameaux, des vrilles et des grappes. Ce processus résulte de modifications anatomiques (formation de liège), de l'accumulation de lignines et de réserves amylacées, et d'une diminution synchrone de la teneur en eau des tissus du bois. Ces réserves proviennent des feuilles qui, après l'arrêt de la croissance, se sont progressivement vidées de leur contenu. L'aoûtement se poursuit jusqu'en novembre, avant les premières gelées, et prépare ainsi le développement de la vigne pour l'année suivante [10]. Dans le courant du mois d'octobre et du mois de novembre, la vigne commence à perdre ses feuilles : c'est la **défeuillaison**, signe de la fin du cycle végétatif actif de la vigne (stade F).

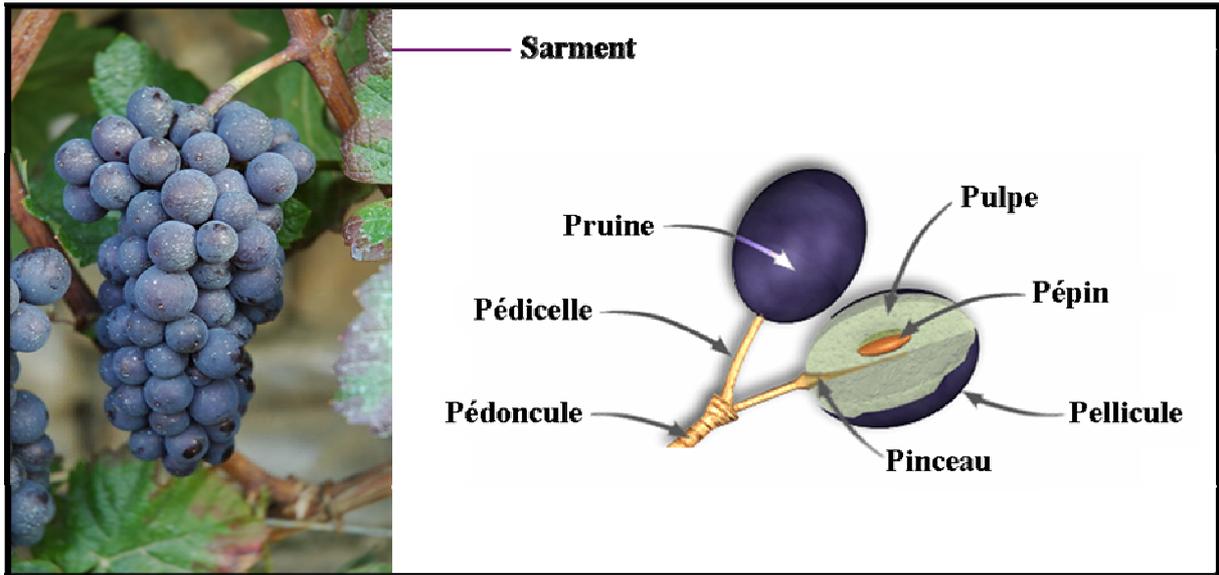


Figure 5 : Organisation d'une baie de raisin

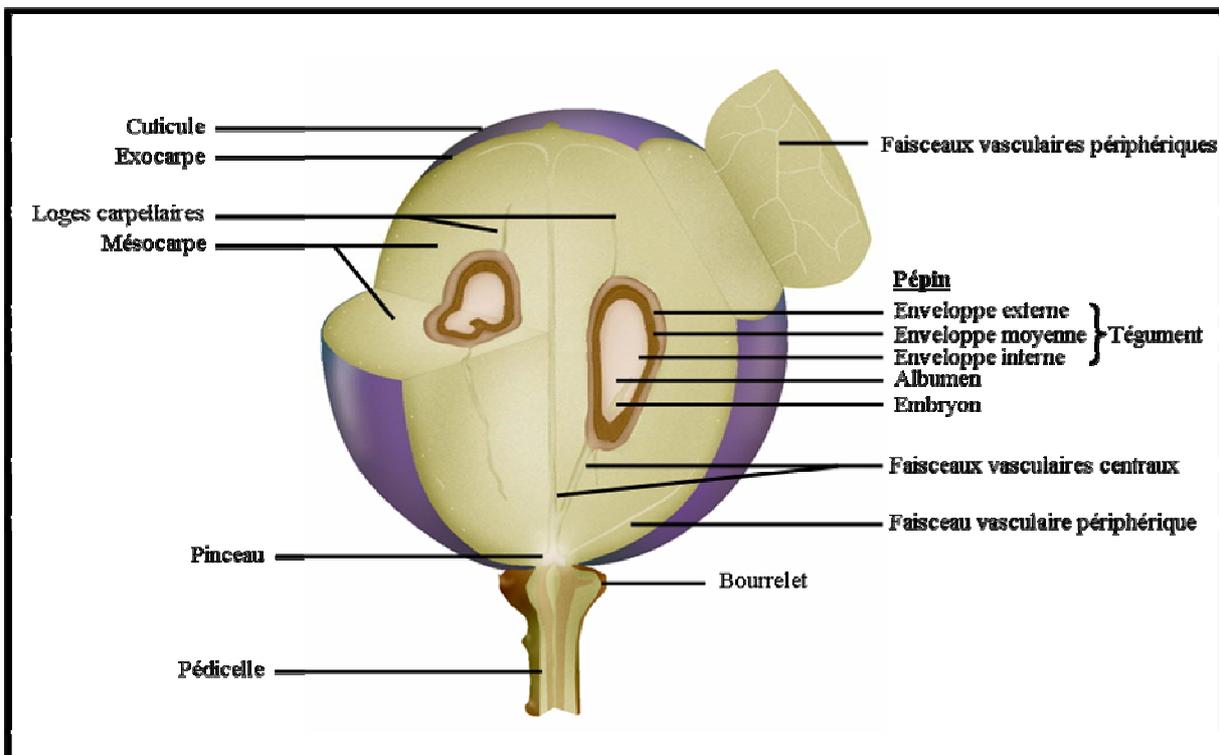


Figure 6 : Structure détaillée d'une baie de raisin (d'après [16, 17]).

1) Morphologie et anatomie de la baie de raisin

Le raisin est une baie classée dans le groupe des fruits charnus à pépins. Ce fruit est regroupé en grappe (figure 5). Cette dernière est composée d'un pédoncule qui la fixe au sarment, d'un rachis ou rafle, partie ligneuse plus ou moins ramifiée dont les ultimes ramifications (les pédicelles) portent les baies. La rafle est essentiellement constituée d'eau, de fibres, de tannins et de matières minérales. Le pinceau constitue le prolongement du pédicelle à l'intérieur de la baie.

La structure et l'ultrastructure de la baie sont liées à sa fonction de puits. Une coupe dans une baie permet de distinguer de l'extérieur vers l'intérieur (figure 6) [16]:

- une couche externe ou exocarpe correspondant à la pellicule ou peau du raisin,
- une couche moyenne ou mésocarpe formant la chair ou la pulpe,
- une couche interne ou endocarpe, réduite à une simple paroi contre le tégument externe de la graine,
- des faisceaux libéro-ligneux nécessaires à l'alimentation de la baie en eau et assimilats,
- les graines ou pépins

a) La pellicule

Les cellules de la pellicule sont caractérisées par la présence de plastides, de vacuoles contenant des composés phénoliques et de nombreuses mitochondries bien développées, révélant une intense activité métabolique.

Elle est constituée de tissus à forte concentration cellulaire, on distingue (figure 6):

- la cuticule, membrane extérieure très mince recouverte d'une matière cireuse appelée pruine. Elle donne un aspect velouté au raisin. Elle assure l'imperméabilité de la pellicule et retient les levures amenées par le vent et les insectes.
- l'épiderme, formé d'une seule assise de cellules régulières sous la cuticule.
- l'hypoderme, tissu constitué de couches de cellules renfermant des granulations de matières colorantes et odorantes, responsables respectivement de la couleur et du fruité du raisin. Généralement ces substances ne se trouvent que dans la pellicule, à l'exception des cépages teinturiers dont la pulpe est colorée et des cépages muscats dont les arômes se trouvent également dans la pulpe [10].

b) La pulpe

La pulpe représente 75 à 85% du poids de la baie. Les cellules de la pulpe contiennent une très grande vacuole, essentielle à leur rôle de stockage. Le contenu vacuolaire fournira le

moût du raisin qui compose la quasi-totalité du poids de la pulpe. Cette dernière est constituée d'eau (70 à 80%), de sucres (100 à 300 g/L), d'acides organiques, de sels minéraux, de substances azotées, de composés aromatiques et phénoliques.

Elle comprend (figure 6):

- une zone externe, peu épaisse, qui tapisse la face interne de la pellicule
- une zone intermédiaire qui est la plus volumineuse. Lorsque la baie est mûre, ces cellules subissent une désorganisation de leur paroi. Lors de la cueillette ou des traitements mécaniques de la vendange, ce sont elles qui libèrent leur jus en premier, et donnent ainsi les meilleurs jus
- une zone interne qui abrite les pépins

c) **Les pépins**

La graine ou pépin résulte du développement de l'ovule fécondé. Le nombre de pépins par baie peut varier de 0 à 4, la moyenne se situant le plus souvent vers 2. Dans certains cas, les raisins n'ont pas du tout de pépins et sont dits apyrènes.

Les pépins sont formés d'une cuticule fine et sensiblement colorée, d'un épiderme et de téguments qui entourent l'albumen et l'embryon [18]. Une coupe longitudinale révèle l'existence de trois types de téguments ou enveloppes (figure 6):

- une enveloppe externe essentiellement composée de cellules parenchymateuses à parois cellulósiques assez molles
- une enveloppe moyenne dure comprenant deux assises de cellules allongées fortement lignifiées
- une enveloppe interne formées par 3 couches de cellules rectangulaires

Les tannins représentent 5 à 8% du poids du pépins et sont synthétisés dans l'épiderme et les cellules du tégument interne [18].

d) **Les faisceaux libéro-ligneux**

Ils partent du bourrelet du pédicelle et assurent la nutrition de la baie. On distingue les faisceaux périphériques, situés à une vingtaine d'assises cellulaires au-dessous de l'épiderme, et un faisceau vasculaire central plus large qui alimente les pépins et la columelle (figure 6). Au début du développement, les faisceaux vasculaires périphériques délimitent l'exocarpe et le mésocarpe. Plus tard, ils sont entourés de cellules de mésocarpe et ne caractérisent plus la frontière exocarpe-mésocarpe [20].

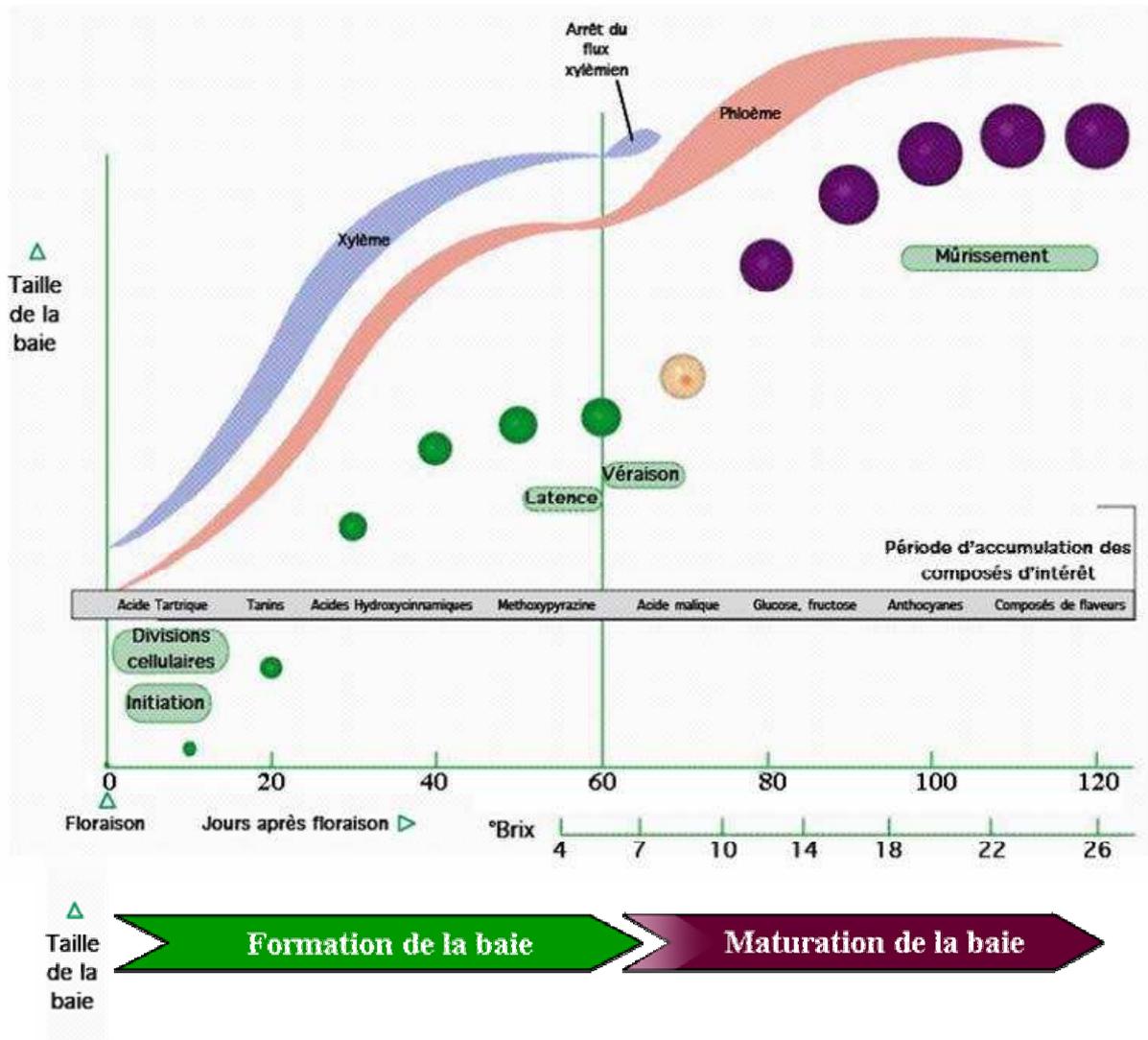


Figure 7 : Schéma représentant le développement de la baie de raisin (d'après [19]).

La taille et la couleur relative des baies sont représentées tous les 10 j à partir de la floraison. Les évolutions des flux xylémien et phloémien sont représentées respectivement par les aires bleues et roses. L'indice Brix, qui correspond à une évaluation de l'alcool que pourrait produire les baies au cours de leur maturation est fonction de la concentration en sucres dans la baie de raisin. Les différents composés qui s'accumulent au cours du développement de la baie sont indiqués.

2) Physiologie de la baie de raisin

a. Cycle de développement de la baie de raisin

La baie de raisin croît selon une double sigmoïde, initialement découpée en trois phases de développement: **phase I** (première période de croissance active), **phase II** (phase de latence ou de ralentissement de la croissance) et **phase III** (seconde période de croissance active) [21-25]. Récemment, les travaux d'Ollat et *al.* ont permis de confirmer les descriptions antérieures faites par Coombe et Staudt, selon laquelle le développement des baies se fait en deux phases seulement : la croissance herbacée et la maturation [26-28] (figure 7). La phase de latence ne correspond pas à un stade physiologique mais seulement à la fin de la première période de croissance. Ainsi, nous décrivons le développement de la baie de raisin selon ces deux phases.

Phase I ou croissance herbacée

La croissance herbacée est déterminée par le nombre de pépins, l'action des régulateurs de croissance, les relations source-puits et les paramètres climatiques [24]. Elle débute à la floraison et dure de 25 à 45 jours (j) suivant les cépages [15]. Pendant cette période, la croissance des baies est liée simultanément à des phénomènes de division et de grandissement cellulaire. La plupart des divisions ont lieu avant l'anthèse, mais le pic de mitoses se situe environ une semaine après celle-ci. Les dernières divisions s'effectuent dans la partie périphérique de la baie une quarantaine de jours après anthèse [28]. Le grandissement cellulaire commence au bout de 10 à 15 j après la fécondation et se poursuit jusqu'à la fin de la phase I [18]. A la fin de cette période, le nombre de cellules dans la baie est établi et il déterminera la taille potentielle de la baie à maturité [19].

Cette phase de croissance rapide aboutit à la formation de la baie ainsi qu'à la production des pépins [18]. La baie est verte, ferme, acide et amère; elle présente une respiration élevée. L'eau est principalement importée par le xylème [24]. Le saccharose, importé par la sève phloémienne, reste peu concentré car il est métabolisé pour assurer le fonctionnement cellulaire. Les acides organiques (majoritairement acides malique et tartrique), les acides hydroxycinnamiques et les tannins sont synthétisés sur place [29]. Des minéraux, des acides aminés (aa), des micronutriments et des composés aromatiques (comme les méthoxypyrazines) s'accumulent également [29-31].

La fin de la phase I est caractérisée par un ralentissement de la croissance qui peut durer de 10 à 20 j selon les cépages. Les baies perdent leur chlorophylle, commencent à devenir translucides et à se pigmenter : c'est la véraison.

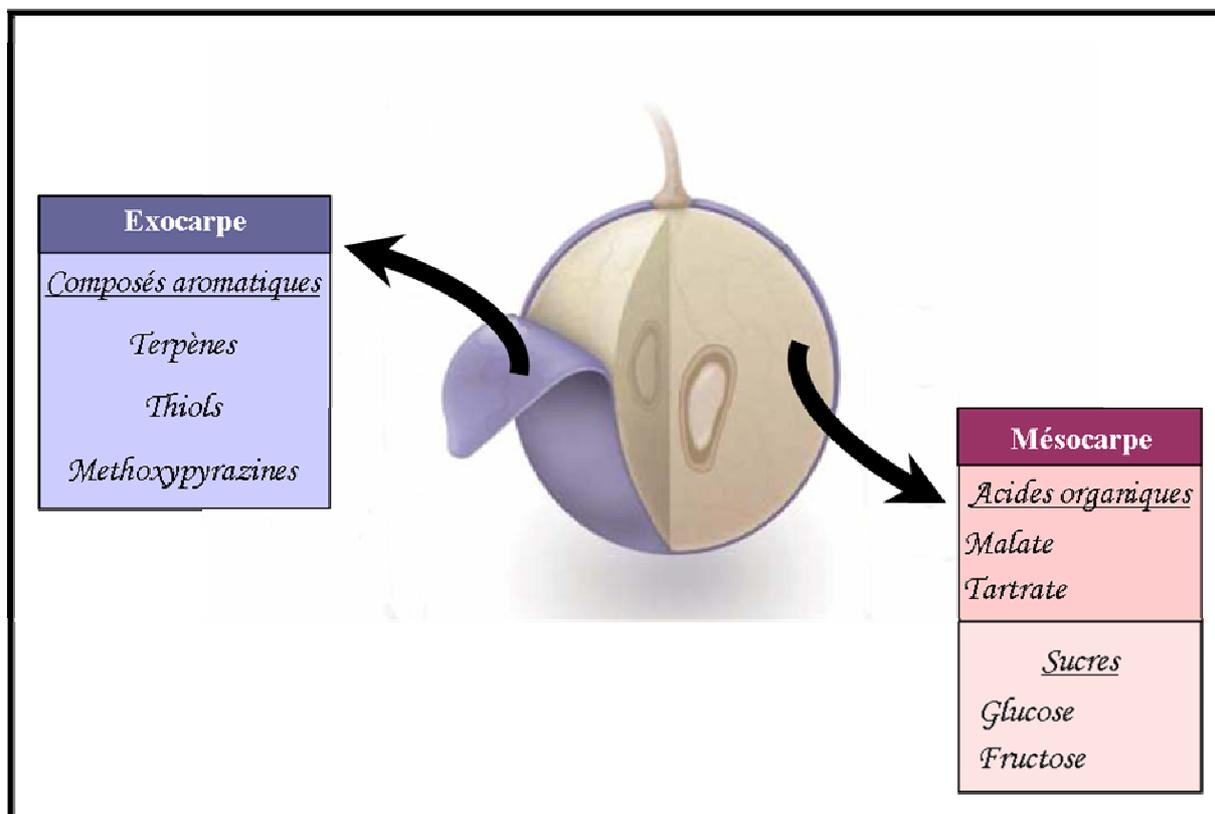


Figure 8 : Répartition tissulaire des composés essentiels à l'élaboration de la qualité organoleptique de la baie de raisin (d'après [34]).

Phase II ou maturation

Le début de la véraison coïncide avec le début de la deuxième période de croissance. Le grandissement cellulaire reprend, entraînant une augmentation du volume des baies. A la fin de la maturation, la taille des baies est 2 à 3 fois supérieure à celle observée pendant la phase de latence [19, 32].

De nombreux changements physiologiques et métaboliques ont lieu pendant cette période qui dure de 35 à 45 jours. La baie se ramollit et sa composition chimique change. La modification de la composition en polysaccharides pariétaux et la distension cellulaire provoquée par l'accumulation massive d'eau et de sucres sont à l'origine de cette plasticité pariétale [17, 33]. La baie devient un véritable organe puits. Les importations de carbone augmentent fortement (3,5 fois en moyenne) [24]. Les cellules de la pellicule sont caractérisées par la présence de plastes, de vacuoles accumulant des composés phénoliques et de nombreuses mitochondries bien développées, ce qui révèle une intense activité métabolique. De façon concomitante à l'accumulation des sucres dans les cellules de la pulpe, l'acidité des baies diminue et le pH augmente [24]. L'acide malique est métabolisé et utilisé comme source d'énergie. Les tannins sont oxydés, remobilisés et les méthoxypyrazines sont dégradés par la lumière [19].

a. Aspects biochimiques du développement de la baie de raisin

La qualité organoleptique des baies est la résultante du contenu en acides, en sucres et en composés phénoliques et aromatiques (figure 8). Ce chapitre résume les avancées réalisées sur la compréhension des mécanismes qui déterminent l'acidité de la baie, l'accumulation des sucres et des arômes. L'accumulation des composés phénoliques et les mécanismes régulateurs de leur biosynthèse seront détaillés dans le chapitre suivant car ce sujet constitue le cœur du travail expérimental présenté dans cette thèse.

Acidité de la baie

Les principaux acides organiques du raisin sont les acides malique et tartrique qui représentent environ 90% de l'acidité totale des baies (figure 7 et 8). Des traces d'acide citrique et d'autres acides organiques plus rares sont également détectés [24]. Pendant la croissance herbacée, les acides malique, tartrique et l'acidité totale augmentent simultanément dans la baie. Leur concentration est maximale à la véraison, et diminue ensuite pendant la maturation [35].

Peu de temps après la floraison, l'acide tartrique est rapidement synthétisé dans les cellules du mésocarpe et de l'exocarpe, à partir du métabolisme de l'acide ascorbique [24]. Après la véraison, l'enzyme clé de la biosynthèse du tartrate, L-IdnDH (L-idonate déhydrogénase) n'est plus exprimée et l'acide tartrique n'est plus accumulé [36, 37]. De fait, pendant la maturation, la concentration en acide tartrique diminue de concert à l'augmentation du volume de la baie. Cette diminution est attribuée à un effet de dilution, puisque le volume des baies augmente alors que la quantité du tartrate par baies reste constante [35].

Contrairement à l'acide tartrique, le malate est un intermédiaire très actif du métabolisme du raisin car il est la source principale en carbone pour la respiration. Sa concentration et sa répartition tissulaire varient énormément au cours du développement et de la maturation des baies. La concentration en acide malique, accumulé dans les cellules de la pulpe à la fin de la première phase de croissance, est maximale à la véraison, puis chute rapidement [25]. En fin de maturation, l'acide malique est majoritairement présent dans les cellules de la pellicule et sa concentration est deux à trois fois inférieure à celle de départ dans la pulpe [35, 38]. Dans les baies non vérées, l'acide malique est synthétisé dans le cytosol à partir de la β -carboxylation de la PEP par l'enzyme PEPC. L'oxaloacétate formé est ensuite réduit par la malate déshydrogénase cytosolique pour donner le malate [24]. A la véraison, l'acide malique est métabolisé par deux enzymes: l'enzyme malique cytosolique et la PEP carboxykinase, permettant ainsi de fournir l'énergie nécessaire à la biosynthèse, la respiration et la néoglucogenèse [35, 39-41]. Une autre voie de dégradation résulterait de la diffusion passive du malate à travers la membrane vacuolaire (malgré l'augmentation d'activité des deux pompes H^+ -PPiase and V-ATPase) puis dans la mitochondrie. La malate déshydrogénase et l'enzyme malique mitochondriale dégraderaient alors le malate respectivement en oxaloacétate et en pyruvate [25, 42]. Les facteurs environnementaux affectent également la teneur en acide malique. L'acidité des fruits à la récolte est corrélée négativement à la température subie pendant la période de maturation. Elle varie selon les régions et les années, avec des niveaux plus élevés en acide malique dans les régions froides [35].

Accumulation des sucres

Produit de la photosynthèse, le saccharose est transporté depuis les feuilles (organe source) vers les baies (organe puits) par le phloème. Le chargement du phloème peut s'effectuer par deux voies : apoplastique ou symplastique. Dans la voie symplastique, le saccharose transite depuis les cellules du mésophylle vers les cellules du complexe cellule compagne – cellule criblée *via* les plasmodesmes, qui assurent la continuité du cytoplasme de ces cellules. Dans la voie apoplastique, le saccharose est libéré des cellules du mésophylle dans l'apoplaste et

ensuite dirigé vers les cellules conductrices du phloème (cellules compagnes ou cellules criblées) grâce à des transporteurs membranaires actifs. Ainsi, un gradient de saccharose est généré entre les organes sources et les organes puits, dirigeant le flux phloémien vers ces derniers.

Dans la baie, le déchargement du saccharose depuis le complexe conducteur jusqu'aux cellules receveuses met en jeu les deux voies décrites ci-dessous. Par la voie symplastique, le saccharose traverse les plasmodesmes et est hydrolysé dans le cytosol ou la vacuole des cellules puits. Par la voie apoplastique, le saccharose est exporté dans l'apoplaste, et il est réabsorbé à travers la membrane plasmique de la cellule receveuse. Des transporteurs spécifiques assurant l'absorption ou l'efflux du saccharose ou des hexoses permettent le franchissement de la membrane plasmique et de la membrane vacuolaire [43]. Pendant la croissance herbacée, le déchargement symplastique est favorisé puis, à partir de la véraison, le déchargement apoplastique prend le dessus [44]. Trois transporteurs de saccharose ont été caractérisés dans la baie de raisin: *VvSUC11*, *VvSUC12* et *VvSUC27* [45-47]. En 1999, l'identification de transporteurs de monosaccharide par Fillion et *al.* suggère que l'importation du sucre dans les cellules receveuses peut se faire aussi sous forme d'hexoses. Le saccharose serait clivé dans l'apoplaste, maintenant ainsi le gradient de sucrose entre le phloème et l'apoplaste. Les hexoses produits seraient alors absorbés dans le cytosol grâce à des transporteurs assurant un symport proton/saccharose à travers la membrane plasmique puis par des transporteurs assurant un antiport proton/hexose dans la vacuole [43]. Six séquences d'ADNc pleine longueur homologues à des transporteurs membranaires d'hexoses ont été clonées à partir de banques d'ADNc de baies, et appelées *VvHT1* à *VvHT6* [25, 48, 49].

Pendant la croissance herbacée, la majorité du sucrose importé dans la baie est métabolisé, et de fait, la concentration en sucres est relativement faible (pas plus de 150 mM d'hexoses) (figure 7 et 8). A partir de la véraison, le saccharose déchargé du phloème, est stocké sous forme d'hexoses (fructose et glucose) dans la vacuole des cellules du mésocarpe. Vingt jours après la véraison, la concentration vacuolaire en hexoses est proche de 1 M, avec un rapport glucose sur fructose de 1 [48]. L'accumulation du fructose et du glucose en proportion égale suggère que la liaison *O*-glycosidique du saccharose est clivée par des invertases plutôt que par la saccharose synthase [24, 48, 50]. Dans la baie de raisin, différentes isoformes d'invertases ont été localisées dans la paroi, le cytoplasme et la vacuole [50]. Deux invertases vacuolaires, *GIN1* et *GIN2*, ont été clonées et caractérisées. Elles sont fortement exprimées dans les étapes précoces du développement de la baie, puis leur expression et leur activité diminuent simultanément à l'accumulation des hexoses [37, 51].

Le déchargement des sucres dans la baie s'arrête à la fin de la maturation, lorsque la quantité en sucres devient constante.

Accumulation des composés aromatiques

La maturité aromatique correspond à une complexité maximale du bouquet odorant se dégageant des raisins, et plus particulièrement de la pellicule et de la pulpe. Il existe plusieurs centaines de composés volatils qui ont un impact olfactif important et qui participent aux arômes. Les caractéristiques de l'arôme variétal, encore appelé arôme primaire, sont attribuées à la présence de certains constituants, parmi lesquels les acides organiques, les proanthocyanidines, les composés terpéniques, les méthoxypyrazines, les dérivés shikimiques et les composés à fonction thiol [34] (figure 8). Ces substances d'origine variétale sont spécifiques du cépage et sont présentes dans le fruit à l'état libre ou sous forme de précurseurs glycosylés [52].

La famille des terpènes a été particulièrement étudiée ces dernières années. Une cinquantaine de composés terpéniques ont été identifiés dans le raisin [25]. Il s'agit de monoterpènes, de quelques sesquiterpènes et des alcools et aldéhydes correspondants. Les monoterpènes les plus odorants se trouvent parmi les alcools monoterpéniques, en particulier le linalol, le géraniol, le nérol et l' α -terpinéol qui développent des odeurs de type floral. Les monoterpènes libres et liés, issus de la voie de synthèse DOXP/MEP, n'ont pas tout à fait la même localisation dans la baie [52]. La pellicule est plus riche en composés terpéniques que la pulpe. Leurs taux augmentent au cours du développement et de la maturation de la baie et même parfois au-delà du stade mûr. Cependant, les terpénols libres n'apparaissent en quantité notable qu'à partir de la véraison. Les tétraterpènes sont à l'origine des caroténoïdes, dont la dégradation oxydative conduit à des composés comme les norisoprénoïdes qui peuvent être odorants. L'ensoleillement des baies pendant la maturation augmente la dégradation des caroténoïdes [25].

En revanche, d'autres composés comme les méthoxypyrazines et les proanthocyanidines sont considérés comme gênants chez certains vins tandis qu'ils donneront une fragrance subtile tant appréciée chez d'autres. Les proanthocyanidines (tannins) sont à l'origine de l'astringence et de l'amertume des vins. Les méthoxypyrazines, produits du métabolisme des acides aminés (aa), sont accumulées sous forme libre pendant la croissance herbacée et sont ensuite progressivement métabolisées au cours de la maturation. Elles donnent le caractère végétal avec des notes de poivrons et de fruits rouges aux vins de Cabernet sauvignon. Leur

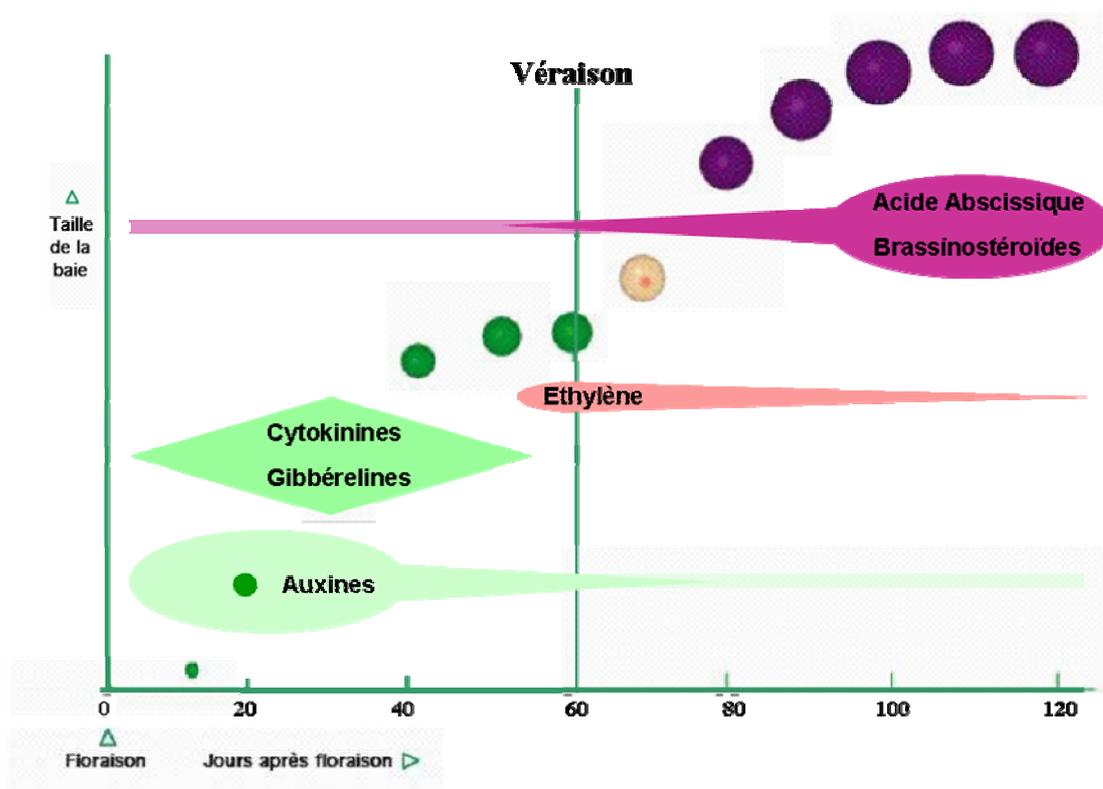


Figure 9 : Régulation hormonal du développement de la baie de raisin (d'après [25]).

La taille et la couleur relative des baies sont représentées tous les 10 j à partir de la floraison. Les variations des teneurs en hormone sont indiquées par des volumes arbitraires. Les deux phases de développement de la baie sont indiquées : la croissance herbacée de 0 à 60 j après floraison et la maturation qui suit la véraison.

dégradation dépend du microclimat environnant et augmente de façon synchrone à l'exposition des baies au soleil [34].

b. Contrôle hormonal du développement de la baie.

Les raisins sont considérés comme des fruits non-climactériques. A l'inverse des fruits climactériques, la maturation de la baie de raisin ne s'accompagne pas d'une augmentation de la respiration associée à une brusque stimulation de la synthèse d'éthylène. Le taux respiratoire évolue peu après la véraison et la synthèse d'éthylène reste faible [21]. La véraison est marquée par un changement significatif de la transcription des gènes (certains seront réprimés tandis que d'autres seront activés). Ces modifications sont synchrones aux changements des taux d'hormones et/ou de leur perception dans la baie [21, 23, 53, 54]. Ces hormones ou phytohormones agissent à de faibles concentrations molaires, fonctionnent souvent de façon concomitante et répondent à divers stimuli environnementaux. Leur action se traduit par de multiples effets permettant ainsi de réguler le développement de la baie, de la fécondation à la maturation (figure 9).

Hormones intervenant dans la mise en place de la baie de raisin

La phase de croissance herbacée de la baie de raisin est majoritairement sous contrôle des auxines et des gibbérellines (GA). Ces phytohormones peuvent être importées dans le fruit, mais elles sont majoritairement produites par les pépins, ou par les tissus maternels (ovules non fécondés) pour les espèces apyrènes [25].

Jusqu'alors, plusieurs travaux suggéraient que l'auxine était impliquée dans le développement précoce de la baie, mais aussi dans l'initiation de la maturation. D'une part, l'application d'IAA ou d'une auxine de synthèse exogène (BTOA) quelques semaines avant la véraison peut inhiber la maturation des baies [53]. D'autre part, chez une espèce américaine (Concord cv *Vitis Labruscana*), la teneur en IAA est maximale avant le plateau herbacé puis décroît rapidement à la véraison, pour atteindre de faibles quantités à la maturation [55]. L'ensemble de ces données suggère que la diminution de la teneur en auxine déclenche la maturation du fruit. Récemment, Symons et *al.* ont constaté que les teneurs en IAA restent faibles durant tout le développement des baies de Cabernet Sauvignon [56]. Ainsi l'auxine jouerait seulement un rôle dans la croissance de la baie et interviendrait dans la régulation de la division et la différenciation cellulaire.

Les gibbérellines sont impliquées dans le grandissement cellulaire. Dans la baie, la teneur en GA est proportionnelle aux nombres de pépins. Dans des baies épépinées, la concentration

endogène en GA est élevée jusqu'à deux semaines après floraison, puis elle atteint des niveaux très bas voire indétectables pendant la maturation [56].

Hormones intervenant au cours de la maturation de la baie de raisin

Le signal hormonal contrôlant la maturation, de la véraison à la récolte, est encore mal connu. Il résulterait de la combinaison de plusieurs signaux hormonaux plutôt que d'un seul. Trois hormones semblent impliquées dans le processus de maturation : l'acide abscissique (ABA), l'éthylène et les brassinostéroïdes (BR).

Dans les fruits climactériques, l'éthylène provoque une brusque augmentation de la respiration appelée crise respiratoire. La baie de raisin est définie comme un fruit non-climactérique car la teneur en éthylène reste faible durant tout le développement [21]. Néanmoins, une augmentation transitoire de la production d'éthylène (2 à 3 fois environ) trois semaines avant la véraison a été constatée dans des baies de Cabernet sauvignon [57]. Cette augmentation faible suggère que l'éthylène pourrait jouer un rôle majeur dans la maturation des baies de raisin. En effet, les travaux de Jeong et *al.*, montrant que l'application d'un inhibiteur des récepteurs à l'éthylène (le 1-MCP) retardait l'augmentation du diamètre de la baie, confortent cette hypothèse [57]. Cette hormone pourrait être également responsable des changements physiologiques qui ont lieu dans la baie au cours de la maturation. Ainsi, l'éthylène serait impliqué dans l'accumulation des sucres et la diminution de la teneur en acides après véraison [57, 58]. Par ailleurs, l'application d'éthylène exogène active, à long-terme, les gènes impliqués dans la synthèse des anthocyanes et induit donc une accumulation plus importante de ces métabolites [59].

L'ABA semble déterminant dans l'initiation de la maturation [21, 53, 55]. De manière similaire à l'éthylène dans les fruits climactériques, il est fortement accumulé dès la véraison et durant toute la maturation [21]. Des traitements pouvant retarder l'accumulation d'ABA entraînent systématiquement un retard de la maturation et inversement, l'application d'ABA exogène (6 à 8 semaines après floraison) réduit la phase de croissance herbacée [21, 53]. Des baies traitées avec de l'ABA à la véraison, mûrissent plus vite, sont moins acides et accumulent plus d'anthocyanes [60-62]. Un retard de l'accumulation d'ABA est accompagné par un retard de l'importation en hexoses, suggérant que l'ABA favoriserait également l'accumulation des sucres dans la baie [53, 62, 63].

Des travaux récents publiés par Symons et *al.* (2006) ont montré que les BR pourraient également jouer un rôle dans la maturation de la baie. De manière similaire à l'ABA, la

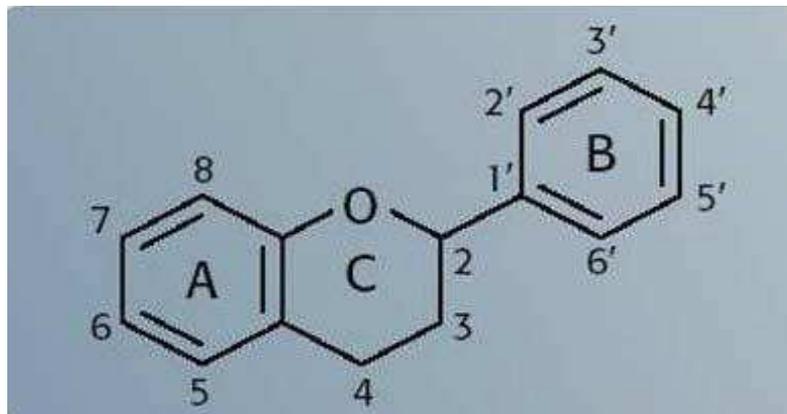


Figure 10 : Structure chimique de base des flavonoïdes

Les deux cycles benzéniques (A et B) sont reliés par un hétérocycle oxygéné (C). A, B et C forment l'unité flavane.

véraison est marquée par une augmentation du niveau de la BR bioactive castasterone (CS) et par des modifications des profils d'expression des gènes qui contrôlent la synthèse des BR. L'application de BR exogène avant la véraison augmente le pourcentage de baies qui sont colorées ainsi que l'accumulation de sucres. A l'inverse, l'application d'un inhibiteur de la synthèse des BR retarderait la véraison [56].

A la véraison, c'est l'action concomitante de ces trois phytohormones qui permet aux baies de mûrir. Les rôles respectifs et les interactions entre ces trois hormones sont encore à déterminer. Des baies traitées avec de l'auxine synthétique montrent non seulement un retard de maturation, mais également un décalage du pic d'accumulation de l'ABA à la véraison [53]. Chez la tomate (fruit climactérique), l'application de BR exogène avance la maturation *via* l'augmentation des teneurs en éthylène [64].

II- Métabolisme des flavonoïdes dans la baie de raisin

Les flavonoïdes sont des composés naturels appartenant à la famille des polyphénols. Ces métabolites secondaires sont considérés comme des pigments universels chez les végétaux. Tous les flavonoïdes sont caractérisés par un squelette à quinze atomes de carbone en C₆-C₃-C₆ : deux cycles benzéniques (A et B) reliés par un hétérocycle oxygéné (C) formant l'unité flavane (figure 10). Présents dans la plupart des plantes vasculaires, ils jouent des rôles biochimiques et physiologiques majeurs dans tous les types cellulaires et organes (racines, tiges, bois, feuilles, fleurs, fruits et graines) où ils sont accumulés. La voie de biosynthèse des flavonoïdes est ubiquitaire dans le règne végétal et produit une variété de composés aussi bien pigmentés que non pigmentés. Ces composés sont impliqués dans de nombreux processus biologiques chez les plantes: attraction d'agents pollinisateurs *via* la pigmentation des organes floraux, germination du tube pollinique, protection contre les rayonnements U.V., et défense contre des insectes et champignons pathogènes en agissant respectivement comme des insecticides et phytoalexines [65]. Chez l'Homme, leurs pouvoirs antioxydants peuvent avoir des effets bénéfiques vis-à-vis différentes pathologies en influençant plusieurs fonctions biologiques dont la synthèse protéique, la différenciation et la prolifération cellulaire, l'angiogenèse et la carcinogenèse [66].

L'étude de la voie métabolique des flavonoïdes a contribué directement ou indirectement à la découverte de nombreux principes biologiques fondamentaux durant les deux derniers siècles. Gregor Mendel, fondateur de la génétique moderne, utilisa la couleur des fleurs et des graines

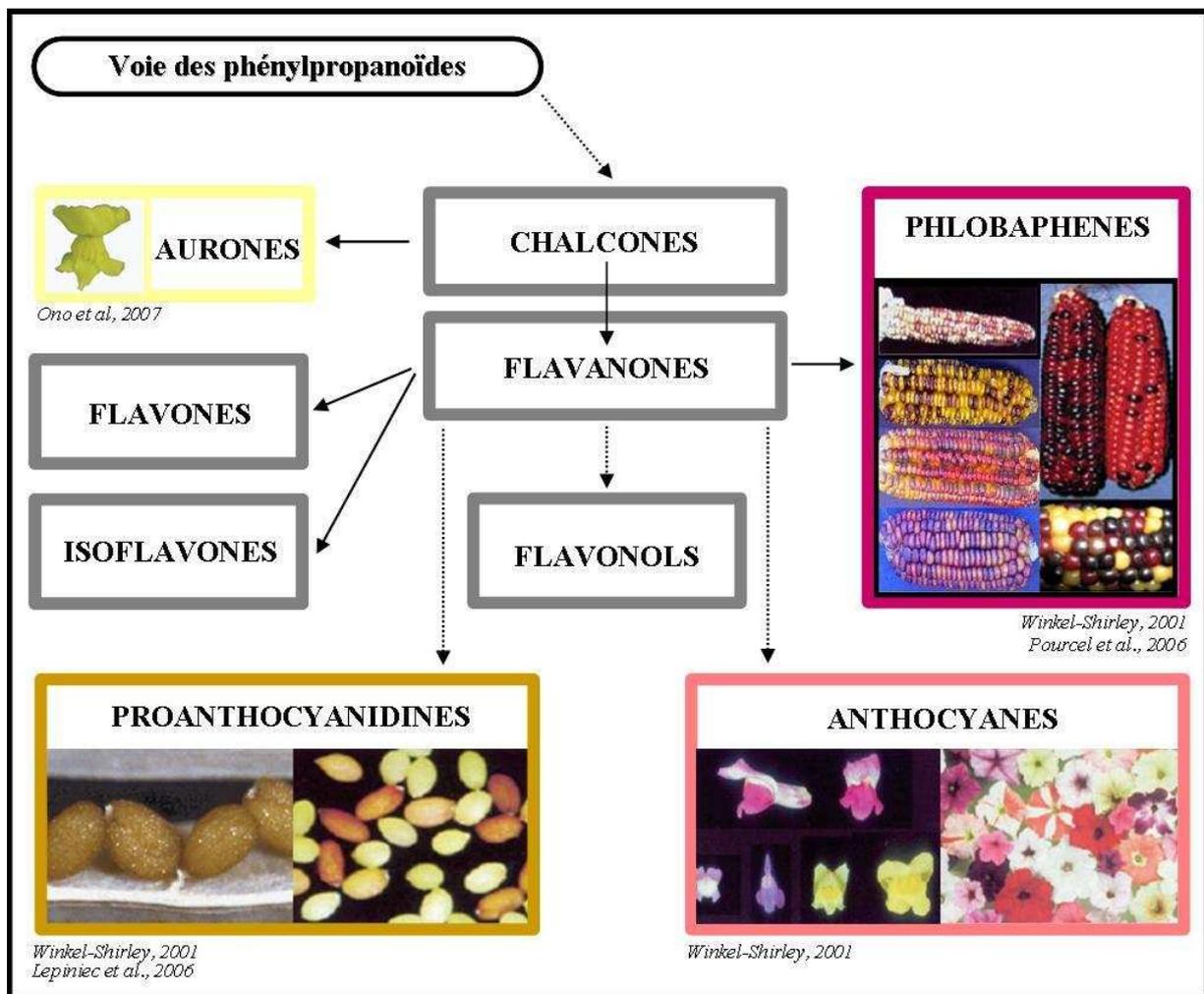


Figure 11 : Représentation schématique des principales classes de flavonoïdes

La biosynthèse des flavonoïdes dérive de la voie des phénylpropanoïdes. Les chalcones sont les précurseurs communs à tous les flavonoïdes. Les photographies illustrent le rôle majeur de certains flavonoïdes dans la pigmentation des plantes. Les références bibliographiques dont sont issues les photographies sont indiquées et correspondent aux numéros [67-70]. Lorsqu'une seule réaction enzymatique permet la synthèse du composé, la flèche est pleine. A l'inverse, quand plusieurs réactions enzymatiques sont nécessaires, les flèches sont en pointillés.

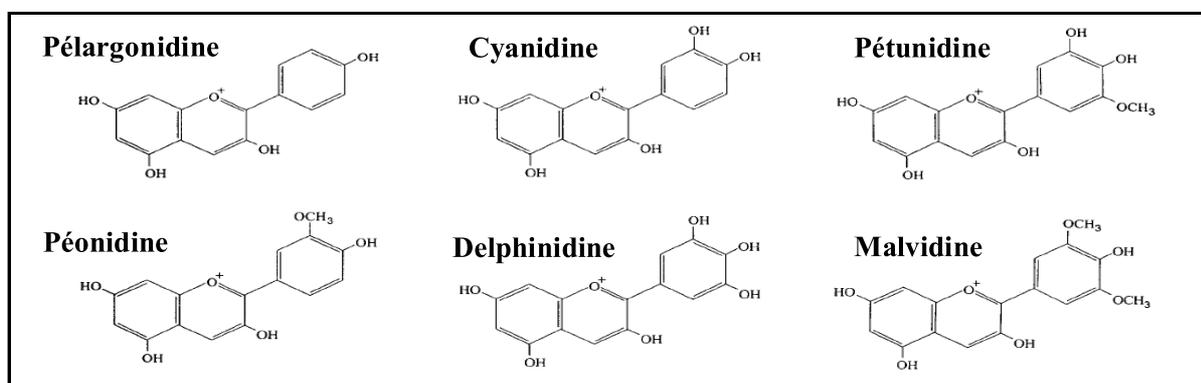


Figure 12: Structure chimique des six anthocyanes majeures de la baie de raisin [71].

de petits pois, parmi d'autres caractères, pour développer sa théorie sur l'hérédité. Le prix Nobel Barbara McClintock, en étudiant la pigmentation des grains de maïs, découvrit les éléments génétiques mobiles. Plus récemment, l'analyse de la coloration des grains de maïs et des tissus végétatifs a permis d'identifier les phénomènes d'épigénétiques connus jusqu'alors sous le nom de paramutations. De la même manière, l'utilisation de plants de Pétunia transgéniques surexprimant des enzymes clés de la voie de biosynthèse des flavonoïdes a révélé les phénomènes de co-suppression/RNAi [67].

1- Flavonoïdes chez les végétaux

Les flavonoïdes constituent un groupe de plus de 9000 composés phénoliques. Ces molécules aromatiques sont toutes issues de la voie des phénylpropanoïdes et peuvent être classées en 6 groupes : les chalcones, les flavones, les flavonols, les flavanones, les anthocyanes et les tannins condensés (ou proanthocyanidines, PA) (figure 11). Les aurones qui constituent un septième groupe moins représenté sont à l'origine du jaune brillant de plusieurs fleurs ornementales dont celles d'*Antirrhinum majus* [72]. Certaines plantes synthétisent aussi des formes spécialisées de flavonoïdes comme les isoflavones, composés incolores présents dans un grand nombre de familles de végétaux, mais les légumineuses (en particulier les *Papilionoideae*) sont les producteurs les plus caractéristiques [69]. Le Shorgo à sucre (*Sorghum bicolor*), le Maïs (*Zea mays*) et le Gloxinia (*Sinningia cardinalis*) sont parmi les rares espèces à produire la 3-déoxyanthocyanine (forme polymérisée du phlobaphène). Les phlobaphènes, polymères de flavan-4-ols aussi connus sous le nom de déoxyflavonoïdes, sont des pigments marrons-rouges présents dans de nombreux tissus de plantes dont le péricarpe de maïs [69, 73].

Les anthocyanes (du grec *anthos* = fleur et *kyanos* = bleu) font l'objet d'une attention particulière car ce sont les matières colorantes les plus importantes et les plus répandues dans les plantes. Ces pigments sont connus pour être responsables, dans les pétales et les feuilles, des couleurs allant du rose au bleu en passant par le rouge, le mauve et le violet. Les anthocyanes correspondent aux anthocyanidines et aux anthocyanines. Au niveau structural, ce sont des molécules formées de deux cycles benzéniques reliés par un ion flavylum (cycle C). Elles diffèrent les unes des autres par le nombre d'acylations, hydroxylations, méthylations, la nature et le nombre de sucres liés à la molécule ainsi que par la nature et le nombre d'acides aliphatiques ou aromatiques liés au sucre. Les structures de base sans sucre attaché au cycle aromatique (aglycones), sont appelées anthocyanidines. Les plus répandues dans les plantes sont: la cyanidine, la delphinidine, la malvidine, la pélagonidine, la péonidine et la pétunidine (figure 12) [71]. Ces molécules sont généralement présentes dans les tissus des végétaux sous forme hétérosidique. Chacune des anthocyanidines peut être

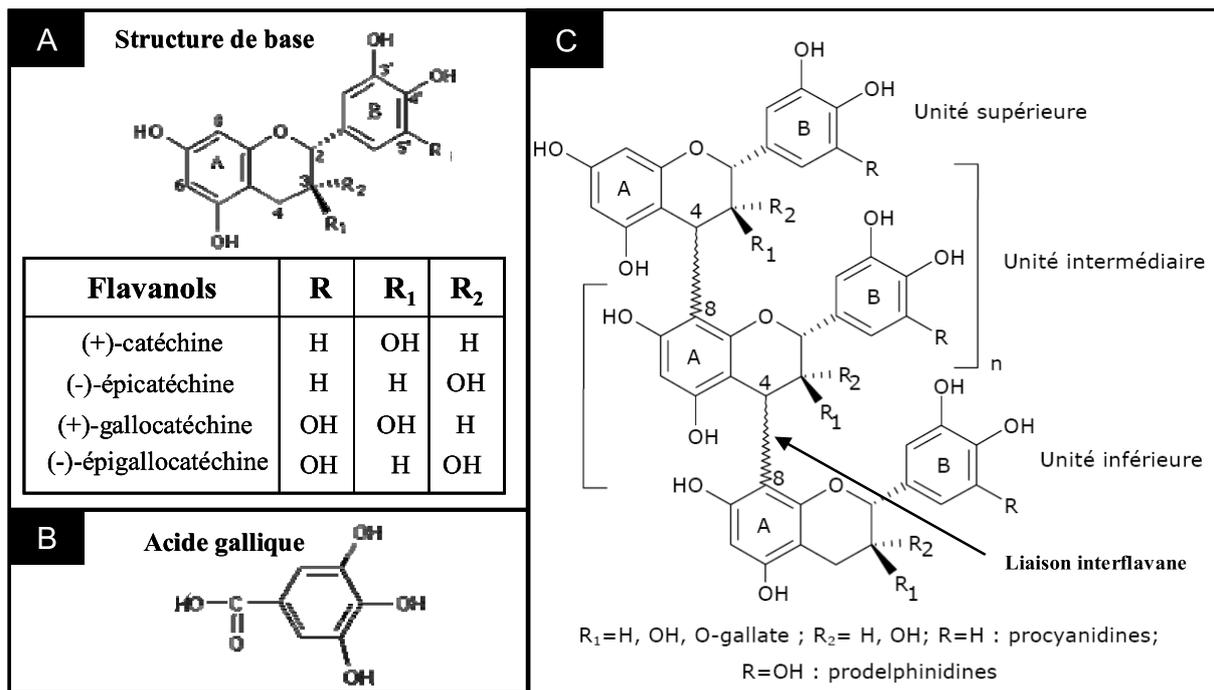


Figure 13: Structure chimique des tannins condensés ou proanthocyanidines.

A) Structure chimique des flavan-3-ols

B) Structure chimique de l'acide gallique

C) Structure chimique des tannins condensés. Le nombre d'unités monomériques " n " peut varier de 1 à 30.

glycosylée et acylée sur différents sites et avec différents groupes sucres et acyls. Une fois glycosylées, les anthocyanidines deviennent des anthocyanines. Les sucres les plus communément liés aux anthocyanidines sont le glucose, le galactose, le rhamnose et l'arabinose [74]. Ces sucres sont eux-mêmes estérifiés par des acides aliphatiques ou aromatiques. La glycosylation et la méthylation participent à la stabilisation des anthocyanes en les protégeant de l'oxydation. L'acylation favorise la co-pigmentation intra et/ou intermoléculaire et augmente également la stabilité des anthocyanes et leur solubilité dans l'eau [75]. La conjugaison aux sucres jouerait un rôle critique dans le transport et le stockage des anthocyanes dans la vacuole. L'ensemble de ces modifications (glycosylations, acylations et méthylations), augmente non seulement le nombre d'anthocyanes (environ 400 identifiées à ce jour), mais il diversifie également leurs activités biologiques [75].

Connus depuis le Moyen Âge pour le tannage des cuirs, les proanthocyanidines sont aussi déterminantes dans la saveur et l'astringence des thés, vins et jus de fruits. La chimie des PA a été étudiée depuis de nombreuses décennies. Leur nom vient de leur aptitude à libérer des anthocyanes en milieu acide, à chaud, par rupture de la liaison intermonomérique [76]. Les tannins condensés sont des polymères de flavan-3-ols (ou flavanols). La condensation des flavanols est réalisée entre une unité flavanol électrophile (appelé unité terminale ou unité supérieure) et une unité flavanol nucléophile (unité d'extension ou unité inférieure) (figure 13C). La structure des PA varie suivant la nature de l'unité terminale et des extensions, la position et la stéréochimie de la liaison avec l'unité inférieure, et le nombre d'unités d'extension (degré de polymérisation). Généralement, la liaison interflavane entre les unités se fait entre la position 4 de l'unité supérieure et la position 8 de l'unité inférieure, mais des liaisons de type C₄-C₆ sont également possibles [76, 77]. Les PA peuvent résulter de la condensation de 1 à 30 unités monomériques. Les monomères diffèrent par la stéréochimie des carbones asymétriques C₂ et C₃ et par le patron d'hydroxylation du noyau B. Les unités terminales de la plupart des PA sont les flavan-3-ols (+)-catéchines (stéréochimie 2-3-*trans*) et les (-)-épicatechines (stéréochimie 2-3-*cis*) [69]. De plus, les monomères peuvent être dihydroxylés (catéchine/épicatechine) ou trihydroxylés (gallocatéchine/épigallocatéchine) (figure 13A). Assez couramment, le groupement hydroxyl en position C₃ des unités flavanols est estérifié, souvent avec de l'acide gallique, formant l'(-)-épicatechine-3-*O*-gallate (figure 13B). L'(-)-épicatechine-3-*O*-gallate libre est l'un des principaux composés phénoliques dans le thé [76]. Les monomères existent en tant que tels ou s'associent en dimères, en oligomères (constitués de deux à dix monomères) et en polymères. En fonction de la composition en monomères, on distingue les procyanidines (polymères de catéchine et

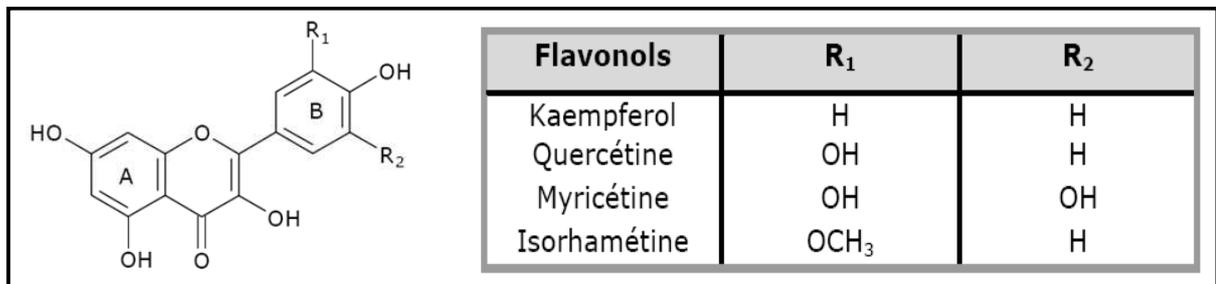


Figure 14 : Structure chimique des principaux flavonols

épicatéchine) des prodelphinidines (polymères de gallocatéchine et d'épigallocatéchine) (figure 13C).

Les flavonols, sont présents dans la plupart des plantes supérieures. Ils interviennent dans la protection contre les rayonnements U.V. dans l'épiderme et comme co-pigments des anthocyanes dans les fleurs, les fruits et les graines [69]. Ils participent également aux interactions plantes-pathogènes [78]. Les principaux aglycones de flavonols libres sont la quercétine, le kaempférol, la myricétine et l'isorhamnétine (figure 14). Dans les plantes, les flavonols sont majoritairement présents sous des formes glycosylées. Généralement, les sites de glycosylation sont les positions C₃ et C₇, et les oses liés, le glucose et le rhamnose.

2- Biosynthèse des flavonoïdes chez les végétaux

Voie commune de biosynthèse des flavonoïdes

L'étude de mutants a permis d'identifier les enzymes intervenant dans la synthèse des différentes classes de flavonoïdes (figure 15). Les mutations des enzymes de cette voie n'affectent pas la viabilité des plantes et sont facilement détectables de part les phénotypes pigmentaires qu'elles engendrent dans les fleurs, les fruits et les graines. Les mutations qui touchent les enzymes des étapes précoces entraînent l'apparition de phénotypes incolores. A l'inverse, celles se produisant dans les étapes tardives affectent le type d'anthocyanes produits. Ces études ont été menées initialement et principalement chez le Pétunia (*Petunia hybrida*), le Muflier (*Antirrhinum majus*) et le Maïs [79]. Récemment, des mutants *transparent testa (tt)* affectés dans la synthèse des flavonoïdes ont été identifiés chez *Arabidopsis thaliana* [69]. La complémentation de mutants *tt* par des gènes orthologues de maïs a montré qu'ils étaient largement conservés dans le règne végétal [80]. De nombreux gènes structuraux ont pu être clonés et identifiés ces dernières années dans les espèces modèles. Cependant, dans cette partie, nous détaillerons seulement les étapes enzymatiques qui mènent à la synthèse des anthocyanes et des proanthocyanidines.

La phénylalanine ammonia lyase (PAL) est la première enzyme impliquée dans la synthèse des flavonoïdes et fut la première enzyme isolée (figure 15) [81]. Elle relie le métabolisme primaire du shikimate, qui conduit aux acides aminés aromatiques, au métabolisme secondaire des phénylpropanoïdes. La PAL, la cinnamate 4-hydroxylase (C4H) et la p-coumarate CoA ligase (4CL) synthétisent, à partir de la phénylalanine, le précurseur commun à tous les flavonoïdes : le 4-coumaroyl-CoA. Le premier flavonoïde (chalcone) est synthétisé par la condensation du p-coumaroyl-CoA et de trois molécules de malonyl-CoA sous l'action de la chalcone synthase (CHS). Les chalcones synthétisées par la CHS sont ensuite

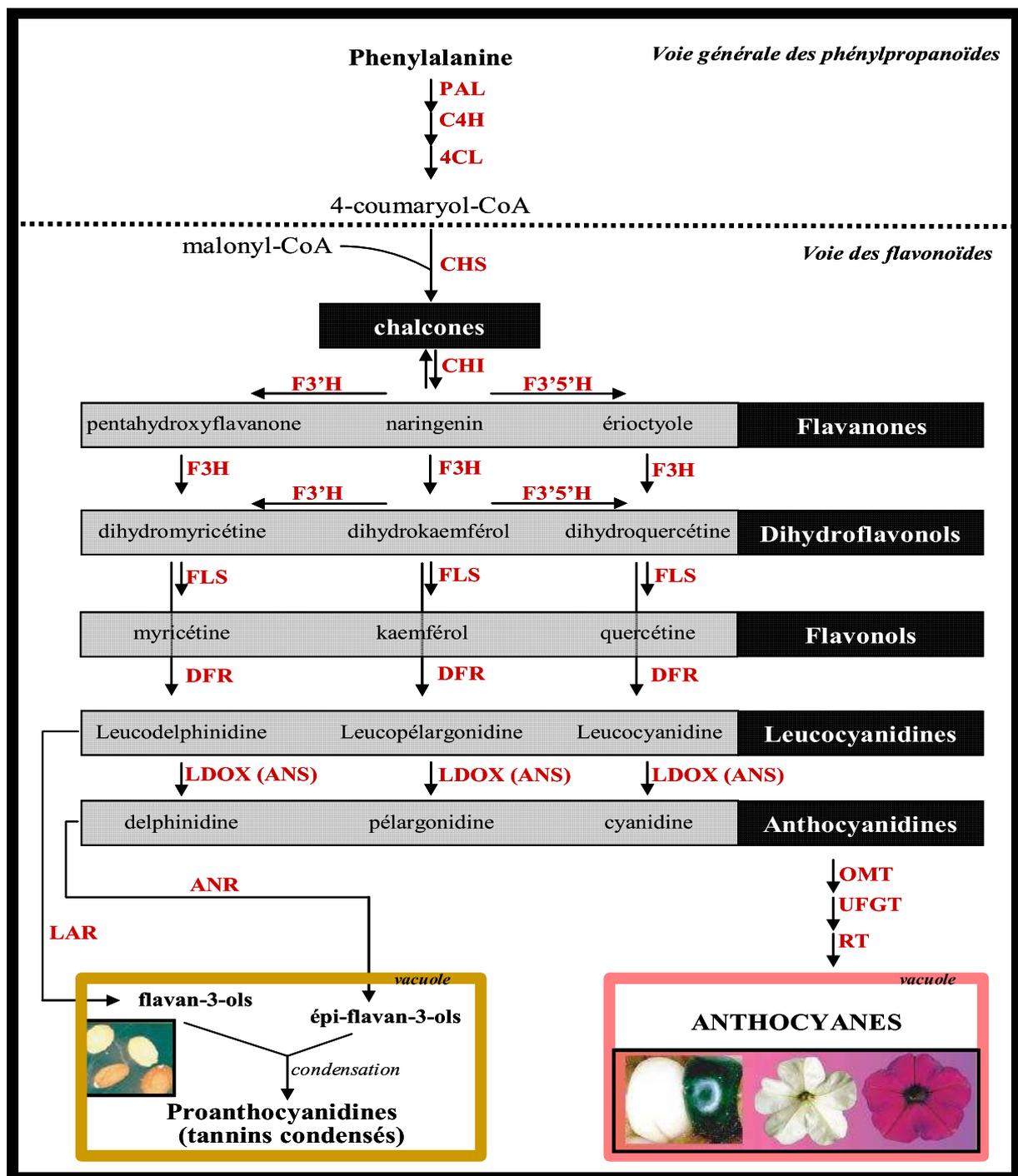


Figure 15 : Représentation schématique des voies de biosynthèse des principaux composés phénoliques présents chez les végétaux d'après [82].

La première étape de biosynthèse des flavonoïdes est catalysée par la chalcone synthase (CHS), qui condense le malonyl CoA et le 4-coumaroyl CoA. Les enzymes, PAL (phénylalanine ammonia lyase), C4H (cinnamate 4-hydroxylase) et la 4CL (4-coumarate CoA ligase), sont communes au métabolisme des phénylpropanoïdes. Les enzymes générales de la voie de biosynthèse des flavonoïdes sont la chalcone isomérase (CHI), la flavanone 3-hydroxylase (F3H), la flavonoïde 3' hydroxylase (F3'H), la flavonoïde 3'5' hydroxylase (F3'5'H), la dihydroflavonol (DFR) et l'anthocyanidine synthase (ANS) ou la leucoanthocyanidine dioxygénase (LDOX). La flavanol synthase (FLS) est spécifique des flavonols, l'anthocyanidine réductase (ANR) et la leucoanthocyanidine réductase (LAR) interviennent dans la synthèse des tannins condensés. Les O-méthyl transférases (OMT), la rhamnolyse transférase (RT) et l'UDP-glucose:flavonoïde 3-O-glucosyltransferase (UFGT) sont spécifiques de la synthèse des anthocyanes.

rapidement métabolisées sous l'action de la chalcone isomérase (CHI) en flavanones. L'hydroxylation des flavanones par la flavanone-3-hydroxylase (F3H) conduit à la formation des dihydroflavonols. Les flavanones ou les dihydroflavones peuvent subir des hydroxylations sur leur cycle B sous l'action de flavonoïde 3'-hydroxylases (F3'H) et de flavonoïde 3'5'-hydroxylases (F3'5'H). Ces enzymes déterminent ainsi le type d'anthocyanes produites. L'activité de la F3'H conduirait à l'accumulation d'anthocyanes de type cyanidine dans les plantes alors que celle de la F3'5'H conduirait aux anthocyanes de type delphinidine. La dihydroflavonol reductase (DFR) catalyse une étape clé dans le métabolisme des flavonoïdes. Elle convertit les dihydroflavonols en leucoanthocyanidines incolores (flavan-3,4-diol), précurseurs communs des anthocyanes et des proanthocyanidines [43].

Biosynthèse des anthocyanes

L'anthocyanidine synthase (ANS), aussi connue sous le nom de leucoanthocyanidine dioxygénase (LDOX) est une enzyme essentielle dans la formation des anthocyanes (figure 15). L'ANS, la F3H et la flavonol synthase (FLS) sont des membres de la famille des dioxygénases-2-oxoglutarate-dépendante, capable de catalyser l'oxydation des leucoanthocyanidines incolores en précurseurs d'anthocyanidines colorées [75]. La caractérisation fonctionnelle de l'ANS dans la synthèse des anthocyanes, a été réalisée pour la première fois chez le maïs. L'expression ectopique du gène *A2* dans un mutant déficient en activité ANS restaurait la synthèse d'anthocyanes [83].

Les anthocyanidines, dont la structure flavylum est labile aux pH physiologiques, sont ensuite stabilisées par des glycosylations en position 3 et 5. La stabilisation des anthocyanidines impliquant l'ajout d'un groupement glucose par l'UDP-glucose:flavonoïde 3-O-glucosyltransférase (UFGT ou 3GT) a été beaucoup étudiée [75]. Le gène *bronze-1* (*bz-1*) de maïs a été la première 3GT clonée [84]. Des anthocyanes contiennent des sucres autres que le glucose et mettent en jeu des enzymes de types 5GT, 7GT et RT (rhamnolyse transférase) [65]. A leur tour, les sucres peuvent être acylés par des acides organiques *via* des acétyltransférases (ACT) ou méthylées (cyanidine et delphinidine) *via* des O-méthyl transférases (OMT).

Biosynthèse des proanthocyanidines

Les sous-unités de base des PA sont fournies par deux voies métaboliques distinctes (figure 15). Deux enzymes, la leucoanthocyanidine réductase (LAR) et l'anthocyanidine réductase (ANR), catalysent la synthèse stéréosélective des monomères flavan-3-ols *trans* et *cis* respectivement [69].

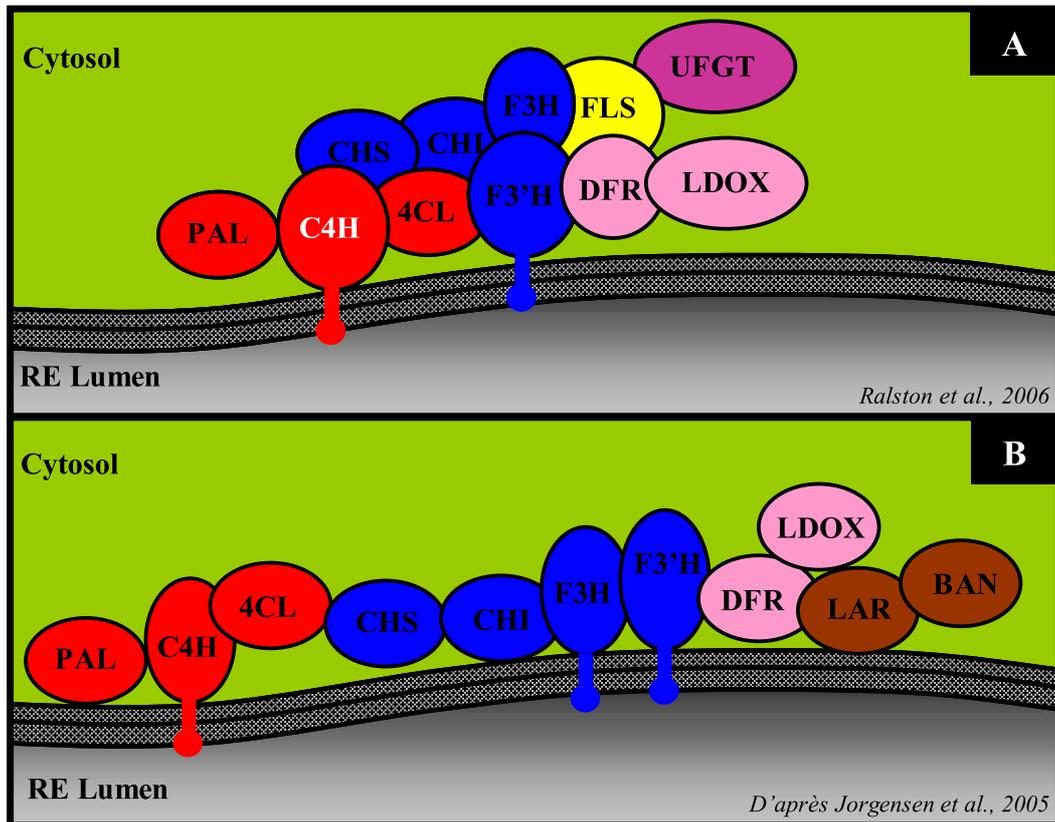


Figure 16 : Modèles d'organisation des enzymes du métabolisme des flavonoïdes en canal métabolique.

Les enzymes communes à la voie des phénylpropanoïdes sont indiquées en rouge et celles communes à la voie des flavonoïdes en bleu. Les enzymes qui sont spécifiquement impliquées dans une voie de biosynthèse sont indiquées par d'autres couleurs. Les enzymes solubles s'associent autour des enzymes cytochromes P450 (C4H, F3H et F3'H) permettant ainsi l'ancrage sur la face cytosolique du réticulum endoplasmique du complexe multi-enzymatique ou métabolon. Le gène BAN code pour l'ANR. Le modèle A est proposé par Ralston et *al.* [90] et le modèle B par Lepiniec et *al.* et Jorgensen et *al.* [69, 91].

Les abréviations sont : PAL (phénylalanine ammonia lyase), C4H (cinnamate 4-hydroxylase), 4CL (4-coumarate CoA ligase) CHS (chalcone synthase) CHI (chalcone isomérase), F3H (flavanone 3-hydroxylase), F3'H (flavonoïde 3' hydroxylase), DFR (dihydroflavonol) et LDOX (leucoanthocyanidine dioxygénase), FLS (flavonol synthase), ANR (anthocyanidine réductase), LAR (leucoanthocyanidine réductase), UFGT (UDP-glucose:flavonoïde 3-O-glucosyltransferase), RE (réticulum endoplasmique).

Dans la première voie, la formation des flavan-3-ols *trans* (les catéchines) est consécutive à l'action de la DFR et de la LAR. L'activité LAR a pu être détectée dans des feuilles issues de légumes riches en tannins mais ce n'est qu'en 2003 que le premier gène *LAR* a été cloné chez *Desmonium uncinatum* [85]. La deuxième voie a été mise en évidence récemment chez *Arabidopsis thaliana* et *Medicago truncatula* [86, 87]. Les plantes d'*Arabidopsis* n'accumulent que des flavan-3-ols *cis* (épicatechines) dans les téguments de leurs graines et ne possèdent pas d'activité LAR [85]. Une approche génétique a permis d'identifier le gène *BANYULS* (*BAN*) qui code pour une ANR, enzyme spécifique de la synthèse des épicatechines [88]. Les travaux récents de *Xie et al.* ont pu démontrer que l'ANR utilise comme substrat les anthocyanidines formées par l'ANS/LDOX [87]. En effet, la mutation des deux allèles ANS/LDOX chez *Arabidopsis*, *tt18* et *tds4* (*tannin deficient seed 4*), bloque la biosynthèse des PA de type épicatechine [89]. Les mutants *ban* ont des phénotypes *tt* caractérisés par l'absence de PA et l'accumulation précoce d'anthocyanes dans les téguments [66]. A l'inverse, l'expression ectopique de *BAN* dans des pétales de fleurs de tabac et des feuilles d'*Arabidopsis* entraîne une perte d'anthocyanes et une accumulation de PA [87].

Les études menées ces dernières années chez *Arabidopsis* et *Medicago* ont permis d'identifier les enzymes responsables de la synthèse des unités terminales des PA. Mais plusieurs questions importantes sur le métabolisme des PA demeurent. À l'heure actuelle, les mécanismes à l'origine de la condensation et la polymérisation des unités monomériques dans la vacuole restent encore inconnus [76]. *TT10*, gène récemment cloné chez *Arabidopsis*, code une polyphénol oxydase de type laccase et serait impliqué dans la polymérisation oxydative des flavonoïdes. Ce gène, impliqué dans la formation des pigments bruns polymères d'épicatechine, pourrait catalyser le brunissement oxydatif des PA incolores. Mais la fonction enzymatique précise de *TT10* reste encore inconnue [70].

Organisation subcellulaire du métabolisme des flavonoïdes

Les enzymes de la voie des flavonoïdes semblent être organisées en complexes multi-enzymatiques, associés à la face cytosolique du réticulum endoplasmique (RER), *via* l'ancrage dans la membrane des protéines cytochromes P450 de la voie (F3H, F3'H, C4H, IFS) (figure 16A) [90, 92-94]. Ces canaux métaboliques permettraient non seulement de séquestrer les intermédiaires toxiques ou les produits instables, mais aussi de diriger le flux dans l'une des multiples branches de la voie qui coexistent simultanément dans la cellule [91, 95]. La CHS, CHI, F3H, la DFR et l'UFGT ont été colocalisées sur la face cytosolique du réticulum endoplasmique (RER) chez *Arabidopsis* et une interaction directe entre la CHS, la

CHI, la F3H et la DFR a été démontrée [96-98]. Ces résultats signifieraient que l'agencement des enzymes dans le complexe ne serait pas linéaire mais plutôt sphérique. Dans ce canal, la CHS serait en contact non seulement avec l'enzyme suivante de la voie, la CHI, mais aussi avec la F3H et la DFR [68]. Concernant la biosynthèse des PA, Lepiniec et *al.* ont proposé que toutes les enzymes, y compris celles impliquées dans la biosynthèse des isomères flavanols seraient localisées sur le RE ou le tonoplaste (figure 16B). En effet, si aucune co-localisation de la LAR et de l'ANR n'a encore été déterminée, il a été démontré que la polymérisation des PA est tributaire de la présence de membranes vacuolaires intactes [69].

Même si la plupart des données convergent vers une synthèse des flavonoïdes strictement localisée dans le cytoplasme, des travaux récents pourraient remettre en question ce modèle. Plusieurs études ont montré la présence des flavonoïdes dans les noyaux de nombreuses espèces (*Arabidopsis*, *Brassica napus*, *Fleaveria chloraefolia*, *Picea abies*, *Tsuga canadensis* et *Tuxus baccata*) [79]. En parallèle, des enzymes spécifiques de la voie des flavonoïdes ont pu être localisées dans des compartiments cellulaires autres que le cytoplasme. L'aureusidine synthase, une polyphénol oxydase qui catalyse la formation des aurones à partir des chalcones dans les fleurs d'*Antirrhinum majus*, a été récemment localisée dans les vacuoles [98]. La CHS et la CHI ont également été détectées dans des noyaux de cellules d'*Arabidopsis* [99]. De fait, de nouvelles perspectives sur des activités nucléaires biosynthétiques ou régulatrices des flavonoïdes peuvent être envisageables [79]. Ils pourraient protéger l'ADN contre les dommages oxydatifs causés par les rayonnements U.V. ou intervenir dans le contrôle de la transcription des gènes nécessaires à la croissance et au développement comme par exemple le transport d'auxine [99-101]. Ces données ont également ouvert de nouvelles perspectives de recherche quant aux mécanismes qui seraient à l'origine de la distribution intracellulaire des enzymes de cette voie. La CHI, protéine de 27 kDa, est de loin la plus petite des enzymes des flavonoïdes et pourrait se déplacer à travers les pores nucléaires par diffusion passive. Cependant seule la CHS, protéine de 47 kDa, possède un signal d'adressage au noyau (NLS: Nuclear Localisation Signal). Ce signal étant situé dans une région topologiquement opposée au domaine de dimérisation de la CHS, il pourrait s'associer à d'autres enzymes de la voie de biosynthèse des flavonoïdes et ainsi permettre leur transport dans le noyau [99].

Transport et compartimentation des flavonoïdes

Une fois synthétisés, les flavonoïdes sont adressés vers leurs compartiments d'activité d'une part pour remplir leurs fonctions biologiques et, d'autre part, en raison de leur toxicité intrinsèque. Les flavonols glycosylés, les anthocyanes et les PA sont surtout accumulés dans la vacuole, alors que les phlobaphènes et les flavonols méthylés sont retrouvés

majoritairement dans la paroi cellulaire [69]. Plusieurs types de transporteurs de la membrane vacuolaire ou plasmique ont été identifiés [102].

Des membres de la famille des glutathione S-transférase (GST) se sont avérés nécessaires à la séquestration des anthocyanes et des PA dans la vacuole chez le maïs (BZ2), le Pétunia (AN9) et *Arabidopsis* (TT19) [103-106]. Les GST sont des enzymes de détoxification cellulaire [107]. Chez les plantes, la détoxification de composés xénobiotiques implique trois phases : **1)** activation des composés phytotoxiques par oxydation ou hydrolyse, **2)** conjugaison avec des molécules hydrophiles (glucose, malonate ou glutathion) et **3)** séquestration des conjugués par des transporteurs membranaires de type ABC (*ATP-binding cassette*) [79]. De manière analogue, les GST peuvent participer au transport cytoplasmique des flavonoïdes avant qu'ils soient transportés à l'intérieur de la vacuole, *via* une protéine du tonoplaste qui nécessite une activité H⁺-ATPase. Cependant, à ce jour, aucun conjugué anthocyan-glutathion n'a été observé *in vivo* et le transport vacuolaire des anthocyanes ne nécessite pas leur glutathionylation par l'enzyme AN9. Ainsi, les GST seraient impliquées dans la liaison et la stabilisation intracellulaire des flavonoïdes plutôt que dans la catalyse de leur glutathionylation [108]. Chez *Arabidopsis*, *AHA10* (*Autoinhibited H1-ATPase isoform 10*) code une pompe H⁺-ATPase qui est impliqué dans le métabolisme des flavonoïdes. Le phénotype des cellules endothéliales du tégument des graines *aha10* révèle une perturbation spécifique dans le stockage des flavonoïdes dans la vacuole et notamment dans l'accumulation des PA [109]. Un autre type de transporteur impliqué dans la translocation des flavonoïdes dans la vacuole a pu être identifié. Il s'agit des transporteurs MRP (Multidrug Resistance associated Protein), une sous-classe de protéines ABC. Chez *Arabidopsis*, deux MRP ont été clonés (*AtMRP1* et *AtMRP2*) et caractérisés chez la levure [110, 111]. Chez le Maïs, les mutants antisens *ZmMRP3* présente des teneurs en anthocyanes plus faibles que dans le type sauvage [112].

Des approches génétiques ont révélé un autre système de transport des flavonoïdes. Chez *Arabidopsis*, les protéines TT12 appartiennent à la famille des transporteurs de type MATE (Multidrug And Toxic compound Extrusion) et sont localisées sur la membrane tonoplastique. Les mutants *tt12* présentent dans les cellules endothéliales une structure vacuolaire aberrante et le transport des PA dans la vacuole est affecté [113]. De même, les mutations *ZmMPP4* (Maïs) et *MTP77* (Tomate) suggèrent que les transporteurs MATE peuvent également intervenir dans le transport des anthocyanes [112, 114].

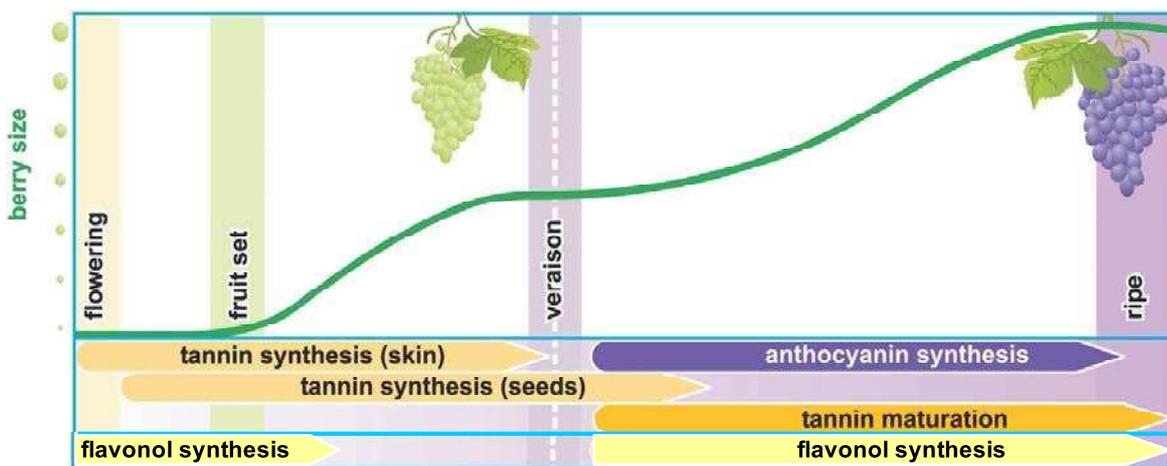


Figure 17: Cinétique d'accumulation des composés phénoliques pendant le développement de la baie de raisin (d'après [118]).

L'évolution de la taille des baies est illustrée par la sinusoïde verte. L'évolution de la couleur des baies est également indiquée. La synthèse des tannins débute dès les premiers stades de développement et se poursuit jusqu'à la véraison dans la pellicule et jusqu'à une à deux semaines après véraison dans les pépins. La maturation des tannins a lieu pendant la phase de maturation de la baie. Après véraison dans les cépages rouges, quand la synthèse des tannins est terminée, les anthocyanes sont spécifiquement accumulées dans la pellicule des baies. Les flavonols sont accumulés avant et après floraison.

Un troisième type de transport impliquerait des structures similaires à des vésicules appelées AVI (Anthocyanic Vacuolar Inclusion). Des analyses microscopiques semblent indiquer que les cellules accumulant des fortes quantités d'anthocyanes utiliseraient des composants de la voie sécrétoire pour un transport direct des anthocyanes dans la vacuole. Après synthèse dans le RE, les anthocyanes seraient séquestrées dans des AVI dérivées du RE qui seraient directement adressées à la vacuole sans passer par le Golgi [115, 116].

3- Composés phénoliques identifiés dans la baie de raisin

En viticulture, les composés phénoliques sont considérés comme des marqueurs essentiels de la qualité organoleptique des vins. Outre leur rôle dans la coloration des baies, ils sont responsables de la structure tannique des vins rouges, particularité gustative très originale qui confère aux grands vins leurs caractéristiques spécifiques.

Les anthocyanes

La présence ou l'absence d'anthocyanes dans la baie différencie les raisins rouges des raisins blancs. Dans les cépages blancs, les baies deviennent translucides au cours de la maturation, tandis que les baies des cépages rouges se nuancent du rose au rouge. Ces pigments sont localisés dans les vacuoles des trois ou quatre premières assises cellulaires de l'hypoderme, et exceptionnellement dans la pulpe des cépages dits teinturiers [117]. La teneur en anthocyanes des raisins augmente durant la maturation, cette accumulation débutant à la véraison (figure 17). Leur synthèse entraîne une augmentation rapide de leur concentration suivie par une phase de stabilisation avant une diminution de la teneur en anthocyanes en fin de maturation et/ou durant la sur-maturité [119].

En dehors de leur rôle pigmentaire dans les baies, les anthocyanes contribuent aux qualités organoleptiques et chimiques des vins. En effet, en interagissant avec d'autres composés du vin (flavonoïdes, protéines et polysaccharides) mais aussi à cause de leur autocondensation, elles jouent un rôle dans l'intensité et la stabilité de la couleur [120, 121]. Cinq anthocyanidines ont été identifiées chez la vigne : la delphinidine, la cyanidine, la péonidine, la pétunidine et la malvidine [71]. Les dérivés de malvidine sont généralement les principales formes présentes et déterminent la couleur des raisins rouges. La quantité totale en anthocyanes et la composition relative des différentes formes sont donc des facteurs importants pour la vinification. Chaque cultivar présente un profil typique d'anthocyanes qui peut être utilisé pour une classification taxonomique [78]. Les variétés rouges de *Vitis vinifera* produisent généralement des dérivés 3-monoglucosides, 3-acétylglucosides et 3-p-coumaroylglucosides [119]. La quantité et la composition des anthocyanes présentes dans

les raisins rouges varient en fonction du cultivar, du degré de maturité, des conditions climatiques, de la région de production et des pratiques culturales [122].

Les proanthocyanidines

Les tannins, localisés dans la pellicule et les pépins, contribuent à la stabilisation, à la tenue, à l'amertume et à l'astringence des vins [123]. Pellicules et pépins sont caractérisés par des types de tannins différents. Dans les pellicules, les tannins sont localisés dans les couches externes de l'hypoderme. Ils sont accumulés soit sous forme libre dans le suc vacuolaire soit sous forme liée à des protéines et des polysaccharides dans les membranes vacuolaires ou à des polysaccharides pariétaux dans les parois [124]. Dans les pépins, ils sont présents dans les enveloppes externe et interne du tégument [18]. Les pépins contiennent des quantités de tannins plus élevées que les pellicules. La proportion des unités galloylées y est également nettement supérieure : 30 % des tannins, contre moins de 5 % en moyenne dans la pellicule. Les tannins des pépins sont moins polymérisés que ceux des pellicules, et le degré de polymérisation moyen évolue différemment suivant le compartiment cellulaire (vacuole ou paroi) au cours du développement de la baie [125]. Les polymères des pellicules contiennent en moyenne 20 à 30 monomères à maturité tandis que ceux des pépins ne contiennent que quatre à six unités [118]. L'unité terminale dans les pépins est représentée par une catéchine, épicatechine ou épicatechine-3-O-gallate, alors que dans la pellicule, c'est principalement une catéchine. Les unités d'extension détectées dans les baies sont constituées de catéchine, épicatechine et épicatechine-3-O-gallate. Jusqu'à présent, l'unité épigallocatechine a été détecté uniquement dans les pellicules [119].

La synthèse des PA débute pendant la formation de la fleur (figure 17). Au début du développement de la baie, l'activité de la voie de biosynthèse des flavonoïdes est dirigée essentiellement vers la synthèse des PA au détriment de la synthèse des flavonols. Puis, les teneurs en PA augmentent de manière synchrone à la formation et la maturation des pépins en développement [126]. L'accumulation des tannins dans la pellicule et les graines est quelque peu différente. Dans la pellicule, les teneurs en tannins, qui augmentent après la mise à fruit, sont maximales une à deux semaines avant véraison. Dans les pépins, la concentration en tannins augmente aussi après la mise à fruit, et devient maximale deux à quatre semaines après la véraison, date à laquelle les pépins sont mûrs et le tégument est brun [123]. Puis, durant la maturation de la baie, les teneurs en tannins déclinent, vraisemblablement du fait de l'oxydation et de la complexation des PA [123, 127]. De fait, le niveau des PA détectés dans le raisin est peut-être dû à un équilibre entre l'accumulation des PA et la diminution de leur extractabilité, et ne serait pas simplement le reflet de leur biosynthèse dans le fruit [126].

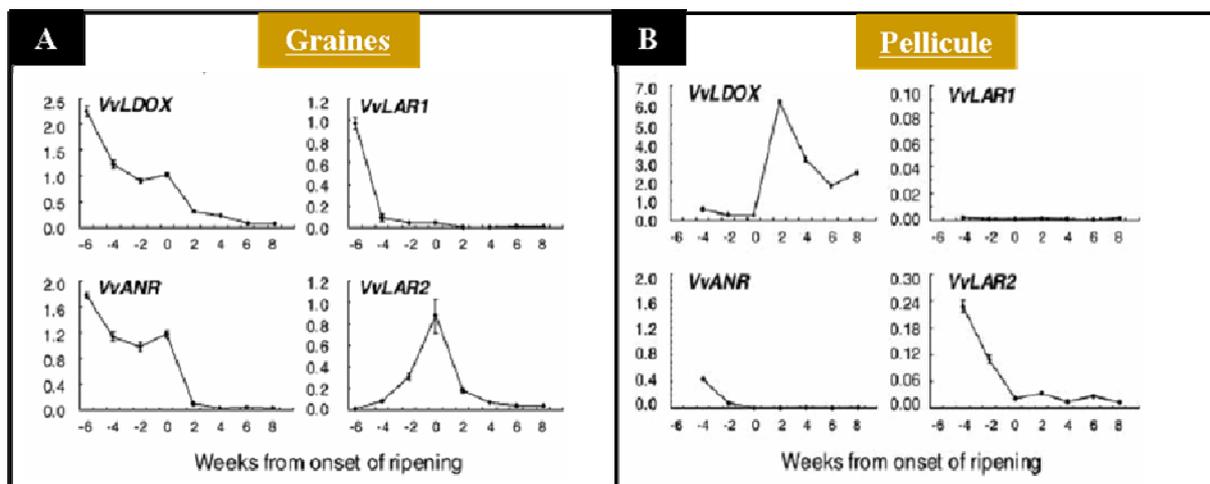


Figure 18 : Profil d'expression des gènes *VvLDOX*, *VvANR*, *VvLAR1* et *VvLAR2* au cours du développement de la baie dans les pépins (A) et la pellicule (B) [126].

Les analyses ont été réalisées par PCR quantitative et les expressions relatives par rapport à *VvUbiquitin1* sont indiquées. Les niveaux d'expression représentent la moyenne des valeurs obtenues pour trois répétitions. La véraison est indiquée par le chiffre zéro. Les abréviations utilisées sont identiques à celles de la figure 15.

Les flavonols

Les flavonols sont des pigments jaunes, synthétisés uniquement dans la pellicule des raisins, aussi bien rouges que blancs. Quatre flavonols sont majoritairement présents dans les baies de raisin: le kaempférol, la quercétine, la myricétine et l'isorhamnéthine. Les dérivés de la quercétine prédominent, quelque soit le cépage. Absente des cépages blancs, la myricétine semble être spécifique des variétés de raisins rouges [78]. Bien que les teneurs en flavonols soient inférieures à celles des anthocyanes et PA, ces composés jouent un rôle important dans la qualité organoleptique des baies et *in fine*, des vins. Ils participent à l'amertume des vins et interviennent dans la coloration des baies et des vins. En effet, les phénomènes de co-pigmentation entre anthocyanes et flavonols stabilisent ou augmentent la couleur du vin rouge [128]. Deux phases d'accumulation des flavonols ont été observées au cours du développement de la baie (figure 17). Ils sont d'abord fortement accumulés à la floraison, puis leur teneur diminue au fur et à mesure que la taille des baies augmente. Après véraison, ils sont à nouveau accumulés et atteignent des concentrations très élevées [128, 129].

1- Biosynthèse des flavonoïdes dans la baie de raisin

L'analyse de l'expression des gènes structuraux biosynthétiques des flavonoïdes a révélé que leur régulation différait selon les espèces. Chez le maïs, tous les gènes structuraux sont activés de manière coordonnée, alors que chez les dicotylédones, on distingue deux groupes de gènes co-régulés : l'induction des gènes de biosynthèse dits précoces précède celle des gènes dits tardifs [130-132]. Dans la baie de raisin, la biosynthèse des flavonoïdes consiste en deux phases séparées. La première, autour de la floraison, coïncide avec la synthèse des flavonols et des tannins, et la seconde, autour de la véraison, concorde avec la biosynthèse des anthocyanes. La synthèse des PA et des anthocyanes est temporellement séparée et la synthèse des anthocyanes a lieu après l'accumulation des PA [126].

Les gènes codant les enzymes spécifiques de la biosynthèse des tannins [*ANR* (anthocyanidine réductase) et *LAR* (leucoanthocyanidine réductase)] et des flavonols (*FLS*, flavonol synthase) n'ont été clonés que récemment chez la vigne [126]. Cinq gènes codant pour des FLS putatives (*FLS1* à *FLS5*) ont été clonés chez le Cabernet sauvignon [128, 129]. Les transcrits des gènes *FLS1*, *FLS2*, *FLS4* et *FLS5* sont accumulés dans les pellicules des baies pendant la croissance herbacée. Autour de la véraison, ces gènes ne sont plus exprimés et les teneurs en flavonols diminuent. Enfin au cours de la maturation, *FLS1*, *FLS4* et *FLS5* sont à nouveau activés et leur expression coïncide avec l'accumulation des flavonols dans les pellicules [128, 129]. Concernant la biosynthèse des PA, deux gènes *LAR* (*VvLAR1* et

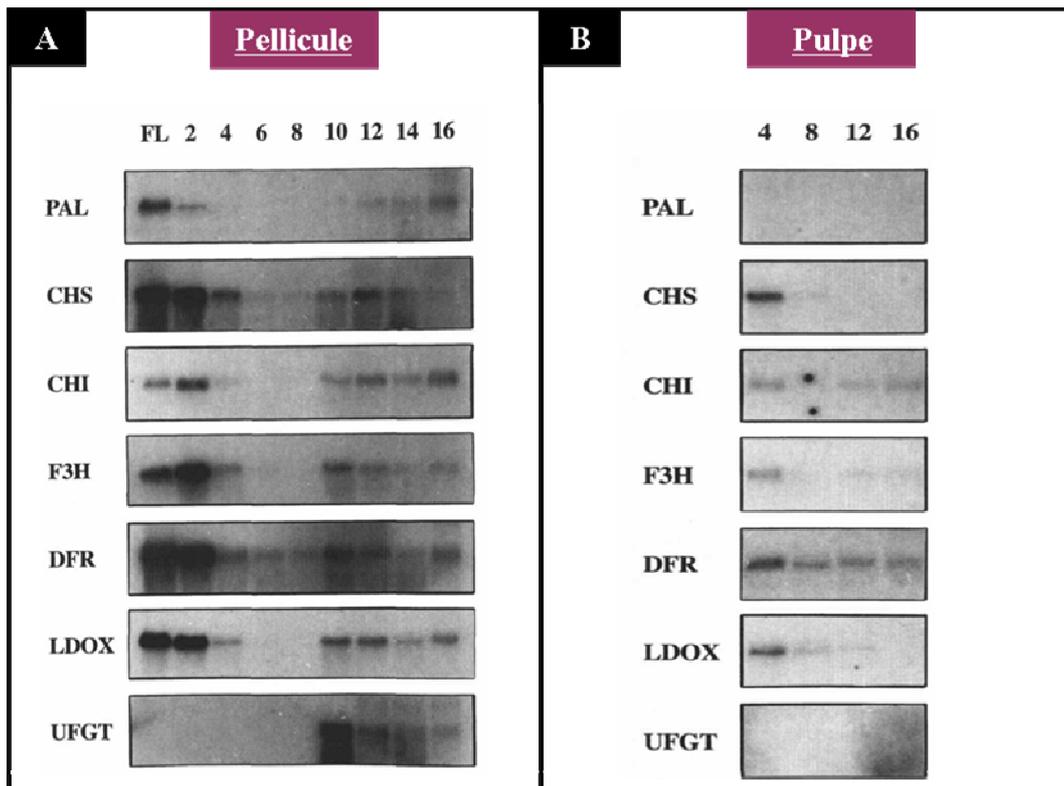


Figure 19 : Profil d'expression des gènes de biosynthèse des anthocyanes au cours du développement de la baie dans la pellicule (A) et la pulpe (B) [126].

Les analyses ont été conduites par RNA blot. Les ARN totaux (4 µg) ont été extraits de fleurs, pellicule et pulpe de baie à différents stades de développement après la floraison. Les chiffres indiquent les semaines après floraison. La véraison a lieu entre la 8^{ème} et la 10^{ème} semaine après floraison. Les abréviations utilisées sont identiques à celles de la figure 15.

VvLAR2) et un gène *VvANR* ont été identifiés et présentent une régulation tissu-temporelle spécifique. *VvANR* et *LDOX* sont exprimés avant la floraison et leur expression augmente après fécondation (figure 18). Deux semaines après la floraison, l'expression de *VvLARI*, *VvLDOX* et *VvANR* augmentent fortement, privilégiant ainsi la synthèse des tannins dans la baie. L'expression de *VvLAR2* augmente également mais dans une moindre mesure [117, 126]. Dans les pépins, les transcrits *VvLARI* s'accumulent fortement jusqu'à six semaines avant véraison, puis plus faiblement tandis que *VvLAR2* présente un pic d'expression à la véraison. Ce pic a lieu juste avant le pic d'accumulation des flavan-3-ols suggérant que les monomères de PA sont encore synthétisés dans les derniers stades de développement des pépins [126]. Dans la pellicule, *VvANR* et *VvLARI* sont exprimés respectivement jusqu'à 2 et 4 semaines avant véraison, puis aucun transcrite n'est accumulé. A l'inverse, *VvLAR2* est exprimé jusqu'à la véraison puis son niveau d'expression reste très faible pendant la maturation. A la véraison, la baisse d'accumulation des PA dans les pellicules est synchronisée à la diminution d'expression des gènes *ANR*, *LAR* et *LDOX* [126].

Un des bouleversements biochimiques majeurs faisant suite à la véraison est la coloration de la baie. La différence entre des baies blanches et rouges/noires tient en l'absence ou présence d'anthocyanes dans les couches cellulaires de la pellicule. Boss *et al.* (2000) ont montré que seule l'expression de l'isoforme 1 du gène *UFGT* était systématiquement associée à la couleur des baies. L'*UFGT* n'est exprimé que dans les pellicules des baies rouges/noires, mais jamais dans celles des cépages blancs (figure 19) [117, 133, 134]. A l'inverse, les gènes codant la *PAL*, la *CHS*, la *CHI*, la *F3H*, la *DFR* et la *LDOX* n'ont pas d'expression tissu-spécifique puisque les transcrits correspondant ont été détectés dans la plupart des tissus de la baie [133]. L'activation des gènes de biosynthèse des anthocyanes comprend deux phases (figure 17). Durant la croissance herbacée, tous les gènes de la voie de synthèse des anthocyanes, à l'exception de l'*UFGT*, sont exprimés puis, réprimés pendant la phase de latence. Suite à la véraison, l'accumulation d'anthocyanes dans la pellicule coïncide avec l'activation coordonnée de tous les gènes de la voie de biosynthèse y compris l'*UFGT* [133].

Une fois synthétisés, les composés phénoliques sont majoritairement stockés dans la vacuole. Les mécanismes moléculaires à l'origine de la compartimentation cellulaire des flavonoïdes dans la baie de raisin n'ont pas encore été identifiés. A l'heure actuelle, un seul transporteur homologue au transporteur de type BLT (bilitranslocase) chez le mammifère a été caractérisé dans la baie de raisin [135]. Néanmoins, plusieurs GST putatives ont été identifiées par des approches classiques de clonage ou par des approches globales transcriptomiques [136, 137]. Ainsi, plusieurs analyses microarrays ont montré que l'isogène *TC69505* serait impliqué dans

la séquestration des anthocyanes. En effet, son expression est plus forte dans les cépages rouges que dans les cépages blancs (60 fois) durant la maturation, et spécifiquement dans les pellicules [37, 137-139]. De plus, cet isogène présente des homologies de séquence avec le gène de *Pétunia AN9* impliqué dans la séquestration des anthocyanes [104]. Cinq autres GST (GST1 à GST5) ont été purifiées à partir de suspension cellulaire de raisins pigmentés. *GST1* et *GST2* complètent les mutants *bz2* et leur expression dans le Syrah augmente après véraison parallèlement à l'accumulation des anthocyanes [136].

I- Régulation transcriptionnelle du métabolisme des flavonoïdes

L'expression des gènes dépend du niveau de la transcription, de la maturation, de la stabilité et du transport des ARNm, de la traduction, et des modifications post-traductionnelles des protéines. Durant de nombreuses années, les phytogénéticiens et les généticiens ont sélectionnés un grand nombre de mutants affectés dans la pigmentation des fleurs et des graines [68]. Les études moléculaires menées ces 20 dernières années ont permis de distinguer deux groupes de mutants. Le premier résulte de mutations dans les gènes structuraux de la voie de biosynthèse des anthocyanes. Le second groupe est caractérisé par une modification de l'expression de plus d'un gène structural, attestant généralement d'une mutation dans un gène régulateur. L'étude de ces mutants au sein de nombreuses espèces permet aujourd'hui de tirer quelques conclusions générales sur la régulation de cette voie [140].

Le contrôle de la transcription semble être l'étape majeure de la régulation parmi les étapes menant à la production des métabolites. Cette régulation est réalisée sur des groupes de gènes structuraux de la biosynthèse des flavonoïdes. Ainsi, dans les pétales de plusieurs fleurs dicotylédones, les premiers gènes structuraux (*EBG* pour "Early Biosynthetic Genes") sont régulés différemment des gènes structuraux de la partie basse de la voie (*LBG* pour "Late Biosynthetic Gene") [140]. Ces groupes de gènes sont variables selon les espèces. *EBG* et *LBG* se séparent au niveau de la F3H si la co-production prédominante est celle des anthocyanes/flavones, ou de la DFR si elle est plus orientée vers les anthocyanes/flavonols. Dans les graines et les tissus végétatifs du maïs et les feuilles de *Perilla frutescens*, ce sont des facteurs de transcription (FT) communs qui régulent les gènes de biosynthèse des anthocyanes de la *CHS* jusqu'à la GST *Bronze2* [141-143]. Dans les baies de raisin, l'UFGT est régulée séparément des autres gènes et son activation est l'étape clé dans la production

d'anthocyanes au cours de la maturation [117, 144]. Enfin, de façon générale, les régulateurs de la voie des anthocyanes ne semblent pas agir sur la synthèse des autres métabolites phénoliques [140].

1) Mécanisme général de la régulation des gènes chez les eucaryotes

Les promoteurs et leurs éléments de régulation

Juste en amont du gène se situe le promoteur basal : séquence d'ADN minimale suffisante pour induire la transcription [145]. Cette région permet le recrutement de la machinerie de la transcription *via* des séquences spécifiques. En effet, leur présence est nécessaire à l'assemblage de l'ARN polymérase II avec des facteurs d'amorçage, pour former le complexe de pré-initiation. La plupart des promoteurs des gènes eucaryotes contiennent une séquence consensus riche en A/T, située à 25-30 nucléotides en amont du site d'initiation de la transcription (TSS). La boîte TATA est le site de fixation de la "TATA binding protein" (TBP), protéine qui assure la liaison à l'ADN du complexe de pré-initiation de la transcription [146]. Pour certains gènes, le site d'initiation de la transcription fait partie d'un élément initiateur (Inr) et le promoteur basal peut alors contenir soit une boîte TATA et un initiateur, soit un seul de ces éléments, soit aucun [147].

Dans la région promotrice des gènes, qui s'étend de quelques centaines à quelques milliers de nucléotides en amont du site d'initiation de la transcription, se trouvent d'autres séquences d'ADN appelées éléments *cis*-régulateurs [147]. Ces éléments, généralement constitués de quelques nucléotides, sont situés en amont du promoteur basal. Entre -200 et -50 pb par rapport au TSS, ils appartiennent au promoteur distal. Des éléments *cis* peuvent même se retrouver dans la séquence codante ou dans les introns d'un gène [148]. Ces éléments de régulation augmentent ou répriment la transcription d'un gène indépendamment de leur orientation et de leur distance par rapport au site d'initiation de la transcription. Ainsi, situés à des milliers de paires de bases du promoteur basal, ils peuvent agir sur la transcription ou perdre leur pouvoir dès qu'ils sont plus distants de plus de 15 à 20 pb. Chaque promoteur contient plusieurs éléments de régulation qui peuvent être regroupés physiquement pour former des modules et réguler de manière cumulative l'expression du gène correspondant [149].

Les facteurs de transcription

La régulation de la transcription s'effectue par la présence de protéines appelées FT, capables de reconnaître et de se lier spécifiquement aux séquences *cis*-régulatrices. Les FT sont des protéines nucléaires dont l'abondance, la disponibilité et l'accessibilité à leurs séquences

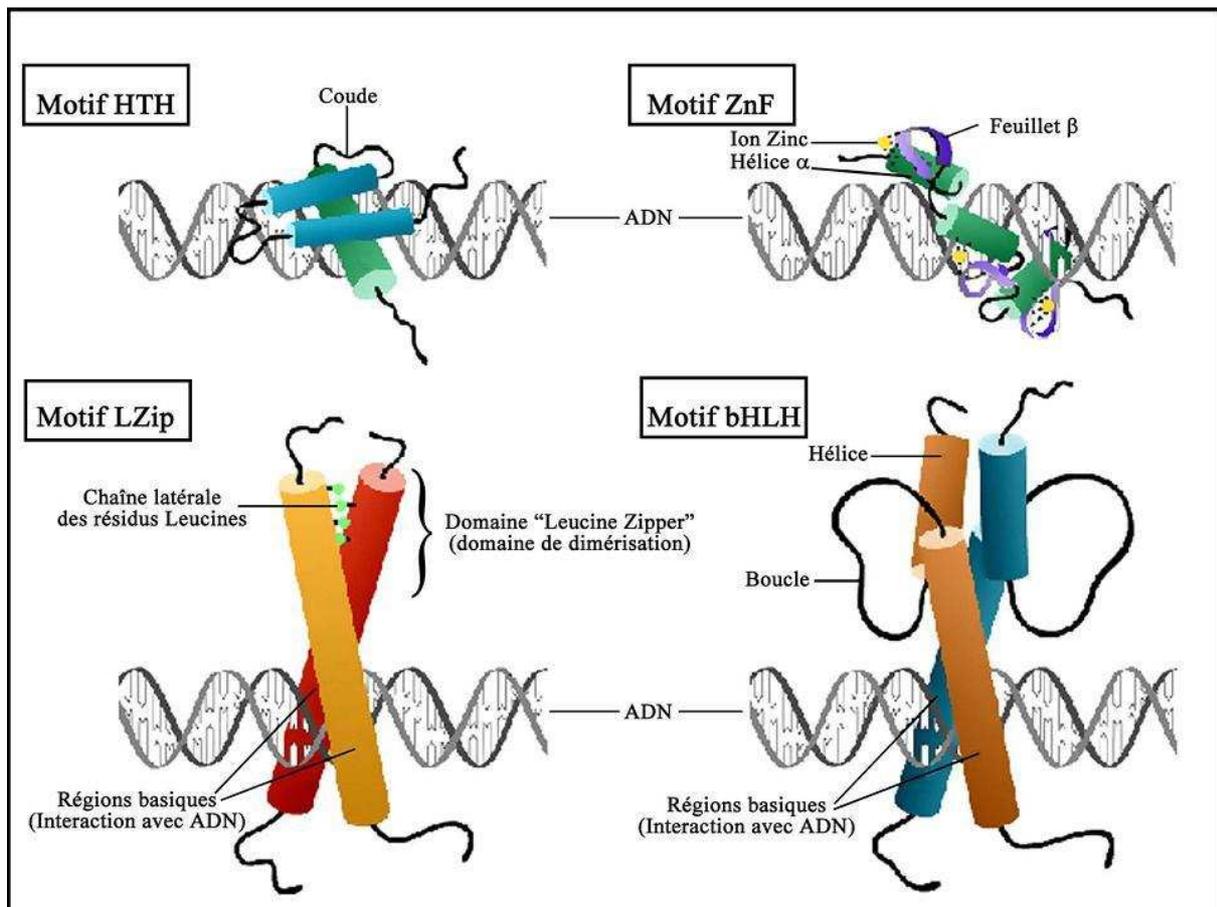


Figure 20 : Principaux domaines de liaison à l'ADN des facteurs de transcription

Représentation schématique de la structure des différentes familles de domaine de liaison à l'ADN des facteurs de transcription et de leur agencement par rapport à l'ADN lors de leur liaison.

HTH : motif de hélice-coude (tour)- hélice

ZnF : motif en doigts de zinc

LZip : motif de fermeture à éclair à leucines

bHLH : motif hélice-boucle-hélice basique (représentation d'un dimère)

cibles vont réguler le niveau de transcription. Les FT présentent des caractéristiques structurales communes, avec au minimum deux domaines : un domaine de fixation à l'ADN et un domaine effecteur de la transcription. Ainsi, ils reconnaissent leurs sites spécifiques et coopèrent avec d'autres protéines pour stimuler ou réprimer la transcription [150]. Les domaines de liaison se replient de façon à présenter une protubérance ou une structure flexible qui entrera en contact avec l'ADN. Ces contacts se font dans le grand sillon de l'ADN, souvent par l'intermédiaire d'une hélice α , avec des liaisons hydrogène et des interactions de Van der Waals.

Les facteurs de transcription sont souvent classés en fonction de la structure de leur domaine de liaison. Il existe 4 types de structures susceptibles d'interagir avec l'ADN (figure 20) [151] :

- structure hélice-tour-hélice (HTH pour Helix-Turn-Helix) présente dans les membres de la famille MYB. Les hélices α se fixent dans le sillon majeur de l'ADN pour former un angle.

- structure en doigt de gant retrouvée dans les protéines zinc finger (pour doigt de zinc). Le pouvoir de lier des atomes de zinc confère à ces protéines une conformation permettant la liaison à l'ADN.

- structure en fermeture éclair par liaison leucine-leucine caractéristique des protéines leucine zippers (bZip). La structure de ces protéines est constituée d'une série de leucines positionnées tous les sept aa autorisant les protéines à se dimériser. Adjacents à ces structures se trouvent plusieurs aa chargés positivement qui sont impliqués dans la fixation à l'ADN.

- structure hélice-boucle-hélice (HLH pour Helix-Loop-Helix) identifiée dans les protéines bHLH (basic Helix-Loop-Helix). Les hélices présentent une face avec des aa hydrophobes et l'autre face avec des aa chargés. Ce motif moléculaire permet la dimérisation entre protéines identiques et la région basique proche de ce motif est en contact avec l'ADN.

2) Aspects généraux sur la régulation de la voie de biosynthèse des flavonoïdes

a. Les éléments *cis*-régulateurs de la voie des flavonoïdes

Les enzymes du métabolisme des phénylpropanoïdes répondent à de nombreux stimuli environnementaux et développementaux. Lors de divers stress de nature biotique (agents pathogènes, blessure, symbiose) ou abiotique (lumière, rayonnements U.V., faible température, carences), les plantes mettent en place un certain nombre de réponses conduisant notamment à l'induction de métabolismes secondaires dont celui des phénylpropanoïdes.

L'induction de la *PAL*, *C4H*, *4CL* et *CHS* au cours des réactions de défense, et suite à d'autres stress biotiques et abiotiques, a été plus particulièrement étudiée. Les enzymes de la voie des flavonoïdes peuvent être induites lors d'une infection par un agent pathogène ou lors d'une surexposition à une source lumineuse [152]. Des études réalisées sur *Arabidopsis* ont montré une augmentation de l'accumulation de flavonoïdes (kaempférol) en réponse aux rayonnements U.V [140]. Des plantes mutées et affectées dans l'expression de la *CHS* et de la *CHI* sont plus sensibles aux rayonnements U.V., ce qui confirme le rôle de ces deux enzymes dans la synthèse de composés intervenant dans la protection des tissus foliaires. Ces voies métaboliques secondaires sont également soumises à une régulation dépendante de l'intensité lumineuse tout au long de la journée. L'expression de la majorité des gènes du métabolisme des phénylpropanoïdes est orchestrée par le cycle circadien. Une étude récente sur l'horloge biologique d'*Arabidopsis*, a montré que la transcription de 6% des gènes étudiés (8000 au total) répond au cycle circadien, et 23 de ces gènes codent pour des enzymes du métabolisme des phénylpropanoïdes [153].

Pour répondre à de nombreux stimuli, plusieurs éléments *cis*-régulateurs activent la transcription des gènes spécifiques de ces signaux. Ces dernières années, les promoteurs des gènes intervenant dans les premières étapes du métabolisme des phénylpropanoïdes ont été plus particulièrement étudiés [140]. Plusieurs séquences *cis*-régulatrices conservées ont été identifiées dans les promoteurs de la *PAL*, *C4H*, *4CL* et *CHS* [81, 154-157]. Trois boîtes très conservées (P, L et A) ont été identifiées dans le promoteur du gène *PAL1* de blé et deux boîtes (H et G) dans le promoteur de la *CHS* du Haricot [158]. Les séquences nucléotidiques des boîtes P, A et L constituent des sites d'interaction avec des protéines en réponse à des éliciteurs fongiques, mais aussi à la lumière blanche et aux rayonnements U.V. [154, 158, 159]. Les séquences des boîtes H et G sont conservées dans les promoteurs de la *PAL* et de la *CHS*. Ces boîtes sont généralement associées à la réponse aux pathogènes et la spécificité tissulaire [160, 161].

En plus des boîtes conservées (P, A, L), d'autres éléments *cis*-régulateurs interviennent dans le signal lumineux. La *CHS* étant impliquée dans la réponse aux rayonnements U.V., son promoteur a été le plus étudié. Chez *Arabidopsis*, la réponse à la lumière implique une région régulatrice appelée LRU (Light Responsive Unit). Deux boîtes sont suffisantes et nécessaires pour induire une activité lumière-dépendante à la *CHS* [162]. Cette unité régulatrice LRU comprend un site de reconnaissance des FT MYB (boîtes MRE^{CHS}) et une boîte contenant une séquence "ACGT" (boîte ACE^{CHS}). Les boîtes ACE sont présentes dans un grand nombre de promoteurs et sont reconnues par des protéines de type bZIP [140]. Des analyses

fonctionnelles menées sur les promoteurs de la *CHS*, *CHI*, *F3H* et *FLS* suggèrent l'implication coordonnée de sites de liaison aux facteurs MYB et aux protéines bZIP dans la régulation de ces gènes par la lumière. De plus, un site de liaison aux facteurs bHLH conférerait une spécificité tissulaire [162].

D'autres éléments *cis* ont pu être identifiés *via* l'étude des facteurs de transcription qu'ils impliquaient. Chez le maïs, les gènes des deux familles de FT, MYB et bHLH, spécifiques de la synthèse des anthocyanes activent fortement l'expression des promoteurs qui possèdent une boîte haPBS (high affinity P1-binding sites). Sans fixation de ces FT, cette boîte n'est pas active [163]. Chez *Arabidopsis*, un fragment de 86 pb dans le promoteur *BAN* agit comme un stimulateur spécifique des PA dans les cellules qui accumulent les tannins [164]. Cette boîte PA contient deux sites de liaison putatifs pour des MYB et des bHLH dont les séquences sont proches des boîtes haPBS. Des structures similaires ont aussi été identifiées dans le promoteur de la *DFR*.

b. Les facteurs de transcription régulateurs de la voie des flavonoïdes

Deux familles de FT sont impliquées dans le contrôle de la synthèse des anthocyanes et des PA : les protéines MYB et bHLH. La production des phlobaphènes n'est sous le contrôle que d'un FT de type MYB [165, 166]. A l'inverse, la production d'anthocyanes et de PA implique que les protéines MYB interagissent avec un ou plusieurs partenaires protéiques pour assurer la régulation transcriptionnelle des gènes cibles [140]. De nombreux complexes MYB/bHLH identifiés chez le maïs, le pétunia et le muflier sont des régulateurs de la synthèse d'anthocyanes [167]. Des études de sur- et sous-expression de gènes chez *Arabidopsis*, couplées à des analyses d'interactions protéine-protéine ont également montré qu'un complexe composé de protéines MYB, bHLH et WD40 (MBW) était à l'origine du contrôle de la transcription des gènes menant à la synthèse des anthocyanes et surtout des PA [162, 168-170].

i- Facteur de transcription MYB

Identité des MYB

« MYB » est un acronyme dérivé du mot MYeloBlastome, décrivant un virus responsable de la leucémie chez le poulet. Le premier gène *MYB* identifié était l'oncogène *v-myb* issu du virus du MyeloBlastome aviaire (AMV) intégré dans le génome du poulet [171, 172]. Introna *et al.* (1994) ont montré que la protéine virale v-MYB était impliquée dans la transformation cancéreuse des cellules hématopoïétiques en dérégulant des gènes importants pour le développement cellulaire [173]. Les FT de type MYB reconnaissent de façon spécifique les

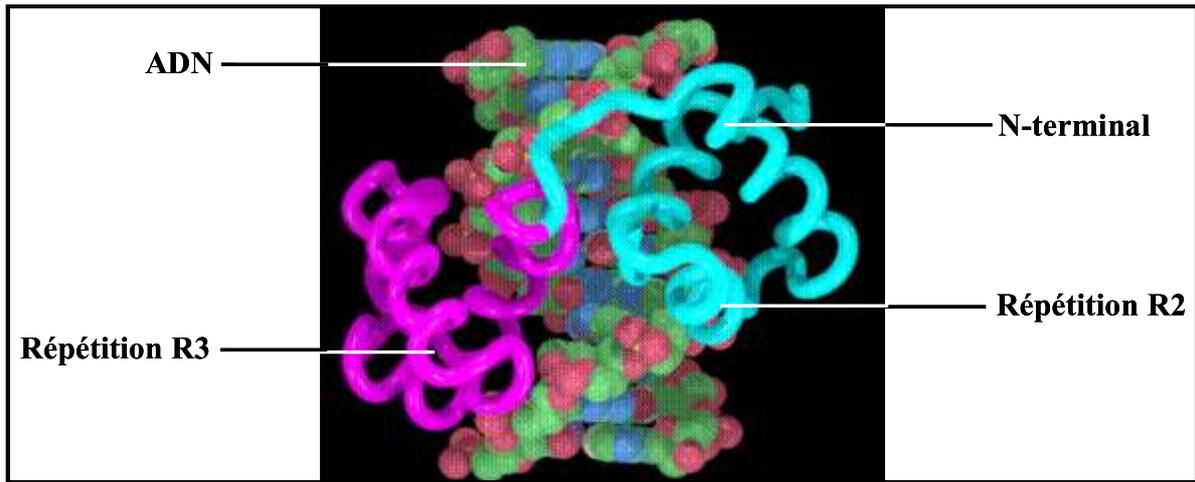


Figure 21 : Structure tridimensionnelle d'une protéine MYB R₂R₃ interagissant avec la molécule d'ADN [126]

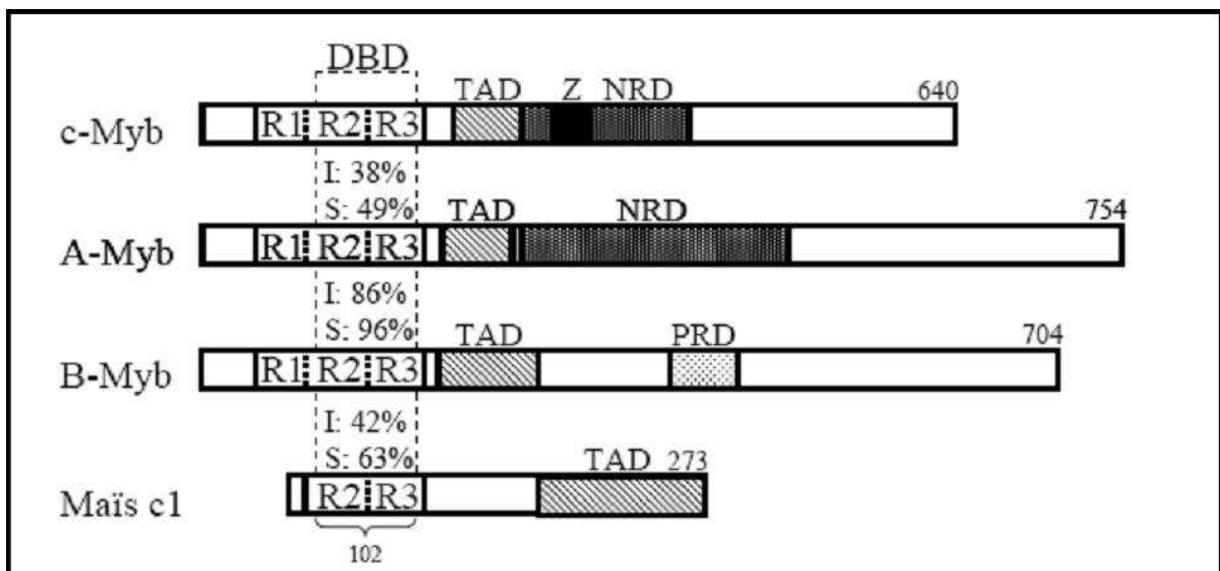


Figure 22 : Représentation des domaines fonctionnels des protéines de type MYB typiques chez les animaux et chez les végétaux [174].

Les différentes répétitions (R) du domaine MYB sont indiquées. DBD: DNA-binding domain, TAD: transcriptional activation domain, NRD: negative regulatory domain, PRD: positive regulatory domain, Z: structure leucine zipper. Les chiffres indiquent le nombre d'acides aminés constituant chaque protéine MYB. Les pourcentages d'identité (I) et de similarité (S) entre les DBD de deux protéines sont calculés avec Bioedit (matrice Blosum62). C-MYB de l'Homme (P10242), A-MYB de *Mus Musculus* (X82327), B-MYB de *Mus Musculus* (X70472) et c1 (colourless 1) de maïs (M37153).

séquences "YAAC(G/T)G" *via* un motif protéique de type hélice-tour-hélice (HTH) [175]. Les protéines MYB se caractérisent par la présence d'un domaine MYB constitué d'environ 50 aa contenant trois tryptophanes conservés et régulièrement espacés de 18 ou 19 aa qui sont à l'origine de la formation d'un cœur hydrophobe nécessaire à la fixation de l'ADN (figure 21) [176]. Plusieurs copies de ce domaine MYB peuvent être répétées dans une seule protéine et permettre l'attachement à l'ADN [177]. Dans les cas de répétitions imparfaites, les tryptophanes peuvent être remplacés par des résidus phénylalanine ou tyrosine et des aa supplémentaires peuvent être observés [178, 179]. La répétition R₁ ne semble pas avoir d'interaction spécifique avec l'ADN, mais elle augmenterait l'affinité de la protéine pour sa séquence cible et la stabilité du complexe protéine-ADN [180-182]. En revanche, les répétitions R₂ et R₃ sont suffisantes pour que la protéine MYB se lie dans le grand sillon de l'ADN de manière séquence-spécifique [183]. Chez les animaux, il n'existe que des protéines MYB à trois répétitions (MYB R₁R₂R₃), alors que chez les plantes et les levures, le nombre de répétitions est variable. De fait, les protéines MYB ont été classées en trois sous-familles suivant le nombre de répétitions : une répétition (MYB R₁), deux répétitions (MYB R₂R₃) ou trois répétitions (MYB R₁R₂R₃) [184].

Chez les animaux, les protéines de type MYB se présentent sous trois formes et on distingue uniquement trois membres chez l'Homme: A-MYB, B-MYB et c-MYB, impliqués dans la division, la différenciation et la mort cellulaire programmée [185]. Elles agissent souvent de concert avec d'autres protéines se liant à l'ADN et des comparaisons de séquences ont permis d'identifier d'autres domaines fonctionnels (figure 22). Parmi ces domaines, ont été identifiés en C-terminal du domaine de liaison à l'ADN, un domaine de transactivation d'une cinquantaine de résidus (hydrophobes et légèrement acides) et un domaine de régulation capable d'inhiber l'activité transcriptionnelle de la protéine. Ce domaine est situé en C-terminal du domaine de transactivation et se compose de deux sous-domaines indépendants (1 et 2) séparés par une structure en forme de leucine zipper (LZ) [180-[186]. Le domaine de fixation des protéines MYB présent en N-terminal est conservé chez la drosophile (D-Myb), la levure (BAS1) ainsi que chez l'homme (A-Myb et B-Myb) les plantes et les levures [187-189]. En revanche, les domaines de transactivation et de régulation négative en C-terminal, capables de moduler l'expression génique, ne sont pas retrouvés dans toutes les protéines MYB [190]. De plus, le domaine de liaison à l'ADN peut se trouver dans la partie C-terminale des protéines comme c'est le cas dans la protéine de levure REB 1 [191].

Les MYB chez les végétaux

Le premier gène *MYB* de plante a été identifié chez le maïs et code la protéine C1 (*colourless1*) [192]. C1 est impliqué dans la synthèse d'anthocyanes dans les aleurones des grains de maïs. Depuis l'identification et la caractérisation de cette protéine MYB à deux répétitions, il apparaît que les FT de type MYB chez les plantes représentent une superfamille de gènes de par la diversité et le nombre élevé de membres. Des études comparatives menées chez *Arabidopsis* et le maïs rapportent respectivement l'identification de 198 et 183 gènes *MYB* [193]. L'analyse complète du génome d'*Arabidopsis* révèle que 126 gènes codent des MYB R₂R₃, 5 des MYB R₁R₂R₃, 64 des "MYB-related" et 3 sont des gènes *MYB* atypiques. Chez *Arabidopsis*, la superfamille des gènes *MYB* est une des classes de FT qui possède le plus grand nombre de membres [194]. Les MYB à deux répétitions sont les plus représentés et ont été classés en 24 sous-groupes [195]. D'autres membres de la famille des gènes *MYB* ont également été identifiés et caractérisés dans d'autres espèces végétales et leur dénombrement n'est pas encore fini (environ 30 membres chez le pétunia, environ 200 chez le cotonnier et 85 chez le riz [196-198]). Par rapport aux gènes MYB de mammifères, ceux de plantes sont plus complexes et diverses, et seulement un petit nombre d'entre eux présentent une expression constitutive et ubiquitaire chez *Arabidopsis*. Ils interviennent dans de nombreux processus physiologiques chez les plantes et peuvent être induits d'une part par des signaux développementaux et, d'autre part, par des stimuli environnementaux [199].

Les fonctions de la plupart des gènes *MYB* de plantes ne sont pas connues. Cependant, les protéines MYB n'ayant qu'une seule répétition (R₁ ou R₃) semblent être impliquées dans le contrôle transcriptionnel des gènes liés aux rythmes circadiens et des gènes impliqués dans le développement de l'endosperme et dans la différenciation des cellules épidermiques [200-205]. Les membres de la famille des MYB R₂R₃ ont révélé que ces protéines participaient au contrôle du métabolisme secondaire, à la régulation de la morphogénèse cellulaire, au développement des organes et à la transduction des signaux en réponses à des stress abiotiques ou à l'attaque de pathogène [206, 207]. Et enfin, les membres du groupe des MYB R₁R₂R₃ sont impliqués dans la régulation du cycle cellulaire [207, 208]. Certaines protéines MYB R₁R₂R₃ interviendraient également dans la tolérance au stress chez le maïs [209].

i- Facteur de transcription bHLH

Identité des bHLH

Les facteurs bHLH se caractérisent par la présence d'un domaine de liaison à l'ADN bipartite d'environ 60 aa [130]. Ce domaine comprend une région composée d'aa basiques et d'une

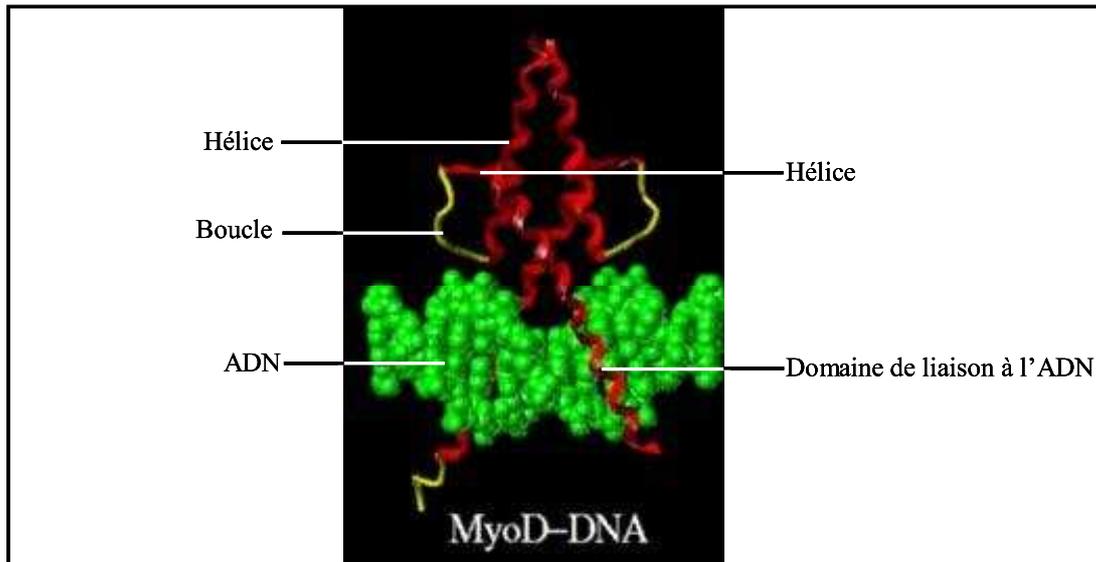


Figure 23 : Structure tridimensionnelle d'une protéine bHLH MyoD interagissant avec la molécule d'ADN.

La photographie est tirée du site web <http://gibk26.bse.kyutech.ac.jp/jouhou/image/dna-protein/hlh/hlh.html>.

région HLH. Le domaine basique est formé d'environ 15 aa situés en N-terminal, capable de se lier à l'ADN au niveau de séquences hexanucléotidiques de type "CANNTG", appelées boîtes E (E-boxes) [210]. Parmi ces boîtes E, la plus commune correspond à la séquence consensus "CACGTG" (boîte G)[211]. Concernant la région HLH, elle contient deux hélices α amphiphatiques séparées par une boucle de longueur variable (figure 23). Les hélices α permettent la dimérisation pour former soit des homodimères avec un autre facteur bHLH soit des hétérodimères avec d'autres familles de FT [131, 212, 213]. La présence d'une variété de boîte E et les possibilités de dimérisation grâce aux régions HLH permettent à ces FT de réguler au niveau transcriptionnel des processus biologiques variés.

Chez les animaux, les protéines bHLH sont impliquées dans la régulation de divers processus développementaux : la neurogenèse, la myogenèse, la détermination du sexe, la prolifération et la différenciation cellulaire [213]. Elles ont été classées *via* des analyses phylogénétiques basées sur les motifs de liaison à l'ADN des bHLH et leurs propriétés fonctionnelles. La famille des bHLH est divisée en 6 groupes répertoriés du groupe A à F [214] :

- les protéines du groupe A peuvent se lier à l'hexanucléotide "CAGCTG" de la boîte E (exemples: Atonal, D, Delilah, dHand, E12, Hen, Lyl, MyoD et Twist)
- les protéines du groupe B ont des fonctions variées et se lient aux séquences "CACGTG" de la boîte G (exemples: protéines Max, Myc, MITF, SREBP et USF)
- les protéines du groupe D possèdent un autre domaine d'interaction protéine-protéine (le domaine PAS) et se lient aux séquences ("NACGTG" ou "NGCGTG")
- les protéines du groupe E possèdent des résidus proline ou glycine dans la région basique et se lient préférentiellement à la séquence "CACGNG" (exemples: E(sp1), Gridlock, Hairy et Hey)
- les protéines du groupe F possèdent un domaine supplémentaire impliqué dans la dimérisation et la liaison à l'ADN

Les bHLH chez les végétaux

Très peu de protéines bHLH ont été caractérisées au niveau fonctionnel chez les plantes. Néanmoins, elles sont impliquées dans la régulation transcriptionnelle de plusieurs processus comme la biosynthèse des anthocyanes, la signalisation liées aux phytochromes, l'expression des globulines, la déhiscence du fruit et le développement des carpelles et de l'épiderme [215]. La première protéine de type bHLH identifiée chez les plantes fût la protéine Lc (Leaf color), produit du gène *R* (*Red*) qui régule la biosynthèse des anthocyanes chez le maïs (tableau I) [216]. Comme la famille des MYB R₂R₃, la famille des bHLH est une superfamille qui compte plus de 162 membres divisés en 21 sous-familles [211, 214, 215,



Figure 24 : Structure tridimensionnelle du domaine WD40 de la sous-unité beta de la protéine G hétérotrimérique.

La structure est composée de feuillets beta formant une hélice de bateau avec 7 pales constituées par les 7 motifs WD40. Chaque pale est représentée par des couleurs différentes. Chaque motif WD40 est constitué de 4 brins.

La photographie est tirée du site web <http://gibk26.bse.kyutech.ac.jp/jouhou/image/dna-protein/hlh/hlh.html>.

217, 218]. La comparaison à la classification animale révèle que la plupart des protéines bHLH de plantes appartiennent au groupe B [214].

Certaines protéines, intervenant dans la régulation du métabolisme des anthocyanes chez le maïs, ont été plus particulièrement étudiées. Ces bHLH, proches des protéines Myc animales, possèdent communément un domaine d'interaction et un domaine d'activation. En position N-terminal, elles possèdent un domaine d'interaction appelé MIR (Myb Interacting Domain). Ce domaine confère aux protéines bHLH la possibilité d'interagir avec les domaines R₂R₃ des protéines MYB [163]. Récemment, il a également été suggéré que le domaine d'interaction coopère avec le domaine d'activation pour permettre une *trans*-activation de la protéine bHLH [219]. Le domaine d'activation, riche en aa acides, forme une plateforme permettant le recrutement de la machinerie de l'ARN polymérase II et l'initiation de la transcription. Un autre domaine d'activation (domaine ACT) a été identifié récemment dans une protéine bHLH impliquée dans le contrôle transcriptionnel de la synthèse d'anthocyanes. Localisé en C-terminal, ce domaine ACT serait une plateforme pour des interactions protéine-protéine. En effet, à la différence des Myc animales qui ne peuvent pas former d'homodimères, ces bHLH de plantes peuvent s'homodimériser pour réguler plus finement la biosynthèse des anthocyanes chez le maïs [220].

b. Les WD40, une autre famille de régulateur de la voie des flavonoïdes

La caractéristique commune qui définit ces protéines est le motif WD (aussi appelé Trp-Asp ou WD40) (figure 24). Ce motif consiste en une région d'environ 40 aa peu conservés mais définie par un dipeptide glycine-histidine (GH) et un dipeptide tryptophane-aspartate (WD) [221]. Ce motif est répété en tandem de 4 à 10 fois dans la même protéine. La protéine WD40 la plus étudiée est la sous-unité β de la protéine hétérotrimérique G participant entre autres à la signalisation cellulaire.

Les protéines à répétitions WD40 sont des acteurs majeurs participant à des mécanismes variés comme la transduction des signaux, la dynamique du cytosquelette, le trafic des protéines, l'export nucléaire ou la maturation des ARN. Elles interviennent également dans les modifications de la chromatine et les mécanismes transcriptionnels. Ces protéines sont également intimement impliquées dans la division cellulaire, la cytokinèse, l'apoptose, la vision et la mobilité cellulaire. Au niveau cellulaire, les protéines WD40 sont des composants du cytoplasme ou du nucléoplasme, reliées au cytosquelette ou associées à la membrane *via* leur interaction avec des protéines membranaires ou avec des domaines auxiliaires

Espèce	Famille de protéine	Protéine		Référence
Maïs	MYB	P	Pericarp color	[224, 225]
	MYB	C1	Colorless1	[192, 226]
	MYB	Pl	Purple plant	[227]
	bHLH	B	Booster	[228]
	bHLH	R	Red	[229, 230]
	bHLH	Lc	Leaf color	[216, 229]
	WD40	PAC1	Pale aleurone color	[231]
Muflier	MYB	ROSEA1		[232]
	MYB	ROSEA2		[232]
	MYB	VENOSA		[232]
	bHLH	DEL	Delila	[233]
	bHLH	MUT	Mutabilis	[233]
Petunia	MYB	AN2	Anthocyanin 2	[234, 235]
	MYB	AN4	Anthocyanin 4	[234, 235]
	bHLH	AN1	Anthocyanin 1	[234, 236]
	bHLH	JAF13		[234]
	WD40	AN11	Anthocyanin 11	[237]
<i>Arabidopsis</i>	MYB	PAP1	Production of anthocyanin pigment 1	[238]
	MYB	PAP2	Production of anthocyanin pigment 2	[238]
	MYB	AtMyb60		[239]
	bHLH	GL3	Glabra3	[240-243]
	bHLH	EGL3	Enhancer of Glabra3	[240, 242, 243]
	bHLH	TT8	Transparent testa 8	[131, 242]
	WD40	TTG1	Transparent testa Glabra3	[244-246]
Vigne	MYB	VvMybA1		[247]
	MYB	VvMybA2		[248]
	MYB	VvMybA3		[248]
	MYB	VvMybA4		[248]
	MYB	VvMyb5a		[249]
	MYB	VvMyb5b		[250]

Tableau I: Protéines MYB, bHLH et WD40 impliquées dans le contrôle de la synthèse des anthocyanes chez les espèces modèles et dans la baie de raisin

L'activité de ces protéines dans la régulation du métabolisme des anthocyanes a été confirmée par la caractérisation de mutants génétiques ou de plantes transgéniques.

d'interaction. Les protéines WD40 connues sont de tailles variables pouvant être supérieure à 400 KDa [222].

Les domaines WD40 facilitent les interactions protéine-protéine et n'ont pas de fonctions enzymatiques intrinsèques. Ce sont des plateformes d'interaction protéique. Ces interactions sont essentielles à l'activité de ces protéines. D'abord, les motifs WD40 fournissent les sites de liaison pour une ou plusieurs protéines et facilitent ainsi des interactions transitoires entre les protéines. C'est le cas des protéines hétérodimériques GTPases où la sous-unité β participe fonctionnellement à l'activité intrinsèque de la protéine du fait de son association avec plusieurs récepteurs membranaires. Les protéines WD40 peuvent également être un membre actif d'un complexe protéique comme par exemple le facteur de transcription général TFIID. Enfin, les WD40 se comportent comme des domaines d'interactions modulaires dans les protéines les plus grosses. Dans ce dernier cas, les protéines WD40 permettent d'adresser les protéines associées aux domaines auxiliaires vers leur cible. Ce mode d'action a été détaillé dans les protéines intervenant dans la signalisation lumineuse comme COP1 [223].

Chez *Arabidopsis*, 237 protéines contenant 4 copies du motif WD40 ou plus, ont été identifiées. Une classification fondée sur leur structure et leur fonction a permis de distinguer 143 familles WD40, dont 49 contiennent plus d'une répétition WD. Approximativement 113 familles présentent une homologie avec les protéines WD40 identifiées chez les autres eucaryotes. La caractérisation fonctionnelle chez *Arabidopsis* des protéines WD40 conservées a révélé qu'elles interviennent dans les mécanismes fondamentaux des processus spécifiques des plantes. En effet, des membres de la famille WD40 sont impliqués dans la floraison, le développement floral, l'organisation méristématique et la signalisation lumineuse [222].

1) Gènes impliqués dans la régulation des gènes de biosynthèse des anthocyanes

Le tableau I présente les gènes *MYB*, *bHLH* et *WD40*, identifiés ces dernières années chez différentes espèces végétales, en tant que régulateur de la voie de biosynthèse des anthocyanes.

Régulation de la biosynthèse des anthocyanes chez les espèces modèles

Activateurs de la biosynthèse des anthocyanes

Chez le maïs, les protéines R (Red, bHLH) et C1 (Colourless 1, MYB), induisent de façon coordonnée les gènes structuraux menant à la synthèse d'anthocyanes dans les couches d'aleurones des grains. L'expression ectopique de *R* et *C1* dans des cultures cellulaires non-pigmentées de maïs entraîne l'activation de la plupart des gènes de biosynthèse et

l'accumulation des anthocyanes [165, 166]. Dans les autres parties de la graine, ce sont d'autres membres des deux familles (Pl, Purple plant et B, Booster) qui assurent le contrôle transcriptionnel [142, 216]. Ainsi, la régulation fine des gènes de cette voie par différents complexes MYB-bHLH permet une accumulation tissu-spécifique des anthocyanes.

Chez le mufler, les mutants *del* (Delila, bHLH) sont caractérisés par une perte de pigmentation dans les tubes floraux [233]. Cette perte de fonction diminue faiblement l'expression des *EBG* (*CHS* et *CHI*) mais inhibe fortement les gènes *F3H*, *DFR*, *LDOX* et *UFGT* [179]. D'autres gènes *MYB* et *bHLH* ont été identifiés et leurs actions combinées assurent le contrôle complexe de la production spatiale et temporelle des anthocyanes dans les pétales de mufler [251, 252]. Cependant à ce jour, aucune interaction entre ces MYB et ces bHLH n'a pu être démontrée [79].

Chez le Pétunia, le système de régulation semble plus complexe. Les gènes *AN1* (Anthocyanin 1, bHLH), *JAF13* (bHLH), *AN2* (Anthocyanin 2, MYB), *AN4* (Anthocyanin 4, MYB) et *AN11* (WD40) sont tous nécessaires pour activer l'expression des *LBG* [234, 236, 253, 254]. En dehors du fait que *JAF13* et *AN1* présentent des séquences en acides aminés différentes, ces protéines possèdent des fonctions divergentes dans la cascade de régulation. *JAF13* ne peut compenser la perte d'expression d'*AN1* dans les pétales des mutants *an1*. De plus, à l'inverse de *JAF13*, *AN1* doit interagir physiquement avec *AN2* pour activer les gènes de biosynthèse des anthocyanes. Pourtant, des expériences d'expression transitoire ont montré que *AN1* et *JAF13* (bHLH) peuvent former un complexe avec *AN2* et induire l'activité du promoteur *DFR* [253]. De plus, en système double hybride chez la levure, *AN1* et *JAF13* peuvent former aussi bien des homodimères que des hétérodimères. Afin de mieux comprendre la fonction de *JAF13*, un mutant contenant un transposon en amont du gène *JAF13* a été isolé par criblage de mutants d'insertion. Le mutant homozygote présente une réduction des transcrits *DFR*, de la teneur en anthocyanes dans les pétales d'environ 50% et des tâches "révertantes". Le phénotype de ce mutant a permis d'émettre deux hypothèses: soit la fonction de *JAF13* est partiellement redondante, soit sa présence dans le complexe protéique n'est pas nécessaire mais amplifierait l'activité du complexe MYB-bHLH [255]. *AN11* fut la première protéine de type WD40 à avoir été identifiée en tant que partenaire protéique dans le complexe MYB-bHLH [237]. Des mutations dans le gène *AN11* entraînent une perte de pigmentation des fleurs mais n'affectent pas l'expression des gènes régulateurs de la voie. De plus, ces mutants ne peuvent être que partiellement complétés par la surexpression des gènes *AN2* et *AN1*. De vetten et al. ont alors suggéré que *AN11* modulerait l'activité des facteurs MYB et bHLH spécifiques de cette voie, par des

modifications post-transcriptionnelles. Depuis, des orthologues d'*AN11* ont été identifiés chez le maïs (*PAC1*), *Arabidopsis* (*TTG1*) et *Perilla frutescens* (*PFWD*), mais la fonction de ces protéines semble être plus pléiotropique que celle d'*AN11*, qui est restreinte à la pigmentation des fleurs [231, 244, 256].

Chez *Arabidopsis*, les protéines MYB (*PAP1*, *PAP2*, *Myb113* et *Myb114*), bHLH (*EGL3* et *GL3*) et WD40 (*TTG1*) régulent la synthèse des anthocyanes [195, 238, 257, 258]. L'action des facteurs MYB est strictement *TTG1*- et bHLH-dépendante [170]. Des simple et double mutants montrent qu'*EGL3* jouerait un rôle plus important que *GL3* dans le contrôle des anthocyanes [242]. La surexpression indépendante de chacun des MYB stimule l'accumulation d'anthocyanes suite à l'activation par ces FT des gènes *LBG*. De plus, les plantes transformées avec une construction RNAi pour les gènes *PAP1*, *PAP2*, *Myb113* et *Myb114* présentent une diminution de l'expression des gènes *LBG* et consécutivement, une diminution de l'accumulation des anthocyanes mais pas des PA. D'ailleurs, les gènes structuraux *LBG* affectés dans ces plantes RNAi sont les mêmes que ceux qui ont été identifiés dans les mutants *ttg1*, *egl3* et *gl3*.

Répresseurs de la biosynthèse des anthocyanes

Quelques protéines MYB ont été caractérisées comme des répresseurs de certaines branches du métabolisme des phénylpropanoïdes [259]. Seulement deux répresseurs du métabolisme des anthocyanes ont été identifiés : *FaMYB1* chez la fraise et *AtMyb60* chez *Arabidopsis*. Les tabacs qui surexpriment *FaMYB1* présentent une diminution significative de la teneur en anthocyanes dans les fleurs et les étamines ainsi qu'une réduction de la teneur en flavonoïdes. Ce phénotype est la conséquence d'une diminution de l'expression de deux gènes de la partie basse de la voie : l'*ANS* et la *GT*. *FaMYB*, comme les répresseurs des phénylpropanoïdes, présente dans sa région C-terminale le motif conservé LNL[E/D]L [260]. Récemment chez *Arabidopsis*, des travaux ont montré que l'expression ectopique de la protéine *AtMYB60* inhibe l'accumulation d'anthocyanes chez la laitue en réprimant l'expression des gènes codant la DFR [239].

Régulation de la biosynthèse des anthocyanes dans la baie de raisin

Au niveau biochimique, les raisins rouges diffèrent des raisins blancs par l'accumulation d'anthocyanes après véraison dans les pellicules. Au niveau moléculaire, cela se traduit par l'expression de l'*UFGT* après la véraison, dans les pellicules des baies des cépages rouges mais pas dans celles des cépages blancs [133]. Au niveau génétique, ce phénotype rouge/blanc est contrôlé par un seul locus [261]. Une cartographie génétique a permis de

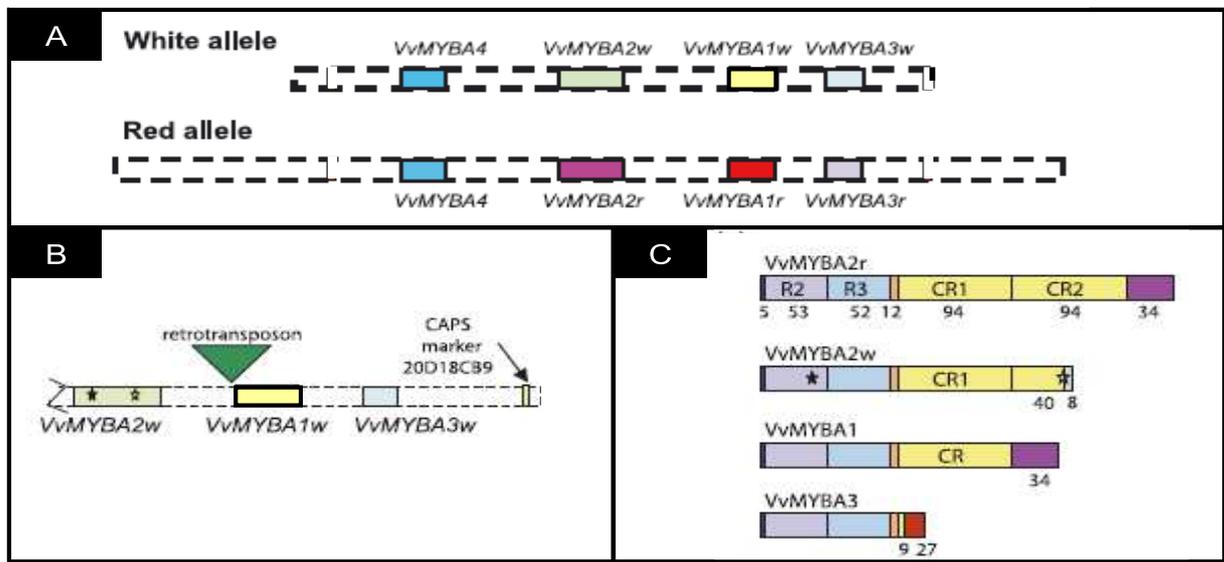


Figure 25 : Représentation schématique du locus qui détermine la couleur des baies de Cabernet sauvignon [248]

A- Séquences des allèles blancs et rouges obtenues par des banques BAC. Les séquences des gènes *VvMyba* sont représentées le long du chromosome par des boîtes de couleurs différentes disposées dans l'ordre. Les boîtes vertes et jaunes représentent l'allèle blanc, les boîtes rouges et pourpres indiquent les allèles rouges polymorphes et les boîtes bleues identifient les séquences identiques entre les deux allèles.

B- Région de 140 Kb du BAC 20D18 qui contient 3 gènes très similaires, *VvMyba2w*, *VvMyba1w* et *VvMyba3w*. L'ordre des gènes sur le chromosome est conservé. Les étoiles représentent des changements de séquences non-conservés identifiés dans *VvMyba2w* par rapport à *VvMyba2r*.

C- Différences dans les séquences protéiques prédites des *VvMyba*. Le nombre d'acides aminés dans chaque domaine est indiqué. Les régions identiques sont présentées par les mêmes couleurs. *VvMyba2* existe sous deux formes: *VvMyba2r* est la protéine fonctionnelle identifiée dans les raisins rouges et *VvMyba2w* (codé par l'allèle blanc) est altéré par une délétion indiquée par l'étoile. Les répétitions MYB sont identiques pour les 4 protéines à l'exception d'un changement dans *VvMyba2w* indiqué par une étoile. Le domaine C-terminal est répété dans *VvMyba2r* (CR1 et CR2).

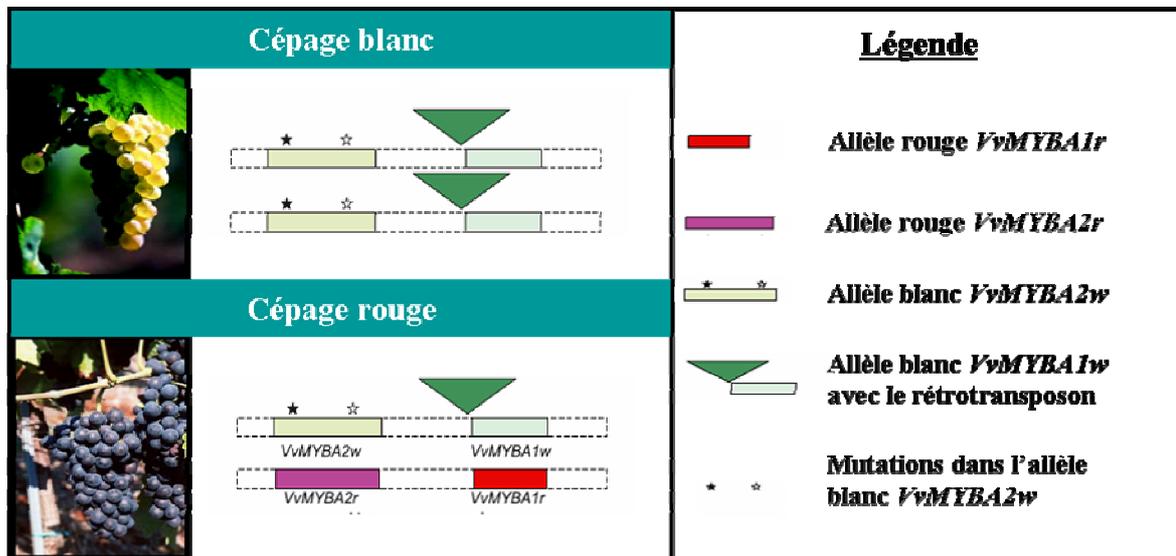


Figure 26 : Différences génétiques à l'origine des cépages rouges et blancs (d'après [248]).

Les cépages blancs sont tous homozygotes pour les allèles blancs *VvMyba1w* et *VvMyba2w*. Les cépages rouges sont hétérozygotes et possèdent les allèles rouges *VvMyba1r* et *VvMyba2r* et les allèles blancs. Les protéines *VvMyba1r* et *VvMyba2r* activent transcriptionnellement l'expression de l'*UFGT* à la véraison, enzyme clé de la synthèse des anthocyanes.

localiser quatre gènes MYB (*VvMYBA1*, *VvMYBA2*, *VvMYBA3* et *VvMYBA4*) (figures 25A et 26)[144, 248]. Deux gènes, *VvMYBA1* et *VvMYBA2*, peuvent réguler la couleur dans la baie de raisin (figure 25B). En effet, *VvMYBA1* et *VvMYBA2*, sont systématiquement mutés dans les raisins blancs [247, 248]. Le gène *VvMYBA1* présente deux allèles : le muté *VvMYBA1a* et le non muté *VvMYBA1b*. *VvMYBA1a* n'est pas transcrit dans les cépages blancs du fait de l'insertion d'un rétrotransposon *Gret1* dans son promoteur (figure 25C) [247]. De même, le gène *VvMYBA2* possède deux allèles : l'allèle blanc *VvMYBA2w* et l'allèle rouge *VvMYBA2r* (figures 25C et 26) [248]. L'allèle blanc *VvMYBA2w* dans les raisins blancs est inactivé par deux mutations non conservées dans sa séquence codante. La première mutation mène à la substitution d'un résidu arginine en résidu leucine (R⁴⁴L) et l'autre entraîne un changement de cadre de lecture à l'origine d'une protéine tronquée à cause de la délétion des deux nucléotides CA en position 258. Chacune des mutations rend inefficace le régulateur *VvMYBA2* et de fait, empêche l'accumulation d'anthocyanes. Tous les cépages blancs sont donc homozygotes pour *VvMybA1a* et pour *VvMYBA2w*. A l'inverse, les cépages rouges sont hétérozygotes et possèdent un allèle blanc (*VvMybA1a* et *VvMYBA2w*) et un allèle rouge non muté (*VvMybA1b* et *VvMYBA2r*) (figure 26).

Deux autres gènes *MYB*, *VvMyb5a* et *VvMyb5b*, régulent également la synthèse des anthocyanes, mais de manière moins spécifique que *VvMYBA1* et *VvMYBA2* [249, 250]. La surexpression de *VvMyb5a* et *VvMYB5b* dans le tabac affecte le métabolisme des anthocyanes, des flavonols, des lignines et des PA. Ceci suggère que ces gènes réguleraient les différentes branches de la voie des phénylpropanoïdes et ne seraient pas spécifique d'une voie en particulier.

2) Gènes impliqués dans la régulation des gènes de biosynthèse des PA

Le tableau II présente les gènes *MYB*, *bHLH* et *WD40*, identifiés ces dernières années dans les différentes espèces végétales, en tant que régulateur de la voie de biosynthèse des PA.

Régulation de la biosynthèse des PA chez *Arabidopsis*

Chez *Arabidopsis*, le gène *BANYULS* (*BAN*) qui code une ANR est directement impliqué dans la biosynthèse des tannins. L'analyse fonctionnelle de son promoteur a permis de démontrer que la spécificité d'expression de ce gène dans les cellules productrices de tannins situées au niveau des téguments est contrôlée principalement au niveau transcriptionnel par le complexe MBW, composé des protéines régulatrices TT2 (Transparent testa 2, MYB), TT8 (bHLH) et TTG1 (Transparent testa Glabra3, WD40) [131, 168, 244, 263]. Le mécanisme de

Espèce	Famille de protéine	Protéine		Référence
Maïs	bHLH WD40	Lc PAC1	Leaf color Pale aleurone color	[216, 229] [231]
<i>Arabidopsis</i>	MYB	TT2	Transparent testa 2	[263]
	MYB	PAP1	Production of anthocyanin pigment 1	[238]
	MYB	PAP2	Production of anthocyanin pigment 2	[238]
	bHLH WD40	TT8 TTG1	Transparent testa 8 Transparent testa Glabra3	[131, 242] [244-246]
Vigne	MYB	VvMybPA1		[262]

Tableau II: Protéines MYB, bHLH et WD40 impliquées dans le contrôle de la biosynthèse des proanthocyanidines chez les espèces modèles et dans la baie de raisin.

L'activité de ces protéines dans la régulation du métabolisme des anthocyanes a été confirmée par la caractérisation de mutants génétiques ou de plantes transgéniques.

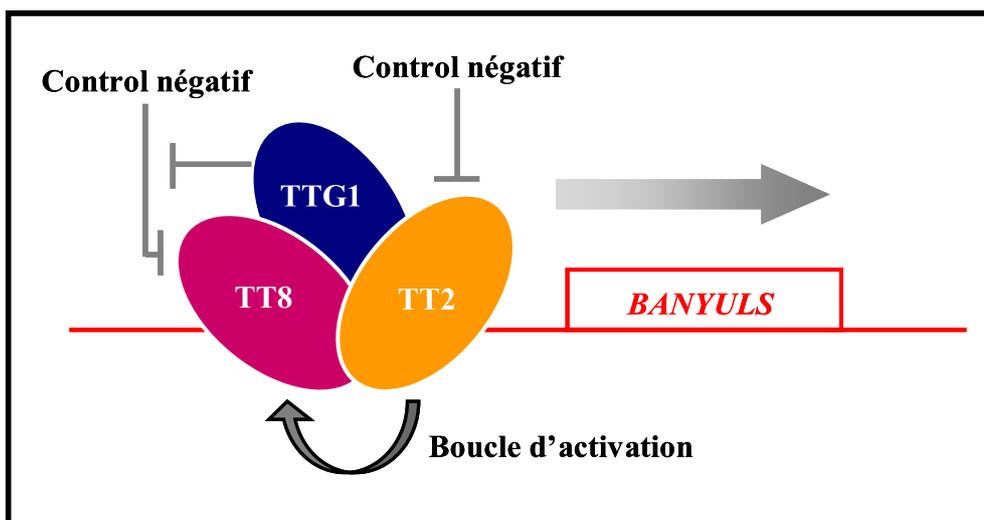


Figure 27 : Représentation schématique de l'activité du complexe MYB-bHLH-WD40 impliqué dans la biosynthèse des proanthocyanidines dans les graines d'*Arabidopsis* d'après [69]

TT2 (MYB), TT8 (bHLH) et TTG1(WD40) forme un complexe transcriptionnel capable d'activer directement le promoteur *BAN*. TT2 est responsable de la reconnaissance spécifique du promoteur *BAN*, en combinaison avec TT8. TTG1 régule l'activité de ces protéines *in planta*, en agissant sur la stabilité de TT8. La spécificité et les niveaux d'expression de *BAN* résulteraient d'un équilibre entre un contrôle développemental de l'expression de *TT2*, *TT8* et *TTG1* et d'un régulateur post-transcriptionnel négatif.

contrôle est similaire à celui décrit pour la régulation de la synthèse des anthocyanes chez le pétunia [163] (figure 27).

TT2 reconnaît spécifiquement l'ADN cible et permet la fixation du complexe TT2/TT8 sur le promoteur de *BAN* [69]. TT8 peut être remplacée par EGL3 et GL3 pour activer la transcription de *BAN*, mais l'activation transcriptionnelle sera moins forte. L'interaction de TTG1 avec le complexe TT2/TT8 permettrait alors d'augmenter son activité transcriptionnelle. De plus, TTG1 serait également impliqué dans la régulation de l'activité de TT8 par un mécanisme post-transcriptionnel, *via* une interaction directe avec cette dernière [168]. Enfin, des analyses récentes ont montré que le gène *TT8* est régulé par ce complexe protéique MBW *via* une boucle de rétroaction [264]. Cette boucle d'autoactivation serait initialement contrôlée par deux FT MYB (TT2 et PAP1) et par TTG1. Ce contrôle transcriptionnel permettrait de synchroniser l'expression de *TT8* avec celle de ces partenaires pour réguler plus finement l'accumulation des PA dans les graines d'*Arabidopsis*.

Régulation de la biosynthèse des PA dans la baie de raisin

Dans la baie de raisin, la synthèse des PA implique deux enzymes, l'ANR et la LAR [126]. Un seul facteur de transcription MYB, *VvMYBPA1*, a été caractérisé et régulerait le métabolisme des PA [262]. L'expression de *VvMYBPA1* est corrélée à l'accumulation des PA au cours du développement de la baie. En expression transitoire, *VvMYBPA1* active les promoteurs de la *LAR* et *ANR* ainsi que ceux de la *CHI*, la *F3'5'H* et la *LDOX*. En revanche, *VvMYBPA1* n'est pas capable d'activer le promoteur de l'*UFGT*. De plus, *VvMYBPA1* complémente le mutant *tt2* d'*Arabidopsis* et restaure l'accumulation des PA dans les cotylédons, les méristèmes végétatifs, les poils racinaires et les racines.

3) Gènes impliqués dans la régulation des gènes de biosynthèse des flavonols

Chez *Arabidopsis*, les protéines MYB appartenant au sous-groupe 7 sont impliquées dans le contrôle transcriptionnel du métabolisme des flavonols [195, 265, 266]. Ces trois protéines présentent des homologies de séquence avec le facteur de transcription MYB (P) qui régule la biosynthèse des phlobaphènes chez le maïs sans interagir avec un partenaire bHLH [163]. MYB11, MYB12 et MYB111 activent les gènes structuraux menant à la synthèse des flavonols, à savoir la *CHS*, la *CHI*, la *F3H* et la *FLS*. Les plantules des triples mutants *myb11*, *myb12* et *myb111* n'accumulent pas de flavonols et la synthèse des anthocyanes n'est pas affectée. Ces protéines agiraient de manière synergique pour réguler la synthèse spatio-temporelle des flavonols chez *Arabidopsis*. En effet, elles présentent des activités

spatiales différentes. MYB12 contrôle principalement la biosynthèse des flavonols dans les racines alors que MYB111 plutôt au niveau des cotylédons [265].

Chapitre 1

Caractérisation des mécanismes régulateurs de l'expression du gène *VvMyb5a* et de l'activité de la protéine correspondante

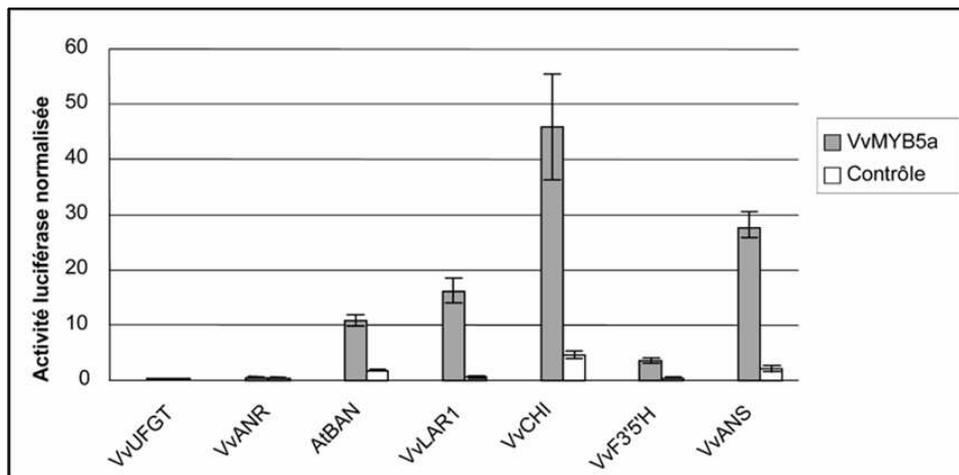


Figure 28 : Activation des promoteurs des gènes codant les enzymes de la voie de biosynthèse des flavonoïdes chez la vigne par *VvMyb5a* [250].

Le contrôle indique l'activité de chaque promoteur en l'absence de *VvMyb5a*. Chaque transfection est réalisée avec une construction *35S::EGL3* codant la protéine bHLH EGL3 d'*Arabidopsis thaliana* (numéro d'accèsion Genbank : NM20235) et un contrôle interne, le plasmide de la luciférase *Renilla* (pRLuc). L'activité normalisée de la luciférase correspond au rapport entre l'activité luciférase de la luciole et de *Renilla*. Chaque colonne représente la valeur moyenne de trois répétitions indépendantes. Les abréviations sont les suivantes: *VvUFGT* (*Vitis vinifera* UDP-glucose : flavonoïde 3-O-glucosyltransférase), *VvANR* (*Vitis vinifera* Anthocyanidin réductase), *AtBAN* (*Arabidopsis thaliana* Banyuls), *VvLAR1* (*Vitis vinifera* Leucoanthocyanidin réductase 1), *VvCHI* (*Vitis vinifera* Chalcone isomérase), *VvF3'5'H* (*Vitis vinifera* Flavonoïde 3'5' hydroxylase) et *VvANS* (*Vitis vinifera* Anthocyanidin synthase).

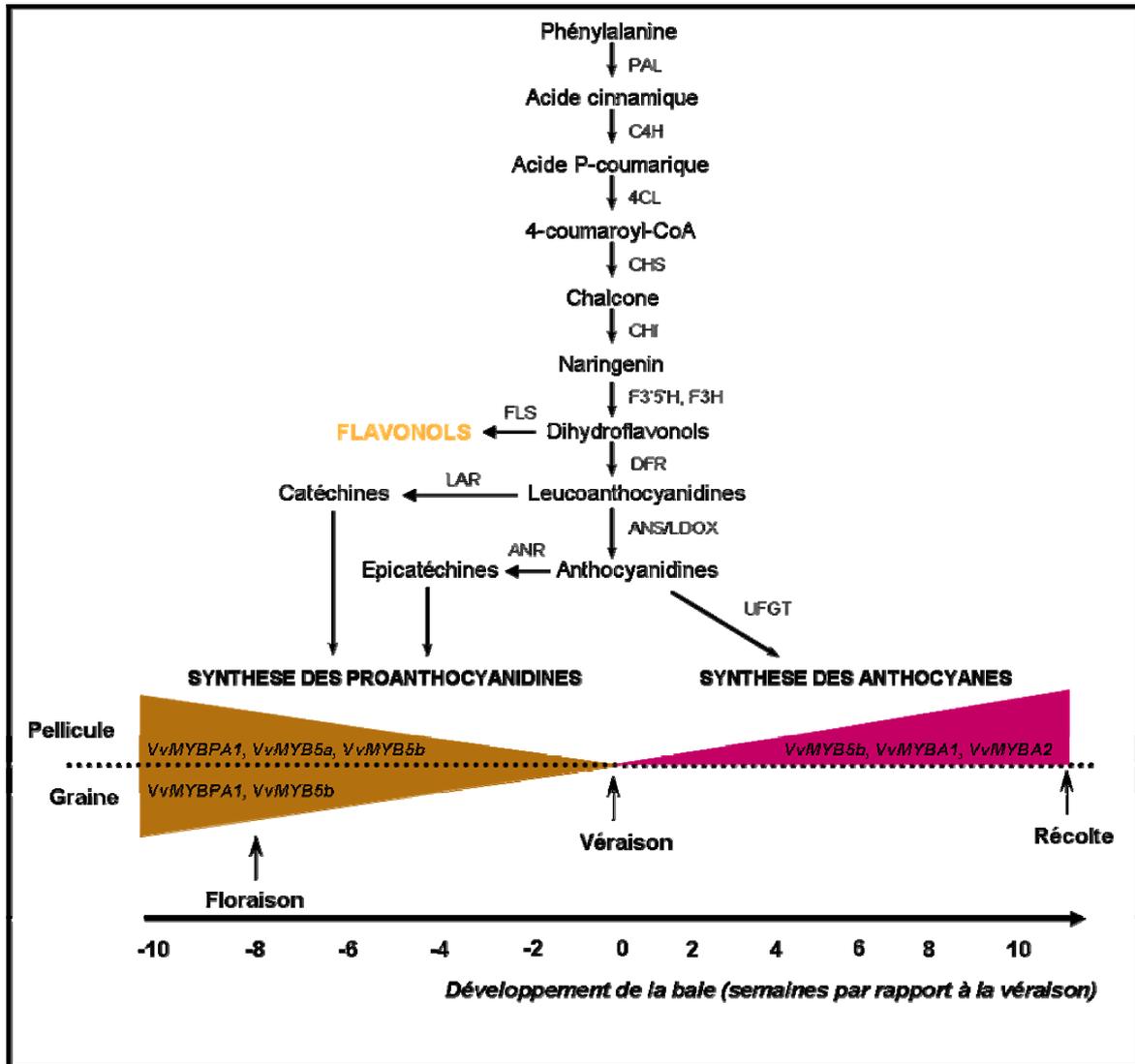


Figure 29: Implication des facteurs de transcriptions MYB R_2R_3 dans les mécanismes régulateurs de la biosynthèse des flavonoïdes au cours du développement de la baie de raisin. [250]

Avant la véraison, les facteurs MYB VvMybPA1, VvMyb5a et VvMyb5b apparaissent particulièrement impliqués dans la régulation de la synthèse des PA dans la pellicule et les pépins. A la véraison, la synthèse des PA se termine et celle des anthocyanes débute dans la pellicule. L'expression de l'*UFGT* est alors spécifiquement régulée par les facteurs de transcription VvMybA1 et VvMybA2 alors que les gènes codant les enzymes de la voie générale de biosynthèse des flavonoïdes seraient contrôlés au moins en partie par VvMyb5b.

Les abréviations sont les suivantes : PAL (phénylalanine ammonia lyase), C4H (cinnamate 4-hydroxylase), 4CL (4-coumarate CoA ligase), CHS (chalcone synthase), CHI (chalcone isomérase), F3'H (flavanone 3-hydroxylase), F3'5'H (3' hydroxylase), F3'5'H (flavonoïde 3'5' hydroxylase), DFR (dihydroflavonol), ANS (anthocyanidine synthase), LDOX (leucoanthocyanidine dioxygénase), ANR (anthocyanidine réductase), LAR (leucoanthocyanidine réductase) et UDP (UDP-glucose:flavonoïde 3-O-glucosyltransferase).

I- Introduction : Etat de l'art sur *VvMyb5a*

Lors de mon arrivée au laboratoire, des travaux antérieurs avaient permis d'identifier et de caractériser le gène *VvMyb5a*, qui code un facteur de transcription MYB de type R₂R₃ [249]. Dans les baies, *VvMyb5a* est fortement exprimé dans les phases précoces du développement. À partir de la véraison, la quantité de transcrits *VvMyb5a* diminue fortement dans les différents tissus de la baie. Pour identifier les gènes cibles de *VvMyb5a*, la capacité de ce facteur de transcription à activer les promoteurs des gènes codant les enzymes de la voie de biosynthèse des flavonoïdes a été étudiée par des expériences d'expression transitoire dans des cellules de vigne [250]. Les résultats obtenus sont présentés dans la figure 28. En présence d'une protéine bHLH, *VvMyb5a* est capable d'activer les promoteurs des gènes *VvLARI* (27 fois), *VvANS* (12,5 fois), *VvF3'5'H* (12 fois) et *VvCHI* (7 fois). En revanche, aucune activation du promoteur *VvUFGT* n'a été mise en évidence. Ainsi, ces données indiquent que *VvMyb5a* est capable d'activer non seulement l'expression des gènes codant les enzymes de la voie générale de biosynthèse des flavonoïdes mais également celle du gène *VvLARI*, spécifiquement impliqué dans la synthèse des tannins. Dans leur ensemble, ces résultats confirment ceux obtenus chez le tabac où la surexpression de *VvMyb5a* s'accompagne d'une accumulation d'anthocyanes et de tannins condensés dans les organes reproducteurs [249]. Ainsi, en tenant compte de son profil d'expression dans les différents tissus de la baie et des résultats de caractérisation fonctionnelle, *VvMyb5a* serait particulièrement impliqué dans la régulation de la biosynthèse des PA et plus spécifiquement dans le contrôle de la synthèse des catéchines dans la pellicule. Ces travaux ont donc permis d'attribuer une fonction précise à *VvMyb5a* et d'intégrer son action au schéma général de régulation du métabolisme des flavonoïdes dans la baie par les facteurs MYB (figure 29). Toutefois, les résultats obtenus chez plusieurs espèces modèles indiquent que l'activité biologique des facteurs MYB apparaît fortement dépendante des interactions avec d'autres protéines de types bHLH, WD40 ou EMSY. À l'heure actuelle, ces protéines n'ont pas encore été identifiées et caractérisées chez la vigne.

Dans ce premier chapitre, nous rendons compte des travaux engagés pour approfondir les connaissances relatives à l'implication de *VvMyb5a* dans la biosynthèse des flavonoïdes dans la baie de raisin. Ce chapitre fait la synthèse de deux approches menées en parallèle : l'étude de la régulation de l'expression du gène *VvMyb5a* et celle de l'activité de la protéine correspondante.

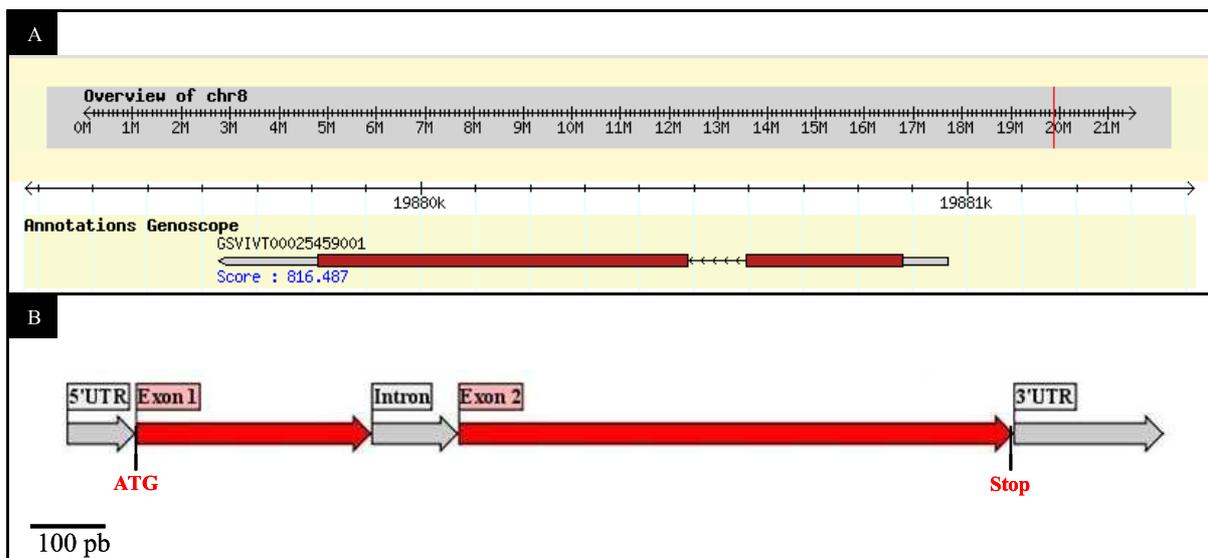


Figure 30: Analyse bioinformatique du locus *VvMyb5a*.

A) Localisation du transcrit *VvMyb5a* sur le chromosome 8.

B) Représentation schématique de la structure intron/exon du gène *VvMyb5a*. Le gène *VvMyb5a* est constitué de deux exons (rouge) et un intron (gris). Les régions non-codantes en 3' et 5' sont également indiqués en gris. Les codons d'initiation (ATG) et de terminaison (Stop) de la traduction sont représentés par des traits noirs. Cette représentation schématique a été réalisée à partir des données disponibles de la base du Génoscope. Echelle, 100 paires de bases (pb).

II- Recherche de partenaires protéiques de la protéine VvMyb5a

1) Analyse *in silico* de la séquence VvMyb5a

Suite à une collaboration entre la France (Inra, Génoscope) et l'Italie (Universités de Milan, Udine et Padoue), le séquençage du génome de la vigne a été achevé et rendu public le 27 août 2007. L'alignement de la séquence de *VvMyb5a* sur l'ensemble des séquences génomiques de vigne montre que ce gène est porté par le chromosome 8 entre les positions 19879628 et 19880963 (figure 30A). *VvMyb5a* est répertorié sous la référence *GSVIVT00025459001* dans la banque de données Génoscope. L'analyse de la séquence nucléique a montré que le gène *VvMyb5a* est composé de deux exons (287 pb et 676 pb) et d'un intron (106 pb) (figure 30B). Elle a également confirmé que *VvMyb5a* présente une région 5' non codante de 84 pb et une région 3' non codante de 186 pb. La partie codante révèle un cadre de lecture ouvert de 963 pb qui code une protéine de 321 aa. Le logiciel ProtParam¹ prédit une masse moléculaire de 35,8 kDa et un point isoélectrique de 6,90.

La recherche de motifs et de domaines signatures dans *VvMyb5a* a été effectuée grâce aux logiciels InterProscan² et ScanProsite³. Ces deux programmes ont mis en évidence deux motifs hélice-tour-hélice qui définissent le domaine MYB (InterProScan001005, PFOO249), un domaine de liaison à l'ATP et au GTP (Prosite00017) et un domaine riche en glutamine (Prosite50322). *VvMyb5a*, comme la majorité des facteurs MYB de plante, contient seulement deux répétitions imparfaites R₂R₃ proches de l'extrémité N-terminale et correspondant au domaine DBD (DNA Binding Domain). Le DBD permet la liaison du facteur MYB aux séquences d'ADN cibles. Cette analyse a également révélé différents sites potentiels de modifications post-traductionnelles: trois signatures de myristoylation (débutant en position aa 53, 212, 216), sept sites de phosphorylation par la protéine kinase C (débutant en position aa 9, 60, 72, 122, 134, 165, 294), six sites de phosphorylation par la caséine kinase II (débutant en position aa 26, 109, 173, 223, 224, 297) et un site de phosphorylation par la tyrosine kinase (débutant en position aa 108).

Des séquences protéiques correspondantes à des MYB homologues de *VvMyb5a* ont été recherchées chez d'autres plantes, dans les banques de données du TAIR⁴ (The Arabidopsis

¹ ProtParam: <http://expasy.org/tools/protparam.html>

² InterProscan: <http://www.ebi.ac.uk/InterProScan>

³ ScanProsite : <http://www.expasy.org/tools/scanprosite/>

⁴ TAIR: <http://www.arabidopsis.org/index.jsp>

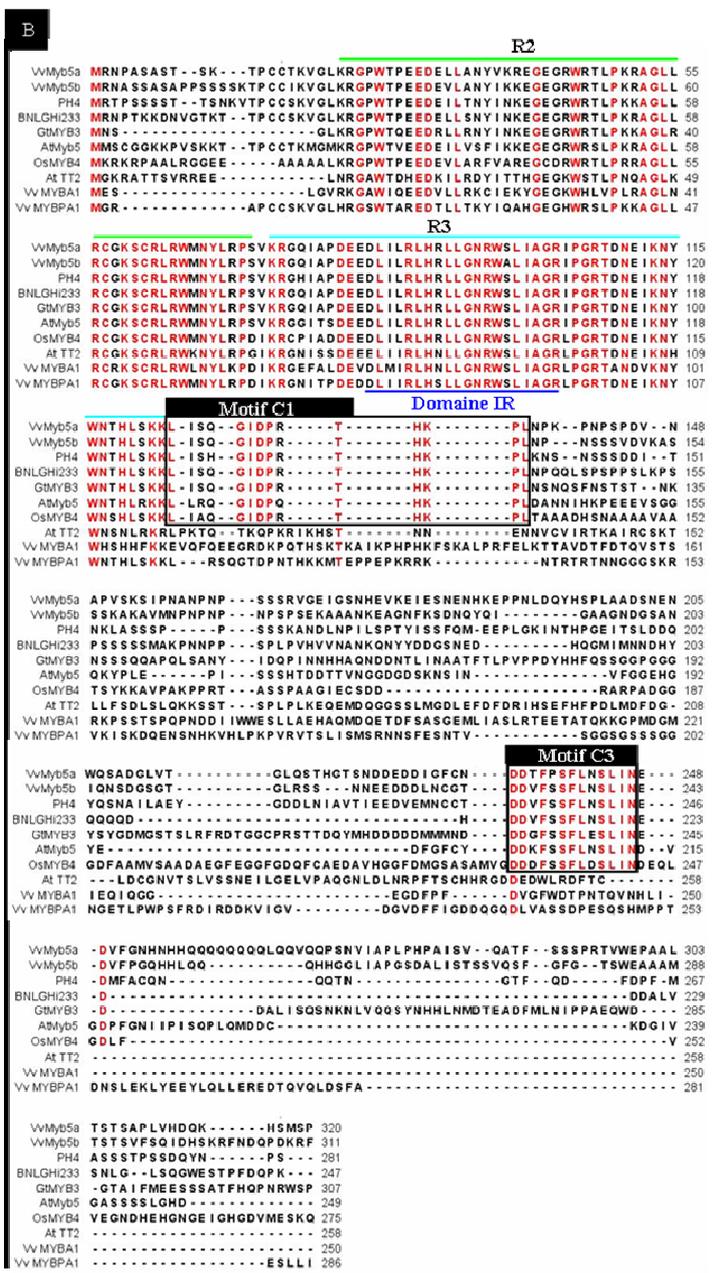
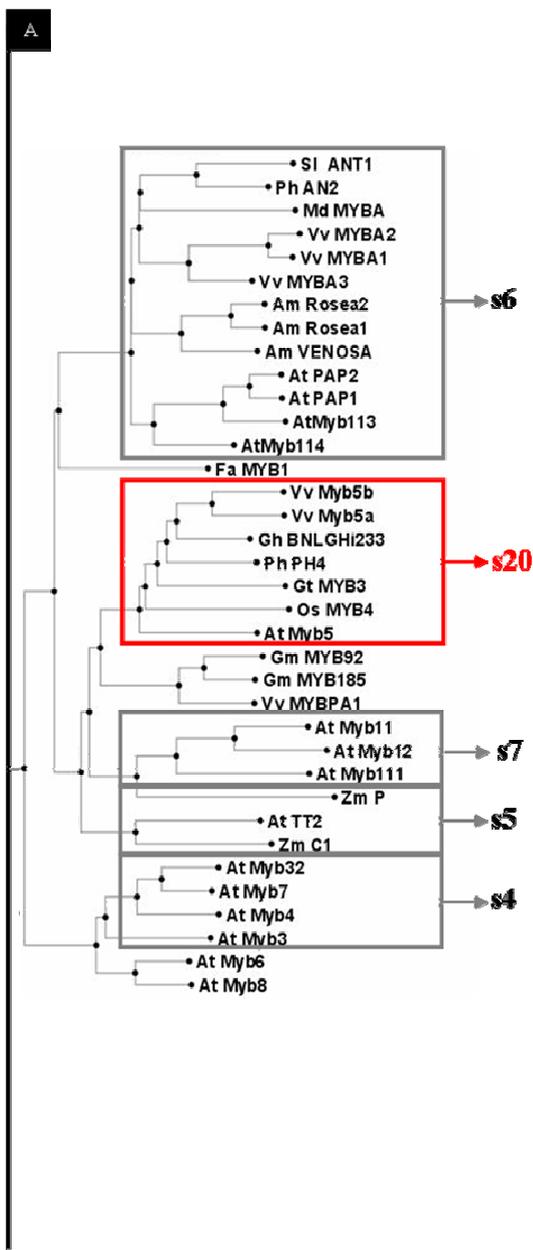


Figure 31 : Analyse des similarités observées entre VvMyb5a et d'autres protéines MYB de plantes.

A) Arbre phylogénétique des séquences protéiques des protéines MYB de *Vitis vinifera* et d'autres espèces végétales. Les séquences protéiques complètes ont été alignées selon un algorithme de type ClustalW et l'arbre phylogénétique a été construit grâce au logiciel Vector NTI. Les cadres gris indiquent les sous-groupes "s" de protéines MYB R₂R₃ définis selon la classification de Stracke [195]. Le cadre rouge indique le nouveau sous-groupe (s20) identifié par Quattrocchio [267]. Les numéros d'accèsion dans les banques de données sont SIANT1 (AAQ55181), PhAN2 (AAF66727), MdMYBA (BAF80582), VvMYBA1 (BAD18977), VvMYBA3 (BAD18979), VvMYBA2 (BAD18978), AmRosea2 (ABB83827), AmRosea1 (ABB83826), AmVENOSA (ABB83828), AtPAP2 (AAG42002), AtPAP1 (AAG42001), AtMyb113 (NP_176811), AtMyb114 (NP_176812), FaMYB1 (AAK84064), VvMyb5b (AAX51291), VvMyb5a (AAS68190), GhBNLGH233 (AAK19611), GtMYB3 (BAF96933), OsMYB4 (BAA23340), AtMyb5 (AAC49311), GmMYB92 (ABH02844), GmMyb185 (ABH02844), VvMybPA1 (CAJ90831), AtMyb11 (NP_191820), AtMyb12 (NP_182268), AtMyb111 (NP_199744), ZmP (AAB67720), AtTT2 (CAC40021), ZmC1 (AAA33482), AtMyb32 (NP_195225), AtMyb7 (NP_179263), AtMyb4 (NP_850879), AtMyb3 (NP_564176), AtMyb6 (NP_192684), AtMyb8 (NP_849749).

B) Alignement des protéines du sous-groupe s20. Les acides aminés identiques sont indiqués en rouge. Les régions conservées correspondant aux motifs C1 et C3 sont encadrées. Le domaine IR indiqué en bleu correspond aux acides aminés homologues aux résidus de la protéine C1 de maïs et nécessaires pour une interaction physique avec les protéines bHLH. Les deux répétitions R2 et R3 qui forment le domaine MYB sont indiqués respectivement par les traits vert et bleu.

Information Resource) et du NCBI⁵ (National Center for Biotechnology Information). Généralement, les protéines MYB présentent de fortes homologies de séquences en N-terminal dans le domaine MYB. Cependant, l'identification de motifs conservés en dehors du domaine MYB a permis de classer les MYB R₂R₃ en 19 sous-groupes [195, 198]. Ces sous-groupes pouvant donner une indication quant à la fonction d'un gène, les séquences protéiques homologues à VvMyb5a et l'ensemble des protéines MYB R₂R₃ d'*Arabidopsis* ont été utilisées pour établir un arbre phylogénétique. Seule une partie de l'analyse de phylogénie a été représentée dans la figure 31A, mais le "clustering" des MYB R₂R₃ identifié par des études antérieures, a été retrouvé. Cette analyse montre que VvMyb5a appartient à un nouveau sous-groupe proche du groupe des protéines caractérisées comme des régulateurs de la voie de biosynthèse des anthocyanes ou des PA. Les sept protéines MYB homologues du sous-groupe présentant les similarités les plus fortes avec VvMyb5a sont impliquées dans divers processus physiologiques et développementaux. La protéine de riz OSMYB4 confère une résistance au froid [268], AtMyb5 serait impliquée dans la détermination cellulaire et le développement des trichomes chez *Arabidopsis* [269], PH4 intervient dans l'acidification vacuolaire chez le pétunia [267], VvMyb5a et VvMyb5b sont des régulateurs de la synthèse des phénylpropanoïdes chez la vigne [249, 250]. Les deux derniers membres de ce groupe, les protéines BNLGHi233 chez le coton (*Gossypium hirsutum*) et GtMyb3 chez la gentiane (*Gentiana triflora*), n'ont pas encore été caractérisées. L'analyse détaillée de l'alignement de l'ensemble de ces protéines montre qu'elles présentent des motifs peptidiques communs (figure 31B). En dehors du domaine DBD très conservé chez les MYB, trois motifs distincts ont été identifiés dans les séquences analysées. Le premier "[D/E]LX₂[R/K]X₃LX₆LX₃R" présent dans la majorité des protéines MYB R₂R₃, correspond au domaine ID (Interacting Domain) et permet l'interaction spécifique avec les protéines bHLH [195, 270]. Les deux autres motifs sont caractéristiques du sous-groupe 20 récemment identifié (figure 31A)[267]. Le premier motif appelé C1 est situé à proximité du domaine MYB et correspond à la signature protéique "Lx3GIDPxTHKPL". Le motif C1, initialement décrit par Kranz et al. (1998) dans les protéines MYB du sous-groupe 4, est présent dans les 7 protéines appartenant au même groupe que VvMyb5a [271]. Le deuxième motif C3 localisé en C-terminal est caractérisé par les aa "DDxF[S/P]SFL[N/D]SLIN[E/D]". Contrairement au motif C1 retrouvé dans d'autres sous-groupes de protéines MYB, C3 est spécifique du sous-groupe 20 mais sa fonction reste inconnue.

L'analyse *in silico* a permis l'identification des homologues les plus proches de VvMyb5a et la classification de ce dernier dans un nouveau groupe de protéines MYB R₂R₃, sur la base de la présence des motifs conservés C1 et C3. Cependant, VvMyb5a est la seule protéine de ce

⁵ NCBI: <http://www.ncbi.nlm.nih.gov/>

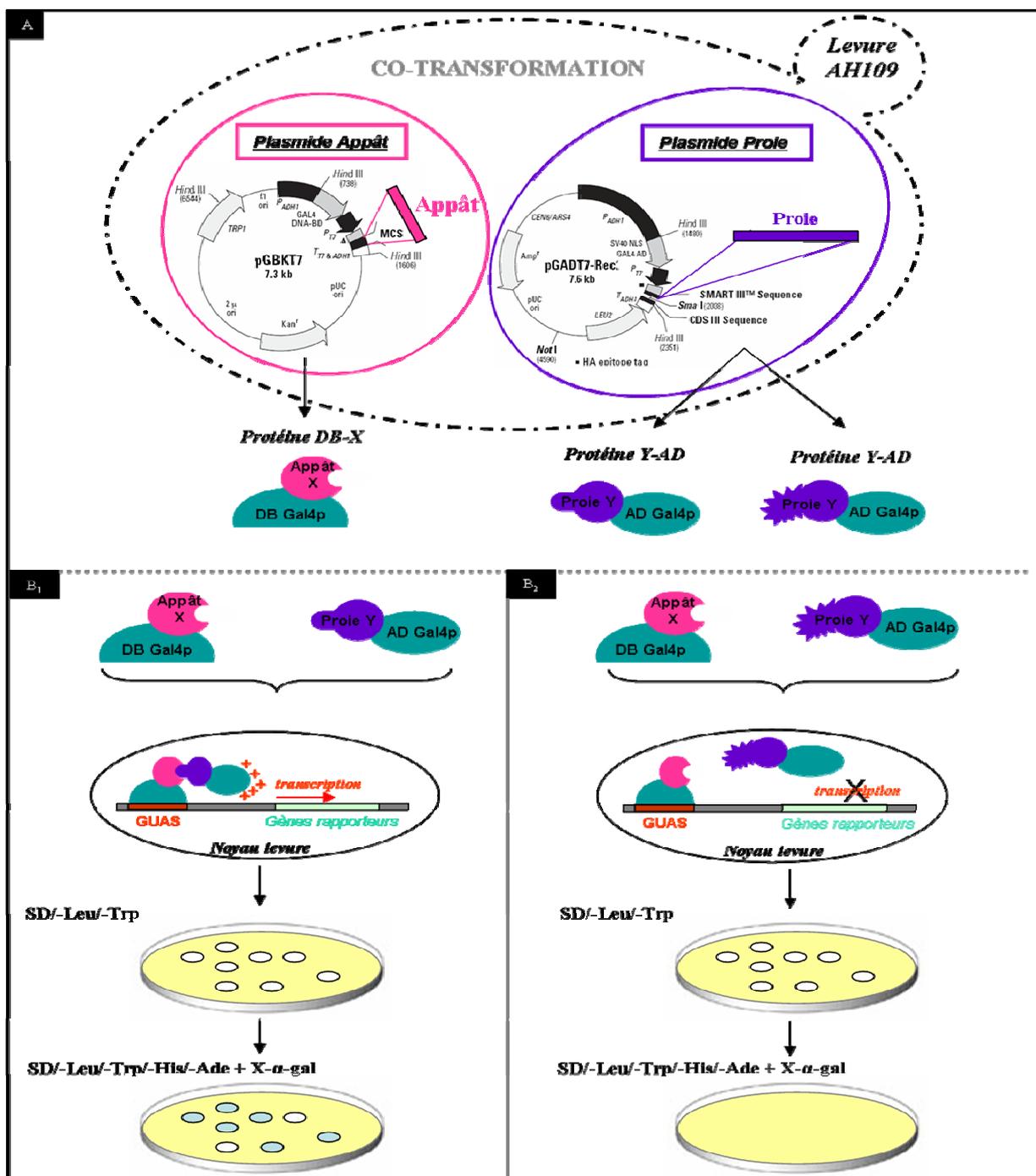


Figure 32 : Principe de la technique du double hybride chez la levure.

A) Une cellule hôte est co-transformée avec les plasmides appât et proie. Le vecteur proie pGADT7-Rec, dans lequel ont été insérés les ADNc de la banque à cribler, présente les séquences supplémentaires suivantes: la séquence codant le domaine d'activation transcriptionnel Gal4p (AD) précédée d'un signal de localisation nucléaire (NLS) et le gène *trp1* permettant la croissance sur milieu carencé en tryptophane. Le vecteur appât pGBKT7, dans lequel a été insérée la protéine ou le domaine protéique d'intérêt, présente les caractéristiques suivantes: une séquence codant le domaine de liaison à l'ADN Gal4p (DB) et le gène *Leu2* permettant la croissance sur milieu carencé en leucine. Les protéines appât (X-AD) et proie (Y-AD) sont chimériques car elles sont fusionnées aux domaines DB ou AD de Gal4, respectivement.

B) L'interaction entre la protéine appât et une cible de la banque d'ADNc reconstitue la protéine Gal4 fonctionnelle qui peut alors activer la transcription des gènes rapporteurs (*His3*, *Ade2*, *LacZ* et *Mell1* dans le système Matchmaker, Clontech). Les gènes *His3* et *Ade2* permettent une sélection sur milieux carencés en histidine et adénine. Les gènes *LacZ* et *Mell1* permettent une sélection chromogénique par détection de l'activité β -galactosidase et α -galactosidase, respectivement. Les levures, dans lesquelles X et Y interagissent physiquement, se développent sur milieu carencé en histidine, adénine, leucine et tryptophane et deviennent bleues (B1). Sans interaction, les levures se développent sur milieux sans tryptophane et leucine (B2).

groupe à posséder un domaine riche en glutamine QQQQQQQQLQQVQQ (domaine GRD, Glutamin Rich Domain). Des séquences riches en glutamine de taille plus ou moins variables ont été identifiées dans plusieurs protéines eucaryotes activatrices ou inhibitrices de la transcription [272, 273]. Au début du travail de recherche qui m'avait été confié, aucune fonction biologique n'avait pu être assignée à ce type de motif. Néanmoins, plusieurs travaux suggéraient que ces motifs étaient d'une part des domaines d'activation de la transcription et, d'autre part, des plateformes potentielles pour des interactions protéines-protéines [274, 275]. De fait, comme VvMyb5a se distingue des autres protéines du sous-groupe 20 par son domaine GRD et qu'aucun motif riche en glutamine n'avait encore été étudié chez les protéines MYB, nous avons engagé sa caractérisation par une approche double hybride chez la levure.

2) Recherche d'interacteurs protéiques du domaine GRD de VvMyb5a

Afin d'identifier des protéines pouvant interagir avec le domaine GRD, la technique de double hybride en levure a été utilisée. Ce système, développé depuis 1989, est l'une des méthodes les plus utilisées pour détecter des interactions physiques entre deux partenaires protéiques ou entre un domaine protéique et une protéine [276].

a. Principe de la technique du double hybride chez la levure

La méthode de double hybride repose sur deux propriétés intrinsèques aux facteurs de transcription. Premièrement, les facteurs de transcription eucaryotes ont typiquement une structure modulaire constituée de deux domaines distincts: un domaine de liaison à l'ADN noté DB (DNA-Binding) et un domaine d'activation de la transcription AD (Activation Domain). De fait, il est possible de créer un FT hybride en combinant le domaine DB d'une protéine et le domaine AD d'une autre [277]. Deuxièmement, ces domaines fonctionnent indépendamment et ne doivent pas nécessairement co-exister au sein de la même chaîne polypeptidique pour entraîner la transcription d'un gène. En effet, quand les domaines sont exprimés séparément puis rapprochés les uns des autres par des interactions non-covalentes, ils peuvent fonctionner collectivement et recréer la fonction transcriptionnelle de la protéine intacte [278, 279].

Le principe du double hybride chez la levure est illustré dans la figure 32. Il s'agit d'un essai *in vivo* d'interaction qui repose sur l'utilisation d'un facteur de transcription issu de la levure *Saccharomyces cerevisiae*, la protéine Gal4p. Dans les levures sauvages en présence de galactose, cette protéine est produite et active la transcription du gène *Gall*, responsable de la digestion du galactose. Le domaine DB de Gal4p se fixe sur une séquence *cis*-régulatrice,

GUAS (*Gall* upstream activating sequence) pour activer la transcription du gène *Gall*. Il a été démontré que dans la levure, la partie intermédiaire de la protéine Gal4p ne modifie pas son action. *Gall*, ainsi que tout gène comportant la séquence GUAS en amont de son promoteur, sera activé par la présence de DB et AD maintenus à proximité l'un de l'autre. Ainsi, si X et Y sont deux protéines qui interagissent, il est possible de reconstituer un facteur de transcription fonctionnel équivalent à Gal4p en exprimant dans la cellule les deux protéines hybrides DB-X et Y-AD. Dans ces conditions, la levure pourra dégrader le galactose si et seulement si X et Y interagissent. Comme la capacité de dégradation du galactose est difficile à observer car la levure dispose de plusieurs mécanismes métaboliques, les souches de levure utilisées dans les systèmes double hybride ont été modifiées par ingénierie génétique. Les cellules de ces levures ne produisent plus la protéine Gal4p et des gènes rapporteurs sélectifs ont été placés sous contrôle de séquences GUAS.

Expérimentalement, les levures sont transformées simultanément par deux plasmides permettant la production de protéines recombinantes (figure 32). La protéine ou le domaine protéique appât (dont on veut identifier les interacteurs) est fusionné(e) au domaine DB de Gal4p (aa 1 à 147). Les protéines proies (candidats potentiels) sont fusionnées au domaine AD de Gal4p (aa 768 à 788). Les plasmides appât (pGBKT7) et proie (pGAD-Rec) portent respectivement les gènes *trp1* et *leu2* permettant ainsi aux levures co-transformées de se développer en milieu de croissance carencé en tryptophane et leucine (SD/-Leu/-Trp). La protéine proie peut se fixer sur les séquences GUAS, mais elle ne peut pas activer la transcription. L'activation des gènes rapporteurs n'a lieu que lorsque le complexe proie-appât se fixe sur une séquence GUAS. Dans le système Matchmaker (Clontech), les séquences GUAS ont été clonées en amont de quatre gènes rapporteurs: *HIS3*, *ADE2*, *MEL1* et *lacZ*.

L'interaction entre l'appât et un partenaire protéique potentiel active l'expression des gènes permettant ainsi aux levures co-transformées de se développer également en milieu de croissance dépourvu d'histidine et d'adénine (SD/-Leu/-Trp/-His/-Ade ou QDO pour Quadruple dropout medium). Deux autres gènes rapporteurs sont activés en cas d'interaction: le gène codant la β -galactosidase (*lacZ*) et le gène codant l' α -galactosidase (*MEL1*). Le gène *lacZ* permet d'effectuer des expériences contrôles de coloration lacZ sur levures issues de milieux de cultures solides ou liquides. L' α -galactosidase est une enzyme sécrétée dont l'activité peut facilement être détectée par l'ajout de son substrat (X- α -Gal) dans les milieux de cultures solides. Si *Mel1* est exprimé et X- α -Gal présent, les colonies de levures se colorent en bleu (figure 32).

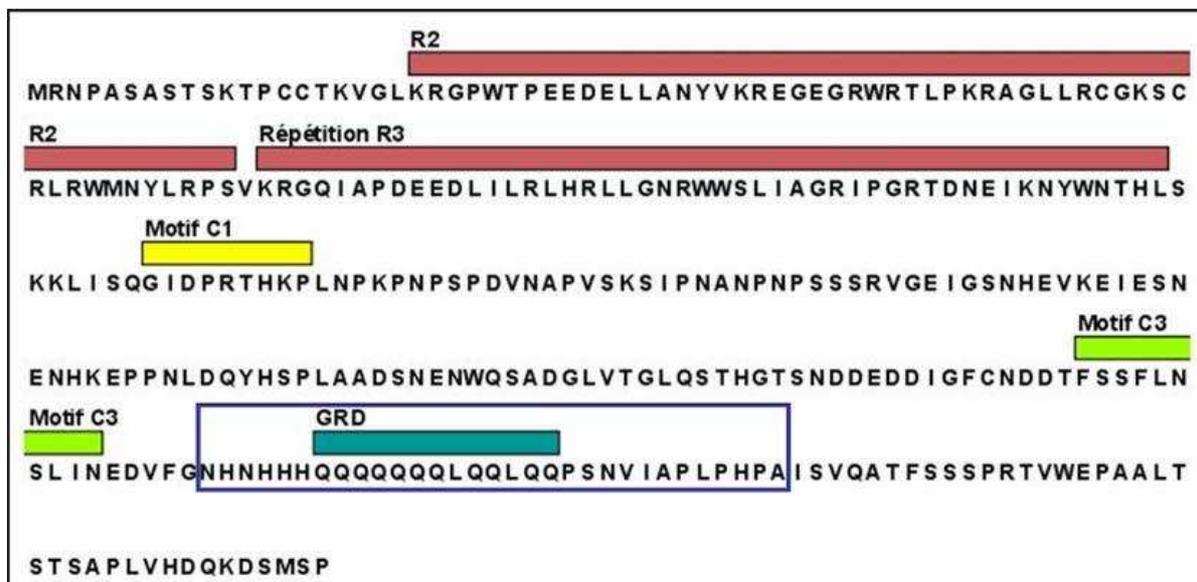


Figure 33 : Représentation schématique des domaines protéiques conservés dans la protéine VvMyb5a et du peptide cible utilisé pour l’approche double hybride.

Les répétitions R2 et R3 du domaine MYB sont colorées en rouge, le motif C1 en jaune, le motif C3 en vert et le domaine GRD (domaine riche en glutamine) en bleu. Le rectangle bleu qui encadre le domaine GRD correspond au peptide appât de 31 aa utilisé pour le double hybride.

Construction	pGBKT7	pGBKT7-VvMyb5b	pGBKT7-GRD
Milieu sélectif			
SD/-Trp/+X-α-gal	+/-	+/+	+/-
SD/-His/-Trp/+X-α-gal	-	+/+	-
SD/-Ade/-Trp/+X-α-gal	-	+/+	-

Tableau III: Représentation schématique des résultats des tests d’autoactivation.

Les différentes constructions utilisées et les résultats de croissance des levures transformées sur différents milieux de sélection sont indiqués dans le tableau. Les contrôles positif et négatif sont respectivement les constructions pGBKT7-*VvMyb5b* et pGBKT7 vide. La construction appât testée pour une autoactivation est pGBKT7-GRD. Le milieu SD/-Trp sélectionne les levures transformées avec le vecteur pGBKT7 ou la construction appât. Les milieux sans histidine ou sans adénine sélectionnent les constructions auto-activatrices. Le X-α-gal permet de sélectionner visuellement les levures où le gène rapporteur *Mel1* est activé car celles-ci se colorent en bleu. Les symboles utilisés sont les suivants : +/- (présence de levure blanche), - (absence de levure), +/+ (présence de levure bleues). Les abréviations utilisées sont : SD pour Synthetic Dropout, Trp pour tryptophane, His pour histidine, Ade pour adénine, X-α-Gal pour 5-Bromo-4chloro-3-indolyl-α-D-galactopyranoside et GRD pour glutamin rich domain.

b. Clonage du domaine GRD dans le plasmide proie et tests préliminaires

Dans le système Matchmaker (Clontech), la souche de levure utilisée est AH109. Cette souche est Ade-, His-, Leu- et Trp- et ne peut se développer sur un milieu dépourvu de ces aa. Les levures qui co-expriment les deux protéines proie et appât sont sélectionnées sur milieu carencé en tryptophane et leucine. Les levures co-transformées où les protéines appât et proie interagissent sont sélectionnées sur milieu QDO. Enfin, les levures AH109 modifiées, offrent la possibilité d'augmenter la stringence du milieu par une sélection chromogénique *via* les gènes *lacZ* et *Mell*. De fait, le milieu le plus sélectif est QDO + X- α -gal.

i. Clonage du domaine GRD

Afin de cibler les interacteurs spécifiques du domaine GRD de VvMyb5a, nous avons choisi d'étendre le domaine appât à six aa situés en amont du domaine GRD et à douze aa situés en aval. La recherche de partenaires a donc été réalisée en utilisant un peptide de 32 aa (figure 33). Pour cloner l'ADNc correspondant à ce peptide, des sites de restrictions ont été ajoutés aux extrémités de chaque amorce. L'amorce sens utilisée comporte le site EcoRI (GRDsens) et l'amorce antisens le site BamHI (GRDAS) (annexe 1). L'ADNc amplifié à partir du vecteur pGEMT[®]-easy-VvMyb5a a été purifié, cloné dans le vecteur pGEMT[®]-easy (annexe 5), multiplié dans des bactéries DH5 α et séquencé. Après vérification, le produit digéré a été cloné en phase avec l'extrémité 3' du domaine DB de Gal4, entre les sites EcoRI et BamHI du site multiple de clonage (MCS) du vecteur pGBKT7 (annexe 6). La protéine chimère est exprimée sous le contrôle du promoteur constitutif ADH1 (P_{ADH1}). Le plasmide pGBKT7-GRD constitue l'appât.

ii. Test préliminaires: toxicité et auto-actvation

Avant de commencer le crible double hybride, des tests préliminaires ont été effectués pour s'assurer que cette technique était adaptée pour notre type d'appât, le domaine GRD de VvMyb5a.

Le peptide GRD est-il toxique pour les levures transformées ?

Dans certains cas, des interactions entre proie et appât ne sont pas identifiées lorsque l'un ou l'autre des partenaires présentent des problèmes de toxicité. Il a fallu écarter la possibilité que la protéine appât (le domaine GRD) produite soit toxique pour la levure. Pour se faire, une comparaison des vitesses de croissance de levures transformées avec le vecteur pGBKT7 vide ou le vecteur pGBKT7-GRD a été effectuée par dosage spectrophotométrique à une longueur d'onde de 600 nm. Après 20 h de culture, l'absorbance était identique (1,520 pour pGBKT7

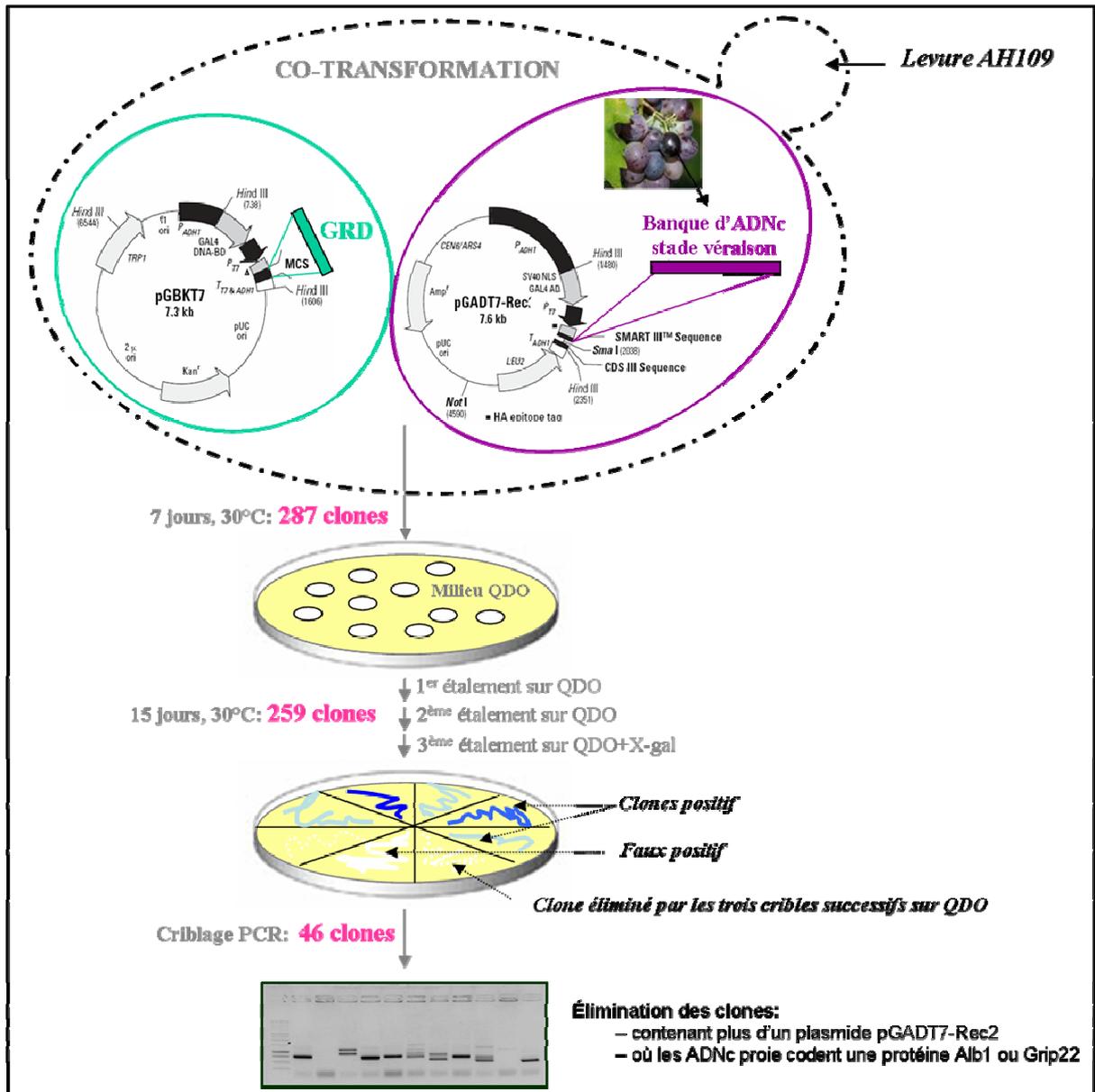


Figure 34 : Représentation schématique du crible effectué pour identifier les clones positifs en double hybride contre le domaine GRD de VvMyb5a.

Les levures AH109 sont co-transformées avec les plasmides appât (pGBKT7-GRD) et proies (pGADT7-Rec2-ADNc de la banque). L'appât est le domaine GRD (glutamin rich domain) de la protéine VvMyb5a. La banque utilisée pour ce crible a été amplifiée à partir d'ARNm extraits de baies du cépage Cabernet sauvignon récoltées au stade véraison et épépinées. Les levures co-transformées dans lesquelles il y a interaction entre l'appât GRD et une protéine de la banque sont sélectionnées sur milieu QDO (SD/-His/-Trp/-Ade/-Leu). Au bout de 7 jours, les levures qui s'étaient développées ont été ré-étagées 2 fois sur milieu QDO puis une fois sur QDO + X- α -gal pour augmenter la stringence du crible. Après le crible nutritionnel, un criblage par PCR a été effectué afin d'éliminer les clones ayant intégré plus d'un plasmide proie et d'identifier les clones récurrents qui exprimaient les protéines Alb1 et Grip22. En parallèle, les contrôles positifs et négatifs du kit Matchmaker ont été utilisés.

vide et 1,595 pour pGBKT7-GRD) et supérieure à 0,8. Donc, nous avons pu conclure que la protéine proie n'était pas toxique pour les levures AH109.

Le peptide GRD est-il un auto-activateur ?

Certains appâts sont des auto-activateurs, c'est-à-dire qu'ils peuvent activer la transcription en l'absence d'interacteur. De ce fait, la capacité du domaine GRD à *trans*-activer les gènes rapporteurs a également été testée. Si ce peptide est un auto-activateur, il activera alors l'expression des gènes rapporteurs *ade2*, *his3*, *mell* et *lacZ*. Les levures exprimant le domaine GRD ont été étalées sur SD/-Trp+X- α -gal, SD/-His/-Trp+X- α -gal et SD/-Ade/-Trp+X- α -gal. Le vecteur vide pGBKT7 a été utilisé comme contrôle négatif et la construction pGBKT7-*VvMyb5b*, agissant comme un FT MYB auto-activateur (Hichri I., communication personnelle) a été utilisée comme contrôle positif. Une représentation synthétique des résultats est présentée dans le tableau III. Les levures transformées avec la construction GRD se développent sur un milieu SD/-Trp+X- α -gal mais pas sur les milieux SD/-His/-Trp+X- α -gal et SD/-Ade/-Trp+X- α -gal. De plus, les colonies qui se développent ne se colorent pas en bleu, indiquant que le gène rapporteur *Mell* n'est pas activé. Le contrôle négatif, à savoir les levures transformées avec le vecteur vide pGBKT7, ne se développe pas sur les milieux SD/-His/-Trp+X- α -gal et SD/-Ade/-Trp+X- α -gal alors que les levures transformées avec le contrôle positif (pGBKT7-*VvMyb5b*) se développent et sont bleues. Ces résultats indiquent que le domaine GRD n'est pas capable d'activer les gènes rapporteurs. Un criblage contre une banque d'ADNc a donc été réalisé afin d'identifier les partenaires protéiques.

c. Criblage de la banque

La procédure de criblage choisie a été la co-transformation des levures AH109 avec la construction appât et les plasmides proies (pGADT7-Rec-ADNc, annexe 7) dans lesquels ont été insérés les ADNc représentatifs de la banque par recombinaison homologue. Pour le criblage de partenaires protéiques, nous avons utilisé une banque d'ADNc produite à partir d'ARNm extraits de baies de raisin récoltées au stade véraison et épépinées. Les levures co-transformées ont été étalées directement sur le milieu sélectif QDO. Le nombre de colonies sur ces boîtes après 7 j d'incubation à 30°C s'élevait à 286. Afin d'éliminer les faux positifs, plusieurs cribles ont été effectués et les résultats obtenus sont indiqués dans la figure 34. Les 286 clones ont d'abord été étalés successivement deux fois sur un milieu QDO puis une fois sur un milieu QDO+X- α -gal. Après quatre jours d'incubation à 30°C, les clones qui n'étaient pas colorés en bleu ont été considérés comme des faux positifs. Vingt sept clones ont ainsi pu être éliminés. Sur les 259 clones restants, un criblage PCR a été réalisé afin d'éliminer ceux

N° Clone	Taille du fragment PCR (pb)	En phase avec AD Gal4?	Taille ORF proie	Références Génoscope	Taille de la protéine Génoscope	Positionnement du peptide proie sur la protéine du Génoscope	Identité putative du gène
3 et 28	948 pb	Oui	175 aa	GSSVIV/P00034774001	317 aa	40-314	late embryogenesis abundant family protein
21 et 255	844 pb	Oui	194 aa	GSSVIV/P00036641001	266 aa	104-266	Cyclase family protein
22	984 pb	Oui	187 aa	GSSVIV/P00019194001	273 aa	86-273	Beta-expansine
30	891 pb	Oui	25 aa	GSSVIV/P00010914001	455 aa	Pas d'alignement	putative 41 kD chloroplast nucleoid DNA binding protein
33	966 pb	Oui	26 aa	GSSVIV/P00010880001	149 aa	Pas d'alignement	Grip 68
36	947 pb	Oui	34 aa	GSSVIV/P0000605001	251 aa	220-251	tonoplast intrinsic protein
37	720 pb	Oui	122 aa	GSSVIV/P00032598001	122 aa	1-122	ubiquinol-cytochrome C reductase complex 14 kDa protein
48	1009 pb	Oui	228 aa	GSSVIV/P00003204001	458 aa	230-458	protein kinase
49	886 pb	Oui	157 aa	GSSVIV/P00001103001	222 aa	97-222	VYTL1 (thaumatin-like protein)
58	892 pb	Oui	68 aa	GSSVIV/P00036563001	735 aa	Pas d'alignement	CXE carboxylesterase
61	1131 pb	Oui	142 aa	GSSVIV/P00032600001	241 aa	31-241	small soluble GTP-binding protein
66	964 pb	Oui	144 aa	GSSVIV/P00014370001	359 aa	248-359	family II extracellular lipase 3
68	758 pb	Oui	136 aa	GSSVIV/P00006205001	272 aa	167-272	ubiquinol-cytochrome C reductase iron-sulfur subunit
72	1119 pb	Oui	248 aa	GSSVIV/P00034644001	264 aa	49-264	class IV chitinase [Vitis vinifera]
81	943 pb	Oui	159	GSSVIV/P00018199001	348 aa	220-348	chorismate synthase
83	874 pb	Oui	227	GSSVIV/P00011932001	393 aa	197-393	chloroplast enzyme sedoheptulose-1,7-bisphosphatase (SBPase)
86	694 pb	Non	/	/	/	/	/
96	480 pb	Non	/	/	/	/	/
97	996 pb	Oui	142	GSSVIV/P00003202001	600 aa	392-600	dehydration-responsive protein
165	346 pb	Oui	17 aa	GSSVIV/P00026043001	384 aa	Pas d'alignement	isocitrate lyase
167 et 230	1055 pb	Non	136 aa	GSSVIV/P00014875001	262 aa	126-262	cysteine proteinase
177	911 pb	Oui	251 aa	GSSVIV/P00015566001	350 aa	130-350	WD-40 repeat family protein
183	560 pb	Oui	163 aa	GSSVIV/P00033399001	650 aa	560-617	unknown protein
205	1019 pb	Oui	104 aa	GSSVIV/P00029445001	252 aa	179-252	Expansin Gene Family
209	1199 pb	Oui	326 aa	GSSVIV/P00001079001	727 aa	435-727	copper-containing amine oxidase
249	1161 pb	Oui	146 aa	GSSVIV/P00034064001	771 aa	656-771	subtilase family protein
257	1043 pb	Non	40 aa	GSSVIV/P00028751001	368 aa	332-368	unknown protein
283	1300 pb	Oui	96 aa	GSSVIV/P00022551001	673 aa	610-673	unknown protein

Tableau IV : Interacteurs protéiques possibles du domaine GRD de la protéine VvMyb5a identifiés par la technique du double hybride en levure

Ce tableau présente les protéines de vigne identifiées par la technique de double hybride pouvant potentiellement interagir avec le domaine GRD de VvMyb5a. Le numéro du clone, la taille du fragment PCR amplifiés par les amorces T7DH et 3AD et les peptides proies en phase avec le domaine d'activation de Gal4p sont indiqués. Les références ont été obtenues par un blasin réalisé dans la base de données de *Vitis vinifera* (Génoscope). Un alignement a été réalisé entre les peptides proies et les protéines du Génoscope correspondantes pour s'assurer que les ADNc étaient clonés dans le bon cadre de lecture. Quand elles n'étaient pas annotées dans le Génoscope, les fonctions hypothétiques des protéines proies ont été identifiées par le logiciel Blast dans NCBI.

qui contenaient plus d'un plasmide proie. Les ADNc de la banque codant des protéines potentiellement interactrices du domaine GRD ont été amplifiés par des amorces spécifiques du vecteur pGADT7-Rec (T7DH et 3AD, annexe 3) et les produits d'amplification ont été analysés par électrophorèse en gel d'agarose (1,2%). Les clones contenant plus d'un plasmide pGADT7-Rec et ceux pour lesquels la taille du produit d'amplification était d'environ 250 pb (indiquant que le vecteur était vide) ont également été éliminés. Sur 259 clones, 114 ont pu ainsi être éliminés. Les produits d'amplification obtenus pour 20 clones ont été ligaturés dans le vecteur pGEMT[®]-easy, multipliés dans des bactéries DH5 α , puis séquencés. Deux séquences codant pour des protéines connues sont apparues fortement représentées parmi ces 20 clones : il s'agit de séquences codant la protéine Alb1 (2S albumine precursor) et la protéine Grip22 (Grape ripening induced protein 22). Afin d'éliminer ces clones récurrents, un deuxième crible PCR a été réalisé en utilisant des amorces spécifiques des ADNc codant les gènes *Alb1* (2Ssens et 2SAS) et *Grip22* (Grip22sens et Grip22AS) (annexe 3). Ce crible a permis d'éliminer 51 clones représentatifs de l'ADNc *Alb1* et 45 de l'ADNc *Grip22*. Ainsi, aux termes de ces différents cribles, 46 clones positifs ont été sélectionnés.

Pour identifier les ADNc présents dans chaque clone détecté, les plasmides ont été isolés et séquencés. Cette ultime étape nécessite l'extraction des plasmides proies des clones de levures et la multiplication de ces plasmides dans des bactéries.

a. Résultats du séquençage des clones positifs

Sur les 46 clones positifs sélectionnés, seuls 31 plasmides ont pu être extraits des levures. Pour les 15 clones récalcitrants, les ADNc ont été amplifiés et les produits de réactions PCR purifiés puis directement séquencés. Cependant, les résultats du séquençage se sont avérés de qualité très moyenne et n'ont pu être exploités à ce jour. Le tableau IV présente les résultats de l'analyse des séquences des 31 clones par le programme informatique BLAT¹ disponible sur le site web du Génoscope. Après cette analyse, certains clones ont pu être directement éliminés. Les clones 86 et 96 contenaient des fragments ADNc hors phase avec la séquence codant le domaine d'activation de Gal4p et aucun cadre de lecture ouvert (ORF, "Open Reading Frame") n'a pu être identifié. Les clones 30, 33, 58 et 165 contenaient des fragments ADNc en phase avec le domaine d'activation de Gal4p et leurs séquences nucléotidiques correspondaient à celle d'un gène identifié dans le génome de la vigne. Cependant, les séquences en aa de la protéine partielle proie et de la protéine correspondante du Génoscope ne s'alignaient pas. En fait, la recombinaison de l'ADNc dans le vecteur proie avait modifié le cadre de lecture et aucun ORF non fusionné au domaine d'activation n'a pu être identifié

¹ BLAT Génoscope : <http://www.Génoscope.cns.fr/blat-server/cgi-bin/vitis/webBlat>

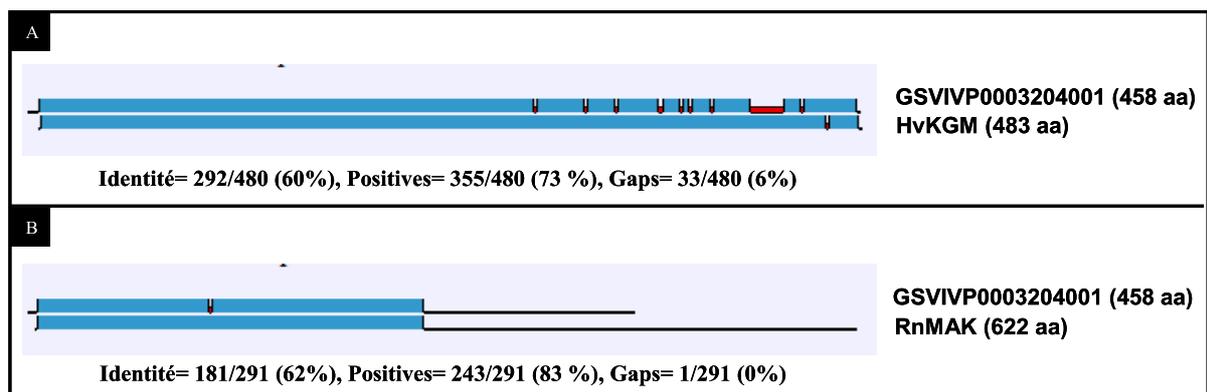


Figure 35 : Comparaison des séquences protéiques de GSVIVP0003204001 avec deux membres de la famille des protéines Mak kinase.

Cette figure a été obtenue avec le programme BLAST 2 SEQUENCES qui permet un alignement des séquences protéiques deux à deux. Le nombre d'acides aminés est indiqué entre parenthèse. Les cadres bleus représentent les zones homologues. Les décalages sont représentés en rouge. L'identité représente le nombre d'acides aminés en commun entre les deux séquences sur le nombre total d'acides aminés de la zone en question. "Positives" représente le nombre d'acides aminés du même groupe. "Gaps" représente le nombre de décalage d'acides aminés dans la séquence pour un meilleur alignement. Ces différentes valeurs sont aussi exprimées en pourcentage. La double lettre préfixe indique l'origine de la protéine Hv pour orge (*Hordeum vulgare*) et Rn pour rat (*Rattus norvegicus*). Les numéros d'accèsion sont AY167561 pour HvKGM et NP_037268 pour RnMAK.

dans le reste de la séquence. Parmi les clones restants, les autres protéines identifiées sont indiquées dans le tableau IV. Parmi ces protéines, deux ont particulièrement retenu notre attention : une protéine kinase et une WD40.

Le clone 48 contenait un fragment d'ADNc correspondant à 873 pb de la séquence du gène annoté *GSVIVT00003204001* dans la base de données du Génoscope (693 pb de la séquence codante en C-terminale et 183 pb de l'extrémité 3' non codante). Le gène *GSVIVT00003204001* complet code une protéine de 458 aa présentant une forte homologie de séquence avec la protéine kinase KGM (Kinase associated with GAMYB) de blé [280]. Cette protéine appartient à un sous-groupe de sérine/thréonine kinase, les Mak kinases (Male germ cell Associated Kinase) dont le premier gène nommé *mak* a été cloné chez le rat [281]. Isolée également par un crible double hybride, la protéine KGM interagit physiquement avec la protéine GAMYB et régule négativement son activité [280]. Les deux séquences protéiques, correspondant aux protéines KGM de blé et KGM hypothétique de vigne, ont été comparées deux à deux par le programme "blast 2 sequences" (figure 35). Les deux protéines présentent une homologie de séquence sur toute leur longueur. L'identité globale est de 60 % et la similarité de 73 %. A l'inverse, la comparaison avec la protéine mak de rat montre une homologie de séquence seulement dans la région correspondant aux sous-domaines catalytiques avec 62 % d'identité et 83% de similarité (figure 35).

Enfin, le clone 177 présente un cadre de lecture ouvert codant pour une protéine de type WD40. L'ADNc inséré correspondait à une séquence de 828 pb du gène *GSVIVT00015566001* (666 pb de la séquence codante C-terminale et 162 pb de la région 3'non-codante). Certaines protéines de type WD40 ont déjà été caractérisées comme des régulateurs de la voie de biosynthèse des flavonoïdes (p42). Cependant, la recherche d'homologues dans la banque de données NCBI n'a pas montré de similarités significatives avec des protéines WD40 déjà identifiées et impliquées dans la régulation de la biosynthèse des composés phénoliques.

II- Etude fonctionnelle du promoteur de *VvMyb5a*

La compréhension du rôle de *VvMyb5a* dans la voie de biosynthèse des phénylpropanoïdes au cours du développement de la baie de raisin implique l'analyse de son expression. Le contrôle de l'expression des gènes se fait en partie au niveau de la transcription. Des signaux physiologiques, générés à l'extérieur de la cellule, sont transmis jusqu'au noyau où les ARN

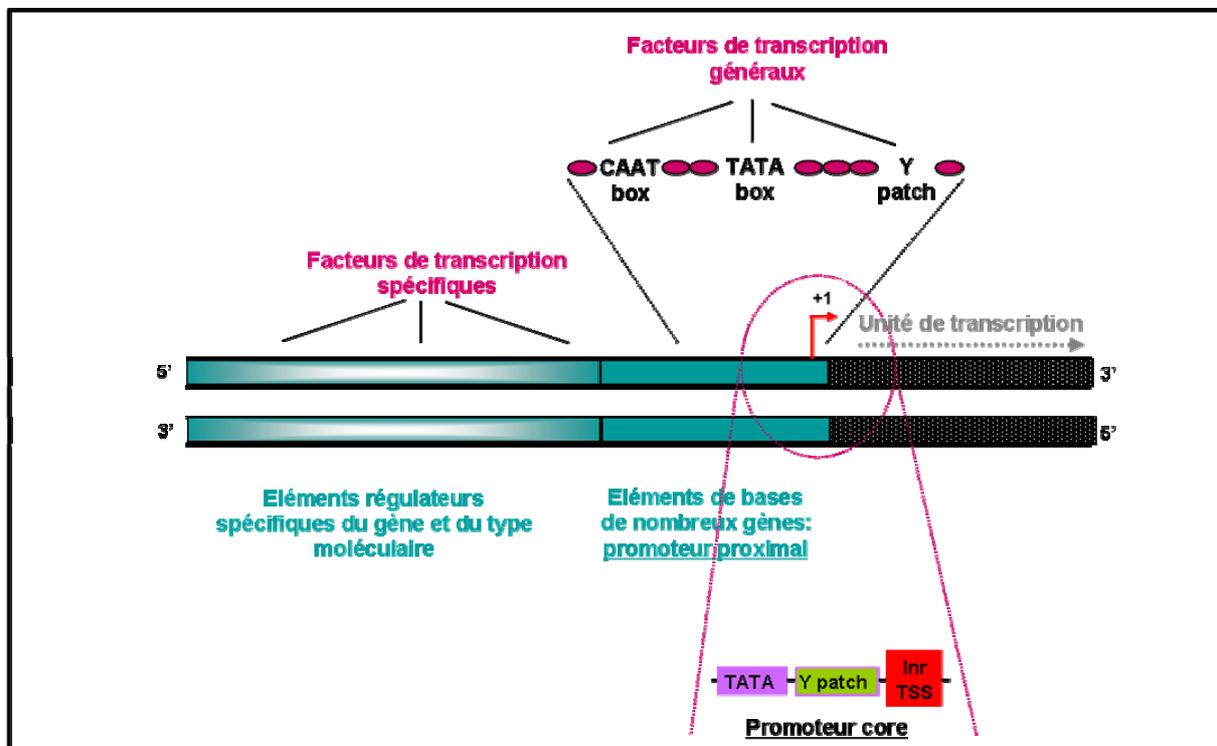


Figure 36: Représentation schématique de la structure des promoteurs de plantes.

Le promoteur est situé en amont de la séquence d'un gène. Le promoteur core et le promoteur proximal sont caractérisés par des motifs conservés et participent directement à l'initiation de la transcription. Cette région comprend le site initiateur (Inr) et le site d'initiation de la transcription (TSS), la TATA-box et une séquence riche en pyrimidine nommée Y patch. Dans la plupart des promoteurs de plantes, le motif Y patch est situé entre la TATA-box et le TSS/Inr [282, 283]. Le TSS est indiqué par une flèche rouge noté +1. Ces séquences sont reconnues spécifiquement par les facteurs de transcription généraux afin de guider l'ARN polymérase II. En dehors de ces motifs obligatoires, d'autres motifs *cis* spécifiques du gène sont présents, généralement en amont du promoteur proximal. Ces séquences permettent la régulation fine du gène en fonction de sa spécificité d'action *via* la reconnaissance spécifique par des facteurs de transcription.

polymérase déclenchent la synthèse des molécules d'ARN. Ces signaux transmis peuvent amplifier ou atténuer l'expression des gènes en modifiant notamment les interactions des FT avec leurs cibles. Ainsi, l'analyse de la région promotrice de *VvMyb5a* était incontournable pour comprendre sa spécificité d'action au cours du développement de la baie de raisin. Durant le travail de recherche de L. Deluc, la séquence promotrice avait été obtenue par la technique de PCR inverse. Après avoir vérifié qu'il s'agissait de la séquence promotrice de *VvMyb5a*, nous avons réalisé une analyse *in silico*. Ces analyses bioinformatiques ont permis de rechercher les éléments de régulation présents dans la région promotrice de *VvMyb5a*. En parallèle, nous avons entrepris la dissection fonctionnelle de la région promotrice à l'aide de délétions progressives en 5'. Enfin, nous avons recherché les facteurs de transcription pouvant réguler en *trans* l'activité d'un fragment du promoteur *VvMyb5a* par une approche simple hybride chez la levure.

1) Définition biologique du promoteur reconnu par l'ARN polymérase II

Le promoteur, région située en amont d'un gène, contient les éléments permettant de réguler sa transcription et d'assurer un taux maximal d'expression pour autant que les conditions nécessaires soient réunies (présences de protéines adéquates)[284]. S'il est parfois difficile de s'accorder sur la limite 5' d'un promoteur, sa limite 3' en revanche est définie par la région initiateur de la transcription (Inr). L'Inr recouvre le motif consensus TSS (Transcription Start Site) qui correspond au site +1 de la transcription, l'endroit où débute la transcription (figure 36). Le promoteur eucaryote, long de plusieurs centaines de bases, est composé de courts motifs sur lesquels viennent se fixer les protéines impliquées dans la transcription. Un premier type de motifs regroupe les éléments non spécifiques, trouvés sur la plupart des promoteurs et utilisés de façon constitutive. Il s'agit de sites de fixation des facteurs généraux de la transcription, préférentiellement impliqués dans l'initiation même de la transcription. Un second type de motifs regroupe les éléments spécifiques à une classe de promoteurs, et dont l'utilisation est régulée. Ce sont eux qui régulent à proprement parler la transcription en fonction des tissus, des stades développementaux, des stimuli environnementaux...

Le premier type de motifs (motifs constitutifs) est localisé dans la région de contrôle située au voisinage immédiat du site d'initiation de la transcription. D'une longueur d'environ 100 pb (+ 40 à -60 par rapport au TSS), cette région correspond au promoteur core. Elle représente la région minimale capable d'initier la transcription basale d'un gène (figure 36). Plusieurs éléments fonctionnels du promoteur core ont été identifiés : TSS, TATA-box, Inr, DPE (Downstream Promoter Element), BRE (TFIIB Recognition Element) et îlots CpG. Ces séquences permettent la fixation du complexe d'initiation formé de l'ARN polymérase

associée aux facteurs généraux de la transcription (TFII-A, TFII-B, TFII-D, TFII-E et TFII-F) [282]. Ces sites de fixation sont supposés être présents sur tous les promoteurs mais ils ne sont pas obligatoires [145]. D'ailleurs, l'architecture du promoteur core chez les mammifères et les plantes est différente [283]. Les promoteurs core de plantes sont caractérisés par la TATA-box, la région Inr et Y-Patch (patch pyrimidine). La TATA-box, motif fonctionnel conservé dans les promoteurs eucaryotes, est localisée entre 45 et 25 pb en 5' du TSS. Elle permet la fixation du facteur TFIID, qui est en fait un complexe entre la protéine TBF (TATA binding protein) et quatorze protéines liées à la TBF. Certains gènes sont appelés "TATA-less promoters" car ils sont dépourvus de TATA-box. C'est le cas par exemple de gènes de ménages (codant les protéines ribosomiques des plastes [285]) et de gènes liés à la photosynthèse (gènes nucléaires codant le photosystème I [286]). Concernant les motifs Y-patch, ces séquences ont été identifiées récemment suite à une comparaison des promoteurs de gènes de plantes (*Arabidopsis* et maïs) et de mammifères (homme et rat)[282, 283]. Ces motifs spécifiques des plantes et très conservés sont composés de séquences riches en pyrimidines (C et T) et sont localisés entre 1 et 100 pb en 5' du TSS. Cependant, la fonction biochimique de ces motifs n'est pas encore connue. Un autre motif consensus, généralement retrouvé dans le promoteur proximal (jusqu'à 200 pb en 5' du TSS), est la CAAT-box. Cette boîte influencerait la fréquence d'initiation de la transcription.

Le deuxième type de motifs (motifs spécifiques) regroupe de courtes séquences (généralement 6 à 8 nucléotides mais parfois jusqu'à une vingtaine) dispersées le long de la région située en 5' du gène et appelées éléments *cis*-régulateurs ou CARE (*cis*-acting regulatory element) ou REG (Regulatory Element Group) (figure 36). Ces motifs constituent les séquences distales du promoteur et sont conservés entre les espèces. On les retrouve parfois en aval du TSS dans la région non traduite de l'ARN (5' UTR). Des facteurs dits *trans*-régulateurs vont se fixer sur ces séquences, modifiant brusquement le taux de transcription en l'augmentant (régulation positive) ou en le diminuant (régulation négative). Les séquences qui les séparent ne semblent pas importantes, mais la distance entre ces éléments peut l'être.

La complexité de la régulation de l'expression des gènes ne s'arrête pas là. En effet, les régions intergéniques sont très longues (en fait, associées aux introns, elles couvrent la majeure partie du génome) et il existe des éléments régulateurs situés très en amont du site de la transcription. Ces éléments sont appelés "enhancer" (activateurs) et "silencer" (répresseurs)[147]. Ils agissent sur la transcription des gènes grâce à la superstructure que forme l'ADN en se repliant sur lui-même pour former des boucles, rapprochant ainsi deux régions très éloignées dans la séquence primaire de l'ADN. La mise en évidence de ces

éléments de régulation lointains du TSS nécessite une dissection fonctionnelle de la région promotrice.

2) Clonage et analyse *in silico* du promoteur *VvMyb5a*

a. Clonage du promoteur *VvMyb5a*

La séquence clonée par la technique de PCR Inverse par L. Deluc présentait une queue polyA de 400 pb en aval de l'ATG de la séquence codante de *VvMyb5a*. Ainsi, afin de confirmer que cette séquence correspondait spécifiquement à la région promotrice de *VvMyb5a*, une réaction PCR a été réalisée sur de l'ADNg du cépage Cabernet sauvignon avec le couple d'amorce PromVvMyb5asens (complémentaire de l'extrémité 5' de la région promotrice de *VvMyb5a*) et 5'VvMyb5aAS (complémentaire du début de la séquence codante N-terminale de *VvMyb5a*) (annexe1). Le produit d'amplification de la réaction PCR d'environ 1400 pb a été cloné dans le vecteur pGEMT[®]-easy puis envoyé à séquencer. La séquence inconnue correspondant à la région promotrice était bien flanquée en 5' de la séquence codante du gène *VvMyb5a*. Par la suite, la publication du génome de la vigne a permis de vérifier que la séquence que nous avons obtenue correspondait bien à la région promotrice du gène *VvMYB5a*.

b. Analyse *in silico* de la séquence promotrice de *VvMyb5a*

De nombreux outils bioinformatiques permettent d'identifier les différents éléments de contrôle, consensus et spécifiques, présents dans les régions promotrices. Pour étudier un promoteur, il faut sélectionner des banques de données spécialisées. Ces banques doivent être capables d'analyser et de comparer la séquence promotrice cible pour identifier les éléments *cis*-régulateurs les plus pertinents du point de vue biologique. De fait, pour comprendre la nature de la fonction du promoteur *VvMyb5a*, une analyse *in silico* détaillée a été réalisée sur la région promotrice située 1337 pb en amont du codon d'initiation. Une recherche des éléments *cis*-régulateurs potentiels et des FT susceptibles de s'y fixer a été effectuée dans la banque de motifs PLACE² (Plant Cis-Acting Regulatory DNA Elements) et par le programme MatInspector (logiciel Génomatix³) [287, 288] [289, 290].

PLACE est une base de donnée japonaise qui répertorie les motifs activateurs et répresseurs de séquences nucléotidiques *cis*-régulatrices de plantes. Cette base de données référence les motifs de bases présents dans des familles de promoteurs de plantes déjà identifiées dans des publications scientifiques. Pour modéliser un motif *cis*-régulateur, PLACE se sert d'une représentation consensus dégénérée. Cette représentation utilise une nomenclature IUPAC

² PLACE: <http://www.dna.affrc.go.jp/PLACE/signalscan.html>

³ Génomatix : <http://www.genomatix.de/index.html>

Outils de prédiction	Nb de motifs détectés	Nb de TATA-box	Nb de CAAT-box
PLACE	143	17	35
MatInspector	169	1	7

Tableau V : Analyse comparative des motifs *cis*-régulateurs identifiés dans le promoteur *VvMyb5a* par les outils de prédiction PLACE et MatInspector.

La taille du promoteur *VvMyb5a* soumis à ces outils de prédictions est de 1400 pb. Le nombre de motifs détectés correspond à l'ensemble des motifs *cis*-régulateurs (consensus et spécifiques) prédits par PLACE (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>) et MatInspector (logiciel, Génomatix <http://www.genomatix.de/index.html>). Le nombre de motif consensus TATA et CAAT-box est également indiqué.

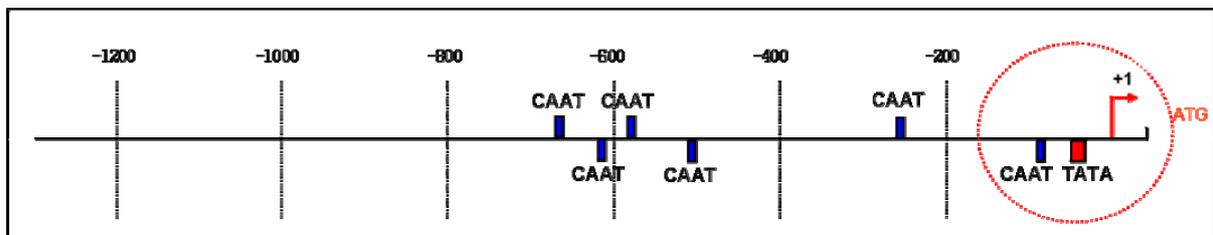


Figure 37 : Représentation schématique de la position des motifs consensus du promoteur proximal *VvMyb5a*

Le promoteur *VvMyb5a* est représenté par une ligne noire. La flèche notée +1 indique le TSS (site d'initiation de la transcription). La TATA-box et les CAAT-box sont représentées respectivement par des rectangles rouges et bleus. Les boîtes situées sur le brin + du promoteur sont au dessus de la ligne et ceux sur le brin - en dessous de la ligne. Le cercle rouge indique la position du promoteur proximal hypothétiques de *VvMyb5a*. Le +1 est situé 39 pb avant l'ATG de la séquence codante du gène *VvMyb5a*. La TATA-box et la CAAT-box sont respectivement localisées aux positions -29 et -79 pb. Les motifs ont été prédits par la base de données PLACE et le programme MatInspector (logiciel Génomatix).

(International Union of Pure and Applied Chemistry), qui permet de prendre en compte une certaine forme de variabilité dans les séquences en considérant qu'une position peut être décrite par un seul, deux, trois ou quatre nucléotides différents.

Le programme MatInspector, lui, utilise une grande bibliothèque de descriptions matricielles de sites de liaisons de facteurs de transcription (TRANSFAC⁴) pour localiser les motifs d'ADN cibles. TRANSFAC est une base de données d'éléments eucaryotes *cis*- et *trans*-régulateurs. La plupart des données présentes dans TRANSFAC sont extraites d'une compilation bibliographique.

3) Recherche de motifs consensus du promoteur proximal de *VvMyb5a*

La plupart des analyses de promoteurs débutent avec la recherche des motifs consensus du promoteur core à savoir le TSS, la TATA-box et la CAAT-box. Pour amplifier l'ADNc pleine longueur correspondant à *VvMyb5a*, une banque d'ADNc orientée produite à partir de baies récoltées à différents stades de développement du cépage Cabernet sauvignon a été utilisée comme matrice [291]. L'extrémité 5' non-codante de *VvMyb5a* qui a été amplifiée était de 39 pb et la dernière base amplifiée a été considérée dans cette étude comme le +1 de la transcription du gène *VvMyb5a* (figure 37).

De nombreux éléments régulateurs hypothétiques du promoteur proximal de *VvMyb5a* ont été identifiés par les outils de détection PLACE et MatInspector (tableau V). Plusieurs TATA-box putatives ont été localisées mais une seule boîte consensus dans PLACE et MatInspector était située entre 20 et 30 nucléotides en 5' du TSS (figure 37). De même, parmi toutes les boîtes CAAT putatives prédites, plusieurs motifs étaient consensus, mais une boîte CAAT était localisée 50 à 130 pb en amont du +1 de la transcription prédit (figure 38). Finalement, les recoupements entre les données issues des différents logiciels utilisés ont permis de localiser une TATA-box et une CAAT-box hypothétiques, respectivement à -29 pb et -79 pb du TSS. De plus, aucun des logiciels utilisés n'a prédit la présence de motif Y-patch.

a. Recherche des motifs spécifiques du promoteur *VvMyb5a*

Outre les éléments de régulation canonique, ces logiciels ont également prédit un certain nombre de séquences régulatrices hypothétiques déterminantes pour la nature du promoteur *VvMyb5a*. La plupart des FT se liant à des zones très courtes (5-25 bases) constituées de motifs dégénérés facilement retrouvés au hasard du génome, les sites identifiés ne restent que hypothétiques. Un des inconvénients de ces outils est donc l'abondance de faux négatifs et de

⁴ TRANSFAC: <http://www.biobase-international.com/pages/index.php?id=transfac>

Régulation d'expression	Eléments régulateurs PLACE	Eléments régulateurs MatInspector
Acide abscissique	ABREATCONSENSUS ABREATRD22 ABRELATERD1 ABREATCAL ACGTABREMOMFA2OSEM DPBFCOREDC3 MYB1AT* MYB2CONSENSUAT* MYBATRD22* MYCCONSENSUAT*	P\$ABRE P\$DPBF* P\$MYCL* P\$CE1F P\$CE3S
Acide gibbérellique	WRKY71OS* MYBGAHV* PYRIMIDINEBOXHVEPB1 PYRIMIDINEBOXOSRAMY1A GAREAT GADOWAT CARE TATCAOSAMY	P\$WBXF P\$MYBL*
Aux/SA	ASFIMOTIFCAMV	P\$OCSE
Ethylène	AGCBOXNPGLB	P\$GCCF

Tableau VI : Liste des éléments *cis*- et *trans*-régulateurs de réponse aux hormones identifiés dans le promoteur de *VvMyb5a* par les outils de prédiction PLACE et MatInspector.

Ce tableau récapitule le nom des motifs attribués par chacun des outils de prédiction. Les motifs ayant des noms différents mais correspondant aux mêmes séquences consensus sont indiqués en gras. Les astérisques indiquent des sites de reconnaissance de facteurs de transcription. Les abréviations sont Aux pour auxine et SA pour acide salicylique.

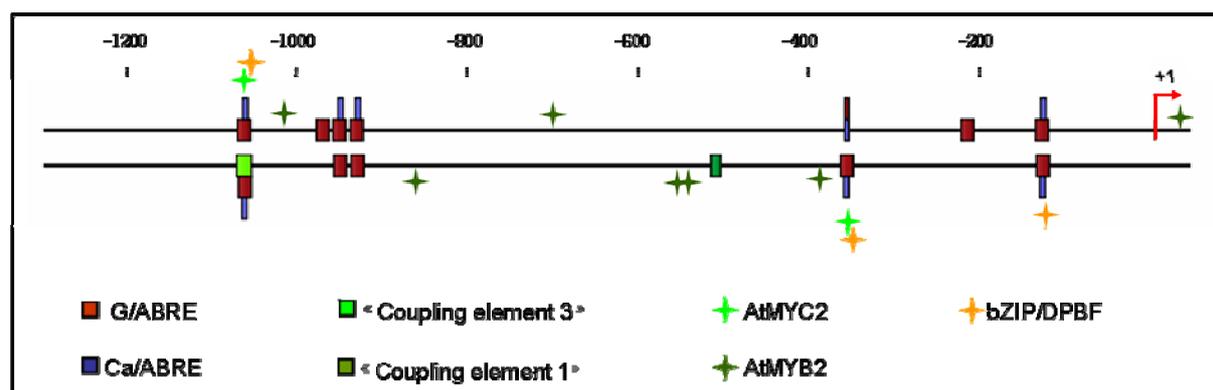


Figure 38 : Représentation schématique de la position des éléments de réponse à l'acide abscissique identifiés dans le promoteur *VvMyb5a*.

Les motifs *cis*-régulateurs sont représentés par des rectangles et les sites reconnus par des facteurs de transcription par des étoiles. Pour plus de lisibilité, les deux brins du promoteur sont indiqués par une double ligne, la ligne supérieure représentant le brin +.

faux positifs. Ainsi, seuls les types de motifs communs (lumières, sucres, ABA) identifiés par PLACE et MatInspector ont été considérés. Selon les motifs prédits par les outils utilisés, la régulation génique de *VvMyb5a* semble être majoritairement sous contrôle des hormones, des sucres et de la lumière.

b. Boîtes de réponses aux hormones

PLACE et MatInspector ont prédit des motifs *cis*- et *trans*-régulateurs communs susceptibles de réguler le promoteur *VvMyb5a* par des hormones: l'acide abscissique (ABA), l'acide gibbéréllique (GA), l'éthylène, l'auxine et l'acide salicylique (SA) (tableau VI).

Motifs de régulation par l'ABA

Les motifs ABRE (ABA Responsive Element) identifiés par les logiciels sont regroupés dans trois régions promotrices de *VvMyb5a* bien distinctes (figure 38). Douze boîtes ABRE présentes dans le promoteur correspondent à des motifs de type G/ABRE. Les G/ABRE possèdent une séquence consensus “(C/T)ACGTGGC” proche de celle des G-box et ont été identifiés dans de nombreux promoteurs de gènes régulés par l'ABA et/ou par des stress biotiques (sécheresse, température, stress salin) [292]. La fonction des motifs ABRE peut être spécifiée car la base de données PLACE a identifié les motifs ABRELATERD1 et ABREATRD22. Ces deux motifs ABRE interviennent dans l'activation des gènes *erd1* (Early Responsive to Dehydratation) et *rd22* (Responsive to Dehydratation) suite à un stress hydrique [293]. Deux autres motifs différents des ABRE, intervenant dans la régulation par l'ABA des gènes *HVA1* et *HVA22* chez l'Orge sont également présents dans le promoteur *VvMyb5a*. Il s'agit des motifs CE1 (Coupling Element 1) et CE3 (Coupling Element 3) [294, 295]. Ces motifs, associés à un motif ABRE, forment le complexe régulateur ABRC (ABA Response Complex) et sont nécessaires et suffisants pour activer les promoteurs des gènes *HVA1* et *HVA22* par l'ABA [292]. Des motifs ABRE (Ca/ABRE) pouvant être impliqués dans la signalisation calcique ont aussi été identifiés dans le promoteur *VvMyb5a* [296].

En dehors des motifs ABRE, des sites de liaisons pour des ABF (ABRE binding factors), des bZIP, des MYB et des MYC ont également été identifiés dans le promoteur de *VvMyb5a* [297]. Trois sites de liaisons pour une sous-famille de protéines bZIP (DPBF-1 et DPBF-2 pour Dc3 Promoter Binding Factor-1 et -2), deux sites de liaisons reconnus par le facteur AtMYC2 et sept sites de liaison pour le FT AtMyb2 ont été localisés. Les FT DPBF contrôlent l'expression du gène *Dc3* induite par l'ABA chez la carotte [298]. Les protéines AtMYC2 et AtMYB2 sont des activateurs transcriptionnels de l'expression des gènes inductibles par l'ABA en conditions de stress hydrique, comme le gène *rd22* chez

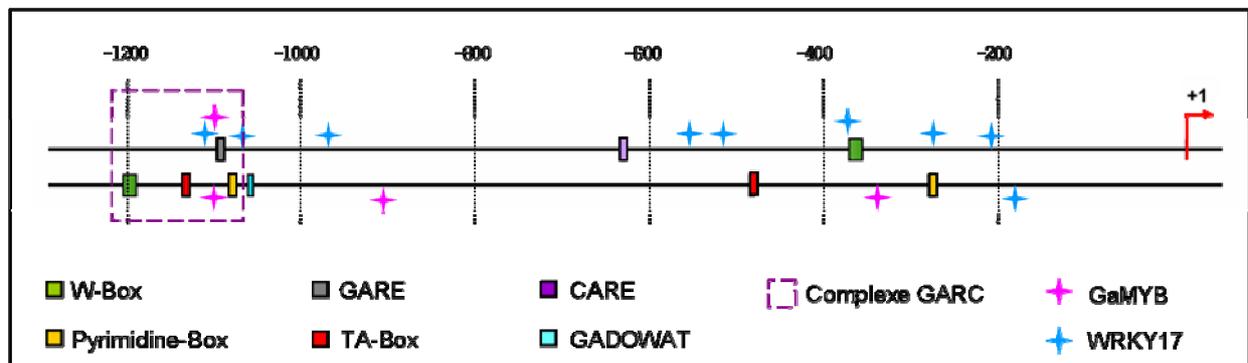


Figure 39 : Représentation schématique de la position des éléments de réponse aux gibbérellines identifiés dans le promoteur *VvMyb5a*.

Les motifs *cis*-régulateurs sont représentés par des rectangles et les sites reconnus par des facteurs de transcription par des étoiles. Pour plus de lisibilité, les deux brins du promoteur sont indiqués par une double ligne, la ligne supérieure représentant le brin +.

Arabidopsis [299, 300] [301].

Coopération de motifs cis-régulateurs pour une signalisation par les gibbérellines

Chez le riz, les α -amylases (enzymes hydrolysant l'amidon quand la germination des graines débute) sont soumises à une régulation tissu-spécifique. Dans les embryons des graines en germination, leur expression est activée par une carence en sucres et inversement. Dans l'endosperme des graines en germination, l'expression des amylases est activée par les GA. Des motifs conservés dans les promoteurs des α -amylases (α -AMY) permettent cette expression tissu-spécifique. Il s'agit des complexes GARC (Gibberellic acid responsive complex) et SRC (Sugar Response Complex) et ces derniers coopèrent pour moduler l'expression des α -AMY par les sucres et les GA [302]. GARC est un complexe régulateur capable d'activer fortement les promoteurs des α -AMY. Cet élément est constitué par des boîtes de types O2S/W-box et pyrimidine-box ("C/TCTTTT"), par un motif GARE (Gibberellic Acid Responsive Element) et enfin par une boîte TA-box ("TATCCA") [303]. Les FT pouvant interagir avec ces éléments cis-régulateurs ont été identifiés et leurs fonctions analysées : chez le blé, GaMYB se fixe sur les motifs GARE, les protéines DOF de riz se fixent sur les boîtes pyrimidines, les facteurs WRKY sur les W-box et les facteurs MYBS sur la TA-box [303]. GaMYB, FT régulé positivement par les GA active la transcription des gènes codant pour les α -amylases GA-dépendante [304]. A l'inverse, la fixation du facteur WRKY71 réprime l'expression des α -AMY GA-dépendante dans les cellules d'aleurones chez le riz [305].

Dans le promoteur du gène *VvMyb5a*, un élément GARC a été identifié (tableau VI, figure 39). Le promoteur *VvMyb5a* possède neuf sites de liaisons pour le facteur de transcription WRKY71 et deux W-box. Cependant, un seul motif cis-régulateur W-box correspond à un site de liaison du facteur *trans* WRKY71. Quatre motifs de liaisons reconnus par le facteur GaMYB ont également été identifiés par PLACE, mais un seul motif correspond à un élément GARE. Enfin, deux TA-box sont présentes, mais aucun site de reconnaissance par les MYBS n'a pu être identifié.

En dehors du complexe GARC, deux autres éléments régulateurs CARE (CAACTC regulatory elements) et GADOWAT ont été localisés. Le motif CARE est connu pour induire l'expression d'une protéinase par les GA dans les graines de riz [306]. Le motif GADOWAT a été identifié dans un ensemble de gènes régulés négativement par les GA.

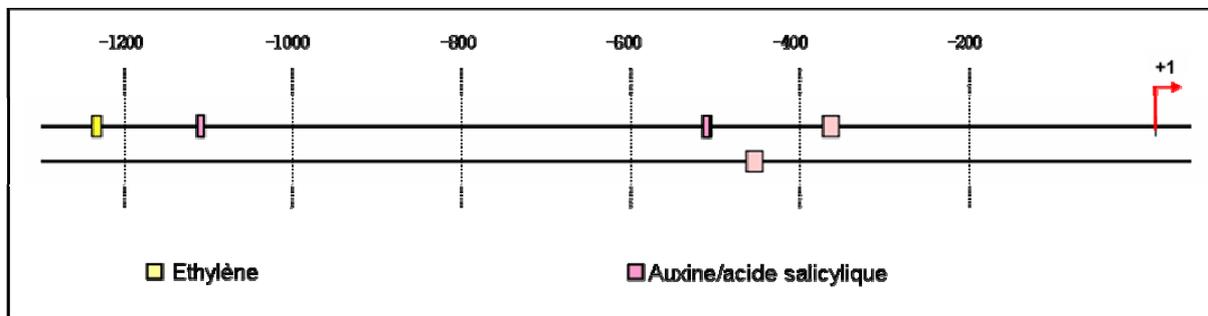


Figure 40 : Représentation schématique de la position des éléments de réponse aux hormones éthylène, auxine et acide salicylique identifiés dans le promoteur *VvMyb5a*.

Les motifs *cis*-régulateurs sont représentés par des rectangles et les sites reconnus par des facteurs de transcription par des étoiles. Pour plus de lisibilité, les deux brins du promoteur sont indiqués par une double ligne, la ligne supérieure représentant le brin +.

Régulation d'expression	Eléments régulateurs PLACE	Eléments régulateurs MatInspector
Eléments <i>cis</i> -régulateurs	SORLIP1AT SORLIP2AT TBOXGAPB GATABOX LRENPCABE IBOX IBOXCORE IBOXCORENT CAGGTGMOTIF	P\$IBOX P\$GAPB P\$LREM
Sites de reconnaissance d'éléments <i>trans</i> -régulateurs	GT1CONSENSUS GT1CORE	P\$GBOX P\$MYCL

Tableau VII : Liste des éléments *cis*- et *trans*-régulateurs impliqués dans le signal lumière identifiés par les outils de prédiction PLACE et MatInspector.

Ce tableau récapitule le nom des motifs attribués par chacun des outils de prédiction. Les motifs ayant des noms différents mais correspondant aux mêmes séquences consensus sont indiqués en gras.

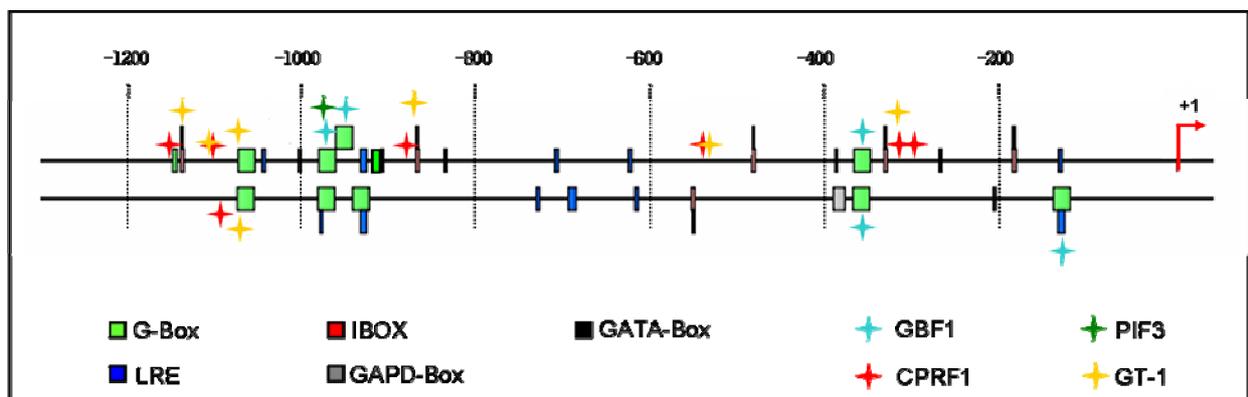


Figure 41 : Représentation schématique de la position des éléments de réponse à la lumière identifiés dans le promoteur *VvMyb5a*.

Les motifs *cis*-régulateurs sont représentés par des rectangles et les sites reconnus par des facteurs de transcription par des étoiles. Pour plus de lisibilité, les deux brins du promoteur sont indiqués par une double ligne, la ligne supérieure représentant le brin +.

Motif de régulation par d'autres hormones

Un motif correspondant à la séquence “ TGAC ” impliqué dans l'activation transcriptionnelle de plusieurs gènes par l'auxine et SA, a été identifié à 4 reprises dans le promoteur *VvMyb5a* (tableau VI, figure 40)[307].

Un seul élément de réponse à l'éthylène “ AGCCGCC ” ou GCC-box a été identifié dans le promoteur *VvMyb5a* par les outils PLACE et MatInspector (tableau VI et figure 40). Les facteurs de transcriptions EREBP (Ethylene-responsive element-binding protein), largement conservés chez les plantes, interagissent spécifiquement avec les GCC-box pour activer ou réprimer la transcription des gènes cibles [308]. Les GCC-box ont été identifiées dans les promoteurs de nombreux gènes de défense inductibles par l'éthylène et par l'acide jasmonique [309] [310].

c. Boîtes de réponse aux stress environnementaux

Motifs impliqués dans le signal lumière

L'analyse par les outils de prédiction a montré que le promoteur *VvMyb5a* possédait de nombreux motifs LRE (Light Response *cis*-Element) (tableau VII). Les motifs LRE sont essentiels au contrôle de l'activité transcriptionnelle par la lumière [311, 312]. Ces motifs ont pu être identifiés grâce à des approches classiques (délétion, mutagenèse) menées sur des promoteurs de gènes connus pour être régulés par la lumière et, plus récemment, grâce à des approches comparatives de promoteurs de gènes co-régulés ou différenciellement exprimés en fonction des conditions de lumière [313]. Malgré l'identification de nombreux motifs LRE et de facteurs de transcription associés, aucun élément simple commun à tous ces promoteurs n'a pu être mis en évidence. Ces observations ont conduit à l'hypothèse que c'est une combinaison de motifs LRE et non pas un seul de ces éléments qui serait à l'origine de la régulation de l'expression de certains gènes par la lumière.

La figure 41 indique la position des différents motifs présents dans le promoteur *VvMyb5a*. Deux types d'éléments SORLIP (Sequences Over-Represented in Light-Induced Promoters), SORLIP1 et SORLIP2, ont été détectés. Les motifs SORLIP ont été identifiés par une analyse comparative des séquences promotrices de gènes régulés par le phytochrome A [313]. D'autres motifs (G-box, I-box et GATA-box) sont aussi présents dans le promoteur de *VvMyb5a*. La combinaison de ces différentes boîtes module la perception du signal lumineux. Il a été montré que les promoteurs qui ne contiennent que des éléments G-box ou GATA-box ne sont pas sensibles à la lumière rouge [314]. Cependant, les promoteurs qui possèdent la combinaison GATA-box et G-box peuvent répondre à un spectre lumineux plus large. De la

Régulation d'expression	Eléments régulateurs PLACE	Eléments régulateurs MatInspector
Stress hydrique et froid	CBFHV* CRTDREHVCBF2*	DREB*
Toucher	AGMOTIFNTMYB2 WBOXNTERF3	AGP1

Tableau VIII : Liste des éléments *cis*- et *trans*-régulateurs impliqués dans différents stress abiotiques identifiés par les outils de prédiction PLACE et MatInspector.

Certains motifs ayant des noms différents mais correspondant aux mêmes séquences consensus sont indiqués en gras. Les étoiles indiquent les sites de reconnaissance de facteurs de transcription.

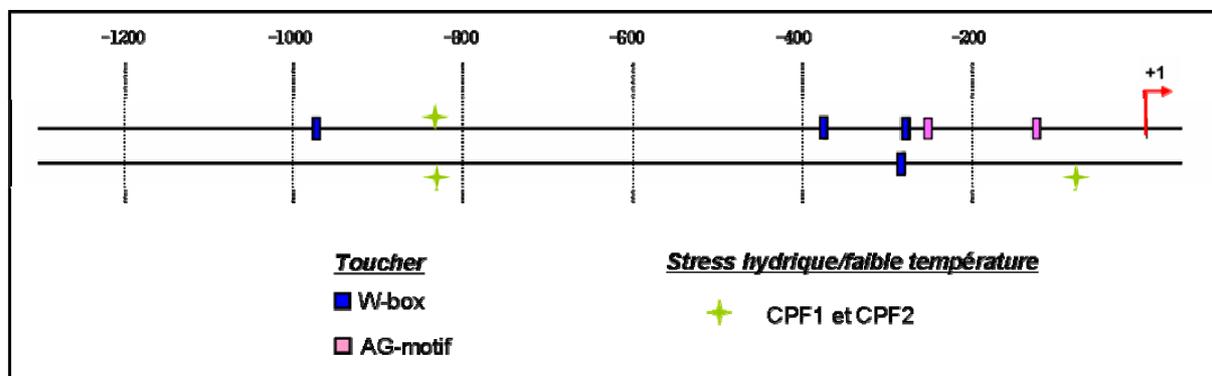


Figure 42 : Représentation schématique de la position des éléments de réponse aux facteurs abiotiques dans le promoteur *VvMyb5a*.

Les motifs *cis*-régulateurs sont représentés par des rectangles et les sites reconnus par des facteurs de transcription par des étoiles. Pour plus de lisibilité, les deux brins du promoteur sont indiqués par une double ligne, la ligne supérieure représentant le brin +.

même façon, la présence des G-box et des I-box est nécessaire et suffisante pour induire la transcription des gènes en réponse à la lumière [314]. Le promoteur de *VvMyb5a* possède un autre élément *cis*-régulateur, appelé CAGGTGMOTIF dans PLACE, qui fait partie d'un module de motifs conservés dans le gène *CHS* de Persil et qui est essentiel pour sa réponse à la lumière [315]. Un autre motif LRE appelé GAPD-box dans MatInspector ou TBOXATGGAPB dans PLACE a également été identifié. Cet élément *cis*-régulateur participe au contrôle transcriptionnel par la lumière des promoteurs des gènes *GAPDH* (glyceraldehyde-3-phosphate deshydrogenase) [316].

De nombreux FT reconnaissant les motifs LRE ont été identifiés. Ces facteurs peuvent être régulés par un type défini de lumière ou par un spectre plus large. Dans le promoteur *VvMyb5a*, des sites de fixation pour les facteurs GT-1, GBF1 (Golgi-Specific Brefeldin A resistance Factor 1) et CPRF1 (Common Plant Regulatory Factor 1) ont été identifiés (figure 41). Les sites de liaisons GT-1 présents dans de nombreux gènes régulés par la lumière interviennent aussi dans la stabilisation du complexe des facteurs généraux de la transcription TFIIA-TBP-DNA [317, 318]. La fixation de GBF1, facteur de transcription de type bZIP, sur des motifs G-box réprime ou active les gènes sensibles à la lumière bleue [313, 314]. CPRF1, dont l'expression est induite par la lumière, régule l'expression des gènes impliqués dans la photomorphogénèse en se fixant sur les séquences G-box des promoteurs cibles [314, 319]. Enfin, un site de fixation pour un FT de type bHLH appelé PIF3 (Phytochrome Interacting Factor 3) a été identifié par MatInspector. Ce facteur serait un intermédiaire transmettant les signaux précoces du phytochrome vers les gènes cibles impliqués dans le processus d'étiollement [320]. Il serait également un régulateur négatif de la photomorphogénèse [314].

Autres motifs impliqués dans les facteurs abiotiques

En dehors des motifs impliqués dans le signal lumière, l'analyse du promoteur *VvMyb5a* a révélé une forte proportion de motifs intervenant dans le stress hydrique. De nombreux motifs de réponse à l'ABA ont été identifiés dans le promoteur *VvMyb5a* (tableau VIII, figure 42). Bien que le déficit hydrique s'accompagne généralement d'une augmentation de la teneur en ABA intracellulaire, certains gènes sensibles au stress hydrique sont régulés de façon ABA-indépendante. Ces gènes possèdent le plus souvent un ou plusieurs motifs de type DRE (Dehydration Responsive Element) et DREB (Dehydration responsive element binding factors) impliqué(s) dans la réponse au stress hydrique. Ainsi, chez *Arabidopsis* la coopération entre les motifs DRE et DREB régule l'expression du gène *rd29* en réponse à la sécheresse et un fort stress salin [321]. Sur le promoteur *VvMyb5a*, aucun motif DRE n'a été identifié, mais trois sites de fixation DREB ont été localisés par les logiciels MatInspector et

Régulation d'expression	Eléments régulateurs PLACE	Eléments régulateurs MatInspector
Eléments <i>cis</i> -régulateurs	SREATMSD	SUCROSE-Box Module P\$GBOX-P\$MYBS
Sites de reconnaissance d'éléments <i>trans</i> -régulateurs	WBOXHVISO1	

Tableau IX : Liste des éléments *cis*- et *trans*-régulateurs impliqués dans la réponse aux sucres identifiés par les outils de prédiction PLACE et MatInspector.

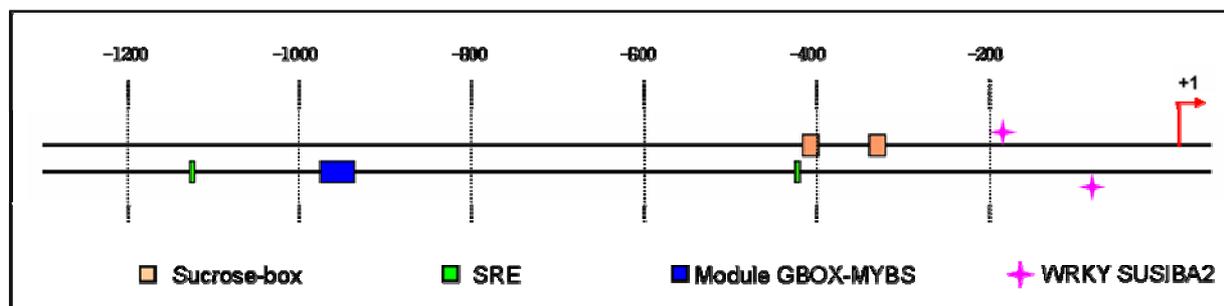


Figure 43 : Représentation schématique de la position des éléments de réponse aux sucres dans le promoteur *VvMyb5a*.

Les motifs *cis*-régulateurs sont représentés par des rectangles et les sites reconnus par des facteurs de transcription par des étoiles. Pour plus de lisibilité, les deux brins du promoteur sont indiqués par une double ligne, la ligne supérieure représentant le brin +.

PLACE (tableau VIII). Chez l'orge, ces motifs permettent la fixation des protéines CBF1 et CBF2 (C-repeat Binding Factor 1 et 2) impliquées dans la réponse au déficit hydrique et au froid [322]. De plus, les liaisons des protéines CBF2 sur les séquences d'ADN cible sont elle-même régulées par la température [323].

Le promoteur de *VvMyb5a* possède également des éléments *cis*-régulateur impliqués dans la réponse au toucher de types W-box et AG-motif (tableau VIII, figure 42). Le motif AG est un site de liaison pour le facteur de transcription à doigt de zinc de type GATA nommé AGP1 (AG-motif binding protein 1). Ce facteur active la transcription du gène de tabac *NtMYB2* inductible par le toucher [324]. Enfin, les boîtes W-box sont impliquées dans l'activation de la transcription du gène *ERF3* chez le tabac en réponse au toucher [325].

d. Boîtes de réponse aux sucres

Le programme MatInspector a localisé deux motifs SUCROSE-box identiques à ceux qui ont été localisées dans le promoteur du transporteur de monosaccharide de vigne *VvHT1*, un gène régulé par les sucres (tableau IX, figure 43) [326]. La base de données PLACE a également permis de localiser plusieurs motifs de régulation sensibles aux sucres. Il s'agit de deux motifs de régulation négatifs par les sucres SRE (Sugar Repressive Element), et de deux motifs de reconnaissance par un FT de type WRKY. Les motifs SRE ont été mis en évidence par une analyse comparative des promoteurs de gènes impliqués dans l'initiation de la formation du bourgeon auxiliaire chez *Arabidopsis*. Le motif WBOXHVISO1 est un site de liaison du facteur de transcription WRKY appelé SUSIBA2 et sa fixation sur des éléments de réponses au sucre SURE active l'expression du gène codant l'isoamylase1 (*iso1*) chez l'Orge [327].

Enfin, dans le but d'identifier des modules régulateurs connus dans le promoteur *VvMyb5a*, nous avons utilisé le programme ModelInspector (logiciel Génomatix⁵). La séquence promotrice de *VvMyb5a* a été comparée à l'ensemble des promoteurs végétaux présents dans leur base de données. Un module relatif à une régulation par les sucres, nommés GBOX-MYBS a été identifié. La présence de ce module suggère que l'expression de *VvMyb5a* pourrait être régulée par les sucres de la même manière que les α -amylase (tableau IX, figure 43).

⁵ Génomatix: <http://www.genomatix.de/index.html>

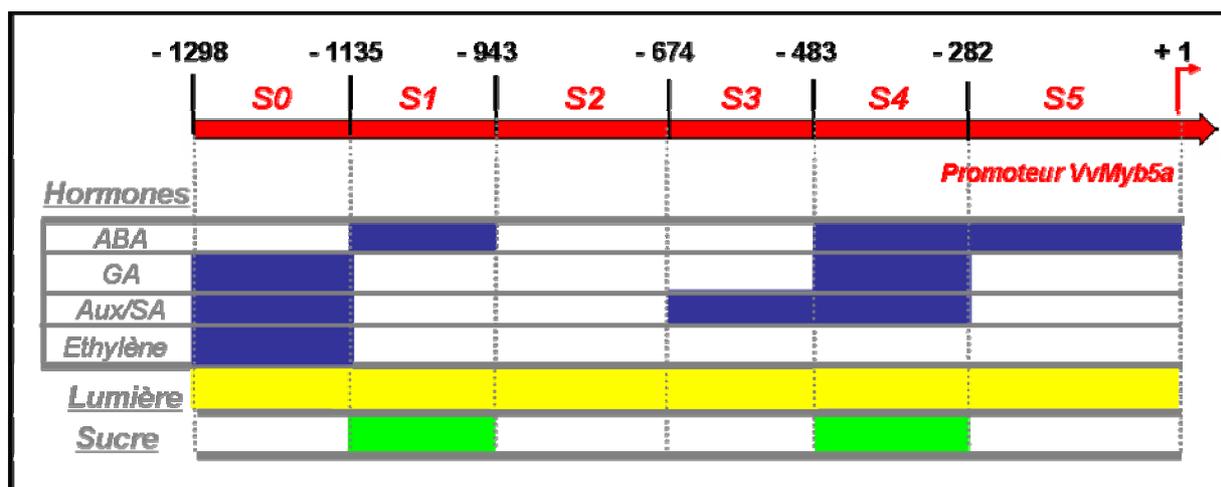


Figure 44 : Représentation schématique des régions promotrices de *VvMyb5a* impliquées dans la réponse aux hormones, aux sucres et à la lumière.

Le promoteur *VvMyb5a* (brin positif et négatif) est représenté par la flèche rouge. Le site d'initiation de la transcription est indiqué (+1). Six secteurs de réponse, notés S₀ à S₅, ont pu être identifiés sur la base de l'analyse *in silico*. Les régions promotrices répondant aux hormones sont indiquées en bleues, celles répondant à la lumière en jaunes et celles répondant aux sucres en vert. Les abréviations utilisées sont ABA (acide abscissique), GA (acide gibbéréllique), Aux (auxine) et SA (acide salicylique).

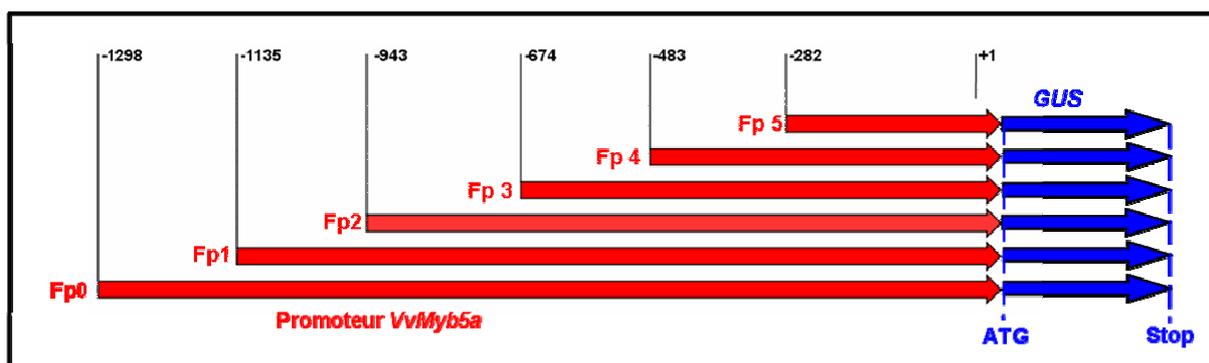


Figure 45: Fusions transcriptionnelles entre divers fragments du promoteur du gène *VvMyb5a* délété en 5' et le gène rapporteur de la β -glucuronidase (*GUS*).

Le fragment Fp0 correspond à la séquence entière du promoteur *VvMyb5a*. Le fragment Fp1 est délété de la séquence S₀ (figure 44), Fp2 est délété de S₀ et S₁ et ainsi de suite. Chacun de ces fragments a été cloné en amont de l'ATG du gène rapporteur *GUS* dans le vecteur d'expression transitoire pAM35.

4) Dissection fonctionnelle du promoteur *VvMyb5a*

L'analyse par les outils PLACE et MatInspector a révélé un nombre très important d'éléments *cis*-régulateurs potentiels. Cependant, la présence des éléments *cis*- et *trans*-régulateurs prédits dans le promoteur *VvMyb5a* ne reste que supposée. En effet, certains motifs ne possèdent seulement que quatre bases, et donc la probabilité de les retrouver dans un promoteur donné est forte. Aussi, la validité des prédictions informatiques vis-à-vis des éléments *cis*-régulateurs ne peut être affirmée qu'après la réalisation d'analyses fonctionnelles. Deux approches sont envisageables pour réaliser l'analyse fonctionnelle des promoteurs: l'expression transitoire en protoplastes et l'expression stable *in planta*. Dans les deux cas, la première étape consiste à construire des gènes chimériques où le promoteur d'intérêt, modifié ou non, est fusionné à un gène rapporteur. Il s'agit ensuite d'analyser les conséquences des modifications apportées au promoteur sur son fonctionnement. Pour des raisons de commodité et d'espace disponible dans les serres, nous avons choisi l'expression transitoire dans des protoplastes d'*Arabidopsis thaliana*.

En se fondant sur la localisation des motifs hypothétiques identifiés, une sectorisation du promoteur *VvMyb5a* a pu être réalisée. En effet, nous avons pu constater, par l'analyse *in silico*, que les boîtes étaient généralement présentes de façon préférentielle dans certaines régions promotrices, à l'exception toutefois des motifs de réponses à la lumière dispersés sur toute la séquence promotrice. Six secteurs notés S₀ à S₅ ont pu être délimités dans le promoteur *VvMyb5a* selon leurs profils de réponses hypothétiques (figure 44). Tous les secteurs possèdent des motifs *cis*- et *trans* régulateurs impliqués dans la réponse à la lumière. Concernant la réponse aux hormones, le secteur S₀ présente des motifs de réponses aux hormones GA/Aux/SA/Ethylène, S₁ à l'ABA, S₃ à l'Aux/SA, S₄ à l'ABA/GA/SA et S₅ à l'ABA. S₂ ne possède pas de motif de réponse aux hormones. De plus, S₀, S₁ et S₄ peuvent intervenir dans la réponse aux sucres et plus particulièrement S₀, caractérisé par la présence du motif de réponse négative par les sucres SRE. Enfin le segment S₅ pourrait regrouper les motifs consensus nécessaires à l'activité basale du promoteur de *VvMyb5a*.

La définition de ces secteurs a permis la réalisation d'une dissection fonctionnelle à l'aide de délétion en 5' (figure 45). Afin de visualiser le profil d'activation du promoteur en expression transitoire dans des protoplastes d'*Arabidopsis*, nous avons voulu cloner les six régions promotrices de *VvMyb5a* en amont de la séquence codante du gène rapporteur *GUS* codant la β -glucuronidase. Les six constructions ont été nommées Fp0 à Fp5 (Fragment promoteur). Fp0 correspond à une séquence promotrice *VvMyb5a* native qui avait été clonée par PCR inverse. Dans Fp1, le secteur S₀ a été délété ; dans Fp2, les secteurs S₀ et S₁ ont été délétés, et

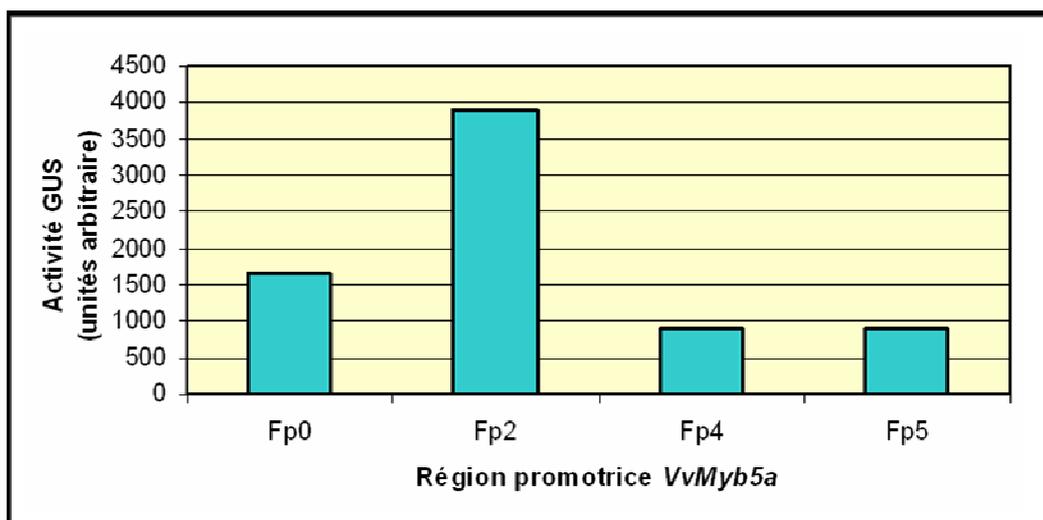


Figure 46: Activité transcriptionnelle du promoteur *VvMyb5a* et des fragments déléts en 5' dans des protoplastes *d'Arabidopsis thaliana*.

Ce graphique présente les mesures d'activité glucuronidase relevées après une heure d'incubation avec le substrat X-Gluc (5-bromo-4-chloro-3-indol-glucuronide) des protoplastes *d'Arabidopsis* transformés de façon transitoire par différentes constructions plasmidiques qui sont indiquées dans la figure 45. Les activités GUS présentées dans ce graphique correspondent à la différence entre l'activité GUS basale avant incubation et l'activité relevée après 1 heure d'incubation.

ainsi de suite. Le fragment Fp5 pourrait correspondre au promoteur minimum. Ces constructions chimériques promoteur (modifié ou non) -*GUS* nous permettront d'une part de valider les analyses *in silico* et d'autre part, de mettre en évidence les zones fortement actives du promoteur.

a. Clonage des régions promotrices dans le vecteur d'expression transitoire pAM35

Pour cloner les différentes séquences promotrices dans le MCS du vecteur d'expression transitoire pAM35 (annexe 12), des sites de restriction HindIII et PstI ont été ajoutés aux extrémités des amorces utilisées pour la réaction PCR. L'amorce antisens est identique pour tous les fragments (PrgusPstIAS), mais les amorces sens sont différentes (PrgusHindS, PrgusHindS1, PrgusHindS2, PrgusHindS3, PrgusHindS4 et PrgusHindS5) (annexe 1). Les produits d'amplifications des réactions PCR, obtenus à partir d'une matrice plasmidique pGEMT[®]-easy-prom*VvMyb5a*, ont été purifiés, clonés dans le vecteur pGEMT[®]-easy, multipliés dans les bactéries *E. coli* DH5 α et enfin séquencés. L'ADN du clone positif a été digéré par HindIII et PstI, et l'insert résultant a été cloné dans le vecteur pAM35. Les ADN des clones positifs Fp1 et Fp3 n'ont jamais pu être clonés dans le vecteur pAM35.

b. Tests préliminaires

Dans le vecteur d'expression transitoire pAM35, seul le promoteur cible intervient dans l'activation transcriptionnelle du gène *GUS*. Il a donc fallu d'une part vérifier que le promoteur *VvMyb5a* est actif dans les protoplastes d'*Arabidopsis* et d'autre part cartographier les régions promotrices pouvant conférer une forte induction de la transcription du gène *GUS* en absence de stimuli. Si aucune des constructions utilisées n'activait l'expression du gène *GUS* dans des protoplastes, alors il aurait fallu utiliser un vecteur d'expression transitoire qui possède un promoteur minimum 35S en amont du gène rapporteur *GUS*. Faute de temps, une seule série de transformation de protoplastes a été réalisée, et il convient donc de considérer les résultats obtenus comme préliminaires. Les dosages d'activité GUS ont été réalisés 16 h après transformation des protoplastes d'*Arabidopsis* (figure 46).

Les clonages des séquences Fp1 et Fp3 ayant échoué, les transformations de protoplastes ont été réalisées avec les constructions Fp0, Fp2, Fp4 et Fp5. Toutes les constructions sont suffisantes pour induire l'expression du gène rapporteur *GUS*. Cependant, le promoteur entier (Fp0) et surtout le fragment Fp2 révèlent un pouvoir activateur plus important que celui des fragments Fp4 et Fp5 (au moins deux fois plus élevé). La construction Fp2 confère l'activation transcriptionnelle la plus forte (figure 46).

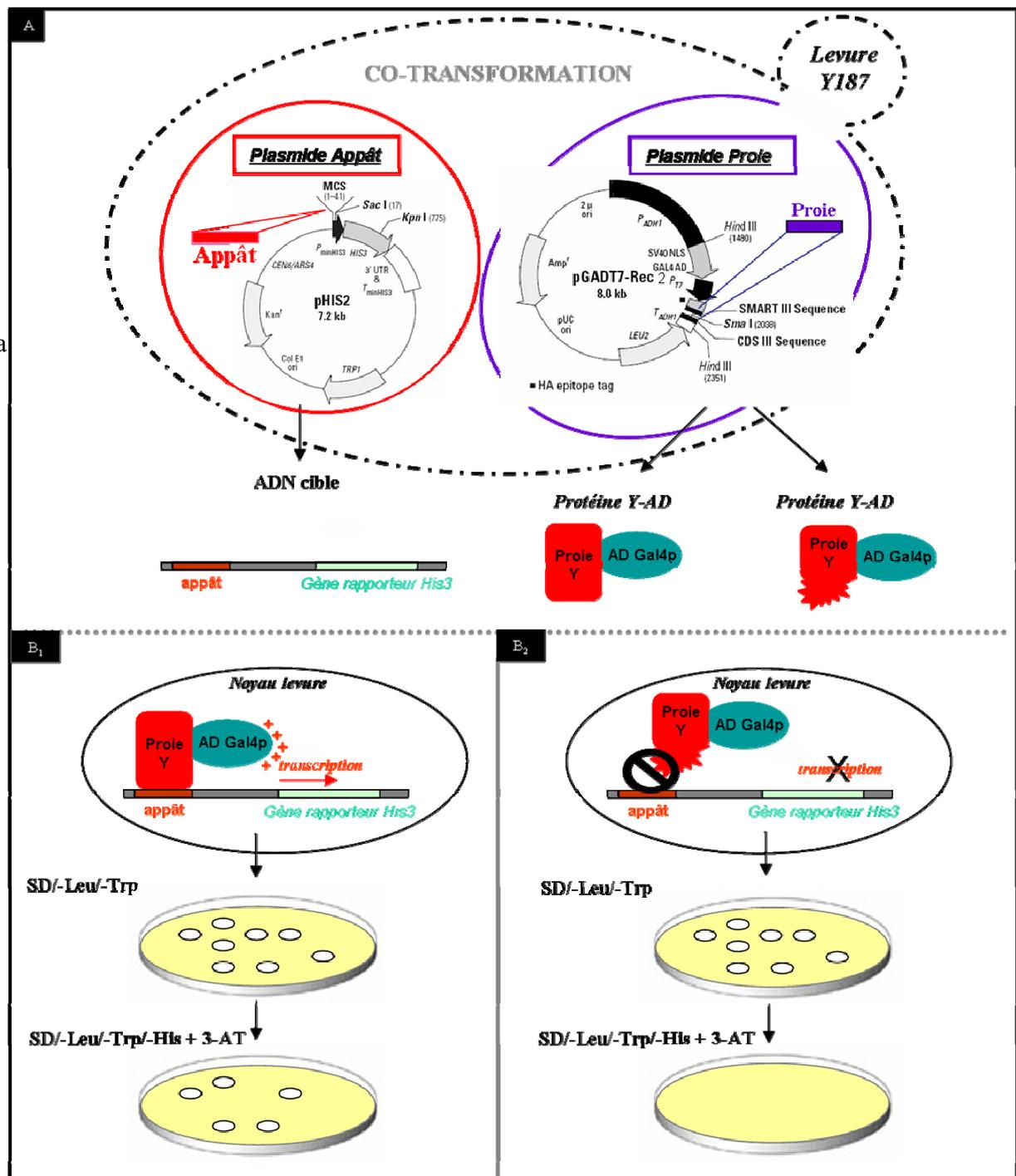


Figure 47 : Principe de la technique de criblage simple hybride chez la levure pour identifier des facteurs *trans*-régulateurs.

A) Une cellule hôte est co-transformée avec les plasmides appât et proie. Le vecteur proie pGADT7-Rec2, où ont été insérés les ADNc de la banque à cribler, présente les séquences supplémentaires suivantes: la séquence codant le domaine d'activation transcriptionnelle Gal4p (AD) précédée d'un signal de localisation nucléaire (NLS) et le gène *trp1* permettant la croissance sur milieu carencé en tryptophane. Le vecteur appât pHIS2, où a été insérée la séquence d'ADN cible présente les caractéristiques suivantes: un promoteur minimum du gène *His3* et le gène *Leu2* permettant la croissance sur milieu carencé en leucine. La protéine proie (Y-AD) est une protéine chimérique car elle est fusionnée au domaine AD de Gal4.

B) La fixation d'une protéine proie sur la séquence d'ADN cible positionne le domaine AD suffisamment proche du promoteur minimum d'*His3* pour activer la transcription du gène rapporteur *His3*. Les levures sont sélectionnées sur un milieu carencé en histidine. Le 3-AT (3-amino-1,2,4-triazole) est un inhibiteur compétitif de la protéine *His3* endogène de levure.

5) Identification d'éléments *trans*-régulateurs par la technique du Simple

Hybride chez la levure

Les FT jouent des rôles majeurs dans la régulation des gènes durant le développement de la plante à travers leur interaction avec des éléments *cis*-régulateurs et/ou d'autres FT. Ces interactions protéine-ADN ou protéine-protéine permettent de réguler finement l'expression spatio-temporellement des gènes. Afin d'identifier les acteurs moléculaires contrôlant le profil temporel d'expression de *VvMyb5a* (figure 29), nous avons voulu rechercher les facteurs *trans* qui régulent son expression. L'approche retenue pour isoler les FT qui interagissent avec les régions promotrices est la technique du simple hybride chez la levure.

a. Principe de la technique du simple hybride chez la levure

Le système simple hybride développé chez la levure *Saccharomyces cerevisiae* est une technique qui permet de détecter l'interaction physique entre des FT dits *trans*-régulateurs et un ADN cible. Son principe repose sur la capacité des FT à reconnaître les séquences d'ADN cibles *via* leur domaine de liaison à l'ADN puis à activer la transcription du gène cible *via* leur domaine d'activation.

Le principe du simple hybride en levure est illustré par la figure 47. Comme pour la technique du double hybride (figure 32), le simple hybride est un essai *in vivo* d'interaction qui repose sur l'utilisation du facteur de transcription Gal4p. Seul AD de Gal4p est utilisé dans la technique du simple hybride : il est fusionné aux protéines proies pour permettre la production de protéine chimère X-AD. Dans ce système, l'appât (Y) est constitué soit par une région promotrice, soit par des motifs *cis*-régulateurs. L'ADN cible est cloné en amont d'un promoteur minimum qui contrôle la transcription du gène rapporteur. Quand X interagit physiquement avec Y, alors AD de la protéine X est suffisamment proche du promoteur minimum pour induire la transcription du gène rapporteur sélectif.

Expérimentalement, les levures sont transformées simultanément par deux plasmides. Le premier permet d'introduire la séquence d'ADN cible appât (dont on veut identifier les facteurs *trans*-régulateurs). Dans le système Matchmaker (Clontech), l'ADN cible est cloné en amont du promoteur minimum du gène rapporteur *HIS3*. Le deuxième plasmide permet de produire les protéines proies (facteurs *trans* potentiels) fusionnées au domaine AD de Gal4 (aa 768 à 788). Les plasmides appât (pHIS2, annexe 9) et proie (pGADT7-Rec2) portent respectivement les gènes *trp1* et *leu2* permettant ainsi aux levures co-transformées de se développer en milieu de croissance carencé en tryptophane et leucine (SD/-Leu/-Trp). L'interaction entre l'appât et un FT potentiel active l'expression du gène rapporteur

permettant ainsi aux levures co-transformées de se développer sur un milieu dépourvu d'histidine (SD/-Leu/-Trp/-His) (figure 47).

b. Clonage des régions promotrices cibles dans le vecteur d'expression pHIS2

Pour identifier les acteurs moléculaires de l'expression de *VvMyb5a* et confirmer les analyses *in silico*, nous avons décidé d'utiliser cinq régions promotrices comme appât correspondant aux secteurs S0 à S5 (figure 44). Pour cloner les cinq fragments dans le MCS de pHIS2, les sites de restrictions SacI et EcoRI ont été ajoutés respectivement aux extrémités des amorces sens et antisens. Le nom et la séquence des amorces utilisées sont indiqués dans l'annexe 1. Chacun des fragments d'ADN, d'une taille d'environ 200 pb, a été amplifié à partir de la matrice plasmidique pGEMT[®]-easy-prom*VvMyb5a*. Les produits d'amplification ont été purifiés, clonés dans le vecteur pGEMT[®]-easy, multipliés dans des bactéries DH5α et séquencés. Après vérification, les vecteurs pGEMT[®]-easy-S₀, S₁, S₂, S₃, S₄ et S₅ ont été digérés et les produits digérés clonés entre les sites SacI et EcoRI du MCS du vecteur pHIS2. Le fragment S₄ fut le premier cloné dans le vecteur pHIS2 et nous avons donc débuté l'analyse simple hybride avec la construction appât pHIS-S₄.

c. Tests préliminaires

La réussite du système simple hybride repose sur le fait qu'aucun facteur de transcription endogène de levure ne doit reconnaître la séquence d'ADN appât. Pour cette raison, il est important de quantifier au préalable l'activité *his3* obtenue avec l'appât en absence de protéines proies.

Le plasmide appât (pHIS-S₄) a été transféré dans une souche de levure modifiée Y187 qui est His⁻, Leu⁻ et Trp⁻. De fait, Y187 ne peut pas se développer sur un milieu dépourvu de ces aa. Les plasmides appât et proie portent les gènes *trp1* et *leu2*, permettant ainsi aux levures transformées de se développer en milieu de croissance carencé en tryptophane (SD/-Trp) et en leucine (SD/-Leu). L'interaction entre la séquence promotrice cible et un partenaire protéique potentiel active l'expression du gène *His3* et permet aux levures co-transformées de se développer également en milieu de croissance dépourvu d'histidine (SD/-Leu/-Trp/-His). Les levures Y187 ont donc été transformées avec le plasmide pHIS-S₄ seul. Au bout de 5 j, certaines levures s'étaient développées sur un milieu dépourvu d'histidine, car la protéine HIS3 de levure était capable de se fixer sur le fragment S₄ du promoteur *VvMyb5a* et d'induire la transcription du gène sélectif *His3*. Pour diminuer ce bruit de fond, les milieux sélectifs doivent être supplémentés d'un inhibiteur compétitif de la protéine HIS3 de levure,

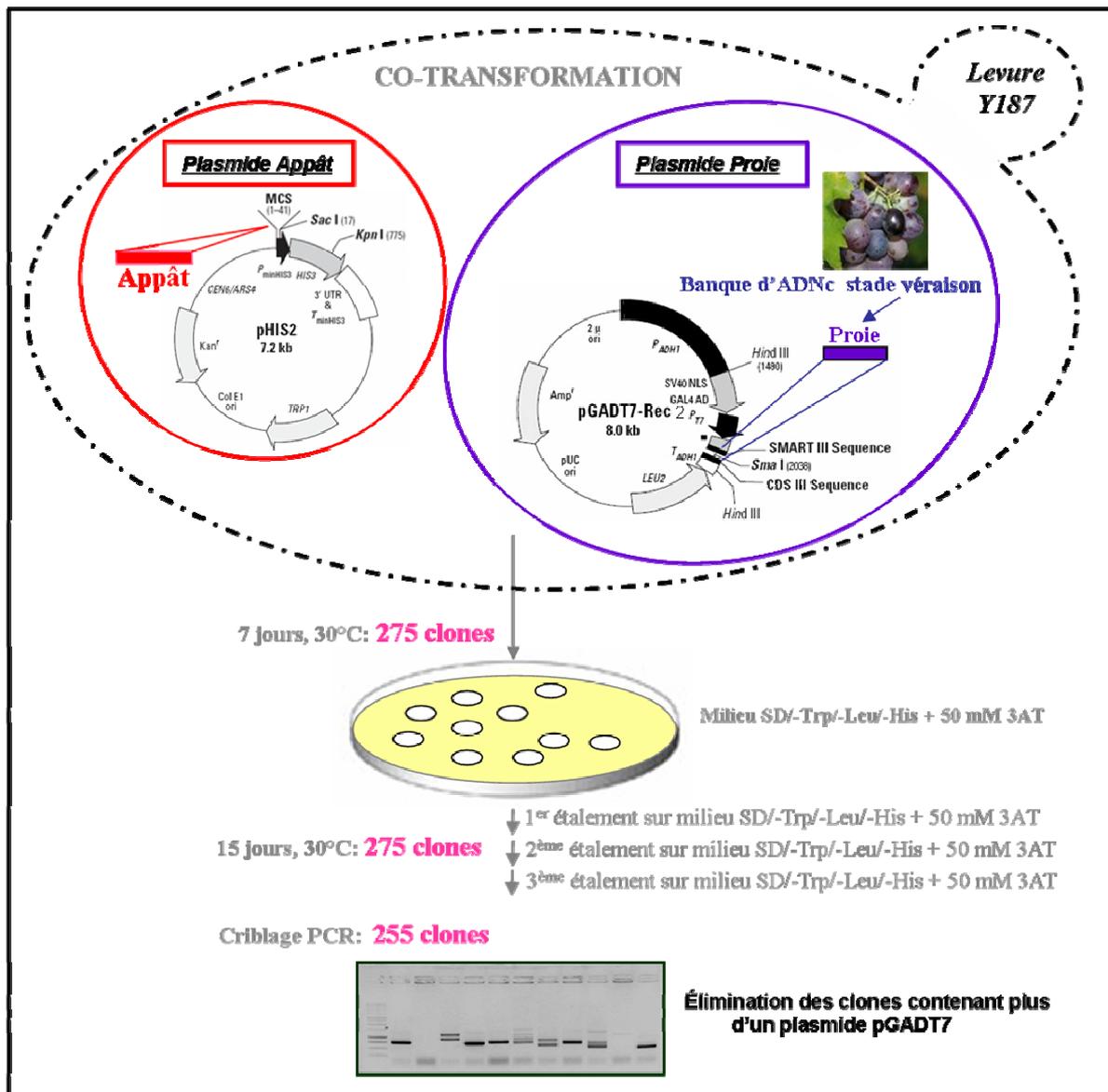


Figure 48 : Représentation schématique du crible effectué pour identifier les clones positifs en simple hybride contre le fragment S₄ de *VvMyb5a*.

Les levures Y187 sont co-transformées avec les plasmides appât (pHIS2-S₄) et proies (pGADT7-Rec2-ADNc de la banque). L'appât est le fragment S₄ du promoteur de *VvMyb5a* (figure 44). La banque utilisée pour ce crible a été amplifiée à partir d'ARNm extraits de baies du cépage Cabernet sauvignon récoltées au stade véraison et épépinées. Les levures co-transformées dans lesquelles il y a fixation de la proie sur le fragment S₄ sont sélectionnées sur milieu SD/-His/-Trp/-Leu + 50 mM 3-AT. Au bout de 7 jours, les levures qui s'étaient développées, ont été ré-étagées 2 fois sur le même milieu sélectif. Après le crible nutritionnel, un criblage par PCR a été effectué afin d'éliminer les clones ayant intégré plus d'un plasmide proie. En parallèle, les contrôles positifs et négatifs du kit Matchmaker (Clontech) ont été utilisés. Le 3-AT (3-amino-1,2,4-triazole) est un inhibiteur compétitif de la protéine His3 endogène de levure.

le 3-AT (3-amino-1,2,4-triazole). Pour déterminer la concentration optimale en 3-AT, les levures Y187 ont été transformées avec pHIS- S_4 puis étalées sur un milieu SD/-His/-Trp contenant 10, 20, 30, 40 ou 50 mM de 3-AT. Au bout d'une semaine, seules les levures étalées sur le milieu SD/-His/-Trp contenant 50 mM de 3-AT ne se sont pas développées. En conséquence, le criblage de la banque a été réalisé avec des milieux contenant systématiquement 50 mM de 3-AT.

d. Criblage de la banque

Comme pour le double hybride, la procédure de criblage choisie a été la co-transformation des levures Y187 avec la construction appât et les plasmides proies (pGADT7-Rec-ADNc) dans lesquels s'insèrent les ADNc représentatifs de la banque par recombinaison homologe (annexe 10). Pour le criblage de partenaires protéiques, nous avons utilisé une banque d'ADNc générée à partir d'ARNm extraits de baies de raisin récoltées au stade véraison et épépinées. Les levures co-transformées ont été étalées sur le milieu sélectif SD/-Leu/-Trp/-His+50mM 3-AT. Le nombre de colonies sur ces boîtes a été compté après 5 et 7 j d'incubation à 30°C. Après 5 j d'incubation, 43 colonies s'étaient développées et après 7 j, 232 colonies. Parmi ces 275 colonies, 75 colonies étaient de petite taille. Pour éliminer les faux positifs, plusieurs cribles ont été effectués, et les résultats obtenus sont résumés dans la figure 48. Ainsi, les 275 clones ont été étalés successivement deux fois sur un milieu SD/-Leu/-Trp/-His+50mM 3-AT. Aucun clone n'a pu être éliminé par ces étalements successifs. Un criblage par PCR a été réalisé afin d'éliminer ceux qui contenaient plus d'un plasmide proie. Les ADNc de la banque codant pour des facteurs *trans* potentiellement interacteurs ont été amplifiés par des amorces spécifiques du vecteur pGADT7 (T7DH et 3AD) et les produits d'amplification déposés sur un gel d'agarose (1,2%). Les clones contenant plus d'un plasmide pGADT7-Rec2 et ceux pour lesquels la taille du produit d'amplification était d'environ 250 pb (indiquant que le vecteur est vide) ont également été éliminés. Vingt clones ont été éliminés à l'issue de ce criblage PCR.

e. Résultats

Pour déterminer la nature des gènes présents dans chaque clone positif, nous avons entrepris d'extraire les plasmides proies. Cependant, par manque de temps, seul dix-huit plasmides ont pu être transférés dans les bactéries et envoyés à séquencer. Les résultats sont présentés dans le tableau X. L'analyse des séquences a été effectuée par le programme BLAT disponible sur le site web du Génoscope ou par le programme BLAST du site NCBI.

N° Clone	Taille du fragment PCR (pb)	En phase avec AD Gal4?	Taille ORF proie	Références Génoscope	Taille de la protéine Génoscope	Positionnement du peptide proie sur la protéine du Génoscope	Identité hypothétiques du gène
SH2	970 pb	Vecteur vide					
SH3	762 pb	Non	122 aa	GSVIV/P00032598001	122 aa	1-122	ubiquinol-cytochrome C reductase complex 14 kDa protein
SH6	826 pb	Oui	121 aa	GSVIV/P00028970001	121 aa	1-121	thioredoxin H
SH7	548 pb	Oui	83 aa	GSVIV/P00025163001	804 aa	1-83	photosystem I assembly protein ycf3
SH10	549 pb	Oui	69 aa	GSVIV/P00025451001	605 aa	536-605	Eukaryotic translation initiation factor 3C
SH13	1006 pb	Oui	190 aa	GSVIV/P00033164001	190 aa	1-190	VHL binding protein
SH14	359 pb	Vecteur vide					
SH17	914 pb	Oui	219 aa	GSVIV/P00018735001	965 aa	747-965	60S ribosomal protein L8 (RPL8C)
SH18	934 pb	Oui	816 aa	GSVIV/P00026101001	816 aa	1-816	PEP carboxylase (PEPC) (Vitis vinifera)
SH19	946 pb						
				Protéine de levure			
SH20	822 pb	Oui	103 aa	GSVIV/P00025451001	309 aa	502-605	Eukaryotic translation initiation factor 3C
SH21	893 pb	Oui	119 aa	GSVIV/P00013061001	421 aa	302-421	G protein-coupled receptor-like protein
SH22	540 pb	Oui	87 aa	GSVIV/P00004134001	545 aa	458-545	SCARECROW gene regulator
SH23	577 pb	Oui	113 aa	GSVIV/P00025871001	173 aa	32-145	putative ripening-related protein Grip 31 (Vitis vinifera)
SH25	912 pb	Oui	158 aa	GSVIV/P00033076001	158 aa	1-159 aa	pathogenesis-related protein 10 (Vitis vinifera)
SH26	974 pb						
SH27	905 pb						
				Vecteur vide			
SH30	631 pb	Non	74 aa	/	/	/	Ubiquitin-like protein

Tableau X: Eléments *trans*-régulateurs interagissant avec le promoteur *VvMyb5a*, identifiés par la technique du simple hybride chez la levure.

Ce tableau présente les facteurs de transcription identifiés par la technique du simple hybride pouvant potentiellement réguler le promoteur *VvMyb5a*. Le numéro du clone, la taille du fragment PCR amplifiées par les amorces T7DH et 3AD et les peptides proies en phase avec le domaine d'activation de Gal4p sont indiqués. Les références ont été obtenues par des blastn et blastp réalisés dans la base de données de *Vitis vinifera* (Génoscope). Un alignement a été réalisé entre les peptides proies et les protéines du Génoscope correspondantes pour s'assurer que les ADNc étaient clonés dans le bon cadre de lecture. Quand elles n'étaient pas annotées dans le Génoscope, les fonctions hypothétiques des protéines proies ont été identifiées par le logiciel Blast dans NCBI.

En dépit des cribles effectués, quatre clones se sont avérés ne pas contenir d'insert. Un autre clone a été éliminé car il codait une protéine endogène de levure, une déshydrogénase. Un seul facteur de transcription, nommé SCARECROW (SCR), a pu être identifié lors de ces premiers essais de séquençage. Le gène *SCR* code un facteur de transcription de la famille GRAS (Gibberellin-insensitive Repressor of gal-3 Sacrecrow). Ce gène est exprimé pendant l'embryogenèse au niveau des cellules qui deviendront le centre quiescent (QC). Son expression a également été détectée dans les cellules initiales impliquées dans la formation et la différenciation du cortex et de l'endoderme [328, 329].

I- Discussion et perspectives

Régulation transcriptionnelle du gène *VvMyb5a*

Avant le début de ce travail de thèse, la caractérisation fonctionnelle de deux facteurs de transcription MYB, *VvMyb5a* et *VvMyb5b*, avait été engagée au laboratoire par L. Deluc [249, 250]. Ces travaux avaient permis de montrer l'implication de *VvMyb5a* et *VvMyb5b* dans la régulation de la voie de biosynthèse des phénylpropanoïdes chez la vigne. Des analyses du profil d'expression par RT-PCR semi-quantitative ont également mis en évidence une régulation spatiale et temporelle de l'expression du gène *VvMyb5a* au cours du développement de la baie de raisin. *VvMyb5a* est exprimé préférentiellement durant la première phase de développement de la baie, et ce quelque soit le tissu considéré. Juste après floraison, les transcrits *VvMyb5a* sont abondants, puis leur quantité diminue au cours de la phase herbacée jusqu'à la véraison. Durant la phase de maturation, *VvMyb5a* est très faiblement exprimé dans la pellicule et les graines et il n'est pas détectable dans la pulpe. Une partie de mon travail de thèse a donc porté sur l'identification des facteurs responsables de la régulation de l'expression de *VvMyb5a*. Un fragment d'environ 1300 pb correspondant au promoteur de *VvMyb5a* a été cloné par la technique de PCR inverse. L'analyse *in silico* réalisée avec les logiciels PLACE et MatInspector a révélé la présence d'un promoteur proximal ainsi que de nombreux éléments *cis*-régulateurs potentiels. L'analyse par le logiciel PLACE est fondée sur des homologies avec des séquences consensus de 4 à 8 bases. Le logiciel MatInspector recherche des sites à partir de matrices de poids et estime leur probabilité d'apparition. Dans un souci de rigueur, seuls les éléments *cis*-régulateurs communs identifiés par les deux logiciels ont été analysés. Il convient toutefois de considérer avec précaution ces éléments *cis*-régulateurs présumés car leur identification repose sur de courtes homologies de séquence nucléique. Cette analyse bioinformatique peut être considérée comme une première approche visant à identifier un éventail de motifs régulateurs

de la transcription dans un promoteur, mais la fonctionnalité et l'activité de ces motifs devront être confirmées expérimentalement. À l'heure actuelle, les motifs de régulation hypothétiques identifiés dans le promoteur de *VvMyb5a* suggèrent que l'expression de ce gène pourrait être régulée principalement par les hormones et les sucres. Durant la phase herbacée du développement des baies, l'auxine, les cytokinines et les GA sont les seules hormones présentes dans la baie (figure 9). Au stade vert qui suit la floraison, l'auxine est quasiment la seule hormone détectable. C'est durant ces premières étapes de développement que le niveau d'expression de *VvMyb5a* est le plus fort dans la baie. La présence dans le promoteur de *VvMyb5a* d'éléments *cis*-régulateurs potentiels impliqués dans l'activation transcriptionnelle des gènes en réponse à l'auxine (tableau VI et figure 40) suggère que cette hormone pourrait induire l'expression de ce gène durant les stades précoces de développement de la baie. À partir du stade mi-vert, la quantité des cytokinines et de GA augmente fortement, parallèlement à la diminution de la quantité d'auxine. L'analyse *in silico* du promoteur de *VvMYB5a* a également permis d'identifier un complexe GARC et plusieurs motifs CARE, impliqués dans la régulation positive de l'expression des gènes en réponse aux GA (tableau VI, figure 39). La forte expression de *VvMyb5a* pendant le stade herbacé du développement des baies pourrait être liée à l'action conjuguée de l'auxine et des GA. À partir de la véraison, la quantité de transcrits *VvMyb5a* diminue fortement et reste très faible jusqu'à la maturité des baies. Pendant cette phase de maturation, les sucres et l'ABA s'accumulent progressivement dans les baies alors que l'auxine et les GA ne sont plus détectables. La faible expression de *VvMYB5a* après la véraison pourrait donc être liée à l'absence d'auxine et de GA combinée à une répression de l'expression par les sucres. En effet, un motif de régulation négatif par les sucres a été identifié dans le promoteur de *VvMyb5a* (tableau IX, figure 43) et les premières analyses d'activation transcriptionnelle du gène *GUS* dans des protoplastes d'*Arabidopsis thaliana* ont mis en évidence un rôle répresseur de la transcription de la région contenant ce motif (figure 47).

Cette analyse bioinformatique semble donc indiquer que l'expression de *VvMyb5a* serait principalement sous contrôle hormonal durant le développement de la baie. Cependant, tant que ces données n'auront pas été confirmées expérimentalement, ces conclusions restent hypothétiques. De ce fait, il serait intéressant d'utiliser les différentes constructions promoteur-*GUS* (FP₀ à FP₅) maintenant disponibles pour vérifier l'activation du promoteur *VvMyb5a* en réponse à différents traitements hormonaux. Une approche d'expression transitoire dans des protoplastes de cellules de vigne semble la plus adaptée pour mener rapidement à bien ces expériences. De plus, afin d'évaluer la participation de chaque motif de régulation potentiel identifié, des expériences de mutagenèse dirigée pourront être effectuées.

Dans un premier temps, le motif répresseur SRE (figure 44) pourrait représenter une cible intéressante pour ce type d'approche

Enfin, une étude de l'impact des processus écophysiologicals et climatiques sur la régulation de *VvMyb5a* dans la baie de raisin pourrait également être menée puisque des boîtes de réponse à la lumière ont été identifiées (tableau VII, figure 41). En effet, il est connu que les conditions environnementales modifient la composition des flavonoïdes dans les baies de raisin [330]. Ces changements sont dépendants du cépage mais d'une manière générale la température, l'ensoleillement et l'intensité du rayonnement solaire influencent l'accumulation des flavonoïdes [331]. Des températures supérieures à 32°C ou inférieures à 15°C ont un impact négatif sur l'accumulation des anthocyanes et des températures nocturnes trop élevées provoquent des perturbations de la synthèse des composés phénoliques [332]. La lumière influence positivement l'accumulation d'anthocyanes [333-335]. De plus, l'alternance jour/nuit est un facteur majeur dans l'accumulation des anthocyanes. Les conditions idéales sont une égalité de la durée jour/nuit et des températures comprises entre 20 et 25°C [336]. Dans le cadre du nouveau projet de recherche sur "les régulations trophiques et microclimatique du métabolisme secondaire dans la baie de raisin", qui a été mis en place cette année dans le laboratoire, il serait intéressant d'analyser par PCR quantitative l'expression de *VvMyb5a* à partir des différents échantillons : baies entières, demi-baies et pellicules ayant été exposées à des contraintes thermiques et/ou lumineuses variées en serre ou au vignoble.

Cependant, même si plusieurs études ont montré que les conditions climatiques modifient la composition et les teneurs en flavonoïdes, les mécanismes intrinsèques moléculaires et cellulaires à l'origine de ces perturbations ne sont pas encore connus. A ce stade, il nous apparaît donc important de signaler un élément relatif aux mécanismes de régulation de l'expression de *VvMYB5a*, et plus globalement de la plupart des gènes *MYB* de vigne identifiés dans notre laboratoire. En effet, les études menées en collaboration avec plusieurs laboratoires étrangers (CSIRO Plant Industry en Australie et Pontificia Universidad Católica au Chili) indiquent une surprenante conservation des profils d'expression des gènes *MYB* au cours du développement des baies de raisin. Ainsi, les gènes *VvMyb5a*, *VvMYB5b*, *VvMybPA1* et *VvMyb24* (chapitre 2) présentent des profils d'expression tout à fait similaires dans des baies de Cabernet sauvignon cultivées dans le bordelais et dans des baies du cépage Shiraz récoltées dans les vignobles de la région d'Adélaïde, en Australie [250, 262]. De même, dans des baies de Cabernet sauvignon cultivées à Santiago au Chili, l'expression de *VvMyb24* est tout à fait comparable à celle que nous avons observée dans le même cépage en région bordelaise (figure 72 A/B). Cette forte conservation des profils d'expression en dépit

de conditions environnementales et de pratiques culturelles disparates semble indiquer une faible influence des facteurs abiotiques sur l'activité de ces gènes *MYB* dans les baies. Ces observations pourraient indiquer un contrôle à dominante développementale de l'expression de ces gènes, et apparaissent en accord avec les données issues de l'analyse du promoteur *VvMyb5a*. Si cette hypothèse se confirmait, l'impact des paramètres environnementaux sur la teneur en flavonoïdes des baies n'impliquerait pas les gènes régulateurs de la famille MYB mais plutôt les réactions de biosynthèse, la stabilité des enzymes et des métabolites produits, ou encore les mécanismes de stockage de ces composés. Par exemple, certaines études ont montré que les molécules d'anthocyanes pouvaient être sujettes à dégradation dans des conditions de température élevée [337].

Les approches de criblage à grande échelle...

En complément de l'analyse bioinformatique menée pour identifier les éléments pouvant réguler l'expression de *VvMyb5a*, une recherche des FT pouvant se fixer sur ces éléments régulateurs a été effectuée par une approche simple hybride. L'ensemble de la séquence promotrice de *VvMyb5a* devait être criblé, mais l'analyse a débuté avec le fragment S₄ situé entre les nucléotides -715 et -524 pb. Sur cette courte séquence de 200 pb, plus de 250 clones potentiellement positifs ont été identifiés. Le criblage PCR n'a pas permis d'éliminer les clones potentiellement redondants et l'étape d'extraction des plasmides a alors été entreprise. Que ce soit pour le double ou simple hybride, nous avons rencontré des problèmes pour l'extraction des plasmides proies des clones de levures. Une fois le protocole d'extraction mis au point, les plasmides doivent être transférés dans des bactéries, criblés, puis l'insert doit être séquencé. Le transfert des plasmides dans les bactéries est une étape longue et, à l'heure actuelle, seule une trentaine de clones ont pu être séquencés. Ces clones ne seront considérés comme positifs qu'une fois que leur capacité à auto-activer les gènes rapporteurs chez la levure aura été à nouveau testée.

Les études menées sur le promoteur du gène *VvMyb5a* devaient nous permettre de mieux comprendre ses mécanismes de régulation transcriptionnelle. Cependant, il était également important d'étudier les propriétés de la protéine VvMyb5a en identifiant les protéines capables d'interagir avec elle et éventuellement de moduler son activité de contrôle de la transcription des gènes de la voie de biosynthèse des flavonoïdes. Nous nous sommes particulièrement intéressés à un domaine riche en glutamine (GRD) présent dans la région C-terminale de VvMyb5a. Peu de données étant disponibles dans la littérature quant à la fonction potentielle de ce type de domaine, nous avons décidé d'établir une carte de l'interactome du domaine GRD de VvMyb5a par la technique de double hybride chez la

levure. Ce système présente l'avantage d'étudier *in vivo* des interactions protéiques transitoires ou instables, et ce indépendamment du niveau d'expression naturel des protéines [276]. Cette approche de criblage à grande échelle a permis l'identification de nombreuses protéines potentiellement capables d'interactions physiques avec VvMyb5a. Ces protéines apparaissent impliquées dans diverses fonctions et localisées dans divers compartiments intracellulaires (tableau IV). De fait, il s'avère difficile de conclure quant au degré de probabilité de ces interactions. Dans un premier temps, il conviendrait d'analyser la localisation cellulaire de la protéine VvMyb5a par des expériences d'expression transitoire avec des constructions VvMyb5a-GFP. En effet, en tant que facteur de transcription, la localisation de VvMyb5a est présumée nucléaire, mais il est possible que cette protéine puisse être adressée ou transiter dans d'autres compartiments intracellulaires. Chez l'homme, il a été montré récemment qu'une protéine MYB, Mybbp1a (Myb-binding protein 1a), pouvait transiter par les pores nucléaires pour passer du noyau au cytoplasme [338].

Toutefois, et même si un grand nombre de clones isolés lors du crible double hybride semble correspondre à des faux positifs en raison de la nature et de la localisation intracellulaire hypothétique des protéines codées par ces clones, deux ont retenu notre attention. Il s'agit d'une protéine kinase et d'une protéine de type WD40:

- La protéine kinase identifiée dans cette expérience présente de fortes homologies de séquence avec la protéine kinase KGM (Kinase associated with GAMYB) de l'orge [280]. Dans les cellules de l'orge, KGM est localisé dans le noyau et le cytosol et des expériences de double hybride ont permis de mettre en évidence que KGM interagissait physiquement avec GAMYB pour réprimer son activité. Le facteur de transcription GAMYB active les gènes hydrolytiques qui interviennent dans la mobilisation des réserves de sucres au cours du développement des plantules. Cependant, même si l'interaction physique entre ces deux protéines a pu être démontrée, aucun essai *in vitro* n'a pu mettre en évidence la phosphorylation de GAMYB par la kinase KGM pour inhiber son activité.

- La deuxième protéine à laquelle nous nous sommes intéressés est une protéine de type WD40. En effet, comme indiqué dans le contexte bibliographique, les protéines WD40 coopèrent avec les protéines MYB et bHLH pour réguler la voie de biosynthèse des tannins chez *Arabidopsis* [69]. Cette protéine WD40 apparaît fortement homologue à la protéine At1g20540 d'*Arabidopsis* qui appartient à la famille des transducines, mais son rôle physiologique n'a pas encore été clairement établi.

Bien que ces interactions paraissent vraisemblables en tenant compte des données de la littérature, elles nécessitent d'être confirmées *in vivo* dans un système autre que la levure. Pour ce faire, l'ADNc pleine longueur de *VvMyb5a* et l'ADNc tronqué de la partie

C-terminale contenant le domaine GRD (contrôle négatif) ont été clonés dans un vecteur d'expression transitoire. Des expériences de co-expression transitoire de la cible VvMyb5a ou VvMyb5a tronquée avec la proie (KGM-like ou WD40) dans des protoplastes d'*Arabidopsis* ou de vigne, suivies d'une co-immunoprécipitation pourront permettre de tester ces interactions éventuelles. En effet, les biais de la technique double hybride sont multiples. Tout d'abord, elle met en jeu des fusions protéiques susceptibles de modifier la structure et les propriétés des protéines cibles. Ainsi, dans le cadre du criblage d'une banque, cette technique peut conduire à l'identification d'un nombre important de faux positifs (protéines auto-activatrices ou protéines dites "collantes"), mais également à la non détection des faux négatifs [339, 340]. De plus, seules les interactions entre deux protéines peuvent être mises en évidence par cette technique, ce qui s'avère être un problème important dans le cadre de cette étude. En effet, chez *Arabidopsis*, l'analyse fonctionnelle du promoteur du gène *BANYULS* (*BAN*), qui code une ANR, a permis de démontrer que sa spécificité d'expression dans les cellules productrices de tannins était contrôlée au niveau transcriptionnel par un complexe nommé MBW composé des protéines régulatrices TT2 (MYB), TT8 (bHLH) et TTG1 (WDR) [69]. D'autres études menées chez le pétunia et le maïs ont confirmé l'implication du complexe MBW dans la régulation de l'expression des gènes codant les enzymes de la voie de biosynthèse des anthocyanes [79]. Enfin, une étude récente chez le maïs a permis d'identifier un nouveau partenaire protéique appelé RIF1 (R-interacting factor 1) qui serait un nouveau co-facteur essentiel du complexe transcriptionnel MBW. Cette protéine interagirait spécifiquement avec la protéine bHLH R (Red) de maïs et serait impliquée dans le remodelage de la chromatine au niveau du promoteur du gène *AI*. La transcription de ce gène, qui code une enzyme de type DFR, serait donc contrôlée par un complexe formé de quatre protéines [341]. De fait, le système du double hybride chez la levure, qui permet de détecter seulement les interactions physiques binaires, ne semble pas être le plus adapté à notre modèle d'étude. D'autres techniques permettant soit la purification de complexe par la méthode du TAP-tag (tandem Affinity Purification) couplée à la spectrométrie de masse [342], soit l'identification de l'appât dans un complexe protéique par la méthode du FRET (Fluorescence resonance energy transfer) [343], doivent donc maintenant être envisagées.

Chapitre 2

Identification de nouveaux facteurs de transcription MYB par analyse de mutants naturels de la vigne

Le génome de la vigne présente un taux élevé d'hétérozygotie, ce qui génère une variabilité génétique importante au cours de la reproduction sexuée. Aussi, les viticulteurs, soucieux de préserver les cépages issus d'une longue tradition, propagent la vigne par multiplication végétative *via* des techniques diverses (bouturage, marcottage, greffage). Ce mode de propagation permet de maintenir les caractéristiques génétiques du cultivar et d'obtenir des plantes qui présentent des phénotypes très similaires. Pourtant, dans de rares occasions, quelques cultivars peuvent avoir un phénotype instable et présenter des caractéristiques phénotypiques différentes, visibles sur une partie ou une branche entière de la plante. Quand ces branches (appelées "bud-sports") sont végétativement propagées par des techniques clonales, le nouveau phénotype est généralement maintenu menant à une nouvelle variété qui se distingue des parents par un seul caractère phénotypique différent [344]. Ainsi, la viticulture dispose d'un grand nombre de mutants naturels et les plus communs sont les mutants de couleurs. En effet, les modifications du métabolisme polyphénolique dans les baies de raisin sont facilement observables et n'ont généralement aucun effet préjudiciable sur la viabilité de la vigne. Disposant de cette ressource originale dans les domaines viticoles de l'INRA à Bordeaux (Grand Parc, Latresnes), nous avons entrepris leur caractérisation afin d'identifier de nouveaux gènes impliqués dans le contrôle transcriptionnel des gènes codant les enzymes de la voie de biosynthèse des anthocyanes.

I- Approche gènes candidats sur le Pinot Noir et deux de ses mutants naturels, le Pinot blanc et le Pinot gris.

Le groupe variétal "Pinot" compte 6 variétés : gris, noir, meunier, teinturier, blanc et moure. Dans ce chapitre, nous nous intéresserons seulement aux cépages Pinot noir, gris et blanc. Dans le cadre du travail de recherche qui m'a été confié durant mon master, une approche gène candidat a été entreprise en utilisant la technique de la RT-PCR semi-quantitative. Le but de cette étude était de rechercher des FT qui présentaient de fortes variations d'expression entre les cépages Pinots noir, gris et blanc et au cours du développement des baies. Des séquences présentant des homologies avec des FT des familles MYB R₂R₃, bHLH et WD40 identifiés chez les espèces modèles et potentiellement impliqués dans le métabolisme polyphénolique ont été recherchées dans la base de données GrapEST¹. Cette base de données regroupe l'ensemble des EST issues du Programme 'Lignome Vigne' initié par l'INRA et dans lequel l'équipe s'était impliquée. Les recherches avaient été réalisées en se basant sur les annotations fonctionnelles des EST et par recherche d'homologies avec des séquences codant des MYB, bHLH et WD40 caractérisées chez les espèces modèles en

¹ GrapEST: <http://web.ensam.inra.fr/GrapEST>

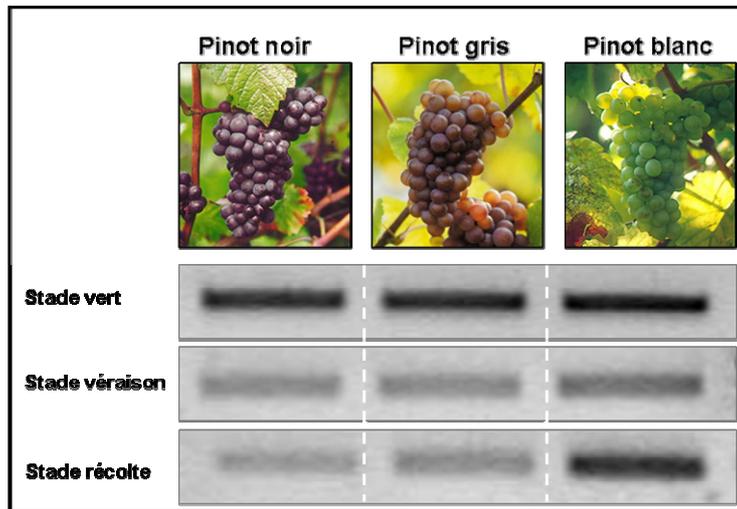


Figure 49 : Analyse par RT-PCR semi-quantitative de l'expression de *VvMybPA1* au cours du développement des baies de Pinot noir, gris et blanc.

Les RT-PCR semi-quantitatives ont été réalisées à partir d'ARN totaux extraits de baies des cépages Pinot noir, gris et blanc, récoltées à trois stades de développement (vert, véraison et récolte) et épépinées. Les ADNc correspondant aux ARNm codant le facteur d'élongation EF1 γ ont été utilisés comme standard interne et ont permis de vérifier que des quantités équivalentes d'ARN totaux avaient bien été utilisées pour chaque réaction de RT-PCR (résultat non montré).

utilisant le programme BLAST². Neuf EST avaient été retenus: six codant des protéines de la famille MYB R₂R₃, deux de la famille bHLH et un de la famille WD40. Cette approche nous a permis d'identifier un gène *MYB* dont le profil d'expression apparaissait particulièrement intéressant. Ce gène a été appelé *VvMyb3* puis renommé au cours de ma thèse *VvMybPA1*, suite à une publication d'une équipe australienne.

1. Analyse de l'expression différentielle du facteur MYB R₂R₃ *VvMybPA1* par la technique de la RT-PCR

Les profils d'accumulation des transcrits *VvMybPA1* ont été analysés dans les baies des trois cépages prélevées à trois stades de développement (stade vert, stade véraison et stade récolte) par la technique de la RT-PCR. De façon à conserver un aspect semi-quantitatif, les réactions de PCR ont été réalisées en utilisant un nombre de cycles d'amplification approprié et constant quel que soit le stade de développement considéré. Une séquence d'ADNc correspondant à l'extrémité 5' du gène *VvMybPA1* a été amplifiée avec le couple d'amorce M5FS1 et M5FAS (annexe 2). Le gène *EF1γ* (Elongation factor 1 gamma) exprimé constitutivement chez la vigne a été utilisé comme référence. Les amorces utilisées pour amplifier *EF1γ* sont EF1sens et EF1AS (annexe 2). La figure 49 montre que l'expression de *VvMybPA1* est très forte dès le stade vert dans les trois cépages puis diminue sensiblement au stade véraison. La quantité de transcrits reste ensuite constante dans les baies de Pinot noir et gris mais elle augmente à nouveau sensiblement chez le Pinot blanc au stade récolte.

2. Clonage et analyse *in silico* de la séquence d'ADNc du gène *VvMybPA1*

a. Clonage de la séquence ADNc du gène *VvMybPA1* par RACE-PCR

Pour amplifier l'ADNc pleine longueur correspondant à *VvMybPA1*, nous avons retenu la technique de la RACE-PCR. Dans la base de donnée GrapEST, l'extrémité 5' du transcrit correspondant au gène *VvMybPA1* était disponible. Afin de pouvoir confirmer cette séquence 5' et identifier le site d'initiation de la transcription de *VvMybPA1*, une amplification de l'extrémité 5' de l'ADNc (5' RACE) a été réalisée. Cette méthode consiste en une transcription inverse réalisée avec une amorce antisens spécifique de *VvMybPA1* (5RaceMyb3AS2, annexe 1) suivie de l'addition d'une queue de dATP à l'extrémité 3' du fragment simple brin de l'ADNc sous l'action d'une transférase terminale. Compte tenu du profil d'expression du gène *VvMybPA1* (Figure 49), des ARN totaux extraits de baies récoltées au stade vert du cépage Cabernet sauvignon et épépinées ont été utilisés comme matrice. Une réaction de PCR "Touch Down" a ensuite été réalisée avec une amorce sens spécifique de la queue polyA ainsi créée (OligodT-anchor primer) et une amorce spécifique

² BLAST: <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>

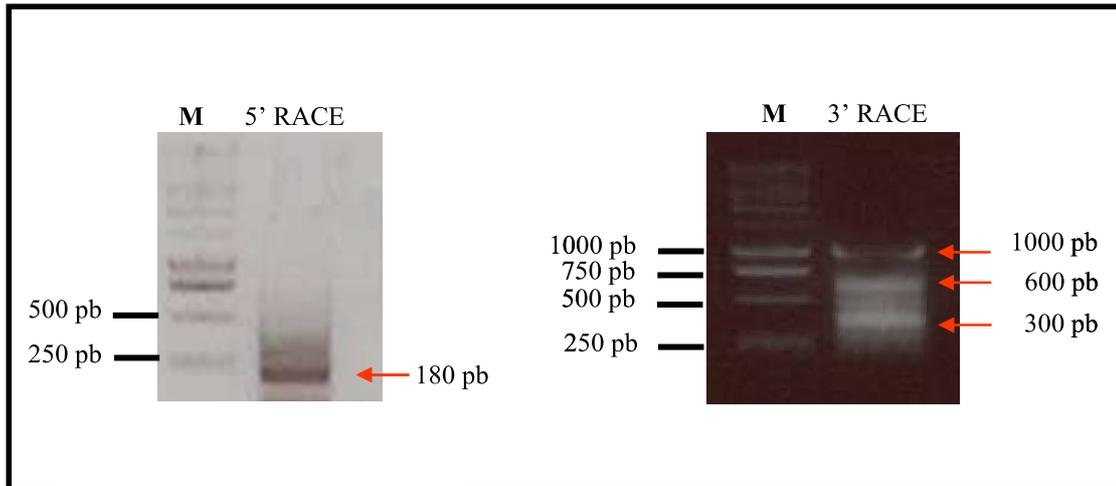


Figure 50: Analyse par électrophorèse en gel d'agarose des amplifications des extrémités 5' et 3' de l'ADNc de *VvMybPA1* par RACE-PCR.

Les amplifications ont été obtenues à partir d'ADNc amplifiés à partir d'ARN totaux extraits de baies épépinées récoltées au stade vert (cépage Cabernet sauvignon). Un fragment de 180 pb a été amplifié du côté 5' de *VvMybPA1* (5'RACE), et trois fragments de 300 pb, 600 pb et 1000 pb du côté 3' (3'RACE). 15 μ l de la réaction PCR ont été déposés par piste. M: marqueur de taille (1 Kb ladder, Promega), pb: paire de base.

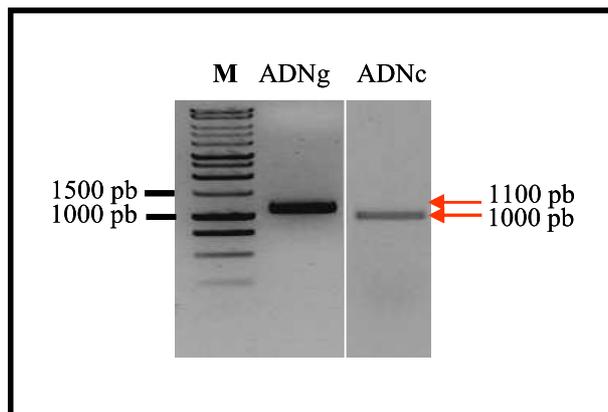


Figure 51 : Analyse par électrophorèse des amplifications de la séquence codante et génomique de *VvMybPA1*.

Les amplifications ont été obtenues à partir d'ADNc synthétisés à partir d'ARN totaux extraits de baies épépinées récoltées au stade vert et d'ADNg extrait de feuille (cépage Cabernet Sauvignon). 15 μ l de la réaction PCR ont été déposés par piste. M: marqueur de taille (1 Kb ladder, Promega), pb: paire de base, ADNg: ADN génomique, ADNc: ADN complémentaire.

de *VvMybPAI* (5RaceMyb3AS1, annexe 1). A l'aide de ce couple d'amorce, un fragment d'environ 180 pb a pu être amplifié (figure 50A). Pour déterminer l'extrémité 3' du transcrit correspondant à *VvMybPAI*, une amplification par 3' RACE a été effectuée. Après l'étape de transcription inverse, une réaction de PCR "Touch Down" à partir des ADNc amplifiés a été réalisée avec une amorce sens spécifique de *VvMybPAI* (3RaceMyb3sens2, annexe 1) et une amorce antisens spécifique de la queue de dTTP (PCR anchor primer). Trois fragments (environ 1000 pb, 600 pb et 300 pb) ont pu être amplifiés (figure 50A). Les fragments obtenus par les réactions de 5' et 3' RACE (respectivement 180 pb et 1000 pb) ont ensuite été clonés dans le vecteur pGEMT[®]-easy, amplifiés dans les bactéries *E. coli* DH5 α puis séquencés. Le séquençage des fragments amplifiés a révélé que *VvMybPAI* présentait une région 5' non codante de 148 pb et une région 3'UTR de 175 pb. L'ADNc pleine longueur de *VvMybPAI* a ensuite été amplifiée par une réaction PCR réalisée sur des ADNc synthétisés à partir d'ARN totaux de baies de cépages Cabernet sauvignon récoltées au stade vert et épépinées avec le couple d'amorce VvMyb3sens (complémentaire de la région 5' non codante de *VvMybPAI*) et VvMyb3AS (complémentaire de la région 3' non codante de *VvMybPAI*) (annexe 1). La partie codante correspond à un cadre de lecture ouvert de 861 pb qui coderait une protéine de 286 aa (Figure 51). Le logiciel ProtParam³ prédit une masse moléculaire de 32,2 kDa et un point isoélectrique de 9,21.

La séquence génomique de *VvMybPAI* a été amplifiée par une réaction PCR réalisée sur de l'ADNg de feuilles du cépage Cabernet sauvignon, avec le couple d'amorce VvMyb3sens et VvMyb3AS (Figure 51). Le produit d'amplification de la réaction PCR de 1100 pb, a été cloné dans le vecteur pGEMT[®]-easy puis séquencé. L'alignement de la séquence de l'ADNc avec la séquence génomique de *VvMybPAI* a révélé l'existence de deux exons (135 pb et 726 pb) et d'un intron de 87 pb (Figure 52). Par la suite, le séquençage du génome de la vigne a permis de vérifier la séquence que nous avons obtenue dans la base de données Génoscope⁴ (Figure 52A). La séquence de *VvMybPAI* s'aligne en fait sur deux références de cette base de données: la région 5' non codante et une partie du cadre ouvert de lecture (nucléotides 1 à 253) sont répertoriées sous la référence *GSVIVT00008645001* (positions 51311721 à 51311321) et le reste de la séquence codante ainsi que la région 3'UTR correspondent à la référence *GSVIVT00008644001* (positions 51311315 à 51310430) (figure 52B). Ces deux références n'ont pas encore été localisées sur un des 19 chromosomes.

³ ProtParam: <http://expasy.org/tools/protparam.html>

⁴ Genoscope: <http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>

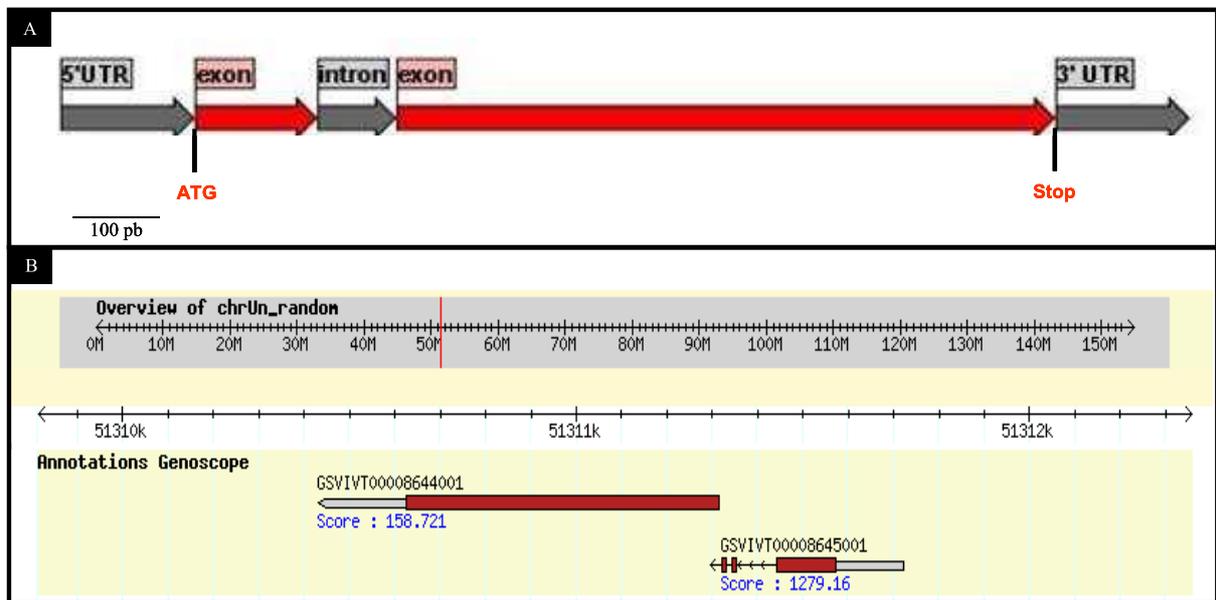


Figure 52: Analyse de la séquence génomique de *VvMybPA1*.

A- Représentation schématique de la structure intron/exon du gène *VvMybPA1*. Le gène codant *VvMybPA1* est constitué de deux exons (rouge) et un intron (gris). Les régions non-codantes en 3' et 5' sont également indiquées en gris. Les codons d'initiation (ATG) et de terminaison (Stop) de la traduction sont représentés par des traits noirs. Cette représentation schématique a été réalisée à partir du logiciel CLC Main Workbench 4. Echelle, 100 paires de bases (pb).

B- Localisation et annotations du transcrit *VvMybPA1*. Le gène *VvMybPA1* n'a pas pu être localisé de façon précise sur un chromosome (chrUn_random). La séquence nucléotidique de *VvMybPA1* s'aligne avec deux séquences identifiées par le Génoscope. Cette représentation schématique a été réalisée à partir de la base de données du Génoscope.

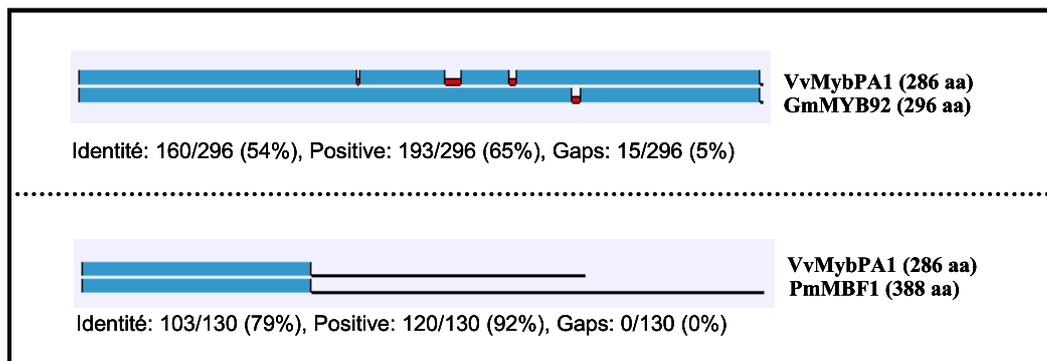


Figure 53 : Comparaison des séquences protéiques de *VvMybPA1* avec deux membres de la famille des protéines MYB.

Cette figure a été obtenue avec le programme BLAST 2 SEQUENCES qui permet un alignement des séquences protéiques deux à deux. Le nombre d'acides aminés est indiqué entre parenthèse. Les cadres bleus représentent les zones homologues. Les décalages sont représentés en rouge. L'identité représente le nombre d'acides aminés en commun entre les deux séquences sur le nombre total d'acides aminés de la zone en question. "Positives" représente le nombre d'acides aminés du même groupe. "Gaps" représente le nombre de décalage d'acides aminés dans la séquence pour un meilleur alignement. Ces différentes valeurs sont aussi exprimées en pourcentage. La double lettre préfixe indique l'origine de la protéine Gm pour le soja (*Glycine max*) et Pm pour pin maritime (*Picea mariana*). Les numéros d'accèsion sont ABH02844 pour GmMYB92 et AAA82943 pour PmMBF1.

b. Analyse *in silico* de la séquence VvMybPA1

La recherche de motifs et de domaines signatures dans la protéine VvMybPA1 a été effectuée en utilisant les logiciels InterProscan⁵ et ScanProsite⁶. Proche de l'extrémité N-terminale, VvMybPA1 contient deux motifs hélice-tour-hélice qui définissent les deux répétitions imparfaites R₂R₃ du domaine MYB (InterProScan001005, PF00249). Le domaine MYB contient le motif [D/E]Lx2[R/K]x3Lx6Lx3R permettant l'interaction avec les protéines bHLH [195, 270]. Le logiciel ScanProsite a aussi identifié un domaine de liaison à l'ATP et au GTP (Prosite00017). Cette analyse a également révélé la présence de différents sites potentiels de modifications post-traductionnelles qui sont les suivants : neuf signatures de myristoylation (débutant aux aa 11, 15, 45, 86, 120, 194, 195, 197, 202), un site de N-glycosylation (débutant à l'aa 184), six sites de phosphorylation par la protéine kinase C (débutant aux aa 18, 52, 113, 125, 151, 211), sept sites de phosphorylation par la caséine kinase II (débutant aux aa 18, 71, 100, 157, 211, 242, 279), un site de phosphorylation par la tyrosine kinase (débutant à l'aa 99) et deux sites de phosphorylation par une protéine kinase AMPc et GMPc dépendante (débutant aux aa 127, 138).

Des protéines homologues à VvMybPA1 ont été recherchées chez d'autres plantes dans la banque de données du NCBI⁷. Le domaine de liaison à l'ADN (DBD) des protéines MYB R₂R₃ étant largement conservé, de nombreuses protéines végétales possèdent un pourcentage d'homologie élevé avec la région N-terminale de VvMybPA1. Cependant, le domaine DBD de la protéine MYB du pin maritime PmMBF1 est le plus proche avec 79% d'identité (figure 53). Toutefois, aucune fonction putative n'a encore été attribuée à PmMBF1. La région C-terminale de VvMybPA1 présente peu d'homologie avec d'autres protéines MYB. Avec 54% d'identité, la protéine de soja GmMYB92 représente l'homologue le plus proche (figure 53), mais cette protéine n'a pas encore été caractérisée fonctionnellement. Enfin, la figure 31 indique que VvMybPA1 est proche des protéines GmMYB92 et GmMYB185. En dehors du domaine DBD, les trois membres de ce sous-groupe ne possèdent pas de domaines conservés qui permettent de les classer dans un des sous-groupes spécifiques proposés dans les classifications de Stracke et Krantz [195, 271].

3. Caractérisation fonctionnelle du gène *VvMybPA1 in planta*

La recherche d'homologue n'ayant pas permis d'attribuer une fonction putative à *VvMybPA1*, nous avons entrepris sa caractérisation fonctionnelle chez *Arabidopsis thaliana*. Deux approches ont été menées en parallèle dans ce travail de thèse: l'étude de mutants d'insertion d'ADN-T et l'étude de lignées sur-expresses. Tout d'abord, nous avons recherché

⁵ InterProscan: <http://www.ebi.ac.uk/InterProScan>

⁶ ScanProsite : <http://www.expasy.org/tools/scanprosite/>

⁷ NCBI: <http://www.ncbi.nlm.nih.gov/>



Figure 54 : Analyse phylogénétique de VvMYB24 et des 126 protéines MYB R2R3 d'*Arabidopsis thaliana*

Les séquences ont été alignées en utilisant le logiciel Muscle [376]. L'arbre phylogénétique initial a été construit suivant le principe du 'Neighbor-joining' [377], et les distances d'évolution ont ensuite été testées en utilisant la méthode d'évolution minimale [378]. La longueur des différentes branches correspond à la distance d'évolution exprimée en nombre de substitutions d'acides aminés par site.

l'homologue de VvMybPA1 chez *Arabidopsis thaliana* par une analyse phylogénétique. Puis, après identification du gène homologue, nous avons recherché les éventuelles conséquences phénotypiques sur les lignées d'insertions d'ADN-T correspondantes. Enfin, nous avons tenté de caractériser des lignées transgéniques d'*Arabidopsis* sur-exprimant le gène *VvMybPA1*.

a. Isolement et caractérisation d'un mutant d'insertion d'ADN-T dans le gène *VvMybPA1*

Au début de ce travail de thèse, nous ne disposions pas dans l'équipe de salles de cultures adéquates pour le tabac ou la vigne. Ainsi, pour déterminer la fonction de *VvMybPA1*, notre choix s'est porté sur *Arabidopsis thaliana*. En raison de ses nombreux avantages (génomme séquencé, petite taille de la plante, rapidité du cycle de vie) et disposant d'une salle dédiée à sa culture, cette plante modèle nous est apparu, dans notre contexte, comme un modèle d'étude de choix.

i. Recherche de l'homologue de *VvMybPA1* chez *Arabidopsis thaliana*

La recherche d'homologues dans la banque de données NCBI par le logiciel BLAST n'ayant pas permis d'identifier clairement l'homologue de *VvMybPA1* chez *Arabidopsis*, une analyse phylogénétique a été réalisée. Les séquences protéiques des 126 protéines MYB R₂R₃ d'*Arabidopsis* ont été recherchées dans la banque de données du TAIR⁸. La séquence protéique de VvMybPA1 et celles de l'ensemble des protéines MYB d'*Arabidopsis* ont ensuite été utilisées pour réaliser un arbre phylogénétique. Cette analyse de phylogénie est présentée dans la figure 54. La protéine homologue la plus proche de VvMybPA1 est AtMyb82 (At5g52600, N° d'accension: NP_680426). AtMyb82 et VvMybPA1 apparaissent relativement proches des groupes de protéines MYB impliquées dans la voie de biosynthèse des anthocyanes et des PA (sous-groupe 5 et 6 respectivement), mais elles sont cependant clairement isolées dans un sous-groupe distinct.

Dans la banque de donnée du TAIR, la séquence nucléique du gène *AtMyb82* était disponible. L'alignement de la séquence de l'ADNc avec la séquence génomique a révélé l'existence de 2 introns (93 et 73 pb) et 3 exons (134, 129 et 343 pb). La partie codante consiste en un cadre de lecture ouvert de 606 pb. La séquence protéique d'AtMyb82, déduite de la séquence de l'ADNc, est formée de 201 aa. Le logiciel ProtParam⁹ prédit une masse moléculaire de 23,3 kDa et un point isoélectrique de 9,11. A notre connaissance, aucun travail de caractérisation fonctionnelle n'a été entrepris sur le gène *AtMyb82*. En fait, seule une expérience de double

⁸ TAIR: <http://www.arabidopsis.org/index.jsp>

⁹ ProtParam: <http://expasy.org/tools/protparam.html>

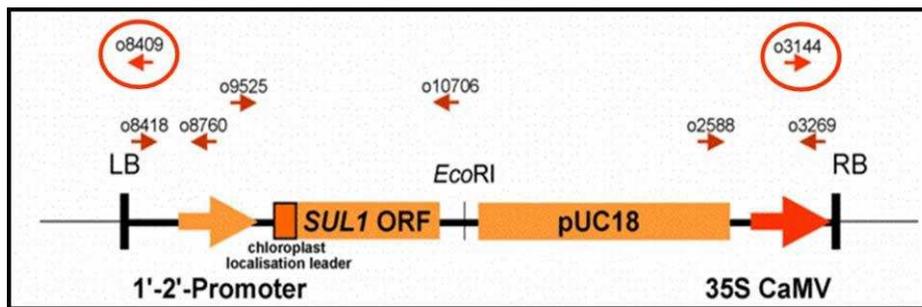


Figure 55 : Représentation schématique de l'ADN-T du vecteur pAC106 utilisé par GABI-kat.

L'ADN de transfert, bordé par les frontières gauche (LB) et droite (RB), contient le marqueur de sélection pour les plantes. Le marqueur *Sul1* (dihydroptéroate synthase) est placé sous contrôle du promoteur fort "1'-2'-Promoter" et confère la résistance à la sulfadiazine. La position et le nom des amorces utilisées pour le génotypage et la détermination du site d'insertion de l'ADN-T dans le gène *AtMyb82* sont entourés en rouge.

hybride chez la levure a pu montrer qu'AtMyb82 interagissait physiquement avec la protéine TT8, impliquée dans la régulation de la biosynthèse des anthocyanes et des PA chez *Arabidopsis* [257].

ii. Obtention du mutant d'insertion

Pour obtenir un mutant « perte de fonction » pour le gène *AtMyb82*, nous avons effectué un criblage *via* internet des collections de plantes d'*Arabidopsis* mutagénisées avec un élément stable (ADN-T) [345]. L'ADN-T est une séquence d'ADN, portée par un plasmide bactérien, qui est intégrée dans le génome nucléaire de la plante lors de la transformation par *Agrobacterium tumefaciens* [346]. Des ADN-T sauvages ont été modifiés puis utilisés pour générer des banques de mutants aléatoires. Il est donc possible d'identifier des mutants pour un gène d'intérêt en criblant ces collections de mutants par PCR. En effet, les insertions sont repérables car elles sont définies par une région flanquante (FST, Flanking Sequence Tag) correspondant au gène d'intérêt. La base de données SIGnAL¹⁰ (Salk Institute Genomic Analysis Laboratory) regroupe les FST disponibles, définies à partir de la bordure gauche du transgène de 130 000 lignées d'insertion ADN-T d'*Arabidopsis* (écotype Columbia, Col-0).

En 2007, les recherches dans cette base de données nous ont permis d'identifier un mutant présentant une insertion dans le premier intron du gène *AtMyb82*. Ce mutant provient d'une lignée produite par le programme GABI-kat¹¹ en Allemagne [347]. Les graines des plantes T3 correspondant à la lignée d'insertion 057A04 ont été commandées au NASC¹² sous la référence N405380. Le NASC (Nottingham *Arabidopsis* Stock Center, Royaume Uni) est un centre de ressources génétiques qui conserve et distribue entre autre l'ensemble des lignées GABI-Kat. L'ADN-T utilisé pour la transgénése dans la banque GABI-Kat possède un seul gène de sélection, le gène *sul1* (Sulfadiazine 1) sous contrôle du promoteur fort « 1'-2'-Promoter » (figure 55). *Sul1* code pour la dihydroptéroate synthase, protéine qui confère à la plante qui l'exprime une résistance à l'herbicide sulfadiazine.

iii. Isolement des lignées mutantes homozygotes

Les graines de la lignée 057A04 sont de génération T3. D'après les indications de la base GABI, sur 100 graines de la lignée 057A04 déposées sur un milieu contenant de la sulfadiazine, 92 plantules sont apparues résistantes. Ainsi, au sein des 17 lots de graines reçus, trois génotypes peuvent être trouvés: homozygotes pour l'insertion de l'ADN-T dans *AtMyb82*, hémizygotés et sans ADN-T. Une recherche par PCR a donc été réalisée pour

¹⁰ SIGnAL: <http://signal.salk.edu/cgi-bin/tdnaexpress>

¹¹ GABI-Kat: <http://www.gabi-kat.de/>

¹² NASC: <http://arabidopsis.info/>

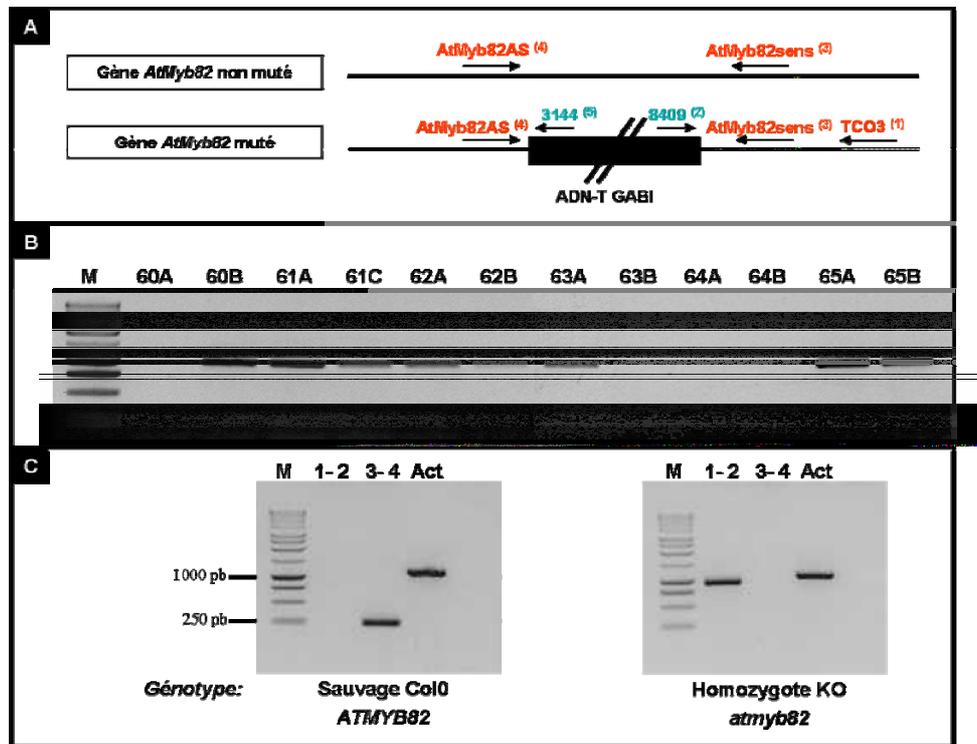


Figure 56: Analyse par PCR du génotype des mutants d'insertion ADN-T pour le gène *AtMyb82* (*At5g52600*).

A- Représentation schématique de la position des amorces utilisées pour le génotypage des mutants KO62B et KO65A. Le couple d'amorces 3-4 permet l'amplification du gène *AtMyb82* non muté. Un produit d'amplification avec le couple 1-2 ou 4-5 révèle la présence de l'ADN-T dans le gène muté *AtMyb82*.

B- Résultat du criblage PCR réalisé sur différents lots de graines. Les amplifications ont été obtenues à partir d'ADNg extrait de feuille avec le couple d'amorce 1-2. Sept µl de la réaction PCR ont été déposés dans chaque piste. M: marqueur de taille (1 Kb ladder, Promega).

C- Résultat du génotypage pour le mutant KO62B. L'amplification d'une partie de l'ADNc de l'*actine2* d'*Arabidopsis thaliana* est utilisée comme contrôle. Sept µl de la réaction PCR ont été déposés dans chaque piste. pb: paire de base, Act: couple d'amorce de l'Actine2.

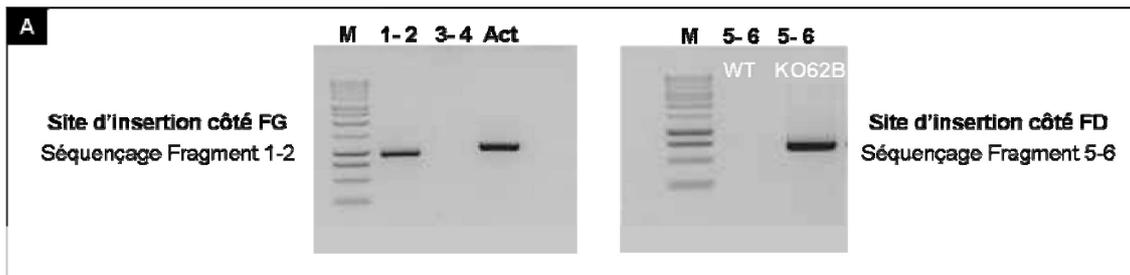


Figure 57: Identification du site d'insertion exact de l'ADN-T dans le mutant KO62B.

A- Analyse par électrophorèse en gel d'agarose de l'amplification obtenue avec les couples d'amorce 1-2 et 5-6 sur de l'ADNg de feuille du KO62B. Sept μ l de la réaction PCR ont été déposés dans chaque piste. La position des amorces est indiquée dans la figure 56. Les deux fragments amplifiés avec ces amorces ont été purifiés puis envoyés à séquencer. M: marqueur de taille (1 Kb ladder, Promega).

B- Site exact de l'insertion de l'ADN-T dans le mutant KO62B après analyse des fragments séquencés. Les séquences en rouge correspondent aux séquences des frontières gauche (FG) et droite (FD) de l'ADN-T. Un nucléotide supplémentaire a été identifié dans la séquence de la frontière gauche et est souligné. Les séquences en lettres minuscules correspondent à l'intron n°1 présent dans le gène *AtMyb82*.

PUBLICATIONS ET COMMUNICATIONS SCIENTIFIQUES

Publications et communications scientifiques

Articles

Deluc L, Bogs J, Walker AR, Ferrier T, Decendit A, Merillon J-M, Robinson SP, Barrieu F: The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. *Plant Physiol* 2008;pp.108.118919.

Communications affichées

Deluc, L., Ferrier, T., Barrieu, F., Descendit, A., And Hamdi S. (2005) Characterization of A NEW MYB transcription factor involved in the regulation of phenylpropanoid metabolism in grape berry. 6^{ème} Colloque Général de la Société Française de Physiologie Végétale. Arcachon, 27-29 Avril 2005.

Ferrier, T., Barrieu, F., And Hamdi S. (2006) Caractérisation Fonctionnelle du Gène VvMyb5a, facteur de transcription MYB impliqué dans le métabolisme des Composés Phénoliques dans la baie de raisin. Journée scientifique de l'Ecole Doctorale Sciences du Vivant, Géosciences et Science de l'Environnement. Bordeaux, 7 avril 2006

Ferrier, T., Barrieu, F., And Hamdi S. (2006). Regulatory mechanisms of the *Vitis Vinifera* MYB transcription factor VvMYB5a: a regulator of the phenylpropanoid pathway in grape berry. XV FESPB CONGRESS (Federation of European Societies of Plant Biology), Lyon, 17-21 juillet 2006 ; p.133.

Ferrier, T., Fouquet, R., Ollat, N., Delrot, S., and Barrieu, F. (2007) Grapevine natural mutants: a tool to study ripening in grape berry. 7^{ème} Colloque Général de la Société Française de Physiologie Végétale. Versailles, 12-14 Septembre 2007.

