HAL
open science

Caractérisation des gènes PR10 chez Vitis vinifera et étude de leur expression durant l'embryogenèse somatique

Sylvain Lebel

▶ **To cite this version:**

Sylvain Lebel. Caractérisation des gènes PR10 chez Vitis vinifera et étude de leur expression durant l'embryogenèse somatique. Sciences agricoles. Université de Haute Alsace - Mulhouse, 2010. Français. NNT : 2010MULH5091 . tel-00591990

HAL Id: tel-00591990

https://theses.hal.science/tel-00591990v1

Submitted on 10 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Haute-Alsace

Ecole Doctorale Jean-Henry Lambert – ED 494

THESE

Pour l'obtention du grade de

Docteur de l'Université de Haute-Alsace

Domaine : Biologie cellulaire

Sylvain LEBEL

Caractérisation des gènes *PR10* chez *Vitis vinifera* et étude de leur expression durant l'embryogenèse somatique

Soutenue publiquement le 13 décembre 2010 devant le jury composé de :

Anne-Françoise ADAM-BLONDON	Directrice de recherches INRA – URGV, Evry	Rapporteur
Pierre COUTOS-THEVENOT	Professeur des Universités – Université de Poitiers	Rapporteur
Laurence DESPONS	Maître de conférences – Université de Strasbourg	Examinatrice
Pascale MAILLOT	Maître de conférences – Université de Haute Alsace	Co-encadrante
Paul SCHELLENBAUM	Maître de conférences – Université de Haute Alsace	Co-encadrant
Bernard WALTER	Professeur – Université de Haute Alsace	Directeur de thèse

Laboratoire Vigne, Biotechnologies et Environnement

EA 3991

Remerciements

Je souhaite tout d'abord bien évidemment remercier toutes les personnes qui prendront le temps de lire le résultat de mon travail, et plus particulièrement les rapporteurs, Anne-Françoise Adam-Blondon et Pierre Coutos-Thevenot. J'espère que cette lecture vous sera agréable.

Je souhaite remercier au moins autant les personnes qui ont encadré ma thèse.
Merci tout d'abord au Professeur Bernard Walter pour m'avoir accueilli dans son équipe, pour la confiance qu'il m'a accordée et pour les corrections qu'il a apportées à mon manuscrit.
Un grand merci à mes co-encadrants, Pascale et Paul pour avoir suivi mon parcours tout au long de ces trois années, dans les hauts comme dans les bas. Merci pour votre disponibilité, votre bonne humeur et toutes les connaissances que vous m'avez apportées, qu'elles soient scientifiques ou non. Je pense malheureusement qu'il n'est pas possible d'exprimer par écrit toute l'estime que je vous porte.

Merci à Christophe Bertsch, directeur du LVBE, qui m'a accueilli au laboratoire. Merci également à toutes les personnes du laboratoire, de l'IUP et de l'IUT pour leur accompagnement au quotidien: Sybille (et sa grand-mère !), Flore, Laurence, Yann, Claude, Julie, Isabelle, Karine J, Carmen, Catherine, Mr et Mme Sediri, Marc, Christine, ceux qui sont partis, loin ou pas : Laurent, Thierry,Jean-Yves, Rosaria, et ceux que j'oublie..

Merci également à ceux qui m'ont accompagné et encouragé durant mes années universitaires, que ce soit à Strasbourg ou à Colmar, avec une mention spéciale pour l'équipe GAV de l'INRA de Colmar et en particulier Pere, pour m'avoir guidé lors de mon stage.

Une grosse pensée pour les thèsards passés, présents et à venir du labo. Ludo, Gaëlle, David, Paul B, Karine B, Mélanie, Noémie, Camille et Flore, bon courage et bonne chance pour la suite.

Enfin, dernier cités mais non des moindre, merci à tous les gens de mon entourage extérieur au labo. Les amis d'abord. Merci à Elodie, Fabien, Thomas Stéphane, Sébastien, Julien, Nelly et tous les autres, pour avoir toujours été là pour me soutenir et pour votre compréhension face à mon absence durant les mois de rédaction. La famille surtout. Que dire d'autre sinon que je les aime ?

Et merci à Nathalie.

Liste des abbréviations

2,4-D : Acide 2,4-Dichlorophenoxyacétique

ABR17 : « ABA-responsive 17 »

ADN : Acide désoxyribonucléique

ADNc : Acide désoxyribonucléique complémentaire

AFP : « Antifungal protein »

AGP : « Arabinogalactan protein »

ARN : Acide ribonucléique

ARNm : Acide ribonucléique messager

auxRE : Élément de réponse à l'auxine

BAK : « BRI1-Associated Kinase »

BAP : 6-Benzylaminopurine, ou Benzyl adénine

BLAST : « Basic Local Alignment Search Tool »

BLASTn : « BLAST nucleotid »

BLAT : « BLAST -Like Alignment Tool »

bv. : biovar

bZIP : « Basic region/leucine zipper »

cDNA-AFLP : Acide désoxyribonucléique complémentaire - « Amplified fragment length polymorphism »

CDS : Séquence codante

Cp : Point de croisement

Cq : Cycle de quantification

CSBP : « cytokinin-specific binding proteins »

Ct : Cycle seuil

cv : cultivar

DNase : Désoxyribonucléase

dNPT : Désoxynucléotide triphosphate

Dof : « DNA-binding with One Finger »

E : Efficacité de la PCR

EBI : « European Bioinformatics Institute »

EDTA : Acide ethylène diamine tétraacétique

EF1α : « Elongation factor 1 α »

EST : « Expressed sequence tag »

GAPDH : « Glyceraldehyde 3-phosphate dehydrogenase »

GRE : Élément de réponse au glutocorticoïdes

IAA : Acide indole 3-acétique

IASMA : « Istituto Agrario San Michele all'Adige »

IPR : « Intracellular PR »

L1L : « Leafy Cotyledon1-Like »

LDR : « Linear dynamic range »

LEC : « Leafy cotyledon »

LTP : « Lipid transfer protein »

MIBBI : « Minimum Investigation for Biological and Biomedical Investigation »

MIQE : « Minimum Information for publication of Quantitative real-time PCR Experiments »

MLN : « Metastatic Lymph Node »

MLP : « Major latex protein »

MLP/RPP : « Major latex protein-Ripening-related proteins »

NAA : Acide naphtalène acétique

NBS-LRR : « nucleotide binding site – leucine rich repeat »

NCBI : « National Center for Biotechnology Information »

NCED : 9-cis-epoxycaroténoïde dioxygènase

NCS : Norco-claurines synthétases

NF-κB : Facteur nucléaire κB

NOA : Acide 2-naphtoxyacétique

ORF : Cadre ouvert de lecture

OxOa : « Oxalate-oxydase »

OxOLP : « Oxalate oxydase-like protein »

P-loop : « Phosphate-binding loop »

PCR : Amplification en chaîne par polymérase

pI : point isoélectrique

PITP : « Phosphatidylinositol transfer proteins »

PM : Poids moléculaire

PR : « Pathogenesis-related »

PRP : Protéine riche en proline

pv : pathovar

PVP : Polyvinylpyrrolidone

RACE-PCR : Amplification rapide des extrémités des ADNc – Amplification en chaîne par polymérase

RDML : « Real-Time PCR Data Markup Language »

RLK : Receptor-like kinase

RLP : Receptor-like protein

RNase : Ribonucléase

RPP : « Ripening-related proteins »

RT-PCR : Amplification en chaîne par polymérase après transcription inverse

RT-qPCR : Amplification en chaîne par polymérase quantitative après transcription inverse

SA : Acide salicylique

SAR : Résistance systémique acquise

SE : Embryogenèse somatique

SEBF : « Silencing element binding factor »

SERK : « Somatic Embryogenesis Receptor Kinase »

sp. : espèce

SSCP : « Single-strand conformation polymorphism »

STAR : « Steroidogenic acute regulatory »

START : « STAR protein-related lipid transfer »

TAE : Tris acétate EDTA

Taq : Thermophilus aquaticus dont est extraite une ADN polymérase

TEV : Virus de la gravure du tabac

Tm : Température de fusion

TMV : Virus de la mosaïque de la vigne

TMVM : Virus de la marbrure des nervures du tabac

TOF : Point de décollage

TSS : Site d'initiation de la transcription

UTR : Région terminale non transcrite

WGD : Duplication totale du génome

XTH : Xyloglucane endotransglucosylase hydrolase

Sommaire

Remerciements

Liste des abréviations

Sommaire

Introduction générale **1**

Chapitre I : Caractéristiques moléculaires et fonctionnelles des protéines PR10 ..... **4**

A. Partie bibliographique **4**

Première partie : Présentation des PR10 ..... **5**

I. Les protéines PR10 sont des « pathogenesis-related proteins » 5

II. Caractéristiques conservées des protéines PR10, membres de la super-famille « Bet v 1-like » 6

III. Rôle physiologique des protéines PR10 8

IV. Structure des gènes *PR10* et organisation génomique 11

Deuxième partie : Origine et intérêt des familles multigéniques **13**

I. Processus de duplication à l'origine des membres des familles multigéniques 13

A. Duplication totale du génome ou polyploïdisation 13

B. Duplications partielles du génome 14

II. Devenir des gènes dupliqués 15

III. Évolution des familles multigéniques 16

B. Partie expérimentale **17**

I. État des connaissances sur les protéines PR10 de *Vitis vinifera* au début de la thèse. 18

II. Présentation de la publication
« Characterisation of the *Vitis vinifera PR10* multigene family » 19
Publication « *Characterisation of the Vitis vinifera PR10 multigene family* », Lebel et al. 2010, BMC Plant Biol **10:** 184 20

Chapitre II : Expression des gènes *PR10* pendant le processus d'embryogenèse somatique 37

A. Partie bibliographique 37

Première partie : L'embryogenèse somatique et ses marqueurs 38
I. L'embryogenèse somatique 38
II. Marqueurs de l'embryogenèse somatique 39
III. Description du modèle d'étude :
le cycle d'embryogenèse somatique secondaire chez *Vitis vinifera* 41

Seconde partie : Étude de l'expression de gènes par PCR après transcription inverse (RT-PCR) 43
I. La RT-PCR semi-quantitative 43
II. La RT-PCR quantitative (RT-qPCR) 44
A. Principe 45
B. Utilité de la RT-qPCR pour l'étude de l'expression de familles multigéniques 46
C. Étapes clés du protocole de RT-PCR quantitative 46

B. Partie expérimentale 48
I. Etat des connaissances sur les marqueurs de l'embryogenèse somatique chez *Vitis vinifera* au début de ma thèse 49
II. Presentation de la publication
« Differential regulation of SERK, LEC1-Like and Pathogenesis-Related genes during indirect secondary somatic embryogenesis in grapevine » 50

Publication « *Differential regulation of SERK, LEC1-Like and Pathogenesis-Related genes during indirect secondary somatic embryogenesis in grapevine* » Maillot et al. 2009, Plant Physiol Biochem **47:** 743-752 52

III. Étude complémentaire de l'expression des gènes *VvPR10* lors de l'embryogenèse somatique – Recherche de marqueurs de l'embryogenèse somatique de la vigne 62
A. Étude de l'expression de VvPR10.1-a à VvPR10.10 par RT-PCR semi-quantitative 62
1. Étude par RT-PCR semi-quantitative de l'expression des gènes *PR10* à différentes étapes du cycle d'embryogenèse somatique secondaire 63
2. Étude par RT-PCR semi-quantitative de l'expression des gènes *PR10* dans des tissus présentant différentes capacités embryogènes lors du traitement 2,4-D 64
B. Étude de l'expression des gènes *VvPR10* par RT-PCR quantitative 65
1. Validation des conditions de RT-qPCR 65
2. Étude par RT-qPCR de l'expression des gènes *PR10* à différentes étapes du cycle d'embryogenèse somatique secondaire. 66
3. Étude par RT-qPCR de l'expression des gènes *PR10* dans des tissus présentant différentes capacités embryogènes lors du traitement 2,4-D. 67
4. Comparaison des résultats de RT-PCR semi-quantitative avec les résultats de RT-qPCR. 68
IV. Discussion 70
A. Expression des gènes VvPR10 pendant l'embryogènese somatique chez la vigne 70
1. Expression durant le cycle d'embryogènese somatique secondaire 70
2. Expression des gènes *VvPR10* lors du traitement inducteur de la capacité embryogénique 71
B. Comparaison des résultats de RT-PCR semi-quantitative et de RT-qPCR 73

Conclusion générale et perspectives 76
I. Identification de la famille multigénique des *PR10* chez la vigne 77
II. Spécialisation et diversification fonctionnelle des gènes *VvPR10* 80
III. Marqueurs de l'embryogenèse chez la vigne 82

Matériels et méthodes 83
I. Matériel végétal 84
A. Embryogenèse somatique secondaire 84
B. Germination d'embryons 84
C. Traitement des tissus au 2,4-D 85
D. Prélèvement des échantillons végétaux 85
II. Biologie moléculaire 85
A. Extraction de l'ADN 85
B. Extraction et traitement de l'ARN 86
1. Extraction 86
2. Contrôle de qualité de l'ARN 86
3. Traitement à la désoxyribonucléase 87
C. Transcription inverse 87
D. Amplification en chaîne par polymérase 88
1. PCR « classique » 88
2. Amplification rapide des extrémités d'ADNc 88
3. PCR semi-quantitative 89
a. Conception des amorces 89
b.Calibrage des conditions de phase exponentielle 89
c. Réaction de PCR semi-quantitative 89
4. RT-PCR quantitative 90
a. Plan d'expérience 90
b. Informations sur les séquences cible et conception des amorces 91
c. Protocole de RT-qPCR 91
d. Validation de la qPCR 93
e. Traitement des résultats 94
E. Analyse des acides nucléiques sur gel d'agarose 95
1. ADN et produits amplifiée 95
2. ARN 95
F. Séquençage des produits de PCR 95
III. Bioinformatique 96
A. Identification in silico des séquences des gènes *PR10* potentiels dans le génome de la vigne 96

B. Recherche d'« Expressed Sequence Tags » (ESTs) correspondant aux séquences *PR10* candidates de la vigne 96
C. Prédiction des structures des gènes *PR10* 97
D. Analyse des séquences protéiques 97
E. Analyse phylogénétique et alignements multiples 97

Références bibliographiques 98

Introduction générale

Au cours des trois années qu'ont duré ma thèse, je me suis interessé à l'embryogenèse somatique chez *Vitis vinifera*, ainsi qu'à rechercher des marqueurs moléculaires de ce processus. De nombreux gènes différents avaient déjà été décrits comme marqueurs, dans diverses espèces végétales et principalement chez *Arabidopsis thaliana*, mais quasiment aucune étude n'avait été publiée dans ce domaine chez la vigne, au moment du début de cette thèse. Dans notre laboratoire, toutefois, la caractérisation de gènes *SERK* et *L1L*, exprimés dans des cultures embryogènes stabilisées de Chardonnay venait d'être réalisée (Schellenbaum et al. 2008).

Pendant ma thèse, j'ai analysé l'expression des gènes *PR10* de la vigne, qui nous ont semblé être de bons candidats marqueurs du processus d'embryogenèse somatique. En effet, bien que ceux-ci ne soient pas habituellement considérés comme tels, plusieurs études, menées sur plusieurs espèces végétales, montrent que leur expression est fortement régulée au cours de ce processus, et suggèrent qu'il puissent y jouer un rôle. Pour ce faire, j'ai utilisé le modèle cultural d'embryogenèse somatique secondaire, qui est un modèle d'embryogenèse récurrente, et qui permet de bien différencier deux phases cruciales du processus : celle supposée d'acquisition de la capacité embryogénique induite par un traitement au 2,4-D et celle d'entrée en différenciation des embryons secondaires.

Par ailleurs, le séquençage complet du génome de la vigne ayant été mis à disposition du public, dès 2007, sur le site du GENOSCOPE, je me suis attelé à identifier l'ensemble des gènes *PR10* présents dans le génome de la vigne. Il s'est agi d'un travail de recherche des séquences, de correction et de comparaison. Ensuite, j'ai cherché à savoir si les séquences trouvées étaient transcrites et j'ai analysé leur expression dans différents organes de la vigne, ainsi que pendant le processus d'embryogenèse somatique secondaire. En étudiant également la régulation de ces gènes dans des tissus présentant des capacités embryogènes différentielles, j'ai essayé de déterminer si certains pouvaient être des marqueurs de cette capacité. Ces études ont été menées à la fois par RT-PCR semi-quantitative et par RT-PCR quantitative, et l'efficacité de ces techniques, leur intérêt et leur pertinence pour cette étude ont été évaluées.

Mon manuscrit de thèse s'articule autour de deux chapitres, intitulés respectivement « Caractéristiques moléculaires et fonctionnelles des protéines PR10 » et « Expression des gènes *PR10* pendant le processus d'embryogenèse somatique ». Chacun de ces chapitres est composé d'une partie bibliographique présentant les sujets abordés au cours de celui-ci et d'une partie expérimentale, présentant les résultats obtenus pendant ma thèse.

La partie bibliographique du premier chapitre traite des protéines PR10 et synthétise les connaissances accumulées à leur sujet depuis 1988, date à laquelle le premier gène codant pour une protéine de cette famille a été caractérisé chez le persil (Somssich 1988), avant de proposer une description succincte des mécanismes à l'origine des familles multigéniques et de leur évolution. La partie expérimentale décrit l'identification et la caractérisation des gènes *PR10* présents dans le génome de la vigne, au travers d'une publication parue dans le journal BMC Plant Biology, en 2010 : « *Characterisation of the Vitis vinifera PR10 multigene family* ».

La partie bibliographique du second chapitre présente le processus d'embryogenèse somatique, et en particulier le processus d'embryogenèse somatique secondaire, puis, de manière résumée, les techniques de RT-PCR semi-quantitative et quantitative et leurs différences. La partie expérimentale est consacrée à la caractérisation moléculaire de l'embryogenèse somatique, d'abord par la présentation d'une publication parue dans le journal Plant Physiology and Biochemistry en 2009 : *« Differential regulation of SERK, LEC1-Like and Pathogenesis-Related genes during indirect secondary somatic embryogenesis in grapevine »*. Une analyse complémentaire, notamment par l'utilisation de la technique de PCR quantitative est détaillée ensuite.

Chapitre I : Caractéristiques moléculaires et fonctionnelles des protéines PR10

A. Partie bibliographique

Classe	Activité biochimique	Membre-type
PR-1	Inconnue	PR-1 (tabac)
PR-2	β-1,3-glucanase	PR-2 (tabac)
PR-3	Chitinase de type I, II, IV, V, VI, VII	P, Q (tabac)
PR-4	Chitinase de type I, II	'R' (tabac)
PR-5	Thaumatin-like	S (tabac)
PR-6	Inhibiteur de protéases	Inhibiteur I (tomate)
PR-7	Endoprotéinase	P_{69} (tomate)
PR-8	Chitinase de type III	Chitinase (concombre)
PR-9	Peroxidase	« Lignin-forming peroxidase » (tabac)
PR-10	Ribonuclease-like	PR1 (persil)
PR-11	Chitinase de type I	Chitinase "classe V" (tabac)
PR-12	Défensine	Rs-AFP3 (radis)
PR-13	Thionine	THI2.1 (arabidopsis)
PR-14	Protéine de transfert de lipides	LTP4 (orge)
PR-15	Oxalate oxidase	OxOa (orge)
PR-16	« Oxalate-oxydase-like »	OxOLP (orge)
PR-17	Inconnue	Prp27 (tabac)

Tableau 1 : Les 17 classes de protéines PR et leur activité biochimique (d'après van Loon et al. 2006).

Classe	Nombre d'allergènes décrits	Exemple
PR-1	1	Cuc m3 (melon)
PR-2	2	Ole e 9 (olivier)
PR-3	19	Pers a 1 (avocat)
PR-4	4	Hev b 6.01, pro-hévéine du latex
PR-5	6	Mal d 2 (pomme)
PR-8	1	Hévamine du latex
PR-9	1	Farine de blé
PR-10	24	Bet v 1 (bouleau)
PR-14	38	Pru p 3 (pêche)

Tableau 2 : Les 9 classes de protéines PR comprenant des protéines allergènes (d'après Malandain et Lavaud 2004).

Première partie : Présentation des PR10

I. Les protéines PR10 sont des « pathogenesis-related proteins »

Les protéines PR10 appartiennent au groupe des protéines PR, pour « pathogenesis-related », initialement définies comme étant exprimées par la plante hôte lors d'une attaque de pathogène (Antoniw et al. 1980). A la suite de leur identification dans des tissus de plantes soumis à des stress biotiques, mais aussi à divers stress abiotiques tels que le froid, les blessures, les métaux lourds, le stress osmotique ou divers traitements chimiques, les protéines PR sont considérées principalement comme des substances de défense spécifiques des plantes (pour revue, van Loon et van Strien 1999, Edreva 2006, van Loon et al. 2006). Cependant, l'accumulation de certaines d'entre elles dans des tissus sains lors d'étapes clés du développement (germination, floraison, maturation des fruits, sénescence) montre l'existence d'une régulation de leur expression en dehors de situations de stress.

Partageant des caractéristiques biochimiques communes telles qu'un point isoélectrique (pI) acide, une résistance à la dégradation protéolytique ou une faible masse moléculaire, les protéines PR sont regroupées dans 17 classes d'après leur structure primaire, leurs relations sérologiques, ainsi que leur activité biologique (Tableau 1). Si la plupart ont des propriétés biochimiques pouvant expliquer, au moins en partie, une action défensive contre divers parasites viraux, bactériens ou fongiques, certaines protéines PR ont un mode d'action encore inconnu, comme celles des classes 1 et 17. Parmi les protéines PR de classe 10, certaines montrent une activité ribonucléasique qui a pu être parfois associée à une activité anti-virale ou anti-fongique (pour revue, Liu et Ekramoddoulah 2006).

Plusieurs classes de protéines PR, comme de nombreuses autres protéines de défense, comprennent des allergènes présents dans le pollen et l'alimentation (pour revue, Malandain et Lavaud 2004 et Tableau 2). La classe PR10 est très représentée parmi les allergènes respiratoires du pollen des fagales ainsi que parmi les allergènes alimentaires des fruits mûrs, justifiant son étude détaillée. Le modèle de ces allergènes, puis, par voie de

Espèce	Nombre de gènes *PR10* identifiés	Référence
Phaseolus vulgaris (haricot)	20	Walter et al. 1996
Pinus monticola (pin argenté)	19	Liu et Ekramoddoullah 2003, Liu et Ekramoddoullah 2004
Malus domestica (pommier)	18	Gao et al. 2005
Fragaria ananassa (fraisier)	8	Mudislowska-Persson et al. 2006
Solanum surattense	6	Liu et al. 2006
Lupinus luteus (lupin jaune)	8	Handschuh et al. 2007
Prunus dulcis (amandier)	8	Chen et al. 2008
Prunus persica (pêcher)	8	Chen et al. 2008
Oriza sativa (riz)	5	Kim et al. 2008
Betula pendula (bouleau)	10	Schenk et al. 2009

Tableau 3 : Famille multigénique des *PR10* chez quelques espèces.

Ligand		Famille de protéines
Lipides membranaires	Phosphatidylcholine	START
		PITP
	Phosphatidylinositol	PITP
	Céramide	START
	Cholestérol	START
Hormones de plantes	Brassinostéroïdes	Bet v 1
	Cytokinines	Bet v 1
Métabolites II	Flavonoïdes	Bet v 1
	Dopamine + 4-hydroxyphénylacétaldéhyde	Bet v 1
	Émodine	Bet v 1
	Polykétides	Polykétides cyclases
	Ènediynes	« CalC-related »
	Hydrocarbures polycycliques aromatiques	Hydroxylases à chaîne alpha
		Hydroxylases homotrimériques

Tableau 4 : Ligands et substrats des membres de la super-famille « Bet v 1-like », déterminés expérimentalement (d'après Radauer et al. 2008).

conséquence, de l'ensemble des protéines PR10, est la protéine Bet v 1 du bouleau, bien que d'autres allergènes tels que les protéines Mal d 1 de la pomme ou Pru av 1 de la cerise aient été également largement étudiées. De même, des protéines LTP (PR14) et thaumatin-like (PR5) sont par exemple respectivement des allergènes bien connus du pollen de l'olivier ou de la pêche, et du pollen du cyprès ou de la pomme (Malandain et Lavaud 2004).

La classe des protéines PR10 a été initialement décrite en 1988 chez le persil *Petroselinum crispum*, à travers deux protéines très similaires, initialement classées parmi les PR1 : PR1-1 et PR1-3 (Somssich et al. 1988). Celles-ci, spécifiquement exprimées en réponse à une infection par le champignon *Phytophtora megasperma*, montraient des caractéristiques biochimiques nouvelles par rapport aux protéines PR habituellement rencontrées, justifiant la création d'une nouvelle classe de protéines PR. Présentant une faible homologie de séquence avec les protéines PR des autres classes, elles étaient, par ailleurs, dépourvues de peptide-signal N-terminal, suggérant une localisation cytoplasmique, contrairement aux autres protéines PR majoritairement retrouvées dans les espaces extracellulaires ou stockées dans des vacuoles au sein de la cellule. Dès cette date, des expériences d'hybridation génomique montraient l'existence de 2 à 6 gènes *PR10* chez le persil. Depuis, des gènes *PR10* ont été identifiés et caractérisés dans de nombreuses espèces ; ils forment de petites familles multigéniques, comme c'est le cas de la plupart des classes de protéines PR (pour revue, Liu et Ekramoddoulah 2006 et Tableau 3).

II. Caractéristiques conservées des protéines PR10, membres de la super-famille « Bet v 1-like »

Les protéines PR10 appartiennent à la super-famille structurale « Bet v 1-like », comprenant plusieurs familles de protéines représentées dans l'ensemble du monde vivant et partageant une grande homologie de conformation (Radauer et al. 2008 et Figure 1). Elles se composent classiquement d'un feuillet β enroulé autour d'une longue hélice α, ménageant ainsi une large cavité hydrophobique capable de lier des ligands lipidiques (Tableau 4). Cette cavité est une caractéristique commune à toutes les protéines de la

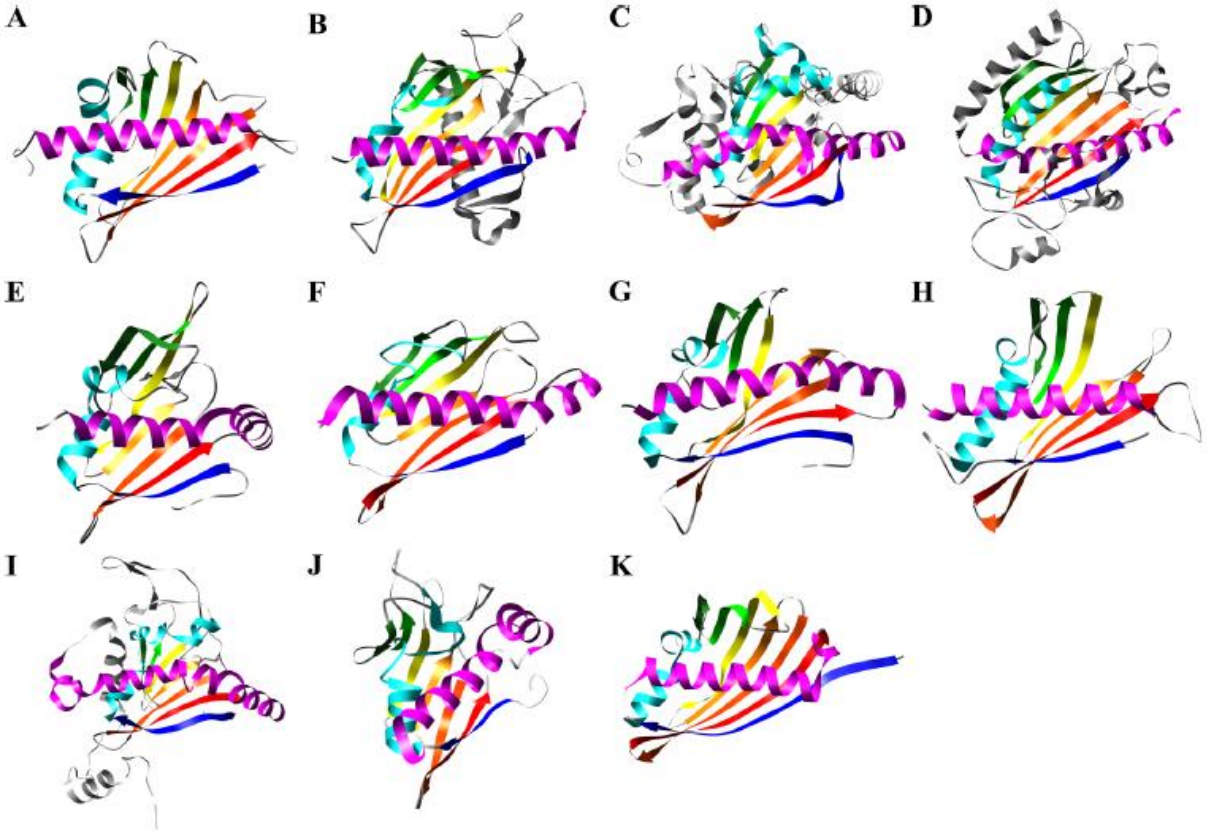


Figure 1 : Structure tridimensionnelle de membres représentatifs de la super-famille "Bet v 1-like". Les structures ont été colorées d'après les éléments de structure secondaire homologues à Bet v 1 (d'après Radauer et al. 2008).

A : Allergène majeur du pollen Bet v 1, *Betula pendula*; **B** : Domaine transporteur de lipides "STAR-related" de MLN64, *Homo sapiens*; **C** : Domaine C-terminal, α-chain, naphtalène 1,2 dioxygenase, *Pseudomonas putida*; **D** : Protéine alpha de transfert de phosphatidylinositol, *Ratus norvegicus*; **E** : Protéine hypothétique NE0264, *Nitrosomonas europaea*; **F** : Protéine hypothétique CC1736, *Caulobacter crescentus*, polykétide cyclase ; **G** : Protéine hypothétique SMU440, *Streptococcus mutans*; **H** : Protéine hypothétique PA1206, *Pseudomonas aeruginosa*; **I** : Domaine C-terminal de la 2-oxo-1,2-dihydroquinoline 8-monooxygènase, *Pseudomonas putida*, « homotrimeric ring hydroxylase » ; **J** Protéine de résistance CalC « self-sacrificing », *Micromonospora echinospora* ; **K** : Protéine hypothétique APE2225, *Aeropyrum pernix*, CoxG.

super-famille « Bet v 1-like », qui, malgré une très faible homologie des séquences, partageraient des similarités fonctionnelles. Ainsi, les membres de cette super-famille proviendraient d'une même protéine hypothétique présente chez le dernier ancêtre commun des différents groupes vivants actuels et ayant évolué indépendamment dans les différents phylums.

Dans cette super-famille, 2 familles sont confinées aux eucaryotes, dont la vaste famille « Bet v 1 » quasi-exclusivement végétale. La famille « Bet v 1 » est elle-même sub-divisée en sous-familles, dont celles des protéines PR10 des gymnospermes et des monocotylédones et la nombreuse famille des PR10 de dicotylédones (Figure 2). Deux petites sous-familles proches des protéines PR10 sont constituées des cytokinin-specific binding proteins (CSBP) spécifiques des Légumineuses et des norco-claurines synthétases (NCS). La sous-famille des major latex proteins (MLP) et ripening-related proteins (RPP) est plus éloignée phylogénétiquement et est très largement représentée chez *Arabidopsis thaliana*, qui curieusement ne contiendrait pas de protéines PR10. Ce dernier groupe des MLP/RPP comprend des protéines qui pourraient être impliquées dans la réponse au stress.

L'allergène du pollen de bouleau Bet v 1 fut la première PR10 cristallisée et purifiée en solution (Gahjede et al. 1996). La détermination de sa structure tridimensionnelle montre une protéine composée de trois hélices α, dont l'une localisée à l'extrémité C-terminale et longue d'environ 25 résidus, et d'un feuillet β formé de sept brins antiparallèles, délimitant les contours d'une cavité hydrophobique en forme de fourche de 30Å de profondeur (Figure 3-A). La structure d'autres protéines PR10 fut décrite par la suite chez le cerisier (Neudecker et al. 2001), le lupin jaune (Biesadka et al. 2002 et Pasternak et al. 2005), le céleri (Schirmer et al. 2005) et la carotte (Markovic-Housley et al. 2009), montrant une grande conservation de la structure tridimensionnelle, bien que de légères variations dans la forme et le volume des cavités hydrophobiques puissent être observées (Figure 3-B).

Les modèles tridimensionnels des protéines PR10 ont permis de mettre en évidence la capacité de la cavité hydrophobique à accueillir et lier de petites molécules. La taille et la configuration de la cavité de Pru av 1 pourraient par exemple lui permettre de lier des phytostéroïdes tels que l'homocastasterone (Neudecker et al. 2001), tandis que la cavité de Bet v 1 pourrait accueillir une vaste gamme de ligands tels que des phytohormones (cytokinines, déhydroergosterol), des acides gras (acide stéarique, acide oléique), et des flavonoïdes (flavone, 4',5,7-trihydroxyflavone) (Mogensen et al. 2002). Il a aussi été

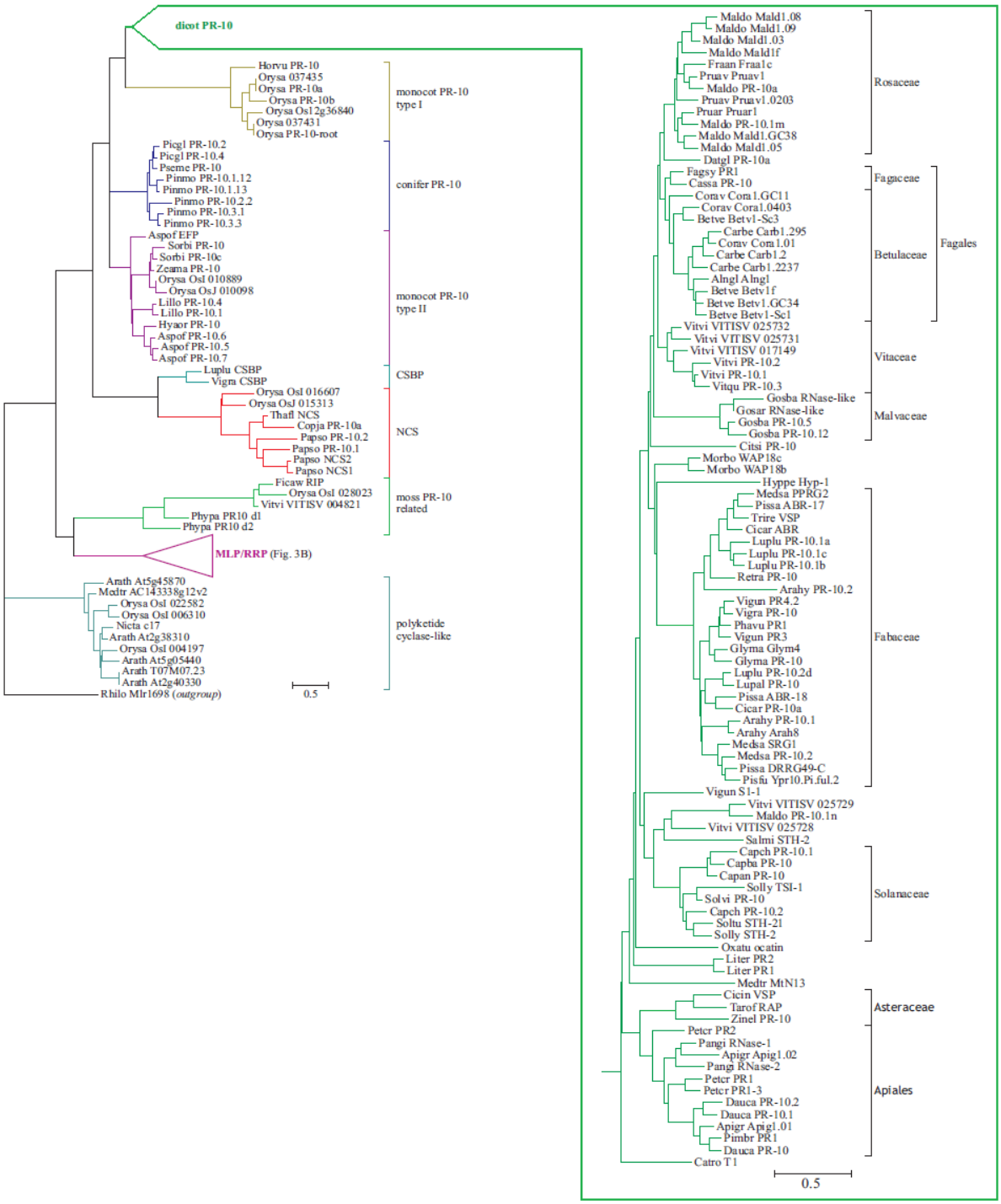
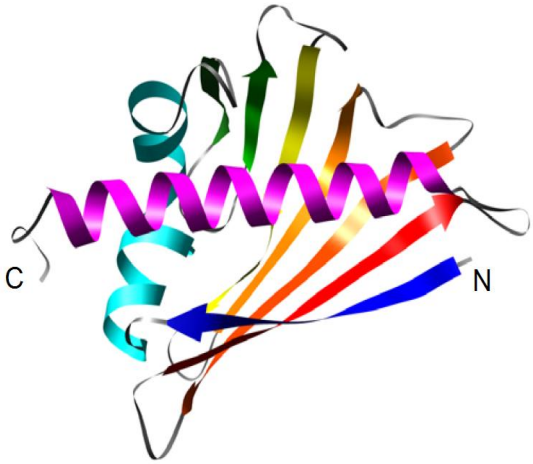
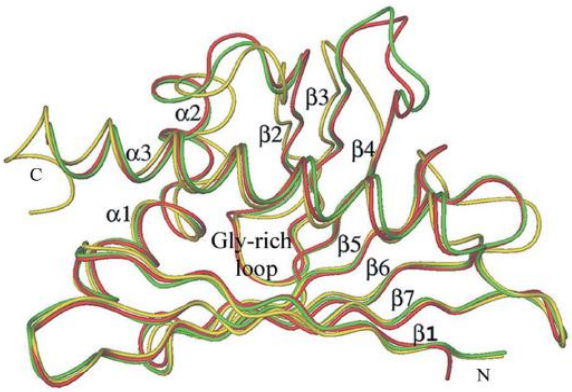


Figure 2 : Arbre phylogénétique « distance-based » de séquences de protéines de la famille Bet v 1 et de la sous-famille des polykétides cyclases-like.

Les séquences sont issues d'organismes végétaux. La seule séquence bactérienne est incluse en tant qu'"outgroup" (d'après Radauer et al. 2009).





A **B**

Figure 3 : Structure tridimensionnelle de protéines de la famille Bet v 1.

A : Représentation « en ruban » de la structure de la protéine Bet v 1. Les hélices α sont indiquées par des spirales tandis que les brins β sont indiqués par des flèches (d'après Radauer et al. 2009).

B : Représentation « en rubans » des structures des protéines Bet v 1 du bouleau (en jaune), Dau c 1 de la carotte (en rouge) et Api g 1 du céleri (en vert).

Espèce	Gène ou protéine induit	Pathogène éliciteur	Référence
Capsicum anuum (piment)	*CaPR-10*	*Xanthomonas campestris pv. vesicatoria* (b), TMV-P0 (v)	Park et al. 2004
Citrus sinensis (oranger)	PR10	*Xanthomonas campestris pv. citri* (b)	Garavaglia et al. 2010
Lupinus luteus (lupin jaune)	LlPR-10.1A, LlPR-10.B, *Ll-pr10.2a, Llpr-10.2b, Llpr-10.2e*	*Pseudomonas syringae* (b)	Sikorski et al. 1999, Handschuh et al. 2007
Malus domestica (pommier)	*Ypr10*	TMV (v), TEV (v), TVMV (v)	Pühringer et al. 2000
Medicago truncatula (luzerne tronquée)	PR10	*Aphanomyces euteiches* (o)	Colditz et al. 2007
Oriza sativa (riz)	*RPR10a, RPR10b, RPR10c*, OsPR-10	*Magnaporthe grisea* (c)	McGee et al. 2001, Kim et al. 2003
Petroselinum crispum (persil)	*PR1*	*Phytophthora megasperma* (o)	Somssich et al. 1989
Pisum sativum (pois)	*PR10.1, PR10.2, PR10.3*	*Fusarium solani* (c)	Tewari et al. 2003
Solanum surattense	*SsPR10*	TMV (v)	Liu et al. 2006
Vitis vinifera (vigne)	*VvPR10.1*	*Plasmopara viticola* (o), *Pseudoperonospora cubensis* (o), *Pseudomonas syringae* (b)	Kortekamp et al. 2006, Polesani et al. 2010, Robert et al. 2001
	VvPR10.1, VvPR10.2	*Erysiphe necator* (c)	Fung et al. 2008
	Homologue de *VqPR10.3*	*Plasmopara viticola* (o)	Polesani et al. 2008, Polesani et al. 2010
Zea mays (maïs)	*ZmPR10.1, ZmPR10.2*	*Erwinia stewartii* (b), *Erwinia chrysanthemi* (b), *Clavibacter michiganensis* (b)	Xie et al. 2010

Tableau 5 : Quelques exemples de protéines PR10 ou leurs gènes, dont l'expression est induite par des pathogènes.

b = bactérie, c = champignon, o = oomycète, v = virus

montré expérimentalement que la cavité hydrophobique de Bet v 1 pouvait contenir deux molécules de déoxycholate et celle de LlPR10 jusqu'à trois molécules de trans-zéatine (Markovic-Housley et al. 2003, Fernandes et al. 2008). Cette capacité à fixer de petits ligands apolaires, conservée chez tous les membres de la super-famille « Bet v 1-like », constitue probablement une des fonctions majeures des protéines PR10. Celles-ci pourraient alors avoir comme rôle de transporter ces ligands au sein de la cellule et/ou d'agir comme des « réservoirs » pour des hormones de plantes ou d'autres petites molécules.

La structure primaire des protéines PR10 est également assez bien conservée. Leur longueur est comprise entre 150 et 162 acides aminés, et leur masse moléculaire est comprise entre 15 et 18 kDa, tandis que leur potentiel isoélectrique, variant de 5,03 à 5,86, est légèrement acide (pour revue, Liu et Ekramoddoulah 2006). Leur séquence arbore deux principaux motifs conservés. Le premier est un motif conservé de 32 à 34 résidus, dont la séquence est G-x(2)-[LIVMF]-x(4)-E-x(2,3)-[CSTAENV]-x(8,9)-[GNDS]-[GS](2)-[CS]-x(2)- [KT]-x(4)-[FY]. Celui-ci est connu comme la signature Bet v 1 [PS00451], et est commun à tous les membres de la super-famille « Bet v 1 » (Osmark et al. 1998). Il est toutefois absent chez certaines protéines PR10 telles que ABR17 du pois ou la RNase 1 du ginseng (Bantignies et al. 2000). Le second motif conservé est à l'origine d'une « phosphate-binding loop » (P-loop). Sa séquence riche en glycines est GXGGXGXXK et se situe au niveau des acides aminés 46 à 54 chez *Lupinus luteus* et *Betula pendula* (pour revue, Liu et Ekramoddoulah 2006). Cette P-loop serait impliquée dans une activité ribonucléasique des PR10 grâce à sa capacité à se lier à des groupements phosphate (Saraste et al. 1990).

III. Rôle physiologique des protéines PR10

Une des caractéristiques fondamentales des protéines PR est que leur expression est induite lors de l'attaque par des pathogènes. Ceci a été démontré dans diverses espèces en présence de champignons, oomycètes, bactéries ou virus, ce qui suggère qu'elles joueraient un rôle dans la réponse aux stress biotiques (Tableau 5). Cependant, cette induction n'est pas systématique et semble varier selon le couple plante-pathogène. En effet, par exemple chez

Espèce	Protéine	Résidus impliqués dans l'activité ribonucléasique (en rouge, les résidus obligatoires)	Référence
Arachis hypogaea (arachide)	AhPR10	K54, F148, H150	Chadha et Das 2006
Astragalus mongholicus (astragale)	AmPR-10		Yan et al. 2008
Capsicum annuum (piment)	CaPR-10		Park et al. 2004
Gossypium arboreum (cotonnier)	GaPR-10	G51, K55, E96, E148, K150	Zhou et al. 2002
Oriza sativa (riz)	JIOsPR10		Kim et al. 2008
Pachyrrizus erosus (pois patate)	SPE-16		Wu et al. 2003
Panax ginseng (ginseng)	RNASE1, RNASE2		Moiseyev et al. 1997
Pinus monticola (pin argenté)	PmPR10	E103, E150, Y152	Liu et al. 2004
Pisum sativum (pois)	PsPR10.1		Srivastava et al. 2006
Solanum surattense	SsPR10		Liu et al. 2006
Theobroma cacao (cacaoyer)	TcPR-10		Pungartnik et al. 2008

Tableau 6 : Quelques protéines PR10 dont l'activité ribonucléasique a été démontrée, ainsi que les espèces dont elles proviennent.

le piment, l'expression du gène *CaPR-10* est induite lors de l'interaction compatible avec le TMV-P0 mais pas lors de l'interaction incompatible avec le TMV-P1,2 (Park et al. 2004). Les différentes PR10 d'une espèce peuvent également être induites de manière différentielle après l'attaque d'un pathogène. Chez le riz, l'expression de *RPR10a* est induite à deux reprises 12 heures puis 48 heures après inoculation par *Magnaporthe grisea* alors que celle de *RPR10b* n'est induite que 48 heures après inoculation (McGee et al. 2001). De plus, il semble qu'il puisse exister une régulation du niveau d'expression d'un même gène après induction par un éliciteur selon le tissu de la plante, comme chez *Solanum surattense* où l'expression de *SsPR10* est induite par le TMV dans les feuilles et les racines, mais pas dans les tiges (Liu et al. 2006).

Le rôle de défense des protéines PR10 pourrait être lié à une activité ribonucléasique observée dans différentes espèces (Tableau 6). Selon plusieurs auteurs, la fonction antivirale des protéines PR10 proviendrait d'un clivage de l'ARN viral, comme chez le piment (Park et al. 2004). De plus, cette activité ribonucléasique a pu parfois être mise en parallèle avec une activité antifongique. C'est par exemple le cas chez *A. hypogaea* et *T. cacao*, où l'activité antifongique est perdue si l'activité ribonucléasique est inhibée (Chadha et Das 2006, Pungartnik et al. 2008). Certaines protéines PR10 ne montrent en revanche pas ou peu d'activité ribonucléasique, comme les protéines LlPR10.1A et LlPR10.1B de *Lupinus luteus*, bien que leur expression soit induite par des pathogènes, permettant aux auteurs de suggérer que l'activité ribonucléasique des PR10 serait accidentelle et n'aurait pas de rôle biologique (Biesiadka et al. 2002). Chez le riz, il a par ailleurs été montré que l'activité antifongique de la protéine JIOsPR10 était indépendante de son activité ribonucléasique (Kim et al. 2008).

L'expression de certaines protéines PR10 ou de leurs gènes est également régulée lors d'interactions symbiotiques. Mais contrairement à ce que l'on observe lors de l'infection par un pathogène, l'expression de ces gènes est réprimée lors de ce type d'interaction. Ainsi, chez le lupin, l'expression des gènes *LlPR10.1A* et *LlPR10.1B* est réprimée dans les nodules jeunes et matures apparaissant suite à l'interaction avec *Bradyrhizobium* sp. *(Lupinus)* (Sikorski et al 1999). Chez *Pisum sativum*, lors de l'interaction symbiotique avec *Rhizobium leguminosarum* bv. *viciae* en carence de bore, l'expression de plusieurs gènes *PR10* est induite, alors que ni l'interaction avec la bactérie, ni une carence en bore n'entraînent seules cette induction. Cette induction pourrait être causée par une défaillance

Espèce	Gène ou protéine	Stress abiotique	Référence
Arachis hypogaea (arachide)	PR10	Stress salin	Jain et al. 2006
Lupinu luteus (lupin jaune)	*Ll-pr10.2a, Llpr-10.2b, Llpr-10.2e*	Blessure	Sikorski et al. 1999, Handschuh et al. 2007
Oriza sativa (riz)	OsPR-10	Blessure, froid	Kim et al. 2003
	RSOsPR10	Sécheresse, stress salin	Hashimoto et al. 2004, Kim et al. 2008
Phaseolus vulgaris (haricot)	*Ypr10*	Obscurité	Walter et al. 1996
Pinus monticola (pin argenté)	*PmPR10-1, PmPR10-2*	Blessure	Liu et Ekramoddoullah 2004
	PmPR10-3	Blessure, froid	
Pinus pinaster (pin maritime)	PR-10	Sécheresse	Dubos et Plomion 2001
Pisum sativum (pois)	*PR10*	Stress salin	Srivastava et al. 2004, Kav et al. 2004
Vitis vinifera (vigne)	RzPR10, PR10	Stress salin	Jellouli et al. 2008, Jellouli et al. 2010
Zea mays (maïs)	*ZmPR10.1, ZmPR10.2*	Obscurité, blessure, froid	Xie et al. 2010

Tableau 7 : Expression de protéines PR10 ou leurs gènes lors de stress abiotiques.

Espèce	Gène ou protéine	Eliciteur chimique	Référence
Gossypium arboreum (cotonnier)	*GaPR-10*	Acide salicylique	Zhou et al. 2002
Malus domestica (pommier)	*Ypr10*	Acide salicylique, glutathion	Pühringer et al. 2000
Oriza sativa (riz)	*OsPR10a*	Acide salicylique, acide abscissique, acide jasmonique, étéphon, H_2O_2	Hwang et al. 2008, Kim et al. 2003
Phaseolus vulgaris (haricot)	*Ypr10*	Acide salicylique, glutathion	Walter et al. 1996
Pisum sativum (pois)	*PsPR10.1*	Chitosane	Tewari et al. 2003
	PsPR10.4, *PsPR10.5*	Acide salicylique, acide abscissique, chitosane	
Solanum surattense	*SsPR10*	Acide salicylique, acide abscissique, acide gibbérellique, méthyl jasmonate, H_2O_2	Liu et al. 2006
Vitis vinifera (vigne)	VvPR10.1, VvPR10.2, Homologue de VqPR10.3	Chitosane, flumioxazine	Castro et al. 2005, Ferri et al. 2009
Zea mays (maïs)	*ZmPR10.1*, *ZmPR10.2*	Acide salicylique, kinétine, acide gibbérellique, acide 1-naphtylacétique, méthyl jasmonate, $CaCl_2$, H_2O_2	Xie et al. 2010

Tableau 8 : Expression de protéines PR10 ou leurs gènes après élicitation par des molécules chimiques.

Espèce	Gène ou protéine	R	F	T	Fl	M	Référence
Capsicum annuum (piment)	*CaPR-10*	nt	+	nt	nt	nt	Park et al. 2004
Gossypium arboreum (coton)	*GaPR-10*	+	-	-	-	nt	Zhou et al. 2002
Lupinus luteus (lupin jaune)	*LlPR-10.a*	+	-	+	nt	nt	Sikorski et al. 1999
	LlPR-10.b	+	+	+	nt	nt	
Malus domestica (pommier)	*Ypr10*	nt	+	nt	nt	nt	Pühringer et al. 2000
Oriza sativa (riz)	JIOsPR10	+	-	+	+	nt	Kim et al. 2008
Phaseolus vulgaris (haricot)	*Ypr10*	+	+	-	nt	nt	Walter et al. 1996
Pinus monticola (pin argenté)	*PmPR10-1.10*	+	+	+	nt	+	Liu et al. 2004
	PmPR10.2, PmPR10.3	-	-	+	nt	-	
Solanum surattense	*SsPR10*	+	+	+	nt	nt	Liu et al. 2006
Vitis vinifera (vigne)	RzPR10	+	+	+	nt	nt	Jellouli et al. 2008, Jellouli et al. 2010
	VvPR10.1	+	+	+	nt	nt	Kortekamp et al. 2006
Zea Mays (maïs)	*ZmPR10.1, ZmPR10.2*	+	+	+	nt	nt	Xie et al. 2010

Tableau 9 : Expression tissu-spécifique de gènes *PR10* dans quelques espèces végétales.

R = racines ; F = feuilles ; T = tiges ; Fl = fleurs ; M = méristème végétatif

(+) : expression détectée ; (-) : expression non détectée ; (nt) : expression non testée

dans la reconnaissance de la bactérie symbiotique, qui, en l'absence de bore, serait reconnue comme pathogène et activerait les défenses de la plante (Reguera et al. 2010).

D'autres facteurs peuvent également causer l'induction de l'expression des protéines PR10, comme des stress abiotiques (Tableau 7) ou l'application de certaines molécules chimiques, qui interviennent lors des mécanismes de défense, comme l'acide salicylique impliqué dans la cascade de signalisation de la SAR (réponse systématique acquise), ou le glutathion impliqué dans le stress oxydatif (Tableau 8). Par exemple chez le lupin, l'induction de l'expression des gènes *LlPR-10.2a, LlPR-10.2b* et *LlPR-10.2e* lors de l'infection par *Pseudomonas syringae* est comparable à l'induction observée lorsque la plante est blessée, suggérant que l'expression de ces gènes ne serait pas une réaction spécifique de l'infection par un pathogène, mais plus généralement une réponse à un stress (Handschuh et al. 2007). A l'opposé cependant, de nombreuses études montrent que l'expression des gènes *PR10* est régulée de manière différentielle en fonction du type de stress subi. Dans les aiguilles de *Pinus monticola,* l'expression de *PmPR10-1*, *PmPR10-2* et *PmPR10-3* est induite lors d'une blessure, alors que seule celle de *PmPR10-3* est induite lors de l'adaptation au froid hivernal (Liu et Ekramoddoullah 2004). De même, chez le riz, l'acide salicylique et l'acide jasmonique induisent l'expression de OsPR10a mais pas de OsPR10 (Kim et al. 2003). Au sein d'une même famille multigénique, les gènes *PR10* montreraient donc une diversification fonctionnelle, comme chez le pois, où l'expression des gènes *PR10.4* et *PR10.5* est induite par l'acide salicylique, l'acide abscissique ou le chitosane, alors que seule le chitosane induit l'expression de *PR10.1,* et qu'aucune induction n'est enregistrée pour *PR10.2* et *PR10.3* (Tewari et al. 2003).

Les protéines PR10 sont également exprimées en l'absence d'éliciteur, parfois en grande quantité dans certains organes. Ainsi, la protéine Bet v 1 est exprimée non seulement dans les tissus somatiques du bouleau mais surtout dans le pollen dont elle constitue un allergène majeur (Hoffmann-Sommergruber et al. 1997, Schenk et al. 2006, Schenk et al. 2009). De même, des protéines PR10 aux propriétés allergènes s'accumulent, en l'absence d'éliciteur, dans les fruits du pommier *Malus domestica* (Beuning et al. 2004). Il semble donc, d'après différentes études, que des gènes *PR10* puissent être exprimés de manière constitutive, mais aussi de manière tissu-spécifique (Tableau 9). Leur expression est le plus souvent détectée dans des feuilles et des racines, mais peut l'être également dans le tapetum des anthères, comme chez *Lilium longiflorum* (Huang et al. 1997) ou dans des graines, comme chez le pois patate, *Pachyrrhizus erosus* (Wu et al. 2003).

Gène	Espèce	Longueur de l'ORF (bp)	Longueur de l'intron (bp)	Référence
Bet v 1	*Betula pendula* (bouleau)	480 à 483	84 à 153	Schenk et al. 2006
Cor a 1	*Corylus avellana* (noisetier)	483	81	Hoffmann-Sommergruber et al. 1997
Llpr-10	*Lupinus luteus* (lupin jaune)	477	214 à 330	Handschuh et al. 2007
Mal d 1	*Malus domestica* (pommier)	480 à 483	111 à 171*	Gao et al. 1997
PmPR10-1.14	*Pinus monticola* (pin argenté)	486	164	Liu et Ekramoddoullah 2003
Pru p 1- Pru d 1	*Prunus persica* et *dulcis* (pêcher et amandier)	483	98 à 233	Chen et al. 2008
Pse m I	*Pseudotsuga menziesii* (mélèze)	486	154	Ekramoddoullah et al. 2000
RPR10a	*Oriza sativa* (riz)	477	366	McGee et al. 2001
VvPR10.1	*Vitis vinifera* (vigne)	477	121	Robert et al. 2001
Ypr10c	*Phaseolus vulgaris* (haricot)	468	103	Walter et al. 1996

*les gènes d'une sous-classe *Mal d 1* sont dépourvus d'intron.

Tableau 10 : Longueur de l'ORF et de l'intron de quelques séquences codant pour des protéines PR10.

IV. Structure des gènes *PR10* et organisation génomique

La structure des gènes *PR10* est largement conservée au sein des diverses espèces (pour revue, Liu et Ekramoddoulah 2006 et Tableau 10). Elle est composée d'une ORF de longueur comprise entre 453 et 486 bp, habituellement interrompue par un intron dont la position est également très conservée. Celui-ci se situe entre les nucléotides 1 et 2 du codon 62 dans le gène *Bet v 1*, le plus souvent employé comme référence (Breiteneder et al. 1989). La longueur de cet intron varie entre 81 et 366 bp, mais on peut noter qu'un groupe de gènes *Mal d 1* chez le pommier en est dépourvu (Gao et al. 2005).

La séquence des gènes *PR10* est également très conservée entre les espèces et encore plus au sein d'une même famille multigénique (Gao et al. 2005, Musidlowska-Persson et al. 2006), faisant de ces gènes d'excellents marqueurs phylogénétiques (Freitas et al. 2003, Finkler et al. 2005). En effet, cette similarité des séquences nucléotidiques (et *a fortiori* des séquences protéiques) est généralement plus forte à l'intérieur qu'à l'extérieur de l'espèce, même si plusieurs clades peuvent être observés dans une même espèce (Figure 4). Plusieurs motifs de régulation plus ou moins spécifiques ont pu être mis en évidence dans les régions promotrices des gènes *PR10* (Tableau 11). Tout d'abord, l'étude des gènes *PR1* du persil a permis de mettre en évidence les éléments W (ou W-box), sites de fixation pour des facteurs de transcription WRKY (Rushton et al. 1996). Par la suite, de nombreux autres motifs de régulation, impliqués soit dans la défense, soit dans le développement, ont été repérés dans les promoteurs de gènes *PR10* de différentes espèces. Ils pourraient être activés en réponse à des stress abiotiques (lumière, élictation par des molécules chimiques, blessure), à des attaques de pathogènes, ou à des traitements hormonaux (auxines, gibbérellines, acide abscissique, éthylène). Ces éléments pourraient aussi expliquer leur expression spécifique dans certains organes ou à certains stades du développement. Cependant, certains gènes dont l'expression a été bien caractérisée ne présentent pas les motifs attendus (Pühringer et al. 2000, Handschuh et al. 2007, Hwang et al. 2008), ce qui montre que ce domaine de connaissances reste encore très exploratoire.

Bien que la séquence génomique complète de nombreuses espèces végétales ne soit pas encore connue, on a pu déterminer que les gènes *PR10* pouvaient former de petites familles multigéniques de 5 à 20 gènes (Tableau 3). De plus, dans quelques cas, il a été montré

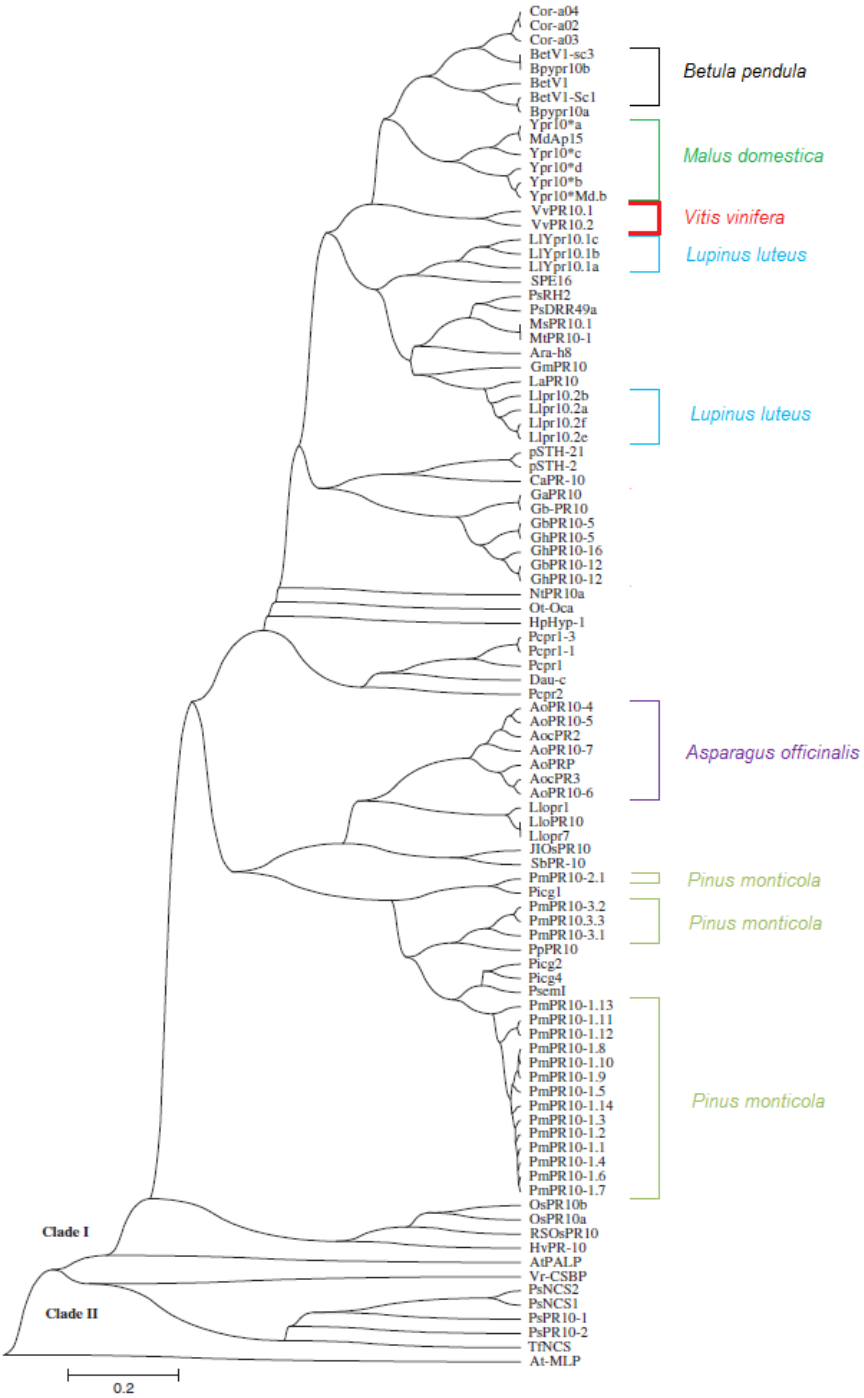


Figure 4 : Arbre phylogénétique construit à partir de l'alignement des séquences protéiques PR10 de diverses espèces végétales.

Certaines familles de protéines PR10 ont été mise en évidence afin de souligner la plus grande variabilité interspécifique qu' intraspécifique des séquences. Une protéine MLP d'*A. thaliana* a été incluse en tant qu'« outgroup » (d'après Liu et Ekramoddoullah 2006).

qu'ils étaient groupés sur un ou deux chromosomes. Chez le pommier, 16 des 18 gènes *Mal d 1* identifiés sont regroupés en tandem et répartis dans deux groupes de liaison (Gao et al. 2005). Les gènes *Pru p 1* du pêcher et *Pru d 1* de l'amandier sont groupés sur un unique groupe de liaison dans leur génome respectif (Chen et al. 2008). Chez le riz, les gènes *RPR10a*, *RPR10b* et *RPR10c* sont organisés en tandem et séparés par des fragments chromosomiques de 2 à 2,5 kb de long (McGee et al. 2001).

Motif identifié	Fonction	Gène	Espèce	Référence
Boîte W	Réponse antifongique, réponse à l'acide salicylique	*PR1-1, PR1-2*	*Petroselinum crispum* (persil)	Rushton et al. 1996
		OsPR10a	*Oriza sativa* (riz)	Hwang et al. 2008
		LlPR10.2b, LlPR10.2f	*Lupinus luteus* (lupin jaune)	Handschuh et al. 2007
		VvPR10.2	*Vitis vinifera* (vigne)	Robert et Buffard 2001
Site de fixation d'un élément SEBF	Réponse à l'auxine	*PR10a*	*Solanum tuberosum* (pomme de terre)	Boyle et al. 2001
Dof	Réponse au stress, à la lumière, aux pathogènes, germination, développement	*PmPR10-10.14*	*Pinus monticola* (pin argenté)	Yanagisawa et al. 2002, Liu et Ekramoddoullah 2003
		LlPR10.2b, LlPR10.2f	*Lupinus luteus* (lupin jaune)	Handschuh et al. 2007
bZIP	Réponse au stress, à la lumière, aux pathogènes, germination, développement	*PmPR10-10.14*	*Pinus monticola* (pin argenté)	Jakoby et al. 2002, Liu et Ekramoddoullah 2003
Boîte H	Expression tissu-spécifique, réponse au stress et aux facteurs fongiques	*LlPR10.2b, LlPR10.2f*	*Lupinus luteus* (lupin jaune)	Dron et al. 1988, Pastori et Foyer 2002, Handschuh et al. 2007
		VvPR10.2	*Vitis vinifera* (vigne)	Robert et Buffard 2001
« Hormone Responsive Element »	Réponse aux gibberellines	*LlPR10.2b, LlPR10.2f*	*Lupinus luteus* (lupin jaune)	Handschuh et al. 2007
GRE	Réponse au glucocorticoïdes	*LlPR10.2f*	*Lupinus luteus* (lupin jaune)	Handschuh et al. 2007
auxRE	Réponse à l'auxine	*LlPR10.2b*	*Lupinus luteus* (lupin jaune)	Handschuh et al. 2007
Boîte GCC	Réponse à l'éthylène	*VvPR10.2*	*Vitis vinifera* (vigne)	Sato et al. 1996, Robert et Buffard 2001
Site de fixation du facteur NF-κB	Réponse à l'acide salicylique	*VvPR10.2*	*Vitis vinifera* (vigne)	Lebel et al. 1998, Robert et Buffard 2001

Tableau 11 : Eléments de régulation identifiés dans les régions promotrices en 5' de quelques gènes *PR10*.

Deuxième partie : Origine et intérêt des familles multigéniques

Dans les différentes espèces étudiées, les gènes *PR10* forment de petites familles multigéniques de faible variabilité intraspécifique, ce qui suggère une certaine redondance génétique, dont il est intéressant de décrire l'origine et l'intérêt. « Les gènes partageant une similarité significative sont présumés avoir évolué à partir d'un même ancêtre commun et font partie de la même famille multigénique » (Huynen et al. 1998). Les membres d'une famille multigénique seraient donc issus de copies d'un unique gène-origine. La formation de familles multigéniques est un élément important de la dynamique évolutive des génomes (Moore et al. 2005). La formation de ces familles passe par des processus de duplication à grande ou petite échelle, créant des paralogues souvent organisés en cluster au sein des génomes (Hurles 2004).

I. Processus de duplication à l'origine des membres des familles multigéniques

A. Duplication totale du génome ou polyploïdisation

La polyploïdisation consiste en une duplication totale du génome (pour revue, Ramsey et Schemske 1998). Les évènements de polyploïdisation ont joué un rôle crucial dans la diversification d'une majorité de lignées eucaryotes animales ou végétales, et des angiospermes en particulier (de Bodt 2005). On estime qu'à l'exception d'*Amborella*, toutes les angiospermes, actuellement diploïdes ou polyploïdes, ont subi au moins une duplication complète du génome pendant leur histoire évolutive (Leitch et Bennett 1997, Ha et al. 2009, Soltis et al. 2009). Le génome d'*Arabidopsis thaliana* a, par exemple, connu trois épisodes de duplication totale au cours de son évolution, dont le plus ancien, produit de la contribution de trois génomes ancestraux, est probablement commun à toutes les angiospermes (Blanc et al. 2003, Bowers et al. 2003). Le génome de *Vitis vinifera* n'échappe pas à la règle et aurait évolué à partir d'un hypothétique ancêtre commun « paléo-hexaploïde » (Jaillon et al. 2007).

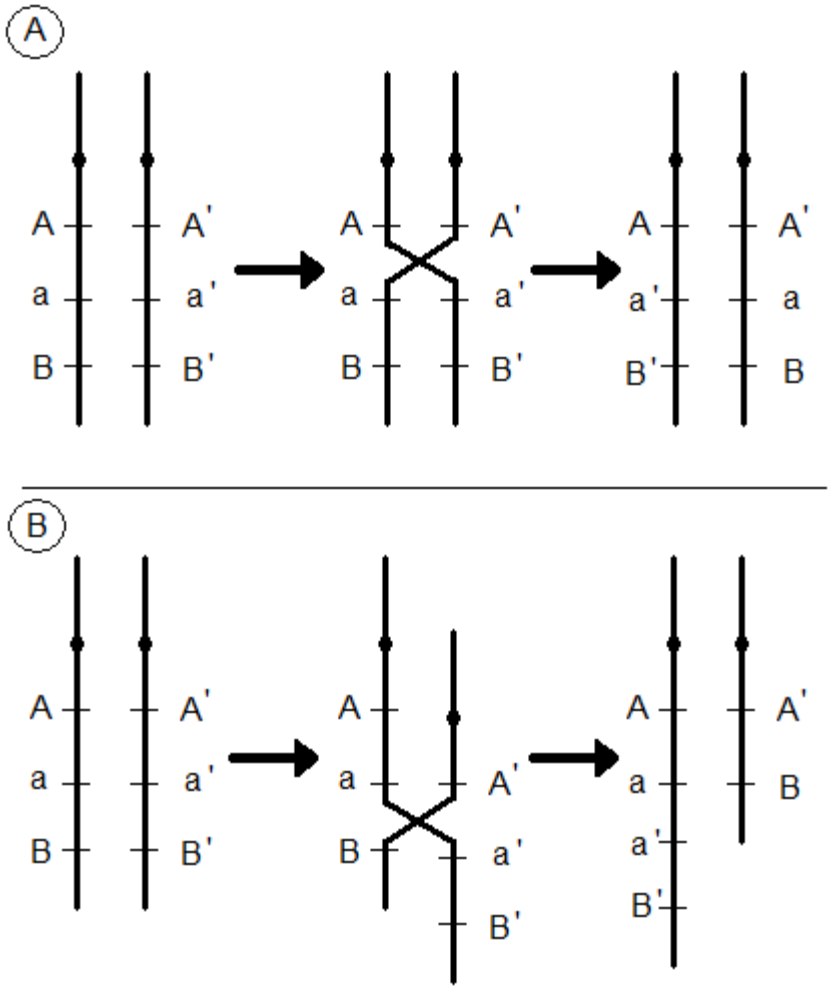


Figure 5 : Modèles de recombinaison homologue (A) et de "crossing-over" inégal (B) entre deux chromosomes homologues. Le crossing-over inégal est rendu possible par la forte similarité entre les gènes A et a.

Un événement de polyploïdisation est habituellement suivi d'une réduction du nombre de gènes, une majorité des gènes revenant à un exemplaire, ainsi que de réarrangements chromosomiques (Adams et al. 2005). Il a été montré, chez *Arabidopsis thaliana,* qu'après la plus récente polyploïdisation de son histoire évolutive, 28,6% seulement des duplicats de gènes ont été conservés (Thomas et al. 2006). La fréquence de rétention des duplicats dépend des familles ontologiques. Les plus fréquemment retenus sont des « connected genes » ou « dose sensitive genes », qui codent des protéines n'intervenant pas seules mais dans des complexes, ce qui implique qu'un ensemble de gènes doit être retenu pour assurer correctement la fonction assignée. Ces gènes peuvent, par exemple, coder des protéines intervenant dans des cascades de signalisation (facteurs de transduction de signal), ou des facteurs de transcription (Blanc et Wolfe 2004, Maere et al. 2005, Thomas et al. 2006).

B. Duplications partielles du génome

Les duplications partielles constituent la seconde force majeure aboutissant à la formation de familles multigéniques. Les duplications segmentales concernent des portions plus ou moins longues des chromosomes pouvant ensuite être dispersées dans le génome, mais les duplications à petite échelle semblent surtout être des duplications d'un (parfois plusieurs) gène(s) aboutissant généralement à la création de copies en tandem. Les duplications en tandem sont plus fréquentes que les duplications totales du génome. Elles surviennent principalement lors de recombinaisons homologues, ce qui implique une corrélation entre le taux de recombinaison d'une région chromosomique et la densité de gènes organisés en tandem ; les régions centromériques présentent ainsi la plus faible densité de duplications en tandem (Zhang et Gaut 2003). Ces duplications peuvent provenir de « crossing over » inégaux produisant des copies en orientation directe (Figure 5), de recombinaisons intrachromosomiques produisant des copies en orientation opposée, ou de glissements de réplication (Walsh 1987, Schuermann et al. 2005 , Rizzon et al. 2006). Le nombre de copies en tandem varie beaucoup selon la famille multigénique considérée. Chez *Arabidopsis thaliana* et *Oryza sativa*, on observe le plus fréquemment des répétitions de 2 ou 3 gènes, les plus longues comprenant respectivement 20 et 33 gènes (Zhang et Gaut 2003, Rizzon et al. 2006).

Certaines familles ontologiques se développent surtout grâce à des duplications en tandem. C'est par exemple le cas de familles de gènes impliqués dans la réponse aux stress biotiques ou abiotiques, comme les gènes *NBS-LRR* ou *MLP* chez *A. thaliana* (Baumgarten et al. 2003, Cannon et al. 2004). Les membres de la famille *Bet v 1* du bouleau sont également issus de duplications en tandem (Schenk et al. 2009). L'inflation du nombre de gènes dans ces familles pourrait correspondre à une

« course à l'armement » contre les pathogènes. De manière plus générale, les membres des familles multigéniques impliqués dans l'adaptation aux changements de l'environnement sont principalement issus de mécanismes de duplication en tandem (Hanada et al. 2008). Contrairement aux polyploïdisations, les duplications de gènes en tandem semblent se produire en continu au cours de l'évolution. De plus, les gènes issus de duplications en tandem présentent globalement plus de variabilité que les gènes issus de duplications complètes du génome, ce qui pourrait être en relation avec leur rôle dans l'adaptation aux conditions de l'environnement et aux pathogènes (Maere et al. 2005, Rizzon et al. 2006 et Hanada et al. 2008).

II. Devenir des gènes dupliqués

Suite aux différents processus de duplication, une faible part des copies est maintenue. Selon le modèle classique de « birth and death », une des copies conservera sa fonction d'origine tandis que de multiples mutations conduiront l'autre copie soit à dégénérer et perdre toute fonction (pseudogénisation ou nonfonctionalisation), soit à acquérir une nouvelle fonction (néofonctionalisation). La néofonctionalisation suppose qu'une des deux copies soit libérée de toute pression de sélection conservatrice jusqu'à l'acquisition d'une nouvelle fonction, moment où elle sera à nouveau soumise à des pressions de sélection (Ohno 1970, Hurles 2004). L'évolution peut aussi suivre d'autres voies. Si le gène d'origine porte plusieurs fonctions, celles-ci peuvent, après duplication, être partagées entre les différentes copies. Il s'agit du modèle de sous-fonctionalisation (Lynch et Force 2000, Duarte et al. 2006). Les deux copies sont par la suite nécessaires pour assurer la fonction totale du gène d'origine. Il est également possible qu'une fonction identique soit conservée pour chaque copie, mais que leur expression soit soumise à une régulation différente, conduisant à l'expression d'une même protéine, soit dans des tissus différents, soit à des stades de développement différents (sous-fonctionalisation spatio-temporelle). Cette expression différentielle est due à une variation dans les séquences régulatrices, pouvant être obtenue après une duplication n'impliquant pas l'intégralité des séquences promotrices (Lockton et Gaut 2005). Cependant, dans la plupart des cas, les copies dégénèrent par accumulation progressive de mutations et forment des pseudogènes. (Hurles 2004, Maere et al. 2005).

III. Evolution des familles multigéniques

Les familles multigéniques semblent devoir évoluer, sous l'effet de la sélection, vers une certaine divergence de leurs membres, condition supposée du maintien des duplicats dans l'organisme. Cette divergence des séquences est à l'origine d'une diversification fonctionnelle des protéines codées, par néo- ou sub-fonctionnalisation. Cependant, secondairement, les membres d'une même famille multigénique peuvent aussi évoluer de manière concertée, ce qui signifie qu'ils évoluent en parallèle, de manière à montrer une plus grande similarité intraspécifique qu'interspécifique (Liao 1999, Ohta 2000). L'évolution concertée au sein d'une famille multigénique ne serait pas due au hasard mais plutôt à la nécessité d'assurer le maintien de l'homogénéité des séquences de la famille (Liao 1999). D'une part, le modèle d'évolution concertée prévoit qu'une mutation favorable survenant sur un des membres d'une famille multigénique puisse être rapidement transmise aux autres, sous l'effet d'une pression de sélection positive (Nei et Rooney 2005). D'autre part, l'évolution concertée au sein d'une famille multigénique pourrait être la conséquence de l'élimination de mutations délétères dans une séquence cruciale d'un des membres de la famille, dans le contexte d'une pression de sélection conservatrice (Liao 1999). Le mécanisme principal aboutissant à une évolution concertée serait l'homogénéisation intrachromosomique, par « crossing-over » inégal ou conversion génique, mais des échanges interchromosomiques, plus rares, existent également. La proximité physique des séquences disposées en tandem, dans le cas de certaines familles multigéniques, est favorable à ce type d'évolution (Rizzon et al. 2006). L'étude phylogénétique des membres des familles multigéniques de *PR10* de différents organismes végétaux suggère qu'ils suivent habituellement un modèle d'évolution concertée (Wen et al. 1997, Finkler et al. 2005, Schenk et al. 2006).

Chapitre I : Caractéristiques moléculaires et fonctionnelles des protéines PR10

B. Partie expérimentale

I. État des connaissances sur les protéines PR10 de *Vitis vinifera* au début de la thèse

Lorsque ma thèse a démarré, deux gènes, *VvPR10.1* et *VvPR10.2*, avaient été identifiés comme induits lors de l'interaction non-hôte avec la bactérie *Pseudomonas syringae* pv. *Pisi*, dans le cépage Ugny blanc (Robert et al. 2001, Robert et Buffard 2001). *VvPR10.1* et *VvPR10.2* sont très similaires, avec respectivement une ORF de 480 ou 477 bp (coupée par un intron de 121 ou 107 bp entre les nucléotides 184 et 185) codant une protéine de 159 ou 158 acides aminés. Quelques autres études publiées ont montré que l'expression des gènes *PR10* de vigne (ou des protéines correspondantes) pouvait être induite en réponse à des stress biotiques ou abiotiques. Dans ces études, les gènes et protéines induits ont été décrits comme similaires aux deux gènes *VvPR10.1* et *VvPR10.2*, tels que référencés dans les banques (*VvPR10.1* : Robert et al. 2001, GenBank: AJ291705.2 ; *VvPR10.2* : Robert et al. 2001, GenBank: AJ291704.2). Il a ainsi été montré que l'oomycète *Plasmopara viticola*, agent du mildiou de la vigne, pouvait induire l'expression de gènes *PR10* dans les cépages Riesling (sensible) et Gloire de Montpellier (résistant), avec toutefois une induction plus précoce dans le cépage sensible, (Kortekamp et al. 2006). Par ailleurs, il a été décrit que l'herbicide flumioxazine entraînait une induction de l'expression de protéines PR10 dans des plantules traitées du cultivar *Chardonnay* (Castro et al. 2005). Dans cette étude protéique, le nombre de spots correspondant à des protéines PR10 observé sur gel était supérieur au nombre de gènes *PR10* caractérisés, suggérant soit l'existence d'une famille multigénique à l'origine des protéines PR10 de vigne, soit l'existence de modifications post-traductionnelles des protéines PR10.1 et PR10.2.

II. Présentation de la publication « Characterisation of the *Vitis vinifera PR10* multigene family »

Au moment du démarrage de ce travail, bien que le génome de peu d'espèces végétales soit connu, l'étude du pommier montrait que les gènes *PR10* forment une famille multigénique dont les membres sont regroupés en tandem sur deux chromosomes (Gao et al. 2005). Étant intéressé par cette famille de gènes *PR10*, classés chez la vigne comme gènes de défense mais potentiellement impliqués dans l'embryogenèse somatique (Imin et al. 2005, Domoki et al. 2006, De Jong et al. 2007), j'ai saisi l'occasion qu'offrait la mise à disposition de la séquence intégrale du génome de *Vitis vinifera*, dès 2008, pour caractériser la famille des gènes *PR10* dans cette espèce.

Dix-sept séquences *PR10* comprenant quatorze séquences complètes et trois pseudogènes ont été identifiées. Elles sont regroupées dans le génome, en un groupe compact de séquences disposées en tandem, long de quatre-vingt dix kb environ. Ces séquences se présentent, de 5' vers 3' : les douze premières et la dernière en orientation directe, et entre, quatre séquences en orientation inverse. Cette disposition reflète probablement une succession d'évènements de duplications en tandem ayant eu lieu pendant l'évolution. Une analyse phylogénétique des séquences, en comparaison avec des gènes *PR10* d'autres espèces, montre une plus faible variabilité intraspécifique qu'interspécifique, suggérant une probable évolution concertée, et/ou de récentes duplications. Une étude de l'expression de dix de ces séquences dans différents organes de la plante : feuilles, tiges, racines, inflorescences, et dans des embryons somatiques, montre la fonctionnalité d'au moins neuf d'entre elles et une certaine spécificité d'expression. L'induction de certains de ces gènes lors du traitement d'embryons au 2,4-D (traitement inducteur de l'embryogenèse somatique secondaire), confirme leur caractère inductible et leur classement comme gènes de défenses.

Characterisation of the Vitis vinifera PR10 multigene family, Lebel S, Schellenbaum P, Walter B, Maillot P, 2010 *BMC Plant Biol* **10:** 184

Lebel et al. BMC Plant Biology 2010, 10:184
http://www.biomedcentral.com/1471-2229/10/184

BMC Plant Biology

RESEARCH ARTICLE Open Access

Characterisation of the *Vitis vinifera* PR10 multigene family

Sylvain Lebel, Paul Schellenbaum, Bernard Walter, Pascale Maillot*

Abstract

Background: Genes belonging to the *pathogenesis related 10* (*PR10*) group have been studied in several plant species, where they form multigene families. Until now, such an analysis has not been performed in *Vitis vinifera*, although three different *PR10* genes were found to be expressed under pathogen attack or abiotic stress, and during somatic embryogenesis induction. We used the complete genome sequence for characterising the whole *V. vinifera PR10* gene family. The expression of candidate genes was studied in various non-treated tissues and following somatic embryogenesis induction by the auxin 2,4-D.

Results: In addition to the three *V. vinifera PR10* genes already described, namely *VvPR10.1*, *VvPR10.2* and *VvPR10.3*, fourteen different *PR10* related sequences were identified. Showing high similarity, they form a single cluster on the chromosome 5 comprising three pseudogenes. The expression of nine different genes was detected in various tissues. Although differentially expressed in non-treated plant organs, several genes were up-regulated in tissues treated with 2,4-D, as expected for *PR* genes.

Conclusions: *PR10* genes form a multigene family in *V. vinifera*, as found in birch, apple or peach. Seventeen closely related *PR10* sequences are arranged in a tandem array on the chromosome 5, probably reflecting small-scale duplications during evolution. Various expression patterns were found for nine studied genes, highlighting functional diversification. A phylogenetic comparison of deduced proteins with PR10 proteins of other plants showed a characteristic low intraspecific variability. Particularly, a group of seven close tandem duplicates including *VvPR10.1*, *VvPR10.2* and *VvPR10.3* showed a very high similarity, suggesting concerted evolution or/and recent duplications.

Background

PR10 proteins belong to the huge family of pathogenesis related (PR) proteins ubiquitous in the plant kingdom. PR proteins were first identified as defence molecules produced in response to pathogen attack and some of them actually display an antimicrobial activity. However, numerous studies have reported their induction under a great variety of abiotic stress conditions as well as possible constitutive or developmentally regulated expression [1]. Sharing common biochemical characteristics (acidic pI, resistance to proteolytic degradation, small molecular mass) PR proteins are divided into seventeen different groups based on their primary structure, serological relationships and biological activity [2]. Most of them are extracellular, but some others are found in the cytoplasm, mainly in the vacuole. PR10 proteins present the specificity to be free in the cytoplasm and are therefore classified as intracellular PR (IPR) proteins. They are closely related to a group of major tree pollen allergens and food allergens, that belong to the Bet v 1-like superfamily [3].

PR10 genes form multigene families with low intraspecific variation and higher interspecific variation that make them interesting phylogenetic markers [4-6]. Some of them were shown to be organized in chromosome clusters [7,8]. Characterised in a number of plant species, most *PR10* genes share an open reading frame (ORF) from 456 to 489 bp interrupted by an intron of 76-359 bp at a highly conserved position [9]. This ORF codes for an acidic small protein with conserved sequence features: three amino acids E96, E148 and Y150 (as positioned in Bet v 1) possibly involved in

* Correspondence: pascale.maillot@uha.fr
Université de Haute Alsace, Laboratoire Vigne, Biotechnologies & Environnement, 33 rue de Herrlisheim, BP 50568, 68 008, Colmar Cedex, France

BioMed Central

© 2010 Lebel et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lebel et al. BMC Plant Biology 2010, 10:184
http://www.biomedcentral.com/1471-2229/10/184

ribonucleasic activity and two other remarkable domains comprising the motif GXGGXGXXK (aa 47-55) that forms a P-loop supposed to have NTPase activity and the Bet v 1 motif [PS00451] characteristic of proteins from the Bet v 1 superfamily [9].

The significance of multiple close copies of a gene in a single plant species has to be clarified. In birch or yellow lupine, some *PR10* genes are constitutively expressed while others are induced under pathogen attack, abiotic stress or during plant development, suggesting functional diversification [10,11]. A significant common feature of PR10 proteins is a large Y-shaped hydrophobic cavity, as shown by the determination of three-dimensional structure [3,12-15]. This internal cavity could be responsible for the intracellular transport of apolar ligands, so diverse as fatty acids, flavonoids, cytokinins or brassinosteroids [10,14,16,17]. Slight modifications of the structure and shape of this cavity would allow to bind different ligands, what could account for the diverse roles hypothesised for PR10 proteins in plant defence and development [12,13].

To date, three different *PR10* genes have been described in *Vitis vinifera*. *VvPR10.1* was shown to be up-regulated during a pathogen interaction with *Pseudomonas syringae* in the cultivar Ugni blanc [18], while the expression of *VvPR10.1*, *VvPR10.2* and *VvPR10.3* was detected during somatic embryogenesis (SE) induction in the cultivar Chardonnay [19]. Moreover, several studies showed a strong and specific production of PR10 proteins in *V. vinifera* under salt or herbicide stress [20,21], as well as after fungal attack [22,23].

The recent publication of the whole genome sequence of *V. vinifera* by the Genoscope [24] and the Instituto Agrario San Michele all'Adige (IASMA) [25] makes genome scale analyses possible. Using these databases, we characterised all sequences of the *PR10* gene family of *V. vinifera* and monitored the expression of nine candidate genes in various tissues and conditions.

Results

In silico identification of *V. vinifera PR10* related sequences

We searched for all *V. vinifera PR10* related sequences in the whole genomic sequence published by the Genoscope. An automatic list of chromosome regions containing *PR10* and *PR10-like* gene annotations was produced. After correction for mis- and redundant annotations, ten sequences with a complete ORF were retained, comprising *VvPR10.1* and *VvPR10.2*, previously identified (AJ291705 and AJ291704 respectively) and the recently described *VvPR10.3* gene (EU379313) [19]. Three additional incomplete sequences corresponded to pseudogenes. All thirteen sequences were arranged in a tandem array on the chromosome 5. On both sides of this cluster, we could identify four more *PR10* related sequences, omitted by automatic annotation. No additional sequence was retrieved by a further homology search, suggesting that all *V. vinifera PR10* genes are grouped in this single cluster, approximately 80,000 pb long and limited by putative genes not related to *PR10* genes. The seventeen *PR10* related sequences were named s1 to s17 and numbered according to their position on the chromosome regardless the strand they were located on (Figure 1): the twelve first (s1 to s12) and the last one (s17) on the minus (-) strand of the chromosome, and s13 to s16 on the plus (+) strand. *VvPR10.1*, *VvPR10.2* and *VvPR10.3* respectively correspond to s16, s10 and s12. All Genoscope references and precise locations are given in the additional file 1. The seventeen sequences collected from the Genoscope (PN 40024) were compared to their counterparts in the IASMA database (Pinot noir ENTAV 115), which are scattered on fifteen different scaffolds [additional file 2]. Each sequence was less than 2% divergent from its counterpart, highlighting a high similarity between the two genomes and databases.

Characterisation and comparison of nucleotide sequences

The putative exon-intron structure of the seventeen *PR10* related sequences was determined (Figure 2a). Each sequence is interrupted by a single intron at a very well conserved position, presenting consensus motifs at the 5′ and 3′ boundaries with exons (Figure 2b). The hypothetical CDS are 204 to 486 bp long. The length of exons is well conserved, except for s6, s7 and s9, which are shortened by a premature stop codon suggesting that they are pseudogenes. In addition, s9 presents a single nucleotide deletion in the first exon, which induces a frame shift definitely compromising a correct translation. The structure of the fourteen sequences with a complete ORF is in accordance with previous reports on *Vitis vinifera* and other plant *PR10* genes [9,18,19].

The CDS were aligned for comparison. The percentages of nucleotide similarity are given in Figure 3. All sequences are very similar and show at least 48% of nucleotide similarity. A very high similarity (≥92%) was found between the six sequences s11-s16, including *VvPR10.1* and *VvPR10.3*, that are also strongly similar to s10-*VvPR10.2* (87-91% of nucleotide similarity), together forming a group of closely related sequences. The sequence of the pseudogene s6 is also remarkably similar to s5 (91%). The intron sequence is more variable and the similarity percentages range from 8 to 100%, being very high for the seven sequences s10-s16 comprising *VvPR10.1*, *VvPR10.2* and *VvPR10.3* on one hand, and for s5 and s6 on the other hand [additional file 3].

Prediction of expression

In order to determine whether the newly identified *PR10* related sequences could be functional genes, we

Lebel *et al. BMC Plant Biology* 2010, **10**:184
http://www.biomedcentral.com/1471-2229/10/184

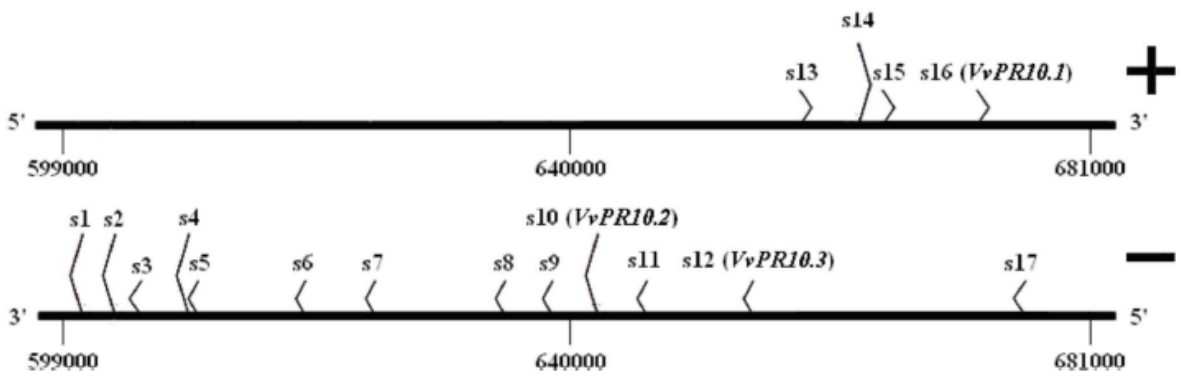


Figure 1 Organisation of *V. vinifera PR10* related sequences in a single cluster on chromosome 5. All sequences are located between the positions 599,000 and 681,000, on the + or the - strand. Nucleotide positions are as referenced on the Genoscope website.

searched for related Expressed Sequence Tags (ESTs) in the databases (Table 1). Consistent with the prediction of s6, s7 and s9 as pseudogenes, no EST was found to match these incomplete sequences. In addition, we found no EST corresponding to s2, s4, s8 nor s17, suggesting that these sequences would not be transcribed, at least in the tissues and conditions reported in the databases. On the contrary, numerous different ESTs (> 100) were found to match s11, s13, s14 and s15 with up to 99% nucleotide similarity. Although some of them also matched the highly homologous *VvPR10.1*, *VvPR10.2* and *VvPR10.3*, the expression of which was reported [18,19], other ESTs were closer to s11, s13, s14 and s15, that would therefore be expressed in varied reproductive and/or vegetative tissues. Several ESTs corresponding to s1, s3 and s5 were also found (1, 52 and 4, respectively). Although a single incomplete EST (27% of coverage) matched s1, its sequence was 100% similar to the sequence of s1. Moreover, this EST covered the last 135 nucleotides of the CDS and 3'UTR, which are usually the most varying parts of a gene sequence, suggesting that this EST is a specific partial transcript of s1. The ESTs matching s3 and s5 with up to 99% similarity and 100% coverage argue for the expression of these

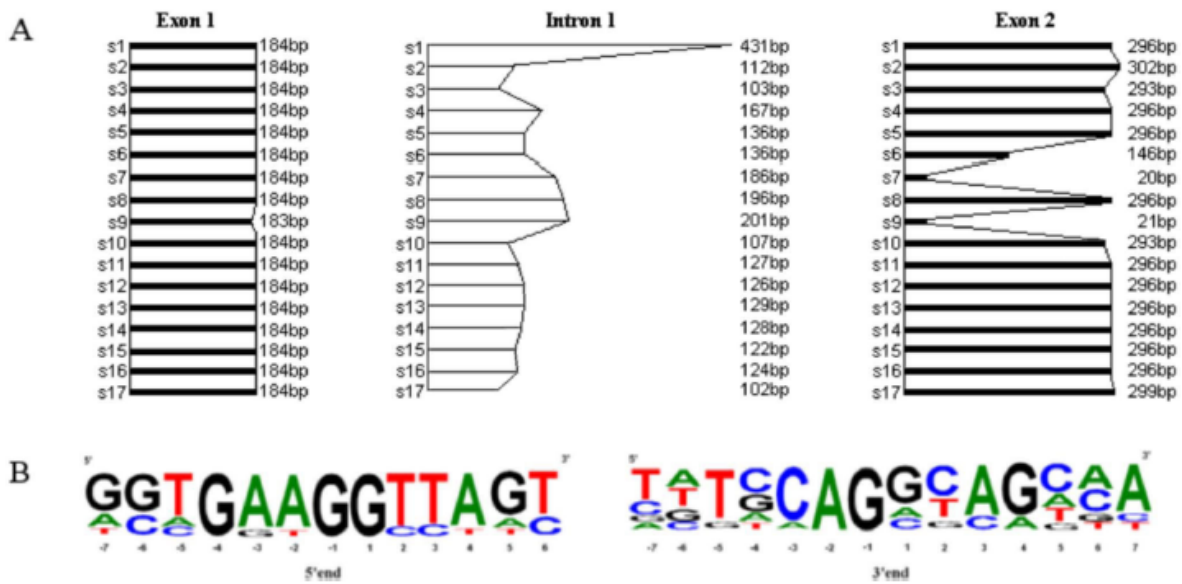


Figure 2 Exon-intron structures of the seventeen *PR10* related sequences. A: length of exons (bold dashes) and introns (thin dashes). B: frequency of nucleotides at the 5' and 3' putative splicing sites of introns. The size of letters is proportional to the nucleotide frequency at each position. The numbers indicate nucleotide position relatively to the first nucleotide of intron 1 (number 1 at the 5' end) or to the first nucleotide of exon 2 (number 1 at the 3' end). *VvPR10.1*, *VvPR10.2* and *VvPR10.3* respectively correspond to s16, s10 and s12.

Lebel *et al. BMC Plant Biology* 2010, **10**:184
http://www.biomedcentral.com/1471-2229/10/184

CDS	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15	s16	s17
s1		66	65	64	66	62	56	61	63	61	63	61	61	60	63	62	66
s2			63	59	60	58	48	56	56	55	60	57	57	57	59	58	65
s3				71	71	71	63	68	71	66	67	68	68	67	68	68	63
s4					85	81	77	80	83	77	77	75	75	74	77	77	61
s5						91	76	81	82	78	80	77	77	77	81	80	62
s6							75	80	80	75	76	75	75	75	77	76	60
s7								76	78	75	72	70	70	70	72	71	60
s8									81	78	82	79	79	78	82	81	60
s9										79	79	78	77	77	79	79	58
s10											90	87	87	88	90	91	58
s11												93	93	94	98	98	55
s12													98	97	92	93	60
s13														97	92	94	60
s14															94	95	59
s15																97	59
s16																	60
s17																	

Figure 3 Percentages of nucleotide similarity between the CDS of the seventeen sequences. High percentages of nucleotide similarity are highlighted in red. *VvPR10.1*, *VvPR10.2* and *VvPR10.3* respectively correspond to s16, s10 and s12. The values were obtained from sequence alignments on ClustalW.

Table 1 ESTs from *Vitis* spp. related to the seventeen *PR10* sequences

Sequence	Number of ESTs	Best results		Best results found in
		Coverage %	Similarity %	
VvPR10.1	> 100	100	100	F, L, R, Be, Bu
VvPR10.2	> 100	100	100	F, L, R, Be, Bu
VvPR10.3	> 100	100	100	F, L, R, Be, Bu
s1	1	27	100	Be
s2	0			
s3	52	100	99	F, L, R, Be
s4	0			
s5	4	100	99	F, L, R, Be
s6	0			
s7	0			
s8	0			
s9	0			
s11	> 100	100	99	F, L, R, Be, Bu
s13	> 100	100	99	F, L, R, Be, Bu
s14	> 100	100	99	F, L, R, Be, Bu
s15	> 100	100	99	F, L, R, Be, Bu
s17	0			

BLAST results were obtained for the putative CDS of the seventeen sequences using the Vitis EST database of NCBI. The number of ESTs indicates the number of matching results for each query. The percentage of coverage indicates the relative length of coverage of the sequence by the best matching EST. The percentage of nucleotide similarity is given for the best matching EST. F = flowers; L = leaves; R = roots; Be = berries; Bu = buds.

two sequences. Using the Gene Finding program, a correct gene structure was predicted for the seven sequences s4, s5, s8, s11, s13, s14 and s15, in addition to *VvPR10.1*, *VvPR10.2* and *VvPR10.3* (Figure 4). All possess the well positioned TATA-box, transcription starting site (TSS) and polyadenylation signal, defining a CDS of 477-480 bp. No promoter was found for s1, s2, s3 and s17. When a manual search was performed, no TSS-motif was detectable for these sequences, but a TATA-box was identified for s1, s3 and s17 at the respective positions -1,214, -699 and -313 nucleotides. On the whole, in addition to the three well characterised *V. vinifera PR10* genes, seven *PR10* related sequences showed the canonical transcription signals, suggesting their possible expression, although no specific EST was found to match s4 nor s8. In contrast, s1 and s3 did not show the features usually described for effective transcription, but some specific ESTs were identified in the databases, indicating their probable expression. Finally, s2 and s17 showed no usual transcription signals, and no EST was found to match these two sequences.

Characterisation of deduced PR10 proteins and phylogenetic analysis

The nucleotide sequences with a complete ORF were named according to usual nomenclature. The highly similar sequences *VvPR10.1*, *s15* and *s11* on one hand, and *VvPR10.3*, s13 and s14 on the other hand (≥97% nucleotide similarity) were named *VvPR10.1-a*, *-b* and

Lebel *et al. BMC Plant Biology* 2010, **10**:184
http://www.biomedcentral.com/1471-2229/10/184

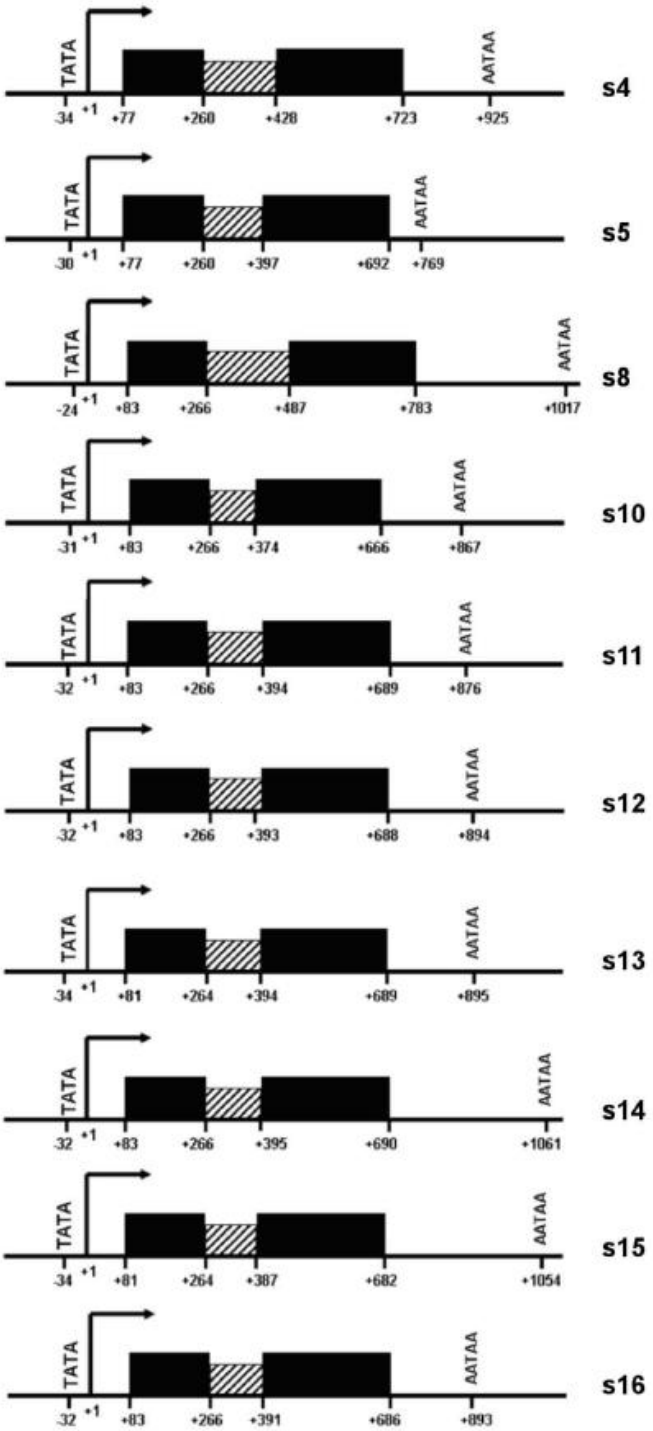


Figure 4 Predicted structures for *PR10* related sequences of *V. vinifera*. Only sequences having canonical transcription signals are shown. The arrow indicates the transcription starting site (+1). Predicted exons are represented as black boxes and deduced introns as dashed boxes. Given positions respectively correspond to the start of the TATA-box, the transcription starting site, the first nucleotide of the CDS, the last nucleotide of exon 1, the first nucleotide of exon 2, the last nucleotide of the CDS and the first nucleotide of the poly-A signal. Predictions were performed using the GeneFinding program. *VvPR10.1*, *VvPR10.2* and *VvPR10.3* respectively correspond to s16, s10 and s12.

Lebel *et al. BMC Plant Biology* 2010, **10**:184
http://www.biomedcentral.com/1471-2229/10/184

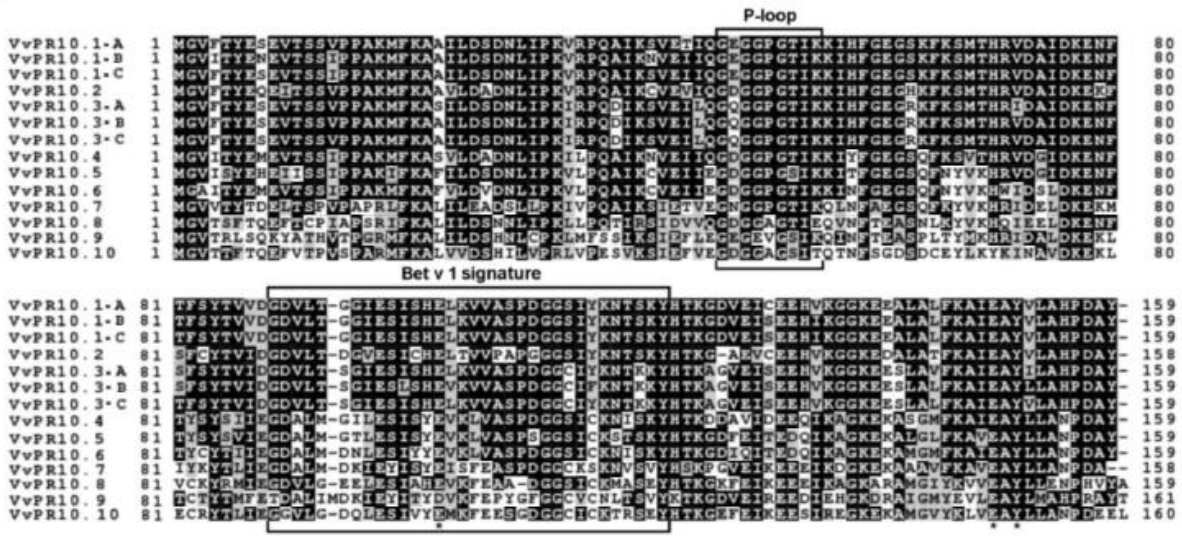


Figure 5 Sequence alignment of deduced PR10 proteins. The P-loop and Bet v 1 signature are framed. A star (*) marks the amino-acids implied in possible ribonucleasic activity (E102, E149 et Y151). Alignments were performed with ClustalW. A *V. vinifera* specific Bet v 1 motif was determined: G-[DG]-[VA]-L-x(4)-E-[SY]-[IL]-[CSATV]-[HY]-[ED]-x-[KST]-x-[VE]-x(3)-[GNDS]-G(2)-[CS]-x(2)-K-x(2)-[SK]-X-Y. In PR10.8 and PR10.10, the last aa varies in the P-loop motif (E and T, respectively); in PR10.9, E_{102} is replaced by D_{102}, the 4th aa in the P-loop motif is E, and the Bet v 1 motif presents 4 differences (out of 34 aa) at the positions 1, 6, 23 and 29; in PR10.6, there is one difference (out of 33 aa) in the Bet v1 motif (position 12).

-*c*, and *VvPR10.3-a*, -*b* and -*c*, respectively. The other sequences were named *VvPR10.4* to *VvPR10.10* for s8, s5, s4, s3, s1, s2 and s17, respectively, following increasing divergence from the group of *VvPR10.1-a* to *VvPR10.3-c* sequences. Putative protein sequences were deduced from nucleotide sequences, analysed and compared. All proteins are 158-160 aa long (Figure 5), have a calculated molecular mass ranging from 17.1 to 18.4 kDa and a theoretical pI from 4.7 to 6.3, consistent with usual plant PR10 features [9]. Apart from few exceptions mentioned in the figure legend, all sequences present the three characteristic amino acids thought to be implied in ribonucleasic activity: E_{102}, E_{149} and Y_{151}, the characteristic P-loop motif (aa 47-55) described in various species and corresponding to the consensus sequence GXGGXGXXK, and the Bet v 1 motif which is the signature for the Pathogenesis-related proteins Bet v 1 family (Prosite accession PS00451)[3,9]. The predictive three-dimensional structure of the deduced proteins was determined. As for PR10.1 (Figure 6), all proteins would be composed of three alpha-helices and a seven-stranded antiparallel beta-sheet arranged to form a large internal hydrophobic cavity. Nine short loops connect the secondary elements. This conformation is in accordance with the crystal structure determined for plant PR10 proteins in cherry, yellow lupine, birch and celery [12-15]. All these observations suggest correct structure and folding for the eleven putative *V. vinifera* PR10 proteins in addition to PR10.1, PR10.2 and PR10.3 [additional file 4].

A phylogenetic tree shows the relationships between *V. vinifera* PR10 proteins and representatives of the main PR10 subfamilies in other plants: Mal d 1 and Pru p 1 subfamilies I-IV in *Malus domestica* [7] and *Prunus*

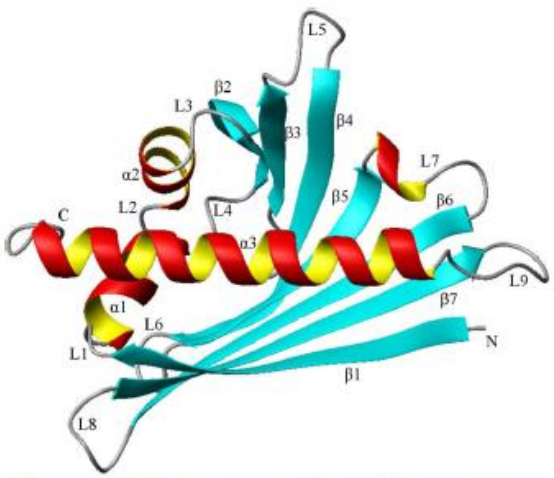


Figure 6 Three-dimensional structure of *V. vinifera* PR10.1 represented by a ribbon diagram. The structure was predicted on an automated comparative protein modeling server using SWISS-MODEL.

Lebel *et al. BMC Plant Biology* 2010, **10**:184
http://www.biomedcentral.com/1471-2229/10/184

persica [8], Bet v 1 subfamilies I-III and V in *Betula pendula* [26] and LlPR10 subclasses I and II in *Lupinus luteus* [11] (Figure 7). In each species, PR10 proteins tend to form one (*Betula pendula, Lupinus luteus*) or two clades (*Malus domestica, Prunus persica* and *Vitis vinifera*), showing a low intraspecific and a higher interspecific variation, as previously described [4-6]. In a first clade, *V. vinifera* PR10 proteins are subdivided into two distinct homogeneous groups: PR10.1-A to PR10.3-C with more than 79% of amino acid identity and PR10.4 to PR10.6 with more than 81% of amino acid identity. These two groups of tightly phylogenetically connected proteins are closer to birch Bet v 1 than to PR10 proteins of other plant species, and are undoubtedly distant from the four other *V. vinifera* PR10 proteins, PR10.7 to PR10-10.

Expression of *V. vinifera* candidate *PR10* genes

The possible expression of ten out of the fourteen complete *PR10* sequences was assessed in various non-treated tissues of *V. vinifera* cv. Chardonnay, as well as during secondary somatic embryogenesis. It was not possible to design specific primers for the detection of *VvPR10.1-b* and *-c* and *VvPR10.3-b* and *-c* transcripts, because of very high similarity with *VvPR10.1-a* and *VvPR10.3-a* transcripts respectively, even in the 5′ and 3′ UTRs.

Nine different genes were expressed in the tissues analysed, *VvPR10.10* transcripts being undetectable in any condition (Figure 8). Moreover, *VvPR10.4* and *VvPR10.9* transcripts were only weakly detected in a few tissues.

In intact tissues, all nine *PR10* genes were expressed, depending on the plant organ. In roots, the expression

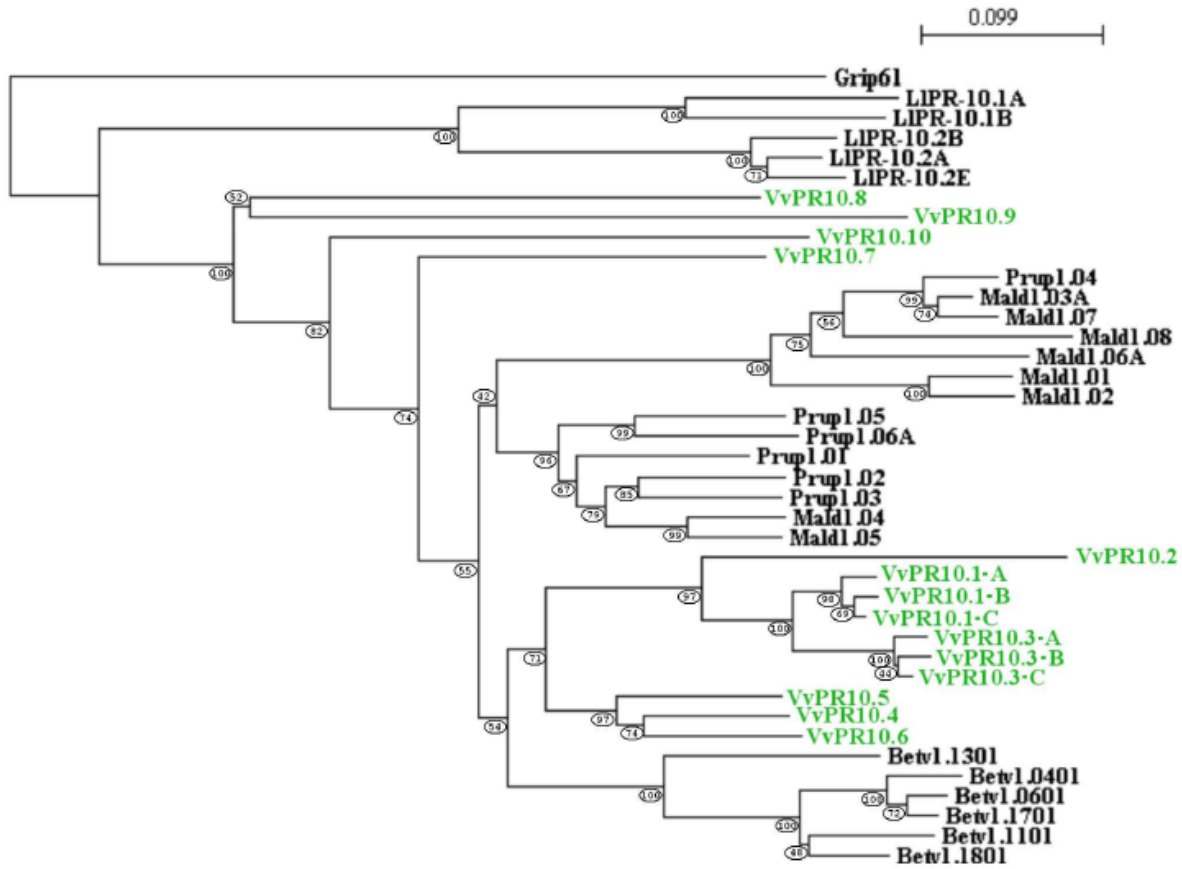


Figure 7 Phylogenetic relationships between *V. vinifera* PR10 proteins and representative PR10 proteins from *Betula pendula, Lupinus luteus, Malus domestica* and *Prunus persica*. GenBank accession numbers are as follows: *Betula pendula* Betv1.0401 (CAA54482), Betv1.0601 (CAA54484), Betv1.1101 (CAA54694), Betv1.1301 (CAA54696), Betv1.1701 (CAA96539) and Betv1.1801 (CAA96540); *Lupinus luteus* LlPR10.1A (CAA56298), LlPR10.1B (CAA56299), LlPR10.2A (AAF77633), LlPR10.2B (AAF77634) and LlPR10.2E (AAP37978); *Malus domestica* Mald1.01 (AAX18288), Mald1.02 (AAX18291), Mald1.03A (AAX18313), Mald1.04 (AAX18294), Mald1.05 (AAX18296), Mald1.06A (AAX18299), Mald1.07 (AAX18307) and Mald1.08 (AAX18310); *Prunus persica* Prup1.01 (ACE80940), Prup1.02 (ACE80942), Prup1.03 (ACE80944), Prup1.04 (ACE80946), Prup1.05 (ACE80948) and Prup1.06A (ACE80952). The *V. vinifera* Grip61 gene (CAB85634) was included as an outgroup representative of another Bet v 1 subfamily [3]. The NJ-tree was generated with the Phylo_win program. The bootstrap value is given for each node.

Lebel et al. BMC Plant Biology 2010, 10:184
http://www.biomedcentral.com/1471-2229/10/184

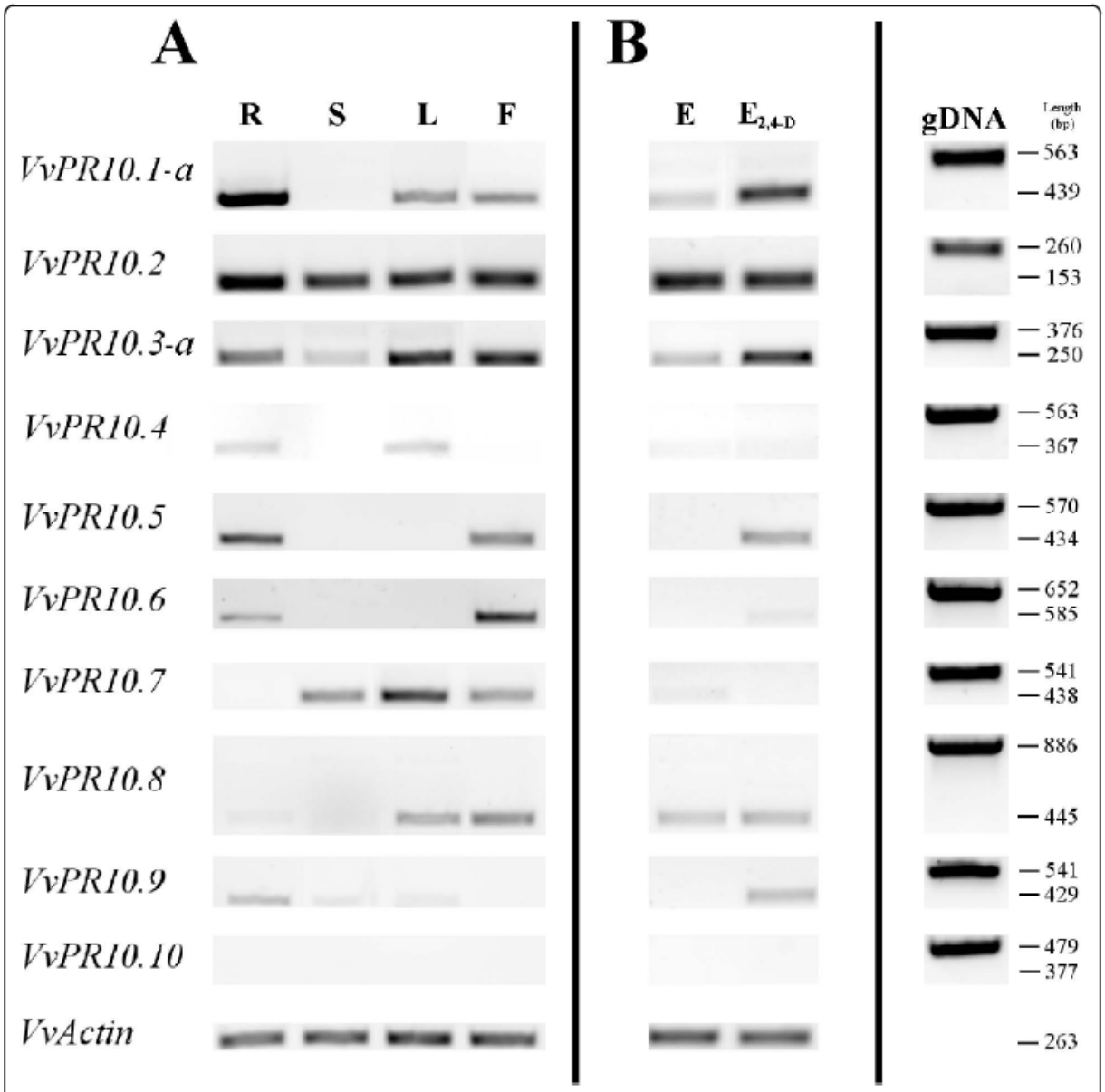


Figure 8 Expression of *V. vinifera PR10* genes. A: tissue-specific expression; R = roots, S = stems, L = leaves, F = flowers. B: expression during secondary somatic embryogenesis induction; E = non-treated somatic embryos at the cotyledonary stage, $E_{2,4\text{-}D}$ = calli obtained from embryos treated with 2,4-D. gDNA = genomic DNA. The length of amplified sequences are given in bp.

of all genes except *VvPR10.7* was detected, although *VvPR10.8* transcription was very weak. At the opposite, *VvPR10.7* transcription was high in leaves, while *VvPR10.5* and *VvPR10.6* transcripts were not detected; all other genes were expressed at varied levels. In stems, *VvPR10.2*, *VvPR10.3-a* and *VvPR10.7* were the only clearly expressed genes. Immature flowers expressed all genes except *VvPR10.4* and *VvPR10.9*. In non-treated somatic embryos, *VvPR10.1-a*, *VvPR10.3-a* and *VvPR10.8* transcripts were weakly detected, while *VvPR10.2* transcription was clear.

The 2,4-D treatment of embryos used for induction of secondary somatic embryogenesis increased the expression of *VvPR10.1-a*, *VvPR10.3-a*, *VvPR10.5* and *VvPR10.9* but not *VvPR10.2* and *VvPR10.8*. The expression of *VvPR10.6* was also weakly stimulated. On the other hand, *VvPR10.7* transcription was weakly detectable in intact embryos but not at all in tissues derived

Lebel et al. BMC Plant Biology 2010, 10:184
http://www.biomedcentral.com/1471-2229/10/184

from 2,4-D-treated embryos. *VvPR10.4* expression was not detected in non-treated or treated embryos.

Discussion

We found a total of seventeen *PR10* related sequences in the whole *V. vinifera* genome. Thirteen unique sequences were retained from an automatic search that initially produced ninety regions, reflecting redundancy of the database as well as annotation errors partly due to wrong homology detection. A manual search allowed the recovery of four additional *PR10* related sequences. All seventeen sequences were found in a single compact cluster on the chromosome 5. Plant *PR10* belong to multigene families. There are at least five *PR10* genes in pea [27], eighteen *Mal d 1* genes in apple [7], ten *Bet v 1* genes in birch [6], eight *Fra a 1* genes in strawberry [28], six *PR10* genes in *Solanum surattense* [29], eight in yellow lupine [11], five in rice [30], and eight *Pru p 1* and *Pru d 1* genes in peach and almond, respectively [8]. They were shown to form physical clusters in apple [7] and peach [8]. Poplar *PR10* genes are also supposed to be grouped on chromosomes [26]. Tandem duplicates are frequent in plant genomes and represent up to 16% of *Arabidopsis* genes [31]. Such gene clusters are thought to be produced by successive single gene or small-scale duplications. We found that thirteen out of the seventeen *V. vinifera PR10* sequences are present on the chromosome in direct orientation, suggesting that most copies were produced by unequal crossing over events, as described in *Arabidopsis* and rice [31].

Following duplication, new copies of a gene may undergo modifications allowing functional diversification, which is a significant source of evolutionary novelty in plants [32]. However, gene duplication mostly creates copies that are rapidly lost through pseudogenisation. As a result, from numerous homologous sequences coexisting in a genome, only a part are functional genes. From the seventeen *V. vinifera PR10* sequences, only fourteen have an intact ORF. In birch, the copy number of *PR10* genes varies from twelve to twenty-five, depending on the species, and pseudogenes represents as much as one-third of *Betula nigra* alleles [6]. In *V. vinifera*, the pseudogenes s7 and s9 share the highest nucleotide similarity with *VvPR10.6* (s4), suggesting that they could derive from its duplication. Likewise, s6 is closer to *VvPR10.5* (s5) than to all other sequences and could therefore originate from its duplication. Moreover, s6 and *VvPR10.5* are also very similar at the intron level (81% of nucleotide similarity), indicative of possible recent duplication or slow pseudogenisation.

Apart from pseudogenes, sequence divergence is generally reduced in *PR10* multigene families. *PR10* genes within a species are generally very similar and more distant from gene copies of other plant species [4]. Such a low intraspecific variability has been observed in *Passiflora* [5] and in the *Betula* genus [6]. The different paralogs are thought to undergo strong concerted evolution which is the tendency of a family of repeated genes to jointly evolve. The close physical proximity of tandem duplicates facilitates gene conversion or unequal crossing-over events leading to concerted evolution [6]. The phylogenetic tree in Figure 7 illustrates probable concerted evolution in the *V. vinifera PR10* family. However, *V. vinifera* PR10 proteins divide into two clades interrupted by PR10 of other plants, indicating partial independent evolution. Interestingly, the seven sequences s10 to s16 corresponding to the very homologous proteins PR10.1-A to PR10.3-C follow one another on the chromosome, what could favor concerted evolution and suggest strong local selection pressure. However, it cannot be ruled out that recent duplication events produced these seven very homologous sequences.

Grapes, as other eudicot species, probably originate from an ancient hexaploid ancestor formed through whole genome duplication (WGD) events [24,33]. Recurrent polyploidisation is tightly linked to evolution in angiosperms, providing raw materials for gene diversification and genome refinement, and coinciding with species radiation [32]. A WGD is followed by incomplete and asymmetric loss of gene copies and chromosome rearrangements allowing the recovery of a diploid-like state compatible with effective reproduction. In *Arabidopsis*, it was possible to track gene pairs released by a recent specific WGD event [34]. Only 28.6% of gene pairs were retained from the transient tetraploid genome, the other pairs having lost a copy from one of the homeologous chromosomes. In *V. vinifera*, *PR10* sequences were solely located on the chromosome 5, suggesting loss of some of the triplicate ancestral copies and/or translocation on a unique chromosome. The seventeen present *PR10* sequences are most probably due to repetitive small duplications having continuously occurred during *V. vinifera* evolution. Subsequent mutations and possible positive selection are responsible for the observed divergence within the sequences. However, at least a part of *V. vinifera PR10* sequences are probably subjected to concerted evolution, reducing variability and hampering the identification of putative triplicate ancestral copies.

We were able to detect the expression of nine out of the fourteen complete *PR10* related sequences in varied non-treated and treated tissues of *V. vinifera* cv. Chardonnay. Three genes, *i.e. VvPR10.1-a*, *VvPR10.2* and *VvPR10.3-a*, were already shown to be expressed under pathogen attack or abiotic stress, and during somatic embryogenesis induction [18,19], whereas expression of

Lebel et al. BMC Plant Biology 2010, 10:184
http://www.biomedcentral.com/1471-2229/10/184

VvPR10.4-*VvPR10.9* was never studied before. *VvPR10.10* transcripts were not detectable in our conditions. This sequence has no corresponding EST in the databases and lacks canonical transcription signals. However, expression could be limited to very specific tissues and/or developmental stages. Although specific ESTs in the databases suggest a probable expression of the four sequences *VvPR10.1-b* and *-c* and *VvPR10.3-b* and -c, the distinction of each transcription product would require more sensitive methods than PCR, such as cDNA-AFLP or SSCP, which were used for discriminate highly homologous sequences in potato, polyploid cotton and barley [35-37]. The expression of *VvPR10.5*, *VvPR10.7* and *VvPR10.8* is not surprising because well matching ESTs were found in the databases. On the contrary, no EST was found to correspond to *VvPR10.4*, *VvPR10.6* and *VvPR10.9*, whose expression is therefore reported for the first time. Interestingly, *VvPR10.7*, *VvPR10.8* and *VvPR10.9* are expressed although devoid of canonical transcription signals.

We found transcripts of all nine *PR10* genes in varying amounts, in the different non-treated tissues analysed, suggesting a role apart from defence. Specific expression profiles were found in intact tissues, except for *VvPR10.2* and *VvPR10.3-a*. Stems and intact embryos expressed a reduced subset of *PR10* proteins. On the contrary, immature flowers expressed a large subset of *PR10* genes, suggesting a possible role of some PR10 proteins during sexual reproduction. No *V. vinifera PR10* gene was solely expressed in calli derived from 2,4-D-treated embryos. However, expression of several genes was enhanced in these tissues, at least weakly. Somatic embryogenesis is generally obtained from tissues subjected to the auxin 2,4-D, which acts as a stress factor able to trigger the reprogramming of plant somatic cells towards embryogenesis [38]. As a consequence, high amounts of defence proteins are produced following a 2,4-D treatment, as shown in grapevine cultures [39]. Consistent with the results reported here, we previously showed that varied *PR* genes including *VvPR10.1-a* and *VvPR10.3-a* are up-regulated during secondary somatic embryogenesis induction in *V. vinifera* [19]. *VvPR10.1* expression was also previously reported to be induced in whole plant leaves challenged with the incompatible bacterium *Pseudomonas syringae* [18]. Interestingly, *VvPR10.4* and *VvPR10.7* were not responsive to 2,4-D treatment.

On the whole, various expression patterns were found for *V. vinifera PR10* genes, indicating functional diversification and possible tissue specificity. Likewise, in other plants, PR10 proteins show expression diversity suggesting various biological activities [9]. A major feature of PR10 proteins is their internal hydrophobic cavity with openings at the protein surface. Several three-dimensional modeling studies showed that this structure is suitable for the binding and transport of diverse hydrophobic ligands as brassinolides [14] or homocastasterone [12]. Some birch and yellow lupine PR10 can bind diverse molecules such as cytokinins, brassinosteroids, fatty acids and flavonoids, suggesting that they could interfere with the trafficking of hormones inside the cell [10,16,17]. Moreover, overexpression of a *PR10* gene in pea led to a change in the ratio between cytokinins and abscisic acid, showing that the PR10 content could be relevant for intracellular hormone regulation [40]. Although all crystallographic models for PR10 proteins share the same canonical fold, their superimposition can reveal structural differences [12]. Remarkably, the volume of the internal cavity and of its openings can show some variability, what could influence the type of transported ligand. In yellow lupine, different shapes for the inducible LlPR10.1A and the constitutive LlPR10.1B could account for their different biological roles [13].

Conclusion

The availability of the complete genome sequence of *V. vinifera* allowed us to characterise the *PR10* gene family. Seventeen different *PR10* related sequences including three pseudogenes were identified and located in a single compact cluster on the chromosome 5, most probably reflecting repetitive small duplications during evolution. A phylogenetic analysis showed a characteristic low variability within the different sequences, especially within seven sequences closely located on the chromosome, suggesting probable concerted evolution. We could analyse the expression of nine genes in various tissues. Different expression patterns indicate functional specialisation. Several genes showed a typical stress induced up-regulation. Further experiments will help to elucidate the differential regulation of *V. vinifera PR10* gene expression.

Methods

In silico identification of *PR10* related sequences in the *V. vinifera* genome and characterization of candidate genes

The grapevine Genoscope database was used to identify any sequence related to *PR10* genes [41]. The Genoscope Blat tool [42] and the ClustalW alignment tool of the European Bioinformatics Institute (EBI) [43] were used to retrieve and compare the sequences. An automatic list produced twenty-three unique (among ninety) putative *PR10* or *PR10-like* gene annotations. Seven annotations with no consistent homology with *PR10* genes were removed. Six other annotations covered three complete sequences, from which each exon was annotated as an independent gene, suggesting that the intron splicing sites were not detected. The last ten annotations comprised seven sequences with a complete ORF, and three with a prematurely interrupted ORF. All

Lebel et al. BMC Plant Biology 2010, 10:184
http://www.biomedcentral.com/1471-2229/10/184

thirteen *PR10* related sequences were found to form a cluster on the chromosome 5, between the positions 604,886 and 675,550. Thoroughly examining this chromosome portion as well as its 5′ and 3′ extensions, we found four additional annotations with nucleotide similarity (48-71% in the ORF) to the thirteen previous sequences. The *PR10* cluster was found to be limited by two putative genes not related to *PR10* sequences GSVIVT00033064001 and GSVIVT00033091001.

All corresponding genome sequences originating from the Istituto Agrario San Michele all'Adige (IASMA) grapevine database were retrieved using the NCBI Blast tool [44]. They were compared to the Genoscope sequences using the ClustalW alignment tool.

The exon-intron composition of *PR10* sequences was automatically determined by the Gaze program on the Genoscope website. When necessary, the position of the intron was corrected relatively to the structure of *VvPR10.1* [18], and according to the 5′ and 3′ splicing consensus sequences NN/GT and AG/NN, respectively.

Vitis sp. ESTs were searched, using the blastn program of the NCBI database [44].

Gene structures were predicted using the Gene Finding program on the University of London bioinformatics web server [45]. The organism parameter was set on "dicots". The FGENESH method showed the position of the TATA-box, the first and last nucleotides of the exons and the position of the polyadenylation signal. The TSSP-TCM method localized the transcription initiation site (+1).

Protein sequence analysis

The putative CDS were translated using the Transeq program [46]. The PM and pI were determined using the COMPUTE program [47]. The PROSITE database [48] was used to find the Pathogenesis-related proteins Bet v 1 family signature. Three-dimensional structures were predicted on an automated comparative protein modeling server using SWISS-MODEL [49].

Phylogenetic analysis and multiple alignments

A phylogenetic tree was built using the Neighbor-Joining method, with the Phylo_win program [50]. Bootstrap values were obtained from 500 replicates. The ClustalW alignment tool was used to compare protein sequences.

Plant material

All expression studies were performed on *V. vinifera* plants or cultures of the cultivar Chardonnay.

We separately collected roots, stems (with nodes) and whole leaves from well developed plantlets (10-12 leaves) obtained from *in vitro* microcutting, as previously described [51].

Immature flowers were harvested from French vineyards, at the stage of separate flower buds.

Recurrent *in vitro* embryogenesis was induced from nodal explants of the *V. vinifera* cv. Chardonnay, as previously described [19]. Young cotyledonary somatic embryos were isolated and transferred onto a medium containing the auxin analog 2,4-D for induction of indirect secondary embryogenesis. After three weeks, an undifferentiated callus was obtained from each embryo.

Tissues were collected, frozen in liquid nitrogen and further conserved at -80°C until RNA extraction. Two independent samplings were performed.

RNA extraction and analysis

Total RNA extractions and semi-quantitative RT-PCR experiments were performed as reported [19]. Primer pairs were designed to overlap the intron, in order to distinguish possible genome DNA contamination from gene expression products (Table 2). All primer pairs were shown to efficiently detect the respective corresponding genomic sequences (gDNA). Semi-quantitative RT-PCR tests were performed in three replicates. PCR products were sequenced to confirm their identity (Genoscreen, France).

Table 2 Sequences of primers used for semi-quantitative RT-PCR

Gene	GenBank	Forward primer 5′→3′	Reverse primer 5′→ 3′
Actin	AF369525	TGCTATCCTTCGTCTTGACCTTG	GGACTTCTGGACAACGGAATCTC
VvPR10.1	AJ291705	GAAATCATACAAGGAGAGGGAGGC	GCCAAACTTATTGAGACTGATAGGTG
VvPR10.2	AJ291704	CGATCACAGTGTAGCGGAATGAGAAT	AAGCTATCAAGTGCGTGGAAGTCATT
VvPR10.3	EU379313	GAAATCCTACAAGGACAGGGAGGT	CGGCCTTGGTGTGGTACTTTT
VvPR10.4	-	ATCCTTCCCCAAGCTATCAAG	GATTTGCCAAGAGGTAAGCC
VvPR10.5	-	ATCCTTGACTCTGATAACCTCA	ATGATATGAGACAAAGGAGTTTC
VvPR10.6	-	GTCCTTGATGTTGATAACCTC	AAGCCAAGCCTTTTAACTG
VvPR10.7	-	ATCGTCCCTCAGGCCATTA	AAGTGATTAAGTGGAGGAGAAGC
VvPR10.8	-	CTCTTGCCCCAGACCATAAG	ACATTGGACAACAGAGAAGTGAC
VvPR10.9	-	CAGTCAAAAGTACGCGACTCA	AAGTATAGGCGCGAGGGTGT
VvPR10.10	-	AATCAGTAAAGAGCATCGAGTT	AAAGTAATCACAACTCCTCGTC

Lebel *et al. BMC Plant Biology* 2010, **10**:184
http://www.biomedcentral.com/1471-2229/10/184

Additional material

Additional file 1: Genome position and annotation reference for *PR10* related sequences, as given on the Genoscope website. (*) a double annotation means that the sequence has been annotated as two individual genes.

Additional file 2: Concordance of the map of the *PR10* cluster obtained from the Genoscope and homologous scaffolds originating from the IASMA genome database. Each coloured band corresponds to a different scaffold, given with GenBank accession number. IASMA sequences were obtained from NCBI.

Additional file 3: Percentage of nucleotide similarity between the introns of the seventeen sequences. High percentages of nucleotide similarity are highlighted in red. *VvPR10.1*, *VvPR10.2* and *VvPR10.3* respectively correspond to s16, s10 and s12. The values were obtained from sequence alignments on ClustalW.

Additional file 4: Three-dimensional structure of deduced *V. vinifera* PR10 proteins represented by a ribbon diagram. The structure was predicted on an automated comparative protein modeling server using SWISS-MODEL. With reference to PR10.1, PR10.8 and PR10.9 have a longer C-terminal end, while PR10.7 and PR10.10 have a shorter C-terminal end. The folding of the regions between α2 and β4 diverges from the model in PR10.5, PR10.6 and PR10.7.

Acknowledgements
This research was supported by the "Université de Haute Alsace" and by a doctoral fellowship from the "Ministère de l'Enseignement supérieur et de la Recherche" to SL.

Authors' contributions
SL performed all *in silico* analyses as well as RT-PCR experiments. PS participated in studies performed *in silico*. PM carried out the tissue cultures. PS, BW and PM designed and coordinated the study. SL and PM wrote the manuscript. All authors read and approved the manuscript.

Received: 29 January 2010 Accepted: 20 August 2010
Published: 20 August 2010

References

1. van Loon LC, Rep M, Pieterse CMJ: **Significance of inducible defense-related proteins in infected plants.** *Annu Rev Phytopathol* 2006, **44**:135-162.
2. Christensen AB, Cho BH, Naesby M, Gregersen PL, Brandt J, Madriz-Ordeñana K, Collinge DB, Thordal-Christensen H: **The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins.** *Mol Plant Pathol* 2002, **3(3)**:135-144.
3. Radauer C, Lackner P, Breiteneder H: **The Bet v I fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands.** *BMC Evol Biol* 2008, **8**:286.
4. Wen J, Vanek-Krebitz M, Hoffmann-Sommergruber K, Scheiner O, Breiteneder H: **The potential of Betv1 homologues, a nuclear multigene family as phylogenetic markers in flowering plants.** *Mol Phylogenet Evol* 1997, **8(3)**:317-333.
5. Finkler C, Glacomet C, Muschner VC, Salzano FM, Freitas LB: **Molecular investigations of pathogenesis-related Bet v 1 homologues in *Passiflora* (Passifloraceae).** *Genetica* 2005, **124**:117-125.
6. Schenk MF, Cordewener JHG, America AHP, van't Westende WPC, Smulders MJM, Gilissen LJW.I: **Characterization of PR-10 genes from eight Betula species and detection of Bet v I isoforms in birch pollen.** *BMC Plant Biol* 2009, **9**:24.
7. Gao ZS, van de Weg WE, Schaart JG, Schouten HJ, Tran DH, Kodde LP, van de Meer IM, van der Geest AHM, Kodde J, Breiteneder H, Hoffmann-Sommergruber K, Bosch D, Gilissen LJWJ: **Genomic cloning and linkage mapping of the *Mal d 1* (*PR-10*) gene family in apple (*Malus domestica*).** *Theor Appl Genet* 2005, **111**:171-183.
8. Chen L, Zhang S, Illa E, Song L, Wu S, Howard W, Arús P, van de Weg E, Chen K, Gao Z: **Genomic characterization of putative allergen genes in peach/almond and their synteny with apple.** *BMC Genom* 2008, **9**:543.
9. Liu JJ, Ekramoddoullah AKM: **The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses.** *Physiol Mol Plant Pathol* 2006, **68**:3-13.
10. Koistinen KM, Soininen P, Vanäläinen TA, Häyrinen J, Laatikainen R, Peräkylä M, Tervahauta AI, Kärenlampi SO: **Birch PR-10c interacts with several biologically important ligands.** *Phytochem* 2005, **66**:2524-2533.
11. Handschuh L, Femiak I, Kasperska A, Figlerowicz M, Sikorski MM: **Structural and functional characteristics of two novel members of pathogenesis-related multigene family of class 10 from yellow lupine.** *Acta Bioch Pol* 2007, **54(4)**:783-796.
12. Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H, Rösch P: **Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1.** *J Biol Chem* 2001, **276(25)**:22756-22763.
13. Biesiadka J, Bujacz G, Sikorski MM, Jaskolski M: **Crystal structures of two homologous pathogenesis-related proteins from yellow lupine.** *J Mol Biol* 2002, **319**:1223-1234.
14. Marković-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O, Breiteneder H: **Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier.** *J Mol Biol* 2003, **325**:123-133.
15. Schirmer T, Hoffmann-Sommergruber K, Susani M, Breiteneder M, Marković-Housley Z: **Crystal structure of the major celery allergen Api g 1: Molecular analysis of cross-reactivity.** *J Mol Biol* 2005, **351**:1101-1109.
16. Mogensen JE, Wimmer R, Larsen JN, Spangfort MD, Otzen DE: **The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands.** *J Biol Chem* 2002, **277(26)**:23684-23692.
17. Fernandes H, Pasternak O, Bujacz G, Bujacz A, Sikorski MM, Jaskolski M: ***Lupinus luteus* pathogenesis-related protein as a reservoir for cytokinin.** *J Mol Biol* 2008, **378(5)**:1040-1051.
18. Robert N, Ferran J, Breda C, Coutos-Thévenot P, Boulay M, Buffard D, Esnault R: **Molecular characterization of the incompatible interaction of *Vitis vinifera* leaves with *Pseudomonas syringae* pv. *pisi*: Expression of genes coding for stilbene synthase and class 10 PR protein.** *Eur J Plant Pathol* 2001, **107**:249-261.
19. Maillot P, Lebel S, Schellenbaum P, Jacques A, Walter B: **Differential regulation of *SERK*, *LEC1-Like* and *Pathogenesis-Related* genes during indirect secondary somatic embryogenesis in grapevine.** *Plant Physiol Biochem* 2009, **47(8)**:743-752.
20. Jellouli N, Ben Jouira H, Skouri H, Ghorbel A, Gourgouri A, Mliki A: **Proteomic analysis of Tunisian grapevine cultivar Razegui under salt stress.** *J Plant Physiol* 2008, **165**:471-481.
21. Castro AJ, Carapito C, Zorn N, Magné C, Leize E, Van Dorsselaer A, Clément C: **Proteomic analysis of grapevine (*Vitis vinifera L.*) tissues subjected to herbicide stress.** *J Exp Bot* 2005, **56(421)**:2783-2795.
22. Fung RWM, Gonzalo M, Fekete C, Kovacs LG, He Y, Marsh E, McIntyre LM, Schachtman DP, Qiu W: **Powdery mildew induces defense-oriented reprogramming of the transcriptome in a susceptible but not in a resistant grapevine.** *Plant Physiol* 2008, **146**:236-249.
23. Polesani M, Desario F, Ferrarini A, Zamboni A, Pezzotti M, Kortekamp A, Polverari A: **cDNA-AFLP analysis of plant and pathogen genes expressed in grapevine infected with *Plasmopara viticola*.** *BMC Genom* 2008, **9**:142.
24. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyere C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pe ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quetier F, Wincker P: **The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla.** *Nature* 2007, **449**:463-467.
25. Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardel B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzersky R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Demattè L, Mraz A, Battilana J, Stormo K, Costa F,

Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando SM, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R: **A high quality draft consensus sequence of the genome of a heterozygous grapevine variety.** *PLoS ONE* 2007, **2(12)**:e1326.
26. Schenk MF, Gilissen LJWJ, Esselink GD, Smulders MJM: **Seven different genes encode a diverse mixture of isoforms of Bet v I, the major birch pollen allergen.** *BMC Genom* 2006, **7**:168.
27. Tewari S, Brown SM, Kenyon P, Balcerzak M, Fristensky B: **Plant defense multigene families: II evolution of coding sequence and differential expression of PR10 genes in *Pisum*.**, arXiv.org/q-bio.PE/0310038.
28. Musidlowska-Persson A, Alm R, Emanuelsson C: **Cloning and sequencing of the Bet v 1-homologous allergen Fra a 1 in strawberry (*Fragaria ananassa*) shows the presence of an intron and little variability in amino acid sequence.** *Mol Immunol* 2007, **44**:1245-1252.
29. Liu X, Huang B, Lin J, Fei J, Chen Z, Pang Y, Sun X, Tang K: **A novel pathogenesis-related protein (SsPR10) from *Solanum surattense* with ribonucleolytic and antimicrobial activity is stress- and pathogen-inducible.** *J Plant Physiol* 2006, **163**:546-556.
30. Kim ST, Yu S, Kang YH, Kim SG, Kim JY, Kim SH, Kang KY: **The rice pathogen-related protein 10 (JIOsPR10) is induced by abiotic and biotic stresses and exhibits ribonuclease activity.** *Plant Cell Rep* 2008, **27**:593-603.
31. Rizzon C, Ponger L, Gaut BS: **Striking similarities in the genomic distribution of tandemly arrayed genes in Arabidopsis and rice.** *PLoS Comput Biol* 2006, **2(9)**:e115.
32. Flagel LE, Wendel JF: **Gene duplication and evolutionary novelty in plants.** *New Phytol* 2009, **183**:557-564.
33. Tang H, Wang X, Bowers JE, Ming R, Alam M, Paterson AH: **Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps.** *Genome Res* 2008, **18**:1944-1954.
34. Thomas BC, Pedersen B, Freeling M: **Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes.** *Genome Res* 2006, **16**:934-946.
35. Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF: **Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development.** *Plant J* 1996, **9(5)**:745-753.
36. Cronn RC, Adams KL: **Quantitative analysis of transcript accumulation from genes duplicated by polyploidy using cDNA-SSCP.** *Biotechniques* 2003, **34(4)**:726-730.
37. Federico ML, Iñiguez-Luy FL, Skadsen RW, Kaeppler HF: **Spatial and temporal divergence of expression in duplicated barley germin-like protein-encoding genes.** *Genetics* 2006, **174**:179-190.
38. Fehér A, Pasternak TP, Dudits D: **Transition of somatic plant cells to an embryogenic state.** *Plant Cell Tissue Organ Cult* 2003, **74**:201-228.
39. Marsoni M, Bracale M, Espen L, Prinsi B, Negri AS, Vannini C: **Proteomic analysis of somatic embryogenesis in *Vitis vinifera*.** *Plant Cell Rep* 2008, **27**:347-356.
40. Srivastava S, Emery RJN, Kurepin LV, Reid DM, Fristensky B, Kav NNV: **Pea PR 10.1 is a ribonuclease and its transgenic expression elevates cytokinin levels.** *Plant Growth Regul* 2006, **49**:17-25.
41. **The Genoscope grape genome browser.** [http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/].
42. **The Genoscope web Blat tool.** [http://www.genoscope.cns.fr/blat-server/cgi-bin/vitis/webBlat].
43. **The European Bioinformatics Institute Clustal W online tool.** [http://www.ebi.ac.uk/Tools/clustalw2/].
44. **The National Center for Biotechnology Information Blast tool.** [http://blast.ncbi.nlm.nih.gov/Blast.cgi].
45. **The University of London bioinformatics web server.** [http://mendel.cs.rhul.ac.uk/mendel.php].
46. **The European Bioinformatics Institute Transeq program.** [http://www.ebi.ac.uk/Tools/emboss/transeq/].
47. **The ExPASy Compute pI/Mw tool.** [http://www.expasy.ch/tools/pi_tool.html].
48. **The ExPASy Prosite database.** [http://www.expasy.ch/prosite].
49. **The ExPASy Swiss-model server.** [http://swissmodel.expasy.org/].
50. Galtier N, Gouy M, Gautier C: **SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny.** *Comput Appl Biosci* 1996, **12**:543-548.
51. Maillot P, Kieffer F, Walter B: **Somatic embryogenesis from stem nodal sections of grapevine.** *Vitis* 2006, **45(4)**:185-189.

doi:10.1186/1471-2229-10-184
Cite this article as: Lebel *et al.*: Characterisation of the *Vitis vinifera* PR10 multigene family. *BMC Plant Biology* 2010 **10**:184.

Submit your next manuscript to BioMed Central and take full advantage of:

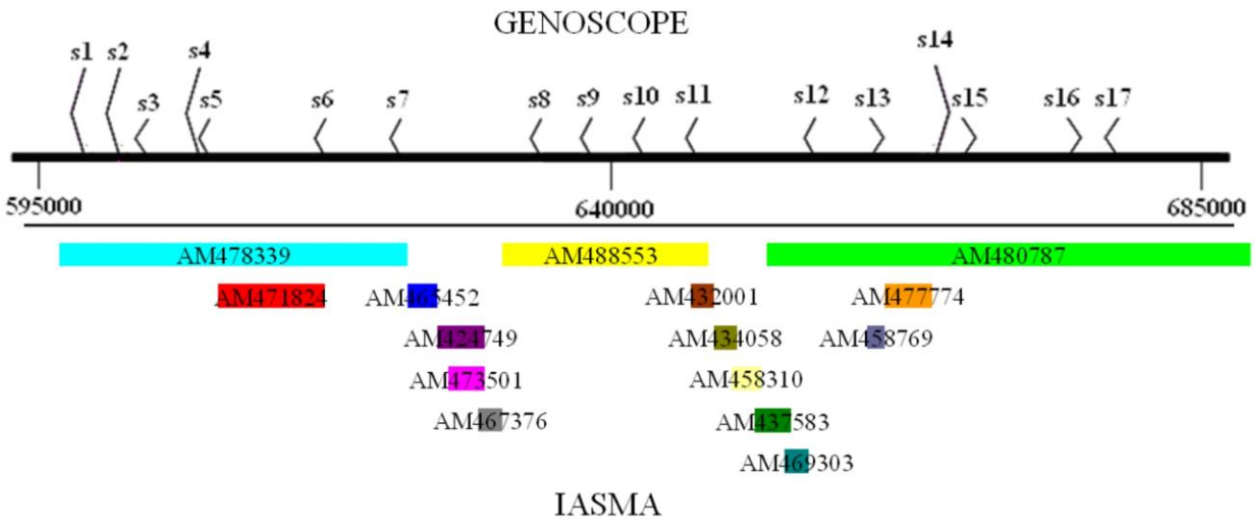
- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central

Sequence	Position	Annotation reference
s1	599,228-600,138	GSVIVT00033065001
s2	601,811-602,408	GSVIVT00033066001
s3	604,886-605,857	GSVIVT00033067001
s4	608,413-609,176	GSVIVT00033068001
s5	610,619-611,471	GSVIVT00033069001
s6	618,372-618,879	GSVIVT00033071001
s7	619,395-625,218	GSVIVT00033072001
s8	635,027-636,027	GSVIVT00033074001
s9	637,811-638,796	GSVIVT00033075001
s10 (*VvPR10.2*)	642,840-643,832	GSVIVT00033076001
s11	646,536-647,481	GSVIVT00033078001
s12 (*VvPR10.3*)	655,517-656,442	GSVIVT00033081001 - GSVIVT00033080001*
s13	659,794-660,761	GSVIVT00033082001 - GSVIVT00033083001*
s14	664,606-665,079	GSVIVT00033084001 - GSVIVT00033085001*
s15	667,370-668,313	GSVIVT00033087001
s16 (*VvPR10.1*)	674,541-675,500	GSVIVT00033089001
s17	678,498-679,082	GSVIVT00033090001

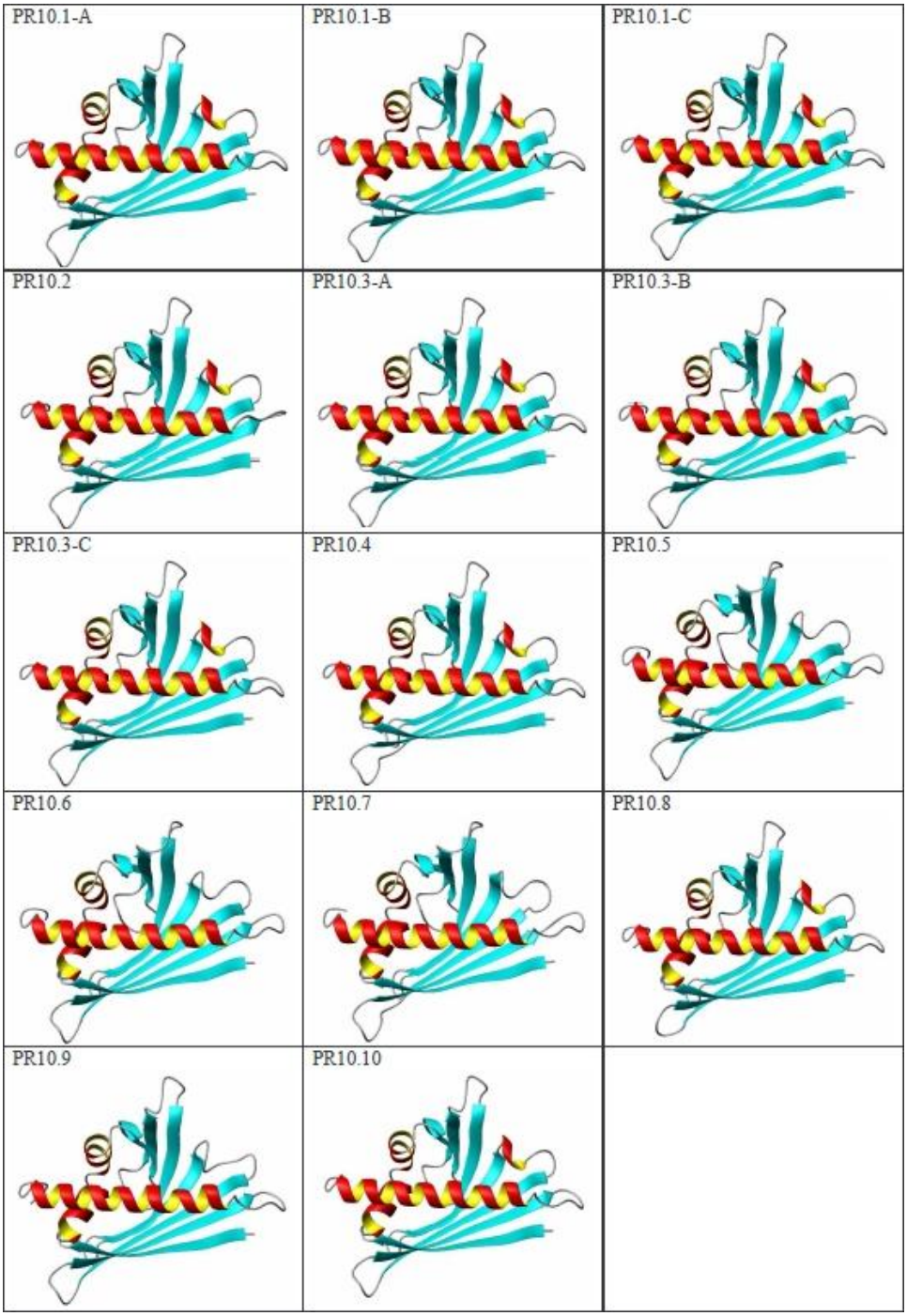
Fichier supplémentaire 1 : Position génomique et annotation apparaissant sur le site du GENOSCOPE pour chacune des séquences *PR10* potentielles.



Fichier supplémentaire 2 : Concordance entre la carte de la région génomique du GENOSCOPE contenant les séquences PR10 potentielles et les fragments homologues de la base de donnée de IASMA.

I	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15	s16	s17
s1		14	21	20	17	19	23	14	19	24	12	11	12	13	13	12	21
s2			24	15	16	8	17	27	9	23	17	33	8	8	17	17	11
s3				17	20	19	15	19	23	14	16	28	24	24	16	55	14
s4					36	30	13	14	21	23	15	26	19	18	35	15	32
s5						81	13	27	50	13	11	33	33	33	8	16	11
s6							15	22	30	13	11	25	28	27	8	10	8
s7								10	11	19	9	15	17	17	9	9	15
s8									15	16	32	19	30	30	33	33	10
s9										33	13	35	16	14	46	12	14
s10											90	85	85	85	86	86	20
s11												84	84	84	99	98	28
s12													95	95	77	82	11
s13														100	77	75	24
s14															78	83	25
s15																97	27
s16																	30
s17																	

Fichier supplémentaire 3 : Pourcentage de similarité nucléotidique entre les introns des 17 séquences.



Fichier supplémentaire 4 : Structure tridimentionnelle des protéines PR10 de *V. vinifera* déduites, représentée par un diagramme en ruban.

Chapitre II : Expression des gènes *PR10* pendant le processus d'embryogenèse somatique

A. Partie bibliographique

Embryogenèse	Directe	Indirecte
Primaire	Initiée à partir d'explants quelconques ; pas de dédifférentiation des cellules	Initiée à partir d'explants quelconques ; dédifférentiation des cellules (cals)
Secondaire	Initiée à partir d'embryons ; pas de dédifférentiation des cellules	Initiée à partir d'embryons ; dédifférentiation des cellules (cals)

Tableau 12 : Résumé des différents types d'embryogenèse somatique.

Première partie : L'embryogenèse somatique et ses marqueurs

I. L'embryogenèse somatique

L'embryogenèse somatique est le procédé par lequel des cellules somatiques différenciées recouvrent leur capacité embryogène. Cette reprogrammation cellulaire est connue sous le terme de transdifférenciation et elle est rendue possible par la totipotence des cellules végétales. L'embryogenèse somatique présente un double intérêt, à la fois sur le plan fondamental et sur le plan appliqué. En effet, les mécanismes mis en œuvre au cours de l'embryogenèse somatique sont très proches de ceux qui sont mis en œuvre au cours de l'embryogenèse zygotique, ce qui en fait un modèle pour l'étude de la régulation des gènes et l'observation morphogénique des étapes précoces du développement embryonnaire (Zimmerman 1993). Sur un plan appliqué, la régénération d'organismes végétaux via l'embryogenèse somatique est une solution efficace pour éviter les problèmes de chimérisme lors de manipulations génétiques (mutagenèse, transformation, fusions cellulaires), car les régénérants sont alors issus d'une unique ou de quelques cellules seulement (Gaj 2004).

Plusieurs processus d'embryogenèse somatique peuvent être distingués. Celle-ci pourra être qualifiée d'embryogenèse somatique primaire ou secondaire, directe ou indirecte. Le Tableau 12 résume ces différents types d'embryogenèse. L'embryogenèse somatique primaire est initiée à partir d'explants tels que des feuilles, racines, anthères, tiges ou cotylédons. Cette liste n'est pas exhaustive et l'efficacité de l'embryogenèse, à savoir le nombre d'explants générant des embryons néoformés, varie suivant les explants et d'une espèce à une autre. L'embryogenèse somatique secondaire est initiée à partir d'embryons et elle est souvent plus efficace que l'embryogenèse somatique primaire (Gaj 2004). L'embryogenèse somatique est qualifiée d'indirecte lorsque les cellules passent par un stade

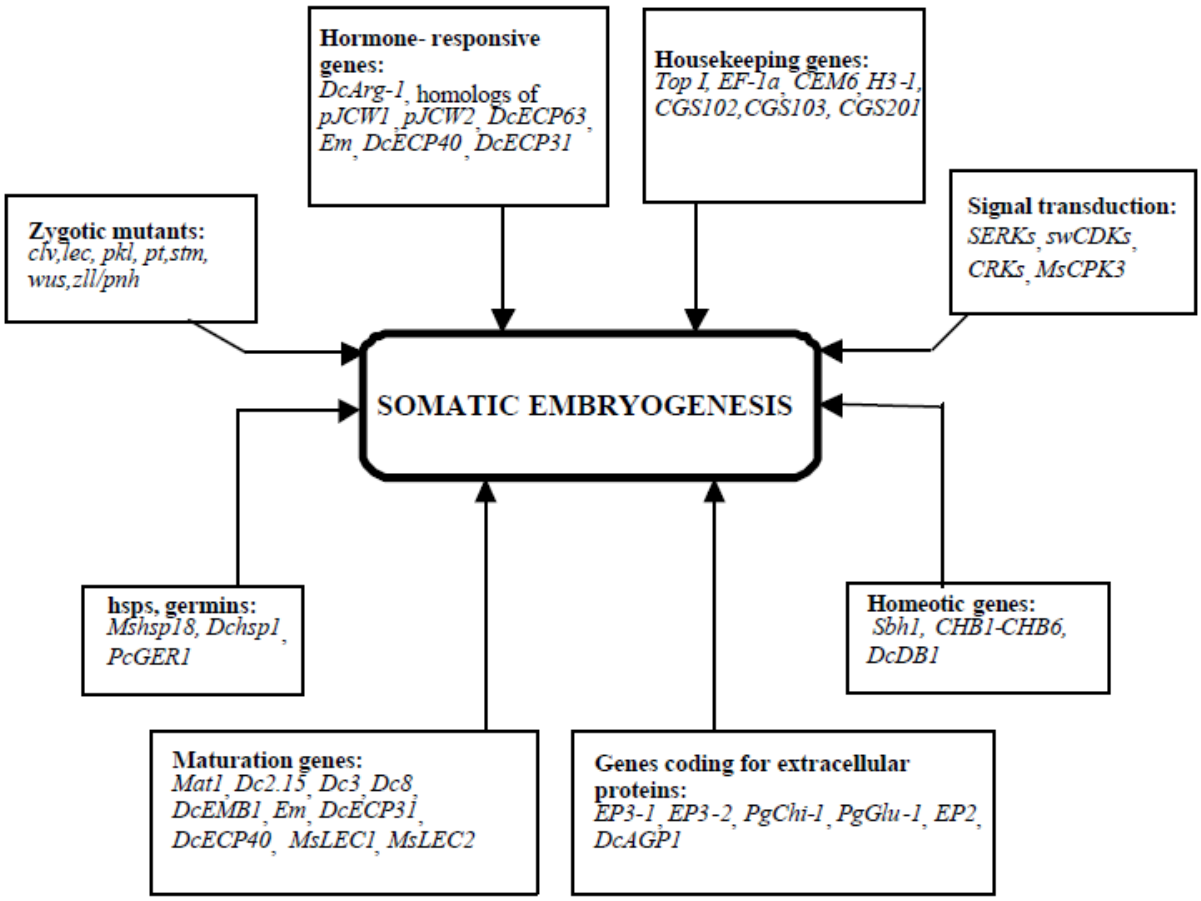


Figure 6 : Représentation schématique de plusieurs groupes de gènes influençant l'embryogenèse somatique chez les plantes supérieures (source : Chugh et Khurana 2002).

de dédifférenciation en cals tandis que les embryons issus d'une embryogenèse somatique directe sont formés directement sur le tissu, à partir de cellules pro-embryogéniques compétentes (Gaj 2004, Quiroz-Figueroa et al. 2006).

Plusieurs facteurs conditionnent l'efficacité de l'embryogenèse somatique (Fehér et al. 2003, Gaj 2004, Jiménez 2005, Quiroz-Figueroa et al. 2006). Outre le type de tissu mis en culture interviennent les sucres, source d'énergie et de carbone dont la concentration doit contribuer à une pression osmotique favorable, la lumière, dont les besoins varient selon l'espèce, les facteurs de stress et les régulateurs de croissance. Ces deux derniers facteurs sont considérés comme les plus importants dans l'acquisition de la capacité embryogène. Le stress subi peut être d'origine variable ; il peut s'agir d'une blessure, d'une exposition aux métaux lourds ou à des mutagènes, d'un choc osmotique, qui vont déclencher l'acquisition de la capacité embryogène. L'application d'hormones, essentiellement auxines, mais aussi cytokinines, constitue une condition importante pour l'obtention de cultures embryogènes. L'ajout de l'auxine 2,4-D au milieu de culture est considéré comme une étape-clé de l'induction de l'embryogenèse somatique, cette molécule agissant à la fois comme une substance à activité auxinique et comme un facteur de stress. Néanmoins, d'autres hormones peuvent être employées pour induire l'embryogenèse somatique, qui peut aussi être obtenue en l'absence de régulateur de croissance. Les connaissances actuelles permettent difficilement de proposer un mécanisme commun pour l'induction de l'embryogenèse somatique dans toutes les espèces végétales, chacune pouvant accepter des conditions de culture différentes (pour revue dans la vigne, Martinelli et Gribaudo 2009).

II. Marqueurs de l'embryogenèse somatique

L'induction de l'embryogenèse somatique implique une reprogrammation cellulaire, durant laquelle l'expression d'un grand nombre de gènes est régulée, dont des gènes impliqués dans la défense, probablement en réponse au stress appliqué (Fehér et al. 2003, Imin et al. 2005, Domoki et al. 2006, Karami et Saidi 2010 et Figure 6). Certains de ces gènes ont un profil d'expression spécifique durant l'embryogenèse somatique et peuvent servir de marqueurs pour identifier les cellules ou les tissus ayant une capacité embryogénique. J'ai choisi de présenter plus en détails les gènes *SERK*, *LEC*, ainsi que ceux qui codent des

Organisme	Nombre de gènes	Référence
Arabidopsis thaliana (arabette)	5	Hecht et al. 2001
Citrus unshiu (mandarinier)	1	Shimada et al. 2005
Cocos nucifera (cocotier)	1	Perez-Nunez et al. 2009
Medicago truncatula (luzerne tronquée)	1	Nolan et al. 2003
Oriza sativa (riz)	2	Hu et al. 2005, Song et al. 2008
Poa pratensis (pâturin des prés)	2	Albertini et al. 2005
Solanum tuberosum (pomme de terre)	1	Sharma et al. 2008
Theobroma cacao (cacaoyer)	1	Santos et al. 2005
Triticum aestivum (blé tendre)	3	Singla et al. 2008
Vitis vinifera (vigne)	3	Schellenbaum et al. 2008
Zea mays (maïs)	2	Baudino et al. 2001

Tableau 13 : Nombre de gènes *SERK* identifiés chez quelques espèces.

chitinases, dont j'ai étudié l'expression au cours du cycle d'embryogenèse somatique chez la vigne pendant ma thèse, en parallèle de l'étude de l'expression des gènes *PR10*.

Les gènes *SERK*, pour « *Somatic Embryogenesis Receptor Kinase* » (pour revue, Maillot et al. 2010), forment de petites familles multigéniques de 2 à 5 membres chez plusieurs espèces végétales (Tableau 13). Ils appartiennent à la famille des « Receptor Like Kinases » (RLK) et possèdent les domaines protéiques caractéristiques suivants : un domaine extracellulaire impliqué dans la fixation de ligands (perception), un domaine transmembranaire et un domaine catalytique cytoplasmique serine/thréonine kinase. Les gènes *SERK* ont été définis comme marqueurs de l'embryogenèse somatique, entre autres chez la carotte, chez *Dactylis glomerata* ou encore chez *Arabidopsis thaliana* (Schmidt et al. 1997, Somleva et al. 2000, Hecht et al. 2001). Cependant, cette petite famille de gènes possède des membres impliqués à la fois dans différents aspects du développement et dans la défense des plantes. Chez *A. thaliana* par exemple, *AtSERK1* et *AtSERK2* semblent avoir des fonctions redondantes dans l'embryogenèse mais aussi dans le développement des anthères (Albrecht et al. 2005), tandis que *AtSERK3/BAK1* et *AtSERK4/BKK1* ont un rôle dans la perception des brassinostéroïdes, tout comme *AtSERK1*, et sont également tous deux impliqués dans la défense (perception de la flagelline, mort cellulaire programmée) (He et al. 2007, Kemmerling et al. 2007). Chez *Oryza sativa*, l'expression du gène *OsSERK1* est induite par différents stress biotiques ou abiotiques, suggérant également un rôle dans la défense (Hu et al. 2005).

Les gènes *LEC*, pour « *LEafy Cotyledon* », codent des facteurs de transcription contenant un domaine de fixation à l'ADN (Harada et al. 2001). Les gènes *L1L*, pour « *Leafy Cotyledon 1 Like* », codent une sous-unité de la protéine HAP3 d'*A. thaliana* proche de LEC1 (Kwong et al. 2003). Ces gènes sont impliqués dans les phases précoce et tardive de l'embryogenèse somatique et contribuent à la mise en place d'un environnement favorable à l'initiation du développement embryonnaire (Kwong et al. 2003, Gaj et al. 2005). L'inactivation des gènes *LEC* chez *A. thaliana* induit une réduction de la capacité embryogénique des tissus, suggérant un rôle majeur dans l'embryogenèse somatique.

Les chitinases sont des enzymes capables d'hydrolyser la liaison β-1,4-glycoside présente dans les biopolymères de N-acétylglucosamine tels que la chitine. Cette molécule étant un constituant de la paroi des champignons, les chitinases sont supposées avoir une action

antifongique par l'intermédiaire de la libération de fragments de chitine ayant un rôle éliciteur des défenses de la plante. Plusieurs classes de protéines PR (PR3, PR4, PR8 et PR11) ont une activité chitinasique (Tableau 1). Des chitinases font partie des protéines secrétées spécifiquement durant l'embryogenèse somatique (Quiroz-Figueroa et al. 2006). L'ajout d'une chitinase dans le milieu de culture d'un mutant de carotte, dont le développement embryonnaire est stoppé au stade globulaire, permet la restauration de sa capacité de régénération, suggérant un rôle primordial de la fonction chitinasique pendant le développement embryonnaire (De Jong et al. 1992).

Lorsque ma thèse a débuté, les gènes *PR10* n'étaient pas considérés comme des marqueurs de l'embryogenèse somatique, bien que leur expression ait été mise en évidence lors de ce processus chez plusieurs espèces. Il avait notamment été montré que chez *Medicago sativa* L., l'expression d'un gène *PR10* est fortement induite pendant le processus d'embryogenèse somatique, tout d'abord lors du traitement initial de protoplastes au 2,4-D, puis, ultérieurement, lors de la différenciation des embryons (Domoki et al. 2006). De même, il avait été décrit que les PR10 font partie des protéines les plus abondamment exprimées durant l'embryogenèse somatique chez *Medicago truncatula* (Imin et al. 2005, De Jong et al. 2007).

III. Description du modèle d'étude : le cycle d'embryogenèse somatique secondaire chez *Vitis vinifera*

L'embryogenèse somatique secondaire permet d'obtenir un grand nombre d'embryons à partir de la culture d'un embryon primaire. Chez la vigne, il est entre autres utilisé pour maintenir des lignées embryogènes, ainsi que pour la production simultanée de quantités importantes d'embryons aptes à être manipulés ou analysés (Martinelli et Gribaudio 2009). L'induction du processus d'embryogenèse somatique secondaire peut être réalisée de manière comparable à celle de l'embryogenèse somatique primaire. Ce modèle de culture, qui peut être schématisé en deux étapes, a été développé au laboratoire sur plusieurs génotypes de vigne (Maillot et al. 2006). La première étape consiste en un traitement qui doit permettre de « relancer » la capacité embryogénique des tissus de l'embryon. Dans notre cas, il s'agira d'un traitement par une auxine forte : le 2,4-D. Les embryons mis en

présence de cet inducteur se dédifférencieront tout d'abord en cals, dans un processus qualifié d'embryogenèse indirecte. Des embryons secondaires se formeront ensuite sur ces cals, devenant alors des cals embryogènes, après plusieurs semaines de culture en l'absence du 2,4-D. Ce type de processus peut être répliqué un nombre indéterminé de fois. Il est qualifié d'embryogenèse récurrente. Ce modèle de culture présente l'intérêt de pouvoir distinguer les deux phases clé du processus : l'induction de la capacité embryogénique par le traitement au 2,4-D et la différenciation des embryons secondaires. L'embryogenèse somatique secondaire constitue le modèle sur lequel nous avons réalisé une grande partie de nos analyses moléculaires.

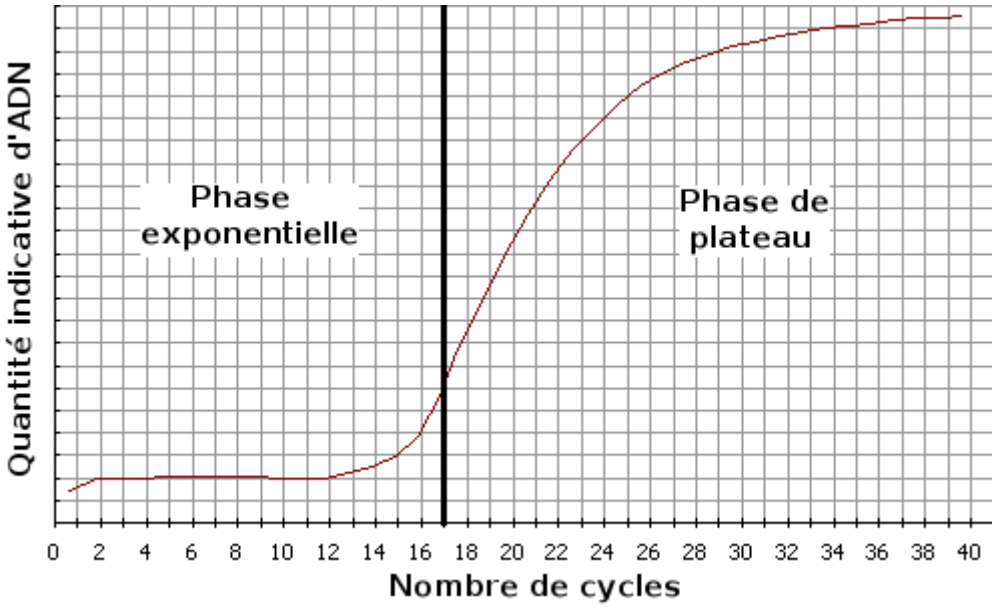


Figure 7 : Courbe d'amplification de l'ADN en fonction du nombre de cycles.

La phase de plateau est atteinte dès que le rendement de la réaction n'est plus exponentiel en raison d'un facteur limitant (épuisement d'un des réactifs libres de la réaction ou diminution d'efficacité de la polymérase).

Seconde partie : Etude de l'expression de gènes par PCR après transcription inverse (RT-PCR)

La technique d'amplification en chaîne (PCR, « Polymerase Chain Reaction ») après transcription inverse (RT-PCR ; « Reverse Transcriptase - PCR ») a été mise au point pour utiliser l'ARN comme matrice d'amplification de la PCR. Cette méthode permet de détecter (voire de quantifier) des ARN messagers (ARNm) et d'étudier les niveaux d'expression de gènes (en fonction du tissu, d'un traitement ...). De l'ADN complémentaire (ADNc) est synthétisé à l'aide d'une transcriptase inverse à partir d'ARN, puis la PCR est réalisée à partir de cet ADNc.

Lors d'une RT-PCR « conventionnelle », en point final (« end-point RT-PCR »), les amplicons sont visualisés sur gel, ce qui permet de déduire si le messager correspondant était présent ou absent de l'échantillon. Mais il n'est pas possible d'évaluer la variation du niveau d'expression du messager : induction ou répression, comme en RT-PCR semi-quantitative, ou de le quantifier, comme en RT-PCR quantitative. Au cours d'une PCR, l'amplification est exponentielle (doublement des amplicons à chaque cycle) avant d'atteindre une phase de plateau (épuisement des réactifs de la PCR) (Figure 7). En fin de réaction, la quantité d'amplicons obtenus n'est donc pas proportionnelle à la quantité de matrice au départ, alors qu'elle l'est durant la phase exponentielle de la PCR. C'est donc dans cette phase qu'il faut se placer pour quantifier l'expression des gènes par les techniques de RT-PCR semi-quantitative ou quantitative.

I. La RT-PCR semi-quantitative

En RT-PCR semi-quantitative, un nombre de cycles d'amplification est déterminé de telle façon à se situer en phase exponentielle de la réaction. Un gène dont l'expression est considérée comme stable dans les différents échantillons est utilisé comme référence pour

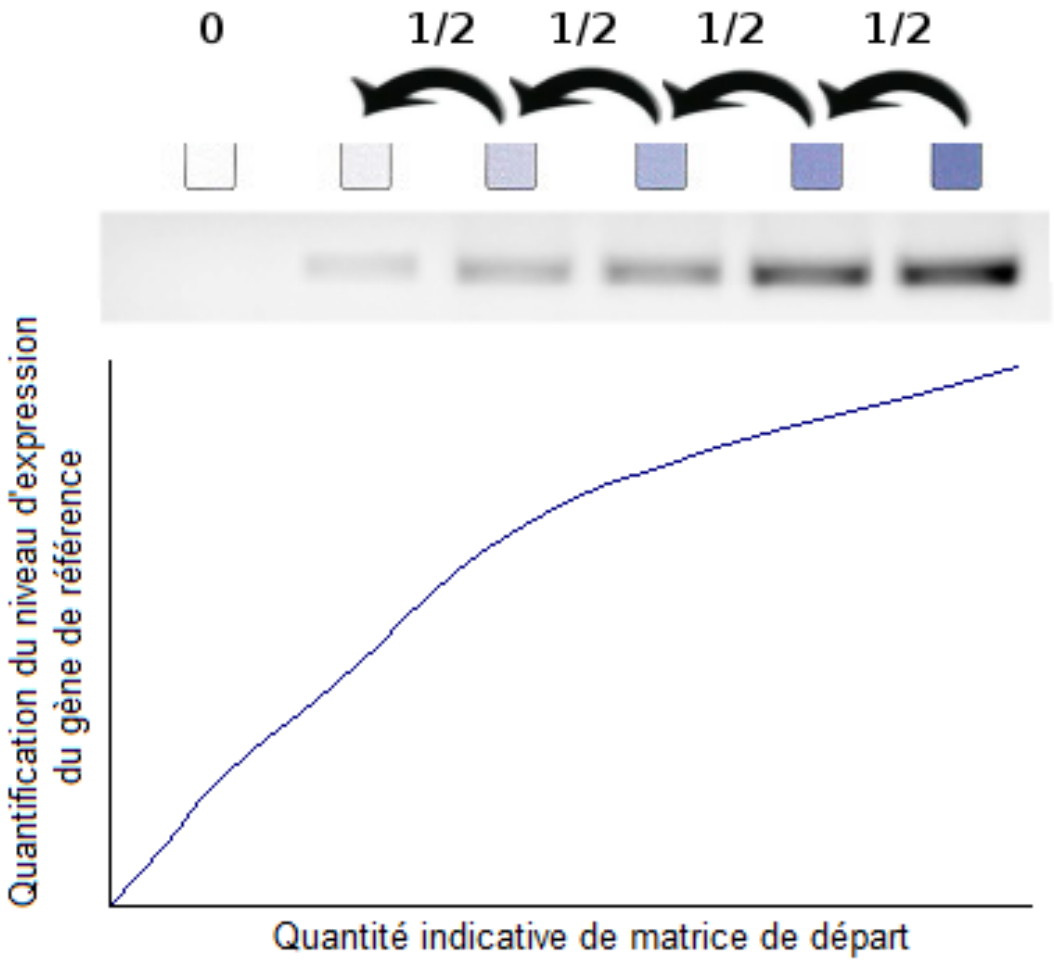


Figure 8 : Calibrage de conditions de PCR.

Afin d'être en conditions semi-quantitatives, il est nécessaire de choisir une quantité de matrice initiale pour laquelle, en fonction du nombre de cycles de réaction, le signal du gène de référence est détectable et non saturé. Cette quantité peut être déterminée par dilutions successives de l'échantillon.

réaliser le calibrage des conditions de la réaction contrôlant la quantité de matrice au départ. Après migration sur gel et révélation de l'ADN par un agent intercalant (ex. bromure d'éthidium), l'intensité des produits d'amplification est analysée et comparée à celle du gène de référence pour permettre une évaluation semi-quantitative (Figure 8). Il devient ainsi possible d'estimer les différences de quantités d'ADNc donc d'ARNm entre les échantillons, pour un même gène.

Toutefois, au cours d'une PCR, les conditions sont variables d'un échantillon à l'autre. En effet, les amorces employées, nécessairement différentes d'un gène à l'autre, impliquent que l'efficacité de la réaction sera systématiquement différente d'un gène à l'autre, rendant difficile la comparaison du niveau d'expression entre gènes différents. L'observation des produits s'effectuant après un nombre de cycles identiques pour tous les échantillons, ceux-ci ne seront pas forcément dans la même phase de la PCR. Pour certains échantillons, l'amplification pourrait être encore dans la phase exponentielle alors que pour d'autres, elle serait déjà en phase de plateau. De plus, la quantification des amplicons obtenus ne peut pas être précise puisque la révélation est assez peu sensible, et quand une bande est discernable, l'étape exponentielle de l'amplification peut être terminée. Ainsi, dépendant de la concentration initiale de matrice et suivant les conditions choisies (principalement le nombre de cycles de réaction et l'efficacité de l'hybridation des amorces), l'évaluation des produits amplifiés pourra entraîner une surestimation ou une sous-estimation du niveau d'expression d'un gène (Ferre 1992).

La RT-PCR semi-quantitative permet donc l'évaluation du niveau d'induction ou de répression d'un gène de manière relative, en fonction d'une référence considérée comme stable dans les différents échantillons.

II. La RT-PCR quantitative (RT-qPCR)

Pour la PCR quantitative après transcription inverse, l'abréviation utilisée dans ce manuscrit est **RT-qPCR** (« reverse transcriptase - quantitative PCR »). En effet, l'utilisation de l'abréviation « RT-PCR » pour la PCR quantitative « en temps réel » (« real time ») est source de confusion avec la PCR (conventionnelle) après transcription inverse (Bustin et al. 2009).

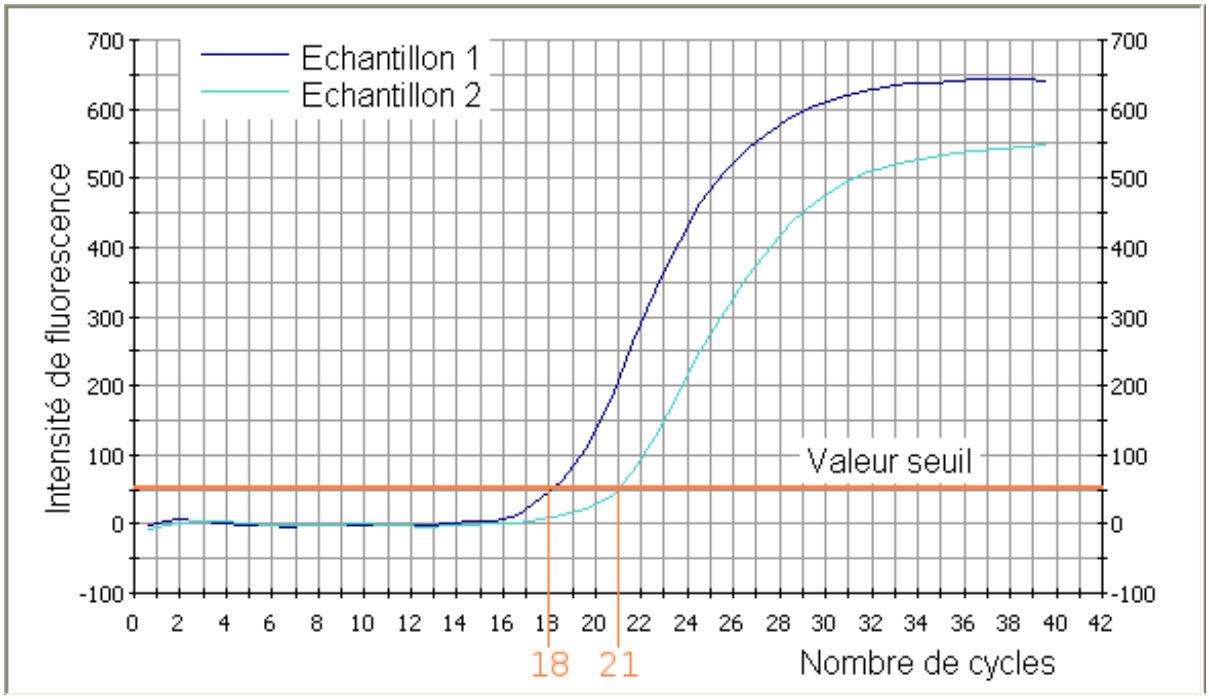


Figure 9 : Exemple de détermination de Cq.

La valeur seuil, ici fixée par l'appareil, se situe dans la phase exponentielle pour les deux échantillons. Le Cq de chaque échantillon est indiqué en orange. Le Cq de l'échantillon 1 (18) est plus faible que le Cq de l'échantillon 2 (21). Cela signifie qu'un plus grand nombre de copies est présent dans l'échantillon 1 que dans l'échantillon 2, au moment du début de la détection.

A. Principe

La technique de RT-qPCR consiste à évaluer la quantité d'ADNc initialement présente dans l'échantillon pour un gène donné. En effet, la RT-qPCR permet de suivre « en temps réel » (à tout moment) la quantité d'ADN produit par l'amplification grâce à la détection d'un signal fluorescent dont l'intensité est directement proportionnelle à cette quantité. Une courbe (niveau de fluorescence en fonction du nombre de cycles) est établie et elle est spécifique à chaque échantillon pour chaque gène. Cette courbe permet d'établir la valeur de seuil, qui est le niveau de fluorescence minimal pour que la courbe se détache du bruit de fond, tout en étant encore dans la phase exponentielle. Le nombre de cycles nécessaires pour atteindre cette valeur seuil constitue le Cq (« quantification cycle ») ou cycle de quantification (Figure 9). Les abréviations Ct (« threshold cycle »), Cp (« crossing point ») ou TOP (« take-off point ») sont également utilisées dans la littérature pour définir ce seuil mais dépendent du fabriquant et du thermocycleur utilisé. Cq sera ici utilisé, en accord avec les standards définis par le consortium RDML (« Real-Time PCR Data Markup Language » ; http://www.rdml.org ; Lefever et al. 2009). Cq est le nombre de cycles de la PCR utilisé pour la quantification de la cible, soit pour établir le nombre initial de copies (quantité absolue), soit pour déterminer le rapport d'expression entre deux échantillons (expression relative) (Tse et Capeau 2003). Il y a une relation inversement proportionnelle entre le Cq et le nombre initial de copies, puisqu'un transcrit présent en un faible nombre de copies nécessitera plus de cycles pour que sa fluorescence atteigne la valeur de seuil et aura un Cq plus élevé (Figure 9). Cette valeur Cq est obtenue en début de phase exponentielle et elle n'est pas influencée par les éléments limitants de la réaction de PCR en phase du plateau. L'établissement d'une droite étalon pour chaque gène permet de prendre en compte le facteur d'efficacité de la réaction et d'estimer le nombre initial de copies présentes dans l'échantillon. Il sera ainsi possible de comparer les niveaux d'expression des différents gènes entre eux.

La RT-qPCR présente un intérêt en terme de rapidité, de sensibilité et de spécificité pour la quantification d'ARNm, par rapport à la RT-PCR semi-quantitative (Bustin et Nolan 2004, Arya et al 2005). La PCR quantitative permet un gain de temps car les résultats sont

obtenus en « temps réel » et il n'est pas nécessaire de réaliser des manipulations supplémentaires (ex. migration sur gel des produits PCR). La spécificité de la PCR quantitative est améliorée car il est possible de vérifier si un seul produit a été synthétisé à l'aide de la courbe de fusion. Enfin, par sa sensibilité, la RT-qPCR permet la détection de produits de gènes présents en un très faible nombre de copies, le minimum théorique étant de 3 copies (Bustin et al. 2009).

B. Utilité de la RT-qPCR pour l'étude de l'expression de familles multigéniques

La rapidité, la sensibilité et la spécificité de la méthode de RT-qPCR la rendent particulièrement adaptée à l'étude de l'expression de gènes au sein de familles multigéniques (Gachon et al. 2004). D'une part, les membres de ces familles présentent des séquences proches les unes des autres, rendant délicate l'évaluation de leur expression par d'autres méthodes. Par exemple, l'expression de la famille multigénique *XTH* d'*A. thaliana,* qui compte 33 membres pouvant partager jusqu'à 90% de similarité nucléotidique, a ainsi pu être évaluée par RT-qPCR (Yokoyama et Nishitani 2001). D'autre part, dans ces familles, certains gènes sont très peu exprimés. Par sa sensibilité, la RT-qPCR permet la détection de produits d'expression présents en un très faible nombre de copies (Bustin et al. 2009). Elle rend ainsi possible l'étude de gènes présentant des niveaux d'expression très différents. Par exemple, la famille *AtNCED* d'*A. thaliana* dont les membres ont des niveaux d'expression variant de 1 à 600 fois a pu être étudiée par RT-qPCR (Tan et al. 2003).

C. Etapes clés du protocole de RT-PCR quantitative

Jusqu'en 2009, il n'existait pas de guide indiquant quelles sont les étapes cruciales pour évaluer une expérience de PCR quantitative puis procéder au traitement et à la publication des résultats. Il y a donc une importante hétérogénéité des méthodes employées pour réaliser la réaction, pour générer et traiter les résultats, aboutissant parfois à une validité discutable des résultats publiés. A cette date, des guides avaient déjà été proposés concernant la bonne utilisation de diverses techniques telles que les puces à ADN, des techniques de protéomique ou l'analyse de séquences génomiques. Le site du MIBBI

POINTS CLES	IMPORTANCE
PLAN D'EXPERIENCE	
Définition des groupes tests et témoins ; effectifs au sein des groupes	**E**
ECHANTILLON	
Description	**E**
Volume et masse	R
Micro- ou macrodissection	**E**
Prélèvement et conservation	**E**
EXTRACTION DES ACIDES NUCLEIQUES	
Protocole	**E**
Nom du kit ; modifications	**E**
Traitements (désoxyribonucléase ou ribonucléase)	**E**
Contrôle de la contamination (ADN ou ARN)	**E**
Quantification des acides nucléiques	**E**
Appareil et méthode	**E**
Pureté (A260/A280)	R
Concentration	R
Intégrité des ARN	**E**
RIN (RNA integrity number) RQI (RNA quality indicator) ou valeur de Cq des transcrits en 3' et 5'	**E**
Profil d'électrophorèse	R
Test d'inhibition (Cq de dilutions, ...)	**E**
TRANSCRIPTION INVERSE (RT)	
Conditions de réaction et quantités de réactifs	**E**
Valeur de Cq avec et sans RT (vérification de l'absence d'ADN contaminant)	R
Conservation des ADNc	R
INFORMATION SUR LA SEQUENCE CIBLE	
Efficacité et limite de détection pour chaque gène	**E**
Numéro d'accession de chaque séquence	**E**
Position des amorces	**E**
Longueur de l'amplicon	**E**
Spécificité des amorces (BLAST, etc)	**E**
Alignement, structure secondaire ...	R
AMORCES	
Séquences	**E**
Position et nature des éventuelles modifications	**E**
Conditions de synthèse (fabriquant, méthode de purification)	R
PROTOCOLE DE qPCR	
Conditions de réaction et quantités de réactifs	**E**
Fabriquant des plaques ou tubes ; numéro de catalogue	R
Règlage du thermocycleur et fabriquant	**E**
VALIDATION DE LA qPCR	
Optimisation (à partir de gradients)	R
Specificité (gel, séquence, courbe de fusion ou digestion)	**E**
Cq des contrôles sans matrice	**E**
Droite étalon	**E**
Efficacité de la qPCR	**E**
Intervalle de confiance de l'efficacité de qPCR ou erreur standard	R
Coefficient de corrélation	**E**
Valeurs des points de la gamme étalon ("linear dynamic range")	**E**
Cq minimal	**E**
Efficacité et limite de détection	**E**
ANALYSE DES DONNEES	
Programme	**E**
Détermination des Cq	**E**
Traitement des valeurs aberrantes	**E**
Résultats des contrôles sans matrice	**E**
Justification du nombre et du choix des gènes de référence	**E**
Méthode de normalisation	**E**
Nombre et concordance des expériences indépendantes	R
Nombre de répétitions techniques (pour la RT ou la qPCR)	**E**
Précision (variation intra-essai)	**E**
Reproductibilité (variation inter-essai, %CV)	R
Analyse statistique	**E**
Logiciel	**E**
Mise à disposition des données brutes en format RDML	R

Tableau 14 : Points clés concernant la validation du protocole de RT-qPCR et des résultats obtenus.

E = information essentielle ; R = informations recommandée (d'après Bustin et al. 2009).

(Minimum Investigation for Biological and Biomedical Investigation, http://www.mibbi.org) offre la possibilité d'accéder à ces guides (Taylor et al. 2008). La publication récente de « The MIQE Guidelines: Minimum Information for publication of Quantitative real-time PCR Experiments » (MIQE) a permis d'établir de manière claire et précise quels sont les points clés d'une bonne analyse par PCR quantitative (Bustin et al. 2009). Une nomenclature standardisée en accord avec le « Real-time PCR Data Markup Language » (RDML) a également été proposée (Bustin et al. 2009, Bustin 2010, Taylor et al. 2010).

Les points les plus critiques décrits par MIQE pour un protocole de RT-qPCR (et permettant, en particulier l'étude d'une famille multigénique) sont présentés dans le Tableau 14. Ils présentent les informations requises à chaque étape pour valider les résultats obtenus. Ils comprennent le plan d'expérience, la préparation des échantillons d'ADNc (incluant le prélèvement des tissus, l'extraction d'ARNm et la transcription inverse), les informations sur l'amplicon et les amorces, les informations sur les conditions de la réaction de qPCR et les contrôles pour la validation de celle-ci et enfin la méthode employée pour l'analyse des données. Pour chacune de ces étapes, MIQE fournit une série de points de contrôle à valider. Ce sont les recommandations que j'ai suivies.

Chapitre II : Expression des gènes *PR10* pendant le processus d'embryogenèse somatique

B. Partie expérimentale

I. Etat des connaissances sur les marqueurs de l'embryogenèse somatique chez *Vitis vinifera* au début de ma thèse.

Au début de ma thèse, aucune étude n'avait été publiée sur les marqueurs de l'embryogenèse somatique chez *Vitis vinifera*. Néanmoins, un travail sur les gènes *SERK* et *LEC* était en cours au laboratoire. Il a permis de caractériser les gènes *VvSERK1*, *VvSERK2*, *VvSERK3* et *VvL1L* et de montrer leur régulation au cours de la différenciation d'embryons à partir de lignées cellulaires embryogènes de Chardonnay (Schellenbaum et al. 2008). Ce premier travail suggérait un rôle de ces différents gènes dans le développement embryonnaire de la vigne.

II. Presentation de la publication « Differential regulation of SERK, LEC1-Like and Pathogenesis-Related genes during indirect secondary somatic embryogenesis in grapevine ».

A la suite de l'identification des gènes *VvSERK1, VvSERK2, VvSERK3* et *VvL1L,* exprimés dans des lignées embryogènes stabilisées du cultivar Chardonnay, nous avons voulu préciser leur implication dans le processus d'embryogenèse somatique chez la vigne. Pour ce faire, nous avons choisi de caractériser leur expression pendant le cycle d'embryogenèse somatique secondaire, en ciblant deux moments clés : la phase dite d'induction de l'embryogenèse, due au traitement au 2,4-D, et qui marque l'entrée dans le cycle, et la phase de différenciation des embryons secondaires qui marque la sortie de ce cycle. Ce modèle d'embryogenèse a été développé, avant mon arrivée dans le laboratoire, de manière à fournir, de manière récurrente, des générations nombreuses d'embryons somatiques au développement synchrone. Le développement de ce modèle fait l'objet de la première partie de la publication. La seconde partie présente les résultats d'analyse de l'expression de gènes impliqués dans le développement (*SERK* et *L1L*) mais aussi de gènes impliqués dans les défenses (*PR*), pendant le processus d'embryogenèse somatique secondaire. Les gènes de défense choisis étaient, d'une part des gènes codant des chitinases (*PR3* et *PR8*), ainsi qu'un gène *PR1* au rôle inconnu, et d'autre part les gènes *PR10.1* et *PR10.2,* seuls gènes de la famille *PR10* caractérisés à l'époque du début de ma thèse. Ma participation a été de contribuer à l'analyse de l'expression de ces gènes par RT-PCR.

Cette étude a tout d'abord permis l'identification d'un troisième gène *PR10* chez *Vitis vinifera,* très similaire à *VvPR10.1,* que nous avons caractérisé et nommé *VvPR10.3* (GenBank : EU379313.1). Les résultats principaux de l'analyse de l'expression sont les suivants :

- l'entrée dans le cycle d'embryogenèse, à la suite du traitement au 2,4-D, est marquée par une répression de l'expression des gènes *VvSERK1, VvSERK2* et *VvL1L,* tandis que l'expression des gènes *VvPR1, VvPR8, VvPR10.1* et *VvPR10.3* est fortement stimulée, suggérant que ces gènes de défense sont induits par le traitement causant l'induction embryogène

- la sortie du cycle est au contraire marquée par la stimulation de l'expression des gènes *VvSERK1*, *VvSERK2* et *VvL1L* dans les embryons néo-formés, et par la répression de celle des gènes *VvPR1*, *VvPR8*, *VvPR10.1* et *VvPR10.3*, montrant l'importance des gènes *VvSERK1*, *VvSERK2* et *VvL1L* dans le développement embryonnaire
- la persistance de l'expression des gènes *VvPR10.1* et *VvPR10.3* pendant plusieurs semaines après le traitement par le 2,4-D, au contraire des gènes *VvPR1* et *VvPR8*, suggère qu'ils pourraient avoir un rôle particulier dans le conditionnement des cultures avant la différenciation des embryons.

Differencial regulation of SERK, LEC1-Like, and Pathogenesis-Related genes during indirect secondary somatic embryogenesis in grapevine, Maillot P, Lebel S, Schellenbaum P, Walter B, 2009

Plant Physiol Biochem **47:** 743-752

Plant Physiology and Biochemistry 47 (2009) 743–752

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Differential regulation of *SERK*, *LEC1-Like* and *Pathogenesis-Related* genes during indirect secondary somatic embryogenesis in grapevine

Pascale Maillot*, Sylvain Lebel, Paul Schellenbaum, Alban Jacques [1], Bernard Walter

Université de Haute Alsace, Laboratoire Vigne Biotechnologies & Environnement, 33 rue de Herrlisheim, BP 50568, 68 008 Colmar Cedex, France

ARTICLE INFO

Article history:
Received 9 December 2008
Accepted 31 March 2009
Available online 15 April 2009

Keywords:
Grapevine
Somatic embryogenesis
Pathogenesis-Related genes
SERK
L1L

ABSTRACT

A culture model was developed in *Vitis vinifera* L., cultivar 'Chardonnay' for studying SE (Somatic Embryogenesis). The auxin 2,4-D (2,4-Dichlorophenoxyacetic acid) was used to induce indirect secondary embryogenesis at a high rate, starting from embryos derived from embryogenic cultures previously obtained. Cotyledonary embryos were shown to be more responsive to SE induction than embryos at the torpedo-stage and were used for molecular analyses. The expression of *SERK* (*Somatic Embryogenesis Receptor Kinase*), *L1L* (*Leafy Cotyledon1 Like*) and a set of *PR* (*Pathogenesis-Related*) genes was monitored during the whole SE process. *VvSERK1*, *VvSERK2* and *VvL1L* were down-regulated by the 2,4-D treatment but expressed in embryonic tissues. On the contrary, *VvPR1*, *VvPR8*, *VvPR10.1* and *VvPR10.3* were strongly up-regulated by the 2,4-D treatment, and their transcripts were not or only weakly detected in clusters of secondary embryos. *VvSERK3*, *VvPR3* and *VvPR10.2* were more stably expressed in all tissues examined. The discussion deals with the putative role of the different genes in grapevine SE.

© 2009 Elsevier Masson SAS. All rights reserved.

1. Introduction

Somatic embryos are the most suitable tissues for transformation of grapevine, transgenesis providing new opportunities for improving this world-wide important crop [44]. Unfortunately, economically important cultivars remain recalcitrant to in vitro biotechnologies. Although SE (Somatic Embryogenesis) can be induced from diverse floral and vegetative explants, embryos are produced at a rather low rate [10,23,39,43,46,51]. Secondary embryogenesis is used to increase the simultaneous production of numerous embryos, however limited to a handful of cultivars [9,42]. Much progress is therefore expected for developing SE in grapevine. The elucidation of molecular events occurring during the embryogenic process would certainly contribute to this purpose.

Due to technical convenience, embryogenesis has especially been studied in model plants such as *Daucus carota* and *Arabidopsis thaliana*, mutants having brought important information about this process [16,21,53]. Numerous genes have been shown to be regulated during zygotic and somatic embryogenesis: genes encoding transcription factors and cell cycle regulators, homeotic genes, hormone-inducible genes, maturation genes, and several genes encoding extracellular proteins such as AGPs (ArabinoGalactan Proteins) or secreted chitinases and lipid transfer proteins [7]. Recently, transcriptomic studies of chicory and wheat embryogenesis highlighted the regulation of a wide set of genes related to several biosynthetic pathways [36,64]. Nevertheless, the specific role of most genes as well as their functional connections remain to be determined. Furthermore, while embryo morphogenesis and maturation become more and more clear, very few is known about the molecular events involved in the shift of a somatic cell to an embryogenic state. *SERK* (*Somatic Embryogenesis Receptor Kinase*) is the best characterized gene involved in early embryogenesis. First identified in carrot, *SERK* was found to be especially expressed in embryonic tissues of many dicots and monocots [24,29,61,63,65]. The ectopic expression of *SERK1* in *Arabidopsis* was shown to result in the enhancement of embryonic cell formation, highlighting its crucial role during embryogenesis induction [24]. *LEC* (*Leafy Cotyledon*) and *L1L* (*Leafy Cotyledon1 Like*) genes that code for transcription factors have also been shown to play a significant role in early steps of plant zygotic and somatic embryogenesis

Abbreviations: 2,4-D, 2,4-Dichlorophenoxyacetic acid; AGPs, ArabinoGalactan Proteins; BAP, 6-BenzylAminoPurine; IAA, Indole-3-Acetic Acid; L1L, Leafy Cotyledon1 Like; NAA, Napthaleneacetic Acid; NOA, 2-NaphtOxyacetic Acid; PR, Pathogenesis-Related; SE, Somatic Embryogenesis; SERK, Somatic Embryogenesis Receptor Kinase.

* Corresponding author. Tel.: +33 3 89 20 23 55; fax: +33 3 89 20 23 57.
E-mail address: pascale.maillot@uha.fr (P. Maillot).
[1] Present address: Botanical Institute of Cologne University, Gyrhofstr. 15, D-50931 Cologne, Germany.

0981-9428/$ – see front matter © 2009 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.plaphy.2009.03.016

[1,15,22,35]. Recently, we characterized three *SERK* genes (*VvSERK1*, *VvSERK2*, *VvSERK3*) and one *L1L* gene (*VvL1L*) expressed in a stabilized embryogenic cell line of grapevine [62].

SE is usually obtained from tissues subjected to varied stresses or treated with hormones, mainly the synthetic auxin 2,4-D (2,4-Dichlorophenoxyacetic acid) [16,21]. This auxin analog is known for its herbicide activity, inducing a general oxidative stress in treated seedlings, followed by protein degradation that probably results in phytotoxicity [60]. A short treatment of plant tissues is able to trigger the reprogramming of somatic cells toward embryogenesis, probably as a consequence of stress induction together with the modification of internal auxin content [53]. As a result, defence proteins can be synthesized, further suggesting that SE is an adaptative process of plant cells to stress [16]. Pathogenesis-Related (PR) proteins are part of the plant defence responses. They are divided among different classes based on their structure, serological relationships and biological functions, and are usually observed upon biotic and abiotic stresses [69]. Although some PR proteins clearly show an anti-microbial activity, many reports deal with a probable additional role in plant development, including SE. For example, most proteins secreted during embryogenesis code for chitinases or glucanases [53]. The accumulation of β-1,3-glucanases, acidic chitinases and osmotin-like proteins is correlated with SE, in chicory [26,27]. Chitinases belong to a large family of enzymes able to hydrolyse the β-1,4-glycoside bond present in biopolymers of N-acetylglucosamine, such as chitin, therefore being thought to degrade fungi at the time of a pathogen attack. In addition, chitinases can cleave plant endogenous substrates such as cell wall AGPs, producing oligosaccharides that could play a role in development as signal molecules [32]. A chitinase has been shown to rescue the carrot mutant *tsl1*, arrested at the globular stage of embryonic development, probably by the release of AGPs derived molecules able to stimulate SE [11,68]. Likewise, although concluding experiments suggest a defence role of PR10 proteins as ribonucleases, these intracellular proteins are probably assigned to additional functions in plant growth and development [38]. Betv1, the major allergenic PR10 protein of birch, is able to bind such diverse molecules as fatty acids, flavonoids, cytokinins and brassinosteroids, probably playing a role in the cytoplasmic transport of hydrophobic ligands and interfering with hormonal regulation in plants [17,33,40,45].

In the following study, we focused on the development of an in vitro culture model allowing the recovery of large amounts of synchronized somatic embryos, with the aim of stuying molecular events of grapevine embryogenesis. We monitored the expression of *VvSERK1*, *VvSERK2*, *VvSERK3* and *VvL1L* on the one hand, and of a set of defence genes, on the other hand: a *PR3* gene and a *PR8* gene respectively coding for a typical basic vacuolar chitinase and an acidic secreted chitinase, three *PR10* genes and a *PR1* gene coding for a secreted protein with an unknown function.

2. Results

2.1. Induction of secondary embryogenesis

Embryogenic cultures previously obtained from the cultivar 'Chardonnay' were used to produce numerous embryos at the same time (Fig. 1A). For inducing secondary embryogenesis, embryos

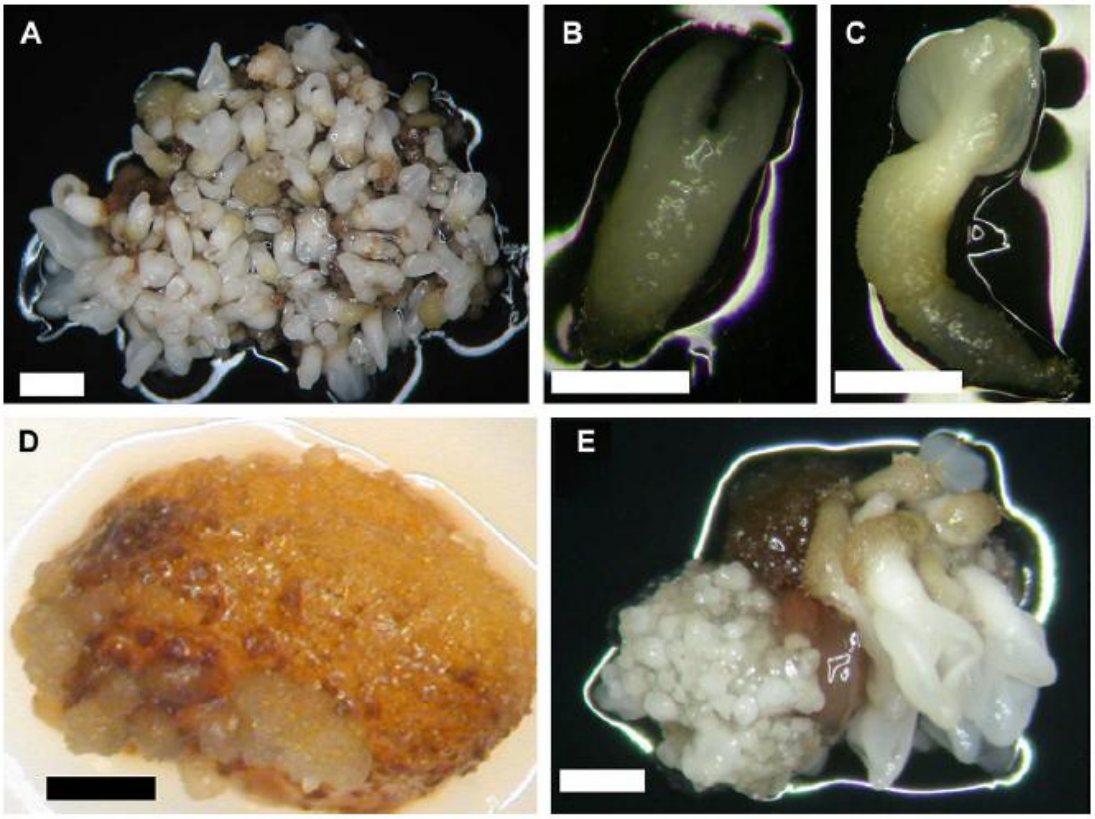


Fig. 1. Induction of secondary embryogenesis from somatic embryos of *Vitis vinifera* cv. 'Chardonnnay'. (A) Starting embryogenic cultures providing numerous embryos at the same time; (B) Embryo at the torpedo-stage; (C) Embryo at the cotyledonary-stage; (D) Compact yellowish callus obtained from an embryo after culture on a medium containing 2,4-D; (E) Secondary embryogenic clusters composed of embryos at varied development stages on the surface of an embryogenic callus. Bars = 2 mm (A), 1 mm (C, D and E) and 0.5 mm (B).

either at the torpedo- (Fig. 1B) or at the cotyledonary-stage (Fig. 1C) were selected. To assess secondary embryogenesis efficiency, four independant experiments were performed involving a total of 346 embryos, separately plated on a culture medium with 2,4-D. After 3 weeks, a compact and yellowish callus was obtained from every embryo (Fig. 1D). After their transfer onto the medium A, calli stopped growing and progressively turned brown. Clusters of secondary embryos appeared on some calli within an additional culture period of 1–4 months on the medium A (Fig. 1E): they were composed of a mix of very small and translucent globular structures and of opaque white embryos at varied developmental stages. In average, 29% of the plated embryos at the torpedo-stage and 68% of the embryos at the cotyledonary-stage gave rise to secondary embryogenic calli (Table 1). In a further experiment, embryos at the cotyledonary-stage were used to induce secondary embryogenesis at a high rate and perform molecular analyses.

2.2. Identification of PCR products

PCR products were sequenced and compared to the whole *Vitis vinifera* genome recently available in the Genoscope databank. When required, a comparison with sequences of other flowering plants was performed, using the BLAST program of the National Center for Biotechnical Information (NCBI) databank. The sequences obtained with primers specific of *VvSERK1*, *VvSERK2*, *VvSERK3* and *VvL1L* cDNAs corresponded to the annotations GSVIVT00001777001 on the chromosome 18, GSVIVT00019412001 on the chromosome 7, GSVIVT00009544001 on the chromosome 12 and GSVIVT00010958001 on the chromosome 10, respectively. They were identical to those previously described [62]. As expected, the sequence amplified with specific primers of a *VvPR1* cDNA was found to be identical to the accession AJ536326; it was located on the chromosome 3. The sequence amplified with specific primers of a *VvPR10.2* cDNA was identical to that of the *VvPR10.2* gene of the cultivar 'Ugni blanc' (AJ291704) and was located on the chromosome 5. The sequence amplified with the primers corresponding to a *VvPR3* cDNA was found to be identical to the *VCHI1b* gene isolated from 'Pinot noir' [4] and to *VvChit1a* from 'Ugni blanc' [58], located on the chromosome 3 and coding for the same basic vacuolar class I chitinase (AJ291505). The sequence amplified with primers corresponding to a *VvPR8* cDNA showed 97% homology with a class III chitinase gene of the cultivar 'Koshu' (AB105374) and 93% with a class III chitinase gene of 'Pinot noir' and 'Ugni blanc' (AJ291507). This *VvPR8* gene was found to be located on a contig not linked to a known chromosome in the *V. vinifera* genome.

Two 244 bp long different sequences were amplified with primers previously described for the specific amplification of a *VvPR10.1* cDNA (AJ291705) [57]. A first sequence showed 99.6% homology with the Genoscope annotation GSVIVT00033089001 related to *VvPR10.1*, while a second sequence showed 100% homology with a genomic sequence located on the chromosome 5 at a location clearly different from that of *VvPR10.1*. At this location, two genomic sequences were annotated as separate very short transcripts, but not as a single *VvPR10* gene. However it corresponded to grapevine ESTs in the banks of the Istituto Agrario di San Michele all'Adige (IASMA), the Genoscope and the Dana-Farber Cancer Institute (DFCI): ISMAAEST003738, CDMVvT00000963001 and TC74380, respectively, and was recognized as the incomplete cDNA of a putative *VvPR10* gene in the NCBI bank. The identification of the complete sequence was performed using a RACE-PCR method. A set of primers was specifically designed for the elongation and amplification of this putative *VvPR10* cDNA. A 799 bp long sequence, colinear with the first sequence of 244 bp, was obtained and sequenced. The complete sequence included a 480 bp long region coding for a polypeptide of 159 amino-acid residues, a 5′-UTR of 80 bp and a 3′-UTR of 239 bp. The full gene corresponding sequence deduced from the Genoscope databank is interrupted by an intron of 129 bp. The nucleotidic sequence and the predicted translation product present the characteristics of genes and proteins related to the birch major pollen allergen Betv1 [28]: a conserved glycine-rich "P-loop" motif GXGGXGXXK at the position 47–55 associated with a putative hydrophobic cavity at the position 89–121 and an intron position at codon 62 (Fig. 2A), as well as a calculated molecular mass of 17.4 kDa and a theoretical isoelectric point of 6.3. The full amino acid sequence shows the maximal percentage of identity (96%) with *Vitis quinquangularis* PR10.3 (ABD78555), a high similarity with *Vitis pseudoreticulata* PR10 proteins and is more distinct from *V. vinifera* PR10.1 and PR10.2 (Fig. 2B and C). According to the nomenclature recommendations, the corresponding gene has been named *VvPR10.3* and the cDNA sequence has been submittted for publication in the GenBank available to EMBL in Europe and the DNA Data Bank of Japan (NCBI): EU379313 coding for the protein ACA58119. The amplification of two distinct *VvPR10* cDNAs with primers previously designed for the specific amplification of a *VvPR10.1* transcript [57] is explained by a high homology of *VvPR10.1* and *VvPR10.3* sequences at the locations of primer hybridization. New primers were then designed for the separate amplification of *VvPR10.1* and *VvPR10.3* (Fig. 3).

Table 1
Efficiency of secondary embryogenesis. Somatic embryos either at the torpedo-stage (experiments A and B) or at the cotyledonary-stage (experiments C and D) were separately treated for inducing secondary somatic embryogenesis.

Embryo stage	Torpedo		Cotyledonary	
Experiment	A	B	C	D
Number of plated embryos	79	117	85	65
Number of secondary embryogenic calli	23	34	65	39
Secondary embryogenesis efficiency %	29.11	29.06	76.47	60.00
Average efficiency[a] %	**29.09** ± 0.04		**68.24** ± 11.65	

[a] Average efficiency = mean of secondary embryogenesis efficiencies obtained in separate experiments and calculated as the number of secondary embryogenic calli recovered from 100 tested embryos, ±standard deviation.

2.3. Gene expression

We recorded the expression of some *PR* genes during the whole process of secondary embryogenesis as well as that of *SERK* and *L1L* genes usually involved in plant embryogenesis (Fig. 4). *Actin* was used as a reference gene. The transcription of all the genes except *VvPR1* was detected in cotyledonary embryos (E) at the beginning of the experiment, although it was very weak for *VvPR8*, which encodes a class III chitinase. Then, the expression of *VvPR10.2* was scarcely increased by the 2,4-D treatment, while that of *VvPR1*, *VvPR8*, *VvPR10.1* and *VvPR10.3* was clearly up-regulated (C1). Calli expressed every *PR* gene at the same level after being transferred onto the medium A for one week, suggesting that gene expression was not immediately affected by the removal of 2,4-D (C2). Later, *VvPR3* and *VvPR10* transcripts were detected in calli, but not *VvPR1* nor *VvPR8* transcripts (NEC and EC). In clusters of secondary embryos detached from embryogenic calli (CSE), the expression of *VvPR3* and *VvPR10* genes was detectable at a lower rate than in embryogenic calli, whereas that of *VvPR1* and *VvPR8* was not detected at all. On the whole, *PR* genes were similarly expressed in secondary embryogenic clusters (CSE) and in embryos at the beginning of the experiment (E). In germinated embryos (GE), these genes were expressed at the same or at a weaker level than in cotyledonary embryos (E) and secondary embryogenic clusters (CSE). Together, these results show that *PR* genes were differently induced by the 2,4-D treatment, *VvPR1*, *VvPR8*, *VvPR10.1* and

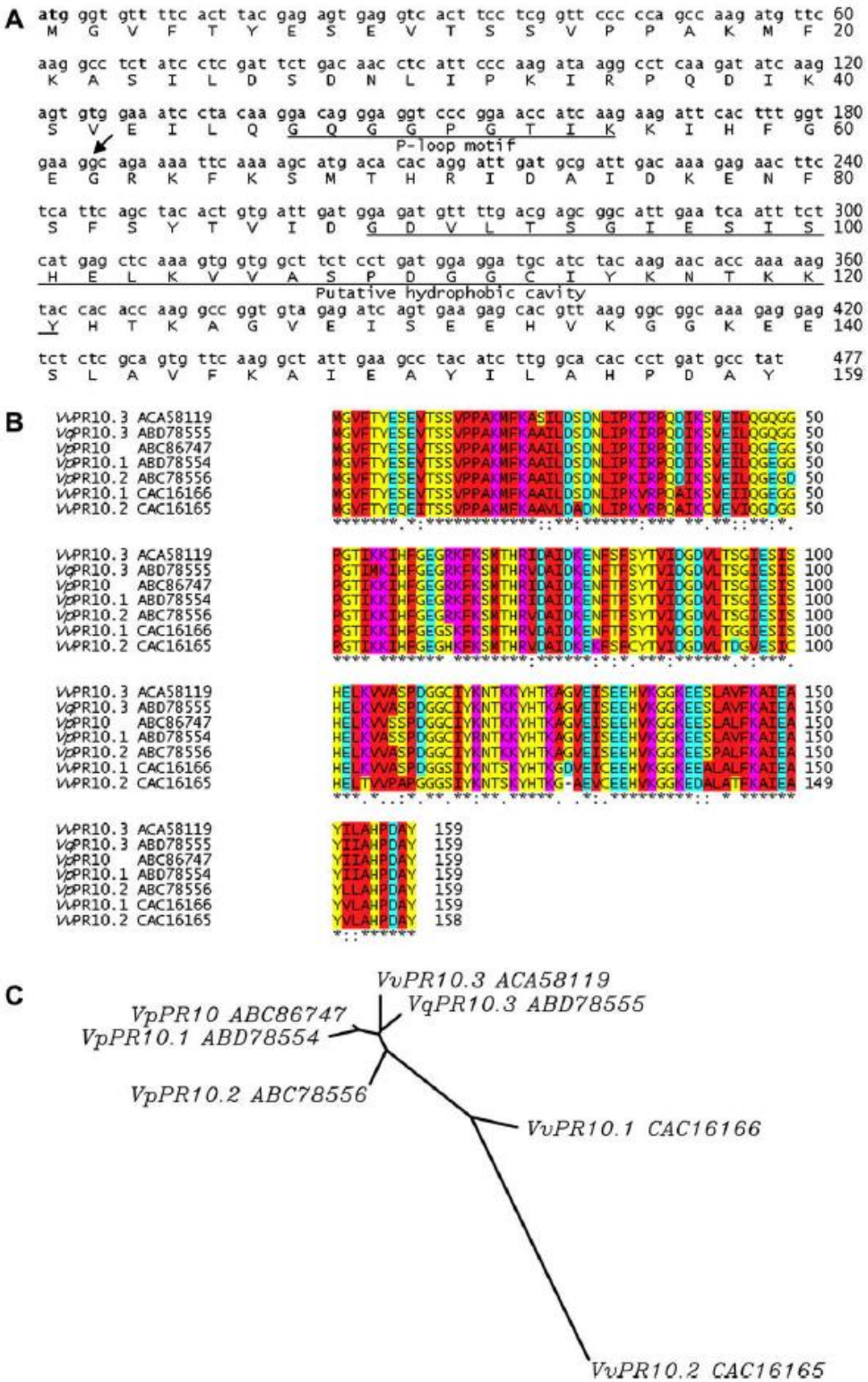


Fig. 2. Sequence analysis of *Vitis vinifera* PR10.3 (ACA58119) and comparison with other PR10 proteins in *Vitis sp.* (A) Nucleotidic sequence and predicted translation product of *VvPR10.3*. The main characteristics of Betv1-related proteins are presented: a P-loop motif at the amino-acid position 47–55, a hydrophobic cavity at the amino-acid position 89–121 and an intron position indicated by an arrow at the codon 62. Bold letters indicate the ATG initiation codon; (B) Alignment with ClustalW of the amino-acid sequence of *Vitis vinifera* PR10.3, with the most related PR10 proteins: *Vitis quinquangularis* PR10.3 (ABD78555), *Vitis pseudoreticulata* PR10 (ABC86747), PR10.1 (ABD78554) and PR10.2 (ABC78556), and *Vitis vinifera* PR10.1 (CAC16166) and PR10.2 (CAC16165); (C) Unrooted dendrogram showing the approximate distance of *Vitis vinifera* PR10.3 with the most related *Vitis* PR10 sequences.

```
PR10.1    TCACCTCAAACCATTCTCTGCAAACCAACCAATCCTCCTCCTCTTCCTCTTTCGATCCTT 60
PR10.3    TTACCTCAAACCACTCTCTGCAAACCAACCAATCCTCCTTGTCTTCCTCTTTTGATCCTT 60

                                                 0276
                                         5'-CGAGAGTGAGGTCACTTCCTCG-3'
PR10.1    TTCATTTCAAACTCTAAGATCATGGGTGTTTTCACTTACGAGAGTGAGGTCACTTCCTCG 120
PR10.3    TTCATT-CAAACTCTAAGATCATGGGTGTTTTCACTTACGAGAGTGAGGTCACTTCCTCG 119

PR10.1    GTTCCCCCAGCCAAGATGTTCAAGGCCGCTATCCTCGATTCTGACAACCTCATTCCCAAG 180
PR10.3    GTTCCCCCAGCCAAGATGTTCAAGGCCTCTATCCTCGATTCTGACAACCTCATTCCCAAG 179

                                         SL101
                                 5'-GAAATCATACAAGGAGAGGGAGGC-3'
PR10.1    GTAAGGCCTCAAGCTATCAAGAGTGTGGAAATCATACAAGGAGAGGGAGGCCCTGGAACC 240
PR10.3    ATAAGGCCTCAAGATATCAAGAGTGTGGAAATCCTACAAGGACAGGGAGGTCCCGGAACC 239
                                 5'-GAAATCCTACAAGGACAGGGAGGT-3'
                                         SL103

PR10.1    ATCAAGAAGATTCACTTTGGTGAAGG/_________/CAGCAAATTCAAAAGCATGACACA 412
PR10.3    ATCAAGAAGATTCACTTTGGTGAAGG/_________/CAGAAAATTCAAAAGCATGACACA 418

                                         0152
                                 3'-CTTCACATTCAGCTACACTGTGGTTG-5'
PR10.1    CCGGGTTGATGCGATTGACAAAGAGAACTTCACATTCAGCTACACTGTGGTTGACGGAGA 472
PR10.3    CAGGATTGATGCGATTGACAAAGAGAACTTCTCATTCAGCTACACTGTGATTGATGGAGA 478

PR10.1    TGTTTTGACGGGCGGCATTGAATCAATTTCTCATGAGCTCAAAGTGGTGGCTTCTCCTGA 532
PR10.3    TGTTTTGACGAGCGGCATTGAATCAATTTCTCATGAGCTCAAAGTGGTGGCTTCTCCTGA 538

PR10.1    TGGAGGATCCATCTACAAGAACACCAGCAAGTACCACACCAAGGGCGATGTAGAGATCTG 592
PR10.3    TGGAGGATGCATCTACAAGAACACCAAAAAGTACCACACCAAGGCCGGTGTAGAGATCAG 598
                            3'-AAAAGTACCACACCAAGGCCG-5'
                                         250

PR10.1    TGAAGAGCACGTTAAGGGTGGCAAAGAGGAGGCTCTGGCATTGTTCAAGGCTATCGAAGC 652
PR10.3    TGAAGAGCACGTTAAGGGCGGCAAAGAGGAGTCTCTCGCAGTGTTCAAGGCTATTGAAGC 658

PR10.1    CTACGTCCTGGCACATCCCGATGCCTATTAAGTAAAATTGCCTGTAGTAATTGAGATATT 712
PR10.3    CTACATCTTGGCACACCCTGATGCCTATTAAGTAAAATTGCCTGTGGTAATTGAGAAATT 718

                                                 439
                                         3'-CACCTATCAGTCTCAATAAGTTTGG-C-5'
PR10.1    AGTCTCTTGAGTCCACCTTCATATTCATTGCACCTATCAGTCTCAATAAGTTTGG-CCTT 771
PR10.3    AGTCTCTTGAGTCCACTTTCATATTCATAGAA----TCAGTCTTAATAAGTTTGGATCTT 774

PR10.1    TTGTGTTTTT-GCCATATGGCCAAAGTAGCCTCGGCTTAGAAATAAGAGTTTCCCAGGTG 830
PR10.3    TTGTGTTTTTTGCCGTATGGCCAAAGTAGCCTCGGCTTAGAAATAAGAGTTTCCTAGGTG 834

PR10.1    TGATCCTTGG-AGTATTTGTGATGTTCTATTGCCAT 865
PR10.3    TGATCCTTGGGAGTACTTGTAATGTTTTATTTCAAT 870
```

Fig. 3. Hybridization of three primer pairs with *VvPR10.1* and *VvPR10.3* cDNAs: '0152–0276' detecting the two cDNAs, 'SL101–439' for the specific detection of *VvPR10.1*, and 'SL103–250' for the specific detection of *VvPR10.3*. Bold types indicate nucleotidic differences at the hybridization sites. The ATG initiation and TAA terminator codons are underlined. Slashes indicate the intron position.

VvPR10.3 being very responsive to the treatment, contrary to *VvPR3* and *VvPR10.2*. Further, only *VvPR3* and *VvPR10* transcripts were detected in calli at the latest culture steps (NEC and EC). Contrary to defence genes, *VvSERK1*, *VvSERK2* and *VvL1L* expression was recorded at the highest level in cotyledonary embryos at the beginning of the embryogenic process (E) and in secondary embryogenic clusters (CSE). It was slightly down-regulated by the 2,4-D treatment (C1) and remained low in calli (C2, NEC and EC). By contrast, *VvSERK3* expression was rather stable during the experiment. In germinated embryos (GE), the expression of *SERK* and *L1L* genes was similar or lowered compared to cotyledonary embryos (E) or secondary embryogenic clusters (CSE).

3. Discussion

3.1. Induction of secondary embryogenesis

Embryogenic cultures provided us with numerous secondary embryos that we brought to an additional embryogenic cycle. Embryos at the torpedo- or cotyledonary-stage were separately plated on a medium containing 2,4-D to form calli. After being transferred onto the medium A, some calli gave rise to clusters of secondary embryos within a period of 1–4 months. SE was therefore an indirect process. As a result, 29–76% of the plated embryos were led to secondary embryogenesis. In previous experiments

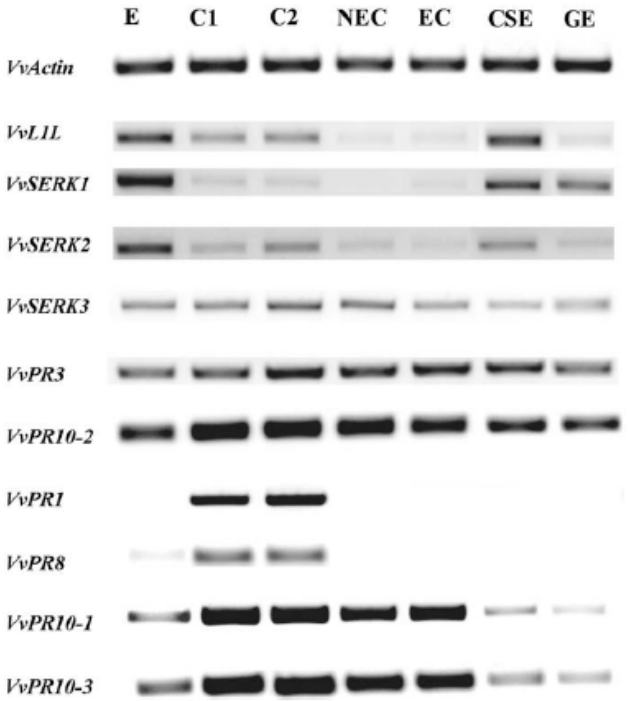


Fig. 4. Expression of *Vitis vinifera Actin*, *L1L*, *SERK*, and *PR* genes, during induction of secondary embryogenesis. E: cotyledonary embryos at the beginning of the experiment; C1: calli developed from embryos on the medium E96; C2: calli after one week subculture on the medium A; NEC: non-embryogenic calli and EC: embryogenic calli cleared of embryogenic clusters, after 12 weeks on the medium A (3 subcultures); CSE: clusters of secondary embryos; GE: germinated embryos.

aimed to induce primary embryogenesis, we showed that only 6% of nodal explants from the cultivar 'Chardonnay' gave rise to embryogenic calli [39]. This result is in accordance with other studies showing that secondary embryogenesis is more effective than primary embryogenesis in grapevine [9,42], just as in other plant species [54]. It further suggests that embryos are more suitable than vegetative organs for SE induction. Futhermore, the use of zygotic embryos has made SE possible for plant species considered to be recalcitrant [21]. Embryos at the cotyledonary-stage gave around 2.3 fold more embryogenic calli than younger embryos, showing that they were more responsive to SE induction. Likewise, immature zygotic embryos of *Arabidopsis*, with fully-developed cotyledons, gave better results than younger embryos in secondary SE experiments [20]. On the contrary, hot pepper embryos at the globular- and heart-stage were found to have a higher potential than cotyledonary embryos for direct SE [48]. Different responses can thus be obtained depending on the species and/or protocol. Our protocol induced indirect SE: a cell dedifferentiation step followed by cell proliferation resulted in a growing callus before the development of secondary embryos. *Arabidopsis* zygotic embryos at the cotyledonary-stage can also be led to indirect SE; secondary embryos develop from proliferating cells arising from the cotyledons, but never from the other parts of the embryo, showing that cotyledons contain those cells able to enter an embryogenic program [55]. This further suggests that well-developed embryos are quite suitable for inducing SE.

3.2. Expression of SERK and L1L genes during embryogenesis

The ability of differentiated tissues to generate embryos, termed 'embryogenic competence', is often induced by a stress or a treatment with hormones [16,21]. Embryogenic competence is generally accompanied by an enhanced expression of *SERK* genes coding for transmembrane somatic embryogenesis receptor kinases. *SERK* expression has thus been recorded in zygotic and somatic embryos of numerous plant species [24,29,61,63,65]. Moreover, *AtSERK1* overexpression has been shown to increase embryogenesis in *Arabidopsis* [24]. *SERK* is therefore seen as a ubiquitous marker of plant embryogenic competence. *LEC* and *L1L* genes have also been shown to interfer with early plant embryogenesis [1,15,22,35]. We previously characterized three *SERK* genes in grapevine: *VvSERK1*, *VvSERK2* and *VvSERK3*, homologous to *AtSERK1*, *AtSERK2* and *AtSERK3*, as well as a *L1L* gene homologous to *AtL1L*: *VvL1L*; these four genes were shown to be expressed in a stabilized embryogenic cell line of Chardonnay [62]. During secondary SE, VvSERK1, VvSERK2 and VvL1L expression was detected in embryos at the beginning of the experiment. It was then down-regulated by the 2,4-D treatment. Their transcription remained low in calli at every culture step, was enhanced in secondary embryogenic clusters and further decreased during conversion of embryos into plantlets. The highest transcription level of *VvSERK1*, *VvSERK2* and *VvL1L* was found in embryos at the beginning of the experiment and in secondary embryogenic clusters, suggesting a role for the encoded proteins in grapevine embryogenesis. On the contrary, *VvSERK3* transcription was rather stable in all tissues examined. These results are in accordance with studies on *Arabidopsis*, which showed that *AtSERK1* and *AtSERK2* act redundantly in embryogenesis [8], while *AtSERK3* is rather involved in plant immunity [25] and brassinosteroid signaling [37,49]. *AtL1L* has also been shown to be specifically expressed in early and late stages of embryo development in *Arabidopsis* [35].

VvSERK1, *VvSERK2* and *VvL1L* expression was identical in non-embryogenic calli and embryogenic calli cleared of embryogenic clusters. Embryogenic competence is thought to be limited to a small subset of plant cells. In carrot, they derive from cytoplasmic rich cells originating from the provascular elements of the hypocotyl and proliferating after the 2,4-D treatment [63]. In *Dactylis*, small cytoplasmic rich competent cells are located close to the vascular bundles [65]. In *Arabidopsis* plant lines expressing a SERK1-YFP fusion protein, a prolonged 2,4-D application on seedling hypocotyls was shown to result in the multiplication of procambium cells specifically expressing *AtSERK1*, as visualized by fluorescence monitoring [34]. Likewise, in our culture model, the shift toward 'embryogenic competence' could firstly concern only a very small part of the callus cells, led to the execution of a new genetic program. These cells, at the origin of embryogenic clusters, could express *SERK* and *L1L* genes as an evidence of embryogenic competence acquisition, while the rest of the callus would not. Therefore, a very weak modification of transcription, due to scarce competent cells into the whole callus, would only be detectable by histological localisation methods.

VvSERK1, *VvSERK2* and *VvL1L* expression slightly decreased following the 2,4-D treatment and remained low in calli. In control experiments, when embryos were continuously cultured on the medium A, being not previously treated by 2,4-D, or when 2,4-D was replaced with a weak auxin such as NAA (NaphtaleneAcetic Acid), no secondary embryogenesis was observed, strongly suggesting that 2,4-D is necessary for inducing SE (data not shown). The synthetic auxin 2,4-D is thought to be crucial for the acquisition of embryogenic competence, acting as a shock triggering cell division before the development of somatic embryos [7,16]. However, although some *SERK* genes have been tightly linked to embryogenic competence, their expression is not always observed in freshly 2,4-D-treated tissues. When wheat leaf bases are submitted to a 2,4-D treatment, *SERK* is firstly down-regulated, being strongly expressed only at the time of embryo differentiation [64]. In other respects,

a high expression of *SERK* is not restricted to embryogenic competent cells. In *Medicago truncatula*, a treatment with NAA and BAP was found to stimulate *SERK* expression in two different cell lines, one being embryogenic and the other not [50]. Likewise, in sunflower, *SERK* is up-regulated in immature zygotic embryos driven either to organogenesis or to somatic embryogenesis, suggesting that *SERK* expression alone is rather correlated to an intense mitotic activity [67]. In our culture system, *VvSERK1*, *VvSERK2* and *VvL1L* expression was the highest in cotyledonary embryos as well as in secondary embryogenic clusters, showing that the corresponding proteins could play a role in the early development of grapevine.

3.3. Characterization of a VvPR10.3 gene

Using primers designed for the amplification of *VvPR10.1*, we detected a *VvPR10.3* transcript. We determined that it probably codes for a protein structurally related to the major food and pollen allergens of trees, close to birch Betv1 [3] and very similar to *V. quinquangularis* PR10.3 (ABD78555). Plant PR10 proteins show a very conservative sequence and a similar three-dimensional (3D) structure. They possess a characteristic P-loop (Phosphate-binding motif) sequence and a hydrophobic cavity, making the 'Betv1-signature' [28]. The P-loop is presumed to function as a binding site for ATP or GTP, involved in RNase and anti-microbial activity, and the Y-shaped hydrophobic cavity is able to bind hydrophobic molecules such as fatty acids, flavonoids, cytokinins and brassinosteroids for their probable intracellular transport [17,33,38,40,45]. We found that the nucleotidic sequence of *VvPR10.3* is very homologous to that of *VvPR10.1* and *VvPR10.2*, suggesting that all three genes could originate from a common ancestral sequence, as proposed for *PR10* genes of birch, apple and passiflora [2,18,70].

3.4. Induction of PR genes by the 2,4-D treatment

Transcription of *VvPR3*, *VvPR8* and of the three *VvPR10* genes was detected at the beginning of the experiment, suggesting a basal expression of these genes in embryos, although at a very low rate for *VvPR8*. However, *VvPR1* transcription was not detected at all. As a result of the 2,4-D treatment, *VvPR1*, *VvPR8*, *VvPR10.1* and *VvPR10.3* were strongly up-regulated, while *VvPR10.2* was only weakly stimulated and *VvPR3* not affected at all. The induction of some defence reactions is not surprising; 2,4-D is known to trigger an oxidative stress in treated tissues [60]. In a recent proteomic study, grapevine cultures strongly expressed varied defence related proteins after a treatment with 2,4-D [41]. Although not really understood, the relation between stress and SE is well documented, SE being also obtained, in some plant species, using non-hormonal inducers like wounding, high salt concentration, heavy metal ions or osmotic shock [16,53]. In fact, SE is thought to be an adaptative process to stress, resulting from the extensive cellular reorganization and the genetic reprogramming of treated tissues [16]. However, we found that *VvPR3*, encoding a class I chitinase, was not very responsive to the 2,4-D treatment, contrary to *VvPR8*, encoding a class III chitinase. Likewise, *VvPR10.1* and *VvPR10.3* were strongly up-regulated after the 2,4-D application, but not *VvPR10.2*. These results highlight a differential response to 2,4-D of *PR* genes encoding proteins with a similar enzymatic function. This further suggests that the two genes encoding chitinases (*VvPR3* and *VvPR8*) as well as the three *VvPR10* genes could be regulated in different ways and have separate roles in grapevine.

PR3 (a class I chitinase) is a basic enzyme located in the vacuole, whereas PR8 (a class III chitinase) is acidic and secreted to the apoplast. Plant chitinases can hydrolyse the β-1,4-glycoside bond present in biopolymers of N-acetylglucosamine, such as fungal chitin, as well as plant endogenous AGPs, releasing oligosaccharides that could play a role as signal molecules involved either in plant defence or in plant development [32]. PR3 and PR8 proteins have already been studied in different grapevine cultivars and their induction tested either upon pathogen attack or chemical treatment. According to these previous studies and despite complex patterns of response depending on the treatment and tissue type, general features may be drawn. PR8 is clearly related to grapevine SAR (Systemic Acquired Resistance), being induced by the chemical inducers SA (Salicylic Acid) and INA (2,6-dichloroIsoNicotinic Acid) [4], as well as during the incompatible interaction with *Pseudomonas syringae* [58]. On the contrary, PR3 seems to be constitutive, being expressed in untreated leaves and induced by wounding but not by SA or INA [6, 58]. This class I chitinase would not therefore be a marker of the SAR, although weakly induced under some pathogenic situations [6]. Consistent with these results, we found that *VvPR3* was continuously expressed during the whole culture, whereas *VvPR8* behaved as a typical defence gene, being induced by the 2,4-D stress.

A different expression profile was also found for *PR10* genes: *VvPR10.1* and *VvPR10.3* were strongly induced by the 2,4-D treatment but not *VvPR10.2*. PR10 is a unique group of intracytoplasmic small proteins often constitutively synthesized and thought to play a role in plant development and growth; they are also produced under pathogen attack and in response to abiotic stress [38]. In grapevine, *VvPR10.1* was shown to be up-regulated during an incompatible interaction of the cultivar 'Ugni blanc' with *P. syringae* [57], while several *PR10* genes were expressed in 'Cabernet-Sauvignon' following a fungal infection [19]. A PR10 protein closely related to *V. quinquangularis* PR10.3 was highly expressed in 'Riesling' leaves infected with *Plasmopara viticola* [52]. Moreover, PR10 proteins were produced at a very high level in grapevine plants submitted to a salt stress [31] or treated with the herbicide flumioxazin [5]. In our experiment, *VvPR10.2* was continuously expressed during the entire SE process, being only weakly up-regulated by the 2,4-D treatment. By contrast, *VvPR10.1* and *VvPR10.3* were strongly induced, as expected for defence genes.

Not detected in cotyledonary embryos at the beginning of the protocol, *VvPR1* expression was highly induced by the 2,4-D treatment. PR1 is a dominant group of PR proteins commonly used as markers of the SAR in many plants, despite a limited antifungal activity [69]. We previously showed that *VvPR1* is induced by SA but not under pathogen attack in whole grapevine plants, suggesting that it is not a good marker of grapevine SAR [6]. Other authors have reported a continuous expression of *VvPR1* in cell cultures and in in vitro plantlets but not in greenhouse plants, suggesting induction by a kind of 'in vitro stress' [71]. In the present study, *VvPR1* transcription was detected in calli after the 2,4-D treatment, but neither in embryos nor in calli at the last culture steps, showing that it was not a response to an 'in vitro stress'. Although both located on the chromosome 3, the *VvPR1* gene we monitored (AJ536326) and the one studied by Wielgoss and Kortekamp [71] (AJ003113) correspond to Genoscope annotations, respectively GSVIVT00038582001 and GSVIVT00038576001, bearing sufficient differences to be considered as separate genes. Three different PR1 proteins have been previously found to be induced by chemicals, wounding, as well as upon pathogen attack, therefore suggesting that several different *VvPR1* genes could be active in grapevine [56].

3.5. Putative role of PR proteins in SE

After the first culture step on a medium containing 2,4-D, all *PR* genes were expressed at a high rate, as a probable sign of oxidative stress. *VvPR1* and *VvPR8* were only transiently expressed after the

2,4-D treatment. Therefore, their role in the embryogenic process remains unclear. By contrast, *VvPR10.1* and *VvPR10.3* were expressed long after the 2,4-D removal, together with *VvPR3* and *VvPR10.2*. The class I chitinase encoded by *VvPR3* as well as the three PR10 proteins could therefore play a role during SE induction. In plant cell cultures, many different proteins are secreted; some of them are chitinase-like proteins, that have been shown to stimulate SE, for instance in grapevine, chicory, carrot, *Coffea arabica* and Norway spruce [13,27,59,68,72]. Plant chitinases may act by hydrolysing endogenous proteoglycans attached to the membranes or cell walls, such as AGPs, producing signal molecules that could have a function in SE [32]. A class IV chitinase, able to release oligosaccharides from cell AGPs, was shown to rescue the carrot *tsl1* mutant, impaired in embryo development [11,68]. Likewise, PR10 proteins would be able to condition cell cultures for SE, probably by interfering with the intracellular transport of hormones [17,33,38,40,45]. The overexpression of a *PR10* gene in a transgenic pea led to an elevated cytokinin content and a decreased abscisic acid level, suggesting that changes in the PR10 content could be relevant for phytohormone regulation [66]. In *M. truncatula*, some PR10 have been shown to be the most abundant proteins induced by a 2,4-D treatment leading to SE [12,30]. Moreover, a *PR10* gene was highly expressed during SE induction, starting from alfalfa protoplasts: a first pulse of transcription followed 2,4-D application and was interpreted as a stress response; the second pulse occurred later, at the time of embryo differentiation, suggesting an additional role of the encoded protein in morphogenesis [14]. However, PR10 expression was also recorded, during SE induction, in a highly (2HA) as well as in a poorly (Jemalong) embryogenic cell line of *M. truncatula*, suggesting that PR10 proteins could contribute to the conditioning of cultures but would not act alone for inducing the shift to embryogenesis [30]. In our SE model, all *PR* genes were expressed in the same way in embryogenic and non-embryogenic calluses, showing that both callus types had similarly been conditioned during the culture. A change was only observed in embryogenic clusters: *VvPR10.1* and *VvPR10.3* were down-regulated, while *VvSERK1*, *VvSERK2* and *VvL1L* were up-regulated, as the sign of the starting up of a new genetic program. In a proteomic analysis of SE, in the grapevine cultivar 'Thompson seedless', two forms of a PR10 protein were also down-regulated in young embryogenic clusters, whereas highly expressed in the underlying callus [41]. We found that *VvPR10.1* and *VvPR10.3* behaved differently from the other *PR* genes in our culture model. They were up-regulated by the 2,4-D treatment, being expressed for a long time in calli before the differentiation of secondary embryos. This suggests a possible specific role of the corresponding proteins in conditioning grapevine cultures for SE.

4. Materials and methods

4.1. Induction of secondary embryogenesis

Embryogenic cultures of *V. vinifera* cultivar 'Chardonnay' were previously obtained from nodal explants and maintained as described [39]. Culture media contained Murashige and Skoog [47] half strength major salts, micro-salts and vitamins, sucrose, and a mixture of growth regulators added before autoclaving. The pH level was adjusted to 5.8 by adding 1N NaOH. Media were solidified by adding 0.7% Bacto-Agar and sterilized by autoclaving at 115° C for 30 min. Culture incubation was performed at 25° C, in the dark. To induce secondary embryogenesis, intact and well-formed embryos were picked-up from embryogenic cultures. To assess secondary embryogenesis efficiency, four independant experiments were carried out, with embryos either at the torpedo-stage (experiments A and B) or at the cotyledonary-stage (experiments C and D). For molecular analyses, secondary embryogenesis was induced from cotyledonary embryos. Embryos were separately plated onto the medium E96 containing 60 g L^{-1} sucrose, 9 μM BAP (6-BenzylAminoPurine) and 4.5 μM 2,4-D. A set of 14 embryos was placed in each Petri dish (60 mm of diameter). Cultures were incubated for 3 weeks to produce calli. Subculture was performed by transferring the calli onto the medium A containing 60 g L^{-1} sucrose, 2.5 g L^{-1} activated charcoal, 20 μM IAA (Indole-3-Acetic Acid), 10 μM NOA (2-NaphtOxyacetic Acid) and 1 μM BAP. Transfer onto a fresh medium occurred every 4 weeks until clusters of secondary embryos appeared on calli. For germination of secondary embryos, embryos at the cotyledonary-stage were transferred, 3 per Petri dish, onto a medium containing 25 g L^{-1} sucrose and 0.2 μM BAP, and incubated for 4–8 days.

4.2. Collecting of tissues for molecular analysis

Tissues were collected at different steps of the secondary embryogenesis process: cotyledonary embryos at the beginning of the experiment (E); calli developed from embryos on the medium E96 (C1); calli after a one week subculture on the medium A (C2); non-embryogenic calli (NEC) and embryogenic calli cleared of embryogenic clusters (EC) after 3 subcultures on the medium A; and clusters of secondary embryos (CSE). Germinated embryos (GE) bearing a visible apical meristem, two developed cotyledons and a differentiated root were also collected. Samples were composed of pools of embryos or calli, carefully harvested and freezed in liquid N_2 for preserving at −80° C.

4.3. RNA extraction

Total RNA was extracted using the 'Nucleospin RNA plant kit' (Macherey-Nagel, Germany) following the manufacturer's instructions with slight modifications: samples corresponding to around 100 mg fresh tissue were ground in liquid nitrogen and transferred into 500 μL buffer RA1 after the addition of 2.5% PVP 40, 1% β-mercaptoethanol and 2% sarkosyl (chemicals purchased from Sigma–Aldrich, Germany). Incubation was performed at 70° C for 10 min. Total RNA was quantified by Optical Density measurement. The absence of contaminating genomic DNA was checked by a PCR test with the total RNA extract as a template and *VvActin* primers.

4.4. Semi-quantitative RT-PCR analysis

RT-PCR amplifications were performed using the 'Superscript II First-Strand Synthesis system' (Invitrogen, USA), starting from 1 μg total RNA. cDNA synthesis was performed at 50° C for 60 min and then 70° C for 15 min. Amplifications were carried out according to the following program: 30 s at 95° C (denaturation), 30 s at 55° C (annealing) and 1 min at 72° C (extension), with a final elongation step at 72° C for 7 min. Prior to expression analysis, *VvActin* was used as a standard to determine the exponential zone of amplification in a PCR test of 28 cycles with 5 serial half-dilutions of each cDNA sample. The appropriate dilutions were then used in PCR amplifications of 32 cycles. Three independent PCR reactions were performed for each gene, giving rise to very reproducible results. We used primer sequences corresponding to the different genes and deduced from the nucleotide sequences available in the banks of the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov) and the Genoscope (http://www.genoscope.cns.fr): forward 5′-TGC TAT CCT TCG TCT TGA CCT TG-3′ and reverse 5′-GGA CTT CTG GAC AAC GGA ATC TC-3′ for *VvActin* (AF369525), forward 5′-GCC TAC GCC CAG AAC TAT GCT AAC-3′ and reverse 5′-CGA ACC ACC ACC CAT TGT TG-3′ for *VvPR1* (AJ536326),

forward 5′-TGT AGC AAC TAA GAC GTT TAT TCA-3′ and reverse 5′-GCC AGC TTA TTA CCA GGT CCA T-3′ for *VvPR3* (AJ291505), forward 5′-ATC ATC GTC TCG GCC ATT AG-3′ and reverse 5′-AGA GCA GTG CCC ATG AAC TT-3′ for *VvPR8* (AJ291507), forward 5′-CAA CCA CAG TGT AGC TGA ATG TGA AG-3′ and reverse 5′-CGA GAG TGA GGT CAC TTC CTC G-3′ for *VvPR10.1* (AJ291705), forward 5′-CGA TCA CAG TGT AGC GGA ATG AGA AT-3′ and reverse 5′-AAG CTA TCA AGT GCG TGG AAG TCA TT-3′ for *VvPR10.2* (AJ291704), forward 5′-TGT TTA AGA GAA CGC CC-3′ and reverse 5′-GAG TTG GAG TGA GGA GCG-3′ for *VvSERK1* (GSVIVT00001777001), forward 5′-GCC TAA GAG AAC GTC CAC CAT-3′ and reverse 5′-ACC TCC TGA CGG ACG ACC-3′ for *VvSERK2* (GSVIVT00019412001), forward 5′-GTT TGA GAG AGC GAG CTG ATG-3′ and reverse 5′-GGT GGG GGA TAT GGT TGT A-3′ for *VvSERK3* (GSVIVT00009544001), forward 5′-CAG CCA TGG AAG ACA CTG AAT GCA-3′ and reverse 5′-CCT TCA AGC TCG CGG TAG TGG TGG-3′ for *VvL1L* (GSVIVT00010958001). The resulting amplification products were analyzed by electrophoresis on a 1.5 % agarose gel and stained with ethidium bromide. Sequencing of the PCR products was performed in order to determine the amplified sequences (Genoscreen, France). Nucleic acid homology search was performed through the databases of the NCBI, the Genoscope, the Istituto Agrario di San Michele all'Adige (IASMA) (http://genomics.research.iasma.it) and the Dana-Farber Harvard Cancer Institute (DFCI) (http://compbio.dfci.harvard.edu). Sequence alignments were performed using the program ClustalW of the European Bioinformatics Institute (EBI) (http://ebi.ac.uk).

4.5. Rapid amplification of cDNA ends (RACE)

Further determination of a cDNA sequence amplified with the *VvPR10.1* primers was performed using a RACE-PCR protocol with oligonucleotides allowing to extend it in both directions (toward the 5′ and the 3′ regions), and leading to slightly overlapping sequences. RACE samples were prepared from the poly(A)$^+$ RNA using the "5′/3′ RACE Kit" (Roche, France), following the manufacturer's instructions. 5′RACE required three antisens specific primers: 5′-GGG ACC TCC CTG TCC TTG TAG-3′ (SP1); 5′-CTT GAG GCC TTA TCT TGG G-3′ (SP2) ; 5′-CGA GGA AGT GAC CTC ACT CTC G-3′ (SP3), whereas 3′RACE required only one specific forward primer: 5′-CTA CAA GGA CAG GGA GGT CCC-3′ (SP5).

References

[1] L. Alemanno, M. Devic, N. Niemenak, C. Sanier, J. Guilleminot, M. Rio, J.-L. Verdeil, P. Montoro, Characterization of *leafy cotyledon1-like* during embryogenesis in *Theobroma cacao* L, Planta 227 (2008) 853–866.

[2] L. Beuning, J.H. Bowen, H.A. Persson, D. Barraclough, S. Bulley, E.A. MacRae, Characterisation of Mal d 1-related genes in *Malus*, Plant Mol. Biol. 55 (2004) 369–388.

[3] H. Breiteneder, K. Pettenburger, A. Bito, R. Valenta, D. Kraft, H. Rumpold, O. Scheiner, M. Breitenbach, The gene coding for the major birch pollen allergen, Bet v. 1, is highly homologous to a pea disease resistance response gene, EMBO J. 8 (1989) 1935–1938.

[4] G. Busam, H.H. Kassemeyer, U. Matern, Differential expression of chitinases in *Vitis vinifera* L. responding to systemic acquired resistance activators or fungal challenge, Plant Physiol. 115 (1997) 1029–1038.

[5] A.J. Castro, C. Carapito, N. Zorn, C. Magné, E. Leize, A. Van Dorsselaer, C. Clément, Proteomic analysis of grapevine (*Vitis vinifera* L.) tissues subjected to herbicide stress, J. Exp. Bot. 56 (2005) 2783–2795.

[6] J. Chong, G. Le Henanff, C. Bertsch, B. Walter, Identification, expression analysis and characterization of defense signaling genes in *Vitis vinifera*, Plant Physiol. Biochem. 46 (2008) 469–481.

[7] A. Chugh, P. Khurana, Gene expression during somatic embryogenesis-recent advances, Curr. Sci. 83 (2002) 715–730.

[8] J. Colcombet, A. Boisson-Dernier, R. Ros-Palau, C.E. Vera, J.I. Schroeder, *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are essential for tapetum development and microscope maturation, Plant Cell 17 (2005) 3350–3361.

[9] P. Croce, C. Vannini, M. Bracale, An easy and convenient method for maintenance of embryogenic cultures of *Vitis vinifera*, Vitis 44 (2005) 197–198.

[10] D.K. Das, M.K. Reddy, K.C. Upadhyaya, S.K. Sopory, An efficient leaf-disc culture method for the regeneration via somatic embryogenesis and transformation of grape (*Vitis vinifera* L.), Plant Cell Rep. 20 (2002) 999–1005.

[11] A.J. De Jong, J. Cordewener, F. Lo Schiavo, M. Terzi, J. Vandekerckhove, A. Van Kammen, S.C. De Vries, A carrot somatic embryo mutant is rescued by chitinase, Plant Cell 4 (1992) 425–433.

[12] F. De Jong, U. Mathesius, N. Imin, B.G. Rolfe, A proteome study of the proliferation of cultured *Medicago truncatula* protoplasts, Proteomics 7 (2007) 722–736.

[13] A. Deloire, E. Kraeva, M.C. Mauro, E. Bonnet, J. Bierne, Immunodetection of chitinase-like and β-1,3-glucanase-like proteins secreted in vitro by embryogenic and non-embryogenic cells of grapevines, Vitis 1 (1997) 51–52.

[14] M. Domoki, J. Györgyey, J. Biro, T.P. Pasternak, A. Zvara, S. Bottka, L.G. Puskas, D. Dudits, A. Fehér, Identification and characterization of genes associated with the induction of embryogenic competence in leaf-protoplast-derived alafalfa cells, Biochim. Biophys. Acta 1759 (2006) 543–551.

[15] M. Fambrini, C. Durante, G. Cionini, C. Geri, L. Giorgetti, V. Michelotti, M. Salvini, C. Pugliesi, Characterization of *LEAFY COTYLEDON1-LIKE* gene in *Helianthus annuus* and its relationship with zygotic and somatic embryogenesis, Dev. Genes Evol. 216 (2006) 253–264.

[16] A. Fehér, T.P. Pasternak, D. Dudits, Transition of somatic plant cells to an embryogenic state, Plant Cell Tissue Organ Cult. 74 (2003) 201–228.

[17] H. Fernandes, O. Pasternak, G. Bujacz, A. Bujacz, M. Sikorski, M. Jaskolski, *Lupinus luteus* Pathogenesis-related protein as a reservoir for cytokinin, J. Mol. Biol. 378 (2008) 1040–1051.

[18] C. Finkler, C. Giacomet, V.C. Muschner, F.M. Salzano, L.B. Freitas, Molecular investigations of pathogenesis-related Betv1 homologues in *Passiflora* (Passifloraceae), Genetica 124 (2005) 117–125.

[19] R.W.M. Fung, M. Gonzalo, C. Fekete, L.G. Kovacs, Y. He, E. Marsh, L.M. McIntyre, D.P. Schachtman, W. Qiu, Powdery mildew induces defense-oriented reprogramming of the transcriptome in a susceptible but not in a resistant grapevine, Plant Physiol. 146 (2008) 236–249.

[20] M.D. Gaj, Direct somatic embryogenesis as a rapid and efficient system for *in vitro* regeneration of *Arabidopsis thaliana* (L.) Heynh, Plant Cell Tissue Organ Cult. 64 (2001) 39–46.

[21] M.D. Gaj, Factors influencing somatic embryogenesis induction and plant regeneration with particular reference to *Arabidopsis thaliana* (L.) Heynh, Plant Growth Regul. 43 (2004) 27–47.

[22] M.D. Gaj, S. Zhang, J.J. Harada, P.G. Lemaux, Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*, Planta 222 (2005) 977–988.

[23] G. Gambino, P. Ruffa, R. Vallania, I. Gribaudo, Somatic embryogenesis from whole flowers, anthers and ovaries of grapevine (*Vitis* spp.), Plant Cell Tissue Organ Cult. 90 (2007) 79–83.

[24] V. Hecht, J.P. Vielle-Calzada, M.V. Hartog, E.D.L. Schmidt, K. Boutilier, U. Grossniklaus, S.C. de Vries, The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture, Plant Physiol. 127 (2001) 803–816.

[25] A. Heese, D.R. Hann, S. Gimenez-Ibanez, A.M.E. Jones, K. He, J. Li, J.I. Schroeder, S.C. Peck, J.P. Rathjen, The receptor-like kinase *SERK3/BAK1* is a central regulator of innate immunity in plants, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 12217–12222.

[26] S. Helleboid, G. Bauw, L. Belingheri, J. Vasseur, J.L. Hilbert, Extracellular β-1,3-glucanases are induced during early somatic embryogenesis in *Cichorium*, Planta 205 (1998) 56–63.

[27] S. Helleboid, T. Hendricks, G. Bauw, D. Inzé, J. Vasseur, J.L. Hilbert, Three major somatic embryogenesis related proteins in *Cichorium* identified as PR proteins, J. Exp. Bot. 51 (2000) 1189–1200.

[28] K. Hoffmann-Sommergruber, M. Vanek-Krebitz, C. Radauer, J. Wen, F. Ferreira, O. Scheiner, H. Breiteneder, Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share introns positions, Gene 197 (1997) 91–100.

[29] H. Hu, L. Xiong, Y. Yang, Rice *SERK1* gene positively regulates somatic embryogenesis of cultured cells and host defense response against fungal infection, Planta 222 (2005) 107–117.

[30] N. Imin, M. Nizamidin, D. Daniher, K.E. Nolan, R.J. Rose, B.G. Rolfe, Proteomic analysis of somatic embryogenesis in *Medicago truncatula*. Explant cultures grown under 6-benzylaminopurine and 1-naphtaleneacetic acid treatments, Plant Physiol. 137 (2005) 1250–1260.

[31] N. Jellouli, H. Ben Jouira, H. Skouri, A. Ghorbel, A. Gourgouri, A. Mliki, Proteomic analysis of tunisian grapevine cultivar Razegui under salt stress, J. Plant Physiol. 165 (2008) 471–481.

[32] A. Kasprzewska, Plant chitinases–regulation and function, Cell. Mol. Biol. Lett. 8 (2003) 809–824.

[33] K.M. Koistinen, P. Soininen, T.A. Venäläinen, J. Häyrinen, R. Laatikainen, M. Peräkylä, A.I. Tervahauta, S.O. Kärenlampi, Birch PR-10c interacts with several biologically important ligands, Phytochemistry 66 (2005) 2524–2533.

[34] M.A.C.J. Kwaaitaal, S.C. de Vries, The *SERK1* gene is expressed in procambium and in immature vascular cells, J. Exp. Bot. 58 (2007) 2887–2896.

[35] R.W. Kwong, A.Q. Bui, H. Lee, L.W. Kwong, R.L. Fischer, R.B. Goldberg, J.J. Harada, LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development, Plant Cell 15 (2003) 5–18.

[36] S. Legrand, T. Hendricks, J.L. Hilbert, M.C. Quillet, Characterization of expressed sequence tags obtained by SSH during somatic embryogenesis in *Cichorium intybus* L, BMC Plant Biol. 7 (2007) 27.

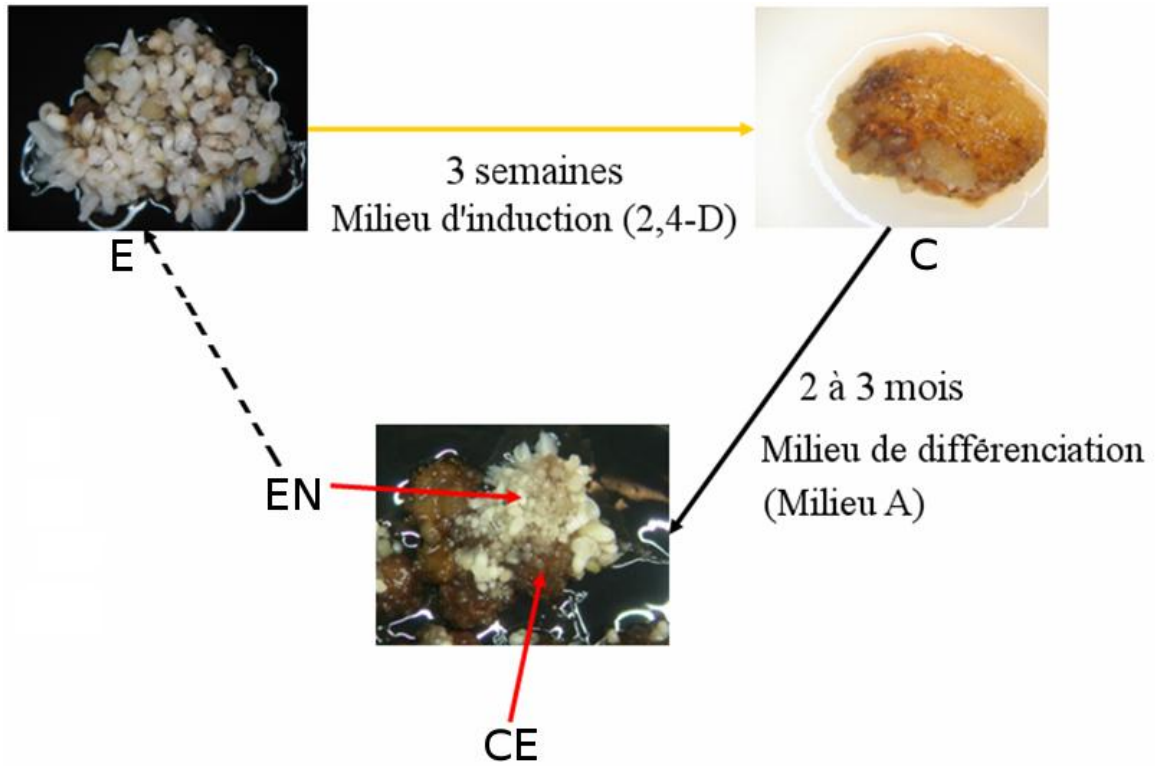


Figure 10 : Tissus analysés aux étapes-clé du cycle d'embryogenèse somatique secondaire de *Vitis vinifera* cultivar Chardonnay.

E = embryons en début de cycle ; C = cals obtenus après un traitement au 2,4-D (3 semaines) ; CE = cals embryogènes sans les embryons néoformés ; EN = masses d'embryons secondaires néoformés en fin de cycle.

III. Étude complémentaire de l'expression des gènes *VvPR10* lors de l'embryogenèse somatique – Recherche de marqueurs de l'embryogenèse somatique de la vigne

A la suite de l'identification de l'intégralité des gènes *PR10* de *Vitis vinifera*, nous avons entrepris une analyse complémentaire de leur expression pendant le processus d'embryogenèse somatique dans le cultivar Chardonnay, d'abord par RT-PCR semi-quantitative, puis par RT-PCR quantitative. Du fait de l'importante similarité entre les séquence *-a*, *-b* et *-c* de *VvPR10.1* et *VvPR10.3*, il n'a pas été possible de concevoir des couples d'amorces spécifiques pour chacun de ces gènes. Un représentant de *VvPR10.1* (*VvPR10.1-a*), un représentant de *VvPR10.3* (*VvPR10.3-a*) ainsi que tous les autres gènes *VvPR10* (*VvPR10.2* et *VvPR10.4* à *VvPR10.10*) ont donc tout d'abord été analysés par RT-PCR semi-quantitative. Sur la base des résultats obtenus, seulement 7 gènes, dont le niveau d'expression était suffisamment élevé, ont été retenus pour l'analyse en RT-PCR quantitative, soit : *VvPR10.1-a*, *VvPR10.2*, *VvPR10.3-a*, *VvPR10.5*, *VvPR10.7*, *VvPR10.8* et *VvPR10.9*.

A. Étude de l'expression de *VvPR10.1-a* à *VvPR10.10* par RT-PCR semi-quantitative

Dans un premier temps, le niveau d'expression des gènes *VvPR10.4*, *VvPR10.5*, *VvPR10.6*, *VvPR10.7*, *VvPR10.8*, *VvPR10.9 et VvPR10.10* a été évalué par RT-PCR semi-quantitative, au moment des deux étapes clés d'entrée et de sortie du cycle d'embryogenèse somatique secondaire, comme cela avait déjà été réalisé pour *VvPR10.1-a*, *VvPR10.2* et *VvPR10.3-a*. Ensuite, l'expression de l'ensemble des dix gènes a été estimée dans 3 types d'explants (embryons, nœuds et feuilles) issus de culture *in vitro*, avant et après un traitement par le 2,4-D, supposé inducteur de l'embryogenèse. Lors de chacune de ces expériences, l'expression du gène codant l'actine, considéré comme stable, a été utilisée comme standard interne.

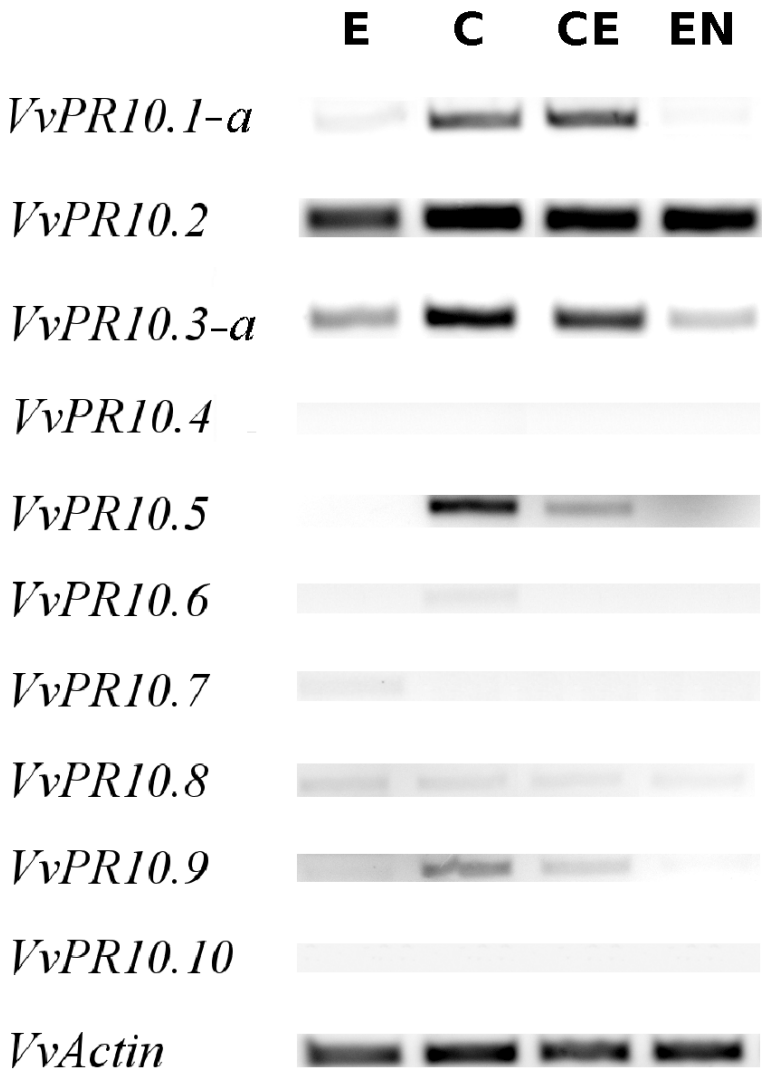


Figure 11 : Expression en RT-PCR semi-quantitative des gènes *VvPR10.1-a* à *VvPR10.10* durant le cycle d'embryogenèse somatique secondaire.

E = embryons en début de cycle ; C = cals obtenus après un traitement au 2,4-D (3 semaines) ; CE = cals embryogènes sans les embryons néoformés ; EN = masses d'embryons secondaires néoformés en fin de cycle.

1. Étude par RT-PCR semi-quantitative de l'expression des gènes *PR10* à différentes étapes du cycle d'embryogenèse somatique secondaire

L'objectif était d'évaluer les différences d'expression des gènes codant des PR10 durant le cycle d'embryogenèse somatique secondaire. A cet effet, les échantillons prélevés correspondent à 3 stades précis du cycle (Figure 10) :

- les embryons en début de cycle (E) qui sont les tissus mis en culture,
- les cals (C) obtenus après un traitement de 3 semaines au 2,4-D, suite à une dédifférenciation des tissus de l'embryon,
- les cals embryogènes (CE) portant des masses d'embryons néoformés (EN), qui correspondent aux cals formés précédemment ayant évolué pendant plusieurs semaines sur un milieu dit de dédifférenciation des embryons.

Les points choisis permettent de différencier les niveaux d'expression des gènes *PR10* après le traitement inducteur de l'embryogenèse (passage du point E au point C), et au moment de la différenciation des embryons secondaires (comparaison des points CE et EN).

La figure 11 présente les résultats obtenus.

A l'exception de *VvPR10.4* et *VvPR10.10*, chaque gène est exprimé à au moins une des étapes observées. *VvPR10.1-a*, *VvPR10.2* et *VvPR10.3-a* montrent des signaux d'expression dont l'intensité est plus élevée que celle des autres gènes. A chaque point de prélèvement, les transcrits de *VvPR10.2* et *VvPR10.8* sont détectés et leurs niveaux d'expression semblent être relativement constants. Les variations des profils d'expression de *VvPR10.1-a*, *VvPR10.3-a*, *VvPR10.5* et *VvPR10.9* sont comparables : augmentation de l'expression lors du traitement 2,4-D (induction de l'embryogenèse) et un niveau d'expression plus élevé dans les cals embryogènes que dans les embryons néoformés qu'ils portent. Il est à noter que les niveaux d'expression dans ces embryons néoformés sont proches de ceux observés dans les embryons en début de cycle. Les transcrits de *VvPR10.6* et *VvPR10.7* sont très faiblement détectés, et seulement dans les cals après traitement 2,4-D pour le premier et dans les embryons en début de cycle pour le second de ces gènes.

De cette analyse, il apparaît que les niveaux d'expression des gènes *VvPR10* sont globalement plus faibles dans les embryons que dans les cals obtenus après traitement au 2,4-D ou les cals embryogènes et que au moins 5 gènes sont induits lors du traitement par

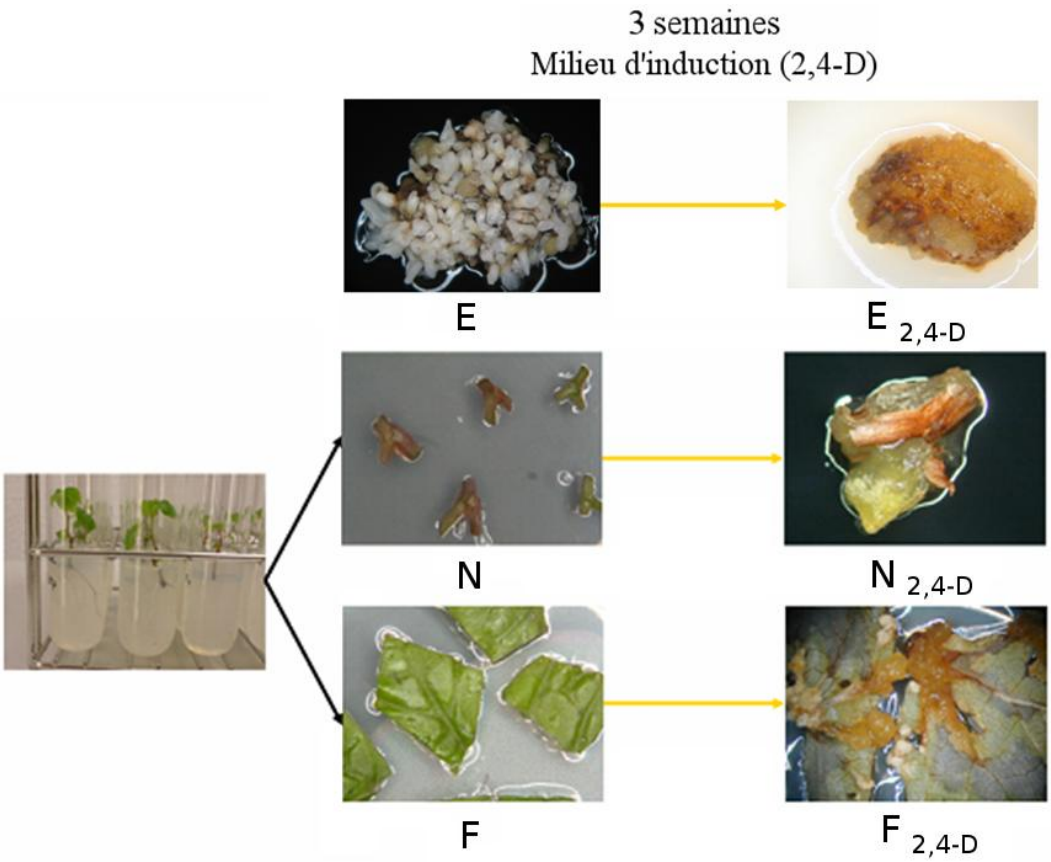


Figure 12 : Tissus analysés lors de l'étude de l'effet du 2,4-D sur l'expression des gènes *VvPR10*.

E = embryons non traités ; $E_{2,4\text{-}D}$ = cals obtenus à partir d'embryons traités au 2,4-D ; N = nœuds non traités ; $N_{2,4\text{-}D}$ = cals obtenus à partir de nœuds traités au 2,4-D ; F = feuilles non traitées ; $F_{2,4\text{-}D}$ = cals obtenus à partir de feuilles traitées au 2,4-D ;

Le 2,4-D : *VvPR10.1-a* et *VvPR10.3-a,* comme précédemment publié, ainsi que *VvPR10.5, VvPR10.6* et *VvPR10.9*.

2. Étude par RT-PCR semi-quantitative de l'expression des gènes *PR10* dans des tissus présentant différentes capacités embryogènes lors du traitement par le 2,4-D

L'objectif était de mesurer l'effet du 2,4-D sur l'expression des gènes *PR10* dans différents tissus ; le 2,4-D est connu pour être un médiateur de l'acquisition de la capacité embryogène par les tissus. Les tissus que nous avons choisis (embryons, nœuds et feuilles) présentent des capacités embryogéniques différentes, antérieurement déterminées au laboratoire, dans les conditions expérimentales décrites plus haut. Les embryons présentent la plus forte capacité embryogénique (de l'ordre de 60-80%), les nœuds ont une capacité embryogénique faible (de l'ordre de 6%), tandis que les feuilles de plantules de Chardonnay ne nous ont jamais permis d'obtenir d'embryons (Figure 12). Les niveaux d'expression des gènes *PR10* dans des échantillons de ces différents tissus avant (E, F et N) et après un traitement de 3 semaines au 2,4-D ($E_{2,4\text{-}D}$, $F_{2,4\text{-}D}$ et $N_{2,4\text{-}D}$,) sont comparés. A ce stade, aucun embryon n'est encore visible sur les cals. Les résultats sont présentés dans la figure 13.

A l'exception de *VvPR10.10,* tous les gènes sont exprimés dans au moins un des échantillons analysés, bien que les signaux correspondant à *VvPR10.4, VvPR10.5, VvPR10.6, VvPR10.8* et *VvPR10.9* soient toujours très faibles et que ceux de *VvPR10.1-a, VvPR10.2* et *VvPR10.3-a* soient à nouveau les plus forts. Les transcrits de *VvPR10.2* sont exprimés dans tous les échantillons étudiés et de manière assez stable, mais on observe une induction par le traitement 2,4-D dans les nœuds traités au 2,4-D. Les profils d'expression des gènes *VvPR10.1-a* et *VvPR10.3-a* sont assez similaires, leur expression étant stimulée par le traitement au 2,4-D, bien que plus fortement pour *VvPR10.1-a* que pour *VvPR10.3-a*. Avec un niveau plus faible, *VvPR10.9* a un profil d'expression comparable à celui des deux derniers gènes. *VvPR10.8* est également exprimé dans chaque tissu mais sans que le traitement n'affecte le niveau d'expression. L'expression des gènes *VvPR10.5* et *VvPR10.6* est observable uniquement dans les embryons traités au 2,4-D et, pour *VvPR10.5,* seulement dans les feuilles traitées au 2,4-D. Enfin, les transcrits de *VvPR10.7* sont détectés dans les nœuds et les feuilles non traités, ainsi que très faiblement dans les

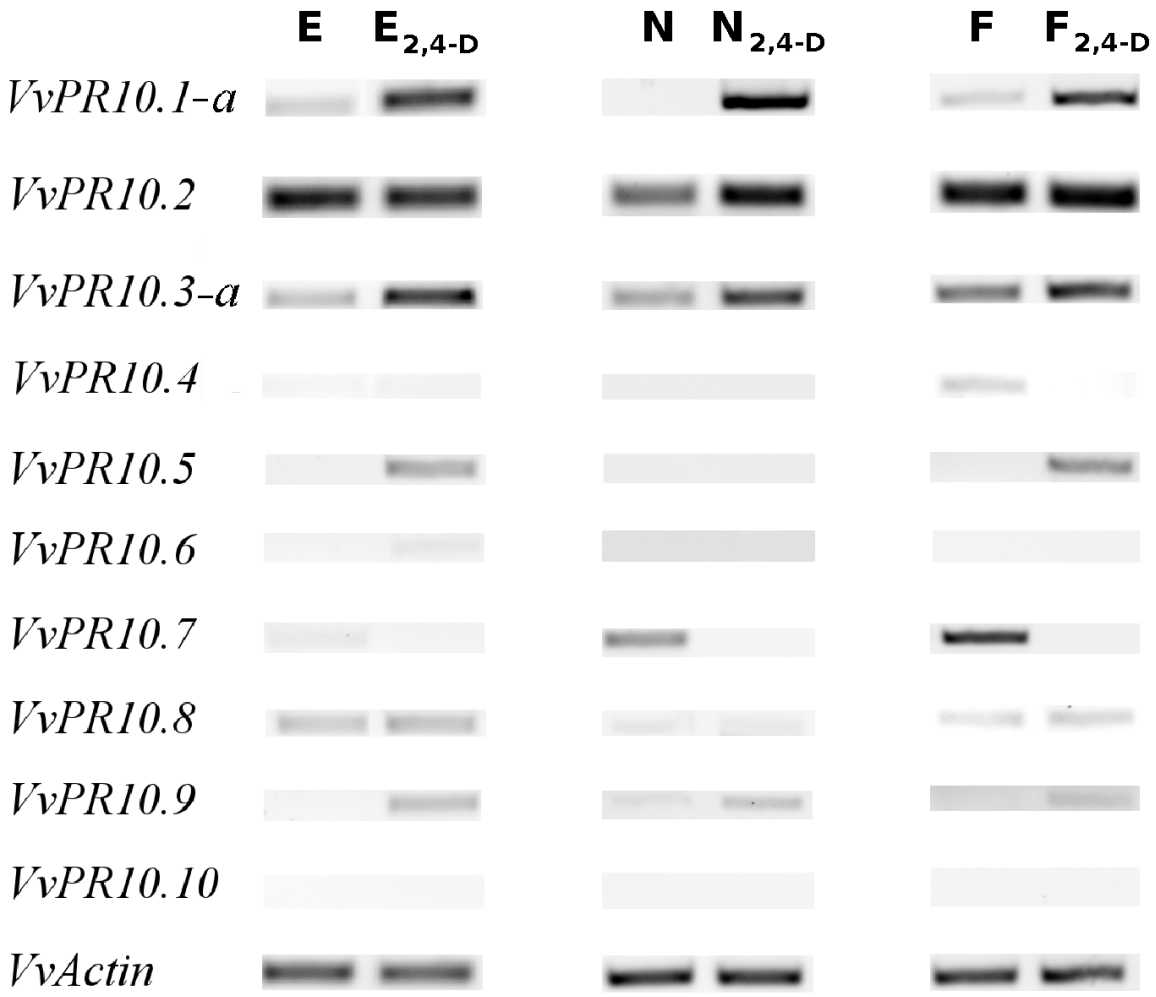


Figure 13 : Expression en RT-PCR semi-quantitative des gènes *VvPR10.1-a* à *VvPR10.10* dans des embryons, nœuds et feuilles avant et après traitement par le 2,4-D.

E = embryons non traités ; $E_{2,4\text{-}D}$ = cals obtenus à partir d'embryons traités au 2,4-D ; N = nœuds non traités ; $N_{2,4\text{-}D}$ = cals obtenus à partir de nœuds traités au 2,4-D ; F = feuilles non traitées ; $E_{2,4\text{-}D}$ = cals obtenus à partir de feuilles traitées au 2,4-D.

embryons, tandis que l'expression de *VvPR10.4* n'est visible que dans les feuilles non traitées. Ces deux gènes présentent donc des profils d'expression totalement différents des autres *VvPR10* puisque leur expression est réprimée lors du traitement 2,4-D et disparaît.

Globalement, le traitement par le 2,4-D a pour effet d'induire l'expression d'une majorité de ces gènes : *VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.6* et *VvPR10.9*, mais de manière variable en fonction du tissu d'origine. Ce traitement semble toutefois ne pas avoir d'effet sur l'expression de *VvPR10.8* et au contraire réprimer l'expression de *VvPR10.4* et *VvPR10.7*.

B. Étude de l'expression des gènes *VvPR10* par RT-PCR quantitative

La technique de RT-PCR quantitative (RT-qPCR) nous a permis d'étudier plus finement les variations des niveaux d'expression des gènes *PR10* lors de l'embryogenèse somatique et de comparer l'expression des différents gènes entre eux. Cependant, sept gènes seulement ont été choisis pour cette analyse : *VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.7, VvPR10.8* et *VvPR10.9*. Les deux gènes dont le niveau d'expression avait été détecté comme très faible lors des expériences de RT-PCR semi-quantitative, *VvPR10.4* et *VvPR10.6*, ainsi que *VvPR10.10*, semblant non exprimé, n'ont pas été étudiés ici.

1. Validation des conditions de RT-qPCR

Avant la quantification, il a été nécessaire de valider les résultats obtenus lors de chaque expérience grâce, d'une part, à une courbe de fusion pour chaque réaction et, d'autre part, à une droite étalon réalisée pour chaque gène.

Les courbes de fusion générées à la fin de chaque réaction ont permis de s'assurer de l'absence de sous-produit d'amplification. Les courbes de fusions obtenues (Figure 14) montrent que les transcrits d'un seul gène ont été amplifiés lors de chaque réaction.

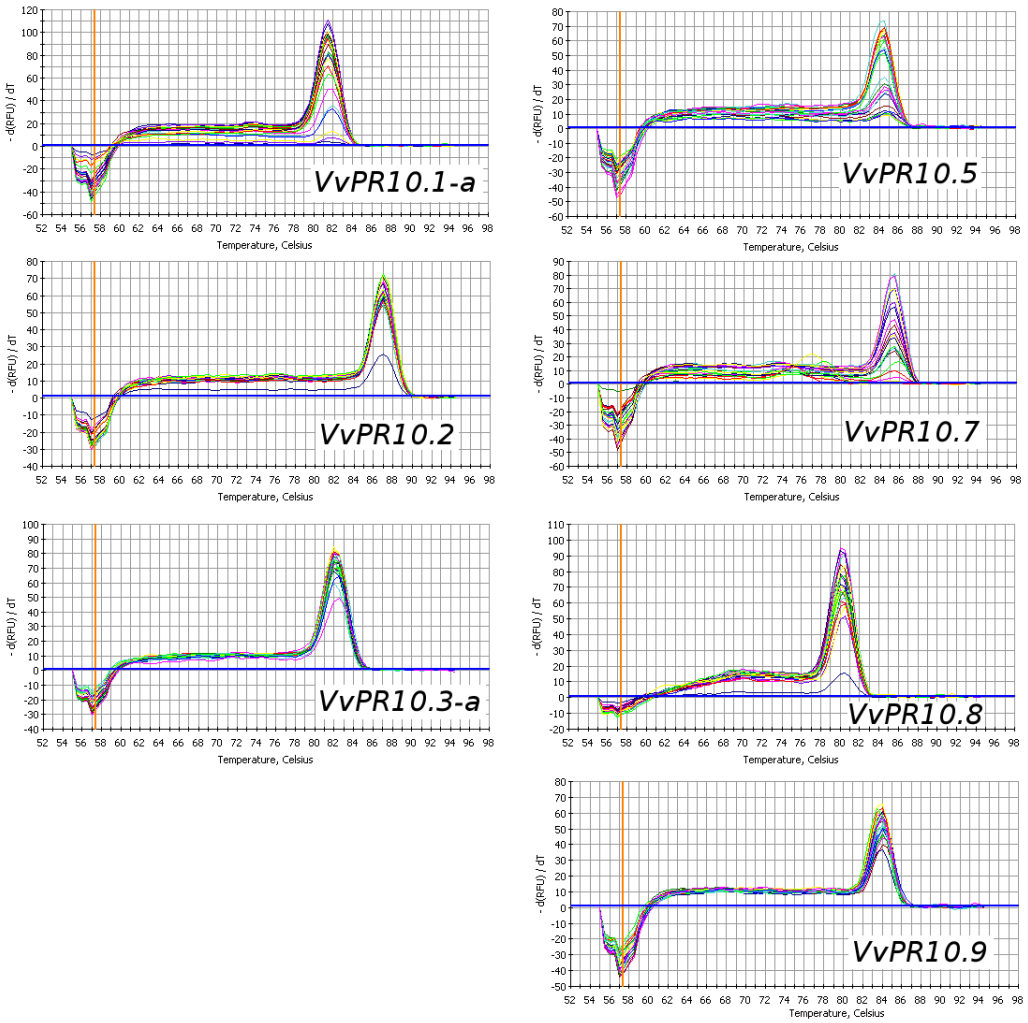


Figure 14 : Courbes de fusion réalisées en fin de réaction de RT-qPCR pour *VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.7, VvPR10.8 et VvPR10.9.*

Les courbes présentent l'opposé de la dérivée de la fluorescence en fonction de la température. Pour chaque gène, 24 courbes de fusions ont été réalisées, correspondant aux 8 échantillons testés en triplicats.

Les droites étalon ont été réalisées. Elle permettent d'une part de comparer les efficacités d'amplification lors de chaque expérience afin d'estimer leur reproductibilité (Tableau 15). Elles sont d'autre part utilisées pour définir les bornes entre lesquelles les Cq seront validés (« linear dynamic range », LDR). Pour un gène donné, les échantillons dont les valeurs de Cq mesurées sont hors de la droite étalon n'ont pas été analysés par la suite (Tableau 16). Par exemple pour *VvPR10.8*, seuls les échantillons de feuilles (avant et après traitement par le 2,4-D) sont quantifiables. Dans tous les cas, les résultats retirés présentaient des valeurs de Cq plus faibles que la limite inférieure de la droite étalon (correspondant à 100 copies pour 12,5 ng d'ADNc).

2. Étude par RT-qPCR de l'expression des gènes *PR10* à différentes étapes du cycle d'embryogenèse somatique secondaire

La première étape des analyses a consisté à étudier l'expression des sept gènes choisis lors du cycle d'embryogenèse somatique secondaire. Des échantillons prélevés aux mêmes étapes clés que celles étudiées en RT-PCR semi-quantitative ont été analysés. La figure 15 représente le niveau d'expression relatif des différents gènes étudiés. La moyenne du niveau d'expression des gènes de référence (*VvActin* et *VvEF1*α) est fixée à 1. Ainsi, les valeurs représentées pour chaque échantillon n'expriment pas le nombre de copies des transcrits des gènes d'intérêt mais un rapport entre le niveau d'expression du gène d'intérêt et le niveau d'expression moyen des gènes de référence.

Dans les quatre échantillons (E, C, CE et EN), les niveaux d'expression de *VvPR10.7* et *VvPR10.8* sont trop faibles et ne pourront pas être analysés, tandis que pour *VvPR10.9*, un seul échantillon (C) est quantifiable. Les profils d'expression des quatre gènes dont les quatre échantillons peuvent être quantifiés présentent une allure similaire, soit *VvPR10.1-a*, *VvPR10.2*, *VvPR10.3-a* et *VvPR10.5*. Ils peuvent être classés par leur niveau d'expression décroissant : *VvPR10.2*, *VvPR10.3a*, *VvPR10.1-a* et *VvPR10.5*. Les trois premiers gènes ont des niveaux d'expression supérieurs au niveau d'expression moyen des gènes de référence, au moins pour certains points (E, C, CE et EN pour *VvPR10.2* ; C et CE pour *VvPR10.1-a* et *VvPR10.3*-a). Pour les quatre gènes, le traitement par le 2,4-D induit fortement l'expression (d'un facteur 10). Le niveau d'expression est toujours élevé dans les cals, aussi bien dans ceux traités au 2,4-D que dans les cals embryogènes. Le niveau est supérieur dans ces derniers pour les gènes *VvPR10.1-a* et *VvPR10.5* alors qu'il est égal

Gène	Efficacité d'amplification moyenne	Écart-type
VvActine	1,97	0,03
VvEF1α	1,97	0,04
VvPR10.1-a	1,79	0,06
VvPR10.2	1,96	0,03
VvPR10.3-a	1,91	0,04
VvPR10.5	1,79	0,04
VvPR10.7	1,99	0,01
VvPR10.8	1,96	0,01
VvPR10.9	1,97	0,02

Tableau 15 : Efficacité d'amplification moyenne et leur écart-type pour chaque gène sur les six expériences indépendantes.

Efficacité : $E = 10^{-1/(\text{pente de la droite étalon})}$, avec un maximum de 2.

Gène	E	C / $E_{2,4\text{-}D}$	CE	EN	N	$N_{2,4\text{-}D}$	F	$F_{2,4\text{-}D}$
VvPR10.1-a	+	+	+	+	+	+	+	+
VvPR10.2	+	+	+	+	+	+	+	+
VvPR10.3-a	+	+	+	+	+	+	+	+
VvPR10.5	+	+	+	+	-	+	-	+
VvPR10.7	-	-	-	-	+	-	+	-
VvPR10.8	-	-	-	-	-	-	+	+
VvPR10.9	-	+	-	-	-	-	+	+

Tableau 16 : Validation des Cq par rapport aux bornes des droites étalons.

+ = résultat validé ; - = résultat non validé

E = embryons en début de cycle ; C / $E_{2,4\text{-}D}$ = cals obtenus après un traitement au 2,4-D (3 semaines) ; CE = cals embryogènes sans les embryons néoformés ; EN = masses d'embryons secondaires néoformés en fin de cycle ; N = nœuds non traités ; $N_{2,4\text{-}D}$ = cals obtenus à partir de nœuds traités au 2,4-D ; F = feuilles non traitées ; $E_{2,4\text{-}D}$ = cals obtenus à partir de feuilles traitées au 2,4-D

pour *VvPR10.2* et inférieur pour *VvPR10.3-a*. Enfin, les niveaux d'expression dans les embryons de départ et dans les embryons secondaires néoformés sont relativement proches pour chaque gène. L'expression du gène *VvPR10.9* n'est détectable que dans les cals obtenus à partir d'embryons traités, ce qui suggère là aussi un effet inducteur du 2,4-D.

3. Étude par RT-qPCR de l'expression des gènes *PR10* dans des tissus présentant différentes capacités embryogènes lors du traitement par le 2,4-D

La seconde étape des analyses a consisté à quantifier l'expression des sept gènes choisis, lors du traitement par le 2,4-D de tissus présentant différentes capacités embryogènes : embryons, nœuds et feuilles. La figure 16 présente le niveau d'expression de chaque gène dans les différents tissus analysés et, lorsque cela est possible, le rapport des niveaux d'expression entre le tissu traité et non traité au 2,4-D.

Tout comme pour l'étude de l'expression durant le cycle d'embryogenèse somatique secondaire, certains gènes étaient trop faiblement exprimés pour être analysés (Tableau 16).

L'expression des gènes *VvPR10.1-a*, *VvPR10.2*, *VvPR10.3-a* et *VvPR10.5* est plus élevée dans les échantillons traités au 2,4-D que dans les échantillons non traités, pour tous les tissus. De plus, pour *VvPR10.1-a*, le niveau d'expression dans les tissus non traités est plus élevé dans les embryons que dans les autres tissus, et suite au traitement par le 2,4-D, ce gène est exprimé à un niveau comparable dans les 3 tissus. De ce fait, l'expression de *VvPR10.1-a* est plus fortement induite par le 2,4-D dans les feuilles (environ 200 fois) et les nœuds (environ 100 fois) que dans les embryons (environ 8 fois). Les niveaux d'expression de *VvPR10.2* et *VvPR10.3-a* sont d'un ordre de grandeur comparable dans les trois tissus non traités (respectivement 10 et 1 fois le niveau d'expression moyen des gènes de référence), ainsi que dans les trois tissus traités (respectivement 100 et 10 fois le niveau d'expression moyen des gènes de référence). Les expressions de *VvPR10.2* et *VvPR10.3-a* sont plus fortement induites par le traitement 2,4-D dans les embryons (respectivement environ 14 et 20 fois) et les nœuds (respectivement environ 17 et 19 fois) que dans les feuilles (respectivement environ 6 et 3 fois). Avant traitement, aucun transcrit de *VvPR10.5* n'est détecté dans les nœuds et les feuilles alors que dans les embryons le niveau

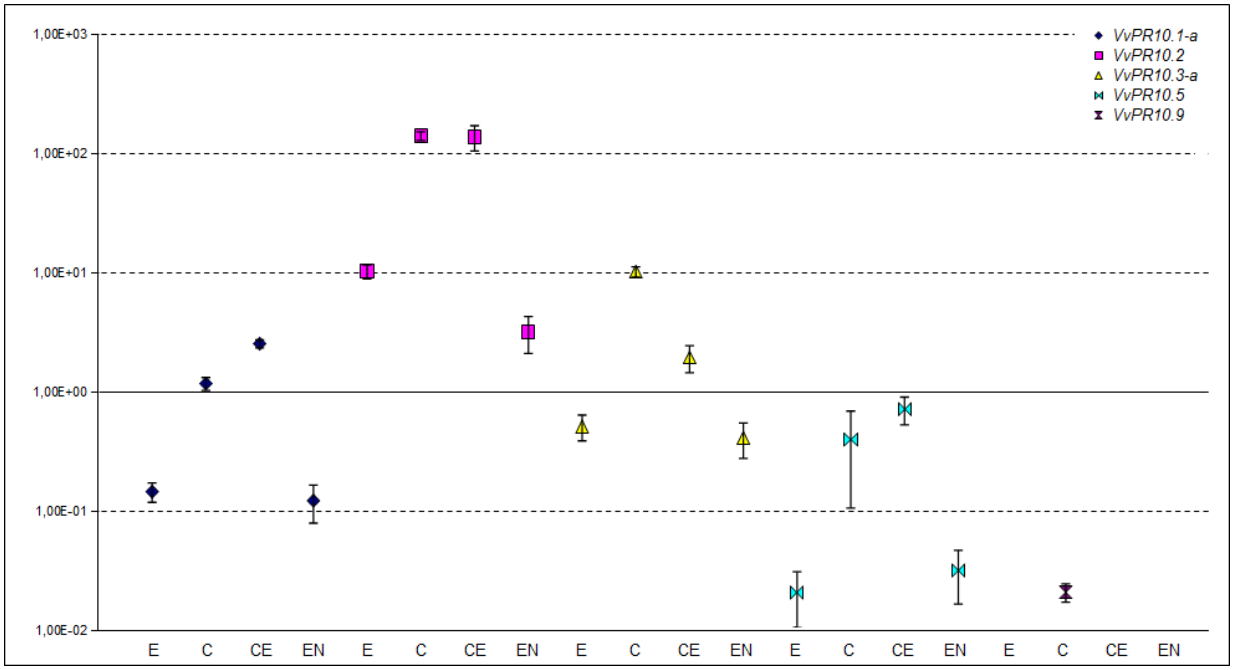


Figure 15 : Expression en RT-qPCR des gènes *VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5* et *VvPR10.9* durant le cycle d'embryogenèse somatique secondaire.

Les niveaux d'expression sont exprimés comme le rapport entre le niveau d'expression du gène d'intérêt et le niveau d'expression moyen des deux gènes de référence, *VvActine* et *VvEF1α*. La valeur 1 correspond donc au niveau d'expression moyen des gènes de référence.

E = embryons en début de cycle ; C = cals obtenus après un traitement au 2,4-D (3 semaines) ; CE = cals embryogènes sans les embryons néoformés ; EN = masses d'embryons secondaires néoformés en fin de cycle.

d'expression est faible (environ 100 fois moins que le niveau d'expression moyen des gènes de référence). Après traitement, il y a une induction de l'expression de ce gène et tous les niveaux d'expression sont d'un ordre de grandeur comparable à celui des gènes de référence. Cependant le facteur d'induction du niveau d'expression par le 2,4-D ne peut être calculé que pour les embryons ; il est d'environ 19 fois.

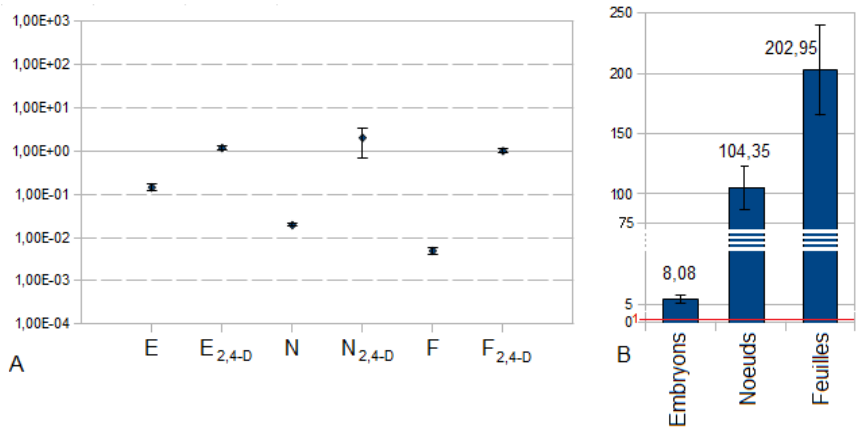
VvPR10.8 et *VvPR10.9* sont tous deux très faiblement exprimés dans les tissus de départ, à peine détectables dans les feuilles (en dehors des bornes de la droite étalon, pour les nœuds et embryons). Le traitement au 2,4-D a peu ou pas d'effet sur la régulation de l'expression de ces gènes dans les feuilles. Cependant, l'expression de *VvPR10.9* est induite dans les embryons traités.

L'expression de *VvPR10.7* n'est détectée que dans deux tissus sans traitement, feuilles et nœuds, à des niveaux environ équivalents, et 10 fois inférieurs au niveau moyen des gènes de référence. A l'opposé de ce qui est observé pour les gènes précédents, il n'y a pas de transcrit de *VvPR10.7* détectable suite au traitement au 2,4-D qui réprime donc l'expression de ce gène.

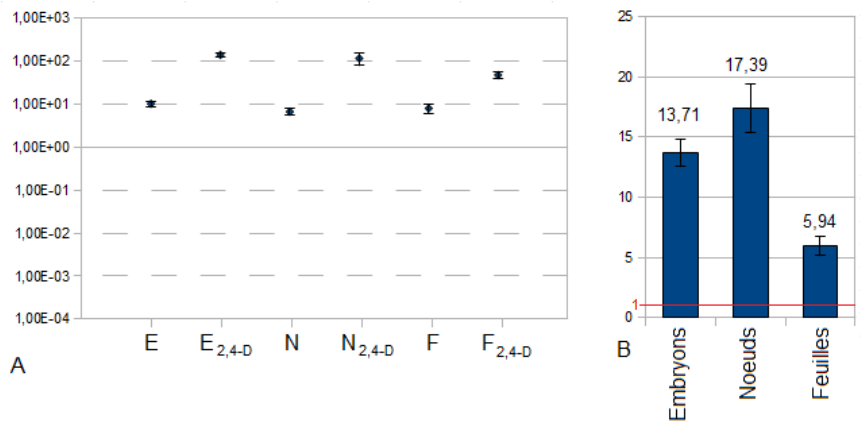
4. Comparaison des résultats de RT-PCR semi-quantitative avec les résultats de RT-qPCR

Lorsque le niveau d'expression a pu être évalué par RT-PCR quantitative, il est possible de comparer ces résultats à ceux obtenus en RT-PCR semi-quantitative, puisque les échantillons ont été prélevés au cours des mêmes expériences.

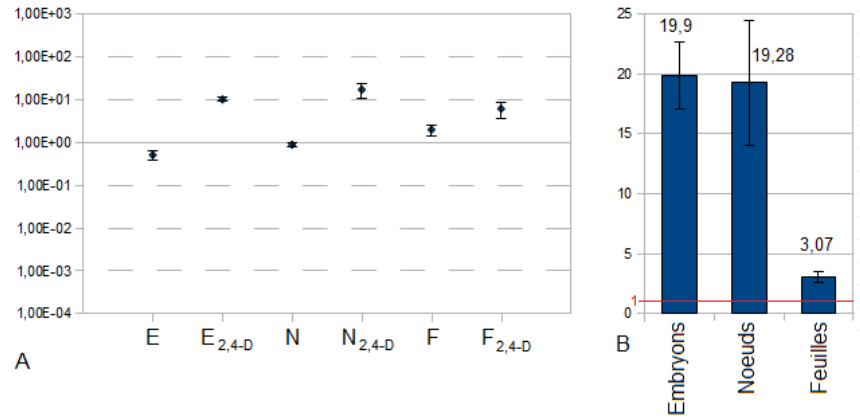
Pour les gènes *VvPR10.1-a, VvPR10.3-a* et *VvPR10.5*, les variations d'expression mesurées par RT-qPCR dans les tissus après traitement 2,4-D sont globalement comparables aux différences observées en RT-PCR semi-quantitative. Cependant dans les nœuds traités au 2,4-D, aucune expression n'a pu être détectée pour *VvPR10.5* en RT-PCR semi-quantitative alors que la RT-qPCR montre que des transcrits de ce gène sont présents. On peut également noter, pour ces trois gènes, que les différences entre les niveaux d'expression dans les cals embryogènes et les embryons néoformés mesurés par les deux méthodes sont relativement comparables.



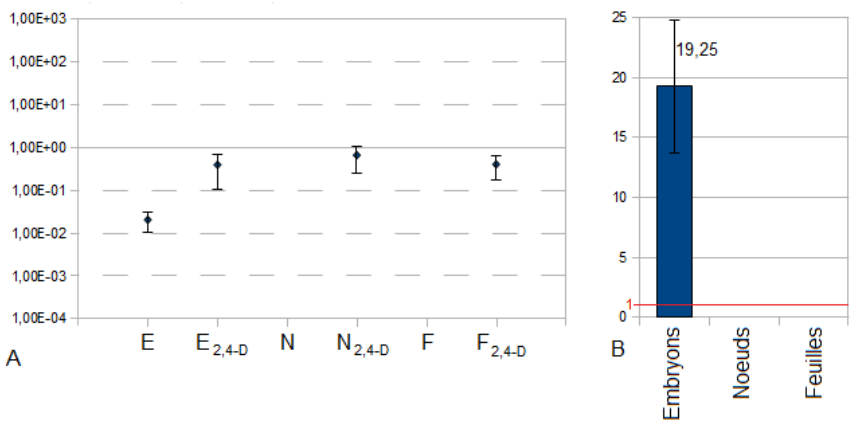
VvPR10.1-a



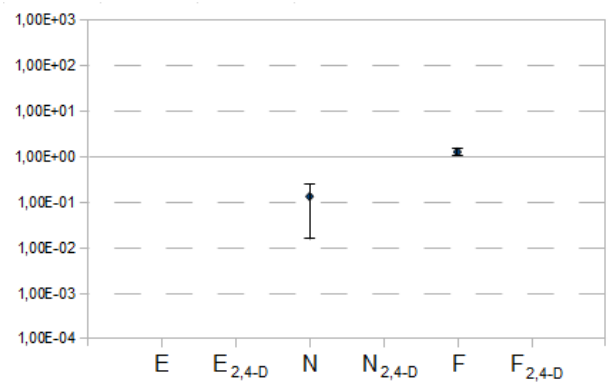
VvPR10.2



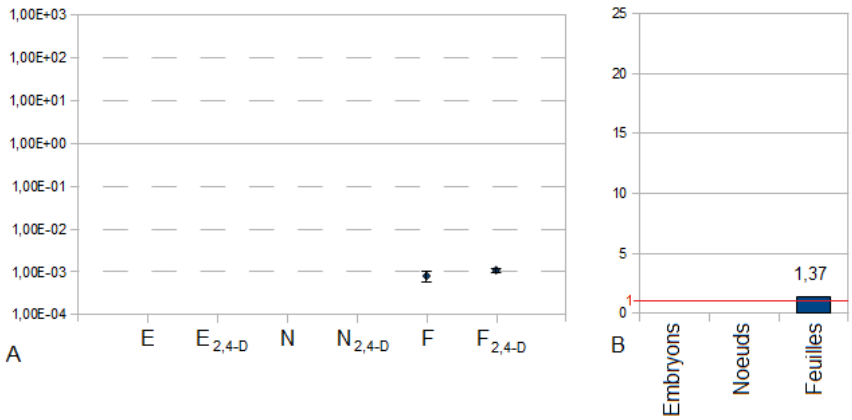
VvPR10.3-a



VvPR10.5



VvPR10.7



VvPR10.8

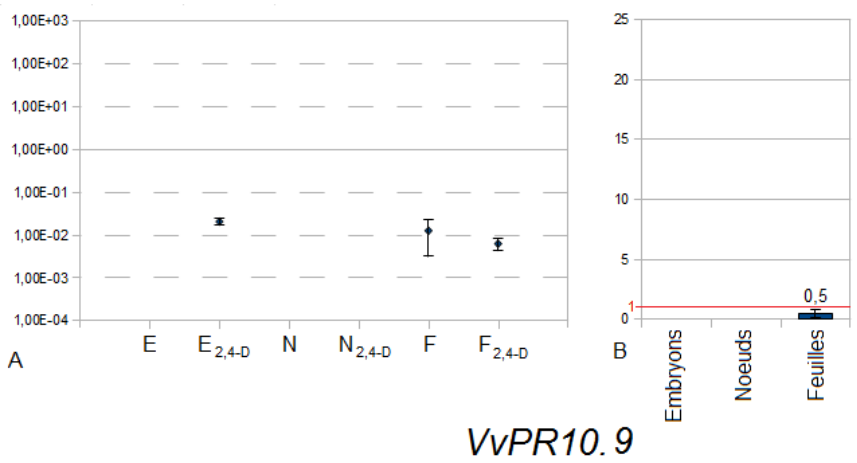


Figure 16 : Expression en RT-qPCR des gènes *VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.7, VvPR10.8* et *VvPR10.9* dans des embryons, nœuds et feuilles avant et après traitement par le 2,4-D.

A : Les niveaux d'expression sont exprimés comme le rapport entre le niveau d'expression de chaque gène et le niveau d'expression moyen des deux gènes de référence, *VvActine* et *VvEF1α*. La valeur 1 correspond donc au niveau d'expression moyen des gènes de référence.

E = embryons non traités ; $E_{2,4\text{-}D}$ = cals obtenus à partir d'embryons traités au 2,4-D ; N = nœuds non traités ; $N_{2,4\text{-}D}$ = cals obtenus à partir de nœuds traités au 2,4-D ; F = feuilles non traitées ; $E_{2,4\text{-}D}$ = cals obtenus à partir de feuilles traitées au 2,4-D.

B : Rapport de l'expression de chaque gène entre l'échantillon traité au 2,4-D et l'échantillon non traité. Ce rapport est calculé pour chaque tissu lorsque cela est possible.

Le gène *VvPR10.2* semble être assez peu régulé lorsqu'on étudie son expression au moyen de la technique de RT-PCR semi-quantitative. Pourtant, grâce à la technique de RT-qPCR, on peut voir que les quantités de transcrits sont plus élevées suite au traitement par le 2,4-D, dans les différents tissus étudiés. Le niveau d'expression est aussi plus élevé dans les cals embryogènes que dans les embryons néoformés. L'image d'expression non régulée de ce gène, obtenue par RT-PCR semi-quantitative, pourrait provenir de la présence d'une quantité importante de transcrits en début d'amplification dans les échantillons, c'est à dire d'une forte expression « de base », amenant rapidement l'amplification en phase plateau. Afin de vérifier cette hypothèse, des amplifications nouvelles ont été réalisées avec un nombre de cycles plus faible. Ces analyses montrent cette fois clairement une différence d'expression de *VvPR10.2* dans des embryons avant (E) et après (C) traitement au 2,4-D, avec 28 ou 30 cycles de PCR (Figure 17).

Avec les deux techniques, les résultats sont similaires pour *VvPR10.7* : les transcrits de ce gène sont présents uniquement dans les feuilles et les nœuds et pas ou très faiblement dans les embryons avant le traitement par le 2,4-D, qui réprime ensuite ce gène.

Dans les embryons et les nœuds non traités, les signaux obtenus par RT-PCR semi-quantitative sont très faibles ou nuls pour *VvPR10.8* et *VvPR10.9*. Cette observation est confirmée par RT-qPCR puisque les Cq mesurés sont en dehors des limites de la droite étalon. Dans les feuilles non traitées, de faibles quantités de transcrits de ces deux gènes sont mesurées à l'aide de la RT-qPCR, mais pas par RT-PCR semi-quantitative, qui est donc moins sensible pour des gènes faiblement exprimés. Après traitement au 2,4-D, les signaux très faibles de *VvPR10.9* obtenus par RT-PCR semi-quantitative suggèrent une légère induction de l'expression de ce gène. Cependant, la RT-qPCR montre que ce gène n'est en fait pas significativement régulé dans les feuilles lors de ce traitement. Par contre, il est possible de quantifier ces transcrits dans les embryons, suite au traitement 2,4-D.

En conclusion, à l'exception de *VvPR10.5* dans les nœuds traités au 2,4-D, les résultats de RT-qPCR confirment globalement ceux observés en RT-PCR semi-quantitative.

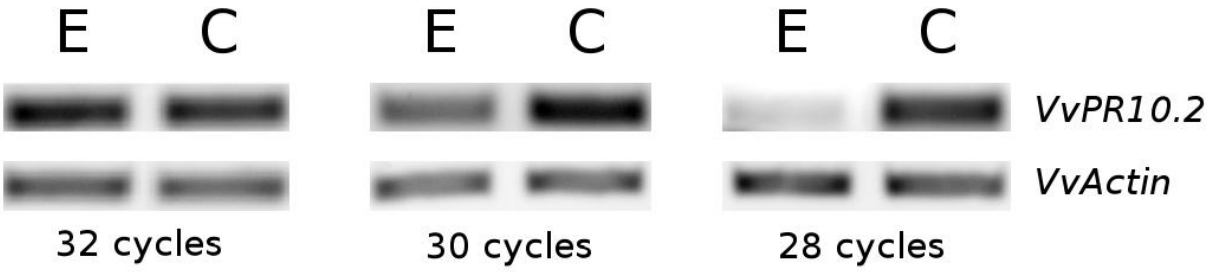


Figure 17 : Expression en RT-PCR semi-quantitative de *VvPR10.2* en fonction du nombre de cycles.

Le gène *VvActin* est utilisé comme standard interne à la réaction.

E = embryons en début de cycle ; C = cals obtenus après un traitement au 2,4-D (3 semaines)

IV. Discussion.

A. Expression des gènes *VvPR10* pendant l'embryogènese somatique chez la vigne

Dans ce travail, l'expression de *VvPR10.10* n'a pu être détectée, dans aucun tissu, traité ou non. De plus aucune EST correspondante n'a pu être trouvée dans les bases de données et aucune expression n'est prédite par les logiciels. Il n'est toutefois pas possible d'affirmer que ce gène ne soit jamais exprimé. Il pourrait l'être très faiblement, dans certains tissus ou organes spécifiques, sous l'action d'éliciteurs particuliers, ou à certains stades du développement, que nous n'avons pas étudiés.

1. Expression durant le cycle d'embryogènese somatique secondaire

L'analyse de l'expression des gènes *VvPR10* durant certaines étapes clé de l'embryogenèse somatique secondaire chez la vigne montre que seuls 7 gènes sur 10 étudiés sont exprimés durant au moins l'une des étapes du cycle. L'expression de l'un d'entre eux (*VvPR10.8*) ne varie pas au cours du cycle, ce qui n'en fait pas un candidat intéressant en tant que marqueur, même si ce gène peut jouer un rôle dans l'embryogenèse somatique. Les six autres (*VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.6, VvPR10.9*) ont une expression stimulée par le traitement 2,4-D, auxine forte supposée induire l'acquisition de la capacité embryogène. Cette stimulation de l'expression est similaire à celle qui est observée dans le cadre d'autres situations de stress. Chez la vigne, l'expression de gènes *VvPR10* peut être induite à la suite d'une infection par *Pseudomonas syringae, Erysiphe necator* ou *Plasmopara viticola* (Robert et al. 2001, Kortekamp et al. 2006, Fung et al. 2008, Polesani et al. 2010), par le stress salin (Jellouli et al. 2008, Jellouli et al. 2010), par un traitement au chitosane (Ferri et al. 2009) ou un traitement à l'herbicide flumioxazine (Castro et al. 2005, Castro et al. 2008). Chez d'autres espèces aussi, l'induction de l'expression de gènes *PR10* semble comparable lors de divers stress biotiques ou abiotiques, ce qui conduit certains auteurs à conclure que cette induction ne serait pas spécifique du type de stress (Handschuh et al. 2007). Toutefois, même si l'induction de l'expression de certains gènes *VvPR10* dans les embryons traités au 2,4-D pourrait être imputée au stress provoqué par l'auxine, on peut constater que, à l'exception de *VvPR10.6*, le niveau d'expression des gènes *VvPR10* continue à être plus élevé dans les cals

embryogènes que dans les embryons de départ, c'est-à-dire jusqu'à un à deux mois après la fin du traitement par le 2,4-D. Cette situation semble particulière, puisque l'analyse de l'expression d'autres gènes de défense comme *PR1* ou *PR8* dans les mêmes conditions, montre que leur expression, provoquée par le traitement au 2,4-D, est seulement transitoire. C'est pourquoi, l'expression élevée de certains gènes *VvPR10* pendant une grande partie du cycle cultural d'induction de l'embryogenèse pourrait suggérer qu'ils jouent un rôle dans ce processus. Les protéines PR10 sont décrites comme pouvant lier de petits ligands apolaires et les transporter dans la cellule (Neudecker et al. 2001, Mogensen et al. 2002, Koistinen et al. 2005, Fernandes et al. 2008). Elles pourraient donc par exemple intervenir dans les processus de régulation des hormones endogènes ayant lieu lors de l'embryogenèse somatique (Gaj 2004, Jimenez 2005).

2. Expression des gènes *VvPR10* lors du traitement inducteur de la capacité embryogénique

L'auxine 2,4-D est une des molécules les plus fréquemment utilisées pour induire l'acquisition de la capacité embryogénique dans des tissus végétaux. Bien que ses effets ne soient pas encore totalement compris, son efficacité pourrait être due à son action double, à la fois comme un facteur de stress et comme une auxine exogène (Fehér et al. 2003, Fehér 2008). Du fait de l'importance que revêt le traitement au 2,4-D dans l'acquisition de la capacité embryogénique, j'ai choisi d'analyser la régulation de l'expression des gènes *VvPR10* en présence de cette hormone dans des tissus présentant des potentiels embryogènes différents. Dans nos conditions de laboratoire, et suivant le protocole de culture utilisé, l'embryogenèse somatique secondaire montre une efficacité largement supérieure à l'embryogenèse somatique primaire réalisée sur des explants nodaux, tandis qu'aucun cal embryogène ne peut être obtenu à partir de feuilles, dans le cultivar Chardonnay. Les trois tissus différents choisis sur la base de leur potentiel embryogène ont donc été des embryons (efficacité embryogène élevée), des nœuds (efficacité embryogène faible) et des feuilles (efficacité embryogène nulle).

Lors de l'étude de l'expression des gènes *VvPR10* avant et après traitement au 2,4-D, aucun gène n'a montré de profil d'expression rigoureusement identique entre les trois tissus analysés. Cette diversité au niveau des profils d'expression a déjà pu être constatée au niveau des gènes *PR10* d'autres espèces : certains gènes montrent une expression tissu-

spécifique ou une régulation durant le développement, d'autres sont exprimés constitutivement ou sont induits par divers stress (Fernandes et al. 2008).

L'expression du gène *VvPR10.8* ne semble pas régulée lors du traitement par le 2,4-D.

Le gène *VvPR10.6* n'est exprimé que très faiblement dans les embryons traités au 2,4-D. Il est en revanche plus fortement exprimé dans les racines et les inflorescences immatures, suggérant une expression tissu-dépendante et une activité dans le développement.

Les gènes *VvPR10.1-a* et *VvPR10.5-a* sont tous deux exprimés plus faiblement dans les nœuds et les feuilles que dans les embryons. Leur expression est induite par le 2,4-D et leur niveau d'expression est comparable dans les trois tissus traités, ce qui implique un facteur d'induction plus élevé dans les nœuds et feuilles que dans les embryons. Ce facteur d'induction n'a pas pu être calculé pour *VvPR10.5* dans les nœuds et feuilles car le niveau d'expression dans ces tissus non traités était en dessous de la limite inférieure de la droite étalon. Pour l'instant, il n'est pas possible d'établir de lien entre le niveau d'expression de ces gènes et la capacité embryogénique des tissus.

Les gènes *VvPR10.4* et *VvPR10.7* sont exprimés tous deux clairement dans les nœuds et/ou les feuilles et sont fortement réprimés lors du traitement par le 2,4-D. Pour le gène *VvPR10.*7, cette répression est telle que le signal est en-dessous de la limite inférieure de la droite étalon. Ce profil de régulation est assez atypique pour un gène *PR10*. Cependant, la répression de l'expression de gènes *PR10* a déjà été mentionnée, dans le cas des gènes *LlPR10.1a* et *LlPR10.1b* du lupin, dans le cadre d'une interaction symbiotique (Sikorski et al. 1999).

L'expression des gènes *VvPR10.2* et *VvPR10.3-a* est induite sous l'action du 2,4-D dans les trois tissus testés. Toutefois, le facteur d'induction est significativement plus élevé dans les embryons et les nœuds, tissus ayant une capacité embryogénique, que dans les feuilles qui n'en montrent pas, ce qui suggère que ces gènes pourraient jouer un rôle dans l'embryogenèse somatique. L'expression du gène *VvPR10.9* est également induite dans les embryons mais pas dans les feuilles sous l'action du 2,4-D. Son niveau d'expression globalement faible ne permet toutefois pas de déterminer quel type de régulation intervient dans les feuilles traitées. De la même manière que pour *VvPR10.2* et *VvPR10.3-a*, cette

régulation différentielle entre des tissus fortement et très faiblement embryogéniques suggère que *VvPR10.9* pourrait jouer un rôle dans l'embryogenèse somatique.

B. Comparaison des résultats de RT-PCR semi-quantitative et de RT-qPCR

L'analyse de l'expression des gènes *VvPR10* a été effectuée à la fois en RT-PCR semi-quantitative et en RT-qPCR, sur les mêmes échantillons, permettant ainsi de comparer les résultats obtenus par ces deux techniques. Mis à part deux exceptions concernant l'expression de *VvPR10.5* dans les embryons et les nœuds, les résultats obtenus avec ces deux techniques sont convergents.

Les recommandations de MIQE (Bustin et al. 2009) permettent de s'assurer que la RT-qPCR a été réalisée dans de bonnes conditions, délivrant ainsi des résultats fiables et reproductibles. Dans ce travail, j'ai veillé à ce que les étapes de récolte et de conservation des tissus végétaux, d'extraction d'ARN, de traitement à la désoxyribonucléase et de transcription inverse soient équivalentes pour tous les échantillons. La récolte d'échantillons homogènes permet de comparer les résultats obtenus pour plusieurs expériences indépendantes. La répétition de la gamme étalon lors de chaque expérience est un niveau de contrôle supplémentaire. Les points correspondant à cette gamme sont toujours alignés sur une droite, attestant la validité des conditions de PCR. De plus, les efficacités d'amplifications sont comparables d'une expérience à l'autre, montrant que les amorces s'hybrident avec une efficacité constante entre les différentes expériences. Toutefois, les efficacités des amorces servant à détecter les transcrits de *VvPR10.1-a* et *VvPR10.5* sont assez faibles (1,79). Des amorces plus performantes pourraient être conçues pour ces gènes afin d'améliorer les conditions de PCR quantitative. Malheureusement, la forte similarité entre les séquences de tous ces gènes laisse très peu de latitude dans le choix de la séquence des amorces. Une trop forte similarité de certaines de ces séquences a d'ailleurs empêché la conception d'amorces permettant d'amplifier spécifiquement certaines des trois formes de *VvPR10.1* ou *VvPR10.3*. Malgré cela, l'absence de sous-produit d'amplification, repérable au niveau des courbes de fusions générées en fin de réaction, prouve que seul l'ADNc recherché a été amplifié. La RT-qPCR permet donc

d'amplifier spécifiquement *VvPR10.1-a* et *VvPR10.3-a*, ce que le séquençage des produits de PCR ne pouvait assurer en RT-PCR semi-quantitative.

La RT-qPCR présente un niveau de précision faisant défaut à la RT-PCR semi-quantitative, qui ne permet pas une quantification des niveaux de variation (induction ou repression). Dans ce travail, elle a par exemple permis de mettre en évidence des facteurs de régulation suite au traitement par le 2,4-D. Ces niveaux de variation dépendent à la fois du gène étudié et du tissu concerné. La RT-PCR quantitative a également montré que l'expression de *VvPR10.2* est régulée au cours de l'embryogenèse somatique, alors qu'en RT-PCR semi-quantitative la quantité de ce transcrit ne semblait pas varier.

Néanmoins, malgré la sensibilité de la RT-qPCR, l'analyse de l'expression des membres de la famille multigénique *VvPR10* peut se montrer délicate, en particulier pour des gènes très peu exprimés. Le rapport d'expression entre le gène le plus exprimé et le moins exprimé est de l'ordre de presque 100 000 fois et, alors que les valeurs de Cq enregistrées pour *VvPR10.2* sont dans la limite supérieure de sa droite étalon, celles qui sont enregistrées pour *VvPR10.8* et *VvPR10.9* sont dans la limite inférieure des leurs, voire fréquemment en dessous. Dans ce cas, la RT-PCR semi-quantitative est utile pour confirmer que ces gènes sont effectivement exprimés dans les embryons et les feuilles. Par ailleurs, dans le cas de gènes faiblement exprimés, c'est la transcription inverse qui peut être limitante si peu de messagers étaient présents dans l'extrait d'ARNs de départ. Le choix d'un autre protocole (autres amorces ou autre transcriptase inverse) pourrait alors s'avérer judicieux. L'utilisation d'un protocole de RT-PCR quantitative employant une quantité initiale de matrice plus importante pourrait également permettre d'évaluer convenablement l'expression de ces gènes faiblement exprimés. Cela nécessiterait l'emploi d'une gamme étalon élargie qui permettrait de les quantifier dans les mêmes conditions que les gènes les plus exprimés, tel que *VvPR10.2*. Les gènes présentant des niveaux d'expression très différents (élevé et faible pour 2 gènes étudiés en même temps) posent également le problème du choix des gènes de référence. Il est conseillé d'employer des gènes de référence dont le niveau d'expression est proche des gènes d'intérêt (Huggett et al. 2005). Si *VvEF1α* et *VvActin* se sont révélés être des choix adaptés à l'étude des gènes *VvPR10* les plus exprimés (*VvPR10.1-a, VvPR10.2, VvPR10.3-a* et *VvPR10.5*), on peut en revanche s'interroger sur le pertinence de normaliser le niveau d'expression d'un gène grâce à des gènes de référence dont les niveaux d'expression lui sont 1000 fois supérieurs. Le choix de

gènes de référence ayant un niveau d'expression beaucoup plus faible que *VvEF1α* et *VvActin* poserait de manière inverse des problèmes pour les gènes les plus fortement exprimés. Le plus judicieux au final serait certainement de réaliser deux groupes, avec d'un côté les gènes fortement exprimés et de l'autre côté ceux qui le sont faiblement. Des gènes de référence pour chacun de ces deux groupes pourraient être choisis. La limite à cette étude est que les mesures dans les deux groupes ne seraient alors plus comparables de manière absolue. Seuls les niveaux d'induction suite au traitement par le 2,4-D par exemple pourraient être comparés.

Au final, il nous est apparu que les techniques de RT-PCR semi-quantitative et de RT-PCR quantitative sont très fortement complémentaires l'une de l'autre. La RT-PCR semi-quantitative donne une vue d'ensemble de l'expression des gènes d'intérêt sur laquelle il est nécessaire de s'appuyer pour pousser l'observation dans le détail au moyen de la RT-qPCR.

Conclusion générale et perspectives

Dans le cadre de cette thèse, deux études ont été menées : d'une part la caractérisation des gènes *PR10* à l'aide d'outils bioinformatiques et des séquences du génome disponibles chez la vigne (Jaillon et al. 2007, Velasco et al. 2007), et, d'autre part, l'étude de l'expression de ces gènes dans différents tissus et durant l'embryogenèse somatique secondaire par RT-PCR semi-quantitative et quantitative.

I. Identification de la famille multigénique des *PR10* chez la vigne

Depuis la deuxième moitié de l'année 2007, deux séquences complètes de génome de vigne de deux accessions (PN40024 et PN115) sont disponibles (Jaillon et al. 2007, Velasco et al. 2007). Ces données, et en particulier les annotations du GENOSCOPE, ont été des outils précieux pour l'identification des gènes *PR10* de la vigne. Ils ont permis de mettre en évidence l'existence de dix-sept séquences apparentées à des gènes *PR10* dans cette espèce. Le logiciel d'annotation automatisé, GAZE (Howe et al. 2002), utilisé par le GENOSCOPE, avait prédit treize séquences parmi les dix-sept comme des *PR10*, et quatre autres séquences étaient reconnues comme des gènes potentiels, mais non comme des *PR10*. Sept autres séquences, au contraire annotées comme des *PR10* ou des *PR10-like*, n'en sont probablement pas, car elles ne montrent qu'une très faible similarité avec les séquences connues. De plus, d'autres erreurs d'annotation concernaient la délimitation de certaines séquences. Pour trois d'entre elles, *VvPR10.3-a*, *VvPR10.3-b* et *VvPR10.3-c*, chacun des deux exons était reconnu comme un gène indépendant, formant ainsi six annotations différentes. Pour six autres séquences, *VvPR10.1-a*, *VvPR10.1-b*, *VvPR10.1-c*, *VvPR10.4*, *VvPR10.8* et *VvPR10.10*, les sites donneur et accepteur d'épissage de l'intron étaient mal identifiés, car décalés de quatre paires de base en aval. Ces problèmes d'annotation des bornes de l'intron suggèrent que l'algorithme du logiciel automatique utilisé par GAZE pourrait avoir des difficultés à reconnaître certains sites d'épissage. De telles erreurs lors de l'annotation automatique de génomes ou de bases de données sont relativement communes (Devos et Valencia 2001, Schnoes et al 2009). Elles ont été mises en évidence pour les gènes de la famille « cinnamyl alcohol dehydrogenase » d'*Arabidopsis thaliana* (Kim et al. 2004) ou plus récemment, pour le génome de *Physcomitrella patens* (Rensing et al. 2008). Elles impliquent la nécessité d'une

vérification systématique des données issues d'annotation automatisée et soulignent l'importance d'un travail coordonné entre les différents membres de la communauté scientifique afin de corriger progressivement les bases de données (Lee et al. 2007, Schnoes et al. 2009).

Les dix-sept séquences apparentées à des *PR10* identifiées dans le génome de la vigne forment une famille multigénique, composée de trois pseudogènes et d'au moins treize gènes exprimés, ce qui est attesté par l'existence d'ESTs correspondantes, dans les banques. L'ensemble des séquences est organisé en un cluster compact de copies disposées en tandem sur le chromosome 5, entre les positions 599228 et 679082, soit sur une portion d'environ 80 kb. Je n'ai pu identifier aucun autre gène sur cette portion du chromosome. De plus, aucun autre gène *PR10* n'a été localisé ailleurs dans le génome, ce qui suggère que mon identification a été exhaustive et que tous les gènes *PR10* de la vigne se trouvent réunis dans ce cluster. Il s'agit, à ma connaissance, de la première famille de gènes *PR10* intégralement identifiée et localisée dans un génome. De précédents travaux relatent l'étude de la famille des *PR10* chez d'autres espèces, mais l'identification des gènes et leur localisation restent partielles, en raison de l'absence de la séquence génomique complète de ces espèces. Les deux études les plus complètes sont celle de la famille *Mal d 1* de *Malus domestica* (Gao et al. 2005), et celle des familles *Pru p 1* de *Prunus persica* et *Pru d 1* de *Prunus dulcis* (Chen et al. 2008). La famille du pommier comprend 18 gènes, dont 16 sont répartis sur deux groupes de liaison (comportant 7 et 9 gènes respectivement) et deux sont indépendants, tandis que celles du pêcher et de l'amandier comptent chacune 8 gènes localisés sur un unique groupe de liaison. Il est possible que les familles de gènes *PR10* de ces trois espèces soient organisées de manière similaire à celle de *Vitis vinifera*. Cette organisation pourrait être caractéristique des familles multigéniques *PR10*, comme suggéré aussi par l'étude de la famille *Bet v 1* du bouleau (Schenk et al. 2009). La connaissance de la famille multigénique *PR10* de la vigne pourrait servir de base pour rechercher des gènes *PR10* dans les autres génomes de plantes déjà séquencés, au moyen d'outils de comparaison génomiques tels que tels que Plaza (Proost et al. 2009, http://bioinformatics.psb.ugent.be/plaza/), FLAGdb++ (Samson et al. 2004, http://genoplante-info.infobiogen.fr/FLAGdb/) ou Phytozome (http://www.phytozome.org). Plusieurs génomes de plantes sont déjà disponibles dans ces bases de données, dont les génomes récemment séquencés *Cucumis sativus* (Huang et al. 2009) ou *Ricinus communis* (Chan et al. 2010).

L'organisation chromosomique des gènes *PR10* de la vigne suggère que la famille s'est agrandie par des duplications en tandem successives, probablement générées majoritairement par des évènements de crossing-over inégal, car 13 copies sur 17 sont en orientation directe. Il serait intéressant, à ce stade, d'essayer de retracer l'histoire évolutive de ce cluster, comme cela a été réalisé pour le locus *MLA* de l'orge, par exemple (Wei et al. 2002). Par ailleurs, nous avons observé une très grande similarité entre les dix-sept séquences, comme c'est aussi le cas au sein des familles multigéniques *PR10* d'autres espèces. L'étude phylogénétique comparative des séquences identifiées dans différentes espèces montre aussi que la variabilité intraspécifique est supérieure à la variabilité interspécifique, ce qui suggère une évolution par des mécanismes d'évolution concertée, comme suggéré par plusieurs auteurs (Wen et al. 1997, Freitas et al. 2003, Finkler et al. 2005, Schenk et al. 2006, Musidlowska-Persson et al. 2007).

La séquence génomique utilisée pour l'identification des gènes *VvPR10* a été assemblée à partir de fragments mis bout à bout, dont la longueur totale équivaut à 8,4 fois la longueur totale du génome (Jaillon et al. 2007). Celle-ci est disponible sur le site du GENOSCOPE sous le nom de « 8X ». Depuis mi-2010, un assemblage de fragments dont la séquence totale atteint 12 fois (12X) la longueur totale du génome est également disponible sur le site du GENOSCOPE. Du fait de cette mise à jour de la base de données, une recherche d'éventuelles séquences supplémentaires de gènes *PR10* est à présent nécessaire. Celle-ci permettra de vérifier, d'une part, que notre identification était exhaustive, et, d'autre part, qu'aucune séquence ne vient interrompre le « cluster » de séquences *PR10*. De plus, une nouvelle carte, présentant les nouvelles références de la base (12X), sera générée pour ce cluster.

II. Spécialisation et diversification fonctionnelle des gènes *VvPR10*

Les premières études sur l'expression de *VvPR10.1* et *VvPR10.2* lors de l'embryogenèse somatique secondaire chez la vigne ont permis de mettre en évidence l'existence d'un gène *VvPR10.3* (Maillot et al. 2009). C'est grâce à un couple d'amorces conçues pour hybrider les ADNc de *VvPR10.1* (Robert et al. 2001) que les transcrits de ce gène ont pu être détectés. L'amplification préférentielle de *VvPR10.3* avec des amorces plus spécifiques de la séquence de *VvPR10.1* a été possible grâce à la très grande similarité entre les deux séquences nucléotidiques (93%), et suggérait que les transcrits de *VvPR10.1* étaient nettement moins abondants que ceux de *VvPR10.3* dans les tissus analysés. Cette hypothèse a été confirmée par les expériences de RT-qPCR montrant une expression de *VvPR10.1* environ 10 fois moindre que celle de *VvPR10.3* dans ces tissus.

A la suite de leur identification dans la base du GENOSCOPE, l'expression de 10 gènes *VvPR10* a été mesurée dans différents tissus tels que des racines, des tiges, des nœuds, des feuilles, des inflorescences immatures et des embryons, par la technique de RT-PCR semi-quantitative. Les niveaux d'expression de 7 de ces gènes, les plus exprimés, ont également été mesurés par RT-qPCR dans des noeuds, des feuilles et des embryons. A l'exception de *VvPR10.2* et, dans une moindre mesure, de *VvPR10.3* (pour les nœuds, feuilles et embryons), aucun gène ne montre d'expression similaire dans tous les tissus analysés. De nombreux gènes *PR10* caractérisés dans d'autres espèces montrent également de tels profils d'expression tissu-dépendant (Walter et al. 1996, Sikorski et al. 1999, Pühringer et al. 2000, Zhou et al. 2002, Liu et al. 2004, Park et al. 2004, Kortekamp et al. 2006, Liu et al. 2006, Jellouli et al. 2008, Kim et al. 2008, Jellouli et al. 2010, Xie et al. 2010). Cette expression tissu-spécifique traduit une certaine spécialisation fonctionnelle.

L'ensemble des résultats d'expression obtenus montre que la famille multigénique des *PR10* de vigne présente une probable diversification fonctionnelle, certains gènes étant plus ou moins inductibles par le 2,4-D et d'autres présentant un profil d'expression atypique pour une « pathogenesis-related protein ». Il serait intéressant, pour prolonger cette étude, de tester l'action de différents éliciteurs biotiques ou abiotiques sur l'expression de ces gènes, afin de préciser l'étendue des phénomènes dans lesquelles ils sont impliqués.

Parallèlement à l'étude de leur induction pendant le processus d'embrogenèse somatique, nous avons montré que l'expression des gènes *PR10* de vigne est le plus souvent tissu-spécifique, suggérant que différents gènes pourraient assurer une même fonction dans différents organes. Les protéines PR10 sont décrites comme ayant deux fonctions majeures, une fonction de transport et une fonction ribonucléasique (Liu et Ekramoddoullah 2006). Leur spécificité est principalement conditionnée par la forme de la cavité hydrophobique pour la première fonction et par les acides aminés impliqués dans l'activité ribonucléasique pour la seconde. Si la gamme de ligands acceptés par certaines cavités hydrophobiques est généralement assez large (Mogensen et al. 2002, Koistinen et al. 2005), une modification dans la forme de cette cavité peut théoriquement la changer. La diversité fonctionnelle des protéines PR10 au sein de la cellule serait ainsi due à la diversité des molécules transportées, selon la conformation de la cavité hydrophobique. Les molécules transportées pourraient elles-mêmes être impliquées dans des processus biologiques différents tels que la défense ou le développement (Wu et al. 2003, Koistinen et al. 2005, De Jong et al. 2007, Fernandes et al. 2008). La détermination expérimentale de la forme des cavités hydrophobiques des différentes protéines PR10 de la vigne, par cristallographie ou résonance magnétique nucléaire par exemple, devrait permettre d'identifier des repliements communs entre les molécules. Des expériences de déplacement de ligand par gradient d'affinité sur des extraits protéiques purifiés, comme celles qui ont été effectuées avec la protéines Bet v 1 de bouleau (Mogensen et al. 2002), permettraient par ailleurs de définir au moins en partie les gammes de ligands pouvant se fixer sur chaque protéine PR10 différente, ce qui fournirait des informations importantes pour déterminer leur fonction spécifique.

III. Marqueurs de l'embryogenèse chez la vigne

Parmi les différents gènes candidats testés comme marqueurs de l'embryogenèse somatique secondaire, *VvSERK1*, *VvSERK2* et *VvL1L* sont exprimés spécifiquement dans les embryons, confirmant leur rôle de marqueurs du développement précoce chez la vigne. Concernant les gènes *VvPR10*, *VvPR10.2* et *VvPR10.3-a* ont une expression induite par le traitement au 2,4-D, facteur supposé inducteur de la capacité embryogénique. Leur niveau d'expression est élevé dans les cals formés après le traitement au 2,4-D et le reste pendant toute la durée de la culture, jusqu'à l'apparition des embryons secondaires. Ces résultats suggèrent que ces gènes pourraient jouer un rôle dans le processus d'embryogenèse somatique, comme cela a été suggéré chez la luzerne (Domoki et al. 2006). Une étude cinétique de l'expression quantitative de ces gènes pendant toute la durée du processus cultural permettrait de savoir si un pic d'expression se produit juste avant la formation des embryons secondaires, comme chez la luzerne. Enfin, *VvPR10.2*, *VvPR10.3-a* et *VvPR10.9*, montrent tous trois une expression plus intense dans les tissus à capacité embryogéniques (embryons et nœuds) que dans les tissus non embryogéniques (feuilles), ce qui pourrait faire d'eux des marqueurs de cette capacité embryogénique

.

Matériels et méthodes

Composition	Concentration (mg/L)
Macroéléments	
KNO_3	950
NH_4NO_3	825
$MgSO_4$, $7H_2O$	185
$CaCl_2$, $2H_2O$	220
KH_2PO_4	85
Microéléments	
$MnSO_4$, $4H_2O$	22,33
KI	0,83
HBO_3	6,2
$ZnSO_4$, $7H_2O$	8,6
CuSO4, $5H_2O$	0,03
$NaMoO_4$, $2H_2O$	0,25
$CoCl_2$, $6H_2O$	0,03
Fer Chélaté	
$FeSO_4$, $7H_2O$	27,8
Na_2 EDTA	37,3
Vitamines et éléments organiques	
Acide nicotinique	0,5
Pyroxydine-HCl	0,5
Thiamine-HCl	0,1
Myo-inositol	100
Glycine	2
Agar	7 000
pH 5,8	

Tableau 17 : Composition du milieu de Murashige et Skoog 1962 modifié (macroéléments réduits de moitié) utilisé comme base nutritive pour les milieux de culture.

I. Matériel végétal

Toutes les études d'expression ont été effectuées à partir de plantes ou de cultures de *Vitis vinifera* cultivar Chardonnay. Les milieux décrits utilisent comme base un milieu Murashige et Skoog (Murashige et Skoog, 1962) modifié, dont la composition est détaillée dans le tableau 17. Les macroéléments, les microéléments, les vitamines, le saccharose et les régulateurs de croissance sont inclus avant autoclavage. Le pH est ajusté à 5,8 par l'ajout de 1N NaOH. Le milieu est solidifié par ajout de 0,7% Bacto-Agar et stérilisé par autoclavage à 115° C pendant 30 minutes. L'incubation des cultures est réalisée à l'obscurité, à une température de 25° C.

A. Embryogenèse somatique secondaire

Des cals embryogènes primaires ont été obtenus par culture d'explants nodaux (Maillot et al. 2006) et les cultures embryogènes qui en sont issues ont été maintenues par embryogenèse secondaire récurrente. Pour induire l'embryogenèse somatique secondaire, des embryons intacts et bien formés, au stade cotylédonaire, sont récoltés sur des cals embryogènes. Les embryons sont déposés séparément, 14 embryons par boîte de Petri (60 mm de diamètre) contenant du milieu E96 (Tableau 18). Les cultures sont incubées durant 3 semaines, les embryons formant des cals par dédifférenciation. Les cals sont ensuite transférés sur milieu A (Tableau 19), cessent de croître pour devenir bruns. Ces cals sont repiqués sur milieu A, toutes les 4 semaines, jusqu'à l'apparition de masses d'embryons secondaires néoformés (en général au cours d'une période de subculture de 1-3 mois).

B. Germination d'embryons

Pour obtenir la germination d'embryons au stade cotylédonaire, il suffit de les transférer (3 par boîte de Petri de 60 mm de diamètre) sur un milieu de germination (Tableau 20) et de les incuber à l'obscurité à 25° C. Après une période de 3 à 8 jours, ils sont transférés progressivement à la lumière pour obtenir des plantules complètement développées

.

Composition	Concentration
Benzyl-adénine (BAP)	9 µM
2,4-D	4,5 µM
Saccharose	60 g/L

Tableau 18 : Ajouts au milieu de base pour le milieu E96.

Composition	Concentration
Charbon actif	2,5 g/L
Acide indole 3-acétique	20 µM
Acide 2-naphtoxyacétique	10 µM
Benzyl adénine (BAP)	1 µM
Saccharose	60 g/L

Tableau 19 : Ajouts au milieu de base pour le milieu A.

Composition	Concentration
Benzyl adénine (BAP)	0,2 µM
Saccharose	25 g/L

Tableau 20 : Ajouts au milieu de base pour le milieu de germination.

C. Traitement des tissus au 2,4-D

Des nœuds et feuilles sont collectés à partir de plantules bien développées (10-12 feuilles) obtenues par micro-bouturage *in vitro* (Maillot et al. 2006). Des embryons au stade cotylédonaire sont prélevés sur des cals embryogènes. Les différents organes sont déposés séparément sur des boîtes de Petri (60 mm de diamètre, 14 embryons, 14 nœuds et 7 fragments de feuille par boîte) contenant du milieu E96 (Tableau 18). Les cultures sont incubées durant 3 semaines, les tissus formant des cals.

D. Prélèvement des échantillons végétaux

Les échantillons végétaux ont été récoltés à diverses étapes : embryons cotylédonaires en début d'expérience (E) ; cals développés à partir d'embryons sur milieu E96 (C1/C/$E_{2,4\text{-}D}$) ; cals une semaine après transfert sur milieu A (C2) ; cals non embryogènes (NEC) et cals embryogènes sans les embryons secondaires (EC/CE) après trois repiquages sur milieu A ; masses d'embryons secondaires néoformés (CSE/EN) ; racines de plantule cultivée *in vitro* (R) ; tiges de plantule cultivée *in vitro* (T) ; feuille de plantule cultivée *in vitro* (F) ; cals développés à partir de feuilles sur milieu E96 ($F_{2,4\text{-}D}$) ; nœuds de plantule cultivée *in vitro* (N) ; cals développés à partir de nœuds sur milieu E96 ($N_{2,4\text{-}D}$) ; embryons ayant germé, portant un méristème apical visible, deux cotylédons développés et une racine différenciée (GE). Les fleurs immatures ont été disséquées à partir d'inflorescences récoltées sur des vignes plantées en extérieur.

Les échantillons regroupaient plusieurs exemplaires du même tissu (plusieurs nœuds, feuilles, embryons, etc.) et ont immédiatement été congelés dans l'azote liquide puis conservés à -80° C.

II. Biologie moléculaire

A. Extraction de l'ADN

L'ADN total a été extrait grâce au « DNeasy Plant Mini Kit » (Qiagen) avec un protocole adapté à la vigne. En effet, la vigne est une plante riche en composés phénoliques qui

interagissent avec les acides nucléiques. Le tampon de lyse a donc été modifié par l'ajout de 2,5% de PVP40 (Sigma-Aldrich). Les échantillons ont été broyés dans l'azote liquide puis incubés dans le tampon de lyse à 65°C durant 10 minutes, avant de continuer avec le protocole décrit par le fabriquant. Les échantillons d'ADN sont ensuite conservés à -20°C. L'ADN a été visualisé sur gel d'agarose 1,5% (II.E.1.) et sa concentration a été déterminée par spectrophotomètrie, en utilisant un marqueur fluorescent de l'ADN et avec l'appareil « Qubit Quantification Platform » (Invitrogen).

B. Extraction et traitement de l'ARN

1. Extraction

L'ARN total a été extrait grâce au « Nucleospin RNA plant kit » (Macherey-Nagel) avec un protocole adapté à la vigne (pour les mêmes raison que pour l'extraction de l'ADN). Le tampon de lyse a donc été modifié par l'ajout de 2,5% de PVP40, 1% de β-mercaptoéthanol et 2% de sarkosyl (Sigma-Aldrich). Les échantillons (80 à 100 mg) ont été broyés dans l'azote liquide puis incubés dans le tampon de lyse à 70°C durant 10 minutes, avant de continuer avec le protocole décrit par le fabriquant. Les échantillons d'ARN sont ensuite conservés à -80°C.

2. Contrôle de qualité de l'ARN

La qualité des extraits d'ARN est primordiale pour la suite des expériences et plusieurs contrôles sont venus s'assurer de celle-ci. L'intégrité des ARN a été vérifiée par migration de 3µl des ARN totaux sur un gel d'agarose 1% (II.E.2.). Il s'agit de contrôler présence des deux bandes correspondant aux sous-unités de l'ARN ribosomi.que et de vérifier que l'intensité de la bande de la grande sous unité est approximativement le double de celle de la petite sous unité. La concentration en ARN a été déterminée par spectrophotomètrie, en utilisant un marqueur fluorescent de l'ARN et avec l'appareil « Qubit Quantification Platform » (Invitrogen). Selon les directives de MIQE, cette méthode pour mesurer l'ARN grâce à des marqueurs fluorescents est parmi les meilleures pour détecter de faibles

concentrations (Bustin et al. 2009). Cette évaluation a permis d'ajuster tous les échantillons à la même concentration, pour avoir la plus grande homogénéité possible lors des étapes ultérieures. De plus, l'établissement du rapport A260/A280, toujours supérieur à 1,8, a permis de s'assurer de l'absence de contamination protéique pouvant gêner lors des étapes ultérieures.

3. Traitement à la désoxyribonucléase

Pour une RT-PCR classique, il est possible d'utiliser des amorces situées de part et d'autre d'un éventuel intron afin de distinguer l'ADNc correspondant au gène d'intérêt de l'éventuel l'ADN génomique contaminant. Lors d'une RT-qPCR, l'ADNc et l'ADNg sont amplifiés et mesurés de la même manière, il est donc important de se débarrasser de toute trace d'ADN contaminant. Un traitement à la désoxyribonucléase a été réalisé sur chaque échantillon grâce à la « RQ1 Rnase-free Dnase » (Promega), sur une quantité d'1µg d'ARN total selon les recommandations du fournisseur. Son efficacité a été contrôlée par PCR sur des échantillons d'ARNm, en utilisant les amorces amplifiant *VvActin* et en vérifiant l'absence d'amplification par migration sur gel d'agarose 1,5 % (II.E.2.). Les échantillons d'ARN traités à la désoxyribonucléase sont ensuite conservés à -80°C.

C. Transcription inverse

La transcriptase inverse nécessite une efficacité maximale pour faire la transcription inverse de tous les ARNm, même ceux présents en très petite quantité. Dans le cadre de la RT-PCR, cette étape peut être incluse dans le protocole de la réaction, on parle alors de RT-PCR « one-step », ou peut être séparée. J'ai choisi cette deuxième solution afin d'avoir un contrôle possible sur cette étape. Le kit « SuperScript II Reverse Transcriptase » (Invitrogen) a été utilisé pour cette étape, en utilisant 3 µl de chaque produit après traitement à la désoxyribonucléase pour un volume réactionnel de 20 µL. A la fin de la réaction, les échantillons d'ADNc sont tous ajustés à 100 µL avec de l'eau stérile (« nuclease-free ») et conservés à –20°C.

Orientation de l'élongation	Nom de l'amorce	Séquence de l'amorce
« RACE » en 5'	SP1	GGGACCTCCCTGTCCTTGTAG
	SP2	CTTGAGGCCTTATCTTGGG
	SP3	CGAGGAAGTGACCTCACTCTCG
« RACE » en 3'	SP5	CTACAAGGACAGGGAGGTCCC

Tableau 21 : Liste des amorces utilisées pour la « RACE-PCR ».

D. Amplification en chaîne par polymérase

1. PCR « classique »

Les amorces de RT-PCR semi-quantitative (II.D.3.a.) et quantitative (II.D.4.b.) ont été testées par PCR « classique » avant d'être utilisées pour ces réactions. Les amplifications en vue de réaliser les gammes étalon de RT-qPCR (II.D.4.d.) ont également été réalisés dans ces conditions de PCR « classique ».

1µl de matrice d'ADN a été utilisé pour chaque réaction de PCR dans 25µl de volume total. L'enzyme « Taq DNA Polymerase Recombinant » (1,25 unité) et son tampon (concentration finale x1) (Invitrogen) ont été utilisés. Les réactions ont été réalisées en présence d'une concentration finale de $MgCl_2$ de 1,5mM. Chaque amorce a été utilisée à une concentration de 100nM. La température d'hybridation a été adapté au Tm des amorces (en général 57°C).

Les réactions de PCR se sont déroulée selon le protocole suivant :

- Dénaturation initiale : 95°C pendant 5 min
- Dénaturation : 95° pendant 30 s
- Hybridation : 57°C pendant 30 s
- Elongation : 72°C pendant 45 s
- Elongation finale : 72°C pendant 10 min

} 36 cycles

Les résultats de PCR sont analysés après migration des produits d'amplification sur gel d'agarose.

2. Amplification rapide des extrémités d'ADNc

L'amplification rapide des extrémités d'ADNc (RACE-PCR :« rapid amplification of cDNA ends ») a permis l'amplification d'un fragment d'ADNc en direction de ses extrémités 5' et 3'. Elle a été réalisée à partir d'ARN totaux grâce au « 5'/3' RACE Kit » (Roche), en suivant les instructions du fabriquant. Des amorces spécifiques décrites dans le tableau 21 ont été utilisées en complément des amorces spécifiques de *VvPR10.1* et *VvPR10.3*. Elle ont permis d'obtenir, à partir des ARN messagers, trois séquences légèrement chevauchantes qui ont été séquencées (II.F).

Gène	GenBank/accession GENOSCOPE	Amorce sens (5'...3')	Amorce antisens (5'...3')
VvActin	AF369525	TGCTATCCTTCGTCTTGACCTTG	GGACTTCTGGACAACGGAATCTC
VvPR10.1-a	AJ291705	GAAATCATACAAGGAGAGGGAGGC	GCCAAACTTATTGAGACTGATAGGTG
VvPR10.2	AJ291704	CGATCACAGTGTAGCGGAATGAGAAT	AAGCTATCAAGTGCGTGGAAGTCATT
VvPR10.3-a	EU379313	GAAATCCTACAAGGACAGGGAGGT	CGGCCTTGGTGTGGTACTTTT
VvPR10.4	FN666569	ATCCTTCCCCAAGCTATCAAG	GATTTGCCAAGAGGTAAGCC
VvPR10.5	FN666570	ATCCTTGACTCTGATAACCTCA	ATGATATGAGACAAAGGAGTTTC
VvPR10.6	FN666571	GTCCTTGATGTTGATAACCTC	AAGCCAAGCCTTTTAACTG
VvPR10.7	FN666572	ATCGTCCCTCAGGCCATTA	AAGTGATTAAGTGGAGGAGAAGC
VvPR10.8	FN666573	CTCTTGCCCCAGACCATAAG	ACATTGGACAACAGAGAAGTGAC
VvPR10.9	FN666574	CAGTCAAAAGTACGCGACTCA	AAGTATAGGCGCGAGGGTGT
VvPR10.10		AATCAGTAAAGAGCATCGAGTT	AAAGTAATCACAACTCCTCGTC
VvPR1	AJ536326	GCCTACGCCCAGAACTATGCTAAC	CGAACCACCACCCATTGTTG
VvPR3	AJ291505	TGTAGCAACTAAGACGTTTATTCA	GCCAGCTTATTACCAGGTCCAT
VvPR8	AJ291507	ATCATCGTCTCGGCCATTAG	AGAGCAGTGCCCATGAACTT
VvSERK1	GSVIVT00001777001	TGTTTAAGAGAACGCCC	GAGTTGGAGTGAGGAGCG
VvSERK2	GSVIVT00019412001	GCCTAAGAGAACGTCCACCAT	ACCTCCTGACGGACGACC
VvSERK3	GSVIVT00009544001	GTTTGAGAGAGCGAGCTGATG	GGTGGGGGATATGGTTGTA
VvL1L	GSVIVT00010958001	CAGCCATGGAAGACACTGAATGCA	CCTTCAAGCTCGCGGTAGTGGTGG

Tableau 22 : Liste des amorces utilisées en RT-PCR semi-quantative.

3. PCR semi-quantitative

a. Conception des amorces

Les amorces ont été conçues au moyen de l'outil en ligne « OligoPerfect » Designer (Invitrogen, http://tools.invitrogen.com/content.cfm?pageid=9716). Les éventuels appariements avec d'autres cibles ont été vérifiés grâce à l'outil en ligne BLAST (« Basic Local Alignment Search Tool ») du site du National Center for Biotechnology Information (NCBI, http://blast.ncbi.nlm.nih.gov) contre l'ensemble du génome de la vigne. Les « self-annealing », « cross-annealing » ou « hairpin » éventuels ont été vérifiés grâce à l'outil en ligne « PCR Primer Stat » (http://bioinformatics.org/sms2/pcr_primer_stats.html). La liste des amorces utilisées pour la RT-PCR semi-quantitative est donnée dans le tableau 22.

b. Calibrage des conditions de phase exponentielle

Pour que l'analyse de l'expression par RT-PCR semi-quantitative soit la plus exacte possible, il est tout d'abord nécessaire de calibrer les conditions afin d'être sur de stopper la réaction en phase exponentielle. Pour ce faire, une amplification de *VvActin* a été réalisée par PCR (II.voir c. Réaction de PCR semi-quantitative) sur une durée de 28 cycles, en utilisant comme matrices une cascade de 5 dilutions au demi à partir de la matrice initiale (ADNc dilué, obtenu après transcription inverse). 5µL de matrice (non diluée ou diluée) sont utilisés pour 25 µL de mélange réactionnel final (II.D.3.c.). Après amplification et migration sur gel d'agarose 1,5 % (II.E.1.), le programme QuantityOne (BioRad) est utilisé pour déterminer l'intensité du signal correspondant à chaque réaction. La courbe (intensité du signal) en fonction de (quantité relative de matrice de départ) construite et permet de déterminer la dilution appropriée pour se placer en phase exponentielle (Figure 8).

c. Réaction de PCR semi-quantitative

5µl de matrice diluée après calibrage ont été utilisés pour chaque réaction de RT-PCR semi-quantitative dans 25µl de volume total. L'enzyme « Taq DNA Polymerase Recombinant » (1,25 unité) et son tampon (concentration finale x1) (Invitrogen) ont été utilisés. Les réactions ont été réalisées en présence d'une concentration finale de $MgCl_2$ de 1,5mM. Chaque amorce a été utilisée à une concentration de 100nM. La température

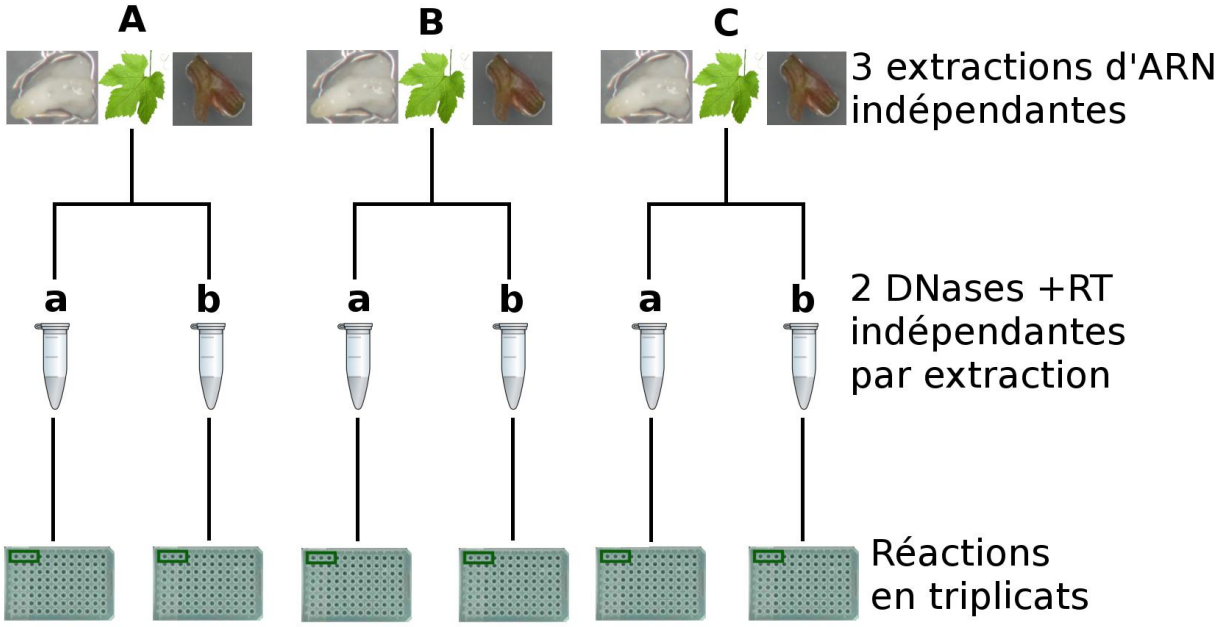


Figure 18 : Plan d'expérience de RT-qPCR.

Trois extractions (d'ARN) indépendantes ont été réalisées. Deux traitements à la désoxyribonucléase (DNases) indépendants suivis de transcriptions inverses (RT) ont été réalisés sur chaque extrait d'ARN. Six expériences indépendantes ont donc été réalisées au total. Les amplifications ont été réalisées en triplicats de dépôt afin de tenir compte de la variabilité intra-essai.

	RT-qPCR (d'après Quellhorst et Rully 2008)	RT-PCR classique (d'après Singh et Kumar 2001)
Taille de l'amplicon	50 à 210 bp	150 à 1000 bp
Taille de l'amorce	19 à 23 nt	17 à 34 nt
Pourcentage de GC	35 à 65 %	45 à 60 %
Tm	60 à 68°C	55 à 56°C

Tableau 23 : Principales différences entre les paramètres idéals de mise au point d'amorces en vue expériences de RT-qPCR ou en vue d'expériences de RT-PCR classique

d'hybridation a été adapté au Tm des amorces (habituellement 57°C). Les réactions de PCR semi-quantitative se sont déroulée selon le protocole suivant :

- Dénaturation initiale : 95°C pendant 5 min
- Dénaturation : 95° pendant 30 s
- Hybridation : 57°C pendant 30 s
- Elongation : 72°C pendant 45 s
- Elongation finale : 72°C pendant 10 min

} 32 cycles

Les résultats de RT-PCR semi-quantitative sont analysés après migration des produits d'amplification sur gel d'agarose. Ils ont été déterminés d'après trois expériences indépendantes.

4. RT-PCR quantitative

Les recommandations de MIQE fournissent une liste d'informations méthodologiques à fournir conjointement aux résultats de RT-qPCR (Bustin et al. 2009). Certains points concernant l'extraction des acides ribonucléiques et la transcription inverse ont été traités plus haut. Les autres seront abordés ci-dessous.

a. Plan d'expérience

Deux parties sont traitées en utilisant la technique de RT-qPCR. La première concerne l'étude de l'expression des gènes *VvPR10* à des points clés du cycle d'embryogenèse (échantillons E, C, CE et EN). La seconde partie traite de l'étude de l'effet du 2,4-D sur des tissus présentant des capacités embryogènes différentes (échantillons E et $E_{2,4\text{-}D}$; Fet $F_{2,4\text{-}D}$; N et $NF_{2,4\text{-}D}$). L'expression dans les échantillons non traités sert de témoin pour comparer l'expression dans les échantillons traités par le 2,4-D.
D'après les conseils de MIQE, des répétitions intra- et inter- expérimentales sont nécessaires pour la validation des résultats. Trois expériences indépendantes ont été réalisées, deux transcriptases inverses indépendantes étant effectuée pour chacune (Figure 18). Les réaction de RT-qPCR (pour chaque gène et chaque échantillon) sont réalisées en triplicats, permettant d'éviter la prise en compte de résultats artefactuels.

```
VvPR10.1-a      AGTGTGGAAACCATACAAGGAGAGGGAGGCCCTGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.1-b      AATGTGGAAATCATACAAGGAGAGGGAGGCCCTGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.1-c      AGTGTGGAAATCATACAAGGAGAGGGAGGCCCTGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.2        TGCGTGGAAGTCATTCAAGGAGACGGAGGCCCTGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.3-a      AGTGTGGAAATCCTACAAGGACAGGGAGGTCCCGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.3-c      AGTGTGGAAATCCTACAAGGACAGGGAGGTCCCGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.3-b      AGTGTGGAAATCCTACAAGGACAGGGAGGTCCCGGAACCATCAAGAAGATTCACTTTGGT 180
VvPR10.4        AATGTTGAGATCATCCAAGGAGATGGAGGGCCTGGAACCATCAAGAAGATCTACTTTGGT 180
VvPR10.5        TGCGTTGAAATAATTGAAGGAGATGGAGGGCCTGGATCCATCAAGAAGATTACTTTTGGA 180
VvPR10.6        TGCGTCGAAATAATTGAAGGAGATGGAGGGCCAGGAACGATCAAGAAGATTAATTTCGGT 180
VvPR10.7        AGTATCGAAACAGTGGAAGGAAATGGAGGCCCTGGAACCATCAAGCAGCTGAACTTTGCT 180
VvPR10.8        AGCATCGATGTTGTCCAAGGAGATGGAGGAGCTGGAACCATTGAACAAGTGAACTTCACC 180
VvPR10.9        AGCATTGAATTTCTTGAAGGGGAAGGAGAAGTTGGGAGCATCAAACAGATCAACTTCACT 180
VvPR10.10       AGCATCGAGTTTGTTGAAGGTGATGGTGGGGCTGGAAGCATTACACAGACGAACTTCTCT 180
s6              TCCACTGAGATCATTGAAGGAGACGGAGGGCCTGGATCAATCAAAAAGATTACTTTTAGA 180
s7              TGTGTTGAGATCATTGAAGGAGATGGAGGGGCTGGAATCATCAAGAAGGTTACTTTTGGT 180
s9              CGTGTTGAAACCATTGAAGGAGGTGGAGGGTCGGGAACCATCAAGAAGATTACTTTTGGT 179
                       **     *  ****    ** *     **    **     *       **

VvPR10.1-a      GAAGGCAGCAAATTCAAAAGCATGACACACCGGGTTGATGCGATTGACAAAGAGAACTTC 240
VvPR10.1-b      GAAGGCAGCAAATTCAAAAGCATGACACACCGGGTTGATGCGATTGACAAAGAGAACTTC 240
VvPR10.1-c      GAAGGCAGCAAATTCAAAAGCATGACACACCGGGTTGATGCGATTGACAAAGAGAACTTC 240
VvPR10.2        GAAGGCCACAAATTCAAAAGCATGACACACCGGGTTGATGCGATTGACAAAGAGAAATTC 240
VvPR10.3-a      GAAGGCAGAAAATTCAAAAGCATGACACACAGGATTGATGCGATTGACAAAGAGAACTTC 240
VvPR10.3-b      GAAGGCAGAAAATTCAAAAGCATGACACACAGGGTTGATGCGATTGACAAAGAGAACTTC 240
VvPR10.3-c      GAAGGCAGAAAATTCAAAAGCATGACACACAGGGTTGATGCGATTGACAAAGAGAACTTC 240
VvPR10.4        GAAGGCAGCCAATTCAAGAGCGTGACACATCGTGTTGATGGGATTGACAAAGAGAACTTC 240
VvPR10.5        GAAGGGAGCCAATTCAATTATGTGAAGCACAGGGTTGATGGAATAGACAAAGAGAACTTC 240
VvPR10.6        GAAGGCAGCCAATTCAACTACGTGAAGCACTGGATCGATTCATTGGACAAAGAGAATTTC 240
VvPR10.7        GAAGGTAGCCAATTCAAATATGTGAAGCATAGGATTGATGAACTAGACAAAGAGAAGATG 240
VvPR10.8        GAAGCTAGCAATCTCAAATACGTGAAGCACCAAATCGAAGAATTGGACAAAGAGAATTTT 240
VvPR10.9        GAAGCTAGTCCTTTAACATACATGAAGCATCGGATCGATGCACTAGACAAAGAGAAACTC 240
VvPR10.10       GGTGATAGTGATTGCGAGTATTTGAAGTATAAGATCAATGCAGTGGACAAAGAGAAGCTA 240
s6              GAAGGGAGCCAATTCAATTATGTAAAGCACAGGGTTGATGGAATAGACAAAGAGAACTTC 240
s7              GAAGACCAGTTAACCTTATAT-TAAGGCA----ATTAAT--ATCGGACAA---GAGCTTG 230
s9              GAAGGTAGTCAATTCAACTACATAAAGCACCGGATAGATGGGATCAACAAAGAGAACTTC 239
                *  *                  * *   *     *  *        ****   **   *

VvPR10.1-a      ACATTCAGCTA-CACTGTGGTTGACGGAGATG---TTTTGACGGGCGGCATTGAATCAAT 296
VvPR10.1-b      ACATTCAGCTA-CACTGTGGTTGACGGAGATG---TTTTGACGGGCGGCATTGAATCAAT 296
VvPR10.1-c      ACATTCAGCTA-CACTGTGGTTGACGGAGATG---TTTTGACAGGCGGCATTGAATCAAT 296
VvPR10.2        TCATTCTGCTA-CACTGTGATCGACGGGGATG---TTTTGACTGACGGCGTTGAATCAAT 296
VvPR10.3-a      TCATTCAGCTA-CACTGTGATTGATGGAGATG---TTTTGACGAGCGGCATTGAATCAAT 296
VvPR10.3-b      TCATTCAGCTA-CACTGTGATTGACGGAGATG---TTTTGACGAGCGGCATTGAATCACT 296
VvPR10.3-c      ACATTCAGCTA-CACTGTGATTGACGGCGATG---TTTTGACCAGCGGCATTGAATCAAT 296
VvPR10.4        ACTTACAGTTA-CAGTATTATCGAAGGTGATG---CTTTAATGGGCATCCTAGAATCAAT 296
VvPR10.5        ACATACAGTTA-CAGTGTGATTGAGGGAGATG---CTTTGATGGGCACCCTTGAATCAAT 296
VvPR10.6        ACATACTGTTA-CACCATAATTGAAGGAGATG---CTTTGATGGACAACCTTGAGTCAAT 296
VvPR10.7        ATCTACAAGTA-TACTTTGATTGAAGGTGATG---CTTTGATGGACAAGATTGAATATAT 296
VvPR10.8        GTGTGCAAGTA-TCGTATGATTGAGGGAGATG---TATTGGGTGAGGAGCTTGAATCAAT 296
VvPR10.9        ACATGCACATA-CACTATGTTTGAGACTGATGCCCTGATAATGGACAAGATTGAATATAT 299
VvPR10.10       GAGTGCAGATA-CACTTTGATTGAAGGTGGTG---TGCTTGGTGATCAGCTTGAATCGAT 296
s6              ACATACACTTA-CAGTGTGATTGAGGGTGATG---CTTTGATGGGCACCCTTGAATCAAT 296
s7              GTTTGAGATCAACAATCTGATTAATAAAT------TCTTAATTA--AACCAAGAATATAC 282
s9              ACTTACAGTTA-CAGTGTGATCGAAGGTGATG---CTTTGATGGGCACCCTTGAATCATT 295
                   *      *      *  *  *              *             ** *

VvPR10.1-a      TTCTCATG-AGCTCAAAGTGGTGGC-TTCTCCTGATGGAGGATCCATCTACAAGAACACC 354
VvPR10.1-b      TTCTCATG-AGCTCAAAGTTGTGGC-TTCTCCTGATGGAGGATCCATCTACAAGAACACC 354
VvPR10.1-c      TTCTCATG-AGCTCAAAGTGGTGGC-TTCTCCTGATGGAGGATCCATCTACAAGAACACC 354
VvPR10.2        TTGTCATG-AGCTCACAGTGGTGCC-TGCTCCAGGTGGAGGATCCATCTACAAGAACACC 354
VvPR10.3-a      TTCTCATG-AGCTCAAAGTGGTGGC-TTCTCCTGATGGAGGATGCATCTACAAGAACACC 354
VvPR10.3-b      TTCTCATG-AGGTCAAAGTGGTGGC-TTCTCCTGATGGAGGATGCATCTTCAAGAACACC 354
VvPR10.3-c      TTCTCATG-AGCTCAAAGTGGTGGC-TTCTCCTGATGGAGGATGCATCTACAAGAACACC 354
VvPR10.4        CTCTTATG-AGGTCAAGTTGGTGGC-ATCTCCTGATGGAGGATCCATCTGCAAGAACATT 354
VvPR10.5        TTCTTATG-AGGTCAAGTTGGTGGC-TTCTCCCAGTGGAGGATCCATCTGCAAGAGCACT 354
VvPR10.6        CTATTATG-AGGTGAAGCTGGTGGC-CTCTCCTGATGGAGGATCCATCTGCAAGAACATC 354
VvPR10.7        AAGTTACG-AGATTTCATTTGAGGC-CAGTCCTGATGGGGGATGCAAGAGCAAGAATGTG 354
VvPR10.8        TGCTCATG-AGGTGAAGTTTGAGG----CAGCAGATGGAGGTAGCATTTGTAAGATGGCA 351
VvPR10.9        TACTTATG-ATGTGAAGTTCGAGCC-ATATGGTTTTGGGGGGTGTGTTTGTAATCTGACA 357
VvPR10.10       TGTTTATG-AGATGAAATTCGAGGA-GTCTGGTGATGGAGGATGCATTTGCAAGACGAGG 354
s6              TTCTTATG-AGGTCAAGTTGGTGGCCTTCTCCTGATGGAGGATCCATCTGCAAAAGCACC 355
s7              TTTTCACGGAACCTACACCTCTTTGAGGTAACTATTGGATTAATTAAGAATCAAAATGTT 342
s9              CTCTTATG-AGGTCAAGTTGGTG--------------GAGGATCCATCTGCAAGAACACC 340
                   * * * *                            *                *

VvPR10.1-a      AGCAAGTACCACACCAAGGGCGATGTAGAGATCTGTGAAGAGCACG--TTAAGGGTGGCA 412
VvPR10.1-b      AGCAAGTACCACACCAAGGGCGATGTAGAGATCAGTGAAGAGCACA--TTAAGGGTGGCA 412
VvPR10.1-c      AGCAAGTACCACACCAAGGGCGATGTAGAGATCAGTGAAGAGCACA--TTAAGGGTGGCA 412
VvPR10.2        AGCAAGTACCACACCAAGGGCGCTG---AGGTCTGTGAAGAGCACG--TTAAGGGTGGCA 409
VvPR10.3-a      AAAAAGTACCACACCAAGGCCGGTGTAGAGATCAGTGAAGAGCACG--TTAAGGGCGGCA 412
VvPR10.3-b      AAAAAGTACCACACCAAGGCCGGTGTAGAGATCAGTGAAGAGCACG--TTAAGGGCGGCA 412
VvPR10.3-c      AAAAAGTACCACACCAAGGCCGGTGTAGAGATCAGTGAAGAGCACG--TTAAGGGCGGCA 412
VvPR10.4        AGTAAGTACCACACAAAAGATGATGCAGTGATCGATGAAGAGCAAA--TTAAAGCTGGCA 412
VvPR10.5        AGCAAGTACCATACCAAAGGTGACTTTGAGATCACAGAAGACCAAA--TTAAGGCTGGCA 412
VvPR10.6        AGCAAGTACCATACCAAAGGCGACATCCAGATCACAGAAGATCAAA--TCAAGGCTGGCA 412
VvPR10.7        AGCGTGTATCATTCCAAACCCGGTGTGGAGATCAAAGAAGAGGAAA--TCAAGGATGGCA 412
VvPR10.8        AGCGAGTATCACACCAAGGGAAAGTTTGAGATTAAGGAAGAAGAGA--TAAAAGCTGGTA 409
VvPR10.9        AGCGTCTACAAAACAAAAGGGGATGTAGAGATCAGAGAAGAGGACA--TAGAGCATGGCA 415
VvPR10.10       AGTGAGTATCACACTAAGGGGGAGTTTGAGATTAAGGAGGAGAGTA--TCAGAGAAGGGA 412
s6              AGCAAGTACCATACCAAAGGTGACTTTGAGATCACAGAAGACCAAA--TCAAGGCTGGCA 413
s7              -GAAGTTGTAGTTGCAATGCTAGTTTTGCTCTTAATTGATGGCAGACTTCAAAGTTTCCA 401
s9              AGTAAATACCACACCAAAGGGGACATTGAGATCACAGAAGACCAGA--TCAAGGCTGCCA 398
                      *        **            *              *          *
```

```
VvPR10.1-a      AAGAGGAGGCTCTGGCATTGTTCAAGGCTATTGAAGCCTACGTCCTGGCACATCCCGATG 472
VvPR10.1-b      AAGAGGAGGCTCTGGCATTGTTCAAGGCTATTGAAGCCTACGTCCTGGCACATCCCGATG 472
VvPR10.1-c      AAGAAGAGGCTCTGGCATTGTTCAAGGCTATTGAAGCCTACGTCCTGGCACATCCCGATG 472
VvPR10.2        AAGAGGATGCTCTCGCAACGTTCAAGGCTATTGAAGCCTACGTCCTGGCCCATCCTGATG 469
VvPR10.3-a      AAGAGGAGTCTCTCGCAGTGTTCAAGGCTATTGAAGCCTACATCTTGGCACACCCTGATG 472
VvPR10.3-b      AAGAGGAGTCTCTCGCATTGTTCAAGGCTATTGAAGCCTACCTCTTGGCACACCCTGATG 472
VvPR10.3-c      AAGAGGAGTCTCTCGCATTGTTCAAGGCTATTGAAGCCTACGTCCTGGCACATCCCGATG 472
VvPR10.4        AAGAAAAGGCTTCTGGGATGTTCAAGGCTATTGAGGCTTACCTCTTGGCAAATCCTGATG 472
VvPR10.5        AAGAGAAGGCTTTGGGCCTGTTCAAGGCTGTTGAAGCCTATCTCTTGGCAAATCCAGATG 472
VvPR10.6        AAGAGAAGGCTATGGGCATGTTCAAGGCCATTGAAGCCTATCTCTTGGCCAATCCTGATG 472
VvPR10.7        AAGAAAAGGCAGCAGCTGTCTTCAAGGCTGTGGAAGCCTACCTCTTAGCTAATCCTGATG 472
VvPR10.8        AGGCTAGAGCTATGGGGATATACAAAGTTGTGGAAGCCTATCTATTGGAGAATCCCCATG 469
VvPR10.9        AGGACAGAGCAATTGGGATGTATGAAGTGTTGGAAGCCTATCTCATGGCACACCCTCGCG 475
VvPR10.10       AGGAAAAAGCCATGGGAGTGTATAAGCTGGTTGAGGCCTACCTTCTTGCTAACCCTGACG 472
s6              AAGAGAAGGCTATGGGCCTGTTCAAAGCTGTTGAAGCCTACCTCTTGGCACATCCTGATG 473
s7              TCAGTACTTGTATTGCTGCCACCCAGTGTTATGGTGGTTGTTACTGGTAAAATTGATATG 461
s9              AAGAGAAGGCCATGGGCATGTACAAGGCTGTTGAAGCCTACCTCTTGGCAAATCCTGATG 458
                               *         *       *  *  *            *       *

VvPR10.1-a      CCTATTAAGT-----AAAACTGCCTGTAGTAATTGAGATATTAGTCTCT---TGAGTCCA 524
VvPR10.1-b      CCTATTAAGT-----AAAATTGCCTGTAGTAATTGAGATATTAGTCTCT---CGAGTCCA 524
VvPR10.1-c      CCTATTAAGT-----AAAATTGCCTGTAGCAATTGAGATATTAGTCTCT---TGAGTCCA 524
VvPR10.2        CCTATTAAGTTTCCTGTCACGTTCTATTGAGTAAGAGATACTAGTCTCT---TGAGTG-- 524
VvPR10.3-a      CCTATTAAGT-----AAAATTGCCTGTGGTAATTGAGAAATTAGTCTCT---TGAGTCCA 524
VvPR10.3-b      CCTATTAAGT-----AAAATTGCCTGTAGTAATTGAGATATCAGTCTCT---TGAGTCCA 524
VvPR10.3-c      CCTATTAAGT-----AAAATTTCCTGTAGTAATTGAGATATTAGTCTCT---TGAGTCCA 524
VvPR10.4        CTTATTAAGC-----AGTATATCATTGAGATATTATTAGTCTCGTGTCTCCCTATATCTA 527
VvPR10.5        CCTATTAAA------AAACTCCTTTGTCTCATATCATCAGTTAAAGTTG---TGGGTTGG 523
VvPR10.6        CCTATTAAGT-----AAAATTGACCGTATTGAGATTATCTTTAATTTGTG--TCATTC-- 523
VvPR10.7        CCTGAGCGT------AATGCATCCTGTTGAGAGTTTCGAACCTGTATTGCTTCTCCTCCA 526
VvPR10.8        TCTATGCTTAGT---TTC--TATCTTCCTCTATATACATATTCACACATAT-TTGATTGT 523
VvPR10.9        CCTATACTTGAA---AAAAACCTCTCTGATTATTTTGAGAATTGTATGTTT-TTTTTCCT 531
VvPR10.10       AGGAGTTGTGAT---TAC--TTTTTCTAATGTAATGGACCCTCATAAATTCATTCCTTGG 527
s6              CCTACTAAGT-----GAAATGGGATATATTG--TGACAGTCTCGTGT-----TGTGTC-- 519
s7              CAATTCAAGG-----TAAATTCTATGGGAATGATAATTTGGTTGTTGGTTC-TAATTTCA 515
s9              CCTATTAAAC-----AGAGTGGTCCATATTCAATATAGTACTAGTCTCT---TGTGTCTC 510
                                                                         *

VvPR10.1-a      CCTTCATATTCATTGCATCTATCAGTCTCAATAAGTTTGGGCTTTTGTGTT--TTTG-CC 581
VvPR10.1-b      CCTTCATATTCATAGCATCA-----------------------TTTGTGTT--TTTG-CC 558
VvPR10.1-c      CCTTCATATTCATAGCATCA-----------------------TTTGTGTT--TTTG-CC 558
VvPR10.2        -CTCTACCTTCATAGCATCAGTC-------AAAGTGTGGAACTTTTGTGTT--TTTG-CC 573
VvPR10.3-a      CTTTCATATTCATAGAATCAGTCTTAATA---AGTTTGGATCTTTTGTGTTT-TTTG-CC 579
VvPR10.3-b      CCTTCATATTCATAGCATCAGTC---TTAATAAGTTTGGATCTTTTGTGTT--TTTTTCC 579
VvPR10.3-c      CCTTCATATTCATAGCATCA-----------------------TTTGTGTT--TTAG-CC 558
VvPR10.4        TATATATATTGGTGAAA--GGTTTTACT------TCTACATATTT-GTTTTT-TCTG-CT 576
VvPR10.5        GTTTCATATATATTTCTCCGA------------------------CGTCCGG-CTTG--- 555
VvPR10.6        -CTTCATATACATCAGTTAAAAG------------GCTTGGCTTTGATATAAATCCTACC 570
VvPR10.7        CTTAATCACTTGTTGCTTGCTGC-----------------CTTTTTGTTTTCCTTCCACC 569
VvPR10.8        CACTTCTCTGTTGTCG--AATGT-------------TTTCTCTATTATA--TACTTT--- 563
VvPR10.9        AATATGTTTAGCTACCTTTTTTT-------------TTTTTCTTTTTTTCTTTTTTCTTT 578
VvPR10.10       CTCTGCTCTGAAAACCTTATTTA-------------TTTATTTATTTTGGCTGGTTTTTG 574
s6              ---TCAATTTCATCAGTTCAAG-----------------------TTGTGTCC-TTTATAT 553
s7              ACCATCTTTTCCAAACTTTAGGA----------------AATTTGTATTCAAATTTGTAG 559
s9              CCTTCATGTCATCAGTTAAAG--------------GTGTGGATTTTATATTG-TTTCTGC 555
                                                                 *
```

Figure 19 : Position des amorces employée en RT-qPCR pour les gènes *VvPR10.1-a, VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.7, VvPR10.8* et *VvPR10.9.*

La séquence codante à partir du nucléotide 61 et le début de la séquence non traduite en 3' (3'-UTR) sont présentées pour chaque séquence. Les nucléotides correspondant aux bornes de l'intron sont en gras, tandis que ceux correspondant au codon stop sont soulignés. Les amorces sens sont représentées en jaunes, les amorces antisens sont représentées en rouge.

Gène	GenBank	Amorce sens (5'...3')	Amorce antisens (5'...3')
VvActine	AF369525	TGCTATCCTTCGTCTTGACCTTG	GGACTTCTGGACAACGGAATCTC
VvEF1α	EC959059	AATGGCTATGCCCCTGTTCTG	CGCCTGTCAATCTTGGTCAGTAT
VvPR10.1-a	AJ291705	AGCCTACGTCCTGGCACAT	GCCAAACTTATTGAGACTGATAGATG
VvPR10.2	AJ291704	AAGGGCGCTGAGGTCTGT	CTTACTCAATAGAACGTGACAGGAA
VvPR10.3-a	EU379313	GATTCACTTTGGTGAAGGCAGA	TCGTCAAAACATCTCCATCAAT
VvPR10.5	FN666570	ACCATACCAAAGGTGACTTT	ATGATATGAGACAAAGGAGTTTC
VvPR10.7	FN666572	GCGTGTATCATTCCAAACC	CAGGATGCATTAGGCTCAG
VvPR10.8	FN666573	TAGAGCTATGGGGATATACAAAGTTG	AATAGAGAAAACATTGGACAACAGA
VvPR10.9	FN666574	GATGTGAAGTTCGAGCCA	CAACACTTCATACATCCCAA

Tableau 24 : Liste des amorces utilisées en RT-PCR quantitative.

b. Informations sur les séquences cibles et conception des amorces.

Du fait des conditions particulières de la RT-qPCR, il n'est pas indiqué d'employer des amorces de RT-PCR classique. Les principaux paramètres de mise au point d'amorces de RT-qPCR comparés aux paramètres de mise au point d'amorces de RT-PCR classiques sont résumés dans le tableau 23.
La principale différence vient de la taille de l'amplicon significativement plus courte dans le cas de la RT-qPCR. Ainsi, une efficacité maximale de la réaction est assurée le plus longtemps possible. Du fait de la grande similarité entre les séquences *VvPR10*, la latitude dans le choix des amorces était très faible (Figure 19). Celles-ci furent cherchées dans l'extrémité 3'UTR, qui est la région présentant les plus de variation entre les différentes séquences. La liste des amorces utilisées pour la RT-qPCR est donnée dans le tableau 24. Les amplicons générés ont une taille comprise en 113 et 142 paires de bases.

c. Protocole de RT-qPCR

La concentration en matrice déposée dans chaque puits doit être constante, et choisie de manière à permettre la détection des transcrits les moins abondants tout en évitant d'avoir un trop grand nombre de copies de l'un d'entre eux, la technique n'étant pas adaptée pour détecter des transcrits présents en un grand nombre de copies (Kursteiner et al. 2003). En l'absence de données concernant la quantité absolue de chaque transcrit, j'ai choisi un protocole standard consistant à utiliser 12,5 ng d'ADNc pour chaque réaction. La concentration de chaque échantillon a été déterminée le plus précisément possible, grâce à 3 mesures indépendantes sur un Qubit Quantification Platform (Invitrogen).
Les milieux réactifs contenaient 5µl de solution, 12,5 µL de « SybrGreen Supermix 2x » (Biorad) et

200 nM de chaque amorce, pour un volume final de 25 µL. Les amplifications ont été réalisées au moyen d'un « iCycler Real-Time PCR Detection System » (BioRad).

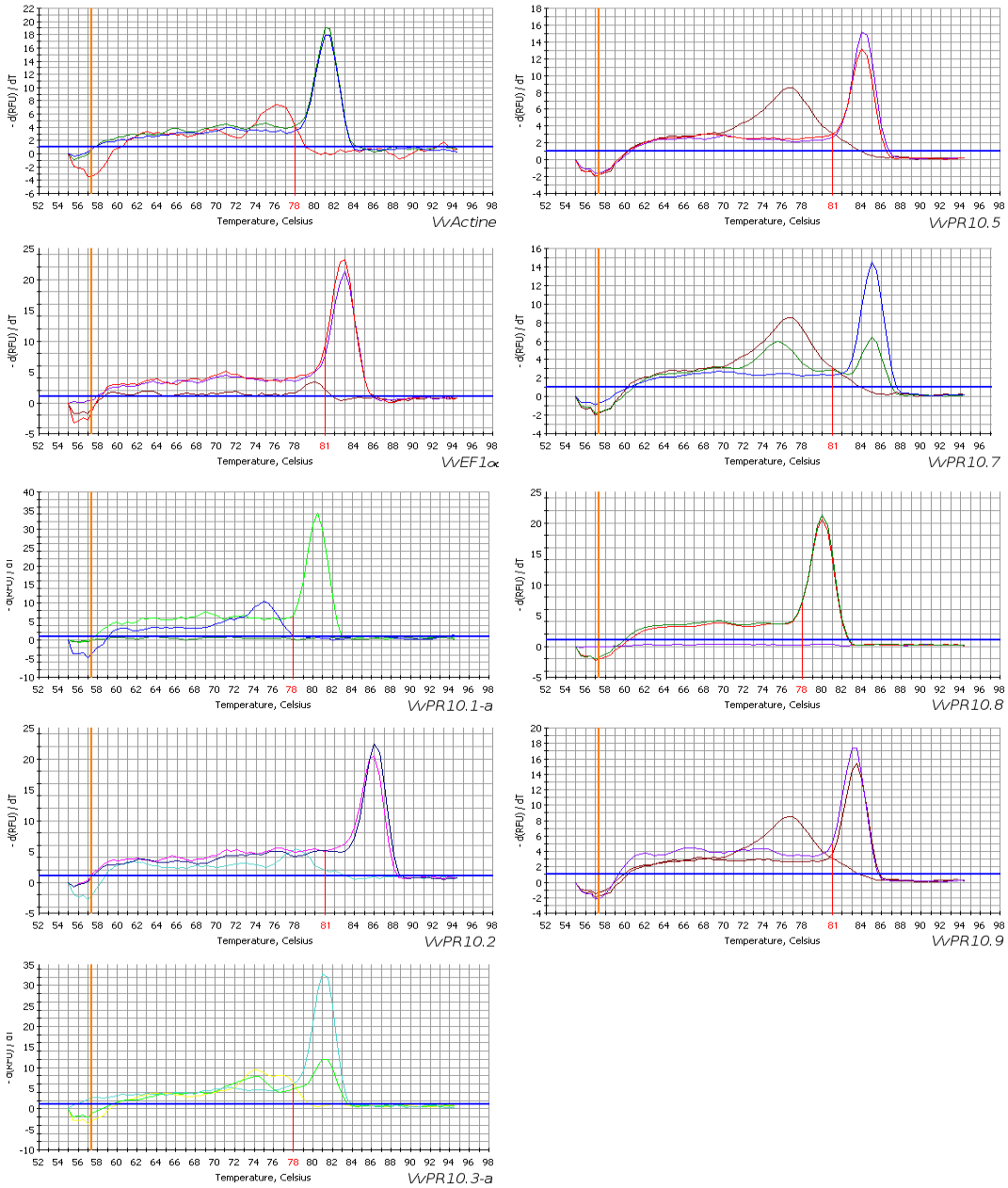


Figure 20 : Courbes de fusions obtenues avec les amorces *VvActine, VvEF1α, VvPR10.1-a* , *VvPR10.2, VvPR10.3-a, VvPR10.5, VvPR10.7, VvPR10.8* et *VvPR10.9.*
Les courbes représentent l'opposée de la dérivée de la fluorescence en fonction de la température, c'est-à-dire la disparition de la fluorescence en fonction de la température. Pour chaque gène, trois courbes sont réalisées. La première est obtenue à partir d'une réaction sans matrice et ne donne aucun pic, ou un pic aux températures les plus faibles. Les deux autres sont obtenues en utilisant un ADNc obtenu après transcription inverse à partir d'un ARNm extrait respectivement d'embryons et de cals après traitement 2,4-D. La température d'acquisition de signal choisie est indiquée en rouge pour chaque gène.

Le programme d'amplification était le suivant :

- Dénaturation initiale : 95°C pendant 30 s
- Dénaturation : 95° pendant 15 s
- Hybridation + elongation : 57°C pendant 15 s
- Acquisition du signal : X°C pendant 10 s

} 40 cycles

- Elongation finale : 72°C pendant 10 min
- Courbe de fusion : de 55°C à 95°C, pas de 0,5°C toutes les 10 s.

La température d'acquisition du signal dépend de la nature de l'amplicon.

Au départ, la spécificité des couples d'amorces a été vérifiée par la réalisation d'une courbe de fusion, après une amplification utilisant le programme de RT-qPCR, sans phase d'acquisition du signal. Les matrices utilisées pour ces courbes étaient des ADNc produits à partir de 2 types d'échantillon : embryons (E) et cals après traitement par le 2,4-D (C). Lorsque la température de dénaturation spécifique de l'amplicon est atteinte, une chute de fluorescence est enregistrée (Figure 20). Si des produits de PCR issus d'hybridations non spécifiques ou des dimères d'amorce sont présents, leur présence sera détectée par la formation, sur la courbe, d'un pic de perte de fluorescence distinct correspondant à leur propre température de dénaturation. Ces produits indésirables sont à éviter, puisqu'ils entraînent une perte d'efficacité de la réaction et éventuellement une mauvaise estimation de l'expression du gène d'intérêt. Néanmoins, les séquences des gènes de la famille *VvPR10* sont très proches les unes des autres, et la mise au point d'amorces spécifiques respectant les critères propre à la RT-qPCR laisse très peu de latitude. Des amorces spécifiques ont été mises au point pour chaque gène testé, mais il a été impossible d'éviter les dimères d'amorces visibles en l'absence de matrice (Figure 20). Les courbes de fusion permettent alors de choisir les températures d'acquisition du signal, entre la fin de la dénaturation de ces dimères et le début de la dénaturation des amplicons d'intérêt, pour permettre de réaliser la quantification. Deux températures d'acquisition furent déterminées : 78°C pour *VvActin*, *VvPR10.1-a*, *VvPR10.3-a* et *VvPR10.8* et 81°C pour *VvEF1*α, *VvPR10.2*, *VvPR10.5* et *VvPR10.7*.

d. Validation de la qPCR

Pour chaque gène lors de chaque expérience, trois contrôles ont été réalisés : un témoin sans matrice, une courbe de fusion et une gamme étalon.

Le témoin sans matrice a permis de vérifier la présence d'une contamination (par des acides nucléiques) et une variation de signal pouvait être parfois détectée dans ce témoin. Elle peut être ignorée si la différence de valeur de Cq entre ce témoin sans matrice et l'échantillon dont le Cq est le plus élevée est assez grande. Par exemple, une différence de Cq de cinq cycles est suffisante (D'haene et al 2010).

La courbe de fusion a permis de s'assurer que chaque réaction s'est déroulée correctement, le pic de dissociation spécifique devant apparaître pour chaque gène afin de vérifier que celui-ci a été correctement amplifié et qu'aucune amplification indésirable n'a eu lieu.

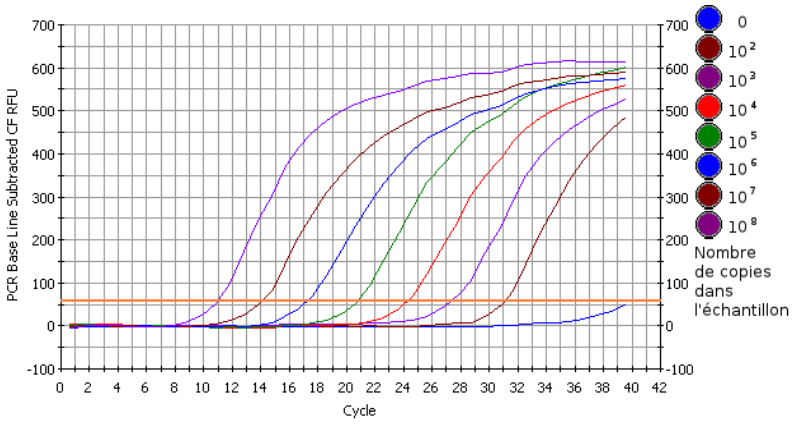
La gamme étalon a permis de valider les conditions de réaction d'une expérience à l'autre. Une gamme étalon représentant le Cq en fonction du nombre initial de copies a été réalisée pour chaque gène testé ainsi que pour les gènes de référence. Elle a permis d'une part de calculer l'efficacité de la PCR et d'autre part de déterminer le nombre initial de copies en fonction des valeurs de Cq.
Pour réaliser la gamme étalon, une amplification en RT-PCR « classique » a été effectuée pour chaque gène dans le but de disposer d'un grand nombre de copies. Après migration sur gel d'agarose, la bande correspondant au produit de PCR est découpée et purifiée. La concentration d'ADN, déterminée par spectrophotométrie, a permis d'évaluer le nombre de copies présentes dans l'échantillon purifié.

Le nombre de copies a été calculé de la manière suivante :

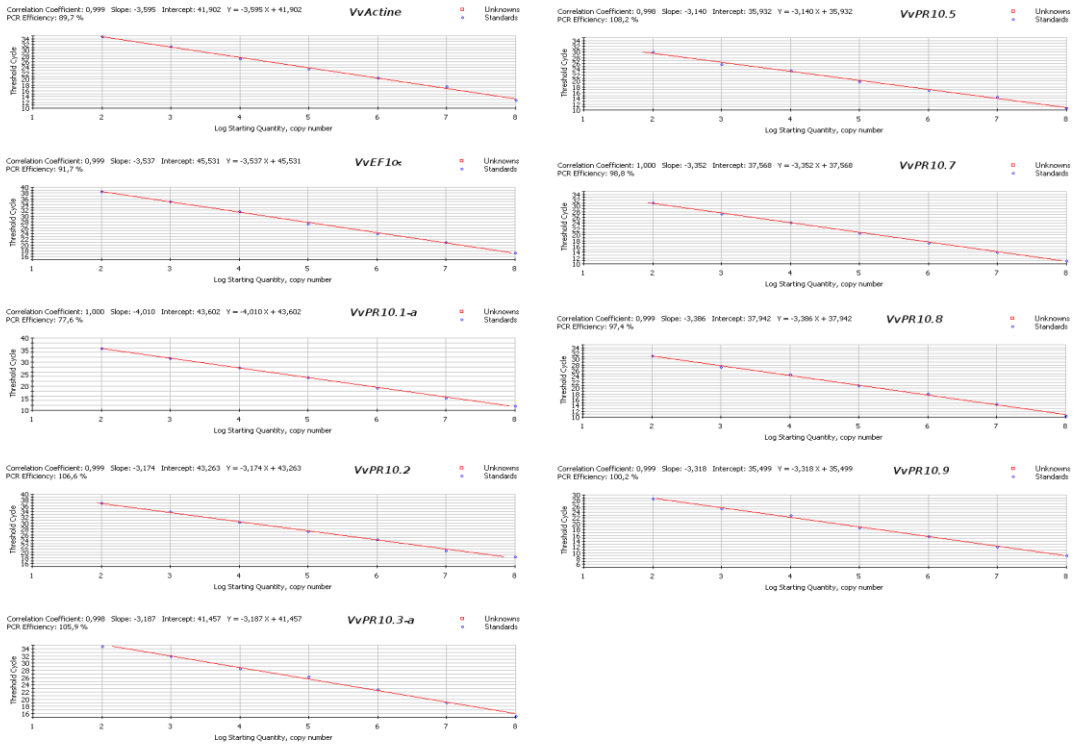
$$\text{Nombre de copies/}\mu\text{L} = \frac{6.10^{23}\ [\text{copies/mol}] \times \text{Concentration}\ [\text{g/}\mu\text{L}]}{\text{Masse moléculaire}\ [\text{g/mol}]}$$

$$\text{Masse moléculaire}\ [\text{g/mol}] = \text{nombre de paire de bases} \times 660\ \text{Daltons}\ [\text{pour une base}]$$

Le nombre a été ajusté à 10^8 copies pour 5 µl et constitue le premier point de la gamme étalon. Les autres ont été obtenus par dilutions successives au dixième jusqu'à 10^2 copies pour 5µl.



A



B

Figure 21 : Obtention des droites étalon.

A : Exemple de courbes d'amplification des différents points d'une gamme étalon. Chaque courbe représente la fluorescence enregistrée en fonction du nombre de cycles. La distance en nombre de cycles entre deux courbes consécutives est constate.

B : Droites étalons (Cq en fonction du log du nombre de copies initial de chaque transcrit) obtenues pour les gènes *VvActine, VvEF1α, VvPR10.1-a, VvPR10.2, VvPR10.3, VvPR10.5, VvPR10.7, VvPR10.8* et *VvPR10.9*. Pour chaque droite, le logiciel calcule le coefficient de corrélation, l'équation ainsi que l'efficacité de la réaction de RT-qPCR.

Dans la fourchette de concentrations en matrices utilisées pour l'établissement de la gamme étalon, la différence de Cq entre deux points successifs de la gamme est constante (Figure 21-A). Cela implique que les points représentant le Cq en fonction du logarithme du nombre initial de copies sont alignés sur une droite (Figure 21-B). Un coefficient de corrélation estime la qualité de la gamme étalon. Celui-ci est de 1 si tous les points mesurés sont alignés sur la droite et il diminue avec le nombre de points non alignés et la distance par rapport à la droite. Il doit rester supérieur à 0,990 pour que la gamme étalon soit utilisable, ce qui a toujours été le cas.

L'efficacité de la PCR peut être mesurée grâce à la pente de la droite étalon. En effet, plus la PCR sera efficace, plus la différence de Cq entre deux points successifs de la gamme étalon, et donc le coefficient directeur de la droite étalon, sera faible. La formule permettant de définir l'efficacité de la PCR est : $E=10^{(1/|\text{coefficient directeur de la droite étalon}|)}$.

Les valeurs des Cq des gènes d'intérêt doivent être comprises entre les limites supérieure et inférieure de la gamme étalon pour pouvoir être validées (Bustin et al. 2009). Dans cet intervalle, la droite étalon permet de lier pour chaque gène le Cq mesuré au nombre de copies initial.

La matrice utilisée pour la gamme étalon provenant de la même dilution d'une expérience à l'autre, les conditions de sa réalisation sont identiques. Une variation de plus d'un Cq entre deux mêmes points de la gamme étalon dans deux expériences différentes indique donc un problème survenu, soit au niveau de l'expérience, soit une dégradation au niveau de la matrice. Une telle variation n’a jamais été constatée dans les gammes réalisées.

e. Traitement des résultats

Les gènes de référence sont des gènes dont le niveau d'expression est stable dans les différents tissus étudiés sous les conditions analysés. Ils sont employés pour normaliser les résultats. Il recommandé d'employer au moins deux gènes de référence distincts, afin d'avoir une normalisation plus précise (Reid et al. 2006 , Hellemans et al. 2007). Les travaux de Reid et al. 2006 ont mis en évidence chez la vigne plusieurs gènes de référence potentiels, dont les plus indiqués sont *VvActin*, *VvEF1α*, *VvGADPH* ou *VvSAND*. Par souci d'homogénéité avec les expériences de RT-PCR semi-quantitative, *VvActin* fut sélectionné. Le second gène choisi fut *VvEF1α*, dont l'expression semblait parmi les plus stables dans les travaux de Reid et al (2006).

Le programme « MyiQ » (BioRad) a été utilisé pour traiter les données générées lors de la réaction de RT-qPCR.
Seuls les échantillons dont au moins deux mesures (sur les trois réplicats) différent au maximum d'un Cq ou moins ont été pris en compte. Les mesures correspondant à ce critère ont été utilisées pour établir un Cq moyen pour chaque échantillon. Les résultats ont ensuite été analysés grâce aux formules décrites par Hellemans et al. (2007). Les niveaux d'expression des gènes d'intérêt a ensuite été exprimé par rapport au niveaux d'expression moyens des deux gènes de référence.

E. Analyse des acides nucléiques sur gel d'agarose

1. ADN et produits amplifiés

L'ADN ou les produits de PCR ont été déposés sur gel d'agarose (Euromedex) 1,5 % (p/v) dissous dans un tampon d'électrophorèse TAE 1x (Tris Acetate EDTA). La migration par électrophorèse a été réalisée dans le même tampon (20 minutes, 110 Volts). Ils sont ensuite colorés au bromure d'éthidium et visualisés et photographiés en lumière UV par le système « Gel Doc Quantity One » (BioRad).

2. ARN

Les ARN été déposés sur gel d'agarose « RNase free » (Euromedex) 1 % (p/v) dissous dans un tampon d'électrophorèse TAE 1x « RNase free ». La migration par électrophorèse a été réalisée dans le même tampon TAE 1x (20 minutes, 110 Volts). (20 minutes, 110 Volts). Ils sont ensuite colorés au bromure d'éthidium et visualisés et photographiés en lumière UV par le système « Gel Doc Quantity One » (BioRad).

F. Séquençage des produits de PCR

Pour vérifier l'identité des produits de PCR, les échantillons ont été envoyés pour séquençage à la société Genoscreen.

III. Bioinformatique

A. Identification *in silico* des séquences des gènes *PR10* potentiels dans le génome de *Vitis vinifera*

Les gènes PR10 potentiels ont été identifiés grâce à la séquence complète du génome (8,4x équivalents génomiques) de la vigne (accession PN40024) disponible dans la base de données du site du GENOSCOPE (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/). Les séquences ont été trouvées grâce à l'outil « BLAT » (« BLAST (Basic Local Alignment Search Tool)-Like Alignment Tool », http://www.genoscope.cns.fr/blat-server/cgi-bin/vitis/webBlat) et comparées entre elles grâce à l'outil d'alignement « ClustalW », disponible sur le site de l'« European Bioinformatics Institute » (EBI, http://www.ebi.ac.uk/Tools/clustalw2). Les séquences issues du séquençage du génome de la vigne (Pinot Noir ENTAV 115) réalisé par l'« Istituto Agrario San Michele all'Adige » (IASMA) et qui correspondent aux séquences de gènes *PR10* potentiels identifiées, ont été recherchées grâce à l'outil « BLAST » (Basic Local Alignment Search Tool) du site du NCBI (http://blast.ncbi.nlm.nih.gov), puis comparées grâce à l'outil d'alignement « ClustalW ».
La composition des exons et des introns a été déterminée grâce à l'outil automatisé GAZE (Howe et al. 2002) sur le site du GENOSCOPE et corrigée manuellement lorsque c'était nécessaire, sur la base de la structure de *VvPR10.1* (Robert et al. 2001) et des séquences d'épissage consensus en 5' et 3' (NN/GT et AG/NN, respectivement).

B. Recherche d'« Expressed Sequence Tags » (ESTs) correspondant aux séquences *PR10* candidates de la vigne.

Les ESTs de vigne correspondant aux séquences *PR10* candidates identifiées ont été recherchées grâce au programme « BLASTn » (« BLAST nucléotide ») contre la base de données d'ESTs des espèces *Vitis* disponible sur le site du NCBI (http://blast.ncbi.nlm.nih.gov).

C. Prédiction des structures des gènes *PR10*

Les structures exon-intron de chaque séquence codant pour un gène PR10 potentiel ont été recherchées grâce au programme en ligne « Gene Finding » sur le serveur bioinformatique en ligne de l'« University of London » (http://mendel.cs.rhul.ac.uk/mendel.php). Les paramètres de la recherche ont été fixés sur « dicot ». La methode de recherche « FGENESH » a permis d'identifier les positions des boites TATA, des premier et dernier nucléotide de chaque exon, ainsi que du signal de polyadénilation. La méthode « TSSP-TCM » a permis de localiser le site d'initiation de la transcription.

D. Analyse des séquences protéiques

Les séquences codantes putatives (CDS) ont été transcrites grâce au programme en ligne « Transeq », sur le site de l'EBI (http://www.ebi.ac.uk/Tools/emboss/transeq/). Leur poids moléculaire (PM) et leur point isoélectrique (pI) ont été déterminés grâce au programme « COMPUTE » du site « Expert Protein Analysis System » (ExPASy, http://www.expasy.ch/tools/pi_tool.html). Sur ce site, la base de données « PROSITE » a été utilisée pour retrouver la signature Bet v 1 (http://www.expasy.ch/prosite). Puis, les structures tridimensionnelles de protéines putatives PR10 ont été prédites sur un serveur automatisé de modélisation protéique utilisant « SWISS-MODEL », en utilisant les paramètres par défaut (http://swissmodel.expasy.org/).

E. Analyse phylogénétique et alignements multiples

Les arbres phylogénétiques ont été construits en utlisant un méthode « Neighbor-Joining » avec le programme « Phylo_win » (Galtier et al. 1996). Les valeurs de « bootstrap » ont été déterminées pour 500 réplicats. Les comparaisons entre les séquences nucléotidiques et protéiques ont été réalisées grâce à l'outil d'alignement « ClustalW ».

Références bibliographiques

Adams KL, Wendel JF: **Polyploïdy and genome evolution in plants.** *Curr Opin Plant Biol* 2005, **8**:135-141

Albertini E, Marconi G, Reale L, Barcaccia G, Porceddu A, Ferranti F, Falcinelli M: ***SERK* and *APOSTART*. Candidate genes for apomixis in *Poa pratensis*.** *Plant Physiol* 2005, **138:**2185-2199

Albrecht C, Russinova E, Hecht V, Baaijens E, de Vries S: **The Arabidopsis thaliana SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis.** *Plant Cell* 2005, **17:**3337-3349

Alemanno L, Devic M, Niemenak N, Sanier C, Guilleminot J, Rio M, Verdeil JL, Montoro P: **Characterization of *leafy cotyledon1-like* during embryogenesis in *Theobroma cacao* L**, *Planta* 2008, **227:**853-866

Antoniw JF, Ritter CE, Pierpoint WS, Van Loon LC: **Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV.** *J Gen Virol* 1980, **47:**79-87

Arya M, Shergill IS, Williamson M, Gommersall L, Arya N, Patel HRH: **Basic principles of real-time quantitative PCR.** *Expert Rev Mol Diagn* 2005, **5(2):**209-219

Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF: **Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development.** *Plant J* 1996, **9(5):**745-753

Bantignies B, Séguin J, Muzac I, Dédaldéchamp F, Gulick P, Ibrahim R: **Direct evidence for ribonucleolytic activity of a PR-10-like protein from white lupin roots.** *PlantMol Biol* 2000, **42**:871-881

Baudino S, Hansen S, Brettschneider R, Hecht VFG, Dresselhaus T, Lörz H, Dumas C, Rogowsky PM: **Molecular characterisation of two novel maize LRR receptor-like kinases, which belong to the *SERK* gene family.** *Planta* 2001, **213:**1-10

Baumgarten A, Cannon S, Spangler R, May G: **Genome-level evolution of resistance genes in *Arabidopsis thaliana.*** *Genetics* 2003, **165:**309-319

Beuning LL, Bowen JH, Persson HA, Barraclough D, Bulley S, MacRae EA: **Characterization of Mal D 1-related genes in *Malus*.** *Plant Mol Biol* 2004, **55**:369-388

Biesiadka J, Bujacz G, Sikorski MM, Jaskolski M.: **Crystal structures of two homologous pathogenesis-related proteins from yellow lupine.** *J Mol Biol* 2002, **319**:1223-1234

Blanc G, Hokamp K, Wolfe KH: **A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome.** *Genome Res* 2003, **13:**137-144

Blanc G, Wolfe KH: **Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes.** *Plant Cell* 2004, **16:**1667-1678

Bowers JE, Chapman BA, Rong J, Paterson AH: **Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events.** *Nature* 2003, **422:**433-438

Boyle B, Brisson N: **Repression of the defense gene *PR-10a* by the single-stranded DNA binding protein SEBF.** *Plant Cell* 2001, **13:**2525-2537

Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, Schneider O, Breitenbach M: **The gene coding fort he major birch pollen allergen *BetvI*, is highly homologous to a pea disease resistance response gene.** *EMBO* 1989, **8(7)**:1935-1938

Busam G, Kassemeyer HH, Matern U: **Differential expression of chitinases inVitis vinifera L. responding to systemic acquired resistance activators or fungalchallenge.** *Plant Physiol* 1997, **115:**1029-1038

Bustin SA, Nolan T: **Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction.** *J Biomol Tech* 2004, **15(3):**155-166

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments.** *Clin Chem* 2009 , **55(4):**611-622

Bustin SA: **Why the need for qPCR publication guidelines?-The case for MIQE.** *Methods* 2010, **50:**217-226

Cannon SB, Mitra A, Baumgarten A, Young ND, May G: **The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana.*** *BMC Plant Biol* 2004,**4**:10

Castro AJ, Carapito C, Zorn N, Magné C, Leize E, Van Dorsselaer A, Clément C: **Proteomic analysis of grapevine (*Vitis vinifera L.*) tissues subjected to herbicide stress.** *J Exp Bot* 2005, **56(421)**:2783-2795

Castro AJ, Saladin G, Bézier A, Mazeyrat-Gourbeyre F, Baillieul F, Clément C: **The herbicide flumioxazin stimulates pathogenesis-related gene expression and enzyme activities in *Vitis vinifera*.** *Physiol Plantarum* 2008, **134(3):**453-463

Chadha P, Das RH: **A pathogenesis related protein, AhPR10 from peanut : an insight of its mode of antifungal activity.** *Planta* 2006, **225**:213-222

Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-Berhan A, Jones KM, Redman J, Chen G, Cahoon EB, Gedil M, Stanke M, Haas BJ, Wortman JR, Fraser-Liggett CM, Ravel J, Rabinowicz PD: **Draft genome sequence of the oilseed species *Ricinus communis*.** *Nature Biotech* 2010, **28:**951-956

Chen L, Zhang S, Illa E, Song L, Wu S, Howard W, Arús P, van de Weg E, Chen K, Gao Z: **Genomic characterization of putative allergen genes in peach/almond and their synteny with apple.** *BMC Genom* 2008, **9**:543

Chong J, Le Henanff G, Bertsch C, Walter B: **Identification, expression analysis and characterization of defense signaling genes in Vitis vinifera.** *Plant Physiol Biochem* 2008, **46:**469-481

Christensen AB, Cho BH, Naesby M, Gregersen PL, Brandt J, Madriz-Ordeñana K, Collinge DB, Thordal-Christensen H: **The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins.** *Mol Plant Pathol* 2002, **3(3):**135-144

Chugh A, Khurana P: **Gene expression during somatic embryogenesis – recent advances.** *Curr Science* 2002, **83:**715-730

Colcombet J, Boisson-Dernier A, Ros-Palau R,. VeraCE, Schroeder JI: A**rabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are essential for tapetum development and microscope maturation.** *Plant Cell* 2005, **17:**3350-3361

Colditz F, Niehaus K, Krajinski F: **Silencing of PR-10-like proteins in *Medicago truncatula* results in an antagonistic induction of other PR proteins and in an increased tolerance upon infection with the oomycete *Aphanomyces euteiches.*** *Planta* 2007, **226(1):**57-71

Croce P, Vannini C, Bracale M: **An easy and convenient method for maintenance of embryogenic cultures of *Vitis vinifera.*** *Vitis* 2005, **44:**197-198

Cronn RC, Adams KL: **Quantitative analysis of transcript accumulation from genes duplicated by polyploidy using cDNA-SSCP.** *Biotechniques* 2003, **34(4):**726-730

D'haene B, Vandesompele J, Hellemans J: **Accurate and objective copy number profiling using real-time quantitative PCR.** *Methods* 2010, **50:**262-270

Das DK, Reddy MK, Upadhyaya KC, Sopory SK: **An efficient leaf-disc culture method for the regeneration via somatic embryogenesis and transformation of grape (*Vitis vinifera* L.).** *Plant Cell Rep* 2020, **20:**999-1005

De Bodt S, Maere S, Van de Peer Y: **Genome duplication and the origin of angiosperms.** *Trends Ecol Evol* 2005, **20(11):**591-597

De Jong AJ, Cordewener J, Lo Schiavo F, Terzi M, Vandekerckhove J, Van Kammen A, De Vries SC: **A carrot somatic embryo mutant is rescued by chitinase.** *Plant Cell* 1992, **4:**425-433

De Jong F, Mathesius U, Imin N, Rolfe BG: **A proteome study of the proliferation of cultured *Medicago tuncatula* protoplasts.** *Proteomics* 2007, **7:**722-736

Deloire A, Kraeva E, Mauro MC, Bonnet E, Bierne J:, **Immunodetection of chitinase-like and b-1,3-glucanase-like proteins secreted in vitro by embryogenic and non-embryogenic cells of grapevines.** *Vitis* 1997, **1:**51-52

Devos D, Valencia A: Intrinsic errors in genome annotation. *Trends Genet* 2001, **17(8):**429-431

Dron M, Clouse SD, Dixon R, Lawton M, Lamb C: **Gluthatione and fungal elicitor regulation of a plant defense gene promoter in electroporated protoplasts.** *Proc Natl Acad Sci USA* 1988, **85:**6738-6742

Domoki M, Györgyey J, Biró J, Pasternak TP, Zvara Á, Bottka S, Puskás LG, Dudits D, Fehér A: **Identification and characterization of genes associated with the induction of embryogenic competence in leaf-protoplast-derived alfalfa cells.** *Biochim Biphys Acta* 2006, **1759:**543-551

Duarte JM, Cui L, Kerr Wall P, Zhang Q, Zhang X, Leebens-Mack J, Ma H, Altman N, dePamphilis CW: **Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis.*** *Mol Biol Evol* 2006, **23(2):**469-478

Dubos C, Plomion C: **Identification of water-deficit responsive genes in maritime pine (*Pinus pinaster* Ait.) roots.** *Plant Mol Biol* 2001, **51(2):**249-262

Edreva A: **Pathogenesis-related proteins: research progress in the last 15 years.** *Gen Appl Plant Physiology* 2005, **31(1-2):**105-124

Ekramoddoullah AKM, Yu, X, Sturrock R, Zamani A, Taylor D: **Detection and seasonal expression pattern of a pathogenesis-related protein (PR-10) in Douglas-fir (*Pseudotsuga menziesii*) tissues.** *Physiol Plantarum* 2000, **110:**240-247

Fambrini M, Durante C, Cionini G, Geri C, Giorgetti L, Michelotti V, Salvini M, Pugliesi C: **Characterization of *LEAFY COTYLEDON1-LIKE* gene in Helianthus annuus and its relationship with zygotic and somatic embryogenesis.** *Dev Genes Evol* 2006, **216:**253-264

Federico ML, Iñiguez-Luy FL, Skadsen RW, Kaeppler HF: **Spatial and temporal divergence of expression in duplicated barley germin-like protein-encoding genes.** *Genetics* 2006, **174:**179-190

Fehér A, Pasternak TP, Dudits D: **Transition of somatic plant cells to an embryogenic state.** *Plant Cell Tissue Organ Cult* 2003, **74**:201-228

Fehér A: **The initiation of somatic embryogenesis: what we know and what we don't.** *Acta Biol Szeg* 2008, **52(1):**53-56

Fernandes H, Pasternak O, Bujacz G, Bujacz A, Sikorski MM, Jaskolski M: ***Lupinus luteus* Pathogenesis-Related protein as a reservoir for cytokinin.** *J Mol Biol* 2008, **378(5)**:1040-1051

Ferre F: **Quantitative or Semi-Quantitative PCR: Reality Versus Myth.** *Genome Res* 1992, **2:**1-9

Ferri M, Tassoni A, Franceschetti M, Righetti L, Naldrett MJ, Bagni N: **Chitosan treatment induces changes of protein expression profile and stilbene distribution in *Vitis vinifera* cell suspensions.** *Proteomics* 2009, **9**:610-624

Finkler C, Giacomet C, Muschner VC, Salzano FM, Freitas LB: **Molecular investigations of pathogenesis-related Bet v 1 homologues in *Passiflora* (Passifloraceae).** *Genetica* 2005, **124**:117-125

Flagel LE, Wendel JF: **Gene duplication and evolutionary novelty in plants.** *New Phytol* 2009, **183**:557-564

Freitas LB, Bonatto SL, Salzano FM: **Evolutionary implications of intra- and interspecific molecular variability of pathogenesis-related proteins.** *Braz J Biol* 2003, **63(3)**:437-448

Fung RWM, Gonzalo M, Fekete C, Kovacs LG, He Y, Marsh E, McIntyre LM, Schachtman DP, Qiu W: **Powdery mildew induces defense-oriented reprogramming of the transcriptome in a susceptible but not in a resistant grapevine.** *Plant Physiol* 2008, **146**:236-249

Gachon C, Migam A, Charrier B: **Real-time PCR: what relevance to plant studies?** *J Exp Bot* 2004, **55:**1445-1454

Gaj MD: **Direct somatic embryogenesis as a rapid and efficient system for in vitro regeneration of *Arabidopsis thaliana* (L.) Heynh.** *Plant Cell Tissue Organ Cult* 2001, **64:**39-46

Gaj MD: **Factors influencing somatic embryogenesis induction and plant regeneration with particular reference to *Arabidopsis thaliana* (L.) Heynh.** *Plant Growth Regul* 2004, **43:**27-47

Gajhede M, Osmark K, Poulsen FM, Ipsen H, Larsen JN, Joost van Neerven PJ, et al.: **X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy.** *Nat Struct Biol* 1996, **3**:1040-1045.

Galtier N, Gouy M, Gautier C: **SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny.** *Comput Appl Biosci* 1996, **12**:543-548

Gao ZS, van de Weg WE, Schaart JG, Schouten HJ, Tran DH, Kodde LP, van de Meer IM, van der Geest AHM, Kodde J, Breiteneder H, Hoffmann-Sommergruber K, Bosch D, Gilissen LJWJ: **Genomic cloning and linkage mapping of the *Mal d 1* (*PR-10*) gene family in apple (*Malus domestica*).** *Theor Appl Genet* 2005, **111**:171-183

Garavaglia BS, Thomas L, Gottig N, Dunger G, Garofalo CG, Daurelio LD, Ndimba B, Orellano EG, Gehring C, Ottado J: **A eukaryotic-acquired gene by a biotrophic phytopathogen allows prolonged survival on the host by couteracting the shut-down of plant photosynthesis.** *PloS One* 2010, **5(1):**e8950

Ha M, Kim ED, Chen J: **Duplicate genes increase expression diversity in closely related species and allopolyploids.** *PNAS* 2009, **106(7):**2295-2300

Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu SH: **Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli.** *Plant Physiol* 2008, **148**:993-1003

Handschuh L, Femiak I, Kasperska A, Figlerowicz M, Sikorski MM: **Structural and functional characteristics of two novel members of pathogenesis-related multigene family of class 10 from yellow lupine.** *Acta Bioch Pol* 2007, **54(4)**:783-796

Harada JJ: **Role of *Arabidopsis LEAFY COTYLEDON* genes in seed development.** *J Plant Physiol* 2001, **158:**405-409

Hashimoto M, Kisseleva L, Sawa S, Furukawa T, Komatsu S, Koshiba T: **A novel rice PR10 protein, RSOsPR10, specifically induced in roots by biotic and abiotic stress, possibly via the jasmonic acid signalling pathway.** *Plant Cell Physiol* 2004, **45:**550-559

He K, Gou X, Yuan T, Lin H, Asami T, Yoshida S, Russell SD, Li J: **BAK1 and BKK1 regulate brassinosteroid-dependent growth and brassinosteroid-independent cell-death pathways.** *Curr Biol* 2007, **17:**1109-1115

Hecht V, Vielle-Calzada JP, Hartog MV, Schmidt DL, Boutilier K, Grossnicklaus U, de Vries SC: **The Arabidopsis *SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1* gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture.** *Plant Physiol* 2001, **127:**803-816

Heese A, Hann DR, Gimenez-Ibanez S, Jones AME, He K, Li J, Schroeder JI, Peck SC, Rathjen JP, **The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants.** *Proc Natl Acad Sci USA* 2007, **104**:12217-12222

Helleboid S, Bauw G, Belingheri L, Vasseur J, Hilbert JL: **Extracellular b-1,3-glucanases are induced during early somatic embryogenesis in *Cichorium*.** *Planta* 1998, **205:**56-63

Helleboid S, Hendriks T, Bauw G, Inzé D, Vasseur J, Hilbert JL: **Three major somatic embryogenesis related proteins in *Chicorium* identified as PR proteins.** *J Exp Bot* 2000, **51:**1189-1200

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J: **qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data.** *Genome Biol* 2007, **8:**R19

Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O, Breiteneder H: **Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions.** *Gene* 1997, **197**:91-100

Howe KL, Chothia T, Durbin R: **GAZE: A generic framework for the integration of gene-prediction data by dynamic programming.** *Genome Res* 2002,**12:**1418-1427

Hu H, Xiong L, Yang Y: **Rice SERK1 gene positively regulates somatic embryogenesis of cultured cells and host defense response against fungal infection.** *Planta* 2005, **222:**107-117

Huang JC, Chang FC, Wang CS: **Characterization of a lily tapetal transcript that shares sequence similarity with a class of intracellular pathogenesis-related (IPR) proteins.** *Plant Mol Biol* 1997, **34:**681-686

Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P, Ren Y, Zhu2, Jun Li H, Lin K, Jin W, Fei Z, Li G, Staub J, Kilian A, van der Vossen EAG, Wu Y, Guo J, He J, Jia Z, Ren Y, Tian G, Lu Y, Ruan J, Qian W, Wang M, Huang Q, Li B, Xuan Z, Cao J, Asan, Wu Z, Zhang J, Cai Q, Bai Y, Zhao B, Han Y, Li Y, Li X, Wang S, Shi Q, Liu S, Cho WK, Kim JY, Xu Y, Heller-Uszynska K, Miao H, Cheng Z, Zhang S, Wu J, Yang Y, Kang H, Li M, Liang H, Ren X, Shi Z, Wen M, Jian M, Yang H, Zhang G, Yang Z, Chen R, Liu S, Li J, Ma L,12, Liu H, Zhou Y, Zhao J, Fang X, Li G, Fang L, Li Y, Liu D, Zheng H, Zhang Y, Qin N, Li Z, Yang G, Yang S, Bolund L, Kristiansen K, Zheng H, Li S, Zhang X, Yang H, Wang J, Sun R, Zhang B, Jiang S, Wang J, Du1 Y, Li S: **The genome of the cuncumber,*Cucumis sativus* L.** *Nature Genetics* 2009, **41:**1275-1281

Huggett J, Dheda K, Bustin S, Zumla A: **Real-time RT-PCR normalisation; strategies and considerations.** *Genes Immun* 2005, 1-6

Hurles M: **Gene duplication: the genomic trade in spare parts.** *PloS Biol* 2004, **2(7):**900-904

Huynen MA, van Nimwegen E: **The frequency distribution of gene family sizes in complete genomes.** *Mol Biol Evol* 1998, **15(5):**583-589

Hwang SE, Lee IA, Yie SW, Hwang DJ: **Identification of a *OsPR10a* promoter region responsive to salicylic acid.** *Planta* 2008, **227:**1141-1150

Imin N, Nizamidin M, Daniher D, Nolan KE, Rose RJ, Rolfe BG: **Proteomic analysis of somatic embryogenesis in *Medicago truncatula.* Explant cultures grown under 6-benzylaminopurine and 1-naphtaleneacetic acid treatments.** *Plant Physiol* 2005, **137**:1250-1260

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quétier F, Wincker P: **The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla.** *Nature* 2007, **449:**463-468

Jakoby M, Weisshaar B, Dröge-Laser W, Vincente-Carbajosa J, Tiedermann J, Kroj T, Parcy F: **bZIP transcription factors in *Arabidopsis.*** *Trends Plant Sci* 2002, **7(3):**106-111

Jain S, Srivastava S, Sarin NB, Kav NNV: **Proteomics reveals elevated levels of PR 10 proteins in saline-tolerant peanut (*Arachis hypogaea*) calli.** *Plant Physiol Biochem* 2006, **44(4):**253-259

Jellouli N, Ben Jouira H, Skouri H, Ghorbel A, Gourgouri A, Mliki A: **Proteomic analisys of Tunisian grapevine cultivar Razegui under salt stress.** *J Plant Physiol* 2008, **165**:471-481

Jellouli N, Ben Jouira H, Daldoul S, Chenennaoui S, Ghorbel A, Ben Salem A, Gargouri A: **Proteomic and transcriptomic analysis of grapevine PR10 expression during salt stress and functionnal characterization in yeast.** *Plant Mol Biol Rep* 2010, **28(1):**1-8

Jiménez VM: **Involvment of plant hormones and plant growth regulators on *in vitro* somatic embryogenesis.** *Plant Growth Regul* 2005, **47:**91-110

Karami O, Saidi A: **The molecular basis for stress-induced acquisition of somatic embryogenesis.** *Mol Biol Rep* 2010, **37(5):**2493-2507

Kasprzewska A: **Plant chitinases–regulation and function.** *Cell Mol Biol Lett* 2003, **8:**809–824

Kav NTV, Srivastava S, Goonewardene L, Blade SF: **Proteome-level changes in the roots of *Pisum sativum* in response to salinity.** *Ann Appl Biol* 2004, **145(2):**217-230

Kemmerling B, Schwedt A, Rodriguez P, Mazzotta S, Frank M, Qamar SA, Mengiste T, Betsuyaku S, Parker JE, Müssig C, Thomma BPHJ, Albrecht C, de Vries S, Hirt H, Nürnberger T: **The BRI1-Associated Kinase 1, BAK1, has a brassinolideindependent role in plant cell-death control.** *Curr Bio* 2007, **17:**1116-1122

Kim ST, Cho KS, Yu S, Kim SG, Hong C, Bae DW, Nam MH, Kang KY: **Proteomic analysis of differentially expressed proteins induced by rice blast fungus and elicitor in suspension-cultured rice cells.** *Proteomics* 2003, **3:**2368-2378

Kim SJ, Kim MR, Bedgar DL, Moinuddin SGA, Cardenas CL, Davin LB, Kang CH, Lewis NG: **Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*.** *PNAS* 2004, **101(6):**1455-1460

Kim ST, Yu S, Kang YH, Kim SG, Kim JY, Kim SH, Kang KY: **The rice pathogen-related protein 10 (JIOsPR10) is induced by abiotic and biotic stresses and exhibits ribonuclease activity.** *Plant Cell Rep* 2008, **27**:593-603

Koistinen KM, Soininen P, Vanäläinen TA, Häyrinen J, Laatikainen R, Peräkylä M, Tervahauta AI, Kärenlampi SO: **Birch PR-10c interacts with several biologically important ligands.** *Phytochem* 2005, **66:**2524-2533

Kortekamp A: **Expression analysis of defence-related genes in grapevine leaves after inoculation with a host and a non-host pathogen.** *Plant Physiol Biochem* 2006, **44:**58-67

Kursteiner O, Dupuis I, Kuhlemeier C: **The *Pyruvate decarboxylase1* gene of *Arabidopsis* is reuired during anoxi but not other environmental stresses.** *Plant Physiol* 2003, **132:**968-978

Kwaaitaal MACJ, de Vries SC: **The *SERK1* gene is expressed in procambium and in immature vascular cells.** *J Exp Bot* 2007, **58:**2887-2896

Kwong RW, Bui AQ, Lee H, Kwong LW, Fischer RL, Goldberg RB, Harada JJ: **LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development.** *Plant Cell* 2003, **15:**5-18

Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E: **Functional analysis of regulatory sequences controlling *PR-1* gene expression in *Arabidopsis*.** *Plant J* 1988, **16:**223-233

Lee D, Redfern O, Orengo C: **Predicting protein function from sequence and structure.** *Nature Rev Mol Cell Biol* 2007, **8:**995-1005

Lefever S, Hellemans J, Pattyn F, Przybylski DR, Taylor C, Geurts R, Untergasser A, Vandesompele J: **RDML: structured language and reporting guidelines for real-time quantitative PCR data.** *Nucleic Acids Res* 2009, **37:**2065-2069

Legrand S, Hendricks T, Hilbert JL, Quillet MC: **Characterization of expressed sequence tags obtained by SSH during somatic embryogenesis in *Cichorium intybus* L.** *BMC Plant Biol* 2007, **7:**27

Leitch IJ, Bennett MD: **Polyploidy in angiosperms.** *Trends Plant Sci* 1997, **2(12):**470-476

Li J, Wen J, Lease KA, Doke JT, Tax FE, Walker JC: **BAK1, an *Arabidopsis* LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling.** *Cell* 2002, **110:** 213-222

Liao D: **Molecular evolution '99 Concerted evolution: molecular mecanism and biological implications.** *Am J Hum Genet* 1999, **64:**24-30

Liu JJ, Ekramoddoullah AKM: **Root-specific expression of a western white pine PR10 gene is mediated by different promoter regions in transgenic tobacco.** *Plant Mol Biol* 2003, **52**:103-120

Liu JJ, Ekramoddoullah AKM: **Characterization, expression and evolution of two novel subfamilies of Pinus monticola cDNAs encoding pathogenesis-related (PR)-10 proteins.** *Tree Physiol* 2004, **24**:1377-1385.

Liu JJ, Ekramoddoullah AKM: **The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses.** *Physiol Mol Plant Pathol* 2006, **68**:3-13

Liu X, Huang B, Lin J, Fei J, Chen Z, Pang Y, Sun X, Tang K: **A novel pathogenesis-related protein (SsPR10) from *Solanum surattense* with ribonucleotic and antimicrobial activity is stress- and pathogen-inducible.** *J Plant Physiol* 2006, **163**:546-556

Lockton S, Gaut BS: **Plant conserved non-coding sequences and paralogue evolution.** *Trends Genet* 2005, **21(1):**60-65

Lynch M, Force A: **The probability of duplicate gene preservation by subfunctionalization.** *Genetics* 2000, **154:**459-473

Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y: **Modeling gene and genome duplications in eukaryotes.** *PNAS* 2005, **102(15):**5454-5459

Maillot P, Kieffer F, Walter B: **Somatic embryogenesis from stem nodal sections of grapevine.** *Vitis* 2006, **45:**185-189

Maillot P, Lebel S, Dchellenbaum P, Jacques A, Walter B : **Differential regulation of *SERK, LEC1-Like* and *Pathogenesis-Related* genes during indirect secondary somatic embryogenesis in grapevine.** *Plant Physiol Biochem* 2009, **47:**743-752

Maillot P, Walter B, Schellenbaum P: **Somatic Embryogenesis Receptor-like Kinases are related to embryogenesis and other developmental and defence pathways**, in *Plant genetic transformation and molecular markers* 2010, pp.149-167

Malandain H, Lavaud F: **Allergénicité des protéines de défense végétale.** *Rev Fr Allergol Immunol Clin* 2004, **44(5):**469-475

Marković-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O, Breiteneder H: **Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier.** *J Mol Biol* 2003, **325**:123-133

Marković-Housley Z, Basle A, Padavattan S, Maderegger B, Schirmer T, Hoffmann-Sommergruber K: **Structure of the major carrot allergen Dau c 1.** *Acta Cryst* 2009, **D65:**1206-1212

Marsoni M, Bracale M, Espen L, Prinsi B, Negri AS, Vannini C: **Proteomic analysis of somatic embryogenesis in *Vitis vinifera*.** *Plant Cell Rep* 2008, **27**:347-356

Martinelli L, Candioli E, Costa D, Poletti V, Rascio N: **Morphogenic competence of Vitis rupestris S. secondary somatic embryos with a long culture history.** *Plant Cell Rep* 2001, **20:**279-284

Martinelli L, Gribaudo I, Bertoldi D, Candioli E, Poletti V: **High efficency somatic embryogenesis and plant germination in grapevine cultivars Chardonnay and Brachetto a grappolo lungo.** *Vitis* 2001, **40:**111-115

Martinelli L, Mandolino G: **Transgenic grapes (Vitis species**, in: Y.P.S. Bajaj (Ed.), *Biotechnology in Agriculture and Forestry*, Springer-Verlag, Berlin 2001, pp.325-338.

Martinelli L, Gibaudo I: **Strategies for effective somatic embryogenesis in grapevine: an appraisal.** *Grapevine molecular physiology & biotechnology* 2009, 461-493

McGee JD, Hamer JE, Hodges TK: **Characterization of a *PR-10 p*athogenesis-related gene family induced in rice during infection with *Magnaporthe grisea.*** *Mol Plant-Mic Int* 2001, **14(7):**877-886

Mogensen JE, Wimmer R, Larsen JN, Spangfort MD, Otzen DE: **The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands.** *J Biol Chem* 2002, **277(26)**:23684-23692

Moiseyev GP, Fedoreyvera LI, Zhuravlev YN, Yasnetskaya E, Jekel PA, Beintema JJ: **Primary structures of two ribonucleases from ginseng calluses. New members of the**

PR-10 family of intracellular pathogenesis-related plant proteins. *FEBS Lett* 1997, **407**:207-210

Moore RC, Purugganan MD: **The evolutionary dynamics of plant duplicate genes.** *Curr Opin Plant Biol* 2005, **8:**122-128

Morgana C, Di Lorenzo R, Carimi F: **Somatic embryogenesis of *Vitis vinifera* L. (cv. Sugraone) from stigma and style culture.** *Vitis* 2004, **43:**169-173

Murashige T, Skoog F: **A revised medium for rapid growth and bio assays with tobacco tissue cultures.** *Physiol Plantarum* 1962, **15(3):**473-497

Musidlowska-Persson A, Alm R, Emanuelsson C: **Cloning and sequencing of the Bet v 1-homologous allergen Fra a 1 in strawberry (*Fragaria ananassa*) shows the presence of an intron and little variability in amino acid sequence.** *Mol Immunol* 2007, **44**:1245-1252

Nair RR, Dutta Gupta S: **High-frequency plant regeneration through cyclic secondary somatic embryogenesis in black pepper (*Piper nigrum*) L.** *Plant Cell Rep* 2006 **24:**699-707

Nam KH, Li J: **BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling.** *Cell* 2002, **110:**203-212

Nei M, Rooney AP: **Concerted and birth-and-death evolution of multigene families.** *Annu Rev Genet* 2005, **39:**121-152

Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H, Rösch P: **Allergic cross-reactivity made visible.** *J Biol Chem* 2001, **276(25)**:22756-22763

Nolan KE, Irwanto RR, Rose RJ: **Auxin up-regulates *MtSERK1* expression in both *Medicago truncatula* root-forming and embryogenic cultures.** *Plant Physiol* 2003, **133:**218-230

Ohno S: **Evolution by gene duplication.** New York, NY: Springer, 1970

Ohta T: **Evolution of gene families.** *Gene* 2000, **159:**45-52

Osmark P, Boyle B, Brisson N: **Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins.** *Plant Mol Biol* 1998, **38**:1243-1246

Park CJ, Kim KJ, Shin R, Park JM, Shin YC, Paek KH: **Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway.** *Plant J* 2004, **37**:186-198

Pasternak O, Biesiadka J, Dolot R, Handschuh L, Bujacz G, Sikorski MM, Jaskolski M: **Structure of a yellow lupin pathogenesis-related PR-10 protein belonging to a novel subclass.** *Acta Cryst* 2005, **61(1):**99-107

Pastori GM, Foyer CH: **Common components, networks, and pathways of cross-tolerance to stress. The central role of « Redox » and abscisic acid-mediated controls.** *Plant Physiol* 2002, **129:**460-468

Pérez-Nuñez MT, Souza R, Sáenz L, Chan JL, Zuñiga-Aguilar JJ, Oropeza C: **Detection of a *SERK*-like gene in coconut and analysis of its expression during the formation of embryogenic callus and somatic embryos.** *Plant Cell Rep* 2009, **28:**11-19

Perrin M, Gertz C, Masson JE: **High efficiency initiation of regenerable embryogenic callus from anther filaments of 19-grapevine genotypes grown worldwide.** *Plant Sci* 2004, **167:**1343-1349.

Polesani M, Desario F, Ferrarini A, Zamboni A, Pezzotti M, Kortekamp A, Polverari A: **cDNA-AFLP analysis of plant and pathogen genes expressed in grapevine infected with *Plasmopara viticola.*** *BMC Genom* 2008, **9**:142

Polesani M, Bortesi L, Ferrarini A, Zamboni A, Fasoli M, Zadra C, Lovato A, Pezzotti M, Delledonne M, Polverari A: **General and species-specific transcriptional responses to downy mildew infection in a susceptible (*Vitis vinifera*) and a resistant (*V. riparia*) grapevine species.** *BMC Genomics* 2010, **11:**117

Proost S, Van Bel M, Sterck L, Billiau K, Van Parys T, Van de Peer Y, Vandepoele K: **PLAZA: a comparative genomics resource to study gene and genome evolution in plants.** *Plant Cell* 2009, **21:**3718-3731

Pühringer H, Moll D, Hoffmann-Sommergruber K, Watillon B, Katinger H, Laimer M: **The promoter of an apple *Ypr10* gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible.** *Plant Sci* 2000, **152(1):**35-50

Pungartnik C, da Silva AC, de Melo SA, Gramacho KP, de Mattos Cascardo JC, Brendel M, Micheli F, da Silva Gesteira A: **High-affinity copper transport and Snq2 export permease of *Saccharomyces cerevisiae* modulate cytotoxicity of PR-10 from *Theobroma cacao.*** *Mol Plant-Microbe Interact 2009, **22**:39-51*

Quiroz-Figueroa FR, Rojas-Herrera R, Galaz-Avalos RM, Loyola-Vargas VM: **Embryo production through somatic embryogenesis can be used to study cell differentiation in plants.** *Plant Cell Tiss Organ Cult* 2006, **86:**285-301

Radauer C, Lackner P, Breiteneder H: **The Bet v I fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands.** *BMC Evol Biol* 2008, **8**:286

Raemakers CJJM, Jacobsen E, Visser RGF: **Secondary somatic embryogenesis and applications in plant breeding.** *Euphytica* 1995, **81:**93-107

Raghavan V: **Role of 2,4-Dichlorophenoxyacetic acid (2,4-D) in somatic embryogenesis on cultured zygotic embryos of *Arabidopsis*: cell expansion, cell cycling, and morphogenesis during continuous exposure of embryos to 2,4-D.** *Am J Bot* 2004, **91:**1743-1756

Ramsey J, Schemske DW: **Pathways, mechanisms, and rates of polyploid formation in flowering plants.** *Annu Rev Ecol Syst* 1998, **29:**467-501

Reid KE, Olsson N, Schlosser J, Peng F, Lund ST: **An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development.** *BMC Plant Biol* 2006, **6:**27

Reguera M, Bonilla I, Bolaños L: **Boron deficiency results in induction of pathogenesis-related proteins from the PR-10 family during the legume-rhizobia interaction.** *J Plant Physiol* 2010, **167(8):**625-632

Rensing SA, Lang D,. Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, Tanahashi T,, Sakakibara K, Fujita T, Oishi K, Shin-I T, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu SH, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove C, Cuming AC, Hasebe M, Lucas S, Mishler BD,Reski R, Grigoriev IV, Quatrano RS, Boore JL: **The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants.** *Science* 2008, **319:**64–68

Repka V, Kubikova J, Fischerova I: **Immunodetection of PR1-like proteins in grapevine leaves infected with Oidium tuckeri and in elicited suspension cultures.** *Vitis* 2000, **39:**123-127.

Rizzon C, Ponger L, Gaut BS: **Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and rice.** *PLoS Comput Biol* 2006, **2(9)**:e115

Robert N, Ferran J, Breda C, Coutos-Thévenot P, Boulay M, Buffard D, Esnault R: **Molecular characterization of the incompatible interaction of *Vitis vinifera* leaves with *Pseudomonas syringae* pv. *pisi*: Expression of genes coding for stilbene synthase and class 10 PR protein.** *Eur J Plant Pathol* 2001, **107**:249-261

Robert N, Buffard D: **La Vigne et ses pathogènes : Isolement et expression de gènes codant pour différentes « Pathogenesis-related » protéines et leur utilisation pour une amélioration, par transgénèse, de ses capacités de défense.** *Thèse de doctorat No : 01 PA06 6216,* 2001

Robert N, Roche N, Lebeau Y, Breda C, Boulay M, Esnault R, Buffard D: **Expression of grapevine chitinase genes in berries and leaves infected by fungal or bacterial pathogens.** *Plant Sci* 2002, **162:**389-400

Rojas-Herrera R, Loyola-Vargas VM: **Induction of a class III chitinase in foliar explants of Coffea arabica L. during somatic embryogenesis and wounding.** *Plant Sci* 2002, **163:**705-711

Romero-Puertas MC, McCarthy I, Gomez M, Sandalio LM, Corpas FJ, Del Rio LA, Palma JM: **Reactive oxygen species-mediated enzymatic systems involved in the oxidative action of 2,4-dichlorophenoxyacetic acid.** *Plant Cell Environ* 2004, **27:**1135-1148

Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE: **Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes.** *EMBO* 1996, **15(20):**5690-5700

Samson M, Brunaud V, Duchêne S, De Oliveira Y, Caboche M, Lecharny A, Aubourg S: **FLAGdb^{++} : a database for the functional analysis of the *Arabidopsis* genome.** *Nucl Ac Res* 2004, **32(1):**D347-D350

Santos MdO, Romano E, Yotoko KSC, Tinoco MLP, Dias BBA, Aragão FJL: **Characterisation of the cacao *somatic embryogenesis receptor-like kinase* (*SERK*) gene expressed during somatic embryogenesis.** *Plan Sci* 2005, **168:**723-729

Saraste M, Sibbald PR, Wittinghofer A: **The P-loop—a common motif in ATP- and GTP- binding proteins.** *Trends Biochem Sci* 1990, **15(11):**430-434

Sato F, Kitajima S, Koyama T, Yamada Y: **Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC cis-sequence.** *Plant Cell Physiol* 1996, **37:**249-255

Schellenbaum P, Jacques A, Maillot P, Bertsch C, Mazet F, Farine S, Walter B: **Characterization of *VvSERK1, VvSERK2, VvSERK3* and *VvL1L* genes and their expression during somatic embryogenesis of grapevine (*Vitis vinifera* L.).** *Plant Cell Rep* 2008, **27:**1799-1809

Schenk MF, Gilissen LJWJ, Esselink GD, Smulders MJM: **Seven different genes encode a diverse mixture of isoforms of Bet v I, the major birch pollen allergen.** *BMC Genom* 2006, **7**:168

Schenk MF, Cordewener JHG, America AHP, van't Westende WPC, Smulders MJM, Gilissen LJWJ: **Characterization of PR-10 genes from eight Betula species and detection of Bet v I isoforms in birch pollen.** *BMC Plant Biol* 2009, **9:**24

Schirmer T, Hoffmann-Sommergruber K, Susani M, Breiteneder M, Marković-Housley Z: **Crystal structure of the major celery allergen Api g 1: Molecular analysis of cross-reactivity.** *J Mol Biol* 2005 **351**:1101-1109

Schmidt EDL, Guzzo F, Toonen MAJ, de Vries SC: **A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos.** *Development* 1997, **124:**2049-2062

Schnoes A, Brown SD, Dodevski I, Babbitt PC: **Annotation error in public databases: misannotation of molecular function in enzyme superfamilies.** *PloS Comput Biol* 2009, **5(12):**e1000605

Schuermann D, Molinier J, Fritsch O, Hohn B: **The dual nature of homologous recombination in plants.** *Trends Genet* 2005, **21(3)**:172-181

Sharma SK, Millam S, Hein I, Bryan GJ: **Cloning and molecular characterisation of a potato *SERK* gene transcriptionally induced during initiation of somatic embryogenesis.** *Planta* 2008, **228(2):**319-330

Shimada T, Hirabayashi T, Endo T, Fujii H, Kita M, Omura M: **Isolation and characterization of the somatic embryogenesis receptor-like kinase gene homologue (CitSERK1) from *Citrus unshiu* Marc.** *Scientia Hort* 2004, **103(2):**233-238

Sikorski MM, Biesiadka J, Kasperska AE, Kopcińska J, Łotocka B, Golinowski W, Legocki AB: **Expression of genes encoding PR10 class pathogenesis-related proteins is inhibited in yellow lupine root nodules.** *Plant Sci* 1999, **149:**125-137

Singla B, Tyagi AK, Khurana JP, Khurana P: **Analysis of expression profile of selected genes expressed during auxin-induced somatic embryogenesis in leaf base system of wheat (*Triticum aestivum*) and their possible interactions.** *Plant Mol Biol* 2007, **65:**677-692

Singla B, Khurana JP, Khurana P: **Characterization of three somatic embryogenesis receptor kinase genes from wheat, *Triticum aestivum*.** *Plant Cell Rep* 2008, **27:**833-843

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS: **Polyploidy and angiosperm diversification.** *Am J Bot* 2009, **96:**336-348

Somleva MN, Schmidt EDL, de Vries SC: **Embryogenic cells in *Dactylis glomerata* L. (Poaceae) explants identified by cell tracking and by SERK expression.** *Plant Cell Rep* 2000, **19:**718-726

Somssich IE, Schmelzer E, Kawalleck P, Hahlbrock K: **Gene structure and *in situ* transcript localization of pathogenesis-related protein 1 in parsley.** *Mol Gen Genet* 1988, **213(1)**:93-98

Somssich IE, Bollmann J, Hahlbrock K, Kombrink E, Schulz W: **Differential early activation of defense-related genes in elicitor-treated parsley cells.** *Plant Mol Biol* 1989, **12(2):**227-234

Song D, Li G, Song F, Zheng Z: **Molecular characterization and expression analysis of OsBISERK1, a gene encoding a leucine-rich repeat receptor-like kinase, during resistance response in rice.** *Mol Biol Rep* 2008, **35:**275-283

Srivastava S, Emery RJN, Kurepin LV, Reid DM, Fristensky B, Kav NNV: **Pea PR 10.1 is a ribonuclease and its transgenic expression elevates cytokinin levels.** *Plant Growth Regul* 2006, **49**:17-25

Tan BC, Joseph LM, Deng WT, Liu L, Li QB, Cline K, McCarty DR: **Molecular characterization of the *Arabidopsis* 9-*cis* epoxycarotenoid dioxygenase gene family.** *Plant J* 2003, **35:**44-56

Tang H, Wang X, Bowers JE, Ming R, Alam M, Paterson AH: **Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps.** *Genome Res* 2008, **18:**1944-1954.

Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M: **A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines.** *Methods* 2010, **50:**s1-s5

Tewari S, Brown SM, Kenyon P, Balcerzak M, Fristensky B: **Plant defense multigene families: II evolution of coding sequence and differential expression of PR10 genes in *Pisum*.** http://arXiv.org/q-bio.PE/0310038

Thomas C, Meyer D, Himber C, Steinmetz A: **Spatial expression of a sunflower *SERK* gene during induction of somatic embryogenesis and shoot organogenesis.** *Plant Physiol. Biochem* 2004, **42:**35-42.

Thomas BC, Pedersen B, Freeling M: **Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes.** *Genome Res* 2006, **16(7)**:934-946

Tse C, Capeau J: **Quantification des acides nucléiques par PCR quantitative en temps réel.** *Ann Biol Clin* 2003, **61:**279-293

Van Hengel AJ, Tadesse Z, Immerzeel P, Schols H, van Kammen A, de Vries SC: **N-Acetylglucosamine and glucosamine-containing arabinogalactan proteins control somatic embryogenesis.** *Plant Physiol* 2001, **125:**1880-1890

Van Loon LC, Van Strien, EA: **The family of pathogenesis-related proteins.** *Physiol Mol Plant Pathol* 1999, **55**:85-97

Van Loon LC, Rep M, Pieterse CM: **Significance of inducible defense-related proteins in infected plants.** *Annu Rev Phytopathol* 2006, **44**:135-162

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, DemattÃ L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando SM, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R: **A high quality draft consensus sequence of the genome of a heterozygous grapevine variety.** *PLoS ONE* 2, e1326

Walsh JB: **Persistence of tandem arrays: implications for satellite and simple-sequence DNAs.** *Genetics* 1987, **115:**553-567

Walter MH, Liu JW, Wünn J, Hess D: **Bean ribonuclease-like pathogenesis-related protein genes (*Yprl0*) display complex patterns of developmental, dark-induced and exogenous-stimulus- dependent expression.** *Eur J Biochem* 1996, **239:**281-293

Wei F, Wing RA, Wise RP: **Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley.** *Plant Cell* 2002, **14:**1903-1917

Wen J, Vanek-Krebitz M, Hoffmann-Sommergruber K, Scheiner O, Breiteneder H: **The potential of *Betv1* homologues, a nuclear multigene family as phylogenetic markers in flowering plants.** *Mol Phylogenet Evol* 1997, **8(3)**:317-333

Wielgoss A, Kortekamp A: **Comparison of PR1 expression in grapevinecultures after inoculation with a host- and a non-host pathogen.** *Vitis* 2006, **45:**9-13

Wiweger M, Farbos I, Ingouff M, Lagercrantz M, von Arnold S: **Expression of Chia4-Pa chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*): similarities and differencies between gymnosperm and angiosperm class IV chitinases.** *J Exp Bot* 2003, **54:**2691-2699

Wu F, Yan M, Li Y, Chang S, Song X, Zhou Z, Gong W: **cDNA cloning, expression, and mutagenesis of a PR-10 protein SPE-16 from the seeds of *Pachyrrhizus erosus.*** *Biochem Biophys Res Comm* 2003, **312:**761-766

Xie YR, Chen ZY, Brown RL, Bhatnagar D: **Expression and functional characterization of two pathogenesis-related protein 10 genes from *Zea mays.*** *J Plant Physiol* 2010, **167(2):**121-130

Yan Q, Qi X, Jiang Z, Yang S, Han L: **Characterization of a pathogenesis-related class 10 protein (PR-10) from *Astragalus mongholicus* with ribonuclease activity.** *Plant Physiol Biochem* 2008, **46**:93-99

Yanagisawa S: **The Dof family of plant transcription factors.** *Trends Plant Sci* 2002, **7(12):**555-560

Yokoyama R, Nishitani K: A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory region involved in cell-wall construction in specific organs of Arabidopsis. Plant Cell Physiol 2001, **42(10):**1025-1033

Zhang L, Gaut BS: **Does recombination shape the distribution and evolution of tandemly arryed genes (TAGs) in the *Arabidopsis thaliana* genome?** *Genome Res* 2003, **13:**2533-2540

Zhang J, Ma H, Chen S, Ji M, Perl A, Kovacs L, Chen S: **Stress response proteins' differential expression in embryogenic and non-embryogenic callus of *Vitis vinifera* L. cv. Cabernet Sauvignon–A proteomic approach.** *Plant Sci* 2009, **177**:103-113

ZhouXJ, Lu S, Xu YH, Wang JW, Chen XY: **A cotton cDNA (*GaPR-10*) encoding a pathogenesis-related 10 protein with in vitro ribonuclease activity.** *Plant Sci* 2002, **162:**629-636

Zimmerman JL: **Somatic embryogenesis: a model for early development in higher plants.** *Plant Cell* 1993, **5:**1411-1423