N° d'ordre :





Thèse de Doctorat

Université des Sciences et Technologies de Lille I

pour l'obtention du grade de

DOCTEUR EN SCIENCES

Stratégies d'Exploitation des Fonctions Biologiques

^{par} Matthieu DUBAN

Clonage et Caractérisation de deux gènes codant des récepteurs transmembranaires à activité kinase chez *Cichorium intybus* L. Expression au cours des phases précoces de l'embryogenèse somatique et du développement de la graine.

Soutenue le 17 décembre 2004, devant la commission:

M. Jean-Pierre JACQUOTM. Peter ROGOWSKYM. Théo HENDRIKSM. Jean-Louis HILBERTMme Marie-Christine QUILLET

Professeur, Université de Nancy I Directeur de Recherche, ENS de Lyon Professeur, USTL Professeur, USTL Maître de conférences, USTL

Rapporteur Rapporteur Examinateur Examinateur Examinateur 50376 2004 217

à Sophie, à Grégoire et aux autres à venir...

,

Tout d'abord, je voudrais exprimer ma reconnaissance envers toutes les personnes que j'ai rencontré au cours de cette thèse et adresser mes plus profonds remerciements à:

Monsieur le professeur Jacques Vasseur qui a accepté en septembre 1999 que je réalise mon DEA au sein de son laboratoire et de poursuivre cette étude au cours de la thèse. J'associe également à ces remerciements Monsieur le Professeur Jean-Louis Hilbert qui lui a succéder à la tête du laboratoire de Physiologie de la Différenciation Végétale et qui a accepté d'être mon co-directeur de thèse. Je te remercie pour tes conseils, notamment au sujet de l'approche immunologique et la confiance que tu as su me témoigner lors de ces années.

Monsieur Peter Rogowsky pour avoir accepter de juger cette thèse en tant que rapporteur en y consacrant de son temps.

Monsieur Jean-Pierre Jacquot, pour avoir accepté de juger cette thèse et d'avoir suivi mon évolution au cours de ce travail. Merci pour l'aide précieux lors de la purification des anticorps... et les sorties en VTT dans la forêt de Haye.

Marie-Christine Quillet, ma co-directrice de thèse, pour l'encadrement au quotidien, les remarques, les conseils, les discussions scientifiques… et de tout ordre. Merci pour le soutien moral apporté lorsque le doute s'installait.

Anne-Sophie Blervacq, pour sa patience et son dévouement pour m'expliquer la petite histoire de la cytologie de l'embryogenèse. Je tiens à te remercier pour les photos d'exceptionnelle qualité que tu as mis à ma disposition et pour ta disponibilité.

Théo Hendricks, professeur au laboratoire, "découvreur" de l'hémoglobine non symbiotique de chicorée qui fut notamment le sujet de discussions enrichissantes. Merci aussi pour ton aide dans l'écriture en anglais.

J'adresse mes remerciements à l'ensemble des maîtres de conférences du laboratoire (qui sont trop nombreux pour les citer individuellement) pour la considération qu'ils ont su me témoigner en m'intégrant à l'équipe de recherche. Je remercie particulièrement Bruno Delbreil pour les discussions enrichissantes que l'on a eues.

Je voudrais saluer les étudiants que j'ai pu côtoyer au cours de cette thèse, à commencer par ceux qui sont partis lorsque je suis arrivé, Oliv', Jeff et Eric qui m'ont fait passé une année dont je me souviendrais. Mes contemporains, du bureau 305, Jérôme, Franck, David, Sylvain, Clara, Sophie A sans oublier Ben' outre atlantique et particulièrement Sophie Da Silva ma voisine de bureau et de paillasse qui me devance de quelques jours pour la soutenance de sa thèse.

Une pensée toute particulière va à Christelle Lepers-Blassiau qui est arrivée au labo la même année que moi et avec qui ça a toujours été un plaisir de travailler. Merci pour avoir passé mes gènes sur tes descendances et pour l'initiation au Southern blot en sonde froide.

Enfin je voudrais terminer ces remerciements par ceux sans qui tout cela n'aurait jamais été possible: tout d'abord, mes parents dont j'ai conscience des soucis que je leur ai causés et des sacrifices qu'ils ont consentis pour moi, et enfin ma douce Sophie et mon petit bout Grégoire que j'ai un peu délaissés ces temps-ci mais qui ont su m'apporter du réconfort au quotidien. Je crains qu'aucun mot ne puisse exprimer ma gratitude.

INTRODUCTION		
ETUDE BIBLIOGRAPHIC	QUE	4 A 42
I L'EMBRYOGENES	E CHEZ LES VEGETAUX	4
I-A. L'EMBRYOGENESE Z	YGOTIQUE	4
I-B. L'EMBRYOGENESE S I-B.1. Caractéristiques	OMATIQUE générales	5 5
I-B.2. Origine des embi	yons somatiques	6
I-B.3. Les différentes pl	antes modèles de l'embryogenèse somatique	7
I-B.4. L'induction de l'er	mbryogenèse somatique	7
I-B.4.1. Les stress I-B.4.2. Les régulate I-B.5. Les cellules emb	urs de croissance ryogènes	7 9 11
I-B.5.1. La morpholo embryogène I-B.5.2. L'établissem I-B.6. Les gènes et les	gie cellulaire : un marqueur précoce de la compétence ent de la polarité protéines associés à l'embryogenèse somatique	11 12 13
I-B.6.1. Les gènes d I-B.6.2. Les gènes d I-B.6.3. Les gènes à I-B.6.4. L'expression I-B.6.5. Les voies de	e ménages e réponses aux hormones homéodomaine et chromodomaine des protéines extracellulaires signalisation associées à l'embryon	13 14 15 16 17
I-C. L'EMBRYOGENESE S I-C.1. Le genre <i>Cichori</i>	OMATIQUE DE LA CHICOREE	20 20
I-C.2. L'embryogenèse	somatique de la chicorée	20
II LES RECEPTEURS LES PLANTES	S TRANSMEMBRANAIRES A ACTIVITE KINASE (CHEZ 21
II-A. INTRODUCTION		21
II-B.1. CLASSIFICATION DES II-B.1. Classification éve	S RECEPTEURS DE TYPE KINASE CHEZ LES PLANTES Dutive	22 22
II-B.2. Classification fon	ctionnelle	24
II-C. LES RECEPTEURS DI	E TYPE KINASE IMPLIQUES DANS LE DEVELOPPEMENT DES P	LANTES 25
II-C.1. Déterminisme de	l'auto-incompatibilité chez Brassica	26
II-C.2. Développement o	les gamétophytes : exemple de PRK1	26
II-C.3. Morphogenèse e	t différenciation cellulaire	27
II-C.3.1. CRINKLY4 II-C.3.2. WAK II-C.3.3. ERECTA II-C.4. Développement d	du méristème	27 28 29 29

II-C	5. Signalisation hormonale	30
II-C	6. Sénescence	31
II-C	7. Abscission d'organe	31
II-D . II-D	VOIES DE SIGNALISATION DES RECEPTEURS DE TYPE KINASE 1. Le système CLAVATA1	32 32
 -D	D.1.1. CLV3 : ligand de CLV1 D.1.2. Les autres éléments du système CLV1 2. Le système SRK	32 33 36
 -D	D.2.1. Sp11/SCR est le ligand de SRK D.2.2. Les autres éléments du système SRK 3. Le système BRI1	36 37 39
 	D.3.1. Bri1 fixe les brassinolidesD.3.2. Les autres protéines du système BRI1	39 39
	LE RECEPTEUR A ACTIVITE KINASE LIE A L'EMBRYOG	ENESE
		41
III-A.		41
ш-д.	SERREI LEMBRIOGENESE ZIGUIQUE	42
MAT	RIEL ET METHODES	43 A 62
I	MATERIEL VEGETAL	43
I-A.	HISTORIQUE: DU GENOTYPE '474' AUX GENOTYPES K59 ET K28	43
I-B.	EMPLOI DES DIFFERENTS GENOTYPES	44
		11
11_A		
II-B. ORGA	MAINTIEN ET MULTIPLICATION PAR EMBRYOGENESE SOMATIQUE ET OGENESE IN VITRO	PAR 45
II-B	1. Maintien par embryogenèse somatique	45
II-B	2. Maintien par organogenèse	45
II-C.	EMBRYOGENESE ZYGOTIQUE ET MATURATION DES GRAINES	46
	EXTRACTION DES ACIDES NUCLEIQUES ET PROTEINES	S 46
III-A.	EXTRACTION DE L'ADN	46
III-B.	EXTRACTION DES ARN TOTAUX	46
III-C.	EXTRACTION ET DOSAGE DES PROTEINES	46
N7		47

V O	V CLONAGE ET SEQUENÇAGE DE PRODUITS DE PCR					
V-A.						
V-B.	-D. TRANSFORMATION ET SELECTION DE COLONIES BACTERIENNES RECOMBINANTES					
v-c.	SEQUENÇAGE D'ADN	49				
VI (DE DO	CLONAGE DE FRAGMENTS D'ADN CORRESPONDANT A UNE PORTIC MAINE KINASE DE RECEPTEUR DE TYPE KINASE	N 50				
VII	OBTENTION DES SEQUENCES CODANTES PLEINES LONGUEURS	50				
VII-A.	RT-PCR INVERSE	50				
VII-B.	5' RACE PCR	52				
VIII	ANALYSE DES SEQUENCES	54				
IX I	HYBRIDATION SOUTHERN	54				
IX-A.	DIGESTION ENZYMATIQUE DES ADN	54				
IX-B.	TRANSFERT SUR MEMBRANE	54				
IX-C.	IX-C. HYBRIDATION DES MEMBRANES					
IX-D.	REVELATION DES MEMBRANES	55				
хг	MESURE DES NIVEAUX DE TRANSCRITS PAR PCR EN TEMPS REEL	55				
X-A.	DEFINITION DES AMORCES OLIGONUCLEOTIDIQUES	55				
Х-В.	TRANSCRIPTION INVERSE	56				
X-C.	AMPLIFICATION PAR PCR EN TEMPS REEL	56				
XI F	PRODUCTION DE PROTEINES RECOMBINANTES DANS E. COLI	56				
XI-A.	CLONAGE DIRECTIONNEL DANS LE VECTEUR PET-16B	56				
XI-B. XI-B.	PRODUCTION ET PURIFICATION DE PROTEINES RECOMBINANTES 1. Production de protéines recombinantes en système bactérien	57				
XI-B.	 Purification de protéines recombinantes possédant un tag histidine en position 	on.				
N-ter	minale	58				
XII (OBTENTION ET PURIFICATION D'ANTICORPS POLYCLONAUX	58				
XII-A.	IMMUNISATION DE LAPINS	58				
XII-B. XII-B	PURIFICATION D'ANTICORPS POLYCLONAUX .1. Préparation d'une colonne de chromatographie d'affinité	59 59				
XII-B	.2. Purification des anticorps	59				

XIII V	ARIATION DU NIVEAU PROTEIQUE	59					
XIII-A.ELECTROPHORESE DE PROTEINES EN CONDITIONS DENATURANTES (SDS-PAGE)6XIII-A.1.Coloration des protéines au nitrate d'argent6							
XIII-A.	XIII-A.2. Coloration des protéines au bleu de coomassie 64						
XIII-B.	(III-B. TRANSFERT SUR MEMBRANE 61						
XIII-C.	HYBRIDATION AVEC LES ANTICORPS	61					
XIII-D. XIII-D.	REVELATION DES IMMUNOBLOTS .1. Révélation colorée	61 61					
XIII-D.	2. Révélation par chemiluminescence	61					
XIV R	EACTION D'AUTOPHOSPHORYLATION	62					
RESUL	TATS ET DISCUSSION 63 A	112					
PARTIE	I: CLONAGE ET CARACTERISATION DE GENES CODANT DES RECEPTEURS DE TYPE L	.RR-					
RLK APF	PARENTES A SERK CHEZ LA CHICOREE 63	a 81					
I C CHEZ L	LONAGE DE GENES CODANT DES RECEPTEURS DE TYPE <i>SERK</i> .A CHICOREE	63					
I-A. DOMAINE	MISE EN EVIDENCE DE L'EXISTENCE DE GENES CODANT UNE REGION HOMOLOGUE A U E KINASE DE TYPE SERK CHEZ LE CLONE '474' DE CHICOREE	JN 63					
I-B. I-B.1.	IDENTIFICATION DES SEQUENCES CODANTES PLEINES LONGUEURS RT-PCR inverse	66 66					
i-B.2.	Amplification des extrémités 5' des ADNc (RLM-RACE)	68					
I-B.3.	Amplification des ADNc pleines longueurs	69					
I-C. <i>CISERK</i>	I-C. CLONAGE DES SEQUENCES GENIQUES CORRESPONDANT AUX ADNC CISERK1 ET CISERK2 70						
II A DEDUIT	NALYSE DES SEQUENCES NUCLEOTIDIQUES ET PROTEIQUES TES.	71					
II-A. LES ADN	COMPARAISON DES SEQUENCES GENIQUES <i>CISERK1_{K59/K28}</i> et <i>CISERK2_{K59/K28}</i> av IC OBTENUS CHEZ '474'	′EC 71					
II-B.	CARTOGRAPHIE DES GENES CISERK1 ET CISERK2	72					
II-C.	GENES DE TYPE SERK CHEZ LA CHICOREE	72					
II-D.	DESCRIPTION DES PROTEINES CISERK	73					
II-E. II-E.1.	CLASSIFICATION DES PROTEINES DEDUITES CISERK Les protéines CiSERK appartiennent à la famille des RLK-LRRII	76 76					
II-E.2.	CiSERK1 et CiSERK2 appartiennent au groupe des protéines SERK	76					

78

PARTIE	II: EXPRESSION DES GENES CISERK AU COURS DE L'INDUCTION DE L'EMBRYC	GENESE			
SOMATIQU	SOMATIQUE ET DU DEVELOPPEMENT DE LA GRAINE 82 A 111				
I V	ARIATION DES NIVEAUX DE TRANSCRITS	82			
I-A. I-A.1.	PRINCIPE DE LA PCR EN TEMPS REEL ET METHODES D'ANALYSES Définition des amorces	82 83			
I-A.2.	Exploitation des données de PCR en temps réel	84			
I-A.2 I-A.2 I-A.2	 2.1. Méthode du ∆∆Ct 2.2. Evaluation des efficacités d'amplification des ADNc cibles et contrôles 2.3. Méthodologie appliquée 	84 s 85 86			
I-B. L'EMBRYC I-B.1.	EXPRESSION DES GENES CISERK1 ET CISERK2 AU COURS DES PHASES PRECO DGENESE SOMATIQUE Evènements cytologiques observés	DCES DE 88 88			
I-B.2.	Variations des niveaux de transcrits	89			
I-B.2 I-B.2 I-B.2	 2.1. Expression du gène de l'hémoglobine non symbiotique <i>CHI7507</i> 2.2. Expression du gène <i>CiSERK1</i> 2.3. Expression du gène <i>CiSERK2</i> 	89 91 91			
I-C. ZYGOTIQL I-C.1.	EXPRESSION DES GENES CISERK1 ET CISERK2 AU COURS DE L'EMBRYOGENE JE ET DU DEVELOPPEMENT DE LA GRAINE Caractérisation cytologique	SE 92 93			
I-C.1 I-C.1	1.1. Expression du gène <i>CiSERK1</i>1.2. Expression du gène <i>CiSERK2</i>	94 94			
I-D. I-D.1.	NIVEAUX DE TRANSCRITS DANS DIFFERENTS ORGANES Variation des niveaux de transcrits lors de la germination.	95 95			
I-D.2.	Variation des niveaux de transcrits entre différents organes	96			
11 V	ARIATION DE L'ACCUMULATION PROTEIQUE	97			
II-A. CISERK II-A.1.	OBTENTION ET PURIFICATION D'ANTICORPS DIRIGES CONTRE LE DOMAINE KINAS Production et purification de protéines recombinantes	E DE 97 97			
-A. ⁻ -A. ⁻ -A.2.	 1.1. Clonage des séquences codantes dans le vecteur de production pET 1.2. Production et purification des protéines recombinantes 1.3. Fonctionnalité du domaine kinase Obtention et purification d'anticorps dirigés contre le domaine kinase de C 	-16b 97 98 99 SISERK 99			

II-B. ACCUMULATION PROTEIQUE AU COURS DES PHASES PRECOCES DE L'EMBRYOGENESE SOMATIQUE 100

II-C.ACCUMULATION PROTEIQUE AU COURS DE L'EMBRYOGENESE ZYGOTIQUE ET DUDEVELOPPEMENT DE LA GRAINEII-C.1.Variation du profil protéique au cours de l'embryogenèse zygotique et du	101
développement de la graine	101
II-C.2. Immunodétection des protéines de type SERK au cours de l'embryogenèse	
zygotique et du développement de la graine	102
III DISCUSSION	102
CONCLUSION ET PERSPECTIVES	113
REFERENCES BIBLIOGRAPHIQUES	116
ANNEXES	137

Introduction

INTRODUCTION GENERALE

Le travail présenté s'inscrit dans le cadre d'un programme d'étude des mécanismes de changement d'identité cellulaire qui interviennent pour l'obtention d'embryons somatiques à partir de tissus de chicorée cultivée *in vitro*. Une caractéristique majeure des cellules végétales réside dans leur totipotence; toute cellule végétale différenciée peut recouvrer un état de dédifférenciation, aussi qualifié de réactivation cellulaire, qui peut lui permettre de se redifférencier dans un autre type cellulaire. L'embryogenèse somatique est le processus par lequel des cellules somatiques différenciées peuvent, après une phase de dédifférenciation, se développer en embryon, de structure similaire à des embryons zygotiques (Zimmerman, 1993).

Certains génotypes de chicorée présentent, dans certaines conditions, la capacité à produire de telles structures de façon rapide et abondante.

L'étude physiologique et moléculaire d'un processus lié au développement peut être appréhendé selon deux approches:

• une approche qui se veut exhaustive, telle que la protéomique ou les techniques moléculaires de "Différential Display RT-PCR", de cDNA-AFLP ou de puces à ADN, permettant d'aborder simultanément l'ensemble des éléments impliqués dans le processus,

 une approche ciblée, reposant dans un premier temps sur l'étude de l'expression de gènes candidats impliqués dans des processus analogues chez d'autres espèces, avec pour objectif de mettre en évidence ou non les mêmes événements dans l'espèce étudiée.

Chez la carotte, l'acquisition de la capacité d'une cellule dédifférenciée à entrer dans une voie morphogène conduisant à l'obtention d'embryons somatiques a été corrélée à l'expression d'un gène codant un récepteur transmembranaire de type kinase appelé *SERK* (*Somatic Embryogenesis Receptor-like Kinase*) (Schmidt *et al.*, 1997). Chez cette plante ainsi que chez *Arabidopsis*, les gènes *SERK* (respectivement *DcSERK* et *AtSERK1*) (Hecht *et al.*, 2001) sont exprimés de manière transitoire à la fois au cours des phases précoces de l'embryogenèse somatique, dans les cellules qui se développeront en embryons, et au cours de l'embryogenèse zygotique. *SERK* appartient à une superfamille de gènes, les récepteurs de type kinase, qui sont impliqués dans de nombreux processus du développement des plantes notamment l'identité et la différenciation cellulaire.

Ces constats et la connaissance des événements cytologiques qui interviennent au cours de l'embryogenèse somatique de la chicorée ont orienté notre travail vers l'approche gène candidat afin de déterminer si des gènes de type *SERK* étaient présents chez la

chicorée et impliqués dans l'embryogenèse somatique plus particulièrement au cours des phases précoces lorsque l'identité cellulaire est réorientée vers la voie embryogène, ainsi que lors de l'embryogenèse zygotique.

Aussi, afin de déterminer s'il existe un gène fonctionnel équivalent à *DcSERK* chez la chicorée, c'est à dire au moins un gène qui possède à la fois des caractéristiques structurales et fonctionnelles identiques, nous avons entrepris de mettre en évidence la présence de ce type de gènes chez la chicorée puis de déterminer si leur expression est induite au cours des phases précoces de l'embryogenèse somatique, ainsi que lors de l'embryogenèse zygotique et le développement de la graine.

Dans un premier temps, l'amplification au moyen d'amorces dégénérées d'une région correspondant à un domaine conservé a été effectuée afin de mettre en évidence l'existence de gène de type *SERK* chez la chicorée. Ceci a ensuite été confirmé par l'obtention de séquences codantes pleines longueurs.

Le second objectif de ce travail a été d'étudier la variation de l'expression des produits de ces candidats au cours des phases précoces de l'embryogenèse somatique, ainsi que lors de l'embryogenèse zygotique et du développement de la graine. Cet objectif a compris deux volets : dans un premier temps, la variation du niveau de transcrits des candidats lors des différents processus a été réalisée par RT-PCR en temps réel et conjointement, un anticorps dirigé contre ces protéines a été obtenu afin de déterminer les profils d'accumulation des produits de ces gènes par western blot.

L'étude de l'expression génique a été réalisée sur un nouveau matériel sélectionné au laboratoire et parallèlement caractérisé. En effet, dans le but d'appréhender les mécanismes moléculaires impliqués dans le déterminisme de l'embryogenèse somatique, il était important de disposer de génotypes embryogènes et non embryogènes les plus isogéniques possibles de manière à identifier avec une meilleure spécificité les gènes dont l'expression est associée à la réponse embryogène. Par ailleurs dans l'objectif de déterminer l'expression des produits des gènes candidats au cours de l'embryogenèse zygotique, une cinétique de développement de la graine a été réalisée et caractérisée en parallèle.

Le premier chapitre de ce manuscrit est consacré à une présentation bibliographique de l'embryogenèse somatique et à l'implication des récepteurs de type kinase dans le développement des plantes.

Un deuxième chapitre détaillera le matériel végétal ainsi que les méthodes analytiques à la base de ce travail.

Un troisième chapitre fera état des résultats obtenus. Il est articulé en deux parties. La première partie détaillera les différentes étapes de clonage et de caractérisation des gènes de type *SERK* chez la chicorée. La seconde partie s'attachera à la détermination de l'expression des produits de ces gènes au cours des phases précoces de l'embryogenèse somatique et de l'embryogenèse zygotique. Cette partie renferme deux volets : un premier volet sur la détermination de la variation du niveau des transcrits et un second volet consacré à la détermination de l'accumulation des protéines. Les résultats seront discutés à la fin de chaque partie.

Une conclusion générale et des perspectives à ce travail viendront clôturer ce manuscrit.

Etude Bibliographique



Figure 1: Comparaison des stades de l'embryogenèse somatique et zygotique (D'après Zimermann, 1993)

I L'EMBRYOGENESE CHEZ LES VEGETAUX

I-A. L'EMBRYOGENESE ZYGOTIQUE

1

Chez les angiospermes, l'embryon zygotique est le produit d'une double fécondation. Le grain de pollen contient deux noyaux haploïdes : un noyau végétatif et un noyau reproducteur. Au cours de la fécondation, le noyau végétatif contrôle le développement du tube pollinique jusqu'à son arrivée à l'ovule. En parallèle, le noyau reproducteur subit une mitose et génère ainsi deux noyaux fils. Un de ces noyaux reproducteurs féconde l'oosphère contenue dans le sac embryonnaire de l'ovule et donne un zygote diploïde; le second fusionne avec deux noyaux du sac embryonnaire, il en résultera un albumen triploïde et parfois polyploïde, selon les espèces. Ces deux éléments (albumen et embryon) constituent la graine et se développent simultanément. Le zygote issu de la fécondation conduit à l'embryon et l'albumen apporte les nutriments nécessaires aux premiers stades de développement de l'embryon. L'albumen persiste dans la graine mature comme tissu de réserve (graines albuminées) ou est résorbé dans les phases plus tardives du développement (graines exalbuminées) (Lopes et Larkins, 1993; Vijayaraghvan et Prabhakar K., 1984).

La polarité de l'embryon est déterminée dès la formation du zygote selon un axe apico-basal (Howell, 1998). Dans la plupart des cas, la première division du zygote est asymétrique et génère une grande cellule basale à l'origine de la formation du suspenseur, et une petite cellule apicale qui donnera l'embryon (Mansfield et Briarty L.G., 1990) (Figure 1).

Chez la dicotylédone *Arabidopsis*¹, le schéma spécifique de développement de l'embryon prend place au stade globulaire. Il est alors composé de huit cellules, et chacune subit une division péricline qui génère le protoderme, isolant l'embryon de toutes les structures périphériques. Les cellules du protoderme se divisent de manière anticlinale pour former le futur épiderme de l'embryon. Quant aux cellules internes, elles continuent de se diviser dans des directions longitudinales et transversales, générant avec le protoderme, les trois éléments radiaux de l'embryon : pro-cambium, parenchyme et épiderme. La structure radiale de l'embryon est établie lorsqu'il atteint un stade globulaire précoce (64 cellules).

Au cours de la transition du stade globulaire vers le stade cordiforme, l'embryon passe d'une symétrie radiale à une symétrie bilatérale (radiale et longitudinale). Une

¹ Il est à noter que le terme *Arabidospsis* est ici employé pour signifier *Arabidopsis thaliana*. L'utilisation d'autres espèces (ex *Arabidopsis lyrata*) est précisée dans le texte.

élongation est alors constatée et les primordium de cotylédons apparaissent du côté apical de l'embryon. Au stade cordiforme, la région de l'hypocotyle est visible et le méristème racinaire est défini. Ceci marque l'achèvement de ce schéma de symétrie longitudinale défini par l'axe apico-basal de l'embryon qui est formé des cotylédons, de l'hypocotyle et de la radicule. A ce stade, l'embryon est achevé. Il subit une maturation et une dessiccation au sein de la graine pour permettre le plus souvent son entrée dans un état de dormance, et ainsi attendre les conditions idéales à sa germination (Mordhorst *et al.*, 1997).

L'embryon ne définit pas l'architecture future de la plante. En effet contrairement aux animaux, les plantes doivent développer des organes tout au long de leur vie. De plus, l'embryon mature n'est pas une structure autonome: son développement ultérieur en plantule nécessite la présence des tissus nourriciers (albumen ou cotylédons) jusqu'à ce que la jeune plantule soit capable de prélever l'eau et les éléments minéraux du sol qui lui permettront de réaliser la photosynthèse et de devenir autotrophe.

I-B. L'EMBRYOGENESE SOMATIQUE

I-B.1. Caractéristiques générales

Les végétaux ont la capacité de se reproduire par d'autres voies que la voie sexuée. Ces voies de multiplication sont dites végétatives et peuvent intervenir *in vivo* (marcottage, stolons, apomixie), mais aussi *in vitro* par différents processus tels que l'embryogenèse haploïde (androgenèse, gynogenèse) ou l'embryogenèse somatique.

Par exemple, l'androgenèse est le processus qui provoque, dans certaines conditions expérimentales, l'induction du développement du gamète mâle haploïde en sporophyte : plutôt que d'évoluer vers une maturation du pollen, le développement des microspores unicellulaires peut être orienté vers la formation d'embryons. Ces embryons haploïdes ont d'abord été obtenus chez le datura (Guha et Maheshwari, 1964). L'androgenèse a ensuite été induite chez différentes espèces comme le colza (Lichter, 1982) ou le tabac (Heberle-Bors, 1989).

Les embryons peuvent aussi se développer à partir de cellules somatiques de plantes. Ces embryons somatiques peuvent être obtenus de manière spontanée chez certaines espèces comme *Bryophyllum* (Yarbrough, 1932) ou les fougères du genre *Malaxis* (Taylor, 1967). Ce phénomène a été observé pour la première fois *in vitro* dans une suspension cellulaire de carotte (Reinert, 1958). L'embryogenèse somatique a pu, ensuite, être induite chez plusieurs espèces telles que le dactyle (Hanning et Conger, 1982), la chicorée (Guedira *et al.*, 1989), le chêne (El-Maâtaoui *et al.*, 1990) ou encore l'hévéa (Michaux-Ferrière et Schwendiman J., 1992).

L'embryon somatique présente une structure comparable à celle de l'embryon zygotique à partir du stade globulaire (Zimmerman, 1993) (Figure 1). Alors que le zygote est intrinsèquement embryogène, les cellules somatiques ne le sont pas spontanément. Par conséquent, il est nécessaire de permettre le passage de ces cellules de l'état différencié à l'état indifférencié, puis le passage de cet état dédifférencié vers une nouvelle identité cellulaire (*i.e.* cellules embryogènes). Au cours de divisions successives, une cellule ou un groupe de cellules va évoluer en embryons. Deux étapes sont indispensables:

• une étape de dédifférenciation² ou de réactivation cellulaire (phase d'induction) conduisant à l'obtention de cellules compétentes³ pour l'embryogenèse somatique, dont certaines vont évoluer en cellules embryogènes⁴; cette étape n'a pas d'équivalence au cours de l'embryogenèse zygotique;

• une phase d'expression au cours de laquelle certaines cellules embryogènes se développent en embryons somatiques par des divisions successives.

I-B.2. Origine des embryons somatiques

L'obtention d'embryons somatiques peut se faire selon deux modes, le mode direct ou le mode indirect (Figure 1).

• Lorsque l'embryogenèse est de type direct, la cellule somatique subit une dédifférenciation, puis une redifférenciation au sein même du tissu dont elle est issue. C'est le cas de la chicorée où les embryons somatiques sont obtenus à partir de différents types d'explants (Dubois *et al.*, 1990).

• L'embryogenèse somatique de type indirect passe par un état de cal formant une masse de cellules à caractères indifférenciés dont certaines vont donner des embryons. C'est le cas de la carotte (de Vries *et al.*, 1988)

Par ailleurs, les embryons somatiques peuvent avoir deux origines différentes. Ils sont soit issus d'une seule cellule, c'est le cas de la carotte et de la chicorée; l'embryogenèse est dite d'origine unicellulaire (Dubois *et al.*, 1990; Toonen *et al.*, 1994). Soit les embryons sont

² Dédifférenciation: processus provoqué par la modification de l'équilibre physiologique des tissus qui peut être réalisée en plaçant les tissus et/ou les organes dans des conditions inhabituelles. Ceci déclenche dans ces cellules une modification plus ou moins importante de leur détermination. Ce processus peut permettre à la cellule de réacquérir une compétence (Dupire, 1999).

³ Cellules compétentes: cellules en cours de transition entre l'état somatique et embryogène (Toonen *et al.*, 1994) qui ont la capacité de percevoir un signal spécifique, interne ou externe, nécessaire à leur évolution en cellules embryogènes.

⁴ Cellules embryogènes: cellules capables d'évoluer en embryons somatiques, mais qui ne nécessitent plus l'application de stimuli supplémentaires (de Jong *et al.*, 1993).

	Nature des explants somatiques	Conditions d'induction		Conditions d'expression	Références
Carotte	Suspension cellulaire d'hypocotyle	Milieu liquide contenant 2µM de 2,4-D	Amas pro-embryogènes, 5 types morphologiques de cellules isolées	Milieu liquide sans 2,4-D	de Vries <i>et al</i> ., 1988
e	Protoplastes de cellules foliaires	Milieu contenant 10 µM 2,4-D	Groupes de cellules, Amas pro embryogènes	Milieu sans 2,4-D	Pasternak <i>et al.</i> , 2002
Luzern	Explants foliaires	milieu solide contenant 10µM ANA, 4 µM BAP + obscurité	Cals	milieu solide contenant 10 μM ANA, 4μM BAP, 1 μM ABA + obscurité	Nolan <i>et al.,</i> 2003
Maïs	Embryons zygotiques immatures	Milieu liquide supplémenté en 2,4-D, 26°C, obscurité	. Cals	Milieu liquide sans hormone	Baudino <i>et al.</i> , 2001
Tournesol	Embryons zygotiques immatures	Blessure + Milieu solide comprenant 3% saccharose, 6.6 µM BAP, 50 µM ABA + obscurité Blessure + Milieu solide comprenant 12% saccharose, 6.6 µM BAP + obscurité			Thomas <i>et al.</i> , 2004
Chicorée	Explants foliaires Explants racinaires	Blessure + Milieu liquide con	nprenant 2% saccharose, 0. obscurité	1 μM ANA, 2.5 μM 2iP+ 35°C + μ	Dubois <i>et al.</i> , 1991

Figure 2: Conditions de culture permettant l'obtention d'embryons somatiques chez différentes plantes modèles. On distingue deux grands types d'embryogenèses somatiques. Dans l'embryogenèse somatique indirecte, une phase d'induction aboutit à la formation de structures cellulaires indifférenciées compétentes à devenir embryogènes (cals) qui nécessitent une seconde phase permettant leur différenciation en embryons somatiques. C'est le cas de la carotte, de la luzerne et du maïs. Dans l'embryogenèse somatique directe, les cellules somatiques se dédifférencient et se redifférencient en embryons somatiques au sein du tissu en une seule phase de culture. C'est le cas du tournesol et de la chicorée. ANA: Acide Naphtalène Acétique (auxine), 2,4-D: acide 2,4-dichlorophénol acétique (auxine), ABA: acide abscissique, BAP: 6-benzylaminopurine (cytokinine), 2iP: 6-dimethylallylaminopurine (cytokinine).

issus d'un ensemble de cellules formant un complexe pro-embryogène, comme dans le cas du trèfle (Maheswaran et Williams, 1985); l'embryogenèse est alors d'origine pluricellulaire.

I-B.3. Les différentes plantes modèles de l'embryogenèse somatique

L'embryogenèse somatique peut être induite chez de nombreuses plantes. Cependant, même si le processus d'induction est décrit chez de nombreuses espèces, son mécanisme est souvent peu connu. Il existe des espèces pour lesquelles les mécanismes d'induction et de développement de l'embryon somatique sont bien décrits. Les conditions d'induction de l'embryogenèse somatique des plantes pour lesquelles ce processus fait l'objet d'études plus approfondies sont illustrées dans la figure 2.

L'embryogenèse peut être induite à partir d'une grande variété de tissus, allant d'explants foliaires ou racinaires (chez la chicorée ou la luzerne), aux embryons zygotiques immatures (chez le tournesol ou le maïs), en passant par des cultures cellulaires issues de différents organes (chez la carotte).

En plus de se distinguer par la variété d'explants somatiques initiaux, les différents modèles montrent aussi une large palette de conditions d'induction qui peut se réaliser de façon directe (chicorée, tournesol) ou indirecte (carotte, luzerne, maïs), en présence de quantités variables de régulateurs de croissances.

Lors de l'embryogenèse somatique, les phases précoces pendant lesquelles les cellules somatiques recouvrent leur état indifférencié pour se différencier à nouveau et suivre la voie embryogène sont déterminantes. Nous nous attacherons par la suite, à décrire les caractéristiques de ces phases précoces en développant uniquement les différents facteurs qui permettent l'induction de l'embryogenèse somatique, ainsi que les caractéristiques d'une cellule embryogène. Les phases tardives de l'embryogenèse somatique correspondant au développement de l'embryon à partir d'une structure cellulaire typique d'un embryon globulaire ne seront pas décrites, mais seulement évoquées lors de la description des gènes et protéines dont l'expression est liée à l'embryogenèse somatique dans sa globalité.

I-B.4. L'induction de l'embryogenèse somatique

I-B.4.1. Les stress

L'état de différenciation des cellules végétales peut facilement être modifié dans des conditions de culture *in vitro*. Ces conditions exposent les explants à des stress significatifs par leur isolement de leur environnement tissulaire original et leur mise en culture sur un milieu contenant des concentrations non physiologiques en éléments tels que des régulateurs de croissance, sels et composants organiques. Les facteurs comme une blessure (Grosset *et al.*, 1990), de fortes concentrations en sel, en ions, en métaux lourds ou

un stress osmotique influencent positivement l'induction de l'embryogenèse somatique chez diverses plantes (Dudits *et al.*, 1995; Pasternak *et al.*, 2002). Les réponses à ces conditions de stress dépendent de deux paramètres principaux : le niveau d'intensité du stress et l'état physiologique des cellules. Un stress supérieur à la tolérance des cellules entraîne leur mort. Par contraste, un niveau de stress plus faible accroît le métabolisme et induit des mécanismes d'adaptation (Lichtenthaler, 1998). Ces adaptations incluent la reprogrammation de l'expression des gènes, de même que des modifications dans la physiologie et le métabolisme cellulaire. Cet état cellulaire de transition, induit par des conditions de stress, peut être caractérisé par une large réorganisation cellulaire et permet un basculement de développement si des signaux appropriés sont perçus (Feher *et al.*, 2003).

Les stress mécaniques et physico-chimiques

Dans les protoplasmes de feuilles de luzerne, des composants induisant un stress oxydatif augmentent le niveau endogène d'auxine dans les cellules et favorise la dédifférenciation. Les stress ne favorisent pas seulement la dédifférenciation, mais peuvent aussi être utilisés pour induire la formation d'embryons somatiques. Par exemple, des cellules embryogènes peuvent être formées à partir de protoplastes de feuilles de luzerne en réponse à différents composants induisant un stress oxydatif en présence d'auxine et de cytokinines (Pasternak *et al.*, 2002).

Les stress hormonaux

L'influence de l'application d'auxines exogènes sur l'induction de l'embryogenèse somatique est bien documentée (Dudits *et al.*, 1991). Parmi les différents analogues d'auxines utilisés pour induire l'embryogenèse somatique, l'acide 2,4-dichlorophénol acétique (2,4-D) est de loin le plus efficace et de ce fait est utilisé dans plusieurs systèmes de culture de tissus et de cellules embryogènes. Le 2,4-D, au-delà d'une certaine concentration a un double effet sur ces cultures: il agit comme une auxine (directement ou au travers du métabolisme auxinique) et comme un agent de stress (Feher *et al.*, 2001). Le 2,4-D est un herbicide qui possède divers effets associés à une activité phytotoxique, ne pouvant être décrits comme une simple surdose auxinique (Fargasova, 1994). Il a été montré que les herbicides auxiniques interagissent avec les synthèses d'acide abscissique et d'éthylène par augmentation du niveau cellulaire de ces hormones de "stress" (Grossmann, 2000).

L'embryogenèse somatique peut être induite par l'application d'autres hormones exogènes telles que les cytokinines (Sagare *et al.*, 2000) ou l'acide abscissique (Nishiwaki *et al.*, 2000) qui agissent souvent en synergie avec d'autres hormones.

Cependant, le développement d'embryon dans les tissus somatiques a aussi été observé en absence ou en très faible quantité de régulateurs de croissance (Choi *et al.*, 1998),

Au regard de cet éventail d'inducteurs, l'embryogenèse somatique ne peut pas être définie comme une réponse spécifique à l'application exogène d'un ou de plusieurs régulateurs de croissance.

I-B.4.2. Les régulateurs de croissance

Les niveaux d'hormones endogènes sont considérés comme des facteurs majeurs dans la détermination de la spécificité des réponses cellulaires aux stimuli de stress. Les auxines et l'acide abscissique sont les régulateurs de croissance principaux chez les plantes, et sont impliqués dans la régulation de la division et de la différenciation cellulaire.

L'Acide indole acétique endogène

Alors que les cultures *in vitro* nécessitent la présence d'auxines exogènes pour une croissance soutenue, les cellules de plantes produisent néanmoins des auxines endogènes et en particulier l'acide indole acétique (AIA). Une concentration élevée d'AIA endogène est associée à l'augmentation de la réponse embryogène chez de nombreuses espèces (Ivanova *et al.*, 1994; Jimenez et Bangerth, 2001). Chez la carotte, l'application de 2,4-D exogène stimule l'accumulation de grande quantité d'AIA endogène (Michalczuk *et al.*, 1992). Ces auteurs émettent l'hypothèse que la compétence embryogène des cellules de carotte est finement associée à l'augmentation des niveaux endogènes d'AIA en raison de la présence de 2,4-D. Il a été suggéré que ce composant synthétique agit indirectement en désorganisant le métabolisme de l'auxine endogène.

Dans les protoplastes de feuilles de luzerne cultivés en présence de 2,4-D, les niveaux endogènes d'AIA augmentent considérablement pendant les 2 à 3 premiers jours de culture. Cette augmentation est transitoire et comparable dans des conditions embryogènes et non embryogènes. Cependant, le pic d'accumulation d'AIA montre approximativement 1 jour de retard dans les conditions non embryogènes, par rapport aux conditions embryogènes. (Pasternak *et al.*, 2002).

Un pic similaire d'AIA endogène a été observé dans les embryons immatures de tournesol cultivés *in vitro* pour induire la formation d'embryons somatiques (Charriere *et al.*, 1999). Les tissus cultivés dans des conditions embryogènes montrent une augmentation de la concentration en AIA d'un facteur 4 par rapport aux tissus cultivés en conditions organogènes. Le moment d'apparition du pic d'AIA (après 24 heures de culture) est bien

corrélé avec le moment de la détermination irréversible de la réponse morphogène (Thomas, 1993).

Le transport polaire de l'auxine endogène est un facteur important dans la formation des embryons somatiques dans des explants de cotylédons de ginseng, qui ne requiert pas d'application de régulateur de croissance (Choi *et al.*, 1998).

Par ailleurs, chez la carotte, la simple application d'acide abscissique sur les plantules induit efficacement la formation d'embryons somatiques (Nishiwaki *et al.*, 2000). Cependant, cette réponse est dépendante de la présence de l'apex caulinaire, (la source principale d'auxine dans les plantules), et illustre par conséquent l'importance de la source d'auxine et de son transport.

L'ensemble de ces observations suggère que les changements spatio-temporaux des niveaux d'auxine endogènes sont des facteurs importants qui contrôlent le devenir des cellules embryogènes.

L'acide abscissique endogène

L'acide abscissique (ABA) est une hormone de croissance végétale qui régule plusieurs processus lors de l'embryogenèse et de la formation de la graine. Elle est connue pour être accumulée dans d'autres parties des plantes lors de la réponse aux stress abiotiques comme la sécheresse, le froid ou le stress salin (Tabaeizadeh, 1998). L'acide abscissique endogène montre un pic d'accumulation lors de la maturation de l'embryon zygotique et diminue dans la graine dessiquée.

La contribution directe de l'ABA endogène sur la phase d'induction de l'embryogenèse somatique a été montrée chez le tabac (Senger *et al.*, 2001). Plusieurs approches ont été utilisées pour diminuer le niveau cellulaire d'ABA dans ces plantes: des lignées transgéniques homozygotes surexprimant un anticorps anti-acide abscissique ont été produites, des plantes sauvages ont été cultivées en présence de l'inhibiteur de la synthèse d'ABA et des mutants de la synthèse d'acide abscissique (*aba1* et *aba2*) ont aussi été utilisés. Dans tous les cas, la déficience d'ABA perturbe la morphogenèse au stade préglobulaire de la formation de l'embryon. Ceci peut être rétabli par l'application d'ABA exogène.

Les niveaux d'ABA endogènes sont inversement corrélés avec la capacité embryogène de la luzerne (Ivanova *et al.*, 1994). Cette observation semble en contradiction avec le rôle positif de l'ABA exogène sur l'induction de l'embryogenèse somatique (Nishiwaki *et al.*, 2000), mais il peut être expliqué par la différence des effets d'applications exogènes et des hormones endogènes : un niveau d'ABA endogène plus élevé a sans doute pour conséquence une diminution de la réponse aux stress ou aux signaux de l'ABA exogène.

10

Les résultats de différentes expériences décrites précédemment concernant l'implication d'un stress et des régulateurs de croissance soulignent l'importance des interactions entre l'auxine, la signalisation du stress et de l'ABA. L'induction parallèle de ces voies peut mener à des modifications de la morphologie et du développement, observées lors de la transition des cellules somatiques vers leur état embryogène par une rapide dédifférenciation suivie d'une redifférenciation. L'activation simultanée des réponses aux hormones et aux stress semble être un événement clé de l'adaptation cellulaire, impliquant une reprogrammation génétique, des mécanismes métaboliques et physiologiques, qui aboutissent à la compétence embryogène des cellules somatiques végétales.

I-B.5. Les cellules embryogènes

Dans l'optique d'étudier les phases les plus précoces de la transition d'un état de cellules somatiques à celui de cellules embryogènes, il est nécessaire de disposer de marqueurs qui indiquent l'acquisition de la compétence embryogène des cellules induites. Un marqueur doit pouvoir être utilisé pour révéler les processus cellulaires spécifiques liés à la transition étudiée et permettre la description et la détermination d'un état cellulaire.

I-B.5.1. La morphologie cellulaire : un marqueur précoce de la compétence embryogène

Chez la carotte, la formation de cellules embryogènes peut être corrélée avec des changements caractéristiques de la morphologie cellulaire. Des cellules isolées fractionnées par centrifugation montrent que la fraction contenant des petites cellules isodiamétriques et présentant un cytoplasme dense, pouvait initier la formation d'embryons somatiques (Nomura et Komamine, 1985). Le développement de techniques telles que le "video cell tracking" a permis de suivre l'évolution des cellules embryogènes de carotte (Toonen *et al.*, 1994). La fraction de cellules isolées au sein d'une culture cellulaire embryogène peut être classée selon 5 groupes morphologiques. Bien que chacun de ces types cellulaires soit capable de se développer en embryons somatiques avec des efficacités variables, la fréquence la plus élevée a été observée dans le cas de petites cellules sphériques montrant un cytoplasme dense.

Chez la luzerne, les cellules dérivées des protoplastes de feuilles cultivées en présence de différentes concentrations de 2,4-D peuvent se développer en type cellulaire soit embryogène soit non embryogène. Des caractéristiques morphologiques similaires peuvent être observées entre des protoplastes de génotype embryogène et non embryogène

cultivés à la même concentration de 2,4-D. Dans les deux cas, les cellules embryogènes sont petites, sphériques avec un cytoplasme dense alors que les cellules non embryogènes sont allongées et avec une large vacuole (Bogre *et al.*, 1990; Dudits *et al.*, 1991; Pasternak *et al.*, 2002).

Dans les explants foliaires de dactyle, des petites cellules isodiamétriques avec un cytoplasme riche sont capables de se développer en embryons somatiques. (Somleva *et al.*, 2000). De même, les cellules compétentes à l'embryogenèse formées sur les explants de chicorée sont aussi caractérisées par un cytoplasme dense (Blervacq *et al.*, 1995).

Par contraste, dans les cultures de cellules initiées à partir de pétiole de luzerne, les cellules isolées sont incapables de se développer en embryons somatiques, à la différence des amas de cellules présents eux aussi dans la suspension. Ces amas cellulaires embryogènes sont constitués de petites cellules se divisant rapidement (Xu et Bewley, 1992).

Cet ensemble d'observations et d'expérimentations suggère que, dans la plupart des systèmes embryogènes, incluant les gymnospermes, la morphologie des cellules ayant développées les caractéristiques de petites cellules à cytoplasme dense soit associée à la compétence embryogène (Egertsdotter et von Arnold, 1998; Yeung, 1995).

I-B.5.2. L'établissement de la polarité

La division zygotique initiale chez les plantes supérieures est asymétrique (Scheres et Benfey, 1999). Cet événement établit la base de la polarité de la plante qui est déterminante pour le modèle de développement ultérieur (Jurgens *et al.*, 1997).

Lors de l'embryogenèse somatique, l'asymétrie n'est pas toujours évidente. Par exemple, à partir de protoplastes de feuilles de luzerne, la formation de cellules se divisant de manière asymétrique est dépendante de la concentration en 2,4-D (Pasternak *et al.*, 2002), du génotype de la plante (Bogre *et al.*, 1990) et détermine son développement futur. Chez l'hybride interspécifique '474' de chicorée la division asymétrique des cellules embryogènes n'a pas été observée (Blervacq *et al.*, 1995).

Cependant, la distribution inégale de déterminants cellulaires peut être décisive dans l'identification des voies de développement cellulaire après la division. Ainsi, chez la carotte, les embryons somatiques sont induits à partir d'une population hétérogène de cellules. Toutefois, les petites cellules isolées à cytoplasme dense ont la capacité de se développer en embryons après une première division inégale et une synthèse polarisée de macromolécules (Nomura et Komamine, 1985). Sur une sous-population de cellules de carotte caractérisées par leur marquage grâce à un anticorps monoclonal (Pennell *et al.*, 1992), les cellules évoluent vers une division qui résulte dans la formation de cellules filles



<u>Figure 3</u>: Représentation schématique de différents gènes identifiés influençant l'embryogenèse somatique chez les plantes supérieures.

possédant des caractéristiques de marquage à l'anticorps différentes et dépendant de leur développement à suivre. Les cellules filles marquées avec l'anticorps meurent tandis que les cellules dépourvues de cet épitope de surface sont embryogènes et forment des embryons somatiques (McCabe *et al.*, 1997).

Il est vraisemblable que l'établissement de la polarité est un événement important, non seulement pour le zygote, mais aussi pour l'embryon somatique, même si dans ce cas l'asymétrie de la division cellulaire n'est pas toujours une caractéristique morphologique.

I-B.6. Les gènes et les protéines associés à l'embryogenèse somatique

Un processus de développement tel que l'embryogenèse somatique implique non seulement le remaniement de la morphologie mais aussi celui de l'identité cellulaire ce qui suppose une modification importante de l'expression des gènes et de leurs produits (Figure 3). Il est de plus en plus évident qu'un panel important de famille de gènes jouant des rôles à différents stades de développement est impliqué dans l'embryogenèse somatique chez plusieurs espèces (Tableau 1).

I-B.6.1. Les gènes de ménages

De nombreux gènes du fonctionnement général des cellules (les gènes de ménages) voient leur expression modulée.

Plusieurs gènes jouant un rôle significatif lors de la division cellulaire et de la formation de la paroi à différents stades de la différenciation de l'embryon ont été identifiés (Hirt *et al.*, 1991; Sato *et al.*, 1995).

Des gènes impliqués dans la machinerie de l'expression génique, incluant des gènes codant des histones (Kapros *et al.*, 1992), un facteur d'élongation, (Kawahara *et al.*, 1992) ou une topoisomérase (Balestrazzi *et al.*, 1996) ont été identifiés et montrent une induction de leur expression à différents stades de l'embryogenèse somatique.

Des gènes impliqués dans la régulation du cycle cellulaire peuvent aussi avoir un rôle clé dans l'embryogenèse somatique. Ainsi, chez la carotte, un ADNc de cycline a été cloné et est exprimé dans les cultures cellulaires lors de l'embryogenèse somatique (Hata *et al.*, 1991). Chez la luzerne, un ADNc codant une protéine kinase dépendante des cyclines cdc2 a été isolé et montre un niveau de transcrit élevé dans les cultures de suspensions cellulaires embryogènes induites par l'auxine (Hirt *et al.*, 1991). Les mêmes observations ont été faites chez l'épicéa au cours de l'embryogenèse somatique (Footitt *et al.*, 2003).

L'implication d'un système de régulation du métabolisme de l'azote commun à l'embryogenèse somatique et zygotique a été montrée notamment grâce à l'expression de

13

Gène / Protéine	Espèce	Nature / Fonction	Stade d'expression	Référence
Gènes de ménage				· · · · · ·
CEM1	Carotte	Facteur d'élongation 1- α	Stade globulaire	Kawahara <i>et al.</i> , 1992
CEM6	Carotte	Protéine de la paroi	Stade pré-globulaire et globulaire	Sato <i>et al.</i> , 1995
H1, H3-11	Luzerne	Histones		Kapros <i>et al.</i> , 1992
topi	Carotte	Topoisomérase I	Phase de prolifération cellulaire + pic au stade torpille	Balestrazzi <i>et al.</i> , 1996; Balestrazzi <i>et al</i> ., 2001
CGS102, CGS103, CGS201	Carotte	Glutamine synthase	Stades précoces	Higashi <i>et al.</i> , 1998
Cyc A	Carotte	Cycline		Hata <i>et al.</i> , 1991
Cdc2	Luzerne Epicea	Protéine kinase dépendante des cyclines	Stades précoces	Hirt <i>et al.</i> , 1991 Footitt <i>et</i> <i>al.</i> , 2003
CHI3206 <u>Réponse aux hormones</u>	Chicorée	Hémoglobine non symbiotique	Stades précoces	Hendriks <i>et al.</i> , 1998
Dcarg-1, Dchsp-1	Carotte	Liés au stress inductible par l'auxine	Stades précoces	Kitamiya <i>et al.</i> , 2000
cDNA clones 43, 93 DcECP31, DcECP40, DcECP63 C-ABI3	Carotte Carotte Carotte	Inductible par l'auxine Inductible par l'ABA Facteur de transcription	Cellules embryogènes Stade torpille –	Yasuda <i>et al.</i> , 2001 Kiyosue <i>et al.</i> , 1993 Shiota <i>et al</i> ., 1998
Pavp1	Epicea	Facteur de transcription inductible par ABA	Suspension embryogène et pic au stade coltylédonnaire précoce	Footitt <i>et al.</i> , 2003
PgEMB12-15	Epicea	Inductible par l'ABA	Phases précoces	Dong et Dunstan, 1997a
CHI-GST1	Chicorée	Glutathione S transférase inductible par l'auxine	Phases précoces	Galland et al., 2001
Transduction du signal				
SERK	Carotte	Récepteur kinase de l'embryogenèse somatique Protéine kinase dépendante du	Cellules cométentes -> stade globulaire	Schmidt <i>et al.</i> , 1997
swCDPK	Bois de Santal	calcium et indépendante des Calmodulines	Stade globulaire	Anil <i>et al.</i> , 2000
CRK	Carotte	Protéine kinase apparentéè à une CDPK	_	Lindzen et Choi, 1995
MsCa1	Luzerne	Protéine calcium dépendante	_	Dudits D. <i>et al.</i> , 1991
MsCPK3	Luzerne	Protéine kinase similaires aux calmodulines	Stades précoces	Davletova <i>et al.</i> , 2001
<u>Gènes à homéodomaines</u>				
CHB1-CHB6	Carotte	Gènes à homéodomaines HD-Zip	Stade globulaire -> Torpille	Kawahara <i>et al.</i> , 1995
sbh1	Soja	Facteur de transcription	stade cœur -> torpille	Ma <i>et al.</i> , 1994
DcDB1	Carotte	Proteine avec domaine d'interaction avec la chromatine	Stades précoces	Kiyosue <i>et al.</i> , 1998
C-LEC1 WUSCHEL	Carotte Arabidopsis	Facteur de transcription Facteur de transcription	-	Yazawa <i>et al.</i> , 2004 Zuo <i>et al.</i> , 2002

CUS1	Concombre	Facteur de transcription		Filipecki <i>et al.</i> , 1997
AGL15	Luzerne	Facteur de transcription	_	Perry et al., 1999
ZmMADS1, ZmMADS3	Maïs	Facteurs de transcription	Phases précoces	Heuer <i>et al.</i> , 2001
Protéines extracellulaires				
EP3-1, EP3-2	Carotte	Chitinases	_	van Engelen <i>et al</i> ., 1995
DcAGP1	Carotte	Arabinogalactanes	_	Baldwin <i>et al.</i> , 2001
AGP	Chicorée	Arabinogalactanes	_	Chapman et al., 2000
EP2	Carotte	Protéine de transfort de linides	_	Sterk <i>et al.</i> , 1991
AtLTP1	Arabidopsis	Floteine de transient de lipides		Toonen <i>et al.</i> , 1997
PgChi-1	Epicéa	Endochitinase	_	Dong et Dunstan, 1997b
Chia4-Pa	Epicéa	Chitinase	_	Wiweger et al., 2003
PgGlu-1	Epicéa	β1-3 glucanase	_	Dong et Dunstan, 1997b
CG1,CG2, CG3	Chicorée	β1-3 glucanase		Helleboid <i>et al.</i> , 2000
Maturation des embryons		-		
Dc2.15	Carotte	Gène de maturation		Holk <i>et al.</i> , 1996
			lle	Franz <i>et al.</i> , 1989 Dong et
Dc3, Dc8, DcECP31, DcECP40,		Protéines chandentes de	d	Dunstan, 1997a; Goupil <i>et</i>
DcEMB1, MsLec1, MsLEC2, PgEMB12-	Carotte, Luzerne, Epicea	l'ombryogenèse tardiye (LEA)	e to	al., 1992; Kiyosue et al.,
14-15		Tempryogenese laruwe (LLA)	ade	1992; Kiyosue <i>et al.</i> , 1993;
			St	Zimmerman, 1993
EMB-1	Carotte	Gène homologue à Em		Ulrich <i>et al.</i> , 1990

Tableau 1: Gènes relatés dans la littérature et exprimés au cours de l'embryogenèse somatique de différentes espèces.

deux gènes codant des isoformes de la glutamine synthase. En effet, chez la carotte, le niveau de transcrits correspondant à deux gènes (*CGS102* et *CGS20*) augmente au cours des stades précoces de l'embryogenèse somatique et dans les graines en développement (Higashi *et al.*, 1998).

Du fait de la complexité de ce processus, il n'est pas surprenant que des gènes essentiels associés à des activités cellulaires diverses et majeures telles que la régulation du cycle cellulaire et du métabolisme général soient exprimés au cours de l'embryogenèse somatique. Des études du transcriptome lors de l'embryogenèse somatique du soja semblent d'ailleurs indiquer que l'expression de nombreux gènes liés au métabolisme général est modifiée (Thibaud-Nissen *et al.*, 2003).

I-B.6.2. Les gènes de réponses aux hormones

Les gènes inductibles par les auxines

Les auxines sont considérées comme l'un des initiateurs potentiels de l'embryogenèse somatique (Komamine *et al.*, 1990). L'application d'auxine sur des organes excisés, des cultures cellulaires ou des plantes entières, peut mener à une accumulation rapide de nombreux ARNm. Par exemple, chez la carotte, un gène codant une protéine de chocs thermiques (Dchsp-1) est exprimé lors du développement de l'embryon somatique en réponse à des hormones comme le 2,4-D (Kitamiya *et al.*, 2000).

Deux clones partiels d'ADNc (43, 93) ont été isolés de groupes de cellules aux stades les plus précoces de l'embryogenèse somatique de la carotte par la technique de "differential display RT-PCR" (DD-RT-PCR). Les transcrits correspondant s'accumulent préférentiellement dans les groupes de cellules embryogènes formés après le traitement au 2,4-D (Yasuda *et al.*, 2001). Les séquences d'acides aminés déduites des clones d'ADNc 43 et 93 montrent une forte homologie respectivement avec une protéine proche de la thaumatine et un précurseur de la protéine de la paroi Dc2.15.

Les gènes inductibles par l'acide abscissique

Des ADNc codant des protéines spécifiques de l'embryon ou des cellules embryogènes ont été isolés et caractérisés chez la carotte et chez *Arabidopsis*. L'expression des gènes codant ces protéines abondantes notamment au niveau des phases tardives de l'embryogenèse est inductible par l'ABA (Kiyosue *et al.*, 1993).

Le gène *C-ABI3*, homologue au gène *ABI3* d'*Arabidopsis*, isolé et caractérisé à partir d'une banque d'ADNc d'embryon somatique de carotte (Shiota *et al.*, 1998), code un facteur

de transcription lié à la signalisation impliquant l'acide abscissique. Il est détecté dans les cellules embryogènes, les embryons somatiques et la graine en développement.

I-B.6.3. Les gènes à homéodomaine et chromodomaine

Les gènes à homéodomaine ont été identifiés pour la première fois chez la drosophile, et jouent un rôle de régulateurs clés en contrôlant le modèle de formation et la différenciation morphologique des organismes multicellulaires (McGinnis *et al.*, 1984). Ces gènes homéotiques renferment une séquence nucléotidique conservée et caractéristique des facteurs de transcription appelée l'homéodomaine (homeobox). Cette région permet la fixation de ces protéines à l'ADN. Ces gènes sont maintenant bien caractérisés chez les plantes comme le maïs, *Arabidopsis* et le riz (Chan *et al.*, 1998).

Plusieurs gènes homéotiques ont été identifiés et caractérisés à partir de banque d'ADNc d'embryons somatiques. Par exemple, chez le soja, *Sbh1*, montre une augmentation du niveau de ses transcrits lors de la transition entre les stades cœur et torpille, lorsque la différenciation des cotylédons et des tissus provasculaires est déterminée. De plus, les ARNm ne sont pas détectables dans les tissus non embryogènes, ni dans les autres tissus à l'exception de la tige et l'hypocotyle qui présentent néanmoins un faible niveau d'expression. (Ma *et al.*, 1994).

Six gènes à homéodomaine (*CHB1-CHB6*) ont été isolés à partir de banque d'ADNc d'embryons somatiques et d'hypocotyles de carotte (Kawahara *et al.*, 1995). L'expression spatio-temporelle des différents transcrits varie lors de l'embryogenèse somatique, mais aussi dans la plantule. *CHB1* est exprimé de façon constante dans les amas de cellules indifférenciées; l'expression de *CHB2* est augmentée après le stade globulaire et passe par un maximum au stade cœur et au début du stade torpille de la différenciation en embryon somatique. L'expression de *CHB1* et *CHB2* est faible dans les cultures cellulaires non différenciées. Les transcrits de *CHB3*, *CHB4* et *CHB5* sont accumulés préférentiellement dans les cellules corticales les plus internes de l'axe embryonnaire au stade torpille. L'expression de *CHB6* est restreinte aux cellules procambiales aux stades cœur et torpille. Dans les cotylédons embryonnaires et l'hypocotyle des plantules, les ARNm de ces 4 derniers gènes sont associés aux tissus vasculaires. Ces observations indiquent que la régulation de la différenciation des tissus vasculaires est impliquée lors du développement embryonnaire.

DcCB1 a été isolé et caractérisé chez la carotte et code une protéine contenant un chromodomaine, ce motif semble intervenir au niveau de l'interaction protéines-chromatine. L'expression de *DcCB1* est augmentée lors des stades précoces de l'embryon somatique

jusqu'au stade cordiforme alors qu'elle diminue au stade torpille de l'embryon somatique (Kiyosue *et al.*, 1998).

L'expression des gènes à homéodomaine et à chromodomaine à différents stades de l'embryogenèse somatique suggère qu'un mécanisme de régulation de l'expression des gènes liés au développement est impliqué dans l'embryogenèse somatique et le développement de l'embryon.

I-B.6.4. L'expression des protéines extracellulaires

Les protéines extracellulaires jouent un rôle important dans le développement embryogène des Angiospermes (van Engelen et de Vries, 1992). Ces protéines et la variation de leur profil d'expression ont été associées à l'induction de l'embryogenèse somatique chez la chicorée (Boyer *et al.*, 1993; Hilbert *et al.*, 1992).

Les enzymes secrétées

Les cultures cellulaires de carotte secrètent un large spectre de protéines dans le milieu, et ce processus contribue au conditionnement de celui-ci. L'une de ces protéines sécrétées est une chitinase capable de restaurer le développement des embryons somatiques d'une lignée cellulaire mutante (de Jong *et al.*, 1992; Kragh *et al.*, 1996; Van Hengel *et al.*, 1998). De la même façon, une peroxydase lève l'inhibition de l'embryogenèse somatique provoquée par la tunicamycine (Cordewener *et al.*, 1991).

Chez la chicorée, trois ADNc: *CG1*, *CG2* et *CG3*, codant des protéines extracellulaires présentant des homologies avec des β 1,3-glucanases, ont été isolés à partir de fragments de feuilles cultivées en conditions d'induction de l'embryogenèse somatique. *CG1* est exprimé de façon préférentielle dans les tissus embryogènes, tandis que les transcrits *CG2* et *CG3* sont détectés à la fois dans les tissus embryogènes et non embryogènes (Helleboid *et al.*, 1998; Helleboid *et al.*, 2000).

Les Arabinogalactanes

Les protéoglycanes arabinogalactanes (AGP, Arabinogalactan Protein), composés d'environ 10% de protéines et à plus de 90 % de sucres sont caractéristiques des plantes supérieures. Ces protéines, riches en hydroxyproline, alanine, glycine et sérine, sont communément trouvées au niveau de la membrane cellulaire, la matrice extracellulaire et la paroi. Bien que leurs fonctions biologiques restent hypothétiques, certains rôles ont été proposés; par exemple les AGP peuvent être impliquées dans la prolifération (Nothnagel, 1997), et l'expansion cellulaire (Willats et Knox, 1996), mais également dans la régulation du développement de l'embryon somatique (Egertsdotter et Arnold, 1995; Knox, 1995). Les

AGP joueraient un rôle dans les interactions moléculaires et la signalisation cellulaire de surface. Leur importance lors de l'embryogenèse somatique a été montrée par l'addition dans le milieu de culture de réactif de Yariv (β D glycosyl [β GlcY]) qui interagit avec les AGP. Ce composé altère la réponse morphogène (formation de racine au lieu d'embryon chez la carotte) (Thompson et Knox, 1998) ou bloque la formation de l'embryon chez la chicorée (Chapman *et al.*, 2000).

Les protéines de transfert des lipides

Les protéines de transfert de lipides (LTP, Lipid Transfer Protein) sont impliquées dans le transport des phospholipides, et en particulier celui des monomères de cutines. Les gènes codant des LTP sont exprimés au niveau des premiers tissus qui se différencient dans l'embryon somatique de carotte comme par exemple le protoderme. Le premier ADNc codant une LTP a été cloné à partir de culture embryogène de carotte : chez cette espèce, le gène *EP2* est exprimé dans les cultures cellulaires embryogènes, ainsi que dans les embryons somatiques, et son niveau d'expression est directement corrélé au potentiel embryogène de la culture cellulaire. La protéine est présente dans la paroi et le milieu extracellulaire (Sterk *et al.*, 1991). Les protéines LTP, sécrétées dans le milieu de culture embryogène, jouent différents rôles. Ces protéines extracellulaires limitent l'expansion cellulaire et renforcent l'adhésion entre cellules filles issues d'une cellule embryogène et permettent la formation d'une structure cellulaire embryonnaire.

I-B.6.5. Les voies de signalisation associées à l'embryon

L'ensemble des éléments que nous venons d'évoquer est impliqué dans l'embryogenèse somatique et intervient soit en tant que signaux déclencheurs des processus de stress physiques ou hormonaux nécessaires à l'induction, soit comme éléments effecteurs du développement des cellules embryogènes ou au cours des phases plus tardives de l'embryogenèse somatique. Le déroulement de ces différentes étapes implique des mécanismes de transduction des signaux afin d'activer les effecteurs responsables des modifications successives d'identités cellulaires.

Le calcium comme second messager

Le calcium est un second messager jouant un rôle clé dans différents processus cellulaires et physiologiques chez les plantes supérieures (Harper, 2001).

Le rôle du calcium est bien connu dans le cas de l'embryogenèse somatique de la carotte où il est essentiel lors de la transition des cellules indifférenciées vers les embryons somatiques à la concentration de 200 μ M (Overvoorde et Grimes, 1994), au-delà, le calcium n'a pas d'effet sur la viabilité ni sur le potentiel embryogène des cultures. A une concentration plus faible ou par chélation du calcium résiduel par l'EGTA, l'embryogenèse somatique est inhibée. De plus, certains inhibiteurs de canaux calciques, comme le vérapamil ou la néfidipine, ont une influence négative sur la capacité embryogène. Ces observations suggèrent des rôles importants du calcium exogène et du maintient du gradient de calcium dans le développement de l'embryon. Les analyses du calcium associé aux membranes et de la distribution du calcium total, grâce aux agents fluorescents, révèlent des variations de sa distribution lors de l'embryogenèse, sans altération de la concentration en calcium associé aux membranes.

Enfin, pendant l'embryogenèse somatique du bois de santal les auteurs ont montré une augmentation d'un facteur 4 de la concentration calcique lors de la différenciation des masses pro-embryogènes en embryons somatiques. Des agents chélateurs bloquent la formation des embryons somatiques alors que les cellules continuent à proliférer, prouvant ainsi que les voies de signalisation font appel au calcium en tant que second messager (Anil et Rao, 2000).

Présence de protéines d'interaction avec le calcium au cours des phases précoces de l'embryogenèse somatique

En parallèle aux mécanismes de transduction du signal à travers la membrane, des protéines perçoivent l'augmentation du calcium cytosolique. En effet, ces protéines fixent le calcium et deviennent actives par un changement de conformation. Ceci leur confère la capacité d'interagir avec d'autres protéines de régulation (Poovaiah et Reddy, 1993).

La calmoduline (CaM) est un exemple de protéine qui possède ce mécanisme d'action chez les végétaux. Les protéines CaM sont codées par une famille multigénique chez les plantes supérieures (Ling *et al.*, 1991).

Des complexes actifs calcium/calmoduline ont été détectés dans les régions méristématiques des embryons aux stades cœur et torpille chez la carotte. L'expression des gènes de calmoduline augmente lors de l'induction de l'embryogenèse, puis demeure à un niveau constant suggérant ainsi le rôle de régulation des calmodulines dans les régions embryonnaires qui présentent des divisions cellulaires rapides (Overvoorde et Grimes, 1994).

Chez la luzerne, une augmentation du niveau de transcrit d'un gène codant une protéine de fixation du calcium (*MsCa1*) après un traitement au 2,4-D et son accumulation

préférentielle aux stades globulaires précoces, confirment le rôle des protéines liées au calcium dans l'induction de l'embryogenèse somatique (Dudits *et al.*, 1995)

Les protéines kinases

Les protéines kinases nécessitent une activation par phosphorylation et sont impliquées dans la régulation d'éléments de la transduction du signal.

Une famille de protéines kinases dépendantes du calcium, mais indépendantes de la calmoduline (CDPKs) a été caractérisée chez le soja (Harmon *et al.*, 1986). Contrairement aux kinases liées à la calmoduline, les CDPKs sont activées par la fixation directe du calcium. Ces kinases disposent d'un motif de fixation du calcium au niveau de leur région C-terminale analogue à celui des calmodulines, qui permet l'activation du domaine kinase après la modification de leur conformation (Harper *et al.*, 1991).

Deux protéines CDPKs de 55 et 60 kDa (swCDPK) ont été isolées dans les extraits protéiques solubles de cultures embryogènes de bois de santal et leur expression préférentielle au cours de l'embryogenèse somatique et zygotique démontre leur implication dans les processus de différenciation et de développement. En effet, lorsque les cellules embryogènes sont cultivées en présence d'agents chélateurs du calcium, les swCDPK ne sont plus détectées reliant donc leur présence à un rôle dans l'embryogenèse somatique par l'intermédiaire des signaux calciques (Anil et Rao, 2000).

De plus, un ADNc codant une protéine kinase nommée CRK (CDPK related kinase) et apparentée à la famille des CDPK a été cloné chez la carotte. Le niveau d'expression des ARNm de ce gène *CRK* est plus important dans les embryons somatiques que dans les tissus différenciés. La protéine qui semble être phosphorylée par une kinase dépendante des cyclines, joue ainsi un rôle central dans la régulation du cycle cellulaire (Lindzen et Choi, 1995).

L'ADNc codant MsCPK3 (Calmoduline-like Protein Kinase), une protéine kinase similaire à la calmoduline a été isolée à partir d'une culture cellulaire de luzerne. L'expression du gène correspondant augmente lors de la phase précoce de l'embryogenèse somatique. La phosphorylation de MsCPK3 *in vitro* est activée par le calcium et inhibée par un antagoniste de la calmoduline, ce qui confirme son implication dans la voie de signalisation calcique. Par ailleurs, un choc thermique ou un traitement à l'auxine induit l'expression de MsCPK, ce qui suggère sa participation dans la reprogrammation de la voie de développement induite par le stress vers l'embryogenèse somatique (Davletova *et al.*, 2001).
<u>Encadré 1</u>: Les formes cultivées de *Cichorium intybus* L. appartiennent à de nombreux cultigroupes que l'on peut classer selon leur utilisation:

- les chicorées industrielles dont les racines sont torréfiées ou conduisent notamment la production d'inuline,
- les chicorées dont les racines sont récoltées la première année puis cultivées dans des conditions de forçage; c'est la chicorée Witloof ou de Bruxelles (endive), la Trévise (rouge), la Castelfranco (panachée) ou la barbe de capucin,
- les chicorées récoltées pour leurs feuilles et consommées crues comme le pain de sucre (verte), la Chioggia (verte, rouge ou panachée), la Vérone (panachée) ou cuites comme la Catalogne.

Un récepteur transmembranaire à activité kinase

L'expression d'un gène codant un récepteur kinase a été mise en évidence au cours de l'embryogenèse somatique de la carotte. Les caractéristiques de ce gène seront exposées dans un paragraphe qui lui est consacré dans la suite de cette bibliographie.

I-C. L'EMBRYOGENESE SOMATIQUE DE LA CHICOREE

I-C.1. Le genre Cichorium

Le genre *Cichorium* appartient à la famille des Astéracées. Il regroupe une dizaine d'espèces dont les origines se situent dans le bassin méditerranéen et l'Europe continentale. Parmi celles-ci, on recense trois espèces européennes, diploïdes (2n=18) (Tutin *et al.*, 1964).

• *Cichorium spinosum L.* est une espèce non cultivée, présente en Espagne, en Italie et en Grèce. Elle est pérenne et allogame préférentielle.

 Cichorium endivia L. est présente sur le pourtour du bassin méditerranéen, est une espèce annuelle, autogame préférentielle. Ses formes cultivées constituent les cultigroupes Scaroles et Frisées.

Cichorium intybus L. est présente de l'Europe à l'Asie, elle est naturalisée dans plusieurs régions du monde. C'est une espèce pérenne, cultivée comme bisannuelle (Encadré 1), allogame préférentielle.

I-C.2. L'embryogenèse somatique de la chicorée

L'embryogenèse somatique de la chicorée a été mise en évidence pour la première fois sur des cultures de cals dérivés d'embryons matures de *Cichorium endivia L.* (Vasil *et al.*, 1964; Vasil et Hildebrandt, 1966). Dans un milieu liquide, les tissus se dissocient en une fine suspension cellulaire ou de groupes de cellules et de petites structures sphériques désignées comme embryoïdes, cependant ces structures ne se développent pas ultérieurement. Des structures embryoïdes mises en évidence à partir de culture de racines de *C. intybus L.* (Heigwehr *et al.*, 1985) ont la capacité de se développer en plantules.

Néanmoins, dans ces deux cas, l'embryogenèse somatique est indirecte, c'est à dire qu'elle fait intervenir une phase intermédiaire de cal.

Des embryons somatiques ont été obtenus à partir d'anthères et de styles prélevés sur de fleurs immatures d'un hybride de chicorée, *C. intybus* \times *C. endivia*: l'hybride '474' (Dubois *et al.*, 1988; Guedira *et al.*, 1989). D'autre part, les racines et les feuilles de plantules



<u>Figure 4</u>: Rôle d'un récepteur dans la réponse cellulaire.



<u>Figure 5</u>: Représentation schématique d'un récepteur kinase. De part et d'autre du domaine hydrophobe et membranaire se situent le domaine extracellulaire qui permet la liaison à un ligand et à l'extrémité duquel un peptide signal de localisation membranaire peut être clivé, d'autre part un domaine intracellulaire qui porte une activité kinase responsable de la première phase de la voie de signalisation cellulaire.

du génotype '474', âgées de 2 à 3 mois et régénérées à partir des premiers groupes d'embryons obtenus, constituent un matériel présentant de meilleures capacités embryogènes, notamment grâce à la possibilité de générer des embryons *in vitro* tout au long de l'année (Dubois *et al.*, 1990; Dubois *et al.*, 1991). L'embryogenèse somatique du clone '474' présente les caractéristiques d'être directe et unicellulaire (Blervacq *et al.*, 1995; Dubois *et al.*, 1991).

II LES RECEPTEURS TRANSMEMBRANAIRES A ACTIVITE KINASE CHEZ LES PLANTES

II-A. INTRODUCTION

Les cellules réagissent à leur environnement par l'intermédiaire de mécanismes de transduction de signaux externes, aboutissant notamment à la modification de l'expression des gènes (Fletcher et Meyerowitz, 2000; Hardie, 1999). Ces signaux extracellulaires sont de natures diverses et sont dans un premier temps perçus par des récepteurs.

Parmi les différents types de récepteurs mis en évidence, il existe des récepteurs transmembranaires dont le rôle est d'interagir avec des ligands hydrophiles qui n'ont pas la possibilité de passer au travers de la membrane. Ces récepteurs servent ainsi d'intermédiaire entre le milieu environnant des cellules et le milieu intracellulaire. Cette perception des signaux présents dans le milieu extérieur engendre des réactions de la part des cellules (Figure 4).

Ces réactions s'effectuent par des mécanismes de transduction dont certains font appel à des récepteurs à activité kinase. La transduction du signal est alors réalisée par l'intermédiaire d'étapes de phosphorylations et de déphosphorylations d'effecteurs intracellulaires.

Le séquençage du génome d'*Arabidopsis* a révélé plusieurs centaines de gènes codant des récepteurs de type kinase. Leur mode de fonctionnement est inconnu pour la majorité d'entre eux, cependant il apparaît que différents membres de cette famille sont impliqués dans certains aspects du développement et de la défense des plantes. Leur découverte a été significative et a permis de démontrer que de tels mécanismes de signalisation, identifiés chez les animaux, existaient également chez les végétaux (Becraft, 1998). En effet, il était admis qu'en raison de la présence de la paroi, la principale voie de communication des cellules végétales se faisait de manière directe au travers des plasmodesmes.

Ces récepteurs possèdent une activité kinase qui est modulée en réponse à un stimulus. Historiquement, les premiers récepteurs kinases identifiés étaient des protéines membranaires contenant un domaine transmembranaire. De telles protéines possèdent un domaine extracellulaire récepteur (ectodomaine) qui permet la reconnaissance et la fixation spécifique du ligand (peptides, lipides, glycannes...), et un domaine cytoplasmique à activité kinase qui active les réactions d'autophosphorylation des récepteurs eux-mêmes mais aussi de phosphorylations de molécules effectrices dont ils modulent l'activité (Figure 5).

Les génomes de plantes codent un nombre important de protéines avec des topologies prédictives de récepteurs kinase pour lesquels la fonction n'a pas été démontrée; c'est pourquoi ces protéines sont référencées comme des récepteurs de type kinase (RLK, Receptor-Like Kinase) par homologie aux récepteurs à activité kinase identifiés chez les animaux.

De façon générale, les récepteurs kinases sont présents dans la membrane en tant que monomères inactifs (Becraft, 1998). La fixation du ligand induit leur dimérisation, rapproche leurs domaines kinases intracellulaires et leur permet de se transphosphoryler et de s'activer mutuellement.

Il y a des exceptions à ce modèle général. Par exemple, certains RLK de plantes sont des multimères sous forme inactive (Giranton *et al.*, 2000; Trotochaud *et al.*, 1999) et semblent fonctionner au travers d'une activité kinase intramoléculaire (Schulze-Muth *et al.*, 1996). De plus, certains RLK disposent de domaines kinases aberrants qui semblent être inactifs (Valon *et al.*, 1993).

Ainsi, le mode d'action des RLK ne peut pas être prédit mais doit être démontré sur des bases individuelles.

II-B. CLASSIFICATION DES RECEPTEURS DE TYPE KINASE CHEZ LES PLANTES

II-B.1. Classification évolutive

Les données concernant les génomes et les transcriptomes (au travers des bases d'EST, Expressed sequence tag) montrent que les végétaux possèdent un grand nombre de gènes codant des récepteurs de type kinase. Ainsi, l'analyse des séquences génomiques a révélé qu'*Arabidopsis* contient 417 gènes codant des récepteurs de type kinase (Shiu et Bleecker, 2001b). Ceci peut être comparé aux 43 récepteurs (40 tyrosine kinase (RTK) et 3 sérine/thréonine kinase (STK)) chez *Caenorhabditis elegans* (Plowman *et al.*, 1999), 25 (20 RTK et 5 STK) annotées chez *Drosophila melanogaster* (Flybase, 1999), et 70 (58 RTK et 12 STK) chez l'Homme (Smith *et al.*, 1997).

Il existe deux classes majeures de récepteurs protéiques à activité kinase qui montrent une différence de spécificité de substrat du domaine kinase. Les récepteurs



<u>Figure 6</u>: Les récepteurs kinases humains et d'*Arabidopsis* appartiennent à des familles distinctes mais apparentées; les kinases de type Pelle sont les homologues animaux des RLK d'*Arabidopsis*. L'arbre phyllogénétique a été construit à partir des séquences protéiques des domaines kinases de récepteurs animaux (h: *Homo sapiens*, ce: *Caennorbitis elegans* dm: *Drosophila melanogaster*) et végétaux (at: *Arabidopsis thaliana*) ainsi que des protéines kinases cytoplasmiques. Des protéines kinases caractéristiques des plantes et des animaux ont été utilisées pour réaliser cette analyse. L'arbre est ancré avec APH(3')III (aminoglycosid kinase de *Staphylloccus* présentant un faible niveau d'évolution). Il indique que les protéines Raf, RSTK, RTK et RLK forment un groupe cohérent distinct des autres protéines kinases eucaryotes (région encadrée). La valeur de boostrap est indiquée aux nœuds (100 échantillonnages). Les kinases animales de type Pelle (région grisée) se situent dans la même clade que les RLK (branches vertes) (d'après Shiu et Bleecker, 2001).

tyrosines kinase (RTK) phosphorylent les résidus tyrosine, tandis que les récepteurs sérine/ thréonine kinase (STK) phosphorylent les résidus sérine et thréonine. En plus d'un nombre de gènes codant des RLK plus important chez les plantes, une distinction majeure entre les récepteurs kinases des animaux et des plantes est soulignée par la prévalence des STK chez les végétaux. En effet, chez les animaux, il existe 20 classes de RTK basées sur les séquences des ectodomaines (Plowman et al., 1999), alors que la famille du récepteur au TGFβ (Transforming Growth Factor) représente la seule classe des STK. Par contraste, tous les RLK végétaux décrits contiennent des séquences consensuelles, constituant les signatures d'une activité sérine/thréonine kinase. Cependant au moins deux apparaissent comme ayant une double spécificité de substrat de l'activité kinase; PRK1 (Pollen Receptorlike Kinase1) s'autophosphoryle sur des résidus sérine et tyrosine (Mu et al., 1994) et SERK (Somatic Embryogenesis receptor-like kinase) se phosphoryle sur des résidus sérine, thréonine et tyrosine (Shah et al., 2001c). De manière similaire la levure ne contient que des protéines kinases de séquences consensus de type STK, cependant, 30 des 119 protéines kinases testées sont capables de phosphoryler des résidus tyrosine (Zhu et al., 2000). Il n'est pas à exclure que d'autres RLK de plante puissent avoir une activité tyrosine kinase, montrant de nouveau la nécessité de caractériser l'activité biochimique de chaque RLK.

Une analyse récente des kinases du génome d'*Arabidopsis* a révélé que les 417 gènes codant des RLK ont une origine monophylétique et appartiennent à un clade qui contient les kinases cytoplasmiques de type "Pelle" (impliqué dans l'orientation dorsoventrale de l'embryon chez la drosophile) des animaux et IRAK (Interleukine-1 receptor-asociated kinase) de l'homme (Figure 6). Ce clade contient aussi un grand nombre de gènes codant des kinases cytoplasmiques végétales (environ 200 gènes). Les RTK animaux forment une lignée distincte des RLK de plantes suggérant que différentes kinases ont été recrutées de manière indépendante pour fonctionner comme récepteurs dans ces deux règnes. Les protéines groupées selon leur homologie du domaine kinase possèdent fréquemment des ectodomaines similaires, ce qui indique que la plupart des familles ont subi une expansion après un événement initial de fusion des domaines extracellulaires et kinases (Shiu et Bleecker, 2001a).

Le séquençage du génome du riz a montré qu'il existe environ deux fois plus de gènes codant des RLK appartenant au clade des RLK/Pelle que chez *Arabidopsis*. Cette différence dans le nombre de gènes qui n'est pas nécessairement observée dans d'autres familles de kinase, n'est pas une conséquence de la taille du génome du riz supérieur à celui d'*Arabidopsis*. Cette extension du nombre de membres de cette famille peut être attribuée à un mécanisme de duplication des gènes plus important chez le riz. Cette duplication affecte d'avantage les gènes impliqués dans la résistance aux pathogènes plutôt que ceux liés au

23



Figure 7: Organisation des domaines de RLK représentatifs et de leur sous-famille d'affiliation. Sur la base de la présence ou de l'absence de domaines extracellulaires, les membres de cette famille de gènes sont considérés comme des récepteurs de type kinase ou des kinases cytosoliques de type récepteur. La ligne grise figure la membrane plasmique. Les noms de locus ou les noms de gènes d'*Arabidopsis thaliana* sont donnés pour les RLK représentatifs. Les noms des domaines sont présentés selon les bases de données SMART et Pfam. Les sous-familles sont assignées sur la base de la phylogénie des domaines kinases et sont exposées selon l'organisation du domaine de la majorité de leurs membres. Les sous-familles avec plus de 30% de membres dans plus qu'une catégorie majeure de domaine extracellulaires sont signifiées par un astérisque. DUF: domaine de fonction inconnu; EGF: epidermal growth factor; PAN: domaine des protéines du plasminogène, du domaine apple de facteur XI de coagulation et de protéine de nématode; TM: Région transmembranaire; TNFR: tumor necrosis factor receptor (d'après Shiu et Bleecker, 2001).

développement et pourrait être une conséquence des pressions de sélection pour la résistance aux pathogènes qu'a pu subir cette espèce au cours de sa domestication et de sa sélection par l'homme (Shiu *et al.*, 2004).

Les banques d'EST (Expressed Sequence Tag) contiennent des séquences de gènes de RLK pour plus de vingt espèces végétales. A l'heure actuelle, il n'y aurait aucune classe de RLK identifiée chez d'autres espèces qui n'ait pas de relation proche avec des membres identifiés chez *Arabidopsis*.

II-B.2. Classification fonctionnelle

Sur la base d'une phylogénie du domaine kinase, la famille des RLK peut être subdivisée en 30 sous-familles qui représentent 21 classes structurelles d'ectodomaines différentes chez *Arabidopsis* (Shiu et Bleecker, 2001b) (Figure 7).

Le motif extracellulaire le plus commun est le domaine LRR (Leucine Rich Repeat), présent chez 216 des 417 membres des RLK d'*Arabidopsis* (Shiu et Bleecker, 2001a). Le motif LRR est commun aux signaux de transduction qui semblent être impliqués dans les interactions protéines-protéines. Les motifs LRR spécifiques des plantes sont composés de 23 à 25 acides aminés et contiennent une séquences consensus conservée: LxxLxxLxxLxxLxxLxxCxIPxx⁵. Les structures tridimensionnelles ont été définies pour quelques LRR chez les animaux et montrent que chaque unité est constituée d'un feuillet β et d'une hélice α , liés par une boucle (Kobe et Deisenhofer, 1995).

Chez *Arabidopsis*, sur la base des domaines kinases, les LRR-RLK composent un groupe de 13 sous-familles caractérisées par différents nombres et arrangements de motifs LRR, pouvant être associés en 8 classes (Shiu et Bleecker, 2001b).

La seconde classe d'ectodomaine la plus représentée contient des motifs analogues aux lectines. Cette classe inclut 42 protéines "lectin receptor kinase" (LecRK), avec un ectodomaine similaire aux lectines des légumineuses. Des substitutions d'acides aminés dans le site prédictif de fixation des monosaccharides suggèrent que les LecRK ne doivent pas fixer les sucres simples comme le font les véritables lectines de légumineuses (Barre *et al.*, 2002).

Un motif de β -lectine ou d'agglutinine est présent dans l'ectodomaine de 40 RLK contenant le domaine S riche en cystéine (Shiu et Bleecker, 2001a).

⁵ L:Leucine; N: Arginine; G: Glycine; I: Isoleucine; P:Proline; x: acide aminé non conservé

Nom du Gène	Fonction biologique	Référence(s)
Récepteur kinase apparenté à CRINCKLY4		
CRINKLY4	Différenciation cellulaire	Becraft <i>et al.</i> , 1996
Pégantaur kingga I PP		
	dse Lnn	Clark at al. 1002
CLVI	integrite du mensieme apical	
ERECTA	Initiation et elongation des organes	I orii <i>et al.</i> , 1996
HAESA	Abscission des organes floraux	Jinn <i>et al.</i> , 2000
PRK1	Développement du pollen	Mu <i>et al.</i> , 1994
EMS/EXS	Développement du pollen	Canales et al., 2002; Zhao et al., 2002
VH1	Différenciation du tissu provaculaire	Clav et Nelson, 2002
BAK1	Signalisation des brassinostéroïdes	Li <i>et al.</i> , 2002: Nam et Li, 2002
BRI1	Signalisation des brassinostéroïdes	Liet Chory, 1997
PSK1	Signalisation des phytosulfokine	Matsubavashi <i>et al.</i> , 2002
SARK	Sénescence	Hajoui <i>et al.</i> , 2000
SIRK	Sénescence	Bobatzek et Somssich 2001
SERK	Embryogenèse	Hecht et al. 2001: Schmidt et al. 1997
OLIIN	LINDIVOGENESE	Theorit et al., 2001, Ochimat et al., 1997
Récepteur kinase à domaine S		
SRK	Auto-incompatibilité chez Brassica	Giranton et al., 2000; Stein et al., 1991

 Récepteur kinase de type WAK

 WAK
 Expansion cellulaire
 Anderson et al., 2001; Lally et al., 2001

 Tableau 2:
 Les différents gènes codant des récepteurs de type kinase de fonction connue et impliqués dans le développement des plantes.

Le domaine S, similaire à la glycoprotéine du locus *S*, est situé au niveau de l'ectodomaine du récepteur kinase associé au déterminisme de l'auto-incompatibilité chez *Brassica* (Stein *et al.*, 1991).

D'autres motifs présents au niveau des domaines extracellulaires suggèrent la fixation de sucres et incluent un domaine chitinase comme dans le cas de CHRK1 (Chitinase Related receptor-like kinase) (Kim *et al.*, 2000b), un domaine thaumatine (Osmond *et al.*, 2001) ou un domaine lysine (Bateman et Bycroft, 2000). Les motifs extensines suggèrent aussi les interactions avec la paroi.

Plusieurs autres motifs extracellulaires pourraient être impliqués dans les interactions entre protéines. Les kinases associées à la paroi (WAK) contiennent toutes des répétitions de motifs similaires à ceux du facteur de croissance épithélial (EGF) en plus d'autres motifs variables similaires à des protéines de la matrice extracellulaire des métazoaires comme les collagènes, la neurexine et la tenascine (He *et al.*, 1999).

La famille de CRINKLY4 (CR4) contient des répétitions riches en cystéine présentes dans le récepteur du facteur de nécrose des tumeurs (TNF) (Becraft *et al.*, 1996). Le domaine DUF26, aussi appelé CRR (Cystein Rich Repeat), contient 4 cystéines conservées avec une signature C-x8-C-x2-C. Le motif est présent dans au moins 42 RLK d'*Arabidopsis* (Shiu et Bleecker, 2001a), et apparaît être spécifique des plantes; sa fonction est inconnue.

La suite de ce chapitre a pour objectif d'apporter une vision globale des récepteurs kinases et de leur fonctionnement dans la de transduction du signal dans les différents aspects du développement des plantes en excluant leur implication dans les relations plantes-microorganismes.

II-C.LES RECEPTEURS DE TYPE KINASE IMPLIQUES DANS LE DEVELOPPEMENT DES PLANTES

Le nombre de récepteurs de type kinase dont la fonction lors du développement des plantes est démontrée est sans cesse en augmentation (Tableau 2). C'est pourquoi nous avons traité dans ce mémoire les plus caractéristiques, impliqués dans les différentes phases du développement des plantes.

Encadré 2: les incompatibilités génétiques

Chez beaucoup de plantes qui pourraient, morphologiquement, s'autoféconder, des systèmes d'incompatibilité génétique permettent de favoriser ou même de rendre obligatoire la fécondation croisée. Ces systèmes se situent au niveau de la germination ou du développement du grain de pollen sur le stigmate. Ils font appel à des gènes d'incompatibilité (*S*) qui existent sous la forme de nombreux allèles (S1, S2, S3, ..., Sx) et peuvent être classés en deux processus : l'auto-incompatibilité gamétophytique et l'auto-incompatibilité sporophytique (d'après le site http://www.snv.jussieu.fr/bmedia/Pollinisation/).

L'auto-incompatibilité gamétophytique

plantes hétérozygotes diploïdes Les disposent pour le gène S de deux allèles. Le pollen, gamétophyte mâle haploïde, ne contient qu'un de ces allèles. Si l'allèle du pollen est le même que l'un des deux allèles exprimé par le stigmate (diploïde) de l'organe femelle, le développement du tube pollinique sera bloqué et la fécondation ne pourra avoir lieu. Dans ce cas. l'autofécondation est impossible. Le schéma ci-dessous montre un cas dans lequel un allèle diffère entre les deux plantes.



La plante A (S1/S3) produit des grains de pollen S1 ou S3. Seul le pollen S1 pourra germer puis féconder la plante B (S3/S3). En revanche, la plante B (S3/S3) ne produit que des grains depollen S3. Ils ne pourront germer sur la plante A (S1/S3) car le stigmate exprime l'allèle S3

L'auto-incompatibilité sporophytique

Les plantes hétérozygotes disposent pour le gène S de deux allèles. Le pollen, gamétophyte mâle haploïde, ne contient qu'un de ces allèles. Mais une partie des composants de sa paroi a été synthétisée par les cellules nourricières du pollen au cours de sa formation. Ces cellules étaient diploïdes et contenaient donc les deux allèles. Si l'un des deux allèles de la plante diploïde qui a généré le pollen est le même que l'un des deux allèles exprimé (diploïde). le stigmate par le développement du tube pollinique sera bloqué et la fécondation ne pourra avoir lieu. Dans ce cas. l'autofécondation est impossible. Le schéma ci-dessous montre un cas dans lequel un allèle diffère.



La plante A (S1/S3) produit des grains de pollen S1 ou S3 qui contiennent dans leur paroi des produits synthétisés par le tapis nourricier diploïde correspondant aux deux allèles S1 et S3. Aucun grain de pollen ne peut germer sur la plante B (S3/S3) car tous contiennent des produits d'origine S3. Il en est de même pour le croisement inverse.

II-C.1. Déterminisme de l'auto-incompatibilité chez Brassica

L'auto-incompatibilité est un mécanisme qui favorise les fécondations croisées et implique un rejet actif de l'auto-pollen. Ce caractère est contrôlé par l'haplotype d'un locus complexe appelé le locus *S* (Nasrallah, 1997). Le récepteur kinase du locus *S* fonctionne dans les cellules stigmatiques lors de la reconnaissance du pollen au niveau du système d'auto-incompatibilité sporophytique de *Brassica* (Goring *et al.*, 1993) (Encadré 2). Deux gènes exprimés dans le stigmate sont codés au niveau du locus *S*: *SRK* et *SLG* (Stein *et al.*, 1991). Le gène SRK code une protéine à activité sérine/thréonine kinase localisée au niveau de la membrane plasmique (Cabrillac *et al.*, 2001).

Des mutations du gène *SRK* au niveau de la région codant le domaine S sont associées à une perte de la réponse d'auto-incompatibilité (Goring *et al.*, 1993). La suppression de l'activité kinase de SRK grâce à une expérience de transformation de *Brassica napus* par une forme inactive de la protéine supprime l'auto-incompatibilité (Stahl *et al.*, 1998).

Par ailleurs, la transformation d'un haplotype *S60* de *Brassica rapa* par une construction contenant l'allèle *S28* de *SRK* confère une réponse de rejet des stigmates au pollen S28 (Takasaki *et al.*, 2000). Des résultats similaires ont été obtenus avec *Brassica napus* (Silva *et al.*, 2001).

SRK est par conséquent le déterminant femelle de l'auto-incompatibilité sporophytique chez Brassica.

L'ectodomaine de la protéine SRK est similaire à SLG, une glycoprotéine sécrétée (Stein *et al.*, 1991). Les motifs extracellulaires de SRK et la protéine SLG incluent un domaine de type agglutinine, un domaine S riche en cystéine et un domaine PAN (domaine commun au plasminogène, au domaine "Apple" du facteur XI de coagulation et à plusieurs protéines de Nématode) pour lequel des interactions protéines-protéines ou les interactions protéines-sucres ont été démontrées (Tordai *et al.*, 1999).

II-C.2. Développement des gamétophytes : exemple de PRK1

Le gène *PRK1* (Pollen Receptor-like Kinase 1) a été isolé à partir d'une banque d'ADNc de tube pollinique de pétunia. L'expression de PRK1 a été initialement détectée dans les anthères qui contiennent des microspores, et le niveau de transcrits atteint un maximum dans le pollen mature, puis persiste dans les tubes polliniques du pollen en germination (Mu *et al.*, 1994).

De plus, des auteurs ont montré que certaines plantes, transformées par une construction contenant la région codant l'ectodomaine en orientation anti-sens, sont incapables de former du pollen. En effet, les cellules mères du pollen évoluent vers une

Encadré 3: Structure de l'albumen d'un grain de maïs L'albumen représente le tissu de réserve de la graine, fournissant les éléments nécessaires à la croissance de la plantule lors de la germination. L'albumen est un tissu simple composé de 3 types majeurs de cellules: l'albumen amylacé (jaune), la couche de cellules de transfert à la base (bleu) et l'aleurone (rouge). La couche de transfert participe au prélèvement des nutriments depuis la plante mère lors du développement de la graine. L'albumen "amidonné" est le site majeur de stockage de l'amidon et de protéines. Tout en ayant des fonctions de stockage, l'aleurone est un tissu du catabolisme; à la germination, il sécrète des amylases dans l'albumen "amidonné" causant la destruction de l'amidon stocké, fournissant les sucres et l'énergie nécessaires à la croissance de plantule (d'après le site du laboratoire de P.W la Becraft http://www.public.iastate.edu/~becraft/Endosperm.htm).



méiose normale et produisent des microspores; mais le développement postérieur est anormal: la moitié des microspores (sans doute celle possédant le transgène) n'entre pas en mitose, et reste uninucléée pour finalement avorter, suggérant que la fonction de PRK1 est essentielle pour le développement postméiotique du pollen (Lee *et al.*, 1996).

Ces expériences ont montré que les plantes transgéniques présentent des anomalies de développement du gamétophyte femelle. Dans ce cas, les sacs embryonnaires des lignées transformées par la construction anti-sens évoluent habituellement vers des mitoses normales, mais la maturation du sac embryonnaire et la différenciation n'ont pas lieu. Les transcrits de PRK1 sont également détectés dans les ovules de plantes sauvages (Lee *et al.*, 1997). Ainsi PRK1 est un élément intégré à la signalisation et il semble nécessaire à la gamétogenèse.

La protéine PRK1 d'une masse de 69 kDa est constituée de 720 acides aminés. Les 328 acides aminés de l'ectodomaine contiennent 5 motifs LRR. Bien que la séquence du domaine kinase prédit une activité sérine/thréonine kinase, l'autophosphorylation du domaine cytoplasmique se fait uniquement sur des résidus sérine et tyrosine (Mu *et al.*, 1994).

II-C.3. Morphogenèse et différenciation cellulaire

II-C.3.1. CRINKLY4

CR4 a été cloné à partir d'un mutant d'insertion de maïs et code un RLK (Becraft *et al.*, 1996). Les plantes mutées pour *CR4* sont petites, avec des feuilles froissées et forment fréquemment des fusions ressemblant à des greffes. Au niveau des graines de ces mutants, les cellules périphériques de l'albumen ne se différencient pas en cellules de l'aleurone mais restent à un état de cellules de l'albumen amylacé (encadré 3). *CR4* est exprimé dans les régions en croissance de la partie aérienne, et particulièrement dans le méristème apical et les primordia latéraux (Becraft *et al.*, 2001). L'étude d'une série allélique de mutants a montré que *CR4* est actif préférentiellement dans l'épiderme, mais est requis dans divers processus du développement cellulaire dans l'ensemble de la partie aérienne. *CR4* est nécessaire dans la graine lors de la différenciation des cellules de l'aleurone, mais ses fonctions peuvent être étendues à la régulation de la prolifération cellulaire, l'identité cellulaire, la morphogenèse et la différenciation. (Becraft et Asuncion-Crabb, 2000).

La protéine CR4 correspondante est composée de 901 acides aminés et présente une activité sérine/thréonine kinase fonctionnelle dans le domaine cytoplasmique (Jin *et al.*, 2000). L'ectodomaine riche en cystéine contient un motif similaire au domaine de fixation du ligand du récepteur mammifère au facteur de nécrose de tumeur (TNF), et suggère par conséquent que le ligand du récepteur CR4 pourrait être un peptide apparenté au TNF (Becraft *et al.*, 1996).

II-C.3.2. WAK

L'ADNc de *WAK1* (Wall associated kinase) a été isolé par la technique de double hybride (Kohorn *et al.*, 1992). Il a été montré que la protéine WAK1 est localisée dans la membrane plasmique et associée fortement à la paroi (He *et al.*, 1996). La sous-famille de RLK à laquelle appartient WAK1 comprend 5 membres chez Arabidopsis (He *et al.*, 1999). L'association à la paroi est une propriété commune aux protéines WAK car le signal d'immunodétection est détecté uniquement à ce niveau de la cellule.

Les transcrits de *WAK1*, *2*, *3* et *5* ont été détectés par northern blot d'abord dans les feuilles et la tige et à l'état de traces dans d'autres tissus. L'expression de *WAK4* est plus importante dans les siliques (He *et al.*, 1999), mais de faibles niveaux de transcrits ont été décelés par RT-PCR dans les feuilles (Lally *et al.*, 2001). L'hybridation *in situ* et la fusion de la protéine GUS⁶ sous contrôle de leur promoteur endogène ont montré que *WAK1* et *WAK2* sont exprimés dans le méristème apical, les feuilles et les méristèmes latéraux racinaires (Wagner et Kohorn, 2001). Leurs profils d'expression se recouvrent partiellement.

La famille des protéines WAK possède un motif EGF dans l'ectodomaine, mais d'autres motifs de l'ectodomaine spécifiques aux différentes WAK sont apparentés aux protéines de la matrice extracellulaire des métazoaires comme la tenascine, le collagène et la neuroxine (He *et al.*, 1999).

Des auteurs ont proposé un rôle de liaison entre la paroi et le cytoplasme (He *et al.*, 1996). Ce lien pourrait être purement structural, indépendant d'une fonction de signalisation supplémentaire, ou représenter une partie intégrale d'une fonction de signalisation (Anderson *et al.*, 2001). En effet, la suppression des transcrits par stratégie antisens de *WAK2* et *WAK4* a pour conséquence un phénotype nain par diminution de l'expansion cellulaire. Cet effet se manifeste au niveau de l'allongement de la tige et des racines et a un effet plus prononcé sur la taille des feuilles. Les protéines WAK 2 et WAK 4 sont nécessaires à l'expansion normale de la cellule. Dans tous les cas, la séquence utilisée pour supprimer les transcrits de *WAK* comporte des régions conservées, et de ce fait doit annuler l'expression des différents membres de la famille de *WAK* (Lally *et al.*, 2001; Wagner et Kohorn, 2001).

⁶ GUS: β- glucuronidase

Encadré 4: Organisation du méristème apical caulinaire (Mayer et al., 1998)

Le méristème apical caulinaire (MAC) des angiospermes prend la forme d'un petit dôme de cellules présentant des caractéristiques structurales spécifiques. L'organisation du méristème apical établi l'arrangement des cellules en deux régions, la tunica et le corpus. Chez *Arabidopsis* et la plupart des autres dicotylédones, la tunica est constituée d'une couche superficielle épidermale (L1) et une couche sous épidermale (L2), chacune d'une épaisseur d'une cellule. Chaque couche reste distincte de l'autre par des divisions cellulaires anticlinales constantes. Le corpus (L3) se tient en dessous de la *tunica* et est constitué de cellules qui se divisent dans tous les plans.

L'épiderme, les feuilles et les fleurs dérivent de la couche L1, tandis que la couche L2 fournit les tissus mésodermiques et les cellules souches. La couche L3 produit les tissus vasculaires.

L'ensemble de ces trois couches cellulaires participe à la croissance de la tige et à la formation des organes; ainsi, la prolifération et le devenir des cellules pendant le développement doivent être coordonnés. Cette coordination peut être réalisée au travers de la voie symplasmique; de récentes expériences par traceur fluorescent dans l'apex ont montré que les couches composant la tunica d'*Arabidopsis* et du hêtre montrent une connectivité symplasmique, par des plasmodesmes primaires et secondaires joignant les cellules et offrant des connections intracellulaires pour le mouvement probablement régulé de molécules signales. Le corpus apparaît former un domaine symplasmique séparé de celui de la *tunica*, mais il a été proposé qu'il maintient la connexion avec la tunica grâce à des plasmodesmes secondaires dans le but d'intégrer les trois groupes cellulaires clonaux.

Des études morphologiques et histologiques indiquent que le méristème apical peut aussi être divisé en trois domaines distincts: une zone centrale, une zone périphérique et une zone "rib".

La zone centrale est située à l'extrémité du méristème et est constituée de cellules se divisant lentement. La zone périphérique entoure la zone centrale; les cellules se divisent plus rapidement et forment les primordia des organes à la périphérie du méristème. La zone "rib" s'étend en dessous des deux autres et produit les cellules qui forment le volume de la tige.

Les cellules de la zone centrale de la tunica apparaissent être isolées par le symplaste de celles de la zone périphérique; elles doivent définir des gradients dans lesquels des petites molécules diffusibles peuvent porter des informations aux cellules à l'intérieur des régions du méristème apical.



II-C.3.3. ERECTA

L'écotype Erecta est une lignée bien décrite d'*Arabidopsis* dans les laboratoires. Le phénotype *erecta* (*er*) est associé à des inflorescences petites et compactes, une tige affinée, de petits pétioles, des feuilles arrondies et des siliques petites et fines. La fonction majeure du gène *ER* est de réguler la forme finale de différents organes et non pas leur mise en place puisque la phyllotaxie n'est pas modifiée chez les mutants (Lease *et al.*, 2001).

Le gène *ER* a été identifié suite à l'insertion d'ADN-T et code un LRR-RLK (Torii *et al.*, 1996). Il est exprimé dans le méristème apical caulinaire et dans les primordia d'organes (Yokoyama *et al.*, 1998). L'ectodomaine de la protéine ER contient 20 motifs LRR dont chacun est codé par un exon différent dans l'organisation génomique du locus. De plus, le domaine kinase cytoplasmique s'autophosphoryle sur des résidus sérine et thréonine *in vitro* (Lease *et al.*, 2001). Bien que le gène *ER* soit caractérisé, les bases cellulaires de ce phénotype n'ont pas été reportées, mais il semble affecter certains aspects de la division et de l'expansion cellulaire.

II-C.4. Développement du méristème

Les méristèmes apicaux caulinaires renferment une population de cellules souches qui se divisent afin de remplacer les cellules incorporées dans des structures différenciées (encadré 4). La fonction propre du méristème requiert une balance entre la prolifération des cellules souches et la différenciation cellulaire. Le gène *CLAVATA1 (CLV1)* chez *Arabidopsis* est impliqué dans le maintient de l'équilibre du méristème apical caulinaire par inhibition de la prolifération cellulaire; les mutants *clv1* ont un méristème anormalement large car la prolifération des cellules souches surpasse la différenciation. L'augmentation de la taille du méristème chez les mutants aboutit à des effets secondaires tels qu'une phyllotaxie altérée, l'augmentation du nombre d'organes floraux, et des siliques en forme de massue (dont le nom CLAVATA dérive) (Clark *et al.*, 1993).

CLV1 code un récepteur de type kinase et est exprimé dans la région du corpus du méristème (Clark *et al.*, 1997). L'ectodomaine contient 21 motifs LRR et le domaine cytoplasmique s'autophosphoryle sur des résidus sérine (Stone *et al.*, 1998).

OsLRK1 a été isolé du riz et la protéine correspondante possède une identité de séquence de 55% avec CLV1 d'*Arabidopsis* (Kim *et al.*, 2000a). Ce gène est exprimé dans les régions en croissance de la partie aérienne, et la suppression des transcrits par antisens cause la production d'organes floraux supplémentaires constituant un phénotype évoquant celui de *clv1*. La fonction *CLV1* semble être conservée entre monocotylédones et dicotylédones (Fletcher et Meyerowitz, 2000).

II-C.5. Signalisation hormonale

Le gène *BRI1* d'*Arabidopsis* a été identifié par une sélection de mutants insensibles aux brassinolides exogènes (mutants *bri1*: **BR-in**sensitive et mutants *bin*: **B**rassinostéroidinsensitive) (Clouse *et al.*, 1996; Li et Chory, 1997). Les brassinolides appartiennent à un type de brassinostéroïdes, hormones stéroïdes végétales, impliquées dans la régulation de la croissance. Elles favorisent l'élongation cellulaire; ainsi les plantes déficientes dans la synthèse des brassinolides (mutants *det2*) ont un phénotype nain lorsque qu'elles sont cultivées à la lumière; alors qu'à l'obscurité les mutants présentent un hypocotyle épais, des cotylédons ouverts et en expansion, et développent des bourgeons de feuilles primaires (Li *et al.*, 1996; Szekeres *et al.*, 1996). Les mutants *bri1* et *bin1* sont des allèles différents du même locus. Les mutants *bri1/bin1* sont phénotypiquement similaires aux plantes qui ont une déficience dans la synthèse des brassinolides (Clouse *et al.*, 1996; Li et Chory, 1997).

Le gène *BRI1* code un LRR-RLK (Li et Chory, 1997). Son domaine extracellulaire contient 25 motifs LRR avec un îlot unique de 70 acides aminés entre les répétitions 21 et 22. Au moins trois allèles mutants présentent des lésions dans cet îlot, ce qui laisse penser que cette région joue un rôle fonctionnel important (Friedrichsen *et al.*, 2000).

L'expression d'une protéine de fusion BRI1-GFP sous le contrôle du promoteur de *BRI1* révèle qu'elle est ubiquiste et localisée dans la membrane plasmique (Friedrichsen *et al.*, 2000).

Une phosphorylation primaire sur les résidus sérine et une phosphorylation d'un niveau inférieur sur des thréonines ont été observées après l'analyse des acides aminés phosphorylés du domaine kinase recombinant de BRI1 (Oh *et al.*, 2000).

Le phénotype normal du riz, peut devenir nain par mutation du gène *d61*, (un gène homologue à *BRI1*) (Yamamuro *et al.*, 2000). La protéine correspondante est composée de 22 motifs LRR, avec un îlot de 70 acides aminés entre les répétitions 18 et 19. De façon similaire aux mutants *bri1* d'*Arabidopsis*, les mutants *d61* montrent une sensibilité réduite aux brassinolides exogènes et un phénotype nain dû à la faible élongation des internœuds et du mésocotyle à l'obscurité.

BAK1 (Bri1-Associated receptor Kinase-1) code un récepteur de type kinase à domaine extracellulaire LRR. La surexpression de *BAK1* chez des mutants faibles de *bri1* supprime de multiples défauts de développement (Li *et al.*, 2002); par exemple les mutants *bri1-5* sont des plantes naines avec de petites inflorescences, alors que les inflorescences du double mutant *bri1-5 bak1-1D* (bri1-associated receptor kinase-1-1Dominant) sont deux fois plus longues que celles des mutants *bri1-5*. Les mutants d'insertion de *bak1* ont des phénotypes semi-nain et la longueur de leur inflorescence est d'environ 70% de celle d'une

plante de phénotype sauvage. De plus, les plantes mutantes de *bak1* produisent des rosettes plus petites avec des feuilles arrondies et des pétioles plus courts. L'implication de la protéine BAK1 dans la signalisation des brassinostéroïdes est renforcée par le fait qu'elle a été identifiée (grâce à la technique du double hybride) comme une protéine interagissant avec BRI1 (Nam et Li, 2002).

II-C.6. Sénescence

SARK (Senescence Associated Receptor-like kinase) a été mis en évidence lors de la sénescence des feuilles chez le haricot par DD-RT-PCR (Differential Display RT-PCR). Le gène SARK code une protéine de 904 acides aminés de type récepteur transmembranaire à activité sérine/thréonine kinase dont le domaine extracellulaire comprend notamment un peptide signal et trois motifs LRR. Lors du vieillissement des feuilles, les transcrits de SARK sont détectés par northern blot à partir de 20 jours après la germination, atteignent un maximum d'accumulation à 40 jours après germination, puis diminuent. A un stade de développement donné, les transcrits de SARK sont plus abondants dans les feuilles les plus âgées, tandis qu'aucun transcrit n'a été détecté dans les jeunes feuilles. Les transcrits de SARK sont présents également en quantité inférieure dans les racines et l'épicotyle 40 jours après la germination (Hajouj *et al.*, 2000).

Un autre gène codant un LRR-RLK semble être impliqué dans la sénescence des feuilles. *SIRK* (Senescence Induced Receptor-like Kinase) a été identifié par cDNA-AFLP chez des plantes d'*Arabidopsis* surexprimant le gène codant AtWRKY6, un facteur de transcription dont l'expression la plus intense a été observée lors de la sénescence des feuilles (Robatzek et Somssich, 2001). Le niveau de transcrit de *SIRK*, est réduit de manière significative dans les feuilles sénescentes de mutants d'insertion déficients pour le gène *WRKY6*, par rapport à une plante sauvage. *SIRK* semble être une cible de AtWRKY6 et l'activation de la transcription de *SIRK* par AtWRKY6 semble être spécifique de la sénescence des feuilles (Robatzek et Somssich, 2002).

II-C.7. Abscission d'organe

RLK5 a été isolé à partir d'une banque d'ADNc d'Arabidopsis criblée avec le domaine kinase de *ZmPK1* de maïs (Walker, 1993), qui est un gène codant un récepteur de type kinase et dont le domaine extracellulaire est apparenté à celui de SRK et à SLG du locus *S* de *Brassica* (Zhang et Walker, 1993). La suppression par anti-sens des transcrits de *RLK5* a mis en évidence son rôle dans l'abscission des organes floraux; celle-ci étant évitée ou



<u>Figure 8:</u> Voie de transduction du signal impliquant le récepteur transmembranaire à activité kinase CLAVATA 1.

A- complexe inactif composé des protéines CLAVATA1 et CLAVATA2 liées par des ponts dissulfures.

B- complexe régulé négativement: la fixation du ligand CLAVATA3 de nature peptidique (peut être en association avec un facteur inconnu) induit l'oligomérisation du complexe inactif et son autophosphorylation. La protéine phosphatase KAPP est phosphorylée par ce complexe, ceci a pour effet de réguler de façon négative le complexe en le déphosphorylant.

C- complexe actif: le complexe phosphorylé est associé à une protéine ROP (directement ou par un intermédiaire inconnu) dont l'action est la répression de *WUSCHEL* (*WUS*), réprimant de ce fait la différenciation cellulaire au profit de la prolifération des cellules souches. La protéine POLTERGEIST (POL), une phosphatase cytoplasmique, fonctionne comme un régulateur négatif de la voie de transduction de CLV1.

(D'après Tichtinsky et al., 2003)

retardée chez les plantes transformées (Jinn *et al.*, 2000). De ce fait, le gène *RLK5* a été renommé *HAESA*, un mot latin signifiant "qui adhère à". Mais, il n'a pas été montré si c'est la différenciation morphologique d'une zone d'abscission qui fait défaut ou si c'est l'abscission proprement dite qui est bloquée.

HAESA est exprimé dans les zones d'abscission des organes floraux. Son expression coïncide avec les stades d'acquisition de la compétence des fleurs à s'autopolliniser c'est-àdire au moment où les zones d'abscission se différencient, les produits du gène sont également observés à la base du pétiole. La protéine HAESA comporte 999 acides aminés et contient 21 motifs LRR dans l'ectodomaine, son domaine cytoplasmique à activité kinase s'autophosphoryle *in vitro* sur des résidus sérine et thréonine. Cette protéine a été localisée par western blot grâce à des anticorps dirigés contre la protéine au niveau de la fraction membranaire de rosettes matures (Jinn *et al.*, 2000).

II-D. VOIES DE SIGNALISATION DES RECEPTEURS DE TYPE KINASE

Pour comprendre complètement la fonction d'un RLK, il est nécessaire de connaître les autres protéines du système, la façon dont elles interagissent, la nature du signal activateur, et les événements qui sont régulés en aval. Lorsqu'ils sont activés par un ligand, les RLK initient des voies de transduction du signal. Ceci implique typiquement le recrutement de protéines effectrices au niveau du complexe récepteur activé, qui déclenchent ainsi des cascades de régulation de phosphorylation de protéines ou d'autres réactions biochimiques, qui aboutissent finalement à des modifications de l'expression de gènes et des fonctions cellulaires. Des approches génétiques et biochimiques commencent à apporter des informations sur quelques systèmes de transduction des RLK et trois des modèles les mieux détaillés seront décrits au cours du chapitre suivant.

II-D.1. Le système CLAVATA1

La voie de transduction du signal impliquant *CLAVATA1* a été bien étudiée. Ainsi, le ligand potentiel, des éléments de transduction intracellulaire du signal et certains gènes cibles ont été identifiés (Figure 8).

II-D.1.1. CLV3 : ligand de CLV1

Les mutants du gène *CLAVATA3* (*CLV3*) d'*Arabidopsis* présentent le même phénotype d'élargissement du méristème que les mutant *clv1*, plaçant par conséquent ces gènes sur la même voie de signalisation (Clark *et al.*, 1995).

CLV3 a été cloné et code une protéine de 96 acides aminés (Fletcher *et al.*, 1999), elle est supposée être sécrétée et clivée pour engendrer une protéine de 78 acides aminés. La masse moléculaire apparente du monomère de CLV3 est de 6 kDa, cependant il est présent dans les extraits protéiques solubles de plantes sous forme d'un multimère de 25 kDa (Trotochaud *et al.*, 2000) dont la nature n'a pas encore été démontrée. CLV3 appartient à une famille de protéines appelée CLE (CLAVATA3/ESR) qui contient un tronçon conservé de 15 acides aminés à proximité de l'extrémité C-terminale; toutes sont supposées être secrétées (Cock et McCormick, 2001). Le génome d'*Arabidopsis* contient 28 gènes de cette famille et trois protéines ESR (Embryo Surrounding Region) de maïs appartiennent aussi à celle-ci. Leur fonction n'est pas connue, mais ces protéines sont détectées dans l'albumen qui entoure le jeune embryon au cours du développement du maïs (Bonello *et al.*, 2002).

CLV3 est exprimé dans un petit groupe de cellules de la zone centrale qui pourrait représenter la population de cellules souches. Cette zone d'expression se superpose légèrement avec celle de *CLV1* dans la couche du corpus (Fletcher *et al.*, 1999). Ainsi, la prolifération de cellules souches semble être régulée par une signalisation entre couches cellulaires.

La protéine CLV3 se fixe aux cellules de levures exprimant CLV1 et CLV2, elle est éluée par chromatographie de gel filtration dans la même fraction que le complexe actif de CLV1, et co-immunoprécipitée avec CLV1 d'extraits de plantes (Trotochaud *et al.*, 2000). Par ailleurs, le complexe actif de CLV1 ne se forme pas chez les mutants *clv3* (Trotochaud *et al.*, 1999). Le fait que la protéine CLV3 se fixe à CLV1 et qu'elle est requise pour la formation d'un complexe actif suggère que CLV3 est le ligand de CLV1.

Certains auteurs ont décrit que, CLV3 ne se fixe pas à la protéine mutante *clv1-10* dans les plantes et les cellules de levures. Cet allèle contient un acide aminé de substitution, au niveau du domaine kinase, qui annule l'activité kinase. Ces résultats suggèrent que l'activité du domaine kinase cytoplasmique est requise pour la fixation efficace du ligand par le domaine extracellulaire (Trotochaud *et al.*, 2000).

II-D.1.2. Les autres éléments du système CLV1

CLAVATA 2

CLV2 est une protéine de 720 acides aminés apparentée à un récepteur qui est composé d'un ectodomaine à motifs LRR, d'un domaine transmembranaire, mais qui ne possède pas de domaine cytoplasmique (Jeong *et al.*, 1999). Bien que l'interaction directe entre les protéines CLV1 et CLV2 n'ait pas encore été démontrée, des éléments génétiques et biochimiques tendent à confirmer cette hypothèse.

Les allèles mutants des gènes *clv1* et *clv3*, possédant de faibles effets sur le phénotype, montrent des relations épistatiques avec les allèles mutants *clv2* dans le méristème apical caulinaire (Kayes et Clark, 1998). Les allèles faibles de *clv1* intéragissent de façon "allèle spécifique" avec *clv2*. Par exemple les doubles mutants *clv1-7/clv2-2* présentent un méristème beaucoup plus large, comparable à celui des mutants forts de *clv1*. Bien que l'expression des transcrits de *CLV1* soit plus importante chez les mutants *clv2* et concomitante avec l'augmentation de la taille du méristème (Kayes et Clark, 1998), les niveaux de protéines CLV1 diminuent de manière significative (Jeong *et al.*, 1999). Ces données pourraient être interprétées en considérant que CLV2 est une sous-unité d'un complexe hétéromérique de CLV1 et est requis pour la stabilité de la protéine CLV1.

Alors que la protéine CLV1 est active exclusivement dans le méristème apical caulinaire, CLV2 dispose de fonctions supplémentaires, en particulier dans le développement des organes floraux (Kayes et Clark, 1998), pendant lequel CLV2 pourrait s'hétérodimériser avec d'autres RLK.

Lorsque des extraits protéiques sont séparés par chromatographie en gel filtration, la protéine CLV1 est présente dans deux fractions de masses moléculaires de 185 kDa et 450 kDa (Trotochaud *et al.*, 1999).

- Le complexe de 185 kDa est présent chez les mutants *clv3* et chez certains mutants *clv1* (Trotochaud *et al.*, 1999). Ce complexe représente la forme inactive, probablement composée des sous-unités CLV1 et CLV2 liées par des ponts dissulfures (Jeong *et al.*, 1999).

- La fraction de 450 kDa correspond à un complexe récepteur actif puisqu'il est absent chez les mutants ayant perdu les fonctions *clv1* ou *clv3* (Trotochaud *et al.*, 1999). D'autres composants doivent se lier au multimère de 185 kDa pour former le complexe actif.

Protéine ROP

La protéine ROP, une GTPase de type Rho, fait partie du complexe actif de CLV1. Elle est décelée dans le complexe protéique immunoprécipité par un anticorps anti-CLV1. Par ailleurs, un immunoblot des fractions d'élution de masses de 185 kDa et 450 kDa contenant CLV1, met en évidence la présence de la protéine ROP dans la fraction du complexe actif uniquement. La fonction de ROP dans la signalisation de CLV1 n'est pas connue, mais elle pourrait agir de façon analogue aux protéines animales de type Ras en activant une voie de transduction impliquant les MAPK (Mitogen Activated Protein Kinase) (Trotochaud *et al.*, 1999).

Protéine KAPP

La protéine KAPP (Kinase Associated Protein Phosphatase) est un autre élément du complexe actif de CLV1. En effet, KAPP est une protéine phosphatase, isolée la première fois du fait de son interaction avec le domaine cytoplasmique de HAESA (Stone *et al.*, 1994). KAPP fixe uniquement la forme phosphorylée de HAESA. Des interactions ont été rapportées ultérieurement entre des KAPP et plusieurs RLK, dont les protéines WAK1 (Park *et al.*, 2001); OsTMK, un RLK dont l'expression est induite par les gibbérellines chez le riz (van der Knaap *et al.*, 1999); et CLV1 (Trotochaud *et al.*, 1999).

KAPP interagit avec le domaine kinase de CLV1 phosphorylé (Stone *et al.*, 1998), et est co-immunoprécipité avec des anticorps dirigés contre CLV1 à partir d'extraits de plantes (Stone *et al.*, 1998). Par ailleurs, les ARNm du gène codant la protéine KAPP sont exprimés dans toutes les zones du méristème apical caulinaire et dans les tissus provasculaires. Cette distribution inclut le domaine d'expression de *CLV1*, mais n'est pas limitée à celui-ci (Williams *et al.*, 1997). Des études de plantes transgéniques indiquent que la protéine KAPP fonctionne comme un régulateur négatif de la signalisation passant par CLV1. En effet, la surexpression de KAPP aboutit à un phénotype similaire aux mutants faibles *clv1* comme par exemple une augmentation faible du nombre de carpelles par fleur, (Williams *et al.*, 1997). Au contraire, la co-suppression des transcrits endogènes de KAPP mène à la suppression du phénotype mutant *clv1-1* (Stone *et al.*, 1998).

POLTERGEIST

Les mutants *pol* (*POLTERGEIST*) suppriment partiellement les phénotypes mutants *clv1*, *2* et *3* mais ne montrent pas de défaut phénotypique (Yu *et al.*, 2000). *POL* code une protéine phosphatase cytoplasmique qui régulerait la voie de signalisation intracellulaire dans plusieurs systèmes de transduction de signaux, dont celui de CLV1 (Yu *et al.*, 2003).

WUSCHEL

Le gène WUSCHEL (WUS) est une cible du système de transduction du signal assuré par les protéines CLV (Schoof *et al.*, 2000). La protéine WUS favorise la prolifération des cellules souches dans le méristème apical caulinaire. En effet, il a été montré que les méristèmes des mutants *wus* épuisent rapidement leur "pool" de cellules souches (Laux *et al.*, 1996). WUS code une protéine à homéodomaine qui est exprimée dans un petit groupe de cellules dans la région médullaire de la zone centrale, juste au-dessous du domaine d'expression de *CLV3* (Mayer *et al.*, 1998). Chez les mutants *clv*, le domaine d'expression de *WUS*. Ainsi la voie de signalisation des protéines CLV inhibe la prolifération cellulaire au travers d'une régulation négative du gène *WUS*. Une boucle de rétrocontrôle de *WUS* agit inversement



Figure 9: Voie de transduction du signal du mécanisme d'auto-incompatibilité chez *Brassica*. **A- complexe inactif**: le récepteur transmembranaire SRK est présent sous forme monomérique ou multimérique dans la membranaire des papilles stigmatiques. Le récepteur est maintenu à l'état non phosphorylé grâce à l'action inhibitrice de thioredoxines H (THL).

B- complexe actif: la fixation du déterminant pollinique d'un pollen incompatible SCR sur le complexe récepteur SRK, stabilisé par la glycoprotéine extracellulaire SLG, permet la levée de l'inhibition des thiorédoxines et l'autophosphorylation de SRK. Le domaine cytoplasmique à activité kinase de SRK phosphoryle et active la protéine ARC1, dont l'activité ubiquitine ligase permet l'ubiquitinisation (Ub) de protéines intracellulaires inconnues, mais dont la dégradation par le protéasome, permet de déclencher la réponse auto-incompatible et le rejet du pollen. En cas de fixation de pollen compatible, la dégradation des protéines par le protéasome n'a pas lieu et le pollen n'est pas rejeté.

C- complexe régulé négativement: le complexe récepteur SRK autophosphorylé est régulé négativement par l'activité phosphatase de la protéine KAPP associée, qui est ellemême phosphorylée par SRK.

(D'après Tichtinsky et al., 2003 et Stone et al., 2003)

comme un régulateur positif de *CLV3*. La surexpression de *WUS* élargit le domaine d'expression de la protéine CLV3 et celle-ci n'est pas exprimée chez les mutants *wus* (Brand *et al.*, 2000).

II-D.2. Le système SRK

Plusieurs éléments composant la voie de transduction impliquant SRK ont été identifiés. Ils comprennent le ligand et des effecteurs intracellulaires du signal (Figure 9).

II-D.2.1. Sp11/SCR est le ligand de SRK

Il est décrit depuis plusieurs années que le déterminant pollinique, tout comme le déterminant stigmatique du système d'auto-incompatibilité de *Brassica* est codé par le *locus S* (Nasrallah, 1997).

Le gène SCR (S-locus Cystéine Rich) aussi appelé SP11 (S-locus protein 11), composé d'une phase ouverte de lecture riche en cystéine, a été découvert lors des analyses de séquences de la région contenant les gènes SRK et SLG (Schopfer *et al.*, 1999). La protéine codée est similaire à une protéine du manteau du pollen préalablement identifiée, PCP-A1 (PCP, Pollen Coat Protein), qui inhibe l'autopollinisation lorsqu'elle est appliquée sur les stigmates (Stephenson *et al.*, 1997). SP11/SCR est exprimée dans les cellules du tapis des anthères et présente un haut degré de variation allélique (Schopfer *et al.*, 1999).

Des plantes homozygotes *S2S2* pour le locus *S* de *Brassica oleacera* transformées avec un allèle *SCR6*, ne peuvent plus polliniser les plantes *S6S6*, mais sont pleinement fertiles en temps que récepteur de pollen dans des croisements réciproques (Schopfer *et al.*, 1999). L'application de la protéine recombinante SCR déclenche la réponse d'autoincompatibilité lorsqu'elle est appliquée aux stigmates du même haplotype, et empêche de ce fait la germination d'un pollen normalement compatible (Kachroo *et al.*, 2001). Ces résultats placent la protéine SP11/SCR dans le rôle de déterminant pollinique de l'autoincompatibilité.

SP11/SCR est, selon l'allèle, une petite protéine de 73 à 83 acides aminés (Schopfer *et al.*, 1999; Takayama *et al.*, 2000). Après le clivage du peptide signal, la protéine mature possède une masse moléculaire relative de 5,7 kDa et contient 8 résidus cystéine qui forment 4 ponts dissulfures (Takayama *et al.*, 2001).

Il a été démontré dans une fraction microsomale *in vitro* que les protéines PCP peuvent activer la phosphorylation des protéines SRK. L'activation est réalisée de façon haplotype spécifique. Les PCP d'haplotypes compatibles n'induisent pas l'activation de SRK,

alors que les protéines PCP de pollens d'haplotypes auto-incompatibles l'activent (Cabrillac *et al.*, 2001). L'interaction directe entre les protéines SCR et SRK a été démontrée. En effet un ectodomaine de SRK6, produit et purifié à partir de feuilles de tabac, interagit avec une protéine SCR6 produite dans un système bactérien. Cette interaction avec SCR6 est approximativement 10 fois plus forte qu'avec un autre haplotype (SCR13) (Kachroo *et al.*, 2001). Ces résultats confirment que SCR/SP11 se fixe à SRK de façon allèle spécifique et active l'autophosphorylation du récepteur.

II-D.2.2. Les autres éléments du système SRK

La glycoprotéine du Locus S (SLG)

La glycoprotéine SLG est codée par un autre gène du locus *S*. Cette protéine contribue à la réponse à l'auto-incompatibilité, mais sa fonction n'est pas encore élucidée (Shiba *et al.*, 2000). Les mutants auto-compatibles qui expriment des niveaux faibles de *SLG* présentent des niveaux normaux de transcrits de *SRK*, alors que la protéine correspondante n'est pas détectée. L'expression de la protéine SRK seule dans le tabac a pour effet la production d'agrégats aberrants de protéines liées par des ponts dissulfures, qui ne se forment pas lorsqu'il y a expression de SLG et SRK. Ces éléments suggèrent que l'expression de la protéine SLG doit faciliter la maturation ou l'accumulation de SRK (Dixit *et al.*, 2000). Cependant certains haplotypes auto-incompatibles de *Brassica* (Suzuki *et al.*, 2000) ainsi que le locus *S* d'*Arabidopsis lyrata* (Kusaba *et al.*, 2001) ne contiennent pas le gène *SLG*, sa présence ne doit donc pas être indispensable. De plus, d'autres études indiquent que la transformation génétique avec *SRK* seul, est suffisante pour conférer un phénotype d'auto-incompatibilité (Takasaki *et al.*, 2000).

Plusieurs autres éléments de la voie de transduction de SRK ont été identifiés.

Protéine phosphatase associée aux kinases (KAPP)

Il a été montré qu'une protéine KAPP d'*Arabidopsis* interagit avec SRK (Braun *et al.*, 1997). Une protéine KAPP homologue chez *Brassica oleacera* interagit avec le domaine cytoplasmique de SRK et la déphosphoryle, tandis qu'elle est elle-même phosphorylée par SRK (Vanoosthuyse *et al.*, 2003). Comme dans le cas de CLV1, le niveau de phosphorylation de SRK peut être modulé et inhibé par une protéine KAPP.

Les Thioredoxines h

Deux clones THL1 et THL2 ont été isolés par criblage en double hybride avec le domaine cytoplasmique de SRK (Bower et al., 1996). Les protéines THL1 et THL2

interagissent avec la protéine SRK et non avec d'autres RLK testés. La protéine SRK interagit avec deux des 5 thioredoxines-h d'*Arabidopsis* (Mazzurco *et al.*, 2001). THL1 est faiblement phosphorylée par SRK et montre une activité thioredoxine *in vitro* (Bower *et al.*, 1996). L'interaction entre ces deux protéines requiert un site catalytique actif pour THL1, un résidu cystéine spécifique sur SRK, et aboutit à une réaction d'oxydo-réduction entre les deux protéines (Mazzurco *et al.*, 2001). Bien que THL2 soit exprimée de façon préférentielle dans les tissus floraux, aucune des deux thiorédoxines n'est spécifique des pistils, ce qui suggère qu'elles possèdent d'autres fonctions en dehors de la signalisation de la voie de SRK (Bower *et al.*, 1996).

SRK s'autophosphoryle de manière constitutive dans les microsomes isolés à partir de stigmates ou dans les cellules d'insectes, mais seulement en réponse à la pollinisation dans les cellules intactes de stigmates. L'addition d'extraits solubles de stigmates inhibe l'autophosphorylation, ce qui suggère qu'un inhibiteur présent dans ce type cellulaire empêche la phosphorylation en absence de ligand. La suppression des thioredoxines de l'extrait lève l'inhibition ; au contraire l'addition d'une THL1 recombinante inhibe l'autophosphorylation *in vitro* de la protéine SRK. Enfin, les protéines du manteau pollinique (contenant SCR/SP11, le ligand de SRK), ajoutées aux réactions de phosphorylation *in vitro* en présence d'extraits de stigmates inhibiteurs, maintiennent l'autophosphorylation, mais exclusivement lorsqu'elles ont été isolées d'un haplotype auto-incompatible avec le variant allélique de SRK. Ainsi l'autophosphorylation de la protéine SRK est inhibée par les thioredoxines et la fixation du ligand supprime l'effet inhibiteur de ces dernières (Cabrillac *et al.*, 2001).

La protéine ARC1

La protéine ARC1 (Arm Repeat Containing 1) a été identifiée par la technique de double hybride avec le domaine cytoplasmique de SRK. L'interaction requiert un domaine kinase actif de SRK, suggérant qu'elle est dépendante de la phosphorylation et que le domaine kinase de SRK phosphoryle ARC1 *in vitro*. L'interaction se produit au niveau de la région C-terminale de la protéine ARC1, qui contient 5 répétitions d'un motif ARM qui permet les interactions protéines-protéines (Mazzurco *et al.*, 2001). Les transcrits de *ARC1* sont exprimés spécifiquement dans les tissus stigmatiques (Gu *et al.*, 1998). La suppression par antisens de l'expression de *ARC1* interrompt la réponse d'auto-incompatibilité (Stone *et al.*, 1999). L'antisens n'a pas d'effet sur les pollinisations compatibles et la surexpression du gène *ARC1* n'affecte pas les pollinisations compatibles, ni incompatibles. La protéine ARC1 est une ubiquitine ligase E3 qui favorise l'ubiquitination de protéines et leur dégradation par le protéasome après les pollinisations auto-incompatibles. Certains inhibiteurs du protéasome ont pour effet d'interrompre la réponse de rejet du pollen. En présence de SRK,



<u>Figure 10:</u> Voie de transduction du signal des brassinostéroïdes impliquant le récepteur transmembranaire à activité kinase BRI1.

A- complexe inactif: le récepteur BRI1 interagit avec le récepteur transmembranaire à activité kinase BAK1. En absence de ligand signal, ce complexe n'est pas phosphorylé. Il en résulte que la protéine kinase cytoplasmique BIN2 est capable de phosphoryler les protéines BES1 et BZR1. Ces protéines ainsi phosphorylées sont soumises à une dégradation protéasomale et ne peuvent avoir d'action. La phosphatase cytoplasmique BSU1 à la capacité de déphosphoryler BZR1 modulant finement son action.

B- complexe actif: le complexe récepteur BRI1/BAK1 fixe un ligand brassinostéroïde. La carboxypeptidase extracellulaire BRS1 semble avoir un rôle dans la signalisation extracellulaire mais le détail du mode d'action n'est pas connu. La fixation du ligand entraîne l'autophosphorylation du complexe récepteur. Le complexe ainsi phosphorylé a une action inhibitrice sur l'activité de BIN2 au travers d'un mécanisme encore inconnu, permettant l'accumulation nucléaire des protéines BES1 et BZR1 qui à leur tour permettent la régulation des gènes cibles du mécanisme de transduction.

(D'après Tichtinsky et al., 2003 et Mora-Garcia et al., 2004)

la protéine ARC1 est localisée au niveau de protéasomes associés au reticulum endoplasmique, ce qui suggère que SRK doit phosphoryler la protéine ARC1, et explique ainsi la localisation protéasomale de ARC1 pour mobiliser des effecteurs potentiels. Ces effecteurs seraient requis pour l'acceptation du pollen et, mèneraient au rejet du pollen lorsqu'ils sont dégradés (Stone *et al.*, 2003).

Ainsi la protéine ARC1 est un élément essentiel de la voie de transduction de SRK et agit comme un effecteur positif de la réponse auto-incompatible.

II-D.3. Le système BRI1

Certains éléments impliqués dans la transduction du signal par BRI1 ont été identifiés (Figure 10)

II-D.3.1. Bri1 fixe les brassinolides

Il a été démontré que les mutants *bri1* sont insensibles aux brassinolides ce qui indique que ce type de récepteur fonctionne en aval ou au niveau du site de perception de cette hormone (Li et Chory, 1997). L'ajout dans le milieu de culture de brassinolides entraîne des réponses de défense aux pathogènes dans les cellules de riz contenant un RLK chimérique composé du domaine extracellulaire de la protéine BRI1 fusionné au domaine cytoplasmique de Xa21, un RLK de résistance au pathogène de riz (He *et al.*, 2000). BRI1 est par conséquent directement impliquée dans la perception des brassinolides et cette protéine possède un fonctionnement similaire à celui des récepteurs aux stéroïdes. Les fractions membranaires d'*Arabidopsis* fixent des brassinolides tritiés et le nombre de sites de fixation augmente dans les plantes transgéniques surexprimant *BRI1* (Wang *et al.*, 2001). Les fractions membranaires des mutants *bri1* dont la mutation provoque une perte d'intégrité du domaine extracellulaire ne fixent pas les brassinolides.

Par ailleurs, les anticorps anti-GFP co-immunoprécipitent les brassinolides tritiés d'extrait de plantes exprimant une protéine de fusion BRI1-GFP, ce qui montre que la protéine BRI1 fixe les brassinolides directement ou au travers d'un complexe.

II-D.3.2. Les autres protéines du système BRI1

Le récepteur kinase associé aux brassinostéroides BAK1

Le criblage par double hybride a montré que la protéine BAK1 interagit avec BRI1 (Nam et Li, 2002). Cette interaction a été confirmée *in vivo* grâce à l'utilisation de plantes transgéniques d'*Arabidopsis* qui expriment les protéines BRI1 et BAK1 et pour lesquelles BAK1 a été mise en évidence au niveau des complexes immunoprécipités de la protéine BRI1 (Li *et al.*, 2002).

Une carboxypeptidase sécrétée BRS1

Un criblage des suppresseurs génétiques du mutant faible *bri1-5* a conduit à l'identification du mutant *brs1-1D* (*bri1 suppressor1-1Dominant*), qui supprime de façon dominante de multiples aspects du phénotype des mutants *bri1* (Li *et al.*, 2001). La surexpression de *BRS1* supprime le phénotype *bri1-5* tandis qu'elle n'a pas d'effet sur le phénotype de plantes sauvages. De plus la mutation *brs1-1D* ne supprime pas le phénotype associé à *bri1* chez les mutants déficients en brassinolides, et il annule les effets phénotypiques dus aux allèles *bri1* impliquant des lésions au niveau du domaine extracellulaire. Cependant, la mutation *brs1-1D* ne supprime pas ces effets chez les mutants *bri1* dont les allèles engendrent des déficiences au niveau du domaine kinase. La protéine BRS1 apparaît donc affecter une étape précoce et spécifique de la signalisation des brassinolides.

D'autre part, un mutant d'insertion de *brs1* ne produit pas d'effets phénotypiques marqués ce qui suggère que ce gène doit être redondant. Cette hypothèse est confortée par le fait que *BRS1* semble coder une carboxypeptidase secrétée et il existe 36 gènes qui codent ce type de protéine chez *Arabidopsis*, parmi lesquels 5 présentent une forte homologie avec BRS1.

Enfin, les plantes *bri1-5* qui surexpriment le gène codant la protéine *BRS1*, mutée au niveau du domaine catalytique, ne retrouvent pas le phénotype attendu, son activité carboxypeptidase est par conséquent requise pour la suppression du phénotype associé au mutant *bri1-5*. La nécessité de cette activité pour la maturation de BRI1, pour un autre composant du signal de transduction de BRI1, ou pour une autre fonction n'a pas encore été démontrée.

Une kinase cytoplasmique BIN2

Un autre gène mutant provoque une insensibilité aux brassinolides. En effet *Bin2* (*brassinolides insensitive 2*) est une mutation semi-dominante de gain de fonction. Le gène *BIN2* a été cloné et code une protéine kinase cytoplasmique apparentée à la protéine GSK3 (Glycogène Synthase Kinase) des mammifères et la kinase SHAGGY de la drosophile (Nam et Li, 2002). La protéine BIN2 agit comme un régulateur négatif de la signalisation passant par BRI1 (Li et Nam, 2002). L'interaction directe entre BRI1 et BIN2 n'a pas été démontrée à ce jour.

Les protéines nucléaires BES1 et BZR1

Les substrats potentiels de BIN2 ont été identifiés, par un criblage génétique et la technique de double hybride. Ces expériences ont conduit à la mise en évidence des protéines BES1 (BRI1-EMS-suppressor1) et BZR1 (brassinazole-resistant-1). Ces protéines

apparentées, contiennent des signaux de localisation nucléaire et plusieurs sites consensus de phosphorylation par des kinases de type GSK3. Il a été montré que BIN2 interagit avec les protéines BES1 et BZR1 et les phosphoryle. De plus, la protéine BIN2 régule de façon négative l'accumulation *in vivo* de celles-ci, et toutes deux sont prédisposées à la dégradation protéasomale. Enfin, la voie de signalisation des brassinostéroïdes inhibe l'activité de BIN2, induisant la déphosphorylation de BES1 et BZR1 et leur accumulation au sein du noyau dans lequel elles activent l'expression des gènes régulés par les brassinostéroïdes (He *et al.*, 2002; Yin *et al.*, 2002; Zhao *et al.*, 2002b).

Une phosphatase nucléaire, BSU1

Récemment, un mutant suppresseur de *bri1* a été identifié : le gène *BSU1* code une phosphatase nucléaire et est exprimé préférentiellement dans les cellules en élongation. La protéine BSU module l'état de phosphorylation de BES1, et s'oppose à l'action de la kinase BIN2, en augmentant le niveau de déphosphorylation de BES1. Le gène *BSU1* appartient à une petite famille de 4 gènes; la diminution des transcrits par ARN interférence a confirmé le rôle de ces phosphatases notamment dans le contrôle de l'élongation cellulaire. La protéine BES1 est donc exposée à des réactions de phosphorylation/déphosphorylation antagonistes, qui permet de réguler finement l'amplitude de la réponse aux brassinolides (Mora-Garcia *et al.*, 2004).

III LE RECEPTEUR A ACTIVITE KINASE LIE A L'EMBRYOGENESE SOMATIQUE

Un clone d'ADNc a été isolé chez la carotte par l'analyse de l'expression différentielle de cultures de cellulaires embryogènes et non embryogènes. Ce clone code un récepteur transmembranaire à activité sérine thréonine kinase et a été nommé *DcSERK* (*Somatic embryogenesis Receptor-like Kinase*) (Schmidt *et al.*, 1997).

III-A. SERK ET L'EMBRYOGENESE SOMATIQUE

L'expression du gène *DcSERK* a été corrélée à la fois quantitativement et qualitativement avec la présence de cellules isolées compétentes à l'embryogenèse somatique. De plus, *DcSERK* est exprimé dans les cellules embryogènes et la présence des transcrits persiste dans les amas de cellules embryogènes ainsi que dans les embryons somatiques jusqu'au stade globulaire. L'expression de la protéine luciférase sous le contrôle du promoteur de *DcSERK* a été observée pendant l'embryogenèse somatique. La plupart des embryons somatique se développent à partir de cellules isolées ou de petits amas de

cellules (4-6) exprimant la luciférase au premier jour d'induction. Les embryons somatiques au-delà du stade globulaire ne montrent plus l'expression de la luciférase confirmant l'expression transitoire de *DcSERK*.

Des résultats similaires ont été observés chez le dactyle pour lequel l'embryogenèse somatique peut être initiée selon le mode indirect ou direct (Somleva *et al.*, 2000). Chez *Arabidopsis*, un gène orthologue à *DcSERK* (*AtSERK1*) exprimé de façon constitutive augmente la capacité des plantules à former des embryons somatiques. Chez des mutants qui ont la capacité de développer des embryons somatiques et dans lesquels le gène codant la protéine GUS sous le contrôle du promoteur *AtSERK1* a été introduit, les embryons somatique se développent à partir de cellules de cals qui expriment GUS. Ainsi *AtSERK1* est lié à la compétence des cellules à l'embryogenèse (Hecht *et al.*, 2001). Par ailleurs, l'expression de *AtSERK1* est augmentée de façon importante chez le mutant *amp1* capable de former des embryons somatiques ce qui reflète ainsi sa forte capacité de régénération.

Ces résultats illustrent l'implication du gène SERK dans la transition des cellules d'un état somatique à l'état embryogène.

La protéine AtSERK1 a été localisée au niveau membranaire. Elle est composée d'un domaine extracellulaire comportant un domaine Leucine zipper impliqué dans l'oligomérisation de la protéine ainsi que 5 répétitions LRR. Une région riche en proline comportant un motif conservé (SPP) précède le domaine transmembranaire. Le domaine intracellulaire est composé du domaine kinase à activité sérine /thréonine kinase capable aussi de s'autophosphoryler sur des résidus tyrosine (Shah *et al.*, 2001b; Shah *et al.*, 2001a).

III-B. SERK ET L'EMBRYOGENESE ZYGOTIQUE

Les transcrits des gènes de type *SERK* sont aussi impliqués dans le processus de l'embryogenèse zygotique de plusieurs espèces. Chez la carotte, les transcrits de *DcSERK* ont été observés dans les embryons en phases précoces de développement jusqu'au stade globulaire. Chez *Arabidopsis*, la protéine GUS sous le contrôle du promoteur de *AtSERK1* a été détectée dans l'ovule en développement. Les transcrits ont été détectés dans tous les constituants du sac embryonnaire. Dans le développement plus tardif de l'embryon, les transcrits de *AtSERK1* ont été observés dans tous les tissus dérivés de la fusion des gamètes. L'expression d'*AtSERK1* persiste dans toutes les cellules de l'embryon jusqu'au stade cordiforme de développement (Hecht *et al.*, 2001). Des résultats similaires ont été observés chez la marguerite, chez laquelle *SERK* est aussi exprimé au cours de l'apomixie, impliquant aussi *SERK* dans un autre processus de reproduction que l'embryogenèse somatique, indépendant de la fécondation (Tucker *et al.*, 2003).

Matériel et Méthodes

I MATERIEL VEGETAL

I-A. HISTORIQUE: DU GENOTYPE '474' AUX GENOTYPES K59 ET K28

La chicorée est utilisée comme modèle d'étude au sein du laboratoire depuis 1985 (Dubois *et al.*, 1988; Guedira *et al.*, 1989). Il existe, chez la chicorée, une variabilité génétique spontanée pour l'aptitude à produire des embryons somatiques *in vitro*. Jusqu'à récemment, les travaux étaient réalisés à partir d'un hybride interspécifique (clone '474') qui représentait la seule source d'embryogenèse somatique chez la chicorée. Le clone '474' qui résultait d'un croisement entre *Cichorium intybus* et *Cichorium endivia* était maintenu *in vitro* par embryogenèse somatique et organogenèse. Les contrôles non embryogènes utilisés pour les différentes études menées étaient des variétés commerciales de *C. intybus* génétiquement très éloignées de l'hybride embryogène.

Dans le cadre d'études visant à mettre en évidence les éléments du déterminisme moléculaire du processus d'embryogenèse somatique, cette distance génétique entre le génotype embryogène et les génotypes non embryogènes est très importante; Il est nécessaire de posséder des génotypes dont les fonds génétiques sont les plus proches et les plus homogènes possibles afin d'identifier avec une meilleure spécificité les gènes dont l'expression est associée à la réponse embryogène.

Des tentatives permettant d'identifier des individus non embryogènes dans des descendances de l'hybride '474' se sont révélées infructueuses du fait de la forte stérilité tant pollinique que femelle du génotype embryogène.

Dès lors la capacité à former des embryons somatiques a été recherchée au sein d'individus de différentes variétés des espèces dont est issu l'hybride '474'. Aucun génotype testé n'a présenté une réponse à l'induction de l'embryogenèse somatique comparable à celle observée pour l'hybride '474' au sein de l'espèce *C. endivia*.

Le caractère embryogène a ensuite été recherché dans la variété population Koospol à l'origine du parent *C.intybus* du clone '474'. Un génotype (K59) a été identifié au sein de cette population pour sa capacité à former des embryons somatiques et à régénérer des plantules lorsqu'il est soumis à des conditions d'induction décrites pour le clone '474'. Bien que *C. intybus* soit une espèce présentant une forte allogamie, quelques graines issues d'autofécondation de K59 ont permis d'obtenir notamment une plante non embryogène (C15).

Parallèlement, une descendance destinée à étudier le déterminisme génétique de l'embryogenèse somatique de la chicorée a été obtenue par croisement de K59 avec K28, une plante elle aussi issue de la population Koospol mais faiblement embryogène.


<u>Figure 11</u>: Cinétique d'obtention des embryons somatiques à partir de fragments foliaires de plantules de génotype 474 de chicorée.

I-B. EMPLOI DES DIFFERENTS GENOTYPES

Plusieurs génotypes ont été utilisés au cours de cette étude.

Le clonage initial visant à mettre en évidence l'existence de gènes codant des récepteurs de type SERK chez la chicorée a été réalisé à partir d'ADN et d'ARN extraits de feuilles soumises aux conditions d'inductions de l'embryogenèse somatique de l'hybride interspécifique '474'. Cette partie des travaux a été réalisée alors que la caractérisation des génotypes issus de la population Koospol était en cours.

Le clonage des deux gènes de type *SERK*, ayant permis d'avoir accès à la structure génique et au polymorphisme de séquence utilisés pour leur cartographie, a été réalisé à partir d'ADN de K59 et K28.

L'étude de la variation du niveau des transcrits au cours de l'induction de l'embryogenèse somatique a été réalisée à partir d'ARN extraits de feuilles soumises aux conditions d'inductions de l'embryogenèse somatique des génotypes K59 et C15.

L'étude de la variation des transcrits au cours de l'embryogenèse zygotique et du développement de la graine a été réalisée à partir d'ARN extraits d'akènes issus du croisement K28 x K59.

II CONDITIONS DE CULTURE

II-A. EMBRYOGENESE SOMATIQUE A PARTIR DE FRAGMENTS FOLIAIRES

Deux phases de culture sont utilisées pour l'obtention des embryons: une phase d'induction suivie d'une seconde au cours de laquelle les embryons somatiques se différencient (phase d'expression) (Figure 11).

Pour les plantes K59 et C15, les feuilles âgées de 2 mois, prélevées sur des plantes cultivées en serre sont aseptisées 20 min dans une solution de Domestos® à 12%, lavées 3 fois 10 min à l'eau stérile puis coupées en fragments de 1,5 x 0,2 cm. Après homogénéisation afin de minimiser la variabilité intra- et inter-plantes, dix fragments sont répartis par fiole.

Pour le clone '474', l'embryogenèse somatique est induite à partir de fragments foliaires issus de plantules cultivées *in vitro*. L'étape d'aseptisation est alors omise.

Au cours de la phase d'induction, les fragments sont cultivés dans 20 mL de milieu de Murashige et Skoog (1962) modifié et dilué au demi, contenant 20 g/L de saccharose et 330 mM de glycérol (M17S20 gly) (Tableau 3). L'addition de glycérol permet d'obtenir une relative synchronisation de la première division des cellules embryogènes pendant les 4 premiers jours de la culture (Robatche-Claive *et al.*, 1992). La culture dure 4 jours à l'obscurité, à 35°C sous agitation.

Composition des milieux de culture	M17S20gly	M17S20	H5	H5 solide
Solution de macroéléments concentrée 10 x	100 mL	100 mL	100 mL	100 mL
Microéléments de Heller	1 mL	1 mL	1 mL	1 mL
Glutamine	250 mg	250 mg	-	-
Vitamines de Morel et Wetmore	10 mL	10 mL	10 mL	10 mL
Inositol	100 mg	100 mg	500 mg	500 mg
2 iP	0,5 mg	0,5 mg	-	-
ANA	0,02 mg	0,02 mg	-	-
KH₂PO₄	85 mg	85 mg	-	-
Solution de Fe-EDTA	5 mL	5 mL	5 mL	5 mL
Saccharose	20 g	20g	5g	5g
Glycérol	24 mL	-	-	-
Agar	-	-	-	6g
H ₂ O		qsp	1L	

M17	Hollor
141.1.7	Hellel
1850	2500
2200	750
8250	
7500	7500
	6000
	1413
	M17 1850 2200 8250 7500

(mg/m)l	
Panthoténate de Ca	1
Thiamine	1
Acide Nicotinique	1
Pyridoxine	1
Biotine	0,01

Milieu de Bourgeonnement Bg0.5	
CaCl ₂ , H ₂ O	4.4 mg
MgSO ₄ , 7H ₂ O	3.7 mg
NH₄NO ₃	16.5 mg
KNO ₃	19 mg
KH₂PO₄	1,7 mg
Solution Fe-EDTA	5 mL
Microéléments (MS)	1 mL
Glutamine	100 mg
Vitamines de Morel et Wetmore	10 mL
Inositol	100 mg
BAP	0,5 mg
Saccharose	20g
pH	5.7

Microéléments Heller (1953) (mg/ml)		
ZnSO ₄ , 7H ₂ O	1	
H₃BO₄	1	
MnSO₄	0,075	
CuSO ₄ , 5H ₂ O	0,03	
AICI ₃	0,03	
NiCl ₂ , 6H ₂ O	0,01	
KI	0,01	
Microéléments de	Murashige et	
Skoog (MS) (1962)	(mg/L)	
H ₃ BO ₃	620	
CoCl ₂	2.5	
CuSO₄	2.5	
CuSO₄ MnSO₄	2.5 1690	
CuSO₄ MnSO₄ KI	2.5 1690 83	
CuSO₄ MnSO₄ KI ZnSO₄, 7H₂O	2.5 1690 83 860	
CuSO₄ MnSO₄ KI ZnSO₄, 7H₂O Na₂MoO₄, 2H₂O	2.5 1690 83 860 25	
CuSO₄ MnSO₄ KI ZnSO₄, 7H₂O Na₂MoO₄, 2H₂O	2.5 1690 83 860 25	

Solution de Fe-ED)TA (g/L)
Na2EDTA	7.46
FeSO4, 7H2O	5.56

<u>Tableaux 3</u>: Composition des solutions nécessaires à la réalisation des différents milieux culture.



Figure 12: Régénération par embryogenèse somatique de chicorée embryogène à partir de racines.

Au quatrième jour de culture, les fragments foliaires sont transférés sur un milieu M17S20 dépourvu de glycérol. A partir de ce moment, on observe les premières divisions cellulaires. La phase d'expression se poursuit pendant 4 jours dans les mêmes conditions que précédemment, à l'obscurité, à 35°C et sous agitation.

Le matériel végétal est prélevé avant mise en culture (J0), à 12h (J0.5), 24h (J1), 36h (J1.5) et à chaque jour de la phase d'induction (J2, J3, J4), au premier (J5) et au quatrième jours de la phase d'expression (J8). Les échantillons sont rapidement congelés dans l'azote liquide et stockés à –80°C pour l'extraction des acides ribonucléiques et des protéines.

II-B. MAINTIEN ET MULTIPLICATION PAR EMBRYOGENESE SOMATIQUE ET PAR ORGANOGENESE *IN VITRO*

Les plantes embryogènes peuvent être maintenues par embryogenèse somatique et par organogenèse, tandis que les plantes non embryogènes ne peuvent l'être que par organogenèse.

II-B.1. Maintien par embryogenèse somatique

Les plantes de génotypes embryogènes sont régénérées à partir des embryons somatiques issus de fragments racinaires. Ils présentent en effet l'avantage de se détacher facilement du tissu par rapport aux embryons somatiques issus de feuilles (Figure 12).

Les embryons somatiques sont initiés à partir de fragments racinaires de plantes cultivées *in vitro* de la même façon que les fragments foliaires (cf. § II-A) à la différence que le milieu d'induction est exempt de glycérol. Après 10 jours de culture en milieu M17S20 les embryons somatiques sont cultivés 10 à 15 jours dans un milieu de Heller (H5) comprenant 5 g/L de saccharose (Tableau 3) sous agitation, à la lumière pendant 16 h à 24°C et 8 h à l'obscurité à 20°C. Le développement en plantules des embryons somatiques chlorophylliens se poursuit en boîte de pétri sur un milieu H5 (Tableau 3) avec la même photopériode.

Les plantules sont alors repiquées en tube sur le même milieu et cultivées dans les mêmes conditions. Après 6 à 8 semaines, certaines plantules sont acclimatées en serre tandis que d'autres sont utilisées pour la multiplication du génotype.

II-B.2. Maintien par organogenèse

Des feuilles de plantes cultivées en serre sont prélevées et aseptisées comme décrit précédemment (cf § II-A). Des fragments incluant la nervure centrale, zone la plus réactive, sont cultivés 15 à 20 jours en boîte de Pétri sur un milieu de bourgeonnement Bg0.5 (Tableau 3) avec la même photopériode que les plantules (cf § II-B.1). L'enracinement des

bourgeons qui présentent des feuilles d'au moins 1 cm est réalisé en tube contenant du milieu H5 solide. Après 6 à 8 semaines de culture, les plantules sont acclimatées en serre.

II-C. EMBRYOGENESE ZYGOTIQUE ET MATURATION DES GRAINES

Des pollinisations manuelles ont été réalisées par mise en contact des étamines du génotype K59 sur les pistils de plantes de génotype K28. Les capitules sont récoltés à intervalles réguliers jusqu'à 19 jours après pollinisation, et disséqués afin de ne conserver que les akènes en développement qui sont rapidement congelés dans l'azote liquide et conservés à –80°C pour l'extraction des acides ribonucléiques et des protéines.

III EXTRACTION DES ACIDES NUCLEIQUES ET PROTEINES

III-A. EXTRACTION DE L'ADN

L'extraction de l'ADN est réalisée au moyen du kit GenElute Plant genomic DNA kit (Sigma) selon les instructions du fabricant. Environ 100 mg de tissu sont broyés dans l'azote liquide, puis incubés 10 min à 65°C dans 400 μ L d'un mélange de tampon de lyse. Après précipitation des débris cellulaires, des protéines et des polysaccharides, par ajout de 130 μ L d'un tampon de précipitation et centrifugation 5 min à 15000 g, le surnageant est filtré. L'ADN est purifié par fixation sur une colonne préalablement équilibrée puis élué, après lavage de la colonne, par deux fois 100 μ L d'une solution d'élution. La concentration de la solution d'ADN ainsi obtenue est déterminée par spectrophotométrie à 260 nm.

III-B. EXTRACTION DES ARN TOTAUX

L'extraction des ARN est réalisée selon le protocole du kit TRI-REAGENT® (Euromedex). Cent milligramme environ de tissus broyés dans l'azote liquide sont homogénéisés dans 1 mL de TRI-REAGENT®. Après une incubation de 5 min à température ambiante, 200 µL de chloroforme sont ajoutés. Le mélange est vigoureusement agité, incubé 10 min à température ambiante, puis centrifugé à 20000 g, à 4°C pendant 15 min. La phase aqueuse supérieure contenant les ARN est transférée dans un nouveau tube, et les ARN sont précipités par centrifugation à 12000 g, à 4°C pendant 10 min après ajout de 500 µL d'isopropanol. Le surnageant est éliminé et le culot est lavé par 1 mL d'éthanol à 75% dans l'eau DMPC. Après centrifugation 5 min à 7500 g, l'éthanol est éliminé, le culot est séché à l'air et repris dans 20 µL d'eau DMPC. Le dosage est réalisé par mesure de l'absorbance à 260 nm.

III-C. EXTRACTION ET DOSAGE DES PROTEINES

A la suite de l'extraction des ARN, les protéines sont extraites suivant le protocole du kit TRI-REAGENT® (Euromedex). Après élimination de la phase aqueuse supérieure contenant les ARN, 300 μ L d'éthanol absolu sont ajoutés. Après homogénéisation par inversion et incubation 2 min à température ambiante, l'ADN et les débris cellulaires sont précipités par centrifugation à 5000 g à 4°C pendant 5 min. Le surnageant contenant les protéines est transféré dans un nouveau tube. Les protéines sont précipitées par l'ajout de 1,5 mL d'isopropanol, et collectées, après une incubation de 20 min à température ambiante, par centrifugation à 15000 g à 20°C pendant 15 min. Le surnageant est éliminé et le culot protéique est lavé trois fois 20 min dans 300 mM d'hydrochlorure de guanidine dans l'éthanol 95%. L'hydrochlorure de guanidine est éliminé et le culot protéique est lavé trois fois 20 min dans 300 mM d'hydrochlorure de guanidine dans l'éthanol absolu. Le culot protéique peut être conservé dans l'éthanol absolu à –20°C. Après élimination de l'éthanol et séchage du culot, les protéines sont solubilisées dans un tampon UKS modifié (Damerval *et al.*, 1986) composé de 5 mM de carbonate de potassium (K₂CO₃) et contenant 9,5 M d'urée, 2% de Triton X-100, 1,25 % de SDS et 3 mM de DTT.

Le dosage des protéines est réalisé en utilisant le kit RcDc Protein Assay (BioRad) selon le protocole du fabricant. Cette méthode est adaptée de celle décrite par Lowry (Lowry *et al.*, 1951) et permet après réaction de développer une coloration bleue quantifiée par spectrophotométrie à 750 nm et d'intensité proportionnelle à la concentration de protéine.

Pour le dosage, 2 µL d'extraits protéiques sont dilués dans 25 µL d'UKS. Vingt cinq µL d'une gamme de concentration croissante en albumine de sérum de bœuf (BSA) (0,2; 0,5; 0,8; 1; 1,3; 1,5 mg/ml) est réalisée dans le tampon UKS identique à celui dans lequel les protéines de chicorée sont solubilisées. La gamme est traitée simultanément avec les échantillons protéiques à doser. La concentration des extraits protéiques est calculée par référence à la gamme étalon de concentration en BSA.

IV AMPLIFICATION D'ADN

L'amplification d'ADN est réalisée à partir d'environ 2 ng d'ADN génomique ou une dilution au dixième d'un produit de transcription inverse obtenu à partir de 5 µg d'ARN total. La PCR est réalisée dans un tampon adapté 1x en présence de 2,5 mM de MgCl₂, 250 nM de chacune des amorces spécifiques de la cible à amplifier, 0,2 mM de chacun des déoxynucléotides et 0,75 U d'ampliTaq ADN polymérase (Applied Biosystems). Le fragment d'ADN cible est amplifié grâce à une dénaturation initiale de 4 min à 95°C, suivie de 40 cycles à trois étapes composés d'une dénaturation de 30 s à 95°C, une hybridation des amorces de 30 s à la température de fusion, caractéristique de chaque couple d'amorces, suivi d'une élongation des amorces 1 min 30 à 72°C. L'amplification se termine par une élongation terminale de 7 min à 72°C.

47



<u>Figure 13</u>: Caractéristiques du vecteur de clonage pGEM-T (PROMEGA). Le vecteur pGEM-T confère aux bactéries transformées la résistance à l'ampicilline.

Amorces	séquences
M13F	5' GTTTTCCCAGTCACGAC 3'
M13R	5' CAGGAAACAGCTATGAC 3'
T7	5' TAATACGACTCACTATAGG 3'
SP6	5' ACTCAAGCTATGCATCCAACG 3'

<u>Tableau 4</u>: séquences des amorces permettant l'amplification et le séquençage d'un fragment d'ADN produit par PCR et cloné dans le vecteur pGEM-T (PROMEGA).

La totalité ou une aliquote des produits d'amplifications est séparée par électrophorèse en gel d'agarose à 1,2% contenant 5.10^{-3} % (v/v) d'une solution à 10 mg/ml de bromure d'éthydium (BET). Les produits d'amplification sont visualisés sous U.V.

V CLONAGE ET SEQUENÇAGE DE PRODUITS DE PCR

V-A. CLONAGE DANS LE VECTEUR PGEM-T

A la suite des réactions de PCR, les fragments d'ADN obtenus sont purifiés à l'aide du kit de purification de produit PCR (QIAGEN) et élués dans 50 μ L de tampon EB (Tri-HCl 10 mM, pH 8.5). Ils sont par la suite clonés dans le plasmide pGEM-T (PROMEGA) (Figure 13) dans un volume de 10 μ L. Cinquante ng de vecteur sont ajoutés à une quantité d'ADN déterminée auxquels est ajouté 1 μ L d'ADN ligase du phase T4. La quantité de produit de PCR nécessaire à la ligation est dépendante de la taille du fragment à cloner, de la taille et de la concentration du vecteur et peut être déterminée selon la formule:

Quantité insert (ng) = $\frac{3}{1} \times \frac{\text{quantité vecteur (ng)}}{\text{taille vecteur (kb)}} \times \text{Taille insert (kb)}$

La ligation est réalisée une nuit à 4°C.

V-B. TRANSFORMATION ET SELECTION DE COLONIES BACTERIENNES RECOMBINANTES

Des bactéries compétentes (100 μ L) *E.Coli* JM109 (PROMEGA) sont transformées par choc thermique de 45 s à 42°C avec la totalité du produit de ligation. Les bactéries sont ensuite cultivées 1 heure à 37°C dans 1 mL de milieu SOC (2% tryptone; 0,5% extraits de levure; 10 mM NaCl, 2,5 mM KCl, 20 mM Mg²⁺, 20 mM glucose, pH 7.0), étalées sur un milieu LB (1% tryptone, 0,5% extraits de levure, 10% NaCl, 15% agar, pH 7.0) contenant 50 μ g/mL d'ampicilline, 80 μ g/mL de X-Gal et 0,5 mM d'IPTG (Isopropyl- β -D-thiogalactopyranoside) puis cultivées une nuit à 37°C.

Chaque colonie bactérienne transformante sélectionnée est lysée dans 50 μ L de tampon TTE (1% Triton X-100, 20 mM Tris-HCl, 2 mM EDTA) 5 min à 95°C, afin de libérer le plasmide. Une centrifugation de 10 min à 10000 g permet de culotter les débris cellulaires.

Un second crible est réalisé par PCR dans les conditions décrites précédemment à partir de 2 µL de produits de lyse bactérienne (cf § IV), afin de sélectionner les clones dont le fragment inséré correspond à la taille attendue. Cette PCR est réalisée en utilisant les amorces universelles M13 situées de part et d'autre du site de clonage (Figure 13, Tableau 4).

V-C. SEQUENÇAGE D'ADN

Dans le but de préparer la matrice qui servira au séquençage de clones sélectionnés, une PCR initiée par les amorces M13 est réalisée dans un volume de 60 μ L dans les mêmes conditions décrites précédemment. Les produits de PCR sont précipités dans 1/10 de volume d'une solution à 3 M d'acétate de sodium, pH 5.2 et 2,5 volumes d'éthanol absolu pendant 1 heure à -80°C avant d'être centrifugés 20 min à 15000 g. Le surnageant est éliminé et le culot est rincé deux fois à l'éthanol 70%, séché, puis repris dans 25 μ L de tampon EB.

Les produits de PCR purifiés sont utilisés pour le séquençage selon la méthode de Sanger. La réaction est initiée par des amorces universelles SP6 ou T7 marquées par le fluorochrome IRD800 (Infra Red Detection) et dont les sites d'hybridation sont aussi présents sur le vecteur pGEM-T (Figure 13). La réaction de séquençage est mise en oeuvre au moyen du kit "ABI PRISM Dye Primer Cycle Sequencing Core Kit" (Applied Biosystem) selon les conditions du fournisseur. Quatres mélanges réactionnels correspondant à chacune des bases sont réalisés. Chaque réaction contient du tampon 1x, 0,2 µM d'amorce marquée, un mélange de déoxynucléotides incluant une base didéoxynucléotide et l'enzyme Taq polymérase FS. La réaction est exécutée selon le programme suivant:

température	durée	Nombre de cycles
96°C	10 s	
48°C	5 s	1 5
70°C	60 s	
96°C	10 s	15
70°C	60 s	

La réaction est stoppée par ajout d'un demi-volume de tampon d'arrêt (95% formamide, 10 mM EDTA, 0,1% bleu de bromophénol, 0,1% xylène cyanol). Les fragments synthétisés sont dénaturés 3 min à 95°C puis refroidis rapidement dans la glace. Environ 1,5 µL de produit de séquence sont déposés dans un gel d'acrylamide à 3,75% dénaturant et soumis à une électrophorèse. Le séquençage est réalisé par un séquenceur automatique Li-COR 4200 qui détecte la fluorescence émise par le fluorochrome (longueur d'onde 820 nm) excité grâce à un faisceau laser basse tension (longueur d'onde 785 nm).



Figure 14 Principe de la RT-PCR inverse

VI CLONAGE DE FRAGMENTS D'ADN CORRESPONDANT A UNE PORTION DE DOMAINE KINASE DE RECEPTEUR DE TYPE KINASE

L'identification de séquences homologues entre différentes espèces, en utilisant la technique de PCR, repose tout d'abord sur la détermination des amorces nécessaires à l'amplification. Ces amorces doivent en effet permettre d'amplifier les séquences nucléotidiques correspondant à des séquences peptidiques conservées. De manière habituelle, ces amorces sont dégénérées sur toute leur longueur et incluent parfois des inosines qui agissent comme des bases de faible spécificité et permettant de diminuer le niveau de dégénérescence des amorces. Rose *et al.* (1998) ont décrit une nouvelle stratégie de définition d'amorces dégénérées pour amplifier des cibles chez une espèce donnée à partir d'alignements multiples de séquences protéiques identiques chez d'autres espèces. Chaque amorce est constituée d'une région 3' dégénérée et d'une région 5' plus longue, non dégénérée permettant l'ancrage à la séquence cible. Entre 3 et 4 acides aminés fortement conservés sont nécessaires pour définir la région 3' de l'amorce qui est alors stabilisée par la région 5', moins conservée, lors de l'hybridation de l'amorce à la matrice. La définition des amorces est gérée par un algorithme informatique, CODEHOP, disponible sur le réseau internet (http://blocks.fhcr.org/codehop.html).

Dans un premier temps des alignements de séquences ont été réalisés entre des protéines présentant de fortes similitudes avec le produit du gène *SERK* de la carotte. Ces alignements montrent des régions fortement conservées correspondant au domaine kinase et dans lesquelles le logiciel CODEHOP nous a permis de définir les amorces dégénérées.

L'amplification des régions kinases conservées a été réalisée dans les conditions décrites à la section IV, à partir d'1 ng d'ADN total de chicorée génotype '474' en utilisant les amorces décrites dans la partie résultat. Les produits ont été clonés et séquencés comme indiqué auparavant (cf § V).

VII OBTENTION DES SEQUENCES CODANTES PLEINES LONGUEURS VII-A. RT-PCR INVERSE

Cette technique, développée au cours de mon DEA (Duban, 2000) est une adaptation aux séquences codantes de la technique de PCR inverse (Ochman *et al.*, 1988) destinée à amplifier les régions adjacentes à une région connue d'ADN. Le principe est illustré figure 14: à la suite d'une transcription inverse des ARN par l'utilisation d'amorces polydT possédant en 5' un site de restriction, les ADNc sont polyadénylés. Le second brin d'ADNc est alors synthétisé en utilisant l'amorce qui a permis la transcription inverse. L'ADNc double brin ainsi formé est soumis à une digestion enzymatique puis à une ligation lui permettant de se circulariser. Un fragment correspondant aux régions flanquantes de la région connue est obtenu par deux PCR successives (nested PCR).

La synthèse du premier brin d'ADNc est réalisée à partir d'ARN totaux extraits d'explants foliaires du génotype embryogène '474' au cinquième jour d'induction de l'embryogenèse somatique en utilisant le kit "Superscript Preamplification system for first strand synthesis" (GibcoBRL). Dans un volume de 12 µL, 5 µg d'ARN et 0,5 µg d'amorces modifiées oligo(dT) en 5' pour posséder un site de restriction Nsil (5' ACTCAAGCTATGCATCCAACG(T)16 3') sont incubés 10 min à 70°C, puis plongés dans la glace pour éliminer les stuctures secondaires. Le volume réactionnel est complété par le tampon 1x, 2.5 mM MgCl₂, 0,5 mM dNTP, 10 mM DTT. Ce mélange de 19 µL est incubé 5 min à 42°C, puis 200 U de transcriptase inverse sont ajoutées. La synthèse du premier brin d'ADNc est réalisée pendant 50 min à 42°C et la réaction est stoppée par l'inactivation de l'enzyme pendant 15 min à 70°C. L'ajout de 2 U de RNase H dans le mélange et une incubation 20 min à 37°C permettent d'éliminer les ARN. Les ADNc produits sont purifiés au moyen du kit de purification de produit PCR (QIAGEN), élués par 50 µL de tampon EB puis concentrés jusqu'à un volume final de 35 µL.

La polyadénylation de la totalité des ADNc est réalisée pendant 15 min à 37°C dans un volume final de 50 µL comprenant le tampon 1x de la transférase terminale (Biolabs), 25 U d'enzyme, 1,5 mM CoCl2 et 0,2 mM ATP. La transférase terminale est ensuite inactivée par une incubation de 10 min à 70°C.

La synthèse du second brin d'ADNc est réalisée sur la totalité des produits précédemment obtenus dans un volume de 100 µL comprenant le tampon de la Taq polymérase, 2 mM MgCl₂, 0,2 mM dNTP, 0,2 µM d'amorce poly(dT)Nsil et 2.5 U d'AmpliTaq®Gold DNA polymerase (Applied Biosystems). Après 7 min d'activation de l'enzyme à 95°C, les amorces sont hybridées à 36°C pendant 30 s. L'élongation est achevée en 30 min à 72°C après un passage lent de 36°C à 72°C. Les ADNc doubles brins sont purifiés de la même manière que précédemment et élués par 80 µL de tampon EB.

L'amplification des ADNc doubles brins est réalisée à ce stade dans l'objectif d'obtenir une quantité suffisante d'ADN pour réaliser les autoligations. L'amplification est effectuée dans un volume de 20 µL à partir de 1 µL d'ADNc doubles brins, en utilisant 0,2 µM d'amorce ayant permis la synthèse des ADNc, en présence de 4 mM MgCl₂, 0,2 mM dNTP, 1 U d'AmpliTaq®Gold DNA polymerase (Applied Biosystems). Après l'activation de la polymérase 7 min à 95°C, 16 cycles de 3 étapes sont effectués comprenant une étape de dénaturation de 30 s à 95°C, l'étape d'hybridation des amorces 30 s à 60°C et l'étape d'élongation de 2 min 30 à 72°C. Ces 16 cycles sont suivis de 7 autres cycles de 3 étapes comprenant la dénaturation 30 s à 95°C, l'hybridation des amorces 30 s à 60°C et l'élongation 2 min 30 s à 72°C avec une extension du temps d'élongation de 5 s par cycles



Figure 15: Principe de la 5'RLM-RACE PCR utilisée par le kit GeneRacer (INVITROGEN)

afin de favoriser l'amplification des fragments les plus longs. A la suite de cette PCR, les produits d'amplification sont purifiés et élués par 50 µL de tampon EB puis concentrés jusqu'à un volume final de 40 µL.

La digestion des produits de PCR est réalisée à partir de la totalité des produits de PCR dans un volume de 50 µL par 5 U d'enzyme de restriction Nsil (Biolabs) pendant 2 heures à 37°C. Les produits sont ensuite purifiés puis élués comme précédemment.

L'autoligation des produits digérés est obtenue dans un volume de 500 µL par ajout de 70 U d'ADN ligase du phage T4 (Biolabs) pendant une nuit à 16°C.

Deux PCR inverses successives sont ensuite réalisées, la première amplification à partir de 1 μ L des produits d'autoligation dans un volume de 15 μ L en présence de 0,75 U d'AmpliTaq®Gold DNA polymerase (Applied Biosystems), 2 mM MgCl₂, 0,2 mM dNTP et 0,25 μ M de chacune des amorces selon le programme présenté précédemment (cf § IV) avec 25 cycles d'amplification comprenant l'hybridation des amorces à 53°C et l'élongation des fragments pendant 2 min. La seconde PCR est réalisée à partir de 1 μ L de produit d'amplification, issu de la première PCR et dilué 10 fois, dans les mêmes conditions mais avec 35 cycles d'amplification. Les amorces utilisées sont définies à l'intérieur du fragment précédent permettant une meilleure spécificité d'amplification (cf partie résultats).

A l'issue de la seconde PCR, les produits d'amplification sont clonés dans le vecteur plasmidique pGEM-T (PROMEGA) (cf § V-A) et séquencés (cf §V-C).

VII-B. 5' RACE PCR

Cette technique d'amplification des extrémités 5' d'ADNc (RACE, Rapid Amplification cDNA Ends) est utilisée pour le clonage des extrémités 5' des séquences codantes qui n'ont pas été obtenues par RT-PCR inverse. La technique de RLM-RACE (RNA Ligase Mediated – RACE) utilisée réside dans la ligation d'oligoribonucléotides aux extrémités 5' d'ARN messagers dont la coiffe a été éliminée (Maruyama et Sugano, 1994). L'amplification des extrémités 5' est obtenue en utilisant des amorces spécifiques de l'oligoribonucléotide ligué et de la séquence d'intérêt (Figure 15). Le clonage des extrémités 5' des ADNc a été réalisé grâce au kit GeneRacer (INVITROGEN).

Après obtention des ARN messagers purifiés à partir de 500 µg d'ARN totaux au moyen du kit Direct Quick messenger RNA (TALENT), la déphosphorylation de 2 µg d'ARNm permet d'éliminer la possibilité d'une ligation ultérieure de l'oligoribonucléotide sur les ARNm tronqués ne possédant pas de coiffe en 5'.

La déphosphorylation des ARNm tronqués est obtenue à l'issue d'une 1 heure d'incubation à 50°C dans un volume de 10 µL comprenant un tampon 1x, 40 U d'inhibiteur de RNase (RNaseOUT[™]) et 10 U de phosphatase intestinale de veau. Les ARNm ainsi traités

sont précipités une nuit à -20°C dans 2,5 volumes d'ethanol absolu et 10% d'une solution de 3 M d'acétate de sodium pH 5.2, centrifugés 15 min à 15000g à 4°C, lavés 3 fois à l'éthanol 80% dans l'eau DEPC, puis le culot est resuspendu dans 7 µL d'eau DEPC.

La destruction de la coiffe de la totalité des ARNm est réalisée dans un volume de 10 µL comprenant le tampon 1x, 0,5 U de pyrophosphatase acide de tabac et 40 U d'inhibiteur de RNase (RNaseOUT[™]) pendant 1 heure à 37°C. Cette destruction laisse un groupement phosphate libre en 5'. Les ARNm sont ensuite précipités et les culots resuspendus comme précédemment.

La ligation de l'oligoribonucléotide est effectuée grâce au groupement phosphate libre en 5' des ARNm traités. La totalité des ARNm est ajoutée à 250 ng d'oligoribonuclétides lyophilisés et incubée 5 min à 65°C. La ligation est réalisée dans un volume de 10 µL par 5 U d'ARN ligase du phage T4 en présence du tampon adéquat, 1 mM ATP, 40 U d'inhibiteur de RNase (RNaseOUT[™]) pendant 1 heure à 37°C. Les ARNm sont ensuite précipités et resuspendus dans 10 µL d'eau DEPC.

La transcription inverse est opérée à partir de la totalité de la solution d'ARNm. Dans un volume de 10 µL, les ARNm sont incubés 15 min à 65°C en présence de 0,5 µg d'amorce oligo(dT) et de 0,5 mM de dNTP puis mis dans la glace pour éliminer les structures secondaires. Le volume réactionnel est complété par le tampon 1x, 5 mM MgCl₂, 10 mM DTT et 40 U d'inhibiteur de RNase (RNaseOUT[™]). Ce mélange de 19 µL est incubé 5 min à 42°C puis 50 U de transcriptase inverse SuperSript[™]II sont ajoutées. La synthèse du premier brin d'ADNc se déroule pendant 50 min à 42°C et la réaction est stoppée par l'inactivation de l'enzyme pendant 15 min à 70°C. L'ajout de 2 U de RNase H dans le mélange et une incubation 20 min à 37°C permettent d'éliminer les ARNm.

L'amplification des extrémités 5' est réalisée dans 50 µL à partir de 1µL de produits de transcription inverse dilués 10 fois par l'ADN Taq polymérase Platinum® High Fidelity (Invitrogen) dans du tampon 1x en présence de 200 µM de dNTP et de 2 mM MgSO₄ en utilisant 0,6 µM d'amorce directe spécifique de l'oligonucléotide ligué et 0,2 µM d'amorce complémentaire définie sur la partie de la séquence codante connue.

Température	Durée	Nombre de cycles
94°C	2 min	1
94°C	30 s	F
72°C	1 min 30 s	5
94°C	30 s	E
70°C	1 min 30	5
94°C	30 s	
60°C	30 s	25
68°C	1 min 30 s	
68°C	10 min	_ 1

Les conditions d'amplification sont les suivantes:

Une seconde PCR (nested PCR) augmente la spécificité des produits obtenus. Elle nécessite aussi des amorces définies à la fois sur l'oligonucléotide ligué et sur la séquence connue (Figure 15) et est réalisée à partir de 1 µL de produit issu de l'amplification précédente en utilisant le même mélange réactionnel. Seules les conditions d'amplification changent: après une dénaturation initiale de 2 min à 94°C, 25 cycles de trois étapes comprenant 30 s de dénaturation à 94°C, 30 s d'hybridation des amorces à 60°C et 2 min d'élongation à 72°C sont réalisés suivis d'une extension finale de 10 min à 68°C.

Après vérification par électrophorèse en gel d'agarose des produits de PCR, les fragments sont clonés et séquencés comme décrit précédemment (cf § V-A, V-C).

VIII ANALYSE DES SEQUENCES

Les alignements de séquences protéiques sont réalisés au moyen du logiciel ClustalX (Thompson *et al.*, 1997) selon des matrices de comparaison Gonnet (Gonnet *et al.*, 2000). Les arbres phylogénétiques sont construits à partir de cet alignement selon la méthode du neighbour joining (Saitou et Nei, 1987). La robustesse des nœuds identifiés est évaluée par le calcul d'une valeur de boostrap estimée à partir de 1000 procédures d'échantillonnage. Les arbres phylogénétiques sont visualisés au moyen du logiciel TreeView (http://taxonomy.zoology.gla.ac.uk/ rod/rod.html).

IX HYBRIDATION SOUTHERN

IX-A. DIGESTION ENZYMATIQUE DES ADN

Environ 10 μ g d'ADN sont digérés par 100 U de chacune des enzymes de restriction MsII et RsaI (Biolabs) une nuit à 37°C dans un volume de 200 μ L en présence de leur tampon 1x respectif et de 10 mM de spermidine. Les produits de digestion sont ensuite précipités et les culots repris dans 25 μ L de tampon EB avant d'être soumis à une électrophorèse en gel d'Agarose à 0,8% dans un tampon TAE 1x pendant une nuit à 35V.

IX-B. TRANSFERT SUR MEMBRANE

Les acides nucléiques digérés et séparés sont transférés sur une membrane de nylon. Lors du transfert sous vide, les acides nucléiques sont dépurinés 20 min par une solution d'HCI 0,25 M, puis dénaturés 20 min par une solution de NaCI 1,5M, NaOH 0,5 M avant d'être neutralisés 20 min par une solution de Tris-HCI 1M, NaCI 1,5M. Le transfert se poursuit alors pendant 2 heures au moyen d'une solution SSC 20x (NaCI 3M, Citrate de Na 0,3M). Les acides nucléiques sont alors fixés à la membrane par irradiation ultraviolette (1,2 J/M²) pendant 5 min.

IX-C. HYBRIDATION DES MEMBRANES

Les sondes sont préparées par PCR à partir de 1ng de plasmide contenant la séquence d'intérêt au moyen du kit "PCR dig labeling mix" (Roche) selon le protocole du fournisseur. Les fragments d'ADN synthétisés incorporent des dUTP couplés à la digoxygénine.

Afin de saturer les sites de fixation non spécifiques, les membranes sont préhybridées 4 à 6 heures à 68°C dans un tampon SSC 5x contenant 0,02% de SDS, 0,4% de laurylsarcosine et 4 % d'agent bloquant (Roche).

Les membranes sont hybridées pendant 12 heures à 68°C dans 30 mL de solution de préhybridation contennant la totalité de la sonde marquée produite par PCR et dénaturée préalablement 5 min à 95°C.

IX-D. REVELATION DES MEMBRANES

Les membranes hybridées sont lavées 2x 5 min dans 400 mL de solution SSC 2x; 0,1% SDS, puis 2x 5 min dans 400 mL de solution SSC 0,2x; 0,1% SDS à température ambiante. Les membranes sont ensuite incubées 2x 20 min dans une solution SSC 0,2x; 0,1% SDS à 45°C puis rincées dans un tampon I (0,1M Tris, 0,15 M NaCl). Les membranes sont alors incubées 45 min à température ambiante dans le tampon I contenant 1% d'agent bloquant, puis incubées 30 min dans 20 mL de tampon I contenant 1% de SDS et 4 µL d'anticorps anti-digoxygénine couplés à la phosphatase alcaline. Après 2 rinçages de 20 min dans le tampon I, l'hybridation des anticorps est révélée par chemiluminescence: les membranes sont placées pendant 30 min à l'obscurité à 4°C dans 15 mL de tampon 0,1M Tris, 0,1M NaCl, 0,1M MgCl₂ contenant 10 µL de CSPD (le substrat de la phosphatase alcaline). Après élimination de la solution de CSPD, les membranes sont exposées (6 à 24 h) à un film autoradiographique qui est ensuite révélé.

X MESURE DES NIVEAUX DE TRANSCRITS PAR PCR EN TEMPS REEL

X-A. DEFINITION DES AMORCES OLIGONUCLEOTIDIQUES

Un couple d'amorces a été défini pour chaque gène d'intérêt ainsi que pour les gènes de contrôle, grâce au logiciel Beacon Designer (Premier Biosoft). Les amorces permettent l'amplification d'un fragment d'une taille comprise entre 80 pb et 150 pb et leur température d'hybridation est d'environ 60°C afin de diminuer les hybridations aspécifiques et de permettre une amplification en deux étapes par cycle.



<u>Figure 16</u>: caractéristiques du vecteur d'expression protéique pET-16b (NOVAGEN). Ce plasmide confère aux bactéries transformées la résistance à l'ampicilline (Ap). La région de clonage et d'expression est détaillée figure 17.



<u>Figure 17</u>: schéma du clonage dans le vecteur pET-16b d'une séquence codant une protéine à produire. La séquence codant la protéine à produire a été obtenu par amplification au moyen d'amorces possédant des sites de restriction Ndel (pour l'amorce directe) et Xhol (pour l'amorce complémentaire). La digestion par ces deux enzymes de restriction du vecteur plasmidique et du fragment d'ADN produit libère des extrémités cohésives permettant leur ligation. La protéine de fusion produite possèdera ainsi un His-Tag en position N-terminale de la séquence protéique d'intérêt.

X-B. TRANSCRIPTION INVERSE

La transcription inverse est réalisée dans un volume de 20 μ L en utilisant le kit "iScript cDNA synthesis" (BioRad). Un μ g d'ARN total est ajouté à 4 μ L d'un mélange réactionnel composé du tampon et d'un mélange d'amorces oligo(dT) et d'hexamères de séquences aléatoires. La réaction est réalisée en présence de 1 μ L de transcriptase inverse RNAse-H⁺. Les ADNc sont synthétisés par l'incubation du mélange 5 min à 25°C suivi de 30 min à 42°C. La transcriptase inverse est ensuite inactivée par une incubation de 5 min à 85°C.

X-C. AMPLIFICATION PAR PCR EN TEMPS REEL

L'amplification par PCR des cibles d'ADNc est réalisée dans un volume de 20 μ L comprenant 375 nM de chacune des amorces, 10 nM de fluorescéine (référence interne) et 10 μ L du mélange Quantitech SYBR Green concentré 2 fois (QIAGEN). Ce mélange contient du SYBR Green I, 5 mM de MgCl₂ et une Taq polymérase recombinante modifiée pour permettre son activation à une température élevée.

L'amplification est réalisée dans un thermocycler BioRad (icycler iQ) équipé d'un module optique. Le programme d'amplification comprend une étape de dénaturation initiale et d'activation de la polymérase de 15 min à 95°C, suivie de 50 cycles d'amplification constitués d'une étape de dénaturation de 30 s à 95°C et d'une étape d'hybridation et d'élongation des amorces de 1 min à 60°C. L'intensité de la fluorescence du SYBR green I est détectée et mesurée au cours de la phase d'élongation.

En fin d'amplification, une courbe de fusion des produits de PCR est établie par des mesures de la fluorescence tous les 0,5°C de 55°C à 95°C, afin de déterminer la température de fusion du produit d'amplification, la spécificité du fragment amplifié et l'absence de dimères d'amorces. L'analyse est réalisée grâce au logiciel iCycleriQ software (BioRad).

XI PRODUCTION DE PROTEINES RECOMBINANTES DANS E. COLI

XI-A. CLONAGE DIRECTIONNEL DANS LE VECTEUR pET-16b

Les protéines recombinantes sont produites au moyen du le vecteur d'expression bactérien pET-16b (NOVAGEN) (Figure 16). Ce vecteur conduit à l'expression d'une protéine de fusion possédant une extension en position N-terminale. Cette extension, composée de 21 acides aminés, comprend 10 histidines (His-Tag) permettant une purification ultérieure par chromatographie d'affinité. Ce système nécessite de cloner une séquence codante dans la même phase de lecture que le His-Tag. Ceci peut être réalisé grâce au site multiple de clonage présent sur le vecteur, composé de 3 sites de restriction (Ndel, Xhol et BamHI).



<u>Figure 18</u>: Production d'une protéine recombinante au moyen du vecteur pET-16b dans le système bactérien BL21 (DE3). En absence de l'inducteur (IPTG), le gène lacI permet la synthèse du répresseur de l'opéron lactose qui se fixe alors au niveau de l'opérateur (lac o), empêchant la transcription et la traduction du gène codant l'ARN polymérase du phage T7. L'inducteur IPTG forme un complexe avec le répresseur l'empêchant de se fixer au niveau de lac o. La transcription et la traduction de l'ARN polymérase du phage T7 permet alors la transcription et la traduction de la protéine recombinante.

Plusieurs constructions correspondant à différents domaines d'un des gènes de type *SERK* clonés ont été réalisées. Des amorces ont été définies à partir de la séquence codante de *CiSERK1*₄₇₄ pour l'amplification des fragments correspondant au domaine kinase intracellulaire, au domaine extracellulaire et à la séquence entière. Ces amorces ont été modifiées en 5' afin de contenir les sites de restriction nécessaires au clonage des fragments dans le vecteur. Ainsi, les amorces situées en 5' du fragment à amplifier possèdent un site de restriction Ndel, les amorces situées en 3' du fragment à amplifier possèdent un site de restriction Xhol (Figure 17). Les séquences des amorces sont données dans la partie résultats.

Les produits d'amplification issus des PCR (conditions décrites § IV) sont purifiés à l'aide du kit de purification de produit PCR (QIAGEN) et élués par 50 μ L de tampon EB. Environ 200 ng de produits d'amplification sont digérés pendant 3 h à 37°C par 20 U de chacune des enzymes de restrictions NdeI et XhoI (Biolabs) dans un volume de 25 μ L contenant un tampon 1x et 1 mM de BSA. Le produit de digestion est précipité une nuit à – 20°C dans 10% d'acétate de Na 3 M, pH 5.2 et 2,5 volumes d'éthanol absolu, puis centrifugé 20 min à 15000 g à 4°C, lavé 3 fois à l'éthanol 80% et resolubilisé dans 20 μ L d'eau.

Les produits de PCR ainsi digérés sont ligués une nuit à 16°C dans 50 ng de vecteur pET-16b, préalablement digérés par les mêmes enzymes de restriction, en présence d'ADN ligase du phage T4 (PROMEGA).

Après transformation de bactéries compétentes JIM109 (PROMEGA) et sélection des colonies recombinantes par PCR en utilisant l'amorce s'hybridant au promoteur de l'ARN polymérase du phage T7 et une amorce spécifique de l'insert (cf § V-B), une préparation plasmidique de vecteur pET-16b recombinant est réalisée. Pour cela, une culture initiée à partir d'une colonie bactérienne sélectionnée est réalisée dans 25 mL de LB en présence d'ampicilline pendant 1 nuit à 37°C. Le plasmide est purifié en utilisant le kit QIAfilter plasmid midi kit (QIAGEN) selon le protocole du fabricant et élué par 50 µL de tampon EB.

XI-B. PRODUCTION ET PURIFICATION DE PROTEINES RECOMBINANTES

XI-B.1. Production de protéines recombinantes en système bactérien

Les protéines de fusion sont produites dans des souches de bactéries *E.Coli* optimisées pour la production de protéine recombinante. Cette souche bactérienne BL21 (DE3) a la particularité de posséder le gène codant l'ARN polymérase du phage T7 sous le contrôle du promoteur de l'opéron lactose inductible par l'IPTG (Isopropyl-β-D-thiogalactopyranoside). En absence d'IPTG, la transcription du gène de l'ARN polymérase est inhibée par le répresseur de l'opéron lactose codé par le gène *lacl*, lui-même introduit

dans le chromosome bactérien. De ce fait, ce système bactérien permet un contrôle de la production de la protéine recombinante (Figure 18)

Cent μ L de ces bactéries compétentes sont transformées de la même façon que décrit au chapitre V-B avec 50 ng de vecteur purifié. Les colonies recombinantes sélectionnées, dont la séquence de l'insert a été vérifiée, sont ensuite mises en culture à 37°C dans 25 mL de LB en présence de 50 μ g d'ampicilline par mL de culture. Cette préculture est utilisée pour initier une culture plus importante pouvant aller jusque 1 L et contenant l'antibiotique. Lorsque la densité optique de la culture atteint une D.O. de 0,6 à 600 nm indiquant que la culture est dans la phase exponentielle de croissance, la production de la protéine recombinante est induite par 100 μ M d'IPTG pendant 4 heures.

XI-B.2. Purification de protéines recombinantes possédant un tag histidine en position N-terminale

Après production de la protéine recombinante d'intérêt, les bactéries sont centrifugées 10 min à 5000 g à 4°C, puis resuspendues dans un tampon PBS (20 mM phosphate, 500 mM NaCl, pH 7.4) à raison de 50 µL de tampon pour 1 mL de culture. Les bactéries sont ensuite lysées par sonication dans la glace, centrifugées, puis la protéine d'intérêt contenue dans le surnageant est soumise à une purification au moyen du kit His-Trap (Amersham).

Les protéines de fusion contenant un tag histidine peuvent être purifiées par chromatographie d'affinité par chélation avec les ions Ni²⁺. Une colonne de séphorase de 1 mL est chargée avec 500 µL de sulfate de nickel 0,1M. Après lavage et équilibration de la colonne par un tampon PBS, 10 mL d'extrait protéique bactérien brut sont injectés dans la colonne. Après lavage, la protéine recombinante retenue sur la colonne est décrochée de la colonne par une concentration croissante de 10 mM à 500 mM d'imidazole, un compétiteur de l'histidine, en solution dans le PBS.

XII OBTENTION ET PURIFICATION D'ANTICORPS POLYCLONAUX

XII-A. IMMUNISATION DE LAPINS

Deux lapins femelles de race néo-zélandaise ont été immunisés par une injection intramusculaire de 500 µg de protéine purifiée correspondant au domaine kinase de CiSERK1₄₇₄ mélangée à l'adjuvent de Freund, suivie par 2 autres injections de 250 µg à un intervalle de 3 semaines. Trois semaines après la dernière injection, environ 50 mL de sang sont prélevés et le sérum immun est préparé et testé. Avant la première injection, 10 mL de sang ont été prélevés afin de préparer un sérum pré-immun.

XII-B. PURIFICATION D'ANTICORPS POLYCLONAUX

XII-B.1. Préparation d'une colonne de chromatographie d'affinité

Les anticorps dirigés contre le domaine kinase de CiSERK1₄₇₄ ont été purifiés par chromatographie d'affinité en fixant la protéine recombinante sur une matrice de sépharose. La liaison de la protéine à la matrice est faite par réaction entre les groupement NH₂ de la protéine et des groupements bromure de Cyanogène (CNBr) activés, couplés à la matrice.

Un gramme de matrice sépharose-CNBr (Amersham) est suspendu et hydraté dans 25 mL de tampon 50 mM NaHCO₃, 100 mM NaCl. Ce mélange est déposé à la surface d'un verre fritté. Les groupements CNBr sont activés par le passage sous vide de 500 mL d'HCl 1M au travers du verre fritté. La matrice ainsi activée est lavée dans 20 mL de tampon 50 mM NaHCO₃, 100 mM NaCl puis mise en contact avec environ 10 mg de protéine recombinante pendant 1 heure à 30°C.

La colonne est préparée dans une seringue de 5 mL dans laquelle le gel de matrice couplé à la protéine est coulé et décanté. Le blocage des sites activés, laissé libres après le couplage, est réalisé par incubation 1 heure du contenu de la colonne dans une solution de Tris 1M pH8. La colonne est ensuite lavée avec le tampon 50 mM NaHCO3, 100 mM NaCl.

XII-B.2. Purification des anticorps

Le sérum est appliqué sur la colonne et l'éluat est récupéré. Afin d'éliminer les protéines fixées de façon aspécifique, la colonne est lavée jusqu'à ce que l'éluat montre une absorbance à 280 nm très faible témoignant de sa faible charge en protéine. Les anticorps sont élués avec 50 mM d'acide acétique par fraction de 1 mL dans des tubes contenant 200 µL 1 M Tris pH 8.0 pour neutraliser l'acidité de l'éluat. Afin de déterminer la présence des anticorps, le suivi de l'élution est réalisé par spectrophotométrie d'absorption à 280 nm.

Les anticorps ainsi purifiés sont dialysés contre un tampon 50 mM NaHCO₃, 100 mM NaCl. Un spectre d'absorption de la solution d'anticorps balayant les longueur d'onde de 240 nm à 600 nm permet de visualiser l'intensité du pic d'anticorps à 280 nm et l'absence de pic d'hémoglobine à 410 nm.

XIII VARIATION DU NIVEAU PROTEIQUE

La variation du niveau d'accumulation d'une protéine peut être observée par la technique de western blot. Cette technique consiste à révéler au moyen de deux anticorps (primaire et secondaire) la présence d'une protéine particulière au sein d'un mélange de protéines séparées initialement par électrophorèse et transférées sur membrane. L'anticorps

primaire obtenu précédemment est spécifique de la protéine recherchée, tandis que l'anticorps secondaire est dirigé contre le premier anticorps et est couplé à une phosphatase alcaline dont l'activité est visualisée soit par réaction colorée soit par chemiluminescence.

XIII-A. ELECTROPHORESE DE PROTEINES EN CONDITIONS DENATURANTES (SDS-PAGE)

L'électrophorèse en gel d'acrylamide en conditions dénaturantes sépare les protéines en fonction de leur masse moléculaire. Quinze µg d'extraits protéiques sont dilués au demi dans une solution de Laemmli 2x (1970) (125 mM Tris-HCl pH6.8, 2% SDS, 200mM DTT, 20% glycérol, 0,0025% bleu de bromophénol). Les protéines sont dénaturées 3 min à 95°C puis déposées sur un gel d'acrylamide à 12,5%.

La migration est réalisée 1 heure à 100 volts dans in tampon de migration 50 mM Tris, 380 mM glycine, 0.1% SDS dans une cuve à électrophorèse Mini-PROTEAN II (BioRad).

XIII-A.1. Coloration des protéines au nitrate d'argent

La coloration au nitrate d'argent des protéines, séparées par électrophorèse SDS-PAGE, est réalisée selon le protocole suivant:

Etapes	Solutions	Durée
Fixation	Fixateur (Ethanol 95%/Acide acétique/ Eau 50/12/ 38)/ 0.1% formaldéhyde (v/v)	1h
Rinçages	Ethanol 50%	3 fois 20 min
Pré traitement	0.2g/I Sodium Thiosulfate	1 min
Rinçages	Eau ultra pure	3 fois 20 s
Imprégnation	1g/l Nitrate d'argent/ 0.075% Formaldéhyde (v/v)	20 min
Rinçages	Eau ultra pure	3 fois 20 s
Révélation	60g/l Carbonate de sodium/ 0.04g/l Sodium thiosulfate/ 0.05% Formaldéhyde (v/v)	Jusqu'à apparition des bandes protéiques
Arrêt de la révélation	Fixateur	

XIII-A.2. Coloration des protéines au bleu de Coomassie

La coloration au bleu de coomassie des protéines, séparées par électrophorèse SDS-PAGE, est réalisée selon le protocole suivant:

Etapes	Solutions	Durée
Fixation	Ethanol 50%, Acide acétique 15 %	30 min
Coloration	Ethanol 50%, Acide acétique 15 %, Bleu de Coomassie R250 (BioRad) 0.22%	30 min
Décoloration	Ethanol 50%, Acide acétique 15 %	Jusqu'à apparition des bandes protéiques

XIII-B. TRANSFERT SUR MEMBRANE

Après séparation par électrophorèse de 15 µg de protéines par essai, leur transfert sur une membrane de PVDF (PolyVinyliDene Fluoride) est réalisé à 4°C pendant 1 heure à 100 V, en système immergé dans un tampon (25 mM Tris; 190 mM glycine, 20% méthanol). Après transfert, les membranes sont rapidement rincées dans un tampon TBS (10 mM Tris pH 7.5, 100 mM NaCl) avant l'hybridation par les anticorps (cf § XIII-C).

XIII-C. HYBRIDATION AVEC LES ANTICORPS

Les membranes sur lesquelles les protéines sont fixées sont dans un premier temps incubées dans un tampon de blocage (TBS + 0,1% Tween20, 5% poudre de lait écrémé) pendant 1 heure à température ambiante. Les membranes sont ensuite incubées dans le tampon de blocage additionné d'anticorps primaire dilué au $1/_{3000}$ pendant 1 heure à température ambiante. Trois lavages de 15 min par la solution de blocage sont ensuite réalisés. Les membranes sont ensuite incubées dans le tampon de blocage contenant l'anticorps secondaire dirigé contre les IgG de lapins dilué au $1/_{3000}$ pendant 1 heure à température ambiante. Avant la révélation, les membranes sont rincées 3 fois 5 min dans le tampon TBS.

XIII-D. REVELATION DES IMMUNOBLOTS

Deux méthodes de révélation des immunoblots ont été utilisées, une révélation colorée et une révélation par chemiluminescence.

XIII-D.1. Révélation colorée

L'activité phosphatase alcaline portée par l'anticorps secondaire fixé à l'anticorps primaire est révélée par incubation des membranes dans 50 mL de tampon (100 mM Tris, 0,5 mM MgCl₂, pH 9.5) contenant 500 μ L de bleu de nitrotetrazolium (30 mg/mL dans du diméthyl formamide à 70%) et 500 μ L de 5-bromo-4chloro-3-indolyl phosphate (15mg/mL dans du diméthyl formamide à 70%). Après plusieurs minutes, une couleur bleue apparaît au niveau des protéines détectées par les anticorps.

XIII-D.2. Révélation par chemiluminescence

L'activité phosphatase alcaline peut aussi être révélée par chemiluminescence au moyen du substrat CDP-Star qui émet un rayonnement capable d'imprimer un film radiosensible. Les membranes sont incubées environ 10 min dans le substrat Immun-Star AP (BioRad) puis un film autoradiographique est exposé contre la membrane pendant une durée variable de 30 s à 5 min avant d'être développé. Cette méthode présente l'avantage de pouvoir ajuster l'intensité du signal obtenu par la durée des expositions du film.

XIV REACTION D'AUTOPHOSPHORYLATION

Un µg de protéine recombinante purifiée (cf § XI-B.2) correspondant au domaine kinase de CiSERK1₄₇₄ est incubé 1 heure à température ambiante dans 40 µL d'une solution contenant 50 mM HEPES pH 7.5, 10 mM MnCl₂, 1 mM DTT, 0,2 mM ATP et 40 µCi [γ^{32} P]-ATP. La réaction est stoppée par l'addition de 40 µL de solution de Laemmli 2x puis incubée 3 min à 95°C. L'échantillon est soumis à une électrophorèse en gel d'acrylamide en conditions dénaturantes (cf § XIII-A) puis transféré sur membrane de PVDF (cf §XIII-B). Un film autoradiographique est exposé deux jours au contact de la membrane puis révélé.

Résultats

Discussion

Partie I: Clonage et caractérisation de gènes codant des récepteurs de type LLR-RLK apparentés à SERK chez la chicorée

<u>PREMIERE PARTIE</u> : CLONAGE ET CARACTERISATION DE GENES CODANT DES RECEPTEURS DE TYPE LRR-RLK APPARENTES A *SERK* CHEZ LA CHICOREE.

L'objectif de cette partie est d'exposer et de discuter les résultats du clonage, chez la chicorée, de deux séquences codant des récepteurs transmembranaires à activité sérine/thréonine kinase apparentées aux gènes de type *SERK* identifiés chez d'autres espèces.

I CLONAGE DE GENES CODANT DES RECEPTEURS DE TYPE SERK CHEZ LA CHICOREE

I-A. MISE EN EVIDENCE DE L'EXISTENCE DE GENES CODANT UNE REGION HOMOLOGUE A UN DOMAINE KINASE DE TYPE *SERK* CHEZ LE CLONE '474' DE CHICOREE

A partir de la séquence du gène *SERK* de la carotte (*DcSERK*) (Schmidt *et al.*, 1997), une recherche d'homologies de séquences a été réalisée dans les banques de données disponibles en juin 1999. L'alignement des dix séquences les plus homologues à *DcSERK* a montré que les régions les plus conservées se situaient au niveau du domaine intracellulaire portant l'activité kinase. Des amorces dégénérées ont donc été définies sur les parties le plus conservées du domaine kinase au moyen du logiciel CODEHOP. Le biais d'utilisation des codons de la chicorée n'étant pas connu, nous avons fait le choix de définir les amorces selon le biais d'utilisation des codons de l'espèce modèle pour les dicotylédones: *Arabidopsis*. Trois amorces ont ainsi pu être définies, SERK-L1, SERK-R1 et SERK-R2.

SERK-L1: 5'-TGAATGATGGAACTCTGGTTGCNGTNAARMG-3'	dégénérescence ⁷ :64
SERK-R1: 5'-TTTCTCAGTAAACTTTCCAGTAGAAACATAYTCNGGNGC-3'	dégénérescence:32
SERK-R2: 5'-AGTACTAGTTTGGTAGATCGGTHNGGYTTYAG-3'	dégénérescence:48

Ces amorces se situent dans les sous domaines V (SERK-L1), VII (SERK-R2) et VIII (SERK-R1) du domaine kinase (Figure 19). Les PCR ont été réalisées à partir d'ADN total de l'hybride interspécifique *C. intybus* x *C. endivia*, clone '474'.

⁷ La dégénérescence correspond au nombre de combinaisons d'amorces qui sont représentées. Par exemple, pour l'amorce SERK-L1, 4 bases sont dégénérées; N correspond à la combinaison des 4 bases (A, C, G, T), R correspond à la combinaison de 2 bases (A, G) et M correspond à la combinaison de 2 bases (A, G) et M correspond à la combinaison de 2 bases (A, G). La dégénérescence de SERK-L1 est donc de 4(N)×4(N)×2(R) ×2(M)=64. Y correspond à la combinaison de 2 bases (A, C, T). H correspond à la combinaison de 3 bases (A, C, T)



<u>Figure 19</u>: Schéma de la structure d'un récepteur kinase de type SERK et positionnement des amorces dégénérées définies au niveau des différents sous domaines conservés du domaine kinase. Chaque boîte contenue dans le domaine kinase correspond à un sous domaine (l à XI). Les flèches indiquent la position des amorces dégénérées définies. a: SERK-L1, b: SERK-R2, c: SERK-R1.

Dans le but d'accumuler suffisamment de produits de PCR, en vue du clonage des fragments obtenus, différentes quantités relatives de l'amorce SERK-L1 par rapport aux amorces SERK-R1 et SERK-R2 ont été testées, du fait des niveaux de dégénérescences différents entre les amorces.

Les séquences génomiques correspondant au domaine kinase de *DcSERK* et à des protéines présentant des homologies chez *Arabidopsis* contiennent un intron d'environ 100 pb dans la région située entre les amorces utilisées. En absence d'intron, les couples d'amorces SERK-L1/SERK-R1 et SERK-L1/SERK-R2 permettent d'amplifier des fragments de 492 pb et 420 pb respectivement. En présence d'un intron chez la chicorée, la taille attendue de ces fragments doit être supérieure à 492 pb et 420 pb.

La figure 20 présente les produits de PCR obtenus après amplification à l'aide de différentes quantités de chacune des amorces. Les combinaisons qui permettent d'obtenir des fragments d'intensités maximales sont 16 pmol d'amorces SERK-L1/ 8 pmol d'amorces SERK-R1 et 4 pmol d'amorces SERK-L1/ 3 pmol d'amorces SERK-R2, pour un volume réactionnel de 15 µL.

L'amplification avec chacun des couples d'amorces a produit deux groupes de fragments de tailles différentes (Figure 20). Dans chacun des cas, les fragments les plus petits ont la taille correspondant à celle attendue en absence d'intron : environ 500 pb avec les amorces SERK-L1/SERK-R1 et environ 400 pb avec les amorces SERK-L1/SERK-R2. On note, de plus, la présence de bandes de taille d'environ 590 pb obtenues avec les amorces SERK-L1/SERK-R1 et d'environ 500 pb obtenues avec les amorces SERK-L1/SERK-R1 et d'environ 500 pb obtenues avec les amorces SERK-L1/SERK-R1 et d'environ 500 pb obtenues avec les amorces SERK-L1/SERK-R1 et d'environ 500 pb obtenues avec les amorces SERK-L1/SERK-R1 et d'environ 500 pb obtenues avec les amorces SERK-L1/SERK-R2. Ces bandes ont une taille qui peut correspondre à une ou plusieurs séquences contenant un intron.

	10 20	30 40	50	60 70	80	90 100	
E1 7ME 100							0.4
SERK Daucus	ADGILVAVKK LKEERTOG-GEL	OFOTEVEMISMAVHRNLLRL OFOTEVEMISMAVHRNLLRL	RGFCMTPTERLLVYP	YMANGSVASCLRE	RPESQPPLDW	PRRORIALGSARGLA PTRKRIALGSARGLS	94
					n		
SKL1 SL1/SR1 588bp	NDGTLVAVKR LKEERSOG-GEL	QFQ <mark>TEVEMISMAVHRNLLRL</mark>	RGFCMTPTERLLVYP	YMANGSVASCLRD	RP <mark>DTQ</mark> PPLDW	P <mark>IRKR</mark> IALG <mark>A</mark> ARGLA	94
SKL2 SL1/SR1 581bp	NDGTLVAVKR LKEERTQG-GEL	QFQ <mark>T</mark> EVEMISMAVHRNLLRL	RGFCMTPTERLLVYP	YMANGSVASCLR <mark>E</mark>	RP <mark>DTR</mark> EPLDW	IRKRIALG <mark>S</mark> ARGLA	94
SKL47 SL1/SR1 565bp	NDGTLVAVKR LTG-HVHG-GEL	OFOKEVGMLSTVVHRNLLRL	HGFCMTPTERFLVYP	YMANGSVASCLRG	RPDTOPALNW	VTR <mark>K</mark> HIALG <mark>S</mark> ARGLA	93
SKL58 SL1/SR2 565bp	NDGTLVAVKR LTG-HVHG-GEL	QFQKEVGMLSTVVHRNLLRL	HGFCMTPTERFLVYP	YMANGSVASCLRG 📌	RPDTQPALNW	/TR <mark>K</mark> HIALG <mark>S</mark> ARGLA	93
SKL68 SL1/SR2 473bp	NDGTLVAVKR LASLSGQGLR	EF <mark>KN</mark> EVT <mark>V</mark> IAKLQHRNLVRL	LGYC <mark>IS</mark> GEE <mark>KI</mark> LIYE	HMANRSLDSI <mark>I</mark> FD 🗙	RKLSASLD	WG <mark>MRFNIIMGI</mark> ARGL	G 92
SKL76 SL1/SR2 472bp	NDGTLVAVKR LASLSGQELQ	EFKNEVTEIAKLQHRNLVRL	LGYCISGEEKILIYE	HVTNRSLDSIIFD	RKLSASLD	WGMRFNIIMIIARGL	S 92
SKL70 SL1/SR2 482Dp	NDGTLVAVKR LASOSPOCNE	FFINET VISCION NIVEL	HCCCTECDOLLLVYE	VMENNSLOCIIED	CKNCMRLD	NDERFDIIMGIARGE.	A 92
SKL57 SL1/SR2 469bp	NDGTLVAVKR LYVNNKFRAA	DFYNEVNMINSVEHKNI.VRI.	LGCSC <mark>SGPESILVYE</mark>	YLPNMSLDLFTFD	EIKGKTLN	WEKRFVIIIGTAEGL	v 92
SKL11 SL1/SR1 496bp	NDGTLVAVKR MECGVITGKGLA	EFTSEIAVLTKVRHRNLVAL	LGYCLDGNEKLLVYE	YMPOGTLSOHLFNW	PEEGLNPLD	TRRLAIALDVARGV	E 96
SKL45 SL1/SR1 499bp	NDGTLVAVKR MESGVMSEKGLD	EF <mark>KS</mark> E <mark>IAVLT</mark> KVRHR <mark>H</mark> LVAL	LGYC <mark>LD</mark> G <mark>NER</mark> LLVYE	YMPQG <mark>TLS</mark> RFLFDW	Q <mark>E</mark> EGLKPLE	WT <mark>K</mark> RLIIALD <mark>V</mark> ARG <mark>V</mark> I	E 96
SKL10 SL1/SR1 499bp	NDGTLVAVKR MESGVMGTKGLN	EFQAEIAVLTKVRHRHLVAL	lgyc <mark>in</mark> gnerllvye	YMPQG <mark>TLS</mark> QHLF <mark>E</mark> W	S <mark>E</mark> H <mark>K</mark> TPPLS	WK <mark>QRVSIALD</mark> VGRGV	E 96
SKL18 SL1/SR1 481bp	NDGTLVAVKR VSS-GSKQGKKE	YVSEVKIISRLRHRNLVQLV	GWCHEQGDFLLVYEY	MPDGSLDSHLFYS-	KSMLSA	SIRYKIAMG-VASAL	L 90
SKLOU SLI/SK2 421DP SKL62 SL1/SR2 393bb	NDGTLVAVKR LERNGERFERS	FRAEWLAVAHLKHRNLVFLK	GWCVHDDQLHLVIDI ICECSEESHBLLVVD	MPNRSLDRLLFRRV	-ENNGTAVAPHLV	FARERIA ICTARCE	E 98
SKL08 SL1/SR2 399bp	NDGTLVAVKR ILK-ESKHGOKE	FASEISTIGRLRHRNLVOLL	GWCRRKGEFFLVYDF	MANGSLONYIFNNP	KAILTWO	ORFKIIND-VSNG-L	L 91
	CEDK 11	n n de en e maiera marca e mar					
	SERK-LI	100 140	150	1.00 170			
	110 120	130 140	150	160 170			
F17M5.190	YLHDHCDPKIIHRDVKAANILLD	EEFEAVVG DFGLAKLMDYK	DTHVTTAVRGTIGHI	APEYLSTGKSSEK	164		
SERK Daucus	YLHDHCDPKIIHRDVKA <mark>A</mark> NILLD	EEFEAVVG DFGLARL <mark>M</mark> D <mark>Y</mark> K	DTHVTTA <mark>V</mark> RGT <mark>L</mark> GYI	APEYLSTGKSSEK	164		
SKLI SLI/SRI 588bp	YLHDHCDPKIIHRDVKAANILLD	EEFEAVVG DFGLAKLMDYK	DTHVTTAVRGTIGHI	APEYVSTGKFTEK	164		
SKUZ SUI/SKI SOIDP	IDADACDEXTIANDVXAANILLD.	BEFERVVG DEGLARDIDIA	DIRVIIAVRGIIGHI	APEIVSIGREIER	104		
SKL47 SL1/SR1 565bp	YLHDHCDPKIIHRYVKA <mark>TS</mark> ILLD	EEFEAVVS DFGLA <mark>KLM</mark> DYN	DTH VTT D <mark>I</mark> RG <mark>SIGHL</mark>	APEYVSTGKFTEK	163		
SKL58 SL1/SR2 565bp	YLHDHCDPKIIHRYVKA <mark>TS</mark> ILLD	EEFEAVVS DFGLATWFDH		SFRK-R1	134		
SKL68 SL1/SR2 473bp	YLHHDSRLRIIHRDFKTCNILLD	ENMNPKIS DFGLARWFDH		o crist i ta	133		
SKL76 SL1/SR2 472Dp	YLHHDSRLQIIHSDFKTCNILLD	GNMNFKIS DEGLARWEDH	SERK-R2		133		
SKL70 SL1/SR2 480bp	FLHDESRIKTVHRDTKATNVLLD	KDLNPKTS DEGLARWEDH			133		
SKL57 SL1/SR2 469bp	YLHENTKTRIIHRDIKAANILLD	LRFRPKIS DFGLARWFDH			133		
SKL11 SL1/SR1 496bp	YLHGL <mark>AQ</mark> QSFIHRD <mark>LKP</mark> SNILLG	DDMRAKVG DFGLVRLAPEG	KGSVETR <mark>I</mark> AGTFGYL	APEYVSTGKFTEK	166		
SKL45 SL1/SR1 499bp	YLHGL <mark>AQ</mark> QSFIHRD <mark>LKPSNILLG</mark>	DDMRAKVA DFGLVRLAPDG	KAS <mark>LVTRLAGTFGYL</mark>	APEYVSTGKFTEK	166		
SKL10 SL1/SR1 499bp	YLHSLAQQSFIHRDLKPSNILLG	DDMRAKVA DFGLVKNAPDG	KYSVKTRLAGTFGYL	APEYVSTGKFTEK	166		
SKLID SLI/SKI 401DD		SSINAKLG DEGLAREVDHD	LGSQTTV <mark>L</mark> AGTMGYL	APETVSTGKFTEK	130		
SKI.62 SI.1/SR2 393bb	YLHEECBNCITHCDTKPENILLD	ODESAKVS DEGLARWEDH	SFRK-R2	SERK-R1	130		
SKL08 SL1/SR2 399bp	YLHEGWEQTVLHRDIKAGNVLLD	SELNGRLG DFGLARWFDH	Contraction .		132		

<u>Figure 21</u>: Alignement des séquences protéiques de DcSERK et de la séquence protéique déduite du gène *F17M5.190* (*AtSERK3*) d'*Arabidopsis* avec les séquences protéiques déduites des clones génomiques de chicorée, isolés par PCR à partir d'amorces dégénérées. Le nom des clones, le couple d'amorces utilisé pour l'amplification ainsi que la taille (bp) des clones sont indiqués à gauche des séquences. Les clones *SKL* (*SerK-Like*) dont la traduction donne la même séquence protéique ont été regroupés. Les acides aminés consensus sont surlignés en gris, ceux similaires au consensus sont surlignés en jaune. La position des amorces est indiquée par des encadrés. La position de l'intron, lorsqu'il est présent au niveau de la séquence génomique, est indiquée par une étoile. SL1 : SERK-L1 ; SR1 : SERK-R1 ; SR2 : SERK-R2.



<u>Figure 20</u>: Produits de PCR obtenus par amplification d'ADN total du génotype '474' en utilisant deux combinaisons d'amorces dégénérées : SERK-L1/SERK-R1 et SERK-L1/SERK-R2. La quantité respective (en pmol pour un volume réactionnel de 15 μ L) de chacune des amorces (SERK-L et SERK-R) utilisées est indiquée au dessus des lignes de dépôt.

Les 4 groupes de produits de PCR obtenus ont été excisés du gel et clonés. Au total, 27 clones issus des différentes ligations ont été séquencés dans les deux sens. Treize clones correspondent à des produits de PCR obtenus avec les amorces SERK-L1/SERK-R1 et 14 ont été obtenus à partir des amorces SERK-L1/SERK-R2. Parmi ces 27 séguences, les alignements nucléotidiques permettent d'identifier 16 groupes différents de séquences parmi lesquels 11 ne sont composés que d'une seule séquence. Les 5 groupes restant sont composés de 2 à 6 séquences identiques. Parmi ces 16 groupes, 8 ont une taille correspondant à une séquence contenant potentiellement un intron et 8 présentent une taille compatible avec l'absence d'intron dans le fragment cloné. Pour les 8 groupes de séquences contenant potentiellement un intron, une recherche de jonction entre exons а été réalisée à l'aide du loaiciel NetGene2 disponible sur internet (http://www.cbs.dtu.dk/services/NetGene2/). L'élimination des séguences intronigues nous a ensuite permis de déduire les séguences protéigues afin de comparer l'ensemble des 16 séquences protéigues avec la séquence protéigue de DcSERK. L'ensemble des 16 groupes montrent des fortes identités de séguences avec des domaines kinase de récepteur de type kinase de plantes.

Parmi ces 16 séquences protéiques déduites, 2 montrent une identité de séquence supérieure à 90 % avec la partie de région kinase correspondante de DcSERK (Figure 21). Ces séquences ont été dénommées *SKL1* (*SerK-Like 1*) et *SKL2*. D'autre part, pour chacun de ces clones, la position des jonctions exon-exon prédite par le logiciel est identique à celle présente sur le gène *DcSERK* (Figure 21). La figure 22 présente l'alignement des séquences nucléotidiques et protéiques déduites des clones *SKL1* et *SKL2*. Les séquences des régions introniques sont différentes entre elles et difficiles à aligner, c'est pourquoi elles n'ont pas été considérées pour estimer l'homologie au niveau nucléotidique. Ces séquences



Figure 22: Alignement des séquences nucléotidiques et protéiques déduites des deux groupes de clones (*SKL1* et *SKL2*) les plus homologues à la région kinase de *DcSERK*. Les séquences des amorces dégénérées utilisées sont figurées au dessus des alignements. Les nucléotides différents ne sont pas surlignés, les acides aminés similaires sont surlignés en gris.

nucléotidiques ont une identité de 83,9 % et leurs séquences protéiques déduites sont identiques à 97,5 % et présentent 98,8 % de similitude.

Les clones *SKL1* et *SKL2* étant fortement homologues à une partie du domaine kinase de *DcSERK*, la suite de nos travaux concernera l'étude de ces deux séquences.

I-B. IDENTIFICATION DES SEQUENCES CODANTES PLEINES LONGUEURS

L'amplification initiale au moyen d'amorces dégénérées à partir d'ADN génomique a permis de mettre en évidence, dans le génome de la chicorée, la présence de gènes dont la traduction menée *in silico* montre une forte homologie de séquence avec un fragment de domaine kinase de *DcSERK*. Compte tenu du fait que cette région est fortement conservée au sein de la famille des récepteurs de type kinase, l'obtention des séquences codantes pleines longueurs était nécessaire pour confirmer que les fragments obtenus sont bien des gènes codant des récepteurs de type kinase apparentés à *DcSERK*. Pour cela, deux approches ont été utilisées. Nous avons dans un premier temps développé une méthode d'amplification simultanée des régions flanquantes d'une séquence connue d'ADNc que nous avons appelée RT-PCR inverse, dont le principe est exposé dans la partie matériels et méthodes (cf § VII-A). Cette technique une fois mise au point, a permis d'obtenir la séquence codante pleine longueur associée à *SKL1* (dénommée *CiSERK1474*), ainsi que la totalité du domaine intracellulaire et transmembranaire associée à *SKL2* (dénommée *CiSERK2474*). L'obtention de la région 5' de *CiSERK2* a été obtenue par RLM-RACE PCR (RNA Ligase Mediated Rapid Amplification of cDNA Ends) au moyen du kit GeneRacer (Invitrogen).

La figure 23 résume les différentes étapes qui ont été nécessaires pour l'obtention des séquences codantes pleines longueurs.

Dans un souci de clarté, seules les séquences codantes pleines longueurs déduites de chacun des ADNc d'intérêt et issues de l'assemblage par chevauchement des différents fragments seront présentées. Les séquences consensuelles résultant du séquençage de chacun des fragments d'intérêt ainsi que leur chevauchement sont présentés en annexe 1 et 2.

I-B.1. RT-PCR inverse

La technique de RT-PCR inverse est une adaptation aux ADNc de la technique de PCR inverse habituellement utilisée pour déterminer les séquences génomiques flanquantes d'une région connue (Ochman *et al.*, 1988). Les matrices utilisées sont dans notre cas des ADNc doubles brins circularisés.



<u>Figure 23</u>: Schéma récapitulatif des méthodes utilisées et de leur apport pour l'obtention des séquences d'ADNc pleines longueurs *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄. Les petites flèches indiquent la position et l'orientation des amorces utilisées pour la RT-PCR inverse et le clonage par 5'RACE, les grandes flèches indiquent les parties de séquences amplifiées à chaque étape. A chaque couleur correspond une étape de clonage et de séquençage. a, b, c: fragments successivement obtenus par PCR inverse à partir d'autoligations identiques; gr: fragments obtenus par 5' RACE. Pour chaque étape, 2 couples d'amorces sont utilisés; la première PCR est réalisée au moyen des amorces figurées par une flèche vide, la seconde PCR (nested-PCR) est réalisée en utilisant les amorces figurées par des flèches pleines (cf séquences ci-dessous).

	(⇒								
	CISERK1474										
а	1ar2: AACGAGTAGACGTTCGGTTG	1ar1: AAAATCCTCTTAGCCGGAGA	1af1: GGCCTTGCTTATTTACACGA	1af2: TACACAACCTCCGCTTGATT							
b	1br2: ACCTGCAAAAAGTAGAGCTG	1br1: ATGATCTTGTGGTTTTCTGC	1bf1: TACAACCAGCACACAATCCAT	1bf2: CTTATAACCTTAGCAACGAGC							
С	1cr2: TTGCAAAATATTATTCGGATCG	for1: GCTATTCCCAGGGTCACAAG	1bf1: TACAACCAGCACACAATCCAT	1bf2: CTTATAACCTTAGCAACGAGC							
CISERK2474											
а	2ar2: GTAACCGTTCGGTAGGTGTC	2ar1: AAGCACGAAGCAACACTACC	2af1: GAGAGACCCGATACACGAGA	2af2: AGAATTCGAAGCAGTTGTGG							
b	2br2: TTGTCGGTTGCCACTTGTAA	2br1: AGCCGCCCTTTGTAGACTTT	2bf1: GTGGGAAGAATGGCAAAGAG	2bf2: CAGATGATTTGGGGGTTTAGGG							
gr	2gr2: CGGCGGGGGACGTCGAAGAAA	2gr1: CACCCCGCCCGAGAATGTGG	GR1: CGACTGGAGCACGAGGACACTGA	GR2: GGACACTGACATGGACTGAAGGAGTA							

<u>Tableau 5</u>: nom et séquences des différentes amorces utilisées pour obtenir les séquences codantes pleine longueur de *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄. a, b, et c représentent les différentes étapes de RT-PCR inverse, gr correspond à l'étape de RACE PCR. Les différentes couleurs ainsi que les flèches reprennent leur signification de la figure 23.
Dans un premier temps, des amorces spécifiques à chaque séquence ont été définies au moyen du logiciel Primer3 disponible sur internet (http://frodo.wi.mit.edu/cgibin/primer3/primer3_www.cgi). Pour chaque séquence, deux couples d'amorces ont été définis. Deux PCR successives ont été réalisées à partir des produits d'autoligation des ADNc. Ces amorces sont 1ar1 et 1af1 pour *CiSERK1*₄₇₄, 2ar1 et 2af1 pour *CiSERK2*₄₇₄ (Tableau 5). Le second couple, 1ar2/1af2 pour *CiSERK1*₄₇₄ et 2ar2/2af2 pour *CiSERK2*₄₇₄ (Tableau 5), permet d'augmenter la spécificité d'amplification selon la méthode de "nested PCR" (Figure 23).

Les produits d'amplifications issus de la seconde PCR ont une taille d'environ 1,2 kpb pour *CiSERK1*₄₇₄ et environ 1 kb pour *CiSERK2*₄₇₄ (Figure 24). La taille de ces fragments est inférieure à celle d'une séquence codant une protéine du type de DcSERK. Néanmoins, ces 2 fragments ont été clonés et environ 8 clones ont été séquencés dans les deux sens pour chaque ADNc étudié. Une séquence consensuelle déduite de l'ensemble des clones correspondant à chacun des fragments a été réalisée. Chaque base de ces séquences à une position donnée.



<u>Figure 24</u>: Produits de PCR inverse obtenus après la seconde PCR réalisée à partir d'ADNc autoligués en utilisant des amorces spécifiques des clones $CiSERK1_{474}$ (1ar1/1af1 puis 1ar2/1af2) et $CiSERK2_{474}$ (2ar1/2af1 puis 2ar2/2af2). (1): marqueur de taille.

Il s'avère que les séquences consensuelles obtenues chevauchent parfaitement, sur 100 bases au moins, les séquences à partir desquelles les amorces ont été définies. Les séquences correspondent aux régions en 3' des séquences connues jusqu'à la région non traduite de chacun des ADNc, ainsi qu'à une partie des régions en 5' (fragments a, Figure 23). Nous avons obtenu 700 pb et 582 pb en 3' ainsi que 243 pb et 151 pb en 5' de la région connue de *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄ respectivement.

Afin d'obtenir les séquences pleines longueurs, 2 nouveaux couples d'amorces, 1br1/1bf1 et 1br2/1bf2 pour *CiSERK1*₄₇₄ et 2br1/2bf1 et 2br2/2bf2 pour *CiSERK2*₄₇₄ (Tableau 5), ont été définis à partir des séquences nouvellement connues à la manière d'une marche sur le chromosome. Afin de favoriser l'amplification des régions 5', les amorces 1bf1, 1bf2, 2bf1 et 2bf2 ont été définies à proximité de la région non traduite en 3'. Les fragments obtenus après le seconde PCR ont été clonés et séquencés de la même façon qu'à la première étape. Cette seconde PCR inverse a permis d'obtenir 580 pb supplémentaires en 5' de l'ADNc *CiSERK1*₄₇₄, et 220 pb pour l'ADNc *CiSERK2*₄₇₄ (fragments b, figure 23).

L'extrémité 5' manquante (environ 260 pb) et incluant la région non traduite de *CiSERK1*₄₇₄ a été obtenue en réalisant une troisième PCR inverse en utilisant les amorces 1cr1/1cf1 et 1cr2/1cf2 (Tableau 5) définies à partir de la séquence obtenue grâce à la seconde PCR inverse (fragments c, Figure 23).

La séquence de l'ADNc d'une taille de 2228 pb pour *CiSERK1*₄₇₄ (Figure 25) a été établie grâce à cette méthode. Malgré plusieurs essais, le domaine extracellulaire de *CiSERK2*₄₇₄ n'a pu être obtenu par ce moyen.

I-B.2. Amplification des extrémités 5' des ADNc (RLM-RACE)

Pour obtenir l'extrémité 5' de *CiSERK2*₄₇₄, une technique alternative a été utilisée. Le principe du kit GeneRacer (Invitrogen) est exposé dans la partie matériels et méthodes (cf § VII-B). L'avantage de cette technique réside en une amplification des extrémités 5' des ARNm pleine longueur par élimination des transcrits tronqués.

De la même manière que pour la RT-PCR inverse, deux PCR successives ont été réalisées en utilisant d'une part des amorces spécifiques de la cible recherchée (2gr1 et 2gr2) et d'autre part des amorces spécifiques de l'oligoribonucléotide ligué en 5' des ARNm (GR1 et GR2) (Figure 23 et Tableau 5). Les produits issus de la seconde PCR ont été clonés et 8 clones séquencés dans chaque sens. Une séquence consensuelle a été déduite et, grâce à une région chevauchante de 200 pb, alignée avec les séquences obtenues précédemment. Cette technique a permis d'obtenir la région manquante constituée d'environ 800 pb en 5' de l'ADNc *CiSERK2474*. Par la combinaison des deux méthodes, la séquence d'ADNc pleine longueur de *CiSERK2474*, longue de 2162 pb, a été établie (Figure 25).

Figure 25 (verso et recto suivants): Séquences d'ADNc pleines longueurs et protéiques déduites de *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄. Les séquences sont issues de l'assemblage par chevauchement des séquences consensus des différents fragments obtenus par RT-PCR inverse et 5' RACE (Annexe 1 et 2). Sur les séquences nucléotidiques, les bases identiques apparaissent surlignées. Sur les séquences protéiques déduites, les acides aminés différents sont grisés.







<u>Figure 27</u>: Schéma de la structure des gènes *DcSERK* et *AtSERK1* et position des introns. Les rectangles représentent les exons tandis que les tirets verticaux noirs figurent la position des introns. Les fragments amplifiés nécessaires à la reconstitution des séquences des deux gènes de chicorée sont illustrés par des lignes de couleurs.

Р	Amorce directe	Amorce inverse	1
	CISER	RK1	
1-1233	1F1-3: CGGTTTGCTTGAGCTGTGT	1R1-3: AGGAAACGGAGTTCTTGAAGG	1,2,3,
478-1313	1F2-4: CCCAATTAACCGATCCGAATA	1R2-4: TGGAGCGATCCGATTGTAGT	2,3,4
1000-1920	1F4-6: TGGGAACAACATCACTGGAA	1R4-6: GAAACAGGTGATTGTGGAGGA	4,5,6
1723-2636	1F6-7: CGCACAATAATTTAAGTGGGATT	1R6-7: CACCAGCGGCAACTCCTC	6,7
1890-2964	1F7-8: ATTCAAGCTCCTCCTCCACAATC	1R7-8: ACCAGAGAACCATCGGCTAA	7,8
2607-3009	1F8: GAGCTATAGCCGGAGGAGTTG	1R8: AACGAGTAGACGTTCGGTTG	8
2904-3915	1F9-10: GCCGATTAGCCGATGGTTCTC	1R9-10: TAAACCATCCCCTTCAAGCA	9,10
3620-4212	1F10: GACGAGGTCATGTTGCTTGA	1R10: TTAGTGAATGCTTTGGTGCT	10
1256	SBF1: CAACAACAACACTTTGACAGGA	Av. Same	
	CiSER	RK2	(
1-190	2F5utr: GGTGGTGGATTAGATTTGTGG	2R5utr: AGAGACTCGGCGATCCATTT	
107-830	2F1-2: TCTGAATGAGTAATCGATGAGAAA	2R1-2: GTGCAACGAGTTGACCTGAC	1,2
610-1266	2F2-4: CAAAGTTGGGATGCGACTCT	2R2-4: GGAATCGTTCCTGTTAACGTGT	2,3,4
1066-1981	2F4-7: AACATTGGGCAAGCTTCAAA	2R4-7 GCTACTCCTCCCGCGATT	4,5,6,7
1577-2494	2F7-8: GCGGTTTCTCCTCAAGCTC	2R7-8: AGCACGAAGCAACACTACCA	7,8
2381-3164	2F9-10: TGATCAGTATGGCGGTTCAC	2F9-10: CGCATCCACTAATGTCTCCA	9,10
2889-3491	2F10: GGGTACGGTGTCATGCTTCTAG	2R10: ATGCACAAAAACCCAATAAA	10

<u>Tableau 7</u>: Séquences des 8 (*CiSERK1*) et 7 (*CiSERK2*) couples d'amorces utilisées dans le but de déterminer les séquences géniques correspondantes. Le nom des amorces est de la même couleur que le fragment qu'elles permettent d'amplifier à la figure 27. La colonne **P** indique la position des amorces sur les figures 28 (*CiSERK1*) et 29 (*CiSERK2*). La colonne **I** indique les introns potentiellement contenus dans le fragment amplifié sur la base de la conservation de la position des introns chez la chicorée.

SBF1 correspond à l'amorce directe utilisée avec l'amorce 1R4-6 pour générer la sonde utilisée pour l'expérience de Southern blot (cf § IIC)

in the

I-B.3. Amplification des ADNc pleines longueurs

Afin de vérifier que les séquences d'ADNc déduites, et issues de l'assemblage des séquences consensuelles des différents fragments, ne sont pas le produit d'un assemblage chimérique, l'amplification des séquences codantes pleines longueurs a été réalisée en utilisant des amorces situées au niveau des régions 5' et 3' non traduites (Tableau 6)

 CiSERK1474
 CiSERK1F: TCTGGTTGAGTAATCGATGATG
 CiSERK1R: TTAGTGAATGCTTTGGTGCT

 CiSERK2474
 CiSERK2F: TTTAGCTTTACCGACCCGAACCC
 CiSERK2R: CCCAAACCCTAAACCCCAAATCATC

<u>Tableau 6</u>: Séquences des amorces utilisées pour l'amplification des ADNc pleines longueurs de CiSERK1474 et CiSERK2474.

Ces amorces doivent permettre d'amplifier 2065 pb de la séquence d'ADNc de *CiSERK1*₄₇₄ et 1905 pb de la séquence de *CiSERK2*₄₇₄. Les produits des PCR sont présentés figure 26 et sont de la taille attendue. Le fragment correspondant à *CiSERK2*₄₇₄ n'a pas été séquencé. Le séquençage d'un fragment correspondant à la séquence codante de *CiSERK1*₄₇₄ a permis de vérifier la séquence de l'ADNc déduit.



Ainsi, les deux méthodes utilisées de façon conjuguée nous ont amené à identifier deux ADNc pleines longueurs, *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄ (Figure 25). *CiSERK1*₄₇₄ a une longueur de 2228 pb et 1839 d'entre elles permettent la traduction d'une protéine de 613 acides aminés. Cette séquence codante est flanquée d'une région non traduite en 5' de 150 pb et en 3' de 239 pb. L'ADNc de *CiSERK2*₄₇₄ a une longueur de 2177 pb dont 1851 permettent la traduction d'une protéine de 617 acides aminés. Cette séquence codante est flanquée d'une information d'une protéine de 617 acides aminés. Cette séquence codante est flanquée d'une région non traduite en 5' de 150 pb et en 3' de 239 pb. L'ADNc de *CiSERK2*₄₇₄ a une longueur de 2177 pb dont 1851 permettent la traduction d'une protéine de 617 acides aminés. Cette séquence codante est flanquée d'une région non traduite en 5' de 207 pb et en 3' de 119 pb (Figure 25).

Les ADNc *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄ montrent entre eux une identité de séquence nucléotidique de 70% qui atteint 78% d'identité lorsque seul le cadre ouvert de lecture est considéré. Ceci laisse penser qu'une telle différence de séquence nucléotidique témoigne de la présence de deux gènes distincts.

Les protéines déduites des séquences codantes de CiSERK1474 et CiSERK2474 présentent une identité de séquence respectivement de 73 % et 75% avec DcSERK, et la

cggtttgcttgagctgtgttcaaagctagcgtttgggttttaggtttttgaatttcttgatctggttgagtaatcgatgatgaatttaaagcttttagcttgactcgagtttgtgggaaATGGATCGAACAATCTCTGTA	140
ATCG <mark>TT</mark> ATTTGGCTTATTTTCGGGTTCAATTATGTATCAAAGGTTTATGGAAATGCTGAAGgtatgtgttttttgtgattcgtgattgttaatttaatc <mark>a</mark> tgactttgaatgtgttgtcgcttaatcgaatttagtaact IV/AIWLIFGFNYVSKVYGNAE	280
tatcgcaattgagtteettgagteatt <mark>t</mark> teeeaattatagtatettegattetaattatagataetgattgegaagaaeatatatagaaattteaeeaattaaaettetttaattaa	420
ggacgctaatcaaatttcatatgtgtaacatagGTGACGCATTGAATGCCCTAAAGACCCCAATTAACCGATCGAATAATATTTTGCAAAGTTGGGATCCGACCCTGGTTAATCCATGTACATGGGTTCATATTACTTGT G D A L N A L K T Q L T D P N N I L Q S W D P T L V N P C T W V/F H I T C GACGGAGCGAATAGCGTTCTAGGCCTTGTAGGCTTGCAGGTTGAGGTTGAGGTGGGGGG	560
D P/Q G N S V S R L D L G N A G L S G Q L	040
V P E L G Q L S K L Q Y L gtatgaaggtataattttgtagaattttcagtatacaacactacaaaataaagtgaaagtacaacaccataacagaaacaaaaaatgaaagtaggatgccattatac <mark>g</mark> caaataattaaaaactac <mark>a</mark> ttgttattt	980
ctggaacagGGAGCTTTATGGGAACAACATCACTGGAAAAATTCCTGAACAGCTTGGAAACTTAACACAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGAAACATTGGGGAAACTTAACACAATTAGAAGGTGACATTCCAGACACATTGGGGAAACTTCAACAAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGAAACTTCAACAAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGGAAACTTAACACAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGGAAACTTAACACAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGGAAACTTCAACAAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGGAAACTTCAACACAATTAATGAGTTTGGATCTTTACTTAACAAATTAGAAGGTGACATTCCAGACACATTGGGGGAAACTTAACACAATTAATGAGTTTGGATCTTTACTTAAACAAATTAGAAGGTGACATTCCAGACACATTGGGGGAACCTTCAAC	1120
E L Y G N N I T G K I P E Q L G N L T Q L M S L D L Y L N K L E G D I P D T L G N L Q AACTCCGTTTCCTgtaatattcaagattcttaattatgaatttaagataatttacttttaaatcataaattattagttttt O L R F L	1260
AACAACACTTTGACAGGAACTATTCCAAATTCATTAACTACAAT <mark>C</mark> GGATCGCTCCAAGTTCTgtgagtacccgcccattttttgcttatt <mark>ca</mark> gaatttttt <mark>ttgt</mark> ttaaactttaaccaccagaa <mark>t</mark> atattt <mark>t</mark> tttcttt N N T L T G T I P N S L T T I G S L O V L	: 1400
gagttttacaaaaaccattatag <mark>g</mark> ttgaaatttatac <mark>aa</mark> tttaaccacccgaacaatata <mark>a</mark> tttaagatttaaccatacacactttgattttttaataaagaacaattttaacca <mark>a</mark> aaaaatattcttccgatgatttt tgcaattcattaccgggaaacccttacaatattatttaacagaaaaacaaac	: 1540 1680
aatagctaaaattgtatattggtgtttattgcagTGATCTATCGCACAATAATTTAAGTGGGATTGTTCCAATTAATGGGTCATTTTCACATCGATCAGgtttgttttatttaatagtttttgaaattatgt D L S H N N L S G I V P I N G S F S L F T P I S	: 1820
g <mark>t</mark> ggtttetttatgtttgttgaetgeaagettttgttettt <mark>e</mark> agTTAT <mark>C</mark> GTAACCTCAACTTGGTAATTCAAGETCCT <mark>C</mark> CTCCACAATC <mark>A</mark> CCTGTTGCACCTGATTCTCAATCTCCTTCAGgtatttttaatteaec Y H/DR/G N L N L V I Q A P P/A P Q S P V A P D S Q S P S	: 1960
attaaaactttatgcatttatttgtatttaatttttgtagtggttcttactcttgaatctcccgtttctatgtttatatatgaacttggtagtggtcatgtttttttt	2100 2240 2380 2520
tagttttgctcgttagagattttgaacgatgtagcgtcttctaattgtatatgaaactttgatttcagTCAGCAACAGTGCCACCGGAGCTATCGCCGGGAGGAGTTGCCGCCGGTGCAGCTCTACTTTTTGCAGGTCGCC	2660
CGATTGCAATTGCTTGGTATCGGCGCGCAGAAAACCACAAGATCATTTCTTTGATGTACCCCGgttagtccatttaccaatatactcgcctaattcaaaa <mark>tt</mark> tttttttttttttttttttttttt	: 2800
atattttcagcCGAAGAGGATCCAGAAGTCCACCTAGGACAACTAAAGAGAGTTCTCTTTACGCGAATTACAAGTCGCAACCGATACCTTTAGTAACAATAACATTCTTGGCCGAGGTGGATTTGGTAAAGAGTTTACAAAGC	2940
CCGATTAGCCGATGGGTCTCTGGTGGCGGTTAAAAGACTCAAAGAGGAACGTAGTCAAGGCGGCGGAATTGCAGTTTCAAACGGAGGTTGAGATGATCAGTATGGCGGTGCACCGGAACCTTCTCCGGCTAAGAGGATTTT R L A D G S L V A V K R L K E E R S Q G G E L Q F Q T E V E M I S M A V H R N L L R L R G F	3080
GCATGACACCAACCGAACCGCTCTACTCGTTTATCCCTACATGGCCAATGGCCAGCGTCGCATCATGTTTACGAGgtaattaaatcctccgataaaaattagggtttactttatcgggctagtgggctgtcgtttaaataata	3220
gcccaataattcatacaaataattgtagATCGACCTGATACACAACCTCCGCTTGATTGGCCGATAAGGAAACGAATCGCTTTGGGAGCCGCACGTGGCCTTGCTTATTTACACGATCATTGTGACCCGAAAATTATACA	3360
ERPDTQPPLDWPIRKRIALGAARGEARGEAGEAGEAGEAGEAGEAGEAGEAGEAGEAGEAGEAGEAG	3500
CCGAGTATCTTTCCACCGGGAAATCTTCGGAAAAAACCGATGTTTTTGGGTACGGTGTCATGCTTCTAGAACTCATCACCGGTCAACGCGCCTTCGATCTTGCCCGCCTTGCCAATGATGACGAGGTCATGTTGCTTGAT PEYLSTGKSSEKTDVFGYGVGVMLLEL	3640
TGGgtgcgcatcctaccaatttttgtctcattgacatatatgtcattgttactgtcactatgttactgctactgatcggttttgaaattgtttagGTGAAAGGACTATTGACGGATAAGATGCTTGAAAAGTTGGTTG	3780
GAGGATTTAGATGGTGCTTATGTGGACGATGAGGTGGAACAGTTGATACAGGTGGCGCCTATTGTGTACACAAGGGACGGCAATGGAGCGGCCGAAGATGTCGGAAGTTGTTAGAATGCTTGAAGGGGAAGGTGTTAGCAGG	3920
GAGGTGGGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGG	4060
RWEEWQKEEIFRQDYNQAHNHNIDWIIPDSTYNLSNEQLSGPR* catgtgtggtgttgtcgttgtgtatatgtttttattttgtttatacgaggtataggggtatttatgtaaaacgtgtgtaggggggttatgttatttat	4200

similitude atteint 80% et 81% respectivement. Les différentes caractéristiques ainsi que les homologies de séquence des protéines déduites CiSERK1 et 2 seront analysées par la suite (§ II-D).

I-C. CLONAGE DES SEQUENCES GENIQUES CORRESPONDANT AUX ADNC CISERK1 ET CISERK2

Pour obtenir les séquences géniques correspondant aux deux séquences d'ADNc mises en évidence chez le génotype '474', la suite de notre travail a été menée sur les génotypes de chicorée K59 et K28. En effet, K28 et K59 sont les parents d'une descendance étudiée au laboratoire dans le cadre de l'étude du déterminisme génétique de l'embryogenèse somatique de la chicorée.

D'autre part, l'étude de l'expression génique au cours des phases précoces de l'embryogenèse somatique a été réalisée sur le génotype embryogène K59 et sur le génotype non embryogène C15, issu de l'autofécondation de K59. Il était donc important de disposer des séquences de chacun des gènes et des différentes formes alléliques potentielles afin que les amorces définies pour réaliser les PCR en temps réel ne présentent aucun mésappariement entre les deux génotypes.

Une analyse de la structure génique des différents gènes apparentés à *DcSERK* chez *Arabidopsis* a montré que ces gènes sont composés de 11 exons et 10 introns dont la position des jonctions est conservée. Afin de définir les amorces nécessaires au clonage des gènes correspondant aux ADNc *CiSERK*₄₇₄, nous avons émis l'hypothèse que la position des introns était conservée chez les gènes de type *SERK* de la chicorée. Les amorces ont été définies sur chacun des ADNc *CiSERK*₄₇₄ afin qu'elles permettent l'amplification de séquences géniques contenant au moins un intron. Sept ou 8 couples d'amorces ont ainsi été définis à partir des séquences d'ADNc, de telle sorte que les fragments obtenus par amplification soient chevauchants sur au moins 100 pb (Figure 27 et Tableau 7).

Au moins 8 clones de chaque fragment obtenu ont été séquencés dans les deux sens puis alignés. Lorsque des substitutions réitérées de bases s'observaient entre les différents clones à une fréquence supérieure à 50%, elles ont été considérées comme un polymorphisme de séquence (SNP: Single nucleotid polymorphism). Le chevauchement des séquences consensus de chacun des fragments a permis de reconstituer la séquence génique. Les positions des introns ont été déduites par alignement des séquences géniques et des séquences d'ADNc déterminées préalablement à partir du génotype hybride '474'. Les

Figure 28 (ci-contre): Séquence génique de *CiSERK1*, reconstituée à partit du chevauchement des différents fragments obtenus par PCR. Les séquences non codantes (utr et introns) sont en lettres minuscules noires, la séquence codante est figurée en lettres capitales bleues sous laquelle apparaît la séquence protéique déduite. La séquence présentée correspond à la forme allélique commune aux génotypes K59 et K28. Les bases surlignées indiquent la présence de polymorphisme entre les 2 allèles clonés chez le génotype K59, et le cas 70 échéant la modification de l'acide aminé codé.

ggtggtggattagatttgtggacgatatacatagattctgttcgtttgagctggctctaaacctagggtttggtgtttggggttctggggtacttcgaatttcttgatctgaatgagtaatcgatgagaaatctaggtttttagctttacc 150M D R R V S L I S A S V L V W F I L V F N N S S K V H G N A E tctaataatagaccctcgatttctagactgattttaccgttacatctccaaattgatgatatgtatcttctttttacaacgaaaattcaccaaattgatgttttcagGCGATGCATTGAATGCGTTAAAGACTCAATTAGGTGATCCGAAT 600 G D A L N A L K T Q L G D P N AATGTTCTTCAAAGTTGGGATGCGACTCTTGTAAACCCCGTGTACATGGTTTCATGTTACGTGCCAATAACGAGAATAGTGTTACAAGAGTgtaagcttattttcttcatttaagataagattaacatcgctcaattagatattatcattat 750 N V L O S W D A T L V N P C T W F H V T C N N E N S V T R V ttaaqaaaatcgttcctttttggtgtttattatacqcagTGATCTTGGAAAACGCGAATTTGTCAGGTCAACTCGGTGAACTCGGTGAACTCACTAATTTACAGTATTTgtaagttgtttttagtatgcgtgtttgttataaacgtat 900 D L G N A N L S G Q L V A Q L G E L T N L Q Y L aataagaattcgatggtttttaatttatttttaatggatgtgttatttaacagGGAACTTTACAGTAATAACATTACTGGAAGAATTCCTGATGCACTTGGAAACTTAACAAATTTAGTGAGTTTGGATCTTTATCTCAACAAATTAGA 1050 E L Y S N N I T G R I P D A L G N L T N L V S L D L Y L N K L D GGIPDTLGKLQKLRFL tgattttggtgtcattttttgcaagaacagCCGTCTCAATAACAACACGGTTAACGGGAACGATTCCCGATTTCATTGACTACTACTACTACTACTACAAGTTCTgtaagtaccaaaaatgttttttttgtttgttatttttttaaaaatttta 1350R L N N N T L T G T I P I S L T T I T S L O V L tgattgttgaagattaattattatgtttttatgtatctcagTGATCTGTCAAACAACAATTTGAGAGGAGACGTCCCGGTCAACGGTTCCTTTTCGCTTTTTACACCTATAAGgtatgttaaaagtcattttcqccctccatttttttaa 1500D L S N N N L R G D V P V N G S F S L F T P I S aaaatttgaaaccttttttttgacaatgaacttttcttcttttagTTTTGCCAATAATCCCCAGTTGAAAGCCCCTGCGGGTTTCTCCTCAAGCTCCGGGCCCCACCAAATTCCCCCCTCTTCTTCTGgtatttttaacttttactgcaattt 1650 F A N N P Q L K A P A V S P Q A P A P P N S P S S S VGN S A T G A I A G G V A A G A A L L F A G P A I A L A W W R R R K P Q D H F F D V P A E E D P E V H L G O L K R F S L R E L O V A T D N F N N R H ATTCTTGGGCGAGGTGGATTTGGTAAGGTTTACAAAGGGCGGTTGGCGGACGGCCCTTTGGTGGCCGTGAAGCGGTTGAAAGAAGAAGAAGCAGTTGCAGTTGCAGTTTCAGAACGGAGGTGGAGATGATTAGTATGGCGGTTCAC 2400 I L G R G G F G K V Y K G R L A D G S L V A V K R L K E E R T Q G G E L Q F Q T E V E M I S M A V H ${\tt CGGAATTTACTCCGGTTGAGAGGGtTTTTGCATGACACCTACCGAACGGTTGCTTGTTTATCCATACATGGCTAATGGCAGTGTTGCTTCGTGCTTAAGAGgtatatttttgttgcattgtgcaagaaatagcaacgtacttttgttaatt 2550$ R N L L R L R G F C M T P T E R L L V Y P Y M A N G S V A S C L R tqtcccttaatttaattactaacaaaattttaattctaqAGAGACCCGATACACAAGAACCGCTTGATTGGCCAATACGTAAGCGAATTGCACTGGGGTCCGCGAGAGGACTTGCATATTTACATGACCATTGTGACCCAAAAATCATTC 2700 E R P D T Q E P L D W P I R K R I A L G S A R G L A Y L H D H C D P K I I ACCGTGATGTAAAAGCTGCAAATATACTACTAGTAGAAGAATTCGAAGCAGTTGTGGGGAGATTTTGGCTAAACTTATGGCTAAACTTATGGATAAAAGATACACATGTCCACGACGAGCTGTCCGTGGCAAATTGGCCCACGAGTATT 2850 H R D V K A A N I L L D E E F E A V V G D F G L A K L M D Y K D T H V T T A V R G T I G H I A P E Y TATCAACCGGGAAATCATCGGAAAAAACCGATGTTTTTGGGTACGGTGTCATGCTTCTAGAACTTATCACGGGACAACGCGCTTTCGAAGACTTGCCAAGATGATGATGATGATGTTGCTCGATTGCGQTAACtttttgaacct 3000 L S T G K S S E K T D V F G Y G V M L L E L I T G Q R A F D L A R L A N D D D V M L L D W tgaattcattgtacacacttcaataacgtaactttaaatacattttaatctgtaagaaatgccacgtcatcagtttgatcgtatgaagtgagaattataatggttgtaatttatagGTGAAAGGGCTTTTGAGGGAAAAGAAATTGGAGA 3150VKGLLREKKLE CATTAGTGGATGCGGATTTGAAAGGTAATTATATTGATGATGAAGATGAAGATGAACAACTTATACAAGTGGCACTTTTATGCACACAAGGGACGCCATTAGAAAGACCAAAAAATGTCGGAAGGTGATGTGGGAAGGTGATGGGTAAGCGG 3300 T L V D A D L K G N Y I D D E V E Q L I Q V A L L C T Q G T P L E R P K M S E V V R M L E G D G L A AAAGGTGGGAAGAATGGCAAAGAGAAGAAAATGTTTAGACAAGAATTCCAATACCGACACATAATCCAAAATACTGATTATTGCAGACTCAACGTATAACCTTCGTCCTGATGAATTATCTGGTCCCAGATGALttggggtttagggtt ERWEEWQREEMFRQEFNTTHNPNTDWIIADSTYNLRPDELSGPR* tggggttttttttaatggttatttattgggtttttgtgcat 3491

Figure 29: Séquence génique de CiSERK2 reconstituée à partir du chevauchement des différents fragments obtenus par PCR. Les séquences non codantes (utr et introns) sont en lettres minuscules noires, la séquence codante est figurée en lettre capitale bleue sous laquelle apparaît la séquence protéique déduite. La séquence présentée correspond à la forme allélique unique commune aux génotypes K59 et K28



Figure 30: Structure schématique des séquences géniques CiSERK1 et CiSERK2. Les rectangles figurent les exons. Les traits noirs correspondent aux introns. Les lignes de couleurs indiquent les fragments amplifiés et définis à la figure 27

	CiSERK1	CiSERK2		CiSERK1	CiSERK2
Exon1	82	94	Exon7	80	80
Intron1	252	290	Intron7	644	316
Exon2	133	133	Exon8	132	132
Intron2	79	100	Intron8	90	85
Exon3	72	72	Exon9	342	342
Intron3	252	93	Intron9	96	89
Exon4	144	144	Exon10	395	395
Intron4	117	132	Intron10	92	132
Exon5	72	72	Exon11	315	315
Intron5	392	89			
Exon6	72	72			
Intron6	78	82			

<u>Tableau 8</u>: Taille (en pb) des introns et exons constituant la structure génique des séquences *CiSERK1* et *CiSERK2*. Les numéros des exons correspondent à ceux de la figure 30.

jonctions entre exons ont été confirmées par une simulation *in silico* de l'épissage au moyen du logiciel NetGene2.

Les figures 28 et 29 présentent les séquences géniques et protéiques déduites de $CiSERK1_{K59}$ et $CiSERK2_{K59}$ issues des fragments clonés chez le génotype K59. Les deux gènes sont composés de 11 exons séparés par 10 introns. Excepté le premier exon qui est plus long de 12 pb pour *CiSERK2*, la taille des exons et les motifs qu'ils codent sont conservés entre les deux séquences. La taille des introns est beaucoup plus variable entre les deux séquences (Figure 30 et Tableau 8). En particulier, les introns 3, 5 et 7 de la séquence *CiSERK1* sont au minimum deux fois plus long que ceux de la séquence *CiSERK2*.

Le séquençage des fragments amplifiés au moyen des amorces définies à partir de la séquence codante de *CiSERK2* n'a révélé qu'un seul type de séquence quel que soit le génotype. Chez le génotype K59, deux types de séquences issues de l'amplification de *CiSERK1* ont été obtenues, dont un correspond au seul type mis en évidence chez le génotype K28. Le polymorphisme de séquence observé entre les 2 types de séquences géniques *CiSERK1* est de 2,16%. Les introns présentent 2,7% de SNP tandis que la région codante montre un polymorphisme de 0,76%, car d'avantage soumise à une pression de sélection de fonctionnalité. Seulement un tiers des SNP dans la séquence codante induisent une substitution d'acides aminés non similaire à l'acide aminé substitué. Aucune insertion ou délétion de nucléotides n'a été observée.

II ANALYSE DES SEQUENCES NUCLEOTIDIQUES ET PROTEIQUES DEDUITES.

II-A. COMPARAISON DES SEQUENCES GENIQUES *CISERK1_{K59 / K28}* et *CISERK2_{K59 / K28}* AVEC LES ADNC OBTENUS CHEZ '474'

Un alignement a été réalisé à partir des séquences codantes de *CiSERK1*₄₇₄, *CiSERK2*₄₇₄, des différentes séquences codantes déduites des séquences géniques de *CiSERK1*_{K59} (dont l'une est présente chez K28), de *CiSERK2*_{K59} et de la séquence codante de *DcSERK*. L'arbre phylogénétique dérivé de cet alignement (Figure 31) révèle trois groupes de séquences distincts. L'un est composé de la séquence de la carotte, le second est composé des séquences *CiSERK1* et le troisième, des séquences *CiSERK2*. Le tableau 9 regroupe les pourcentages d'identité de séquences entre les différentes formes de *CiSERK1* et *CiSERK2*. Tandis que les séquences d'un même groupe présentent plus de 97 % d'identité, celle entre les deux groupes (*CiSERK1* et *CiSERK2*) n'atteint que 73% environ.



<u>Figure 31</u>: Arbre phylogénétique issu de l'alignement des séquences nucléotidiques codantes de *CiSERK1* et *CiSERK2* mises en évidence chez le clone '474', des formes alléliques (a et b) des séquences codantes déduites des séquences géniques de *CiSERK1* et *CiSERK2* chez K59 et de la séquence nucléotidique codant *DcSERK*. La valeur du bootstrap indique la robustesse du nœud.

9. <u>jan - 19</u>	CiSERK1474	CiSERK1 _{K59} a	CiSERK1 _{K59} b	CiSERK2474
CiSERK1 _{K59} a	98.35 %			<u> </u>
<i>CiSERK1_{к₅9}</i> b	98.24 %	99.23 %		
CiSERK2474	73.13 %	73.74 %	73.02 %	
CiSERK2 _{K59}	73.37%	73.96 %	73.26 %	97.41 %

<u>Tableau</u> 9: Identité des séquences nucléotidiques codant les différents allèles de CiSERK1 et CiSERK2 chez les chicorées '474', K59 et K28.

Ces résultats suggèrent que *CiSERK1* et *CiSERK2* sont deux gènes distincts. Les 3 séquences *CiSERK1* (*CiSERK1*₄₇₄, *CiSERK1*_{K59}a et *CiSERK1*_{K59}b) et les 2 séquences *CiSERK2* (*CiSERK2*₄₇₄, *CiSERK2*_{K59}) sont vraisemblablement des formes alléliques des gènes *CiSERK1* et *CiSERK2* respectivement.

II-B. CARTOGRAPHIE DES GENES CISERK1 ET CISERK2

Le polymorphisme de séquences de *CiSERK1* détecté entre les génotypes K59 et K28 a été utilisé afin de localiser ce gène sur la carte génétique de la chicorée construite, dans le cadre du GIS CARTOCHIC⁸, à partir de la descendance issue du croisement K28 x K59. Le génotype des individus de la descendance a été déterminé par SSCP (Single Stand Conformation Polymorphism). Bien que la séquence *CiSERK2* ne montre pas de polymorphisme entre les génotypes K59 et K28, un polymorphisme a été détecté au sein d'une autre descendance, elle aussi utilisée pour la construction la carte génétique de référence de la chicorée. En utilisant des marqueurs ancres entre les deux cartes, les gènes *CiSERK1* et *CiSERK2* ont été cartographiés sur des groupes de liaison différents. Ces résultats apportent la confirmation que les gènes *CiSERK1* et *CiSERK2* sont portés par deux locus distincts.

II-C. GENES DE TYPE SERK CHEZ LA CHICOREE

Une hybridation d'ADN soumis à une digestion par les enzymes de restriction MsII et RsaI a été réalisée afin de déterminer l'existence d'autres gènes de types *SERK* chez la chicorée. La sonde utilisée correspond à un fragment amplifié au moyen des amorces sB-F1 et 1R4-6 (tableau 7) et contenant les introns 4,5 et 6. Cette sonde inclut les exons 5, 6 et une partie de l'exon7 ainsi que les introns 5 et 6 de la séquence de *CiSERK1_{K59}*. La région

⁸ GIS CARTOCHIC: Groupement d'Intérêt Scientifiques réunissant des intérêts privés (Ets Florimond-Desprez, Hoquet Graines, Leroux-Finaler) et académiques (Physiologie de la Différenciation Végétale, USTL) qui vise à établir une carte génétique de référence de la chicorée dans l'objectif d'une utilisation dans le cadre de l'amélioration variétale, notamment. Les données concernant la cartographie des gènes *CiSERK* ont été aimablement fournies par M. Thierry CADALEN, (Ets Florimond-Desprez) coordinateur du projet.



<u>Figure 32</u>: Southern blot d'ADN K59 (A) et K28 (B) digérés par les enzymes de restriction Msll (I) et Rsal (II). La sonde utilisée correspond à un fragment d'ADN génomique de *CiSERK1_{K59}* incluant les exons 5, 6 et 7 ainsi que les introns 5 et 6.

a à p: signaux d'hybridation. M: marqueur de taille.

- Signaux correspondant à CiSERK1 ou CiSERK2
- Signaux correspondant à d'autres gènes apparentés

correspondante de *CiSERK2* est plus courte du fait de la taille différente des introns et s'étend de la position 1235 à 1600 sur la figure 29. La sonde a été obtenue à partir du fragment cloné chez le génotype K59.

Les signaux d'hybridation obtenus sont présentés figure 32. Les profils d'hybridation sont identiques entre les deux génotypes.

Une simulation de la digestion des deux séquences *CiSERK* par l'enzyme de restrictions MsII conduit à la production pour *CiSERK1* d'un seul fragment de 2819 pb (sites MsII en positions 540 et 3359 sur la figure 28) capable d'hybrider la sonde. Pour *CiSERK2*, un seul fragment de 1769 pb (sites MsII en positions 235 et 2004 sur la figure 29) contient la région correspondant à la sonde. Deux fragments dont les tailles correspondent sont détectés sur le profil d'hybridation (bandes a et b). Le signal d'intensité le plus fort (a) correspond à *CiSERK1*. Huit autres fragments de tailles variables (c à j) sont aussi capables d'hybrider la sonde.

La simulation de la digestion de *CiSERK1* par l'enzyme Rsal conduit à la formation de 2 fragments capables d'hybrider la sonde; un fragment de 1185 pb et un fragment de 419 pb (sites Rsal en positions 905, 1324 et 2509 sur la figure 28). Le fragment de 1185 pb (k) est visible sur l'autoradiographie tandis que le fragment de 419 pb n'est pas détecté.

La digestion de *CiSERK2* par Rsal conduit à la formation de 2 fragments de 163 pb et 388 pb (sites Rsal en positions 1125, 1288 et 2676 sur la figure 29) contenant les régions correspondant à la sonde. Aucune bande de taille inférieure à 600 pb n'est visible sur le profil de Southern blot. Par ailleurs au moins 5 fragments supplémentaires hybrident aussi avec la sonde (bandes l à p).

Ainsi, bien que la sonde possède une forte spécificité due à la présence de 2 introns, au moins 5 à 8 signaux supplémentaires à ceux attendus ont été observés, selon l'enzyme de restriction utilisée. Il existe donc d'autres gènes de type *SERK* en plus des deux identifiés chez la chicorée, présentant une forte homologie de séquence avec le domaine extracellulaire de *CiSERK1*.

II-D. DESCRIPTION DES PROTEINES CISERK

La traduction *in silico* des séquences codantes permet de déduire que *CiSERK1* et *CiSERK2* codent des protéines de 613 et 617 acides aminés respectivement qui ont des masses moléculaires calculées de 67,7 kDa et 68,2 kDa respectivement (Figure 33). Les séquences protéiques sont identiques à 83%; leur similitude est de 90%.

Dans un premier temps, les différents domaines des protéines déduites CiSERK1_{K59} et CiSERK2_{K59} ont été définis au moyen du logiciel Smart disponible sur internet (http://smart.embl-heidelberg.de/). Les deux protéines ont les mêmes caractéristiques. Elles sont composées d'un domaine présentant les caractéristiques d'une activité kinase, d'un

	Site de cliva	ge du peptide s	ignal				
	10	20	30	40	50	60	70
			^ . <u>.</u>				
CiSERK1 _{K59}	MDRTISLIA	AIWLIFGFNYVS	KVYCNAE GDA	LNALKTQLTDE	NNILQSWD	PTLVNPCTWFHI	TCDQG
CiSERK2 _{K59}	MDRRVSLISASVI	VWFILVFNNSS	KVHGNAEGDA	LNALKTOLGDE	NNVLOSWDA	TLVNPCTWFHV	TCNNE
	Peptide s	signal		Leucine	Zipper		
	80	90	100	110	120	130	140
CiSERK1 _{K50}	NSVSRIDLGNAGI	SGOLVPELGOL	SKLOYLELYG	MITGKIPEOI		LDLYLNKLEGDI	PDTLG
CiSERK2	NSVTRUDLONA	SCOLVAOLGEL	TNLOYIELYS	MUTGRIPDAT	M.TNI.VSI	DLYLNKLDGGT	PDTLG
OIDDIRCERSS							DD 2
	150	160	170	180	190	200	210
	1 1	100	1 1	1 1 1	1 20	200	210
C: CEDV1							····
CISERKI _{K59}	NLQQLKFFKLI	TLIGITENSLI	TIGSLQVIDL	SHILLSGIVPI	COSESLE I	I SI DGNLNLVI	QAAPP
CISERKZ _{K59}	KLQKLRFIRLNN	TLTGTIPISLT	TITSLQVLDL	SNNNLRGDVPV	N-SFSLFT	PISFANNPQLKA	PAVSP
		LRR 4		LRR	5		
	220	230	240	250	260	270	280
	· · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·				· •••• · · · · · · · · · · · · · · · ·	
CiSERK1 _{K59}	QSPVAPDSQSPS	SNSATGAIA GG	VAAGAALLFA	GPAIAIAWYRF	RRKPQDHFFI	DVPAEEDPEVHL	GQLKR
CiSERK2 _{K59}	QAPAPPNSPSSSV	GNSATGAIAGG	VAAGAALLFA	G <mark>PAIALAWW</mark> RF	RKPQDHFFI	OVPAEEDPEVHL	GQLKR
	1	Dom	aine transme	mbranaire		1	
	290	300	310	320	330	340	350
CiSERK1 _{K59}	FSLRELQVATDIE	SNNNILGRGGF	GKVYKGRLAD	GSLVAVKRLKE	ERSQGGEL	FQTEVEMISMA	VHRNL
CiSERK2 _{K50}	FSLRELOVATONE	NNRHILGRGGF	GKVYKGRLAD	GSLVAVKRLKE	ERTOGGEL	FOTEVEMISMA	VHRNL
1(5)	~	Signature d'u	in site de fivat	ion de l'ATP	A., .	he Au	
	360	370	380	390	400	410	420
			[].m	1]
CiSERK1 KEO	I.RI.RGFCMTPTER	TLVYPYMANGS	VASCLEREPD	TOPPLOWPIRK	RTALGAAR	T.AYLHDHCDPK	TTHRD
CiSERK2	I PI POFOMT PTER			TOFPLOWPIPK	PTAL COAP	TAXI HDHCDPK	TTURD
CIDENN2K59			VASCUILLATED	I QUE L'DWE I KR	MIALGOAN		ite setif
	120	110	450	160	170	Signature d'un s	ite actir
	430	440	400	400	470	400	490
al apprel							· · · ·
CISERKI _{K59}	VKAANILLDEEFE	AVVGDEGLAKL	MDYKDTHVTT	AVRGTIGHIAE	'EYLSTGKS:	SEKTDVEGIGVM	LLLL
CiSERK2 _{K59}	VKAANILLDEEFE	CAVVGDFGLAKL	MDYKDTHVTT	AVRGTIGHIAF	PEYLSTGKSS	SEKTDVFGYGVM	LLELI
c	activité sérine/thre	éonine kinase		1440 7440 1440	100 July 200	20 A	
	500	510	520	530	540	550	560
						• • • • • • • •	••••
CiSERK1 _{K59}	TGQRAFDLARLAN	DDEVMLLDWVK	GLLTDKMLEK	LVDEDLDGAYV	DDEVEQLIG	VALLCT QG T AM	ERPKM
CiSERK2 _{K59}	TGQRAFDLARLAN	IDDDVMLLDWVK	GLLREKKLET	LVDADLKGNYI	DDEVEQLIG	VALLCTQGTPL	ERPKM
			AN INCOME AND A DECEMPTION				
	570	580	590	600	610		
		.				1	
CiSERK1 _{K59}	SEVVRMI EGDGLA	ERWEEWQKEEI	FRQDYNQAHN	HNIDWIIPDST	WLSNEQLS	B G P R	
CiSERK2	SEVVRMI EGDGLA	ERWEEWOREEM	FROERNETHN	PNTDWITADST	YNLRPDEL	SGPR	
N09			· X - · · · · · · · · · · · · · · ·		and the second s		

<u>Flgure 33</u>: Alignement des protéines CiSERK1_{K59} et CiSERK2_{K59} déduites des séquences géniques correspondantes et caractéristiques remarquables. Les traits pointillés verticaux indiquent la position des jonctions entre les exons. La double flèche verticale rouge indique le site de clivage prédictif du peptide signal. Les doubles flèches horizontales indiquent les différents motifs prédits *in silico*. Les motifs LRR indiqués sont ceux déterminés en tenant compte de la structure génique (voir texte). La région encadrée en noir correspond au domaine trans-membranaire. La région encadrée en gris représente le domaine kinase. Les résidus asparagine entourés en bleu sont des sites potentiels de N-glycosylation.

domaine hydrophobe pouvant correspondre à une hélice transmembranaire et d'un domaine présentant des répétitions de motifs riche en leucines (LRR) (Figure 33).

Un site de clivage d'un peptide signal a été prédit entre les acides aminés en position 28 et 29 au moyen des logiciels SignalP (http://www.cbs.dtu.dk/services/SignalP/) et TargetP (http://www.cbs.dtu.dk/services/TargetP/). De plus, le profil d'hydrophobicité des protéines CiSERK, réalisé par le logiciel TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/), a révélé la présence d'une région hydrophobe aux extrémités N-terminales (position 1 à 28). Ce peptide possède les caractéristiques d'un signal d'adressage de la voie sécrétoire, compatible avec le positionnement des protéines CiSERK au niveau de la membrane plasmique.

Une région "Leucine Zipper" a été localisée dans le domaine extracellulaire à la suite du peptide signal (position 35 à 56). Cette région comprend 4 répétitions, d'un motif dont la séquence consensuelle constituée d'un résidu Leucine suivi de 6 acides aminés quelconques (Lx_6) (Landschulz *et al.*, 1988), et est impliquée dans les interactions protéiques (O'Shea *et al.*, 1989). Cette région est nécessaire à l'oligomérisation de AtSERK1 *in vivo*, vraisemblablement grâce à la présence de 2 cystéines à ce niveau (position 60 et 67) permettant la formation de liaison par pont dissulfure (Shah *et al.*, 2001a)

Une région contenant des motifs riches en leucine formant des répétitions LRR est présente entre le motif "Leucine Zipper" et le domaine transmembranaire. Ces motifs, connus pour être impliqués dans les interactions protéiques, sont dotés d'une séquence consensuelle LxxLxxLxxLxxLxxNxLxGxIPxx définie par Kobe et Deisenhofer (1995). Quatre répétitions peuvent ainsi être identifiées (positions 91-104, 105-138, 139-162, 163-186) et possèdent la séquence consensuelle LGxLxxLxDLxLxxNxLxGxIPxx pour les séquences CiSERK. Chacun des motifs 1, 3 et 4 sont alors chevauchants sur deux exons successifs et interrompus au niveau du 2^e nucléotide codant la leucine à ce niveau; le 2^e motif est inclus entièrement dans l'exon 4. Par ailleurs, 5 répétitions ont été déduites au moyen d'une recherche dans les bases de motif protéique PFAM (positions 72-95, 96-119, 120-143, 144-167, 168-193). La séguence consensus déduite des alignements des motifs LRR ainsi définis chez les protéines CiSERK est xLxxLxDLxxNxLxGxIPxxLGxLx. De la même façon les répétitions 1, 2, 4 et 5 chevauchent deux exons successifs, le 3^e motif est alors codé entièrement par l'exon 4. Dans les deux cas les motifs sont composés de 24 acides aminés. Il s'avère que les exons 3, 5 et 6 codent des motifs composés de 24 acides aminés correspondant à la taille d'un motif LRR, alors que l'exon 4 code un motif d'une taille double. Si on considère qu'un motif LRR peut être codé entièrement par un exon (Baudino et al., 2001) les protéines CiSERK possèdent 5 répétitions LRR contenues pour 3 d'entre elles dans un exon unique (LRR1, LRR4 et LRR5 codés respectivement par les exons 3, 5 et 6) tandis que les répétitions LRR2 et LRR3 sont codées par l'exon 4. Il résulte alors que la

Consens	sus CiSERK	DLxxNxLxGxIPxxLGxLxxLxxL
CiSERK1	LRR1	DL <mark>C</mark> NACLSCQLVPELCQLSKLQYL
CiSERK1	LRR2	ELYCNNITGKIPEQLGNLTQLMSL
CiSERK1	LRR3	DLYUNKLEGDIPDTLGNLQQLRFL
CiSERK1	LRR4	RLNNNTLTGTIPNSLTTIGSLQVL
CiSERK1	LRR5	DLSHNNLSGIVPIN-GSFSLFTPI
CiSERK2	LRR1	DLGNANLSGQLVAQLGELTNLQYL
CiSERK2	LRR2	ELYSNNITGRIPDALGNLTNLVSL
CiSERK2	LRR3	DLYLNKLDGGIPDTLGKLQKLRFL
CiSERK2	LRR4	RLNNNTLTGTIPISLTTITSLQVL
CiSERK2	LRR5	DLSNNNLRGDVPVN-GSFSLFTPI

séquence consensuelle des motifs LRR ainsi définie est DLxxNxLxGxIPxxLGxLxxLxxL pour les protéines CiSERK (Tableau 10).

<u>Tableau 10</u>: Séquences consensuelles issues des 5 répétitions des motifs LRR des protéines CiSERK.

La prédiction d'une hélice transmembranaire a été confirmée par le logiciel TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Cette région est composée de 24 acides aminés hydrophobes (position 230 à 254) et positionne le domaine kinase dans le cytoplasme.

Une analyse du domaine kinase des protéines CiSERK indique qu'il s'étend sur 278 résidus aminés en position 293 à 567 sur la figure 33. Ce domaine possède les 11 sous domaines fortement conservés et caractéristiques des protéines à activité kinase (Hanks *et al.*, 1988). Les signatures d'un site de fixation de l'ATP (acide aminé 299 à 321) ainsi que d'une activité sérine/thréonine kinase (position 416 à 428) ont également été détectés.

Pour les deux protéines CiSERK, 7 sites de N-glycosylation sont prédits ce qui indique que ces protéines peuvent être soumise à des modifications post-traductionnelles. Cinq sites sont situés dans le domaine extracellulaire et deux sites sont situés dans le domaine intracellulaire. Dans le domaine extracellulaire, 4 sites impliquent des résidus asparagine situés à la même position dans les mêmes motifs LRR (2 sites dans le motif LRR2, 1 site dans le motif LRR4 et 1 site dans le motif LRR5), le cinquième site prédit de N-glycosylation est situé au niveau du motif LRR5 et LRR1 pour les protéines CiSERK1_{K59} et CiSERK2_{K59} respectivement. Parmi les deux sites de N-glycosylation prédits dans le domaine intracellulaire, un est situé à la même position chez les deux protéines (position 372 sur la figure 33), le second implique des résidus asparagine différents.

Ces résultats permettent de conclure que les protéines CiSERK possèdent les caractéristiques qui définissent des protéines possédant une configuration de récepteur transmembranaire à activité kinase de type RLK-LRR identiques à celles de type SERK:



Figure 34: Arbre phylogénétique issu de l'alignement des domaines kinase des séquences CiSERK1 et CiSERK2 et des domaines kinases d'un membre représentatif de chacune des 22 sous-famille de RLK selon la classification de Shiu et Bleecker (2001). Le nom des différentes sous-familles est figuré en gras.

• un domaine extracellulaire composé d'un peptide signal, d'une région "Leucine zipper" et de 5 répétitions LRR.

• un domaine transmembranaire unique

• un domaine cytoplasmique caractéristique d'une activité sérine/thréonine kinase.

II-E. CLASSIFICATION DES PROTEINES DEDUITES CISERK

II-E.1. Les protéines CiSERK appartiennent à la famille des RLK-LRRII

Afin de positionner les protéines CiSERK au sein de la famille des RLK, les domaines kinases d'un membre représentatif de chaque sous-famille de RLK et ceux des protéines CiSERK ont été alignés pour construire un arbre phylogénétique (Figure 34). En effet, la séquence du domaine kinase est caractéristique de chaque sous-famille de récepteur de type kinase (Shiu et Bleecker, 2001b). Il apparaît que le domaine kinase des protéines CiSERK forme un groupe robuste (valeur de bootstrap égale à 1000) avec le domaine kinase de la protéine At4g30520, utilisé comme représentant de la sous-famille LRRII.

II-E.2. CiSERK1 et CiSERK2 appartiennent au groupe des protéines SERK

La sous-famille des RLK-LRRII comporte 14 membres chez Arabidopsis (Shiu et Bleecker, 2001b), et cinq appartiennent au groupe SERK (AtSERK1 à AtSRK5) (Hecht *et al.*, 2001). Les séquences protéiques complètes de ces 14 membres ainsi que les protéines de type SERK identifiées chez d'autres espèces comme la carotte (Schmidt *et al.*, 1997), le maïs (Baudino *et al.*, 2001) et la luzerne (Nolan *et al.*, 2003) ont été alignées avec les protéines CiSERK. Un arbre phylogénétique a été déduit de cet alignement (Figure 35). L'ensemble des protéines de type SERK est porté par la même branche et comprend les deux protéines CiSERK. Le groupe des protéines de type SERK peut être subdivisé en 2 sous-groupes distincts (valeur de boostrap égale à 1000); le premier comprend AtSERK1 et 2, MtSERK, DcSERK, ZmSERK1 et 2, le second groupe comprend AtSERK3, 4 et 5 et les 2 protéines CiSERK de chicorée.

Afin de préciser les caractéristiques des protéines CiSERK, les séquences protéiques apparentées à SERK chez *Arabidopsis*, la carotte, le maïs et la luzerne ainsi que les séquences protéiques déduites de *CiSERK1* et *CiSERK2* ont été alignées et les jonctions entre exons ont été positionnées à partir des informations de séquences géniques (Figure 36). Il apparaît que les gènes codant les protéines CiSERK1 et 2 ont la même structure que ceux déjà décrits. Une analyse au niveau nucléotidique a révélé que chacune des jonctions

		CISERK2	DcSERK	ZmSERK1	MtSERK1	AtSERK1	AtSERK2	AtSERK3	AtSERK4	AtSERK5
 Fl	CiSERK1	84,2	73,2	72,7	75,6	74,8	74,4	78,2	72,9	67,7
	CiSERK2		75,9	75,1	78	77,7	77,2	80,4	74,6	69,3
PS	CiSERK1	54,8	3,3	17,2	12,5	9,3	18,7	41,3	32,3	34,4
10	CiSERK2		3,2	22,5	15,6	12,5	18,7	41,9	33,3	25,8
17	CiSERK1	80	73,3	73,3	73,3	71,1	71,1	68,8	64,5	58,3
LZ	CiSERK2		84,4	73,3	82,2	82,2	82,2	82,2	77	64,5
LRR	CiSERK1	80,1	72,7	73,5	73,5	73,5	74,3	74,3	69,4	62,8
	CiSERK2		76,0	78,5	77,6	80,1	79,3	80,9	72,7	66,9
Bro rich	CiSERK1	38,4	16,2	16,2	18,9	21,6	18,9	25	21,4	10,7
FIGHCH	CiSERK2		29,7	27	27	24,3	27	32,1	39,2	25
тм	CiSERK1	93,1	77,2	65,9	81,8	77,2	79,5	84	77,2	54,5
I IVI	CiSERK2		79,5	70,4	81,8	81,8	84	86,3	81,8	59
KD	CiSERK1	93,7	88,4	91,7	89,7	91,0	90,0	92,3	88,0	84,7
	CiSERK2		89,7	92,7	90,3	91,0	90,3	93,0	87,7	84,4
C torm	CiSERK1	73,4	65,3	61,2	55,1	61,2	55,1	62,7	50,9	52,9
C term	CiSERK2		65,3	63,2	51,0	63,2	59,1	60,7	50,9	52,9

<u>Tableau 11</u>: Identité de séquence des protéines CiSERK avec les autres membres du groupe SERK. Les pourcentages d'identité ont été calculés pour les protéines complètes (FI) et les séquences correspondant à chaque domaine: PS: peptide signal, LZ: Leucine Zipper, LRR: Leucine Rich Repeat, Pro Rich: région riche en Proline, TM: domaine transmembranaire, KD: domaine kinase, C term, région C terminale.

entre exon est dans la même phase de lecture, quelle que soit l'espèce. Par exemple, la jonction entre l'exon 3 et l'exon 4 implique, pour l'ensemble des protéines, un résidu leucine codé par un codon dont deux bases sont comprises dans l'exon 3 tandis que la dernière base est contenue dans l'exon 4.

Les séquences protéigues CiSERK montrent une identité de séquence variant de 68% à 78 % avec les autres protéines du groupe SERK (Tableau 11). Cependant ces valeurs d'identité peuvent être très variables selon les motifs codés par les différents exons. En effet, certains domaines semblent moins conservés entre les différentes protéines. La faible identité de séquence correspondant au peptide signal peut être expliquée par la différence de taille entre les protéines. Le second exon code le motif leucine Zipper, pour lequel, AtSERK4 et 5 se distinguent par la présence de 3 acides aminés supplémentaires par rapport aux autres protéines. Les exons 3 à 6 codent les motifs LRR pour lesquels CiSERK1 et 2 montrent une identité globalement égale vis à vis de toutes les autres protéines SERK, à l'exception de AtSERK4 et 5 plus éloignées (70% et 65 % d'identité respectivement). L'exon 7 code une région riche en proline dont la séquence permet de distinguer les sous-groupes 1 et 2 (Figure 35). La présence d'un motif très conservé PCPGPPFSPPPP entre les protéines MtSERK, DcSERK, AtSERK1 et 2, et dans une moindre mesure ZmSERK1 (variation d'un acide aminé) est caractéristique du sous-groupe 1 Ce motif n'est pas conservé dans les protéines CiSERK1 et 2 ainsi que AtSERK3, 4 et 5. Cependant, bien que cette région soit plus courte dans ces séquences, elle présente plusieurs résidus proline mais une seule est conservée. L'exon 8 code le domaine transmembranaire pour lequel AtSERK5 est isolé par l'absence d'acides aminés conservés dans les autres séquences. Les exons 9, 10 et la première moitié de l'exon 11 correspondent au domaine kinase et montrent une forte identité de séquence entre les différentes protéines. La seconde moitié de l'exon 11 correspond à la région C terminale pour laquelle, les protéines du sous-groupe 1 se distinguent de celui formé par AtSERK3, 4 et 5. Les protéines CiSERK sont, par cette région, intermédiaires entre ces deux sous-groupes.

Par la nature des protéines et des motifs codés par les différents exons, les protéines CiSERK appartiennent au groupe des protéines SERK. Au sein de ce groupe, alors que deux sous-groupes peuvent se distinguer, les protéines CiSERK possèdent souvent les caractéristiques communes avec le sous-groupes comprenant AtSERK3, 4 et 5 et plus particulièrement avec la protéine AtSERK3.



<u>Figure 35</u>: Arbre phylogénétique issu de l'alignement des séquences protéiques complètes des différents membres de la famille des RLK-LRRII d'*Arabidopsis* (dont AtSERK1 à 5) et des séquences protéiques de type SERK chez la carotte (DcSERK), le maïs (ZmSERK1 et 2), la luzerne (MtSERK1) et des séquences CiSERK1 et CiSERK2 de chicorée. En vert est figurée la séquence utilisée comme membre des RLK-LRRII pour générer la figure 34. Le groupe comprenant les cinq membres de la famille de SERK chez *Arabidopsis* est encadré en gris. Les valeurs de boostrap indiquant la robustesse de chaque nœud sont indiquées.



Région C-terminale

XI

III DISCUSSION

Méthodes d'identification utilisées

L'objectif de cette partie était de mettre en évidence l'existence de gènes codant des protéines de type SERK chez la chicorée. Dans la mesure ou aucune donnée de séquençage n'existe chez la chicorée, plusieurs stratégies étaient envisageables comme un crible de banque d'ADNc de feuilles soumises à l'induction de l'embryogenèse somatique, ou encore à un crible de banque d'ADN génomique. Cependant, ces deux types de banques ne sont pas disponibles au laboratoire et le crible d'une banque d'ADNc impliquait que les gènes cibles soient exprimés à un niveau suffisant pour être représentés dans une banque, ce qui pouvait ne pas sembler le cas au regard du niveau très faible d'expression de *DcSERK* observé chez la carotte (Schmidt *et al.*, 1997).

C'est pourquoi, nous avons opté pour une stratégie basée sur la PCR au moyen d'amorces dégénérées. L'obstacle de la présence d'un faible niveau de transcrit de gènes apparenté à *SERK* chez la chicorée a été pallié en initiant le clonage à partir d'ADN.

Le fait d'amplifier de l'ADN présente l'avantage, en plus du fait de ne pas être dépendant du niveau d'expression des gènes, d'avoir accès à des éléments de structure génique comme la présence ou non d'introns qui peut alors constituer un critère supplémentaire à l'identité de séquence pour le choix de séquences candidates. Ceci nous a permis de mettre en évidence une grande variété de séquences codant toutes des fragments de domaines kinase chez la chicorée et de réaliser une première sélection des séquences montrant à ce niveau une forte identité de séquence protéique avec DcSERK et d'autre part possédant la même structure génique que *DcSERK*, caractérisée par la présence d'un intron. Cet élément laissait entrevoir que parallèlement à la conservation des séquences protéiques entre espèce, la structure génique pouvait elle aussi être conservée.

Pour réaliser le clonage des ADNc pleines longueurs de *CiSERK1*₄₇₄ et *CiSERK2*₄₇₄ nous avons mis au point une technique originale de RT-PCR inverse qui offre des résultats très satisfaisants puisqu'elle a permis d'obtenir en une seule étape de PCR à la fois la totalité de l'extrémité 3' des séquences codantes mais aussi une séquence de longueur variable en 5' de la région connue. L'obtention des extrémités 5' a été plus limitante dans l'obtention des ADNc pleines longueurs mais il s'avère que ce problème est commun avec les méthodes classiques de 5'RACE PCR.

L'obtention des extrémités 5' des ADNc par RT-PCR inverse peut être limitée par plusieurs éléments qui peuvent réduire l'efficacité de l'étape de transcription inverse. Ainsi, la processivité de l'enzyme transcriptase inverse peut être limitante du fait de la présence de structures secondaires des ARNm. D'autre part, la nature tronquée des ARNm en 5' ne permet pas d'atteindre les extrémités 5', limitant de ce fait leur amplification postérieure. Les ADNc issus des transcrits pleines longueurs sont alors en quantités trop faibles pour être amplifiés majoritairement et subissent une compétition en faveur de l'amplification des fragments plus courts, ceci d'autant plus que les ARNm ciblés sont faiblement représentés. Néanmoins la robustesse de la RT-PCR inverse pour l'obtention de la région 3' et d'une région 5' partielle ou totale d'une cible est attestée par les différentes cibles clonées de cette façon au laboratoire et peut être considérée comme une alternative moins coûteuse aux différents kits de RACE PCR présents sur le marché. Néanmoins, nous avons dû utiliser la technique de RLM-RACE PCR pour cloner l'extrémité 5' de l'ADNc *CiSERK2474*.

Nombre de gènes de type SERK chez la chicorée

L'obtention des séquences géniques de CiSERK1 et CiSERK2 chez les génotypes K59 et K28 a révélé la présence de polymorphismes de séquence pour CiSERK1. Deux formes alléliques ont été clonées chez le génotype K59 tandis qu'une seule forme allélique est présente chez le génotype K28. Le polymorphisme de CiSERK1, ainsi que celui observé par ailleurs pour CiSERK2 chez d'autres génotypes, ont permis de cartographier ces deux gènes et de montrer qu'ils correspondent à 2 loci distincts. Ainsi, nous avons mis en évidences deux gènes codant des récepteurs de type kinase apparenté à DcSERK chez la chicorée. L'existence d'autres gènes de ce groupe chez la chicorée a été montrée par Southern blot. Les gènes de type SERK composent une petite famille multigénique dont 5 membres existent chez Arabidopsis (Hecht et al., 2001): deux sont situés sur le chromosome 1 (AtSERK1 et AtSERK2), deux sont situés en tandem sur le chromosome 2 (AtSERK4 et AtSERK5) et un est situé sur le chromosome 4 (AtSERK3). Chez la luzerne, la présence de 5 membres apparentés à SERK a été montrée par Southern blot (Nolan et al., 2003). Par ailleurs 3 gènes de type SERK ont été cartographiés chez le maïs sur 3 chromosomes différents (Baudino et al., 2001). De plus 4 séquences nucléotidiques partielles correspondant à un domaine kinase apparenté à SERK ont été séquencées chez le tournesol (Thomas et al., 2004). Le clonage et la cartographie de deux gènes distincts de type SERK ainsi que les expériences de Southern blot indiquent que d'autres gènes de types SERK existent chez la chicorée, et formeraient aussi une famille multigénique chez cette espèce. Le nombre de gènes de type SERK chez la chicorée pourrait être déterminé par l'utilisation d'une sonde correspondant au domaine kinase, qui aurait un plus large spectre d'hybridation.

Les gènes CiSERK

Les séquences protéiques déduites des gènes *CiSERK* indiquent qu'ils possèdent une structure de type LRR-RLK comme les gènes de type *SERK* identifiés chez d'autres espèces. De plus, les protéines CiSERK montrent plus de 70 % d'identité de séquence avec la plupart des protéines de type SERK ce qui confirme que les gènes clonés chez la chicorée appartiennent à la famille des gènes de type *SERK*. D'autre part la structure de ces gènes est fortement conservée entre espèces puisque les différents gènes de type *SERK* sont composés de 11 exons et 10 introns ce qui est aussi le cas des gènes *CiSERK*. Enfin la position des jonctions entre exons, mais aussi la phase de lecture de l'acide aminé impliqué dans les jonctions entre exons, est identique pour l'ensemble des gènes de type *SERK*.

Par ailleurs chaque exon code un motif structural correspondant aux caractéristiques des différents gènes de type *SERK* existants (Hecht *et al.*, 2001). Le premier exon correspond à une séquence peptide signal, le second exon code une région "Leucine Zipper" nécessaire à l'oligomérisation montrée pour AtSERK1, les exons 3 à 6 codent 5 motifs LRR. La région riche en proline est codée par l'exon 7 suivie du domaine transmembranaire codé par l'exon 8. Les exons 9 à 11 codent le domaine intracellulaire à activité kinase. La seconde moitié de l'exon 11 correspond à la région C-terminale.

Position des protéines CiSERK au sein du groupe SERK

Si le domaine kinase permet assez bien de prédire la sous-famille de RLK à laguelle appartient une protéine de type récepteur kinase (Shiu et Bleecker, 2001b), ce domaine ne permet pas de déterminer avec précision le degré de parenté avec un membre particulier d'une famille. En effet, sur la base du domaine kinase, il est difficile de discriminer les protéines du groupe SERK du fait d'une identité de séquence supérieure à 90% entre tous les membres y compris CiSERK1 et 2. Les protéines de type SERK forment un groupe au sein de la famille des RLK-LRRII. Néanmoins, il a été proposé que les protéines de type SERK d'Arabidopsis formaient 2 sous-groupes composés d'une part des protéines AtSERK1, et AtSERK2 et d'autre part des protéines AtSERK3, 4 et 5 (Hecht et al., 2001). Notre analyse phylogénétique montre que les différentes protéines de type SERK mises en évidence chez d'autres espèces appartiennent au premier groupe et que les protéines CiSERK sont plutôt rattachées au sous-groupe 2, en particulier pour ce qui est de la région riche en proline codée par l'exon 7. Cependant, l'analyse de l'alignement de l'ensemble des protéines de type SERK indique que les protéines CiSERK possèdent parmi les motifs les moins conservés des motifs apparentés à ceux du sous-groupe 1 comme le motif "Leucine Zipper" de l'exon 2 (commun à AtSERK3) ou la région C-terminale de l'exon 11 (différente d'AtSERK3), et des formes de motifs conservés dans le sous-groupe 2 comme la région riche en proline de l'exon 7. En considérant les deux sous-groupes, les protéines CiSERK semblent d'avantage appartenir au groupe AtSERK3 qu'à celui comprenant AtSERK1.

Parmi les 5 gènes de types *SERK* d'*Arabidopsis*, *AtSERK1* présente une expression importante dans les boutons floraux avant l'anthèse et dans les ovaires 3 jours après la pollinisation. Les ARNm sont détectés dans une moindre mesure dans les autres tissus, dans les feuilles, les ARNm sont dix fois moins abondants que dans les boutons floraux (Hecht *et al.*, 2001). *AtSERK3* est aussi connu sous le nom de *BAK1* (BRI1 associated kinase 1), et est impliqué dans la réception et la transduction du signal lié aux brassinostéroïdes en synergie avec *BRII* avec lequel il interagit (Nam et Li, 2002) et est exprimé de façon ubiquiste. Le rôle ou la fonction des autres protéines apparentées à SERK chez *Arabidopsis* n'est pas connu.

Un moyen de déterminer si les gènes *CiSERK* sont des équivalents fonctionnels de *DcSERK* et *AtSERK1*, est de déterminer leur profil d'expression au cours de processus dans lesquels *DcSERK* et *AtSERK1* sont impliqués.

Nous avons donc entrepris de déterminer les profils d'accumulation des transcrits des deux gènes *CiSERK* au cours des phases précoces de l'induction de l'embryogenèse somatique et lors de l'embryogenèse zygotique et du développement de la graine de chicorée.

Partie II: Expression des gènes *CiSERK* au cours de l'induction de l'embryogenèse somatique et lors du développement de la graine <u>DEUXIEME PARTIE</u> : EXPRESSION DES GENES *CISERK* AU COURS DE L'INDUCTION DE L'EMBRYOGENESE SOMATIQUE ET DU DEVELOPPEMENT DE LA GRAINE

I VARIATION DES NIVEAUX DE TRANSCRITS

L'étude de la variation des niveaux de transcrits des gènes *CiSERK1* et *CiSERK2* a été entreprise afin de déterminer la modulation de l'expression de ces gènes au cours des phases précoces de l'embryogenèse somatique de la chicorée. De même, le profil d'expression des gènes *CiSERK* a été déterminé au cours de l'embryogenèse zygotique et du développement de la graine. Ces résultats pourront nous permettre alors de comparer les profils d'expression des gènes *CiSERK* avec ceux décrits pour d'autres gènes de type *SERK*.

La technique utilisée est la RT-PCR quantitative en temps réel (qRT-PCR) qui permet d'allier la haute sensibilité et la forte spécificité de la PCR.

I-A. PRINCIPE DE LA PCR EN TEMPS REEL ET METHODES D'ANALYSES

En considérant la formule théorique de la PCR,

 $X_n = X_0 \times (1 + E)^n$

où X_n est la quantité de produits formés après n cycles de PCR, X₀ est la quantité de matrice initiale et E est l'efficacité de l'amplification, il existe une relation directe entre la quantité de matrice initiale et la quantité de produits formés. Lorsque le nombre de cycles devient élevé, cette relation n'existe plus; la réaction atteint un plateau du fait d'événements de compétitions et de saturations. Cette proportionnalité n'est observée qu'en début de PCR, c'est à dire pour de faibles nombres de cycles, lorsque ces produits ne sont pas visibles par les méthodes classiques d'observation (*i.e.* par migration sur gel d'agarose suivie d'une coloration des gels par le bromure d'éthidium).

Par la technique de PCR en temps réel, le produit formé à chaque cycle de la réaction d'amplification est visualisé grâce à un fluorochrome et donne accès à la phase exponentielle de la réaction.

Les produits d'amplification ont été visualisés grâce au SYBR green I, un agent intercalant qui se lie spécifiquement à l'ADN double brin. Lorsqu'il est fixé à l'ADN et excité par une radiation de longueur d'onde de 490 nm, il émet une fluorescence de longueur d'onde 530 nm, d'intensité proportionnelle à la quantité de SYBR green I lié aux doubles



Figure 37: Positions des amorces utilisées pour l'étude de la variation des niveaux de transcrits de CiSERK1 et CiSERK2 par PCR en temps réel. Seule la région 3' de la séquence génique de CiSERK1 (exposée figure 27) ainsi que la région 5' de CiSERK2 (exposée figure 28) sont présentées. Les régions encadrées correspondent aux positions des amorces, la région soulignée indique le fragment amplifié.

brins d'ADN. Cette fluorescence émise est mesurée par le bloc optique du thermocycler à chaque cycle d'amplification.

Le traitement informatique des signaux obtenus permet de reconstituer la courbe d'amplification de la réaction de PCR, qui est graphiquement représentée par la fluorescence émise à chaque cycle de la réaction. Il est alors possible d'avoir accès à la phase exponentielle de la réaction.

La technique de qRT-PCR peut être utilisée dans la cadre de l'étude de la variation du niveau de transcrits. Dans ce cas, les matrices sont constituées d'ADNc simples brins obtenus par transcription inverse d'ARN. Le niveau de transcrits d'un gène d'intérêt (cible) est normalisé au niveau de transcrits d'un ou de plusieurs gènes contrôles.

I-A.1. Définition des amorces

Un couple d'amorce a été défini pour chacune des cibles testées et des contrôles utilisés.

Pour les gènes *CiSERK*, les amorces utilisées pour réaliser les amplifications ont été définies dans des régions spécifiques de chaque gène, constituées notamment des régions non traduites des transcrits. Pour le gène *CiSERK1*, dont il existe deux formes alléliques, les amorces ont été définies de telle sorte qu'elles ne montrent pas de mésappariement entre les deux allèles. Le couple d'amorce utilisé permet d'amplifier un fragment de 144 pb (Figure 37); l'amorce directe est située au niveau de la séquence codante et l'amorce complémentaire se situe dans la région 3' non traduite. Le couple d'amorce défini à partir de la séquence *CiSERK2* permet l'amplification d'un fragment de 80 pb (Figure 37); l'amorce directe est située dans la région 5' non traduite tandis que l'amorce complémentaire est située dans la région 5' non traduite tandis que l'amorce complémentaire est située dans la séquence codante correspondant au peptide signal.

Un ADNc partiel d'*actine* de chicorée, dont la séquence protéique déduite présente une forte identité (90%) avec les actines de type 2 et de type 8 d'*Arabidopsis*, a été cloné au laboratoire, à partir d'une suspension cellulaire de chicorée (Poulain, 2003). Ce gène d'actine a été utilisé comme contrôle interne (Annexe 3).

Le second contrôle que nous avons utilisé est le gène codant les ARN *18s*. L'amplification des ADNc issus de ces ARN ribosomiques a été possible car les réactions de transcriptions inverses ont été initiées par un mélange d'amorces contenant, non seulement des amorces oligo(dT), mais aussi des amorces hexamères de séquences aléatoires. Du fait de la très grande conservation des séquences d'ADN *18s*, des amorces ont été définies à partir d'une séquence partielle de tournesol (Annexe 4), botaniquement proche de la chicorée puisque de la même famille (Astéracées).



Figure 38: Cinétique d'amplification de trois répétitions d'un échantillon visualisée par PCR en temps réel. La fluorescence est mesurée à chaque cycle d'amplification. Le seuil est défini dans la phase exponentielle de l'amplification pour une valeur égale à au moins 10 fois l'écart-type de la valeur moyenne mesurée pour la fluorescence de base. La projection sur l'axe des abscisses du point d'intersection de la courbe d'amplification et du seuil de fluorescence détermine la valeur du Ct.



<u>Figure 39</u>: Méthode de calcul du Δ Ct. La courbe rouge correspond à la cinétique d'apparition des produits d'amplification d'un gène cible (X). La courbe bleue correspond à la cinétique d'apparition des produits d'amplification d'un gène contrôle (Y). La projection sur l'axe des abscisses du point d'intersection de la courbe et du seuil de fluorescence indique le nombre de cycles de PCR nécessaires pour atteindre le seuil. Le Δ Ct peut alors être calculé comme indiqué dans l'encadré 5 et correspond à la normalisation du niveau de transcrits par rapport à celui du gène contrôle.

Le gène d'une hémoglobine non symbiotique de chicorée *CHI7507*^e (Annexe 5) a été utilisé comme témoin de la réponse des explants à l'induction de l'embryogenèse somatique. En effet, ce gène est induit de façon transitoire lors des phases précoces de l'embryogenèse somatique de la chicorée hybride '474' (Hendriks *et al.*, 1998).

Les séquences des différentes amorces utilisées sont indiquées dans le tableau 12.

Cible	Amorce directe	Amorce complémentaire	Taille amplicon (pb)
Actine	GACATGAAAGAGAAACTCGCCTACG	CTTTCGGCTCCAATGGTGATGAC	131 pb
18s	GCCTACTATGGTGGTGACGGGTGAC	GCCTGCTGCCTTCCTTGGATGTG	95 pb
CiSERK1 _{K59}	TGTCGGGTCCTAGATGATTGGATTC	CCAAAGCAACCGCCAACCTATATC	144 pb
CiSERK2 _{K59}	TTTAGCTTTACCGACCCGAACCC	ACCAGACAAGAACAGAAGCTGAAATC	80 pb
CHI7507	GTGAGGTTGTGGTTTCTGGTTCTAC	CACTACATTTCTCTCCCATTGCCTTC	140 pb

<u>Tableau 12</u>: Séquences des amorces utilisées pour l'étude de la variation des niveaux des transcrits par RT-PCR en temps réel. La position des amorces sur les séquences des gènes est indiquée en annexe (*actine* Annexe 3; 18s Annexe 4; hémoglobine non symbiotique (*CHI7507*) Annexe 5). La localisation des amorces des gènes cibles $CiSERK1_{K59}$ et $CiSERK2_{K59}$ est indiquée sur la figure 37.

I-A.2. Exploitation des données de PCR en temps réel

La première étape de l'analyse des données générées par PCR en temps réel consiste à définir un seuil de fluorescence de telle sorte que ce seuil se situe dans la phase exponentielle de la réaction et soit suffisamment différent du niveau de fluorescence basal (au moins 10 fois l'écart type de la valeur de fluorescence basale moyenne). Ce seuil coupe la courbe d'amplification et la projection sur l'axe des abscisses du point d'intersection permet de définir le cycle d'amplification permettant d'atteindre le seuil de fluorescence (Ct pour Cycle Threshold) (Figure 38). La valeur du Ct est directement corrélée à la quantité initiale de matrice selon l'équation:



où X_{Ct} est la quantité de produit formée après Ct cycles d'amplification d'une quantité initiale X_0 de cible. Un Ct faible signifie une quantité de matrice initiale élevée tandis qu'un Ct important signifie une faible quantité de matrice initiale.

I-A.2.1. Méthode du ∆∆Ct

Les données de qRT-PCR peuvent être exploitées de deux manières. Une quantification absolue des cibles étudiées peut être obtenue par référence à une gamme

⁹: Le gène de l'hémoglobine non symbiotique correspondant à *CHI7507* (N°ACC AJ007507) correspond au clone 3206 publié sous le même n° d'accession EMBL par (Hendriks *et al.*, 1998).

<u>Encadré 5</u>: Mesure de la variation du niveau de transcrits par PCR en temps réel selon la méthode du $\Delta\Delta$ Ct (Livak et Schmittgen, 2001, Applied Biosystems User Bulletin N°2 (P/N 4303859)).

L'exemple présenté considérera le niveau de transcrits d'un gène cible X normalisé à un gène de contrôle Y entre 2 échantillons biologiques différents A, B.

Normalisation du niveau de transcrit de la cible par rapport à celui du contôle.

La normalisation du niveau de transcrit d'un gène cible par rapport à un gène contrôle

correspond au rapport de la quantité initiale de la cible X0 et de la quantité initiale d'un gène

contrôle Y_0 soit: $\frac{X_0}{Y_0}$

Selon l'équation de la PCR au seuil de fluorescence dans la phase exponentielle d'amplification ce rapport est égal à $\frac{X_{Ct}}{Y_{Ct}} = \frac{X_0 \times (1+E_x)^{-Ct_x}}{Y_0 \times (1+E_y)^{-Ct_y}}$

où $\frac{X_{Ct}}{Y_{Ct}}$ représente le rapport des quantités de produits formés au seuil de fluorescence. Cette valeur est une constante K

Dans la mesure où les efficacités d'amplification de la cible (Ex) et du contrôle (Ey) sont identiques (E), alors:

$$\mathsf{K} = \frac{\mathsf{X}_{\mathsf{Ct}}}{\mathsf{Y}_{\mathsf{Ct}}} = \frac{\mathsf{X}_0}{\mathsf{Y}_0} \times (1+\mathsf{E}) \ {}^{\mathsf{Ct}_x - \mathsf{Ct}_y} = \mathsf{N} \times (1+\mathsf{E}) \ {}^{\mathsf{Ct}_x - \mathsf{Ct}_y} = \mathsf{N} \times (1+\mathsf{E}) \ {}^{\Delta\mathsf{Ct}}$$

Où N est la quantité relative de cible normalisée à celle du contrôle et ∆Ct correspond à la différence des Ct de la cible et du contrôle.

Cette équation peut aussi être posée sous la forme: $N = K \times (1+E)^{-\Delta Ct}$

Expression du niveau de transcrits relativement à un échantillon de référence

Le niveau de transcrits de l'échantillon B est exprimé par rapport à l'échantillon A considéré comme l'état de référence selon la formule:

 $\frac{N_{B}}{N_{A}} = \frac{K \times (1+E)^{-\Delta Ct_{B}}}{K \times (1+E)^{-\Delta Ct_{A}}} = (1+E)^{-\Delta Ct_{B}} + \Delta Ct_{A} = (1+E)^{-\Delta \Delta Ct_{A}}$

Dans la mesure où les efficacités d'amplification (E) des ADNc cibles et contrôles sont considérées comme égales à 100 %, le niveau de transcrits de l'échantillon B, relativement à celui de l'échantillon A, peut être calculé selon la formule:

$$QR = \frac{N_B}{N_A} = 2^{-\Delta\Delta Ct}$$

standard de quantité de cible (nombre de copies ou ng). La quantification des transcrits peut aussi être mesurée de façon relative au niveau de transcrits d'un état de référence. La méthode d'exploitation est similaire à la quantification absolue mais repose sur des formules arithmétiques qui permettent de se dispenser de la quantification absolue initiale des cibles (Encadré 5).

Afin de corriger les variations de quantité d'ADNc induites par le dosage des ARN et l'étape de transcription inverse, le niveau de transcrits d'un gène cible a été exprimé de façon relative à celui d'un gène contrôle. Cette normalisation s'exprime par la différence de Ct au seuil de fluorescence entre la cible et le contrôle (Δ Ct) (Figure 39). La variation du niveau de transcrits ainsi normalisé de la cible étudiée entre plusieurs échantillons se traduit par la variation du Δ Ct entre les échantillons ($\Delta\Delta$ Ct). L'expression du $\Delta\Delta$ Ct est faite par rapport à un échantillon de référence. Le niveau de transcrits relatif de l'état de référence (QR) est alors égal à 1. Cependant l'utilisation de la méthode du $\Delta\Delta$ Ct implique que les efficacités d'amplification des cibles et des références soient considérées comme identiques et égales à 100 %.

I-A.2.2. Evaluation des efficacités d'amplification des ADNc cibles et contrôles

Nous avons évalué les efficacités d'amplification obtenues pour chacun des couples d'amorces définis afin de pouvoir justifier de l'utilisation de la méthode du $\Delta\Delta$ Ct. Pour cela, une gamme de dilution de matrice d'ADNc a été amplifiée (Figure 40) pour chaque couple d'amorces. La pente (p) de la droite de régression déduite d'un graphique exprimant le Ct en fonction du logarithme de la quantité relative initiale, entre les différents points de gamme, est liée à l'efficacité de l'amplification selon l'équation:

$E= 10^{(-1/p)} - 1$

Pour chaque couple d'amorces utilisées, des gammes constituées de l'amplification d'au moins trois dilutions sériées d'ADNc ont été réalisées pour chacun des tissus testés et, le cas échéant, pour chacun des génotypes utilisés. Le tableau 13 indique les valeurs moyennes d'efficacité d'amplification mesurée lors de l'utilisation des différents couples d'amorces.


<u>Figure 40</u>: Exemple d'évaluation de l'efficacité de l'amplification du couple d'amorces correspondant à *CiSERK2*. L'efficacité d'amplification est déduite de la représentation graphique du Ct mesuré en fonction du Log (quantité relative de matrice initiale). La quantité initiale amplifiée correspond à une série de dilution de matrices ADNc. Arbitrairement, la solution la moins concentrée correspond à une quantité initiale égale à 1 (Log =0), et les quantités relatives des solutions plus concentrées sont calculées selon les facteurs de dilutions réalisées dans ce cas 1/4 (Log=0.6), 1/16 (Log=1.2), 1/64 (Log=1.8). L'efficacité d'amplification (E) est une fonction de la pente de la droite de régression selon l'équation E= 1-10^(-1/p), le coefficient de corrélation détermine la qualité de la mesure.

	n	Е	e-tE	r	e-tr	Tableau 13: Efficacités de l'amplification selon
Actine	4	99.88	5.34	0.993	0.008	les couples d'amorces utilisés. E: efficacité moyenne calculée à partir de n mesures
18s	4	100.58	3.53	0.995	0.007	réalisées. e-tE: écart-type des efficacités
CiSERK1	5	101.88	4.11	0.990	0.006	calculé sur n mesures réalisées; le coefficient
CiSERK2	5	100.92	4.48	0.992	0.003	de corrélation indique la qualité des mesures. e-
CHI7507	1	102.90	-	0.998	-	mesurés.

Quel que soit le couple d'amorces testées, les valeurs d'efficacité d'amplification sont très proches de 100 % et nous les avons considérées identiques à 100 %. Dans le calcul final de la quantité relative de transcrits, la valeur de l'efficacité d'amplification (1+E) sera considérée comme égale à 2, la variation du niveau de transcrits par rapport à un état de référence sera alors exprimée selon la formule QR = $2^{-\Delta\Delta Ct}$ (cf encadré 5).

I-A.2.3. Méthodologie appliquée

Pour chaque extrait d'ARN (*i.e.* chaque jour de culture d'induction de l'embryogenèse somatique pour chacun des génotypes et chaque point de cinétique de l'embryogenèse zygotique), 3 transcriptions inverses (RT) ont été réalisées.

Pour chaque RT, 3 amplifications simultanées (triplet) ont été réalisées. Le Ct moyen de chaque triplet a été calculé pour chacun des gènes cibles et contrôles utilisés et le Δ Ct a été calculé par la différence du Ct moyen correspondant au gène cible et du Ct moyen correspondant au gène contrôle. Chaque échantillon et chaque cible est donc représenté par 3 valeurs de Δ Ct.

Les 3 valeurs de Δ Ct, issues des trois transcriptions inverses de l'échantillon utilisé comme état de référence (*i.e.* J0 pour l'embryogenèse somatique et J0ap pour l'embryogenèse zygotique), ont permis de calculer le Δ Ct moyen de l'état de référence. Ce Δ Ct moyen de référence a été soustrait à chacun des 3 Δ Ct de chaque échantillon pour calculer 3 valeurs de $\Delta\Delta$ Ct pour chaque échantillon. Le niveau relatif (Qr) de transcrits correspondant a été calculé à partir de chacune des valeurs de $\Delta\Delta$ Ct. Enfin, la moyenne et l'écart type des niveaux relatifs de transcrits ont été calculés pour chaque échantillon. Cette variation de transcrits est ainsi exprimée de façon relative au niveau moyen de transcrit mesuré à l'état de référence.

Des analyses des variances ont été réalisées selon la procédure GLM (General Linear Model) du logiciel SAS. Les moyennes des niveaux de transcrits ont été comparées en réalisant un test de Student-Newman-Keuls, à un seuil de signification de 5%.

Initialement, les essais ont été réalisés à partir d'ARN traités par de la DNAse afin d'éliminer les traces d'ADN contaminants. Cependant, la reproductibilité des triplets d'amplification à partir de ces échantillons s'est révélée être médiocre. Un protocole sans traitement des ARN avec de la DNase a donc été développé. Le niveau de contamination ADN de tous les échantillons d'ARN a été évalué pour chacune des cibles et des contrôles en réalisant des amplifications à partir des solutions d'ARN non soumises à la transcription inverse, dans les mêmes conditions expérimentales que pour les échantillons transcrits inversement. La différence de Ct entre les échantillons transcrits inversement et non transcrits inversement s'est toujours révélée être supérieure à 7 Ct, ce qui représente une différence de quantité supérieure à 100 fois entre les ADNc et les ADN contaminants. Ce niveau de contamination a été considéré comme négligeable, dans le cadre d'une évaluation de la variation du niveau de transcrits.

Lors de chaque réaction de PCR et pour chaque couple d'amorces testées, des réactions d'amplification ont été réalisées en absence de matrice d'ADNc afin de vérifier l'absence de contaminations ou la formation de dimères d'amorces.

En fin de réaction, une courbe de fusion des produits de PCR a été réalisée afin de déterminer la spécificité du produit d'amplification par sa température de fusion.

Pour réaliser ces expériences et du fait que cette technique était nouvelle au laboratoire, nous avons privilégié la répétition des expériences sur des échantillons biologiques uniques, à la répétition d'échantillons biologiques. Cependant, les échantillons dont ont été extraits les ARN et les protéines ont été caractérisés cytologiquement afin de vérifier la nature des événements provoqués. De la même façon, les observations des embryons en développement dans les akènes ont été effectuées sur les mêmes lots que ceux à partir duquel les analyses ont été réalisées.

Pour l'étude de la variation du niveau de transcrits des gènes *CiSERK1* et *CiSERK2* au cours d'une cinétique d'induction de l'embryogenèse somatique, 3 répétitions de transcription inverse ont été réalisées pour le génotype K59 et 2 répétitions de transcription inverse ont été exploitées pour le génotype C15. Pour le gène *CHI7507*, une seule transcription inverse a été réalisée. Les niveaux de transcrits ont été exprimés en référence au niveau mesuré dans les feuille avant mise en culture (J0).

L'étude de la variation de l'expression des gènes *CiSERK1* et *CiSERK2* au cours de l'embryogenèse zygotique et du développement de la graine a été réalisée à partir de 3 répétitions de transcriptions inverses. Les niveaux de transcrits ont été exprimés relativement au niveau moyen mesuré dans les ovaires avant la pollinisation (J0ap). Dans tous ces cas, 2 gènes de contrôle ont été utilisés (*actine* et *18s*). Les études d'expression à



<u>Figure 41</u>: Caractérisation cytologique de la réaction des génotypes K59 et C15 aux conditions de culture permettant l'induction de l'embryogenèse somatique chez un génotype embryogène.

1: Figures cytologiques de référence observées dans des fragments foliaires du génotype embryogène (K59). Les figures cytologiques caractéristiques sont entourées. 2: Proportion de cellules dans différents états cytologiques observés dans des fragments foliaires de chicorée embryogène (K59) et non embryogène (C15) au cours du processus d'embryogenèse somatique. Pour chaque état cytologique, des lettres différentes indiquent une différence significative. L'absence de lettre indique que la valeur n'est pas statistiquement différente de 0.

CER: Cellules en cours de réactivation; CR: Cellules réactivées; ZD: Zones de divisions denses: ZV: Zones de divisions vacuolisées.

partir de germinations et de boutons floraux ont été réalisées à partir d'une réaction de transcription inverse en utilisant le gène *actine* comme contrôle.

I-B. EXPRESSION DES GENES CISERK1 ET CISERK2 AU COURS DES PHASES PRECOCES DE L'EMBRYOGENESE SOMATIQUE

I-B.1. Evènements cytologiques observés

L'embryogenèse somatique est un processus qui implique des modifications de l'identité cellulaire entraînant une modification de la morphologie des cellules. Certaines cellules passent d'un état différencié de cellules de feuilles à celui de cellules embryogènes capables de se développer en embryon. Ces modifications de la morphologie des cellules sont caractéristiques de l'évolution de l'induction de l'embryogenèse somatique et permettent de corréler la variation des niveaux des transcrits des gènes aux différentes étapes des processus du développement des embryons.

Deux génotypes dont le fond génétique est proche ont été soumis à des conditions de cultures permettant l'induction de l'embryogenèse somatique. Le génotype K59 est très embryogène tandis que le génotype C15 (issu de l'autofécondation de K59) est incapable de développer des embryons somatiques.

La figure 41 montre les différentes structures cytologiques observées dans des tissus foliaires induits au cours de l'embryogenèse somatique (Da Silva, 2004; Blervacq, communication personnelle). Des cellules en cours de réactivation (CER) peuvent être observées. Tandis que les cellules du parenchyme foliaire montrent un noyau périphérique lenticulaire, les CER montrent un noyau qui s'arrondit tandis qu'apparaît le nucléole. Certaines de ces cellules peuvent alors atteindre un état de cellules réactivées (CR). Dans ces cellules, le noyau sphérique migre au centre de la cellule et 1 à 3 nucléoles sont visibles. Le cytoplasme se densifie et la vacuole se fragmente. Ces cellules sont alors au stade préprophasique de la mitose. Certaines CR peuvent entrer en division pour former des zones de divisions vacuolisées (ZV) ou denses (ZD) dont certaines vont aboutir à la formation d'embryons somatiques.

Les explants foliaires du génotype non embryogène C15 soumis aux mêmes conditions d'induction présentent des cellules en cours de réactivation pendant la durée du traitement avec un maximum au niveau du deuxième jour. Aucune autre figure cytologique n'a pu être observée dans ces tissus.

Les explants foliaires du génotype embryogène K59 présentent des cellules en cours de réactivation atteignant le même niveau (≈12% des cellules) que le génotype C15 au deuxième jour de culture. La fréquence des CER augmente jusqu'au quatrième jour de culture. Après le transfert dans un milieu sans glycérol, la proportion de cellules en cours de réactivation diminue et reste constante. Parallèlement, des cellules réactivées apparaissent aux 4^e et 5^e jour de culture. Des zones de divisions sont observées dès le 4^e jour de culture. La fréquence des zones de divisions vacuolisées est maximale à J5 tandis que celle des zones de divisions denses est maximale à J8. Ceci montre la réactivité du génotype K59 à ces conditions d'induction de l'embryogenèse somatique, comparativement au génotype C15 très peu réactif. En effet, des structures cellulaires capables de se diviser et de se différencier en embryons sont visibles uniquement au sein des tissus du génotype embryogène.

Il apparaît que les conditions de culture induisent une réaction des cellules foliaires du fait de la présence de cellules en cours de réactivation quel que soit le génotype. Cependant, la capacité à former de structures cellulaires aptes à se différencier en embryons somatiques n'a été observée que chez le génotype K59. Les processus qui permettent le passage d'un état de CER à celui de CR semblent être bloqués chez C15.

I-B.2. Variations des niveaux de transcrits

La variation du niveau de transcrits des gènes *CiSERK* et *CHI7507* a été mesurée à partir d'ARN extraits de fragments foliaires soumis à des conditions de culture permettant la formation d'embryons somatiques chez le génotype embryogène K59. Les ARN extraits de fragments foliaires du génotype non embryogène C15 soumis au même traitement ont été utilisés afin de déterminer si les variations des niveaux des transcrits sont spécifiques au processus d'embryogenèse somatique ou sont dues aux conditions de culture. Les prélèvements ont été réalisés avant mise en culture (J0), après 12h (J0.5), 24h (J1), 36h (J1.5), 2 jours (J2), 3 jours (J3) et 4 jours (J4) après mise en culture. Au 4e jour les explants cultivés sont transférés dans un milieu exempt de glycérol et sont prélevés 1 jour (J5) et 4 jours (J8) après le transfert.

I-B.2.1. Expression du gène de l'hémoglobine non symbiotique *CHI7507*

La variation du niveau de transcrits d'un gène codant une hémoglobine non symbiotique de chicorée (*CHI7507*) a été mesurée chez les génotypes K59 et C15 cultivés dans des conditions d'induction de l'embryogenèse somatique afin de valider la méthode de mesure de la variation du niveau de transcrits par qRT-PCR.

D'autre part il était important de confirmer l'induction de l'expression de ce gène au cours de l'embryogenèse somatique induite à partir d'un nouveau génotype embryogène



l'embryogenèse somatique (jours)

Figure 42: Variation du niveau de transcrits du gène *CHI7507* au cours des phases précoces de l'embryogenèse somatique.

A: Le gène de contrôle utilisé est l'actine. B: Le gène de contrôle utilisé est le 18s.

Les niveaux de transcrits indiqués au dessus des barres d'histogrammes sont exprimés en référence au niveau de transcrits mesuré à J0 chez le génotype non embryogène C15 (QR=1).

(K59 versus hybride '474'). En effet, l'expression de ce gène révèle par northern blot une augmentation transitoire de son niveau de transcrits avec un maximum au 3^e jour d'induction de l'embryogenèse somatique de la chicorée hybride '474'. Les transcrits de ce gène n'ont pas été détectés par northern blot chez un génotype non embryogène (Hendriks *et al.*, 1998).

En considérant le gène de contrôle a*ctine*, le niveau de transcrits de *CHI7507* chez le génotype C15 augmente régulièrement jusqu'au 4^e jour de culture pour atteindre 30 fois le niveau de transcrits mesuré à J0 puis diminue et reste constant (Figure 42-A). Chez le génotype K59, le niveau de transcrit augmente de façon plus marquée que chez le génotype C15, dès le premier jour de culture, et montre, au 2^e et 3^e jours de culture, une accumulation 90-100 fois supérieure à celle observée à l'état initial puis diminue de moitié au 4^e jour de culture et reste constant.

La même variation est observée en considérant le contrôle *18s* chez les deux génotypes à l'exception du fait que le pic d'accumulation de transcrits remarqué chez K59 ne s'observe qu'au 2^e jour de culture puis le niveau semble chuter plus régulièrement. Par ailleurs, l'intensité de la variation de QR est plus faible que celle observée lorsque l'*actine* est utilisé comme gène de contrôle (Figure 42-B).

L'augmentation des niveaux de transcrits observée chez le génotype K59 intervient dès le premier jour de culture, lorsque seules des cellules du parenchyme et des cellules en cours de réactivation sont présentes. Ce gène est donc probablement exprimé dans l'un ou l'autre de ces types cellulaires. L'expression du gène de cette hémoglobine non symbiotique semble être induite chez le génotype non embryogène C15 de façon régulière jusque J4. Dans ces tissus, seules des cellules parenchymateuses et des cellules en cours de réactivation sont aussi observées. Le fait que ce gène soit induit de façon plus intense chez le génotype K59 suggère que l'expression de ce gène est liée à l'embryogenèse somatique. Il paraît alors vraisemblable que l'expression du gène de cette hémoglobine soit liée à la réactivation cellulaire.

Malgré les différences de matériel végétal, de méthode de mesure de la variation des niveaux de transcrits, les observations de la variation de l'expression du gène de l'hémoglobine *CHI7507* sont similaires entre le génotype embryogène K59 et le clone '474'. Par ailleurs, la sensibilité de la technique de qRT-PCR a permis de montrer que l'expression du gène de l'hémoglobine *CHI7507* était aussi induite chez le génotype non embryogène C15.





A: Le gène de contrôle utilisé est l'actine. B: Le gène de contrôle utilisé est le 18s.

Les niveaux de transcrits sont exprimés en référence au niveau moyen mesuré à J0 chez le génotype non embryogène C15 (QR=1). Les nombres au dessus des barres d'histogrammes correspondent aux valeurs moyennes du niveau de transcrits relatifs calculées sur 2 ou 3 répétitions. Les barres d'erreur correspondent aux écart-types. Les lettres au dessus des barres d'histogrammes indiquent les classes de valeurs définies par un test de comparaison multiples de moyennes. Des lettres différentes indiquent une différence significative au seuil de 5% des valeurs moyennes.

I-B.2.2. Expression du gène CiSERK1

En considérant le contrôle *actine* (Figure 43-A), il apparaît que le niveau de transcrits de *CiSERK1* chute de 60 % et 50 % dès la première demi-journée de culture chez les génotypes K59 et C15 respectivement, puis ne varie plus au cours de l'embryogenèse somatique et ceci quel que soit le génotype.

Lorsque le gène *18s* est le contrôle (Figure 43-B), il apparaît de la même façon que le niveau de transcrit de *CiSERK1* chute de 60 % dès la première demi-journée de culture du génotype K59 tandis que chez le génotype C15, cette chute de niveau de transcrits n'est pas observée. Au-delà de la première journée de culture, le niveau transcrit semble diminuer régulièrement chez le génotype C15 pour atteindre 20 % du niveau de transcrits initial après 8 jours de culture. Le niveau de transcrits de *CiSERK1* chez K59 après 8 jours de culture représente aussi 20% du niveau de transcrits à J0.

Ainsi, les transcrits de *CiSERK1* sont détectés dans les feuilles de chicorée avant la mise en culture (J0) et semblent en quantité plus importante chez le génotype K59. Cette différence de niveau de transcrits dans les feuilles peut être la conséquence d'états physiologiques différents des deux génotypes.

Le fait que cette chute de niveau de transcrits de *CiSERK1* soit observée chez les deux génotypes suggère qu'elle est la conséquence d'événements communs précoces dus aux conditions de mise en culture. Les premiers événements qui interviennent dans les conditions de culture impliquent des stress dus aux blessures infligées aux explants foliaires, à l'hypoxie générée par la culture en milieu liquide, à la température de culture élevée ainsi qu'à l'absence de lumière.

I-B.2.3. Expression du gène CiSERK2

En considérant le contrôle *actine* (Figure 44-A), les niveaux de transcrits de *CiSERK2* chez K59 et C15 ne varient pas de façon significative pendant la durée de culture correspondant aux phases précoces de l'induction de l'embryogenèse somatique.

Lorsque le contrôle correspond au gène *18s* (Figure 44-B), le niveau de transcrit ne varie pas lors de la culture chez le génotype C15. Chez le génotype K59, le niveau initial de transcrit apparaît 2,5 fois plus élevé que chez le génotype C15 et une chute du niveau de transcrits est observée après 1,5 jours de culture pour rejoindre et conserver le niveau de transcrits mesuré chez le génotype C15.

Il résulte de ces expériences que si on considère les résultats par rapport à un même gène contrôle, l'accumulation des transcrits des deux gènes ne varie pas de la même façon au cours de l'induction de l'embryogenèse somatique. Lorsque le gène de contrôle est l'*actine*, le niveau de transcrits de *CiSERK1* chute dès 12 heures de cultures quel que soit le



<u>Figure 44</u>: Variation du niveau de transcrits du gène *CiSERK2* au cours des phases précoces de l'induction de l'embryogenèse somatique.

A: Le gène de contrôle utilisé est l'actine. B: Le gène de contrôle utilisé est le 18s. Les niveaux de transcrits sont exprimés en référence au niveau moyen mesuré à J0 chez le génotype non embryogène C15 (QR=1). Les nombres au dessus des barres d'histogrammes correspondent aux valeurs moyennes du niveau de transcrits relatifs calculées sur 2 ou 3 répétitions. Les barres d'erreur correspondent aux écart-types. Les lettres au dessus des barres d'histogrammes indiquent les classes de valeurs définies par un test de comparaison multiples de moyennes. Des lettres différentes indiquent une différence significative au seuil de 5% des valeurs moyennes.



Durée de culture en conditions d'induction de l'embryogenèse somatique (jours)

<u>Figure 45</u>: Variation du niveau de transcrits du gène 18s au cours de l'embryogenèse somatique en utilisant le gène actine comme contrôle. Les niveaux de transcrits sont exprimés en référence au niveau moyen mesuré dans les fragments foliaires du génotype C15 avant la mise en culture (J0)(QR=1). Les barres d'erreur correspondent aux écart-types. Les lettres au dessus des barres d'histogrammes indiquent les classes de valeurs définies par un test de comparaison multiples de moyennes. Des lettres différentes indiquent une différence significative au seuil de 5% des valeurs moyennes. génotype puis reste stable au cours de la culture tandis que le niveau de transcrits de *CiSERK2* ne varie pas au cours de la culture.

Lorsque le gène de contrôle est le *18s*, le niveau de transcrits de *CiSERK1* chute dès 12 heures de cultures puis diminue plus lentement pendant la culture chez K59, tandis que chez le génotype C15, une chute des niveaux de transcrits régulière est observée après la première journée de culture. Les niveaux de transcrits de CiSERK2 chutent après 1 jour de culture puis restent stables chez K59 alors qu'ils ne varient pas chez C15.

Le fait que les variations de niveaux de transcrits des deux gènes *CiSERK* au cours de l'embryogenèse somatique présentent des différences, selon le gène contrôle utilisé, peut être expliqué par le fait que les niveaux de transcrits des gènes contrôles ne sont pas stables au cours du processus. Des informations récentes, provenant d'expériences d'hybridations sur puces à ADN réalisées au laboratoire, indiquent que l'expression du gène d'*actine* est stable au cours de l'induction de l'embryogenèse somatique (Legrand, communication personnelle). Ainsi, les variations des niveaux de transcrits du gène *18s* normalisés au niveau de transcrits du gène d'actine (Figure 45) peuvent expliquer les variations de profils d'accumulation des transcrits des cibles en fonction du gène contrôle utilisé. Cependant, du fait que les variations du niveau de transcrits des gènes contrôles sont de faibles amplitudes, et statistiquement non significatives, les variations de forte intensité des gènes cibles étudiés (comme par exemple la variation de *CHI7507*) sont faiblement sensibles à ces variations.

Néanmoins, bien qu'une chute du niveau de transcrits de *CiSERK1* a été observée au cours de la première demi journée, ces résultats établissent que les deux gènes *CiSERK* ne montrent pas d'accumulation de leur transcrits au cours des phases précoces de l'embryogenèse somatique.

I-C. EXPRESSION DES GENES CISERK1 ET CISERK2 AU COURS DE L'EMBRYOGENESE ZYGOTIQUE ET DU DEVELOPPEMENT DE LA GRAINE

Des capitules de chicorée ont été récoltés avant la pollinisation (J0), 1 jour (J1ap) après pollinisation ainsi que tous les deux jours jusque 19 jours après pollinisation (J3ap à J19ap) et les akènes ont été collectés. C'est à partir des mêmes lots d'akènes que la caractérisation cytologique du développement embryonnaire et l'étude de la variation du niveau de transcrits des gènes *CiSERK1* et *CiSERK2* ont été réalisées.



.

I-C.1. Caractérisation cytologique

Le développement de l'embryon zygotique et de la graine a été caractérisé par différents stades (Figure 46) (Da Silva, 2004; Blervacq, communication personnelle).

Dès 18 heures après la pollinisation, on observe le premier stade zygote, où la première cellule embryonnaire est visible ainsi que la première cellule suspensoriale. L'albumen n'est pas cellularisé et le noyau 3n est toujours présent.

Entre 24 et 48 heures après la pollinisation, on assiste à la cellularisation de l'albumen et on observe une phase de pro-embryogenèse caractérisée par une segmentation du zygote qui recouvre des stades d'embryons à 4 puis 8 cellules. Au niveau de la partie basale, des recloisonnements transversaux conduisent à la formation d'un suspenseur.

Deux à 3 jours après pollinisation, on observe un embryon globulaire pour lequel le protoderme est mis en place. L'albumen est toujours cellularisé (Figure 46-B).

Entre 4 et 5 jours après la pollinisation, l'embryon est au stade cordiforme (Figure 46-C). Le passage du stade globulaire au stade cordiforme signifie pour l'embryon le changement de symétrie (passage d'une symétrie axiale à une symétrie bilatérale). Ceci est concomitant avec l'initiation de deux primordium cotylédonaires. Le territoire précaulinaire se mettra en place entre ces deux futurs cotylédons. L'albumen cellularisé est en cours de désagrégation.

Le stade torpille est atteint vers le 5 à 6^e jour après la pollinisation (Figure 46-D). L'embryon subit une rotation sur lui même. Trois ou quatre assises cellulaires forment la coiffe. Les procambiums cotylédonaires et hypocotylaires sont mis en place.

A partir du 7^e jour après pollinisation, l'embryon est au stade cotylédonaire (Figure 46-E). Deux états différents peuvent être distingués. Le stade cotylédonaire I, où l'apex des cotylédons n'atteint pas le sommet de l'akène. A ce stade, la mise en place des méristèmes apicaux et racinaires est visible dans ces embryons. Des zones de procambium cotylédonaire se différencient. Ces zones fusionnent par la suite et forment un cordon vasculaire continu avec le procambium hypocotylaire.

Le stade cotylédonaire II est atteint entre 11 et 13 jours après la pollinisation. Les cotylédons atteignent alors le sommet de l'akène et différents territoires sont nettement visibles : coiffe, gemmule, procambium. L'embryon se développe grâce à des divisions anticlines. L'albumen a quasiment disparu. Le contenu des cellules composant les cotylédons présente des structures de types vésiculaires pouvant indiquer la formation de réserves.

D: Embryon au stade torpille, E: Embryon au stade cotylédonnaire I.

Figure 46: (ci-contre): Stades de développement des embryons zygotiques chez la chicorée.

I: Chronologie des stades de développement des embryons chez la chicorée. L'échelle de temps correspond au nombre de jour après la pollinisation. J0 est le jour de l'anthèse, avant la pollinisation. JAP: jours après la pollinisation. Les traits verticaux représentent l'intervalle de temps dans lequel l'embryon est au stade de développement indiqué.

II: Figures cytologiques des embryons aux différents stades caractéristiques du développement et annotation des principales structures.

A: Sac embryonnaire avant la pollinisation, B: Embryon globulaire, C: Embryon au stade cordiforme,



3 .

Figure 47: Variation du niveau de transcrits du gène *CiSERK1* au cours de l'embryogenèse zygotique et du développement de la graine.

A: Le gène de contrôle utilisé est l'actine. B: Le gène de contrôle utilisé est le 18s.

Les niveaux de transcrits sont exprimés en référence au niveau moyen mesuré à J0 (QR=1). Les nombres au dessus des barres d'histogrammes correspondent aux valeurs moyennes du niveau de transcrits relatifs calculées sur 3 répétitions. Les barres d'erreur correspondent aux écart-types. Les lettres au dessus des barres d'histogrammes indiquent les classes de valeurs définies par un test de comparaison multiples de moyennes. Des lettres différentes indiquent une différence significative au seuil de 5%.

I-C.1.1. Expression du gène CiSERK1

Lorsque le contrôle est le gène d'*actine*, le niveau de transcrit du gène *CiSERK1* au cours de l'embryogenèse zygotique et du développement de la graine augmente d'environ 3 fois à partir de 7 Jap (Figure 47-A). Ce niveau de transcrit reste au moins deux fois plus élevé jusqu'à J15ap puis un pic est observé 17 jours après la pollinisation (J17ap). Les transcrits retrouvent à J19ap le niveau observé à J15ap.

L'augmentation des niveaux de transcrits à partir de J7ap, observée en utilisant le contrôle *actine*, n'est pas détectée de façon significative avec le contrôle *18s* mais le pic observé 17 jours après la pollinisation est confirmé avec une intensité plus faible (Figure 47-B).

L'augmentation des transcrits de *CiSERK1* au cours de l'embryogenèse zygotique et du développement de la graine à partir de J7ap et le pic observé à J17ap peuvent être mis en relation avec les événements de développement de l'embryon qui atteint un stade cotylédonaire. A ce stade, des différenciations de tissus interviennent (procambium, méristèmes) ainsi que la formation de réserves cotylédonaires. Par ailleurs, le pic de niveau d'expression de *CiSERK1* à J17ap peut aussi être la conséquence d'événements de régulation du métabolisme et de la physiologie qui n'engendrent pas de modifications dans la morphologie tissulaire et cellulaire de l'embryon.

I-C.1.2. Expression du gène *CiSERK2*

En considérant le contrôle *actine*, le niveau de transcrits de *CiSERK2* ne varie pas de façon significative (Figure 48-A) au cours de l'embryogenèse zygotique et du développement de la graine. Le niveau de transcrits diminue régulièrement au cours du développement de la graine en considérant le gène de contrôle *18s* (Figure 48-B).

De la même manière que précédemment, les profils d'accumulations des transcrits de *CiSERK1* et *CiSERK2* présentent des fluctuations selon le gène contrôle utilisé. Une augmentation du niveau de transcrits de *CiSERK1* est observée après J7ap lorsque le gène de contrôle utilisé est l'*actine* mais n'est pas observée lorsque le contrôle est le gène *18s*. Le niveau de transcrits de *CiSERK2* est stable lors de l'utilisation du gène d'*actine* comme contrôle alors qu'il chute régulièrement lorsque le contrôle est le gène *18s*. Ceci peut être la conséquence de la variation du niveau de transcrits des gènes contrôles au cours de l'embryogenèse zygotique. La figure 49 présente la variation du niveau de transcrits correspondant au gène *18s* normalisé à celui de l'actine. Il apparaît que le niveau de transcrits des gènes *18s* augmente significativement au cours du développement de la graine. Ceci peut être compatible avec une augmentation de l'activité traductionnelle dans



Figure 48: Variation du niveau de transcrits du gène *CiSERK2* au cours de l'embryogenèse zygotique et du développement de la graine.

A: Le gène de contrôle utilisé est l'actine. B: Le gène de contrôle utilisé est le 18s.

Les niveaux de transcrits sont exprimés en référence au niveau moyen mesuré à J0 (QR=1). Les nombres au dessus des barres d'histogrammes correspondent aux valeurs moyennes du niveau de transcrits relatifs calculées sur 3 répétitions. Les barres d'erreur correspondent aux écart-types. Les lettres au dessus des barres d'histogrammes indiquent les classes de valeurs définies par un test de comparaison multiples de moyennes. Des lettres différentes indiquent une différence significative au seuil de 5% des valeurs moyennes.



<u>Figure 49</u>: Variation du niveau de transcrits du gène 18s au cours de l'embryogenèse somatique et du développement de la graine en utilisant le gène actine comme contrôle. Les niveaux de transcrits sont exprimés en référence au niveau moyen dans l'ovaire mesuré à J0 (QR=1). Les barres d'erreur correspondent aux écart-types. Les lettres au dessus des barres d'histogrammes indiquent les classes de valeurs définies par un test de comparaison multiples de moyennes. Des lettres différentes indiquent une différence significative au seuil de 5% des valeurs moyennes.

ces tissus en raison d'une synthèse abondante de protéines. Ainsi, la stabilité du niveau de transcrits de *CiSERK1* observé à J7ap en utilisant le contrôle *18s* ainsi que la chute du niveau de transcrits de *CiSERK2* serait la conséquence de l'augmentation du niveau de transcrits des gènes *18s*.

I-D. NIVEAUX DE TRANSCRITS DANS DIFFERENTS ORGANES

Le niveau de transcrits des gènes *CiSERK1* et *CiSERK2* a été mesuré dans différents organes afin de déterminer s'il existe un profil d'expression spécifique de ces gènes. Les résultats ont été exploités en utilisant le gène de l'*actine* comme contrôle à partir d'une seule transcription inverse d'un lot unique de plantules (pour chacune des conditions de mise en germination) et d'un lot unique de boutons floraux. Les niveaux de transcrits qui concernent les fragments foliaires, induits pour l'embryogenèse somatique, et l'embryogenèse zygotique correspondent aux valeurs moyennes des expériences précédentes.

I-D.1. Variation des niveaux de transcrits lors de la germination.

Au cours de l'embryogenèse somatique, les niveaux de transcrits de *CiSERK1* montrent une chute dès le premier jour de culture. Une hypothèse émise était que cette chute était induite par les stress provoqués par les conditions d'induction comme la blessure, l'hypoxie, la température ou l'obscurité. Afin de déterminer si l'absence de lumière peut induire la chute des niveaux de transcrits, des graines ont germé 4 et 8 jours en absence ou en présence de lumière et nous avons évalué le niveau de transcrits dans ces plantules (Figure 50). On observe que le niveau de transcrits de *CiSERK1* est réduit de moitié dans les plantules qui ont germé 4 jours à la lumière par rapport aux plantules qui ont germé à l'obscurité. Cette différence de niveau de transcrits ne s'observe plus dans les plantules âgées de 8 jours germées à la lumière ou à l'obscurité. L'obscurité semble avoir pour effet une induction plus rapide de l'expression du gène *CiSERK1*. Le niveau de transcrits des gènes que l'absence de lumière seule induise une chute des niveaux de transcrits des gènes *CiSERK*. Cette chute du niveau de transcrits de *CiSERK1* dès le premier jour de culture en conditions embryogènes peut alors dépendre d'autres stress.



Figure 50: Niveaux relatifs des transcrits *CiSERK1* et *CiSERK2* normalisés avec le gène d'a*ctine* dans des plantules germées 4 jours à l'obscurité et à la lumière, et 8 jours à l'obscurité et à la lumière en référence au niveau mesuré dans les plantules germées 4 jours à l'obscurité (QR=1).

1.8





Figure 51: Niveaux relatifs des transcrits *CiSERK1* et *CiSERK2* normalisés avec le gène d'a*ctine* au sein d'un même tissu (**A**) et entre différents organes de la chicorée (**B**). Les niveaux relatifs de transcrits sont exprimés en référence au niveau de transcrits de *CiSERK1*(QR=1) (**A**) ou à celui déterminé dans les ovaires avant la pollinisation (EZ J0) (QR=1) (**B**). BF: bouton floraux (8-10 jours avant l'anthèse); 17JAP: akènes 17 jours après pollinisation ; Germ8jlum: plantules germées 8 jours à la lumière ; F-J0: feuille de génotype K59 avant mise en culture; F-J8: feuille de génotype K59 après 8 jours de culture en conditions d'embryogenèse somatique. Les nombres au dessus des barres d'histogrammes correspondent aux valeurs du niveau de transcrits relatifs à celui de la référence.

I-D.2. Variation des niveaux de transcrits entre différents organes

Les niveaux relatifs de transcrits des gènes *CiSERK1* et *CiSERK2* ont été comparés dans différents organes (Figure 51-A). Il apparaît que le niveau de transcrits de *CiSERK2* est toujours plus important que celui de *CiSERK1* quel que soit l'organe considéré. Le niveau de transcrits de *CiSERK2* peut être 7 fois supérieur à celui observé pour *CiSERK1* dans les feuilles avant la mise en culture (J0), et cette différence peut être supérieure à 140 fois dans les plantules en germination.

Les niveaux de transcrits observés dans différents organes ont été comparés, le niveau de transcrits dans les ovaires avant pollinisation (EZ J0) étant utilisé comme référence (Figure 51-B). Il apparaît que les profils d'expression de *CiSERK1* et *CiSERK2* sont divergents. Les niveaux de transcrits de *CiSERK2* dans les feuilles avant leur mise en culture (J0), après 8 jours de culture et dans les plantules germées 8 jours à la lumière sont semblables et environ deux fois plus abondants que dans les ovaires avant la pollinisation et dans la graine 17 jours après la pollinisation. Le niveau de transcrits dans les boutons floraux 8-10 jours avant l'anthèse est nettement plus faible.

Les transcrits de *CiSERK1* sont les plus abondants dans les feuilles avant leur mise en culture (J0) et sont 4 fois plus abondants que dans les feuilles cultivées pendant 8 jours. Les niveaux de transcrits de *CiSERK1* dans les ovaires avant la pollinisation et dans les plantules germées 8 jours à la lumière sont semblables et 15 fois moins abondants que dans les feuilles à J0. Dix-sept jours après la pollinisation, les transcrits de *CiSERK1* sont 7 fois plus abondants que dans l'ovaire avant pollinisation. Le niveau de transcrits dans les boutons floraux 8-10 jours avant l'anthèse est le plus faible.

La variation de niveau de transcrits de *CiSERK2* observée entre les organes n'est jamais supérieure à 3 fois (à l'exception des boutons floraux). Le niveau de transcrits de *CiSERK2* reste assez stable dans les fragments foliaires cultivés en condition d'induction de l'embryogenèse somatique ou dans les akènes en développement. Le niveau de transcrits de *CiSERK1* est toujours plus faible que celui de *CiSERK2* et montre un niveau de variation plus important, tant entre différents organes, qu'au sein d'un organe en développement (akènes) ou soumis à des conditions de cultures embryogènes (fragments foliaires). L'amplitude de la variation peut aller jusqu'à 15 fois le niveau observé dans les ovaires (exception faite des boutons floraux). L'expression du gène *CiSERK1* semble d'avantage régulée que celle du gène *CiSERK2*.

II VARIATION DE L'ACCUMULATION PROTEIQUE

Parallèlement à l'étude du niveau de transcrits des gènes *CiSERK1* et *2* au cours de l'induction de l'embryogenèse somatique, de l'embryogenèse zygotique et du développement de la graine, l'étude des variations de l'accumulation protéique a été entreprise par la technique de western blot.

Afin de suivre l'accumulation des protéines CiSERK, nous avons souhaité utiliser des anticorps dirigés contre le domaine kinase de CiSERK et d'autres contre le domaine extracellulaire de CiSERK. L'anticorps dirigé contre le domaine kinase aurait une moindre spécificité vis à vis de CiSERK du fait de la forte conservation de ce domaine au sein des protéines du groupe SERK tandis que la spécificité de reconnaissance serait d'avantage obtenue au moyen d'un anticorps dirigé contre le domaine extracellulaire qui est lui moins conservé.

Afin d'obtenir ces deux types d'anticorps, nous avons entrepris de produire et de purifier deux protéines recombinantes partielles correspondant, l'une au domaine extracellulaire, et l'autre au domaine kinase de CiSERK1.

II-A. OBTENTION ET PURIFICATION D'ANTICORPS DIRIGES CONTRE LE DOMAINE KINASE DE CISERK

II-A.1. Production et purification de protéines recombinantes

Trois protéines recombinantes ont été produites. Elles correspondent respectivement à la protéine dans sa totalité, au domaine extracellulaire et au domaine kinase. La protéine recombinante correspondant à la protéine complète sera utilisée comme témoin de spécificité de chacun des anticorps.

II-A.1.1. Clonage des séquences codantes dans le vecteur de production pET-16b

Le site de clonage du vecteur pET-16b est constitué de 3 sites de restrictions Ndel, Xhol et BamH1.

Des amorces (Tableau 14) ont été définies pour chacune des protéines à produire à partir de la séquence codante de l'ADNc *CiSERK1₄₇₄* puis modifiées en 5' par l'ajout d'un site de restriction Ndel dans l'amorce directe et d'un site Xhol dans l'amorce complémentaire. Les fragments amplifiés et digérés par ces deux enzymes de restriction ont été clonés en phase avec la séquence codant une extension His-tag (cf matériel et méthodes § XI-A). Tandis que l'amorce complémentaire nécessaire à l'amplification des



<u>Figure 52</u>: Spécificité du sérum dirigé contre le domaine kinase de CiSERK1.

A : gel électrophorétique (SDS-PAGE 12,5 %) de protéines totales de bactéries produisant le domaine extracellulaire (i), le domaine kinase (ii) et la protéine complète de CiSERK1 (iii) (coloration au bleu de Coomassie). Les flèches indiquent les différentes protéines recombinantes produites.

B Immunodétection des protéines correspondant au gel A (extraits dilués 100 x) par le sérum dirigé contre le domaine kinase de CiSERK1. Le western blot est révélé par réaction colorée.



<u>Figure 53</u>: purification de la protéine recombinante correspondant au domaine kinase.

Les protéines ont été séparées par SDS-PAGE à 12,5%, et colorées au bleu de Coomassie.

a: extrait protéique bactérien total, b: protéines bactériennes non retenues sur la colonne, c fraction d'élution contenant la protéine recombinante purifiée.



Figure 54: Test d'autophosphorylation du domaine kinase recombinant purifié de CiSERK1.

A: gel d'électrophorèse de la protéine purifiée (SDS-PAGE 12.5 %, coloration au bleu de Coomassie).

B: Autoradiographie correspondant au gel A.

séquences pleines longueurs et du domaine kinase possède le codon stop intrinsèque à la séquence, un codon stop a été introduit par l'ajout d'une base dans l'amorce complémentaire permettant l'amplification du domaine extracellulaire.

Domaine à produire	Amorce directe	Amorce complémentaire
Protéine complète	TTTTTT CATATG GATCGAACAATCTCTCT	TTTTT <u>CTCGAG</u> CAA <u>TCA</u> TCTAGGACCCGACAAT
Domaine extracellulaire	TTTTTT CATATG GATCGAACAATCTCTCT	TTTTT <u>CTCGAG</u> GCACT A GTTGCTGACTGAAGG
Domaine kinase	TTTTTT CATATG GCCGAAGAGGATCCAGAAGT	TTTTT <u>CTCGAG</u> CAA <mark>TCA</mark> TCTAGGACCCGACAAT

<u>Tableau 14</u> Séquences des amorces utilisées afin d'amplifier les domaines indiqués dans le but de permettre le clonage dans le vecteur pET-16b et la production des protéines. En gras sur les séquences les sites de restriction Ndel; les sites de restriction Xhol sont soulignés. Les bases encadrées correspondent au codon stop. En gras et italique, la base ajoutée pour introduire un codon stop.

II-A.1.2. Production et purification des protéines recombinantes

Le séquençage de la région codant la protéine de fusion a été réalisé afin de vérifier que la séquence codante d'intérêt se situait bien en phase avec la séquence codant l'extension histidine.

La taille calculée des protéines recombinantes est de 70.1 kDa pour CiSERK1 complet, 26.4 kDa pour le domaine extracellulaire et 42.2 kDa pour le domaine intracellulaire. Les trois protéines recombinantes ont été produites dans un système bactérien (*E.coli* BL21 (DE3)). On observe la présence de bandes majeures de la taille attendue dans chacun des extraits protéiques bactériens correspondant aux protéines recombinantes produites (Figure 52-A).

Les protéines recombinantes correspondant au domaine extracellulaire et à la protéine entière n'ont pas pu être purifiées sur colonne d'affinité car il n'a pas été possible de les solubiliser, malgré plusieurs essais dans des conditions de solubilisation différentes. Ainsi, seul le domaine kinase a été purifié.

Le domaine kinase a été purifié par chromatographie sur colonne de sépharose préalablement chargée avec du sulfate de nickel. L'élution de la protéine recombinante a été réalisée par l'injection de concentration croissante d'imidazole. La protéine recombinante a été éluée avec 300 mM d'imidazole (Figure 53). Le rendement de purification a été d'environ 8 mg de protéine recombinante purifiée par litre de culture bactérienne.



Figure 55: Purification par chromatographie d'affinité des anticorps dirigés contre le domaine kinase de CiSERK1. Les anticorps ont été élués par 50 mM d'acide acétique. L'absorbance à 280 nm a été mesurée dans chacune des fractions d'élution des anticorps. Les fraction 2 à 6, les plus concentrées en anticorps ont été regroupées et utilisées dans les expériences de western blot.

II-A.1.3. Fonctionnalité du domaine kinase

L'activité kinase d'un domaine intracellulaire de récepteur implique qu'il soit capable de transférer un groupement phosphate d'une molécule d'ATP sur une autre protéine ou sur lui-même. Un test d'autophosphorylation de la protéine recombinante purifiée, correspondant au domaine kinase, a été réalisé afin de vérifier son activité catalytique. La protéine recombinante a donc été incubée en présence d' γ [³²P]-ATP puis déposée en gel d'acrylamide dénaturant. Après transfert sur une membrane de PVDF, un film autoradiographie a été exposé (Figure 54). Un signal autoradiographique intense est observé au niveau de la bande protéique. Le domaine kinase produit est donc capable de transférer un groupement phosphate de façon intramoléculaire.

II-A.2. Obtention et purification d'anticorps dirigés contre le domaine kinase de CiSERK

Un test du sérum a été réalisé sur la protéine recombinante purifiée. Tandis que le western blot correspondant au sérum pré-immun ne présente aucun signal (non montré), la membrane incubée en présence du sérum immun révèle bien la protéine recombinante. Afin de tester la spécificité du sérum pour le domaine kinase de CiSERK, un western blot d'extraits protéiques de bactéries induites pour la production des trois protéines recombinantes a été réalisé (Figure 52). L'immunoblot montre que le sérum reconnaît de façon spécifique le domaine kinase à partir duquel il a été produit, la protéine complète dans laquelle le domaine kinase est présent et ne reconnaît pas le domaine extracellulaire. L'aspécificité pour des protéines bactériennes est très faible. D'autre part, ces trois constructions ont en commun le fait qu'elles possèdent une extension His-tag. Il semble que cette extension ne soit pas reconnue par le sérum du fait qu'aucun signal n'est détecté dans les extraits protéiques des bactéries produisant le domaine extracellulaire.

Tandis qu'aucun signal n'a été observé par western blot des protéines bactériennes produisant la protéine recombinante au moyen du sérum pré-immun, le western blot à partir de protéines de chicorée a présenté des signaux (non montré). Cette réactivité du sérum pré-immun vis-à-vis des protéines de chicorée peut être la conséquence de l'alimentation fournie aux lapins. Afin d'éliminer les signaux d'immunodétection imputés au sérum pré-immun, les anticorps ont été purifiés par chromatographie d'affinité à l'aide d'une colonne sur laquelle le domaine kinase purifié a été couplé (Figure 55).





<u>Figure 56</u>: Profils électrophorétique (SDS-PAGE 12,.5 %) des protéines (5µg) extraites des fragments foliaires de K59 (A) et C15 (B) cultivées en condition d'induction de l'embryogenèse somatique. 0, 0.5, 1, 1.5, 2, 3, 4, 5, 8 sont les jours de cultures dans le milieu, M: marqueur de masse moléculaire. Coloration au nitrate d'argent.



<u>Figure 57</u>: Immunodétection des protéines lors de la culture des explant foliaires de K59 (a) et C15 (b) en conditions d'induction de l'embryogenèse somatique. 0, 0.5, 1, 1.5, 2, 3, 4, 5, 8 sont les jours de cultures dans le milieu.

II-B. ACCUMULATION PROTEIQUE AU COURS DES PHASES PRECOCES DE L'EMBRYOGENESE SOMATIQUE

Les western blots de protéines extraites de fragments foliaires cultivés en conditions embryogènes ont été réalisés au moins deux fois. Afin de s'assurer que les signaux sont spécifiques, des contrôles ont été réalisés en absence d'anticorps primaire ou d'anticorps secondaire.

L'électrophorèse des protéines suivie d'une coloration au nitrate d'argent a été réalisée afin de visualiser les modifications dans le profil protéique au cours des phases précoces de l'embryogenèse somatique chez les deux génotypes K59 et C15 (Figure 56). On constate que le profil protéique monodimensionnel varie peu au cours de l'induction de l'embryogenèse somatique. Ainsi, les modifications protéiques qui peuvent être observées au cours des phases précoces de l'induction de l'embryogenèse somatique impliquent vraisemblablement des protéines faiblement présentes.

Les profils d'immunodétection obtenus à partir d'extraits protéiques de fragments foliaires de génotype K59 et C15 cultivés en conditions d'induction de l'embryogenèse somatique sont présentés figure 57. Plusieurs bandes protéiques sont détectées sans doute du fait de la présence d'épitopes communs avec le domaine kinase.

Cependant au moins 3 signaux ont une taille compatible avec celle d'un récepteur de type SERK.

Une bande protéique de 83 kDa est présente dans les feuilles avant la mise en culture (J0) à la fois chez K59 et C15. L'intensité du signal correspondant est plus forte chez le génotype K59. Cette différence d'intensité entre les deux génotypes s'estompe dès le premier jour de culture. L'intensité de cette bande reste ensuite stable sur la durée de la culture. La masse moléculaire de cette bande est plus grande que la taille calculée des protéines CiSERK (environ 68 kDa et 65 kDa lorsque le peptide signal est clivé) mais plusieurs sites de N-glycosylation ont été prédits *in silico* et il est ainsi possible que les protéines CiSERK soient glycosylées et leur masse moléculaire s'en retrouve alors augmentée.

Une protéine de masse moléculaire (MM) de 75 kDa est détectée après 2 jours de culture chez les deux génotypes. Cependant tandis que l'intensité de la bande protéique reste constante chez C15, elle augmente chez le génotype K59. La présence de cette protéine chez les deux génotypes témoigne de son implication dans une réaction commune aux conditions de culture, cependant, le fait que l'intensité de cette bande augmente chez K59 suggère qu'elle puisse être liée à la capacité embryogène de K59. On observe des



<u>Figure 58</u>: Profil électrophorétique (SDS-PAGE 12,5 %) des protéines (5 μ g) des graines en développement issues du croisement de K28 x K59. 0,1,3,5,7,11,13,15,17,19 sont les jours après pollinisation. Coloration au nitrate d'argent.



Figure 59: Immunodétection des protéines au cours du développement des embryons et des graines issus du croisement K28xK59.

0: ovaire avant pollinisation, 1, 3, 5, 7, 11, 13, 15, 17, 19 jours après pollinisation.

cellules en cours de réactivation dans les mêmes proportions chez les deux génotypes pendant les deux premiers jours de culture, cependant, tandis que cette proportion diminue chez C15, elle reste élevée chez K59 (cf Figure 41). Le profil d'accumulation de cette bande de 75 kDa peut être ainsi lié à la présence de cellules en cours de réactivation au sein des tissus.

Une bande protéique de 62 kDa est détectée après 3 jours de culture chez le génotype K59. Son intensité croît jusqu'au 5^e jour de culture puis est stable jusqu'à 8^e jour de culture. Cette bande n'est pas détectée chez le génotype C15. Au niveau cytologique, après 3 jours d'induction quelques cellules réactivées sont observées mais elles ne sont significativement dénombrées qu'à J4 chez le génotype K59. Les cellules réactivées ne sont pas visibles chez le génotype C15 (cf Figure 41). Cette protéine semble donc être exprimée au moment où certaines cellules en cours de réactivation vont atteindre le stade de cellules réactivées chez le génotype K59.

II-C. ACCUMULATION PROTEIQUE AU COURS DE L'EMBRYOGENESE ZYGOTIQUE ET DU DEVELOPPEMENT DE LA GRAINE

II-C.1. Variation du profil protéique au cours de l'embryogenèse zygotique et du développement de la graine

L'électrophorèse des protéines extraites des akènes en développement a été réalisée afin d'observer les variations du profil protéique au cours de l'embryogenèse zygotique et du développement de la graine (Figure 58).

Le profil protéique subit des remaniements importants au cours du développement de la graine (Figure 58). Deux phases principales peuvent être distinguées sur ce profil protéique, la première s'étend jusqu'à 7 jours après pollinisation, la seconde se prolonge de J11ap jusqu'à 19 jours après pollinisation. Ceci peut être corrélé au développement embryonnaire qui atteint le stade cotylédonaire vers 7 jours après pollinisation. Au-delà, des protéines de faibles poids moléculaires (environ 20 kDa) apparaissent. Cette période correspond à la différenciation de tissus et à l'apparition, dans les cotylédons, de vésicules intracellulaires qui peuvent impliquées dans l'accumulation des réserves cotylédonaires. Chez le tournesol, des oléosines (de MM de 15 à 20 kDa) sont présentes dans la graine et sont impliquées dans la formation des corps gras de la graine (Cummins et Murphy, 1992). Ceci tend à corroborer la mise en place de tissus de réserves environ 11 jours après la pollinisation chez la chicorée.

Par ailleurs, on observe l'apparition transitoire de protéines de masses moléculaires différentes (25 kDa, 38 kDa, 80 kDa), sur l'ensemble des stades de développement qui témoignent de profondes modifications cellulaires et métaboliques au sein de la graine.

II-C.2. Immunodétection des protéines de type SERK au cours de l'embryogenèse zygotique et du développement de la graine

Au moins deux répétitions de western blot des extraits des akènes en développement ont été réalisées. Le profil d'immunodétection obtenu à partir d'extrait protéiques d'akènes avant pollinisation (J0) et à différents jours après pollinisation (J1 à J19) est présenté figure 59.

Quatre bandes protéiques sont détectées. Deux d'entre elles n'ont pas des masses moléculaires compatibles avec celles de récepteurs transmembranaires de type SERK. Une bande protéique de forte intensité de 35 kDa est présente 11 jours après pollinisation et son intensité est stable jusque 19 jours après pollinisation et peut correspondre à la bande majeure de cette masse moléculaire visible sur le gel d'électrophorèse coloré à l'argent (Figure 58). Un signal de 40 kDa est présent dans les ovaires avant pollinisation et son intensité reste constante au cours du développement de l'embryon et de la graine.

Une bande protéique de 62 kDa, compatible avec la taille d'un récepteur transmembranaire, est présente dans l'ovaire avant pollinisation. Son intensité semble augmenter jusqu'au 11^e jour après pollinisation puis se stabilise jusque 17 jours après pollinisation. Cette protéine n'est pas détectée à J19. La présence de cette bande à J15 a été observée sur d'autres immunoblots réalisés. Cette bande protéique a la même masse moléculaire que celle qui est détectée dans les fragments foliaires après 3 jours de culture au cours de l'embryogenèse somatique.

Une bande de 55 kDa est présente 11 jours après pollinisation et son intensité reste stable jusque 19 jours après pollinisation. Son apparition est concomitante à la bande protéique détectée à une taille de 35 kDa et aux bandes de 20 kDa observées sur gel d'acrylamide. Ces protéines peuvent être alors mise en relation avec la formation des tissus de réserve dans la graine.

III DISCUSSION

A la suite du clonage de deux gènes de type *SERK*, *CiSERK1* et *2* chez la chicorée, nous avons vérifié si ces gènes pouvaient être les équivalents fonctionnels de *DcSERK* initialement mis en évidence chez la carotte. *DcSERK* est exprimé dans les cellules compétentes à l'embryogenèse et dans les embryons somatiques jusqu'au stade globulaire de leur développement et au cours de l'embryogenèse zygotique et du développement de la

graine. La caractérisation structurale des gènes *CiSERK1* et *2* et des protéines déduites a mis en évidence leur forte identité avec d'autres membres du groupe *SERK*. C'est pourquoi nous avons entrepris d'évaluer la variation des niveaux de transcrits des gènes *CiSERK1* et *CiSERK2* au cours des phases précoces de l'embryogenèse somatique et de l'embryogenèse zygotique de la chicorée.

Au regard du faible niveau de transcrits de *DcSERK* décrit chez la carotte, nous avons opté pour la technique de RT-PCR quantitative en temps réel qui permet d'allier la haute spécificité à la forte sensibilité de la PCR pour la quantification relative des niveaux de transcrits des gènes *CiSERK*.

Pour réaliser les expériences, du fait que cette technique était nouvelle au laboratoire, nous avons privilégié la répétition des expériences sur des échantillons, issus d'une extraction d'ARN, à la multiplication des répétitions biologiques. Les échantillons, dont ont été extraits les ARN et les protéines, ont été cytologiquement caractérisés. Ces observations ont clairement montré que le génotype K59 était réactif et embryogène tandis que le génotype C15, bien que faiblement réactif, était incapable de développer des embryons somatiques. Ceci a permis, en outre, de connaître l'évolution des différentes figures cytologiques observées dans les tissus pour chacun des jours de prélèvement effectué. De la même façon, les observations des embryons en développement dans les akènes ont été effectuées sur les mêmes lots que ceux à partir desquels les analyses d'expression des gènes *CiSERK* ont été réalisées. Il était alors possible de faire un lien entre les événements cytologiques observés au niveau des tissus et les profils d'accumulations de transcrits et de protéines.

La technique de PCR en temps réel implique que l'on dispose de gènes contrôles qui se veulent stables entre différents organes et au sein d'un organe soumis à différentes conditions expérimentales (Bustin, 2000). Aucun gène de ce type n'était connu et disponible chez la chicorée. Aussi notre choix s'est porté d'une part sur un gène d'*actine* cloné au laboratoire et fortement identique aux gènes d'*actine 2* et *8* d'*Arabidopsis* connus pour leur stabilité d'expression c'est-à-dire dont le niveau de transcrits varie faiblement entre organes et dans différentes conditions (An *et al.*, 1996; Charrier *et al.*, 2002; Wolf *et al.*, 2003). D'autre part, nous avons utilisé les gènes codant les ARNr *18s* classiquement utilisés comme référence dans les expériences de northern blot. L'utilisation du contrôle *18s* a été rendue possible par l'étape de transcription inverse initiée par un mélange d'amorces oligo(dT) et d'hexamères de séquences aléatoires. Le fait que les gènes codants les ARNr *18s* soient très conservés entre espèce nous a dispensé de cloner la séquence correspondante chez la chicorée.

Les profils d'expression des gènes *CiSERK1* et *CiSERK2* montrent des variations dans l'accumulation des transcrits parfois différents selon le gène contrôle utilisé ce qui suggère la possible variation d'expression d'au moins un des deux gènes contrôle au cours des processus de développement étudiés. Des données récentes issues d'expériences d'hybridation sur puces à ADN semblent indiquer que l'expression du gène d'*actine* utilisé est stable au cours de l'induction de l'embryogenèse somatique (Legrand, communication personnelle). De ce fait, il nous parait raisonnable de discuter des résultats obtenus avec l'*actine* comme gène de contrôle.

Ainsi, au cours des phases précoces de l'embryogenèse somatique, le niveau de transcrits de *CiSERK1* est plus élevé avant la mise en culture dans les fragments foliaires de K59 par rapport à C15. On observe ensuite une chute du niveau de transcrits chez les deux génotypes au cours de la première journée de culture puis une stabilisation. Le niveau de transcrits de *CiSERK2* ne varie pas au cours de l'induction de l'embryogenèse somatique. Au cours de l'embryogenèse zygotique et du développement de la graine, les transcrits de *CiSERK1* montrent une accumulation à partir de 7 jours après pollinisation puis un pic de transcrits a été observé 17 jours après pollinisation. Le niveau de transcrits de *CiSERK2* ne

Par ailleurs, les données qui vont être générées par les analyses d'hybridation sur puce à ADN actuellement en cours de réalisation au laboratoire vont probablement permettre d'avoir accès à d'autres gènes dont les niveaux transcrits sont stables et constitutifs afin d'être utilisés comme contrôle dans les expériences de qRT-PCR. L'utilisation de plusieurs gènes contrôles permettra d'affiner la normalisation de l'expression des gènes cibles et d'accroître la robustesse des résultats (Vandesompele *et al.*, 2002).

Variation des niveaux de transcrits

Chez Arabidopsis, le groupe des gènes de type SERK est composé de 5 membres parmi lesquels 2 seulement ont des profils d'expression ou des fonctions déterminés. AtSERK1 est l'orthologue de DcSERK (Schmidt et al., 1997) et est exprimé dans les cellules compétentes à l'embryogenèse somatique et dans les embryons somatiques jusqu'au stade globulaire chez Arabidopsis ainsi que dans les ovules en développement et dans les embryons zygotiques jusqu'au stade cordiforme (Hecht et al., 2001). AtSERK3, aussi connu sous le nom de BAK1 (BRI1 Assiciated receptor-like Kinase 1), est impliqué dans la perception et la signalisation des hormones brassinostéroïdes. Son expression est ubiquiste (Nam et Li, 2002). Aussi, afin de déterminer si les gènes CiSERK sont fonctionnellement équivalents à DcSERK/AtSERK1, nous avons évalué leur profil d'expression au cours des



<u>Figure 60</u>: Schéma général des résultats obtenus au cours de l'embryogenèse somatique chez le génotype K59.

phases précoces de l'embryogenèse somatique et de l'embryogenèse zygotique de la chicorée.

Les gènes CiSERK et l'embryogenèse somatique

La phase d'induction étudiée qui s'étend sur 8 jours de culture regroupe les différents stades de modification de la morphologie cellulaire caractéristiques de l'induction de l'embryogenèse somatique et témoigne de la modification de l'identité cellulaire chez le génotype embryogène K59. La réactivation des cellules du parenchyme foliaire aboutit à la formation de cellules réactivées compétentes dont certaines vont se diviser et se développer en embryons somatiques (Figure 60).

Les transcrits des gènes *CiSERK* ont été détectés de façon significative dans les fragments foliaires avant la mise en culture (J0). Chez *Arabidopsis*, les transcrits de *AtSERK1* ont été détectés à un très faible niveau dans les tissus de plantes adultes. Ce résultat est cohérent avec l'expression de la protéine GUS, sous contrôle du promoteur de *AtSERK1*, observée au sein des tissus vasculaires (Hecht *et al.*, 2001). Il n'est pas exclu que les gènes *CiSERK* soient exprimés au niveau des tissus vasculaires dans les feuilles ce qui expliquerait la détection des transcrits dans ces tissus.

Cependant, le fait que ces gènes ne soient pas induits lors des phases précoces de l'embryogenèse somatique de la chicorée n'est pas en adéquation avec le profil d'expression de *DcSERK* au cours de l'induction de l'embryogenèse somatique de la carotte. En effet, les transcrits de *DcSERK* sont détectés dans les cellules compétentes à l'embryogenèse somatique (Schmidt *et al.*, 1997). Bien que les modes d'induction de l'embryogenèse somatique soient différents entre la carotte et la chicorée, des cellules réactivées compétentes sont présentes dans les tissus foliaires du génotype embryogène après 4 jours d'induction. Dès lors, on aurait pu s'attendre à observer une induction de l'expression des gènes *CiSERK1* et/ ou *CiSERK2* chez K59 après environ 4 jours de culture.

La différence de mode d'induction de l'embryogenèse, indirecte chez la carotte et directe chez la chicorée, pourrait expliquer cette différence d'expression des gènes *CiSERK*. Cependant, chez le dactyle où l'embryogenèse somatique est de type direct, les transcrits d'un homologue de *DcSERK* sont détectés dans les cellules compétentes des fragments foliaires induits pour former des embryons somatiques (Somleva *et al.*, 2000).

L'embryogenèse somatique des plantes qui expriment des gènes de type *SERK* au cours du processus embryogène nécessite l'apport dans le milieu d'induction d'une concentration élevée d'auxine. Chez la luzerne, l'expression d'un gène de type *SERK*



<u>Figure 61</u> : Schéma général des résultats obtenus au cours de l'embryogenèse zygotique et du développement de la plante.
(*MtSERK*) est inductible par l'auxine (ANA), indépendamment de la voie embryogène ou rhizogène du développement dans laquelle s'engagent les cellules (Nolan *et al.*, 2003). Ceci suggère que l'expression du gène *DcSERK/AtSERK1* pourrait aussi être induite par les auxines. Les conditions d'induction de l'embryogenèse somatique chez la chicorée nécessitent environ 20 fois moins d'auxine (0,1 µM d'ANA) que chez les espèces (carotte, tournesol) qui expriment un gène de type *SERK* au cours des phases précoces de l'embryogenèse somatique. Cette différence pourrait expliquer l'absence d'induction des gènes *CiSERK* dans les tissus embryogènes de chicorée.

Par ailleurs, indépendamment de la concentration en auxine nécessaire à l'induction de l'embryogenèse somatique, l'expression de gènes de type *SERK* n'est pas systématiquement corrélée à l'embryogenèse somatique. En effet, chez le maïs les gènes *ZmSERK1* et *ZmSERK2* sont exprimés dans les cultures de cals embryogènes, et de cals non embryogène (Baudino *et al.*, 2001).

Nous avons observé une différence de niveaux de transcrits du gène *CiSERK1*, dans les feuilles avant mise en culture, entre les deux génotypes. Les transcrits sont plus abondants chez le génotype K59. Cette différence peut être due à une différence physiologique initiale ou due à une variation expérimentale non contrôlée. La répétition des expériences pourrait confirmer ou infirmer cette différence de niveau initial entre les deux génotypes. Pour nos différents essais, les feuilles avaient subi des stress non négligeables au cours des étapes d'aseptisation et de coupure des fragments foliaires. Afin de minimiser les effets des stress sur les niveaux de transcrits, les fragments foliaires des différents génotypes ont été mis en cultures simultanément. Ces stress peuvent cependant induire l'expression rapide de gènes avec des intensités différentes selon les génotypes. Le niveau de transcrits de *CiSERK1* pourrait être induit de façon transitoire juste après le prélèvement des feuilles en réponses aux stress. Afin de vérifier cette hypothèse, il serait nécessaire de réaliser l'évaluation du niveau de transcrits correspondants sur des fragments foliaires avant le traitement de mise en culture (aseptisation, coupe).

Les gènes CiSERK et l'embryogenèse zygotique et le développement de la graine

L'expression des gènes Dc*SERK* et *AtSERK1* est induite au cours de l'embryogenèse zygotique de la carotte et d'*Arabidopsis*. C'est pourquoi nous avons suivi l'évolution de l'accumulation des transcrits des gènes *CiSERK1* et *CiSERK2* au cours de l'embryogenèse zygotique et du développement de la graine chez la chicorée (Figure 61). L'analyse de la variation de transcrits a été réalisée à partir d'akènes récoltés à intervalles réguliers, couvrant le développement de l'embryon et de la graine. Ces expériences nous ont permis

de montrer que les transcrits des gènes *CiSERK* sont détectés dans les ovaires avant la pollinisation. Cet événement est commun aux différents gènes de type *AtSERK1* qui sont exprimés dans les ovules (Tucker *et al.*, 2003) et dans tous les constituants du sac embryonnaire (Hecht *et al.*, 2001).

Tandis que *CiSERK2* montre un niveau d'expression constant, le niveau de *CiSERK1* augmente 7 jours après pollinisation avec un pic d'accumulation des transcrits 17 jours après pollinisation. L'augmentation de l'expression de *CiSERK1* coïncide ainsi avec la mise en place du stade cotylédonaire de l'embryon. Nos travaux montrent que les transcrits des gènes *CiSERK* sont détectés au-delà des stades de développement pour lesquels les transcrits de *DcSERK* et *AtSERK1* sont accumulés dans les embryons de carotte et d'*Arabidopsis*. En effet, chez la carotte, les transcrits de *DcSERK* sont détectés dans les embryons jusqu'au stade globulaire. Chez *Arabidopsis* et chez la marguerite, il a été montré que la protéine GUS dont l'expression est sous le contrôle du promoteur de *AtSERK1* est exprimée dans la graine en développement dans les phases embryonnaires précoces, jusqu'au stade cordiforme (Hecht *et al.*, 2001; Tucker *et al.*, 2003)

Au cours du stade cotylédonaire du développement, différents tissus, comme les méristèmes, se différencient au sein de l'embryon. Les procambiums cotylédonaires et hypocotylaires se différencient et fusionnent. Le gène *CiSERK1* pourrait avoir un rôle dans les processus de différenciation de ces éléments. En effet, chez le tournesol, au moyen d'une sonde partielle correspondant au domaine kinase d'un gène de type *SERK*, l'hybridation *in situ* de coupes d'embryons zygotiques immatures au stade cotylédonaire montre des signaux d'hybridation dans l'épiderme, dans les tissus provasculaires et dans les méristèmes (Thomas *et al.*, 2004). Cette localisation des transcrits de *SERK* n'est pas mise en évidence dans l'embryon zygotique à ce stade de développement chez d'autres espèces (Hecht *et al.*, 2001; Tucker *et al.*, 2003).

L'utilisation d'une sonde correspondant au domaine kinase conservé a peut être permis d'hybrider d'autres transcrits de type *SERK*, d'autant plus qu'il existe 4 séquences partielles de type *SERK* chez le tournesol. L'existence de transcrits de type *SERK* dans ces tissus corrobore l'augmentation du niveau de transcrits de *CiSERK1* à partir de 7 jours après la pollinisation et l'expression de *CiSERK2* à ce stade de développement.

L'ensemble de ces données suggère que les gènes *CiSERK* ne sont pas les équivalents fonctionnels de *DcSERK/AtSERK1*.

Aussi, nous avons évalué les profils d'expression des gènes *CiSERK1* et *CiSERK2* dans différents organes afin de caractériser d'avantage les modalités d'expression de ces deux gènes.

Expression des gènes CiSERK dans d'autres organes.

Les transcrits de CiSERK1 ont montré une chute significative de leur niveau dès le premier jour de culture embryogène. Afin de tester l'hypothèse d'un effet de la lumière sur la répression des transcrits des gènes CiSERK nous avons expérimenté l'influence de la lumière dans l'expression des gènes CiSERK au cours de la germination de graine. Il s'avère que l'expression de CiSERK1 est plus élevée dans les plantules germées 4 jours en absence de lumière qu'en présence de lumière et que cette différence de niveau d'expression disparaît après 8 jours de germination. Le niveau de transcrits de CiSERK2 ne varie pas ni, en fonction de la lumière, ni au cours du temps. La lumière semble donc limiter l'expression de CiSERK1 de façon transitoire ce qui correspond à l'inverse de l'effet attendu bien qu'il ne s'agisse pas du même organe (l'embryogenèse somatique étant induite à partir de fragments foliaires). Cependant, ce profil d'expression au cours de la germination en fonction de la lumière peut être rapproché de celui de BRI1 chez Arabidopsis codant le récepteur aux brassinostéroïdes. Les transcrits de BRI1 sont plus abondants dans les plantules germées 3 jours à l'obscurité qu'à la lumière et cette différence s'estompe après 5 jours de germination (Li et Chory, 1997). Au regard du fait que BRI1 est associée à BAK1 (aussi appelé AtSERK3) (Li et al., 2002) qui code un récepteur transmembranaire de type SERK, il est envisageable que les transcrits soient régulés de la même manière. Ainsi BAK1 pourrait avoir le même profil d'expression que BRI1. Cet élément suggère que l'expression de CiSERK1 pourrait alors correspondre d'avantage à celle de BAK1/AtSERK3 qu'à celle d'AtSERK1.

Nous avons analysé les niveaux de transcrits de *CiSERK1* et *CiSERK2* dans différents organes afin de déterminer s'ils sont exprimés préférentiellement dans certains organes. Il apparaît que le niveau de transcrits de *CiSERK2* varie peu entre les différents organes testés ; il est identique dans les germinations et les feuilles et seulement deux fois moins important dans les ovaires et au cours du développement de la graine. Par ailleurs le niveau de transcrits de *CiSERK2* est stable dans les tissus soumis à l'induction de l'embryogenèse somatique et au cours de l'embryogenèse zygotique et du développement de la graine. Les niveaux de transcrits de *CiSERK1* sont beaucoup plus variables entre les organes puisque dans les ovaires et dans les feuilles. Les transcrits de *CiSERK1* dans les fragments foliaires sont 4 fois moins abondants après 8 jours de culture dans un milieu d'induction. L'expression de *CiSERK1* varie aussi lors de l'embryogenèse somatique et du développement de la graine. Les deux gènes sont très faiblement transcrits dans les boutons floraux 8-10 jours avant l'anthèse.

Par ailleurs, les niveaux des transcrits de *CiSERK2* sont toujours plus élevés que ceux de *CiSERK1*. Il apparaît que les deux gènes clonés chez la chicorée, bien qu'ils soient très proches structurellement, montrent des niveaux de transcrits relatifs différents au sein d'un même tissu et des profils d'expression assez différentes entre différents organes.

Variation de l'accumulation des protéines

Parallèlement à l'analyse de la variation des niveaux de transcrits, nous avons entrepris d'étudier la variation du niveau protéique correspondant à *CiSERK1* au cours de l'induction de l'embryogenèse somatique et de l'embryogenèse zygotique. Pour cela, nous avons produit un anticorps dirigé contre le domaine kinase de CiSERK1. Du fait de la forte conservation du domaine kinase des protéines CiSERK1 et CiSERK2, il est fortement probable que cet anticorps est capable de reconnaître le domaine kinase des 2 protéines ainsi que celui d'autres protéines kinase apparentées à SERK, appartenant à la famille des RLK de type LRRII.

Ceci est d'ailleurs renforcé par l'analyse des profils d'immunoblots obtenus qui montrent plusieurs bandes, confirmant que l'anticorps reconnaît des épitopes communs à différentes protéines. Cette spécificité élargie de l'anticorps aux domaines kinases d'autres protéines a permis de mettre en évidence des protéines liées à l'embryogenèse somatique et au développement de la graine.

Protéine kinases et embryogenèse somatique

Les analyses par western blot des extraits protéiques issus des fragments foliaires soumis à l'induction de l'embryogenèse somatique montrent plusieurs bandes dont la taille est compatible avec celle de récepteurs kinases de type SERK (Figure 60).

Une bande protéique de 83 kDa observée dans les tissus foliaire à J0 montre un niveau d'accumulation plus important dans les feuilles de K59 que dans celles de C15. Cette différence s'estompe dès le premier jour de culture. Ce profil n'est pas sans rappeler celui observé pour les transcrits de *CiSERK1* qui montre un niveau de transcrits dans les fragments foliaires plus élevé chez le génotype K59 avant la mise en culture que chez le génotype C15; cette différence de niveau de transcrit s'estompe également dès la première journée de culture. De plus, les protéines correspondantes à *CiSERK1* et *CiSERK2* ont des masses moléculaires sensiblement identiques (environ 65 kDa) et possèdent toutes deux le même nombre de sites de N-glycosylation qui augmenterait leur masse dans les mêmes proportions. Il est alors vraisemblable que ces deux protéines ne peuvent pas être distinguées par immunoblot. Dans la mesure où les transcrits de ces deux gènes ont été

détectés dans les feuilles avant la mise en culture, et que le niveau de transcrits stable de *CiSERK2* est beaucoup plus élevé que celui de *CiSERK1*, il est possible que cette bande protéique corresponde au produits des gènes *CiSERK*, même s'il semble difficile de faire un parallèle entre le niveau de transcrits et l'accumulation des protéines correspondantes.

Une bande protéique de 75 kDa est détectée dans les tissus embryogènes et non embryogènes après 2 jours de culture. Alors que le niveau de protéine est constant dans le génotype non embryogène, une augmentation est notée dans le génotype embryogène. Le fait que cette protéine est détectée lors de la phase d'induction chez les deux génotypes suppose qu'elle soit impliquée dans un processus commun aux génotypes embryogène et non embryogène. Après deux jours de culture, les explants foliaires présentent des cellules en cours de réactivation chez les deux génotypes. La présence de cette protéine de 75 kDa pourrait être ainsi corrélée à la dédifférenciation des cellules du parenchyme, d'autant plus que l'intensité de cette bande augmente chez le génotype embryogène. De plus, à partir du 3^e jour de culture, la réactivation cellulaire se poursuit chez K59 et n'augmente plus chez C15.

Une bande protéique de 62 kDa est détectée après 3 jours de culture et s'accumule jusqu'au 5^e jour de culture de façon spécifique dans les tissus du génotype embryogène. Cette bande protéique serait donc liée à la réponse du génotype K59 à la phase d'induction de l'embryogenèse somatique. Le profil d'accumulation de cette protéine semble corrélée à la présence de cellules réactivées observées dès 3 jours de culture dans les tissus embryogènes uniquement. Ces cellules réactivées ont les caractéristiques de cellules dédifférenciées et peuvent s'engager, à se stade, dans une nouvelle voie de différenciation.

Chez la carotte des embryons somatiques se développent à partir de cellules isolées ou de petits amas de cellules dédifférenciées qui expriment la protéine luciférase sous le contrôle du promoteur de *DcSERK* jusqu'au stade globulaire du développement de l'embryon somatique (Schmidt *et al.*, 1997). De plus, chez *Arabidopsis*, les cellules embryogènes compétentes qui se développent en embryons somatiques au sein de cals expriment la protéine GUS sous le contrôle du promoteur de *AtSERK1* tandis qu'aucune coloration due à l'activité GUS n'est observée dans les cellules non embryogènes (Hecht *et al.*, 2001).

Le profil d'accumulation de la protéine de 62 kDa peut être rapproché du profil d'expression de *DcSERK/AtSERK1*, du fait qu'elle n'est détectée que dans les tissus du génotype embryogène et que son expression est corrélée à la présence de cellules réactivées dans les explants.

Protéine kinase et embryogenèse zygotique

Afin de déterminer s'il existe des éléments communs à l'induction de l'embryogenèse somatique et aux stades précoces de l'embryogenèse zygotique, nous avons réalisé des expériences d'immunoblots à partir de protéines extraites d'akènes en développement.

Le nombre de bandes protéiques observées, ayant une taille compatible avec celle d'un récepteur kinase de type SERK, est plus faible par rapport à celui observé dans les feuilles au cours l'embryogenèse somatique (Figure 61). Ceci suggère que l'expression de la protéine de 83 kDa est plutôt spécifique des tissus foliaires (car présente avant l'induction de l'embryogenèse somatique) et que l'accumulation de la protéine de 75 kDa est induite par les conditions d'induction de l'embryogenèse somatique chez les deux génotypes.

Par ailleurs si on considère que la bande de 83 kDa puisse correspondre aux produits des gènes *CiSERK*, le niveau de transcrit de ces gènes dans les ovaires au cours de l'embryogenèse zygotique est plus faible que dans les feuilles et cette différence de représentation des transcrits peut suffire à diminuer l'accumulation des protéines au point qu'elles ne soient plus détectées dans les ovaires et les embryons zygotiques.

Une bande protéique de 62 kDa est détectée dans les extraits d'akènes avant la pollinisation et au cours du développement de la graine. La quantité de cette protéine semble augmenter 11 jours après la pollinisation et n'est plus détectée à J19ap. Cette protéine a une masse moléculaire proche ou identique à celle observée au cours de l'induction de l'embryogenèse somatique induite à partir des fragments foliaires. S'il s'agit de la même protéine, elle intervient à la fois au cours des processus d'embryogenèse somatique et zygotique. Cependant, la prudence s'impose quant à l'identité de ces deux protéines, qui ne pourrait être confirmée que par des analyses approfondies.

Chez *Arabidopsis*, la protéine GUS exprimée sous le contrôle du promoteur d'*AtSERK1* est présente dans les ovules. Par ailleurs, les transcrits de *AtSERK1* sont détectés dans l'embryon jusqu'au stade cordiforme (Hecht *et al.*, 2001) ce qui laisse présager de la présence de la protéine jusqu'à ce stade de développement.

Au contraire, des transcrit de gènes de type *SERK* ont été détectés dans les embryons zygotiques immatures de tournesol au stade cotylédonaire (Thomas *et al.*, 2004). Ainsi, il est probable que des transcrits de gènes *SERK* soient présents à tous les stades de développement de l'embryon ce qui pourrait être corrélé à la présence de la protéine de 62 kDa au cours du développement de la graine de chicorée.

Une bande protéique d'une masse moléculaire de 55 kDa est détectée 11 jours après la pollinisation et semble constante jusque 19 jours après la pollinisation. Sa masse moléculaire est compatible avec celle d'un récepteur de type kinase. Son expression est induite au stade cotylédonaire du développement de l'embryon.

Cette protéine est détectée de façon concomitante à la détection d'une autre bande protéique de 35 kDa dont la MM est moins compatible avec celle d'un récepteur transmembranaire. La présence des ces deux protéines peut être corrélée au stade cotylédonaire de l'embryon mais aussi à la présence des bandes de 20 kDa environ observée sur les gels d'électrophorèse et qui peuvent correspondre à des protéines de réserves. Ces observations suggèrent que des protéines de type kinase peuvent être impliquées dans la mise en réserve des tissus embryonnaires.

Conclusion

Perspectives

L'objectif de ce travail a été de montrer l'existence de gènes codant des récepteurs de type kinase apparentés à *DcSERK* chez la chicorée ainsi que l'implication des produits de ces gènes dans les phases précoces de l'induction de l'embryogenèse somatique, de l'embryogenèse zygotique et du développement de la graine.

Par une approche gènes candidats, nous avons isolé deux gènes présentant les caractéristiques de *DcSERK*. Ces gènes montrent une structure génique identique à celle des 5 membres apparentés à *AtSERK1* chez *Arabidopsis*, mais aussi chez la carotte, le maïs ou la luzerne. Les 11 exons des 2 gènes codent des motifs très conservés entre les différentes protéines de type SERK. Au sein du groupe formé par les gènes de type *SERK*, les gènes *CiSERK* semblent être structurellement plus proches d'*AtSERK3* que d'*AtSERK1*.

L'étude par PCR quantitative en temps réel de la variation du niveau de transcrits des gènes *CiSERK* au cours des phases précoces de l'embryogenèse somatique de la chicorée n'a pas montré une induction de leur expression, contrairement à ce qui a été observé chez d'autres espèces. Du fait que ces gènes appartiennent à des familles multigéniques, il n'est pas exclu que le ou les gènes orthologues de *DcSERK/AtSERK1* (c'est à dire possédant à la fois une forte identité de structure mais aussi un rôle fonctionnel similaire), existent chez la chicorée mais n'ont pas été identifiés lors de ce travail.

Les données plus abondantes de séquences disponibles aujourd'hui permettraient de définir des amorces dégénérées plus spécifiques au sous-groupe auquel appartient *AtSERK1*. Une amorce dégénérée pourrait alors être définie dans la région riche en proline notamment au niveau de la séquence consensuelle PCPGPPFSPPPP que nous avons identifiée comme spécifique au sous-groupe des gènes de type *SERK* comprenant *DcSERK/AtSERK1*. Ceci permettrait de vérifier de manière simple l'existence de gènes de type *DcSERK/AtSERK1* chez la chicorée.

En effet, lors de nos expériences de Southern blot nous avons montré qu'il existe plusieurs gènes de type *SERK* chez la chicorée en plus des 2 identifiés. Ceci a été renforcé par nos expériences de western blot utilisant des anticorps dirigés contre le domaine kinase des protéines CiSERK et qui ont montré la présence d'autres bandes protéiques.

Cependant, l'existence de gènes de type *DcSERK/AtSERK1* chez la chicorée ne préjuge en rien de leur expression au cours des phases précoces de l'embryogenèse somatique. En effet, des hypothèses vont dans le sens d'une induction des gènes de type *AtSERK1* par l'auxine (Nolan *et al.*, 2003; Somleva *et al.*, 2000; Thomas *et al.*, 2004) dont les concentrations sont très faibles pour induire l'embryogenèse somatique chez la chicorée. Par ailleurs, chez le maïs, l'expression des gènes *ZmSERK* n'est pas corrélée au potentiel

embryogène des cals. Néanmoins, bien que les séquences existantes chez les Astéracées comme le tournesol et la marguerite ne sont que des séquences partielles correspondant au domaine kinase très conservé de *SERK*, l'implication de ces gènes dans l'embryogenèse somatique ou zygotique a été décrite chez ces espèces (Thomas *et al.*, 2004; Tucker *et al.*, 2003). Ces éléments sont corroborés par les résultats que nous avons obtenus par western blot qui semblent indiquer que des protéines de type SERK sont exprimées de façon spécifique au cours de l'embryogenèse somatique.

Une méthode qui permettrait de déterminer l'implication potentielle de gènes de type *SERK* au cours de l'embryogenèse somatique serait d'exprimer au sein de tissus de chicorée embryogène et non embryogène un gène rapporteur sous le contrôle du promoteur d'*AtSERK1*, comme cela a été réalisé chez la marguerite (Tucker *et al.*, 2003). Si l'induction de l'embryogenèse somatique nécessite l'expression d'un gène *SERK*, on peut penser que le gène rapporteur sera exprimé. Par ailleurs, en optant pour un gène rapporteur (de type GUS par exemple), les types cellulaires qui expriment ce gène pourraient être identifiés cytologiquement. Cette construction permettrait également de déterminer et de localiser l'expression du gène *SERK* au cours de l'embryogenèse zygotique et dans l'ensemble des tissus de chicorée.

L'analyse des profils d'expression des gènes *CiSERK* entre différents organes a permis de montrer qu'ils sont présents à des niveaux différents et variables selon les organes, le stade de développement des embryons zygotiques et les conditions de culture appliquées aux explants foliaires. Le niveau de transcrits de *CiSERK1* est d'avantage variable que celui de *CiSERK2*. L'étude du niveau de transcrits dans d'autres organes pourrait être envisagée afin d'affiner le profil d'expression de ces gènes. En particulier, le niveau d'expression dans les racines n'a pas pu être déterminé en raison de difficultés à extraire des ARN de bonne qualité de cet organe riche en glycannes.

Cependant, les profils d'expression des gènes *CiSERK* ne permettent pas de déterminer une fonction spécifique à ces gènes. Il serait intéressant d'observer les conséquences phénotypiques de la suppression des transcrits dans les tissus par ARN interférence par exemple, en choisissant, dans la mesure des données disponibles, la séquence de l'ARN interférant de telle sorte qu'elle soit la plus spécifique possible de la cible à supprimer.

Parallèlement, afin de mettre en évidence la présence des protéines CiSERK dans les différents extraits étudiés, nous avons obtenu un immunsérum dirigé contre le domaine kinase de CiSERK1. Plusieurs bandes protéiques de tailles compatibles avec celles de récepteurs de type kinase ont été détectées au cours des phases précoces de l'embryogenèse somatique et du développement de la graine. L'utilisation d'anticorps dirigés contre un domaine kinase conservé a permis de mettre en évidence l'accumulation de bandes protéiques qui peuvent être impliquées dans les processus étudiés. Trois bandes protéiques ont attiré notre attention:

_ une bande protéique de 83 kDa détectée dans les fragments foliaires et dont le profil d'accumulation ne semble pas lié à l'embryogenèse somatique.

_ une protéine de 75 kDa détectée au cours de l'induction de l'embryogenèse somatique de façon corrélée à la réactivation cellulaire.

_ une protéine de 62 kDa détectée au cours de l'induction de l'embryogenèse et corrélée à la présence de cellules réactivées dans les tissus embryogènes. Une bande protéique de taille identique a été détectée au cours de l'embryogenèse zygotique.

Il serait intéressant d'identifier ces deux dernières protéines dont la présence est corrélée à deux états cellulaires différents et successifs au sein des tissus embryogènes (cellules en cours de réactivation et cellules réactivées). La détection de ces protéines a été réalisée au moyen des anticorps dirigés contre le domaine kinase des protéines CiSERK, aussi, une chromatographie d'affinité utilisant les anticorps couplés à une matrice permettrait de piéger spécifiquement les protéines avec lesquelles ils interagissent. Après élution, celles-ci pourraient être séparées selon leur masse moléculaire, par électrophorèse par exemple, puis soumises à une identification et un séquençage par spectrométrie de masse (MS-MS) ou selon la méthode d'Edman.

Ceci permettrait notamment de déterminer si la bande d'une taille de 62 kDa détectée à la fois dans les tissus foliaires au cours de l'induction de l'embryogenèse somatique et au cours du développement de la graine sont identiques et ouvrirait alors des perspectives sur l'identification d'événements communs à l'embryogenèse somatique et zygotique chez la chicorée. Références Bibliographiques

- An, Y. Q., McDowell, J. M., Huang, S., McKinney, E. C., Chambliss, S., et Meagher, R. B. (1996). Strong, constitutive expression of the *Arabidopsis* ACT2/ACT8 actin subclass in vegetative tissues. *Plant J.* 10: 107-121.
- Anderson, C. M., Wagner, T. A., Perret, M., He, Z. H., He, D. Z., et Kohorn, B. D. (2001). WAKs: cell wall-associated kinases linking the cytoplasm to the extracellular matrix. *Plant Molecular Biology* 47: 197-206.
- Anil, V. S., Harmon, A. C., et Rao, K. S. (2000). Spatio-temporal accumulation and activity of calciumdependent protein kinases during embryogenesis, seed development, and germination in sandalwood. *Plant Physiology* 122: 1035-1043.
- Anil, V. S. et Rao, K. S. (2000). Calcium-mediated signaling during sandalwood somatic embryogenesis. Role for exogenous calcium as second messenger. *Plant Physiology* 123 : 1301-1312.
- Baldwin, T. C., Domingo, C., Schindler, T., Seetharaman, G., Stacey, N., et Roberts, K. (2001). DcAGP1, a secreted arabinogalactan protein, is related to a family of basic proline-rich proteins. *Plant Molecular Biology* **45**: 421-435.
- Balestrazzi, A., Toscano, I., Bernacchia, G., Luo, M., Otte, S., et Carbonera, D. (**1996**). Cloning of a cDNA encoding DNA topoisomerase I in *Daucus carota* and expression analysis in relation to cell proliferation. *Gene* **183**: 183-190.
- Balestrazzi, A., Bernacchia, G., Pitto, L., Luccarini, G., et Carbonera, D. (**2001**). Spatial expression of DNA topoisomerase I genes during cell proliferation in *Daucus carota. Eur. J. Histochem.* **45**: 31-38.
- Barre, A., Herve, C., Lescure, B., et Rouge, P. (2002). Lectin Receptor Kinases in Plants. *Critical Reviews in Plant Sciences* 21: 379-399.
- Bateman, A. et Bycroft, M. (2000). The structure of a LysM domain from E. coli membrane-bound lytic murein transglycosylase D (MltD). *J. Mol. Biol.* 299: 1113-1119.
- Baudino, S., Hansen, S., Brettschneider, R., Hecht, V. R. G., Dresselhaus, T., Lorz, H., Dumas, C., et Rogowsky, P. M. (2001). Molecular characterisation of two novel maize LRR receptor- like kinases, which belong to the SERK gene family. *Planta* 213: 1-10.
- Becraft, P. W., Stinard, P. S., et McCarty, D. R. (1996). CRINKLY4: A TNFR-like receptor kinase involved in maize epidermal differentiation. *Science* 273: 1406-1409.

Becraft, P. W. (1998). Receptor kinases in plant development. Trends in Plant Science 3: 384-388.

- Becraft, P. W. et Asuncion-Crabb, Y. (2000). Positional cues specify and maintain aleurone cell fate in maize endosperm development. *Development* 127: 4039-4048.
- Becraft, P. W., Kang, S. H., et Suh, S. G. (2001). The maize crinkly4 receptor kinase controls a cellautonomous differentiation response. *Plant Physiology* 127: 486-496.
- Blervacq, A. S., Dubois, T., Dubois, J., et Vasseur, J. (1995). First divisions of somatic embryogenic cells in *Cichorium* hybrid "474". *Protoplasma* 186: 163-168.
- Bogre, L., Stefanov, I., Abraham, M., Somogyi, I., and Dudits, D. (**1990**) Differences in the responses to 2,4-dichlorophenoxyacetic acid (2,4-D) treatment between embryogenic and non embryogenic lines of alfalfa. Dans, *Progress in Plant Cellular and Molecular Biology* (pp 427-436), par Nijkamp, H. J. J., van der Plas, L. H. W., and Van Aartrijk, J. Editions Kluwer Academic Publishers, Dordrecht.
- Bonello, J. F., Sevilla-Lecoq, S., Berne, A., Risueno, M. C., Dumas, C., et Rogowsky, P. M. (**2002**). Esr proteins are secreted by the cells of the embryo surrounding region. *J. Exp. Bot.* **53**: 1559-1568.
- Bower, M. S., Matias, D. D., Fernandes-Carvalho, E., Mazzurco, M., Gu, T., Rothstein, S. J., et Goring, D. R. (1996). Two members of the thioredoxin-h family interact with the kinase domain of a *Brassica* S locus receptor kinase. *Plant Cell* 8: 1641-1650.
- Boyer, C., Hilbert, J. L., et Vasseur, J. (1993). Embryogenesis-related protein synthesis and accumulation during early acquisition of somatic embryogenesis in *Cichorium. Plant Sci.* 93: 41-53.
- Brand, U., Fletcher, J. C., Hobe, M., Meyerowitz, E. M., et Simon, R. (**2000**). Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by CLV3 activity. *Science* **289**: 617-619.
- Braun, D. M., Stone, J. M., et Walker, J. C. (**1997**). Interaction of the maize and *Arabidopsis* kinase interaction domains with a subset of receptor-like protein kinases: implications for transmembrane signaling in plants. *Plant J.* **12**: 83-95.
- Bustin, S. A. (**2000**). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* **25**: 169-193.
- Cabrillac, D., Cock, J. M., Dumas, C., et Gaude, T. (2001). The S-locus receptor kinase is inhibited by thioredoxins and activated by pollen coat proteins. *Nature* 410: 220-223.
- Canales, C., Bhatt, A. M., Scott, R., et Dickinson, H. (2002). EXS, a putative LRR receptor kinase, regulates male germline cell number and tapetal identity and promotes seed development in *Arabidopsis. Curr. Biol.* 12: 1718-1727.
- Chan, R. L., Gago, G. M., Palena, C. M., et Gonzalez, D. H. (1998). Homeoboxes in plant development. *Biochim. Biophys. Acta* 1442: 1-19.

- Chapman, A., Blervacq, A. S., Vasseur, J., et Hilbert, J. L. (2000). Arabinogalactan-proteins in *Cichorium* somatic embryogenesis: effect of beta-glucosyl Yariv reagent and epitope localisation during embryo development. *Planta* 211: 305-314.
- Charrier, B., Champion, A., Henry, Y., et Kreis, M. (**2002**). Expression profiling of the whole *Arabidopsis* shaggy-like kinase multigene family by real-time reverse transcriptase-polymerase chain reaction. *Plant Physiology* **130**: 577-590.
- Charriere, F., Sotta, B., Miginiac, E., et Hahne, G. (**1999**). Induction of adventitius of somatic embryos on *in vitro* cultured zygotic embryos of *Helianthus annuus* : Variation of endogenous hormone levels. *Plant Physiol. Biochem.* **37**: 751-757.
- Choi, Y. E., Yang, D. C., Park, J. C., Soh, W. Y., et Choi, K. T. (1998). Regenerative ability of somatic single and multiple embryos from cotyledons of Korean ginseng on hormone-free medium. *Plant Cell Rep.* 17: 544-551.
- Clark, S. E., Running, M. P., et Meyerowitz, E. M. (1993). CLAVATA1, a regulator of meristem and flower development in *Arabidopsis. Development* 119: 397-418.
- Clark, S. E., Running, M. P., et Meyerowitz, E. M. (1995). CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. *Development* 121: 2057-2067.
- Clark, S. E., Williams, R. W., et Meyerowitz, E. M. (1997). The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* 89: 575-585.
- Clay, N. K. et Nelson, T. (2002). VH1, a Provascular Cell-Specific Receptor Kinase That Influences Leaf Cell Patterns in *Arabidopsis. Plant Cell* 14: 2707-2722.
- Clouse, S. D., Langford, M., et McMorris, T. C. (**1996**). A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. *Plant Physiology* **111**: 671-678.
- Cock, J. M. et McCormick, S. (2001). A large family of genes that share homology with CLAVATA3. *Plant Physiology* 126: 939-942.
- Cordewener, J., Booij, H., van der Zandt, H., van Engelen, F. A., van Kammen, A., et de Vries, S. (1991). Tunicamycin-inhibited carrot somatic embryogenesis can be restored by secreted cationic peroxidase isoenzymes. *Planta* 184: 478-486.
- Cummins, I. et Murphy, D.J. (1992). cDNA sequence of a sunflower oleosin and transcrit tissue specificity. *Plant Mol. Biol.* 19:873-876.

- Damerval, C., de Vienne, D., Zivy, M., et Thiellement, H. (**1986**). Technical improvements in twodimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis* **7**: 52-54.
- Da Silva, S. (2004). Identification des gènes CiSTM et CiCLV1 chez Cichorium intybus L.. Implication dans les modifications d'identité cellulaire au cours des phases précoces de l'embryogenèse somatique et du développement des embryons zygotique eet de la graine. Thèse Université Lille1.
- Davletova, S., Meszaros, T., Miskolczi, P., Oberschall, A., Torok, K., Magyar, Z., Dudits, D., et Deak,
 M. (2001). Auxin and heat shock activation of a novel member of the calmodulin like domain protein kinase gene family in cultured alfalfa cells. *J. Exp. Bot.* 52: 215-221.
- de Jong, A. J., Cordewener, J., Lo, S. F., Terzi, M., Vandekerckhove, J., van Kammen, A., et de Vries, S. C. (**1992**). A carrot somatic embryo mutant is rescued by chitinase. *Plant Cell* **4**: 425-433.
- de Jong, A. J., Schmidt, E. D., et de Vries, S. C. (1993). Early events in higher-plant embryogenesis. *Plant Mol. Biol.* 22: 367-377.
- de Vries, S. C., Booij, H., Janssens, R., Vogels, R., Saris, L., Lo Schiavo, F., Terzi, M., et van Kammen, A. (1988). Carrot somatic embryogenesis depends on the phytohormone-controlled presence of correctly glycosylated extracellular protein. *Genes Dev.* 2: 462-476.
- Dixit, R., Nasrallah, M. E., et Nasrallah, J. B. (**2000**). Post-transcriptional maturation of the S receptor kinase of *Brassica* correlates with co-expression of the S-Locus glycoprotein in the stigmas of two *Brassica* strains and in transgenic tobacco plants. *Plant Physiology* **124**: 297-311.
- Dong, J. Z. et Dunstan, D. I. (**1997a**). Characterization of cDNAs representing five abscisic acidresponsive genes associated with somatic embryogenesis in *Picea glauca*, and their responses their responses to abscisic acid stereostructure. *Planta* **203**: 448-453.
- Dong, J. Z. et Dunstan, D. I. (**1997b**). Endochitinase and beta-1,3-glucanase genes are developmentally regulated during somatic embryogenesis in *Picea glauca*. *Planta* **201**: 189-194.
- Duban, M. (2000). Clonage et expression de gènes codant des récepteurs de type kinase au cours du processus d'induction de l'embryogenèse somatique de la chicorée. DEA Université Lille1
- Dubois, T., Dubois, J., Guedira, M., et Vasseur, J. (1988). Embryogenèse somatique directe sur les styles de *Cichorium*: effets de la température et origine des embryoïdes. *C. R. Acad. Sci. Paris* 307: 667-675.
- Dubois, T., Guedira, M., Dubois, J., et Vasseur, J. (1990). Direct somatic embryogenesis in roots of *Cichorium*: is callose an early marker? *Ann. Bot.* **65**: 539-545.

- Dubois, T., Guedira, M., Dubois, J., et Vasseur, J. (1991). Direct somatic embryogenesis in leaves of *Cichorium*. A histological and SEM study of early stage. *Protoplasma* 162: 120-127.
- Dudits, D., Bogre, L., et Gyorgyey J. (1991). Molecular and cellular approaches to the analysis of plant embryo development from somatic cells in vitro. *J. Cell Sci.* 99: 475-484.
- Dudits, D., Gyorgyey, J., Bogre, L., and Bako, L. (**1995**) Molecular biology of somatic embryogenesis. Dans, *In Vitro Embryogenesis in Plants* (pp 267-308), par Thorpe, T. A. Editions Kluwer Academic Publishers, Dordrecht.
- Dupire, L. (**1999**) Etude d'un mutant embryogène d'*Asparagus officinalis* L.: description cytologique, approches biochimiques et moléculaires. Thèse Université Lille1.
- Egertsdotter, U. et von Arnold, S. (**1995**). Importance of arabinogalactan proteins for the development of somatic embryos of Norway spruce (*Picea abies*). *Physiologia Plantarum* **93**: 334-345.
- Egertsdotter, U. et von Arnold, S. (**1998**). Development of somatic embryos in Norway spruce. *J. Exp. Bot.* **49**: 155-162.
- El-Maâtaoui, M., Espagnac, H., et Michaux-Ferrière, N. (1990). Histology of callogenesis and somatic embryogenesis induced in stem fragments of Cork oak (*Quercus suber*) cultured *in vitro. Ann. Bot.* 66: 183-190.
- Fargasova, A. (**1994**). Comparative study of plant growth hormone (herbicide) toxicity in various biological subjects. *Ecotoxicol. Environ. Saf* **29**: 359-364.
- Feher, A., Pasternak, T. P., Miskolczi, P., Ayaydin, F., et Dudits, D. (2001). Induction of the embryogenic pathway in somatic plant cells. *Acta Hort.* 560: 293-298.
- Feher, A., Pasternak, T. P., et Dudits, D. (2003). Transition of somatic plant cells to an embryogenic state. *Plant Cell, Tissue and Organ Culture* 74: 201-228.
- Filipecki, M. K., Sommer, H., et Malepszy, S. (**1997**). The MADS-box gene *CUS1* is expressed during cucumber somatic embryogenesis. *Plant Sci.* **125**: 63-74.
- Fletcher, J. C., Brand, U., Running, M. P., Simon, R., et Meyerowitz, E. M. (1999). Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems. *Science* 283: 1911-1914.
- Fletcher, J. C. et Meyerowitz, E. M. (2000). Cell signaling within the shoot meristem. *Curr. Opin. Plant Biol.* 3: 23-30.
- Flybase (**1999**). The FlyBase database of the Drosophila Genome Projects and community literature. The FlyBase Consortium. *Nucleic Acids Res.* **27**: 85-88.

- Footitt, S., Ingouff, M., Clapham, D., et von Arnold, S. (2003). Expression of the viviparous 1 (*Pavp1*) and p34cdc2 protein kinase (*cdc2Pa*) genes during somatic embryogenesis in Norway spruce (*Picea abies* [L.] Karst). J. Exp. Bot. 54: 1711-1719.
- Franz, G., Hatzopoulos, P., Jones, T. J., Krauss, M., et Sung, Z. R. (**1989**). Molecular and genetic analysis of an embryonic gene, DC 8, from *Daucus carota* L. *Mol. Gen. Genet.* **218**: 143-151.
- Friedrichsen, D. M., Joazeiro, C. A., Li, J., Hunter, T., et Chory, J. (**2000**). Brassinosteroid-insensitive-1 is a ubiquitously expressed leucine-rich repeat receptor serine/threonine kinase. *Plant Physiology* **123**: 1247-1256.
- Galland, R., Randoux, B., Vasseur, J., et Hilbert, J. L. (**2001**). A glutathione S-transferase cDNA identified by mRNA differential display is upregulated during somatic embryogenesis in *Cichorium*. *Biochim. Biophys. Acta* **1522**: 212-216.
- Giranton, J. L., Dumas, C., Cock, J. M., et Gaude, T. (**2000**). The integral membrane S-locus receptor kinase of *Brassica* has serine/threonine kinase activity in a membranous environment and spontaneously forms oligomers in planta. *Proc. Natl. Acad. Sci. U. S. A* **97**: 3759-3764.
- Gonnet, G. H., Korostensky, C., et Benner, S. (2000). Evaluation measures of multiple sequence alignments. *J. Comput. Biol.* 7: 261-276.
- Goring, D. R., Glavin, T. L., Schafer, U., et Rothstein, S. J. (**1993**). An S receptor kinase gene in selfcompatible *Brassica napus* has a 1-bp deletion. *Plant Cell* **5**: 531-539.
- Goupil, P., Hatzopoulos, P., Franz, G., Hempel, F. D., You, R., et Sung, Z. R. (**1992**). Transcriptional regulation of a seed-specific carrot gene, *DC8. Plant Mol. Biol.* **18** : 1049-1063.
- Grosset, J., Marty, I., Chartier, Y., et Meyer, Y. (**1990**). mRNAs newly synthesized by tobacco mesophyll protoplasts are wound-inducible. *Plant Mol. Biol.* **15**: 485-496.
- Grossmann, K. (**2000**). Mode of action of auxin herbicides: a new ending to a long, drawn out story. *Trends Plant Sci.* **5**: 506-508.
- Gu, T., Mazzurco, M., Sulaman, W., Matias, D. D., et Goring, D. R. (**1998**). Binding of an arm repeat protein to the kinase domain of the S-locus receptor kinase. *Proc. Natl. Acad. Sci. U. S. A* **95**: 382-387.
- Guedira, M., Dubois-Tylski, T., Vasseur, J., et Dubois, J. (**1989**). Embryogenèse somatique directe à partir de culture d'anthères du *Cichorium* (Asteraceae). *Can. J. Bot.* **67**: 970-976.
- Guha, S. et Maheshwari, S. C. (1964). *In vitro* production of embryo from anthers of *Datura*. *Nature* **204**: 497

- Hajouj, T., Michelis, R., et Gepstein, S. (2000). Cloning and characterization of a receptor-like protein kinase gene associated with senescence. *Plant Physiology* 124: 1305-1314.
- Hanks, S. K., Quinn, A. M., et Hunter, T. (**1988**). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**: 42-52.
- Hanning, G. E. et Conger, B. V. (1982). Embryoïd and plantelet formation from leaf segments of *Dactylis glomerata. Theorical and applied genetics* 63: 155-159.
- Hardie, D. G. (**1999**). PLANT PROTEIN SERINE/THREONINE KINASES: Classification and Functions. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**: 97-131.
- Harmon, A. C., Putnam-Evans, C. L., et Cormier, M. J. (**1986**). Calcium dependent but calmodulin independent protein kinase from soybean. *Plant Physiology* **83**: 830-837.
- Harper, J. F., Sussman, M. R., Schaller, G. E., Putnam-Evans, C., Charbonneau, H., et Harmon, A. C. (**1991**). A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. *Science* **252**: 951-954.
- Harper, J. F. (2001). Dissecting calcium oscillators in plant cells. Trends Plant Sci. 6: 395-397.
- Hata, S., Kouchi, H., Suzuka, I., et Ishii, T. (**1991**). Isolation and characterization of cDNA clones for plant cyclins. *EMBO J.* **10**: 2681-2688.
- He, J. X., Gendron, J. M., Yang, Y., Li, J., et Wang, Z. Y. (**2002**). The GSK3-like kinase BIN2 phosphorylates and destabilizes BZR1, a positive regulator of the brassinosteroid signaling pathway in *Arabidopsis. Proc. Natl. Acad. Sci. U. S. A* **99**: 10185-10190.
- He, Z., Wang, Z. Y., Li, J., Zhu, Q., Lamb, C., Ronald, P., et Chory, J. (2000). Perception of brassinosteroids by the extracellular domain of the receptor kinase BRI1. *Science* 288: 2360-2363.
- He, Z. H., Fujiki, M., et Kohorn, B. D. (**1996**). A cell wall-associated, receptor-like protein kinase. *J. Biol. Chem.* **271**: 19789-19793.
- He, Z. H., Cheeseman, I., He, D., et Kohorn, B. D. (1999). A cluster of five cell wall-associated receptor kinase genes, *Wak1-5*, are expressed in specific organs of *Arabidopsis*. *Plant Mol. Biol.* 39: 1189-1196.
- Heberle-Bors, E. (1989). Isolated pollen culture in tabacco: plant reproductive development in a nutshell. *Sexual Plant Reproduction* 2: 1-10.
- Hecht, V., Vielle-Calzada, J. P., Hartog, M. V., Schmidt, E. D., Boutilier, K., Grossniklaus, U., et de Vries, S. C. (2001). The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is

expressed in developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiology* **127**: 803-816.

- Heigwehr, K. M. G., Banerjee, N., Van Nerum, K., et De Langhe, E. (1985). Somatic embryogenesis and plant regeneration in *Cichorium intybus* L. (Witloof, Compositae). *Plant Cell Rep.* 4: 108-111.
- Helleboid, S., Bauw, G., Belingheri, L., Vasseur, J., et Hilbert, J. L. (**1998**). Extracellular beta-1,3glucanases are induced during early somatic embryogenesis in *Cichorium*. *Planta* **205**: 56-63.
- Helleboid, S., Chapman, A., Hendriks, T., Inze, D., Vasseur, J., et Hilbert, J. L. (**2000**). Cloning of beta-1,3-glucanases expressed during *Cichorium* somatic embryogenesis. *Plant Molecular Biology* **42**: 377-386.
- Heller, R. (**1953**). Recherches sur la nutrition minérale des tissus végétaux cultivés *in vitro. Ann. Sci. Nat. Bot. Biol. Veg.* **14**: 1-23.
- Hendriks, T., Scheer, I., Quillet, M. C., Randoux, B., Delbreil, B., Vasseur, J., et Hilbert, J. L. (**1998**). A nonsymbiotic hemoglobin gene is expressed during somatic embryogenesis in *Cichorium. Biochim. Biophys. Acta* **1443**: 193-197.
- Heuer, S., Hansen, S., Bantin, J., Brettschneider, R., Kranz, E., Lorz, H., et Dresselhaus, T. (2001). The maize MADS box gene *ZmMADS3* affects node number and spikelet development and is coexpressed with *ZmMADS1* during flower development, in egg cells, and early embryogenesis. *Plant Physiology* 127: 33-45.
- Higashi, K., Shiota, H., et Kamada, H. (**1998**). Patterns of expression of the genes for glutamine synthetase isoforms during somatic and zygotic embryogenesis in carrot. *Plant and Cell Physiology* **39**: 418-424.
- Hilbert, J. L., Dubois, T., et Vasseur, J. (1992). Characterization of embryogenesis related proteins and their use as markers during somatic embryo formation in *Cichorium. Plant Physiol. Biochem.* 30: 733-741.
- Hirt, H., Pay, A., Gyorgyey, J., Bako, L., Nemeth, K., Bogre, L., Schweyen, R. J., Heberle-Bors, E., et Dudits, D. (1991). Complementation of a yeast cell cycle mutant by an alfalfa cDNA encoding a protein kinase homologous to p34cdc2. *Proc. Natl. Acad. Sci. U. S. A.* 88: 1636-1640.
- Holk, A., Kaldenhoff, R., et Richter, G. (**1996**). Regulation of an embryogenic carrot gene (DC 2.15) and identification of its active promoter sites. *Plant Mol. Biol.* **31**: 1153-1161.
- Howell, S. H. (1998) Molecular genetics of plant development. Cambridge university press, New York.
- Ivanova, A., Velcheva, M., Denchev, P., Atanassov, A., et Van Onckelen, H. A. (**1994**). Endogenous hormone levels during direct somatic embryogenesis in *Medicago falcata. Physiol. Plant.* **92**: 85-89.

- Jeong, S., Trotochaud, A. E., et Clark, S. E. (**1999**). The *Arabidopsis* CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* **11**: 1925-1934.
- Jimenez, V. M. et Bangerth, F. (2001). *In vitro* culture and endogenous hormone levels in immature zygotic embryos, endosperm and callus cultures of normal and high-lysine barley genotypes. *J. Appl. Bot. Angew. Bot.* **75**: 1-7.
- Jin, P., Guo, T., et Becraft, P. W. (**2000**). The maize CR4 receptor-like kinase mediates a growth factor-like differentiation response. *Genesis*. **27**: 104-116.
- Jinn, T. L., Stone, J. M., et Walker, J. C. (**2000**). HAESA, an *Arabidopsis* leucine-rich repeat receptor kinase, controls floral organ abscission. *Genes Dev.* **14**: 108-117.
- Jurgens, G., Grebe, M., et Steinmann, T. (1997). Establishment of cell polarity during early plant development. *Curr. Opin. Cell Biol.* 9: 849-852.
- Kachroo, A., Schopfer, C. R., Nasrallah, M. E., et Nasrallah, J. B. (2001). Allele-specific receptorligand interactions in *Brassica* self-incompatibility. *Science* 293: 1824-1826.
- Kapros, T., Bogre, L., Nemeth, K., Bak, L., Gyorgyey, J., Wu, S. C., et Dudits, D. (1992). Differential expression of histone H3 gene variants during cell cycle and somatic embryogenesis in alfalfa. *Plant Physiology* 98: 621-625.
- Kawahara, R., Sunabori, S., Fukuda, H., et Komamine, A. (1992). A gene expressed preferentially in the globular stage of somatic embryogenesis encodes elongation-factor 1 alpha in carrot. *Eur. J. Biochem.* 209: 157-162.
- Kawahara, R., Komamine, A., et Fukuda, H. (1995). Isolation and characterization of homeoboxcontaining genes of carrot. *Plant Mol. Biol.* 27: 155-164.
- Kayes, J. M. et Clark, S. E. (1998). CLAVATA2, a regulator of meristem and organ development in *Arabidopsis. Development* 125: 3843-3851.
- Kim, C., Jeong, D. H., et An, G. (2000a). Molecular cloning and characterization of *OsLRK1* encoding a putative receptor-like protein kinase from *Oryza sativa*. *Plant Science* **152**: 17-26.
- Kim, Y. S., Lee, J. H., Yoon, G. M., Cho, H. S., Park, S. W., Suh, M. C., Choi, D., Ha, H. J., Liu, J. R., et Pai, H. S. (2000b). CHRK1, a chitinase-related receptor-like kinase in tobacco. *Plant Physiology* 123: 905-915.
- Kitamiya, E., Suzuki, S., Sano, T., et Nagata, T. (**2000**). Isolation of two genes that were induced upon the initiation of somatic embryogenesis on carrot hypocotyls by high concentrations of 2,4-D. *Plant Cell Rep.* **19**: 551-557.

- Kiyosue, T., Yamaguchi-Shinozaki, K., Shinozaki, K., Higashi, K., Satoh, S., Kamada, H., et Harada, H. (1992). Isolation and characterization of a cDNA that encodes ECP31, an embryogenic-cell protein from carrot. *Plant Mol. Biol.* 19: 239-249.
- Kiyosue, T., Yamaguchi-Shinozaki, K., Shinozaki, K., Kamada, H., et Harada, H. (**1993**). cDNA cloning of ECP40, an embryogenic-cell protein in carrot, and its expression during somatic and zygotic embryogenesis. *Plant Mol. Biol.* **21**: 1053-1068.
- Kiyosue, T., Shiota, H., Higashi, K., Kamada, H., et Shinozaki, K. (1998). A chromo box gene from carrot (*Daucus carota* L.): its cDNA structure and expression during somatic and zygotic embryogenesis. *Biochim. Biophys. Acta* 1398: 42-46.
- Knox, J. P. (**1995**). The extracellular matrix in higher plants. 4. Developmentally regulated proteoglycans and glycoproteins of the plant cell surface. *FASEB J.* **9**: 1004-1012.
- Kobe, B. et Deisenhofer, J. (**1995**). Proteins with leucine-rich repeats. *Curr. Opin. Struct. Biol.* **5**: 409-416.
- Kohorn, B. D., Lane, S., et Smith, T. A. (1992). An Arabidopsis serine/threonine kinase homologue with an epidermal growth factor repeat selected in yeast for its specificity for a thylakoid membrane protein. Proc. Natl. Acad. Sci. U. S. A. 89: 10989-10992.
- Komamine, A., Kawahara, R., Matsumoto, M., Sunabori, S., Toya, T., Fujiwara, A., Tsukahara, M., Smith, J., Ito, M., Fukuda, H., Nomura, K., et Fujimura, T. (1990). Mechanisms of somatic embryogenesis in cell cultures: physiology, biochemistry and molecular biology. *In Vitro Cellular & Developmental Biology - Plant* 28: 11-14.
- Kragh, K. M., Hendriks, T., de Jong, A. J., Lo, S. F., Bucherna, N., Hojrup, P., Mikkelsen, J. D., et de Vries, S. C. (1996). Characterization of chitinases able to rescue somatic embryos of the temperature-sensitive carrot variant *ts 11*. *Plant Mol. Biol.* **31**: 631-645.
- Kusaba, M., Dwyer, K., Hendershot, J., Vrebalov, J., Nasrallah, J. B., et Nasrallah, M. E. (2001). Selfincompatibility in the genus *Arabidopsis*: Characterization of the S locus in the outcrossing *A- lyrata* and its autogamous relative *A-thaliana*. *Plant Cell* **13**: 627-643.
- Laemmli U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685.
- Lally, D., Ingmire, P., Tong, H. Y., et He, Z. H. (2001). Antisense expression of a cell wall-associated protein kinase, wak4, inhibits cell elongation and alters morphology. *Plant Cell* 13: 1317-1332.
- Landschulz, W. H., Johnson, P. F., et McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240: 1759-1764.

- Laux, T., Mayer, K. F., Berger, J., et Jurgens, G. (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in *Arabidopsis*. *Development* 122: 87-96.
- Lease, K. A., Lau, N. Y., Schuster, R. A., Torii, K. U., et Walker, J. C. (2001). Receptor serine/threonine protein kinases in signalling: analysis of the erecta receptor-like kinase of *Arabidopsis thaliana*. *New Phytol.* **151**: 133-143.
- Lee, H. S., Karunanandaa, B., McCubbin, A., Gilroy, S., et Kao, T. H. (**1996**). PRK1, a receptor-like kinase of *Petunia inflata*, is essential for postmeiotic development of pollen. *Plant J.* **9**: 613-624.
- Lee, H. S., Chung, Y. Y., Das, C., Karunanandaa, B., et van Went, J. L. (**1997**). Embryo sac development is affected in *Petunia inflata* plants transformed with an antisense gene encoding the extracellular domain of receptor kinase PRK1. *Sexual Plant Reproduction* **10**: 341-350.
- Li, J., Nagpal, P., Vitart, V., McMorris, T. C., et Chory, J. (**1996**). A role for brassinosteroids in lightdependent development of *Arabidopsis. Science* **272**: 398-401.
- Li, J. et Chory, J. (**1997**). A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction [see comments]. *Cell* **90**: 929-938.
- Li, J., Lease, K. A., Tax, F. E., et Walker, J. C. (2001). BRS1, a serine carboxypeptidase, regulates BRI1 signaling in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 98: 5916-5921.
- Li, J. et Nam, K. H. (2000). Regulation of brassinosteroid signaling by a GSK3/SHAGGY-like kinase. *Science* 295: 1299-1301.
- Li, J., Wen, J., Lease, K. A., Doke, J. T., Tax, F. E., et Walker, J. C. (**2002**). BAK1, an *Arabidopsis* LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell* **110**: 213-222.
- Lichtenthaler, H. K. (1998). The stress concept in plants: an introduction. *Ann. N. Y. Acad. Sci.* 851: 187-198.
- Lichter, R. (**1982**). Induction of haploid plants from isolated pollen of *Brassica napus*. *Z. Pflanzenphysiol.* **105**: 427-434.
- Lindzen, E. et Choi, J. H. (**1995**). A carrot cDNA encoding an atypical protein kinase homologous to plant calcium-dependent protein kinases. *Plant Mol. Biol.* **28**: 785-797.
- Ling, V., Perera, I., et Zielinski, R. E. (**1991**). Primary structures of *Arabidopsis* calmodulin isoforms deduced from the sequences of cDNA clones. *Plant Physiology* **96**: 1196-1202.
- Livak, K. J. et Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402-408.

- Lopes, M. A. et Larkins, B. A. (1993). Endosperm origin, development, and function. *Plant Cell* 5: 1383-1399.
- Lowry, O. H., Rosebrough, N. J., Farr, A. L., et Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193: 265-275.
- Ma, H., McMullen, M. D., et Finer, J. J. (**1994**). Identification of a homeobox-containing gene with enhanced expression during soybean (Glycine max L.) somatic embryo development. *Plant Mol. Biol.* **24**: 465-473.
- Maheswaran, G. et Williams, E. G. (1985). Origin and development of somatic embryos formed directly on immature embryos of *Trifolium repens in vitro*. *Ann. Bot.* 56: 619-630.
- Mansfield, S. G. et Briarty L.G. (1990). Early embryogenesis in *Arabidopsis thaliana*. II. The developing embryo. *Can. J. Bot.* 69: 447-460.
- Maruyama, K. et Sugano, S. (**1994**). Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**: 171-174.
- Matsubayashi, Y., Ogawa, M., Morita, A., et Sakagami, Y. (**2002**). An LRR receptor kinase involved in perception of a peptide plant hormone, phytosulfokine. *Science* **296**: 1470-1472.
- Mayer, K. F., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G., et Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell* 95: 805-815.
- Mazzurco, M., Sulaman, W., Elina, H., Cock, J. M., et Goring, D. R. (**2001**). Further analysis of the interactions between the *Brassica* S receptor kinase and three interacting proteins (ARC1, THL1 and THL2) in the yeast two-hybrid system. *Plant Molecular Biology* **45**: 365-376.
- McCabe, P. F., Valentine, T. A., Forsberg, L. S., et Pennell, R. I. (**1997**). Soluble Signals from Cells Identified at the Cell Wall Establish a Developmental Pathway in Carrot. *Plant Cell* **9**: 2225-2241.
- McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., et Gehring, W. J. (**1984**). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* **308**: 428-433.
- Michalczuk, L., Ribnicky, D. M., Cooke, T. J., et Cohen, J. D. (**1992**). Regulation of indole-3-acetic acid biosynthetic pathways in carrot cell cultures. *Plant Physiology* **100**: 1346-1353.
- Michaux-Ferrière, N. and Schwendiman J. (1992) Histology of somatic embryogenesis. Dans, *Reproductive biology and plant breeding* (pp 247-259), par Dattée Y., Dumas, C., and Gallais A. Editions Springer-Verlag, Berlin.

- Mora-Garcia, S., Vert, G., Yin, Y., Cano-Delgado, A., Cheong, H., et Chory, J. (**2004**). Nuclear protein phosphatases with Kelch-repeat domains modulate the response to brassinosteroids in *Arabidopsis*. *Genes Dev.* **18**: 448-460.
- Mordhorst, A. P., Toonen, M. A. J., et de Vries, S. C. (1997). Plant embryogenesis. *Critical Reviews in Plant Sciences* 16: 535-576.
- Mu, J. H., Lee, H. S., et Kao, T. H. (**1994**). Characterization of a pollen-expressed receptor-like kinase gene of *Petunia inflata* and the activity of its encoded kinase. *Plant Cell* **6**: 709-721.
- Murashige, T. et Skoog, F. (**1962**). Revised medium for rapid growth and bioassays with tobacco tissue culture. *Plant Physiology* **15**: 473-477.
- Nam, K. H. et Li, J. (2002). BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. *Cell* 110: 203-212.
- Nasrallah, J. B. (1997). Signal perception and response in the interactions of self-incompatibility in *Brassica. Essays Biochem.* 32: 143-160.
- Nishiwaki, M., Fujino, K., Koda, Y., Masuda, K., et Kikuta, Y. (2000). Somatic embryogenesis induced by the simple application of abscisic acid to carrot (*Daucus carota* L.) seedlings in culture. *Planta* 211: 756-759.
- Nolan, K. E., Irwanto, R. R., et Rose, R. J. (2003). Auxin up-regulates *MtSERK1* expression in both *Medicago truncatula* root-forming and embryogenic cultures. *Plant Physiology* 133: 218-230.
- Nomura, K. et Komamine, A. (**1985**). Identification and isolation of single cells that produce somatic embryos at a high frequency in a carrot cell suspension culture. *Plant Physiology* **79**: 988-991.
- Nothnagel, E. A. (**1997**). Proteoglycans and related components in plant cells. *Int. Rev. Cytol.* **174**: 195-291.
- O'Shea, E. K., Rutkowski, R., Stafford, W. F., III, et Kim, P. S. (**1989**). Preferential heterodimer formation by isolated leucine zippers from fos and jun. *Science* **245**: 646-648.
- Ochman, H., Gerber, A. S., et Hartl, D. L. (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics* 120: 621-623.
- Oh, M. H., Ray, W. K., Huber, S. C., Asara, J. M., Gage, D. A., et Clouse, S. D. (**2000**). Recombinant brassinosteroid insensitive **1** receptor-like kinase autophosphorylates on serine and threonine residues and phosphorylates a conserved peptide motif in vitro. *Plant Physiology* **124**: 751-765.

- Osmond, R. I., Hrmova, M., Fontaine, F., Imberty, A., et Fincher, G. B. (2001). Binding interactions between barley thaumatin-like proteins and (1,3)-beta-D-glucans. Kinetics, specificity, structural analysis and biological implications. *Eur. J. Biochem.* 268: 4190-4199.
- Overvoorde, P. J. et Grimes, H. D. (1994). The role of calcium and calmodulin in carrot somatic embryogenesis. *Plant and Cell Physiology* 35: 135-144.
- Park, A. R., Cho, S. K., Yun, U. J., Jin, M. Y., Lee, S. H., SachettoMartins, G., et Park, O. K. (2001). Interaction of the *Arabidopsis* receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. *J. Biol. Chem.* 276: 26688-26693.
- Pasternak, T. P., Prinsen, E., Ayaydin, F., Miskolczi, P., Potters, G., Asard, H., Van Onckelen, H. A., Dudits, D., et Feher, A. (2002). The Role of auxin, pH, and stress in the activation of embryogenic cell division in leaf protoplast-derived cells of alfalfa. *Plant Physiology* 129: 1807-1819.
- Pennell, R. I., Janniche, L., Scofield, G. N., Booij, H., de Vries, S. C., et Roberts, K. (1992). Identification of a transitional cell state in the developmental pathway to carrot somatic embryogenesis. *J. Cell Biol.* 119: 1371-1380.
- Perry, S. E., Lehti, M. D., et Fernandez, D. E. (1999). The MADS-domain protein AGAMOUS-like 15 accumulates in embryonic tissues with diverse origins. *Plant Physiology* 120: 121-130.
- Plowman, G. D., Sudarsanam, S., Bingham, J., Whyte, D., et Hunter, T. (1999). The protein kinases of Caenorhabditis elegans: a model for signal transduction in multicellular organisms. *Proc. Natl. Acad. Sci. U. S. A.* 96: 13603-13610.
- Poovaiah, B. W. et Reddy, A. S. (1993). Calcium and signal transduction in plants. *CRC Crit. Rev. Plant Sci.* 12: 185-211.
- Poulain, J. (2003). Etude du rôle de gènes controllant le cycle cellulaire au cours du développement racinaire de Cichorium intybus L. Isolement et Caratérisation d'une cycline mitotique de type B de chicorée. Thèse Université Lille1
- Reinert, J. (**1958**). Morphogenese und ihre kontroll an gewebculturen aus carotten. *Naturwissenschaften* **45**: 344-345.
- Robatche-Claive, A. S., Couillerot, J. P., Dubois, J., Dubois, T., et Vasseur, J. (**1992**). Embryogenèse somatique directe dans les feuilles de *Cichorium* hybride 474: synchronisation de l'induction. *C. R. Acad. Sci. Paris* **314**: 317-371.
- Robatzek, S. et Somssich, I. E. (2001). A new member of the *Arabidopsis* WRKY transcription factor family, AtWRKY6, is associated with both senescence- and defence-related processes. *Plant J.* 28: 123-133.

- Robatzek, S. et Somssich, I. E. (2002). Targets of AtWRKY6 regulation during plant senescence and pathogen defense. *Genes Dev.* 16: 1139-1149.
- Rose, T. M., Schultz, E. R., Henikoff, J. G., Pietrokovski, S., McCallum, C. M., et Henikoff, S. (**1998**). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucl. Acids. Res.* **26**: 1628-1635.
- Sagare, A. P., Lee, Y. L., Lin, T. C., Chen, C. C., et Tsay, H. S. (**2000**). Cytokinin-induced somatic embryogenesis and plant regeneration in *Corydalis yanhusuo* (Fumariaceae) a medicinal plant. *Plant Sci.* **160**: 139-147.
- Saitou, N. et Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
- Sato, S., Toya, T., Kawahara, R., Whittier, R. F., Fukuda, H., et Komamine, A. (1995). Isolation of a carrot gene expressed specifically during early-stage somatic embryogenesis. *Plant Mol. Biol.* 28: 39-46.
- Scheres, B. et Benfey, P. N. (1999). Asymetric cell division in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50: 505-537.
- Schmidt, E. D., Guzzo, F., Toonen, M. A., et de Vries, S. C. (**1997**). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* **124**: 2049-2062.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F., Jurgens, G., et Laux, T. (**2000**). The stem cell population of *Arabidopsis* shoot meristems in maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. *Cell* **100**: 635-644.
- Schopfer, C. R., Nasrallah, M. E., et Nasrallah, J. B. (1999). The male determinant of selfincompatibility in *Brassica. Science* 286: 1697-1700.
- Schulze-Muth, P., Irmler, S., Schroder, G., et Schroder, J. (1996). Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*). cDNA, gene, intramolecular autophosphorylation, and identification of a threonine important for auto- and substrate phosphorylation. *J. Biol. Chem.* 271: 26684-26689.
- Senger, S., Mock, H. P., Conrad, U., et Manteuffel, R. (2001). Immunomodulation of ABA function affects early events in somatic embryo development. *Plant Cell Rep.* 20: 112-120.
- Shah, K., Gadella, T. W., Jr., van Erp, H., Hecht, V., et de Vries, S. C. (2001a). Subcellular Localization and Oligomerization of the *Arabidopsis thaliana* Somatic Embryogenesis Receptor Kinase 1 Protein. J. Mol. Biol. 309: 641-655.

- Shah, K., Schmidt, E. D., Vlak, J. M., et de Vries, S. C. (2001b). Expression of the *Daucus carota* somatic embryogenesis receptor kinase (DcSERK) protein in insect cells. *Biochimie* 83: 415-421.
- Shah, K., Vervoort, J., et de Vries, S. C. (2001c). Role of Threonines in the Arabidopsis thaliana
 Somatic Embryogenesis Receptor Kinase 1 Activation Loop in Phosphorylation. J. Biol. Chem. 276: 41263-41269.
- Shiba, H., Kimura, N., Takayama, S., Hinata, K., Suzuki, A., et Isogai, A. (2000). Alteration of the selfincompatibility phenotype in *Brassica* by transformation of the antisense *SLG* gene. *Biosci. Biotechnol. Biochem.* 64: 1016-1024.
- Shiota, H., Satoh, R., Watabe, K., Harada, H., et Kamada, H. (**1998**). C-ABI3, the carrot homologue of the *Arabidopsis* ABI3, is expressed during both zygotic and somatic embryogenesis and functions in the regulation of embryo-specific ABA-inducible genes. *Plant and Cell Physiology* **39**: **1184-1193**.
- Shiu, S. H. et Bleecker, A. B. (2001a). Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci. STKE.* 2001: RE22
- Shiu, S. H. et Bleecker, A. B. (2001b). Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *Proc. Natl. Acad. Sci. U. S. A.* 98: 10763-10768.
- Shiu, S. H., Karlowski, W. M., Pan, R., Tzeng, Y. H., Mayer, K. F., et Li, W. H. (2004). Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* 16: 1220-1234.
- Silva, N. F., Stone, S. L., Christie, L. N., Sulaman, W., Nazarian, K. A., Burnett, L. A., Arnoldo, M. A., Rothstein, S. J., et Goring, D. R. (2001). Expression of the S receptor kinase in self-compatible Brassica napus cv. Westar leads to the allele-specific rejection of self-incompatible Brassica napus pollen. *Mol. Genet. Genomics* 265: 552-559.
- Smith, C. M., Shindyalov, I. N., Veretnik, S., Gribskov, M., Taylor, S. S., Ten Eyck, L. F., et Bourne, P. E. (**1997**). The protein kinase resource. *Trends Biochem. Sci.* **22** : 444-446.
- Somleva, M. N., Schmidt, E. D. L., et deVries, S. C. (2000). Embryogenic cells in *Dactylis glomerata* L. (*Poaceae*) explants identified by cell tracking and by *SERK* expression. *Plant Cell Rep.* **19**: 718-726.
- Stahl, R. J., Arnoldo, M., Glavin, T. L., Goring, D. R., et Rothstein, S. J. (1998). The self-incompatibility phenotype in *Brassica* is altered by the transformation of a mutant S locus receptor kinase. *Plant Cell* 10: 209-218.
- Stein, J. C., Howlett, B., Boyes, D. C., Nasrallah, M. E., et Nasrallah, J. B. (1991). Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea. Proc. Natl. Acad. Sci. U. S. A.* 88: 8816-8820.

- Stephenson, A. G., Doughty, J., Dixon, S., Elleman, C., Hiscok, S., et Dickinson, H. G. (1997). The male determinant of self-incompatibility in *Brassica oleacera* is located in the pollen coating. *Plant J.* 12: 1351-1359.
- Sterk, P., Booij, H., Schellekens, G. A., van Kammen, A., et de Vries, S. C. (1991). Cell-specific expression of the carrot EP2 lipid transfer protein gene. *Plant Cell* 3: 907-921.
- Stone, J. M., Collinge, M. A., Smith, R. D., Horn, M. A., et Walker, J. C. (**1994**). Interaction of a protein phosphatase with an *Arabidopsis* serine-threonine receptor kinase. *Science* **266**: 793-795.
- Stone, J. M., Trotochaud, A. E., Walker, J. C., et Clark, S. E. (1998). Control of meristem development by CLAVATA1 receptor kinase and kinase- associated protein phosphatase interactions. *Plant Physiology* 117: 1217-1225.
- Stone, S. L., Arnoldo, M., et Goring, D. R. (**1999**). A breakdown of *Brassica* self-incompatibility in *ARC1* antisense transgenic plants. *Science* **286**: 1729-1731.
- Stone, S. L., Anderson, E. M., Mullen, R. T., et Goring, D. R. (**2003**). ARC1 is an E3 ubiquitin ligase and promotes the ubiquitination of proteins during the rejection of self-incompatible *Brassica* pollen. *Plant Cell* **15**: 885-898.
- Suzuki, T., Kusaba, M., Matsushita, M., Okazaki, K., et Nishio, T. (**2000**). Characterization of *Brassica* S-haplotypes lacking S-locus glycoprotein. *FEBS Lett.* **482**: 102-108.
- Szekeres, M., Nemeth, K., Koncz-Kalman, Z., Mathur, J., Kauschmann, A., Altmann, T., Redei, G. P., Nagy, F., Schell, J., et Koncz, C. (**1996**). Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in *Arabidopsis*. *Cell* **85**: 171-182.

Tabaeizadeh, Z. (1998). Drought-induced responses in plant cells. Int. Rev. Cytol. 182: 193-247.

- Takasaki, T., Hatakeyama, K., Suzuki, G., Watanabe, M., Isogai, A., et Hinata, K. (**2000**). The S receptor kinase determines self-incompatibility in Brassica stigma. *Nature* **403**: 913-916.
- Takayama, S., Shiba, H., Iwano, M., Shimosato, H., Che, F. S., Kai, N., Watanabe, M., Suzuki, G., Hinata, K., et Isogai, A. (2000). The pollen determinant of self-incompatibility in Brassica campestris. *Proc. Natl. Acad. Sci. U. S. A* 97: 1920-1925.
- Takayama, S., Shimosato, H., Shiba, H., Funato, M., Che, F. S., Watanabe, M., Iwano, M., et Isogai,
 A. (2001). Direct ligand-receptor complex interaction controls Brassica self- incompatibility. *Nature* 413: 534-538.
- Taylor, R. L. (1967). The foliar embryos of Malaxis paludosa. Can. J. Bot. 45: 1553-1556.

- Thibaud-Nissen, F., Shealy, R. T., Khanna, A., et Vodkin, L. O. (2003). Clustering of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean. *Plant Physiology* 132: 118-136.
- Thomas, C., Meyer, D., Himber, C., et Steinmetz, A. (2004). Spatial expression of a sunflower SERK gene during induction of somatic embryogenesis and shoot organogenesis. *Plant Physiol. Biochem.* 42: 35-42.
- Thomas, T. L. (**1993**). Gene expression during plant embryogenesis and germination: an overview. *Plant Cell* **5**: 1401-1410.
- Thompson, H. J. M. et Knox, J. P. (**1998**). Stage-specific responses of embryogenic carrot cell suspension cultures to arabinogalactan protein-binding ß-glucosyl Yariv reagent. *Planta* **205**: 32-38.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., et Higgins, D. G. (**1997**). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids. Res.* **25**: 4876-4882.
- Tichtinsky, G., Vanoosthuyse, V., Cock, J. M., et Gaude, T. (**2003**). Making inroads into plant receptor kinase signalling pathways. *Trends in Plant Science* **8**: 231-237.
- Toonen, M. A. J., Hendriks, T., Schmidt, E. D. L., Verhoeven, H. A., van Kammen, A., et de Vries, S.
 C. (1994). Description of somatic-embryo-forming single cells in carrot suspension cultures employing video cell tracking. *Planta* 194: 565-572.
- Toonen, M. A., Verhees, J. A., Schmidt, E. D., van Kammen, A., et de Vries, S. C. (**1997**). AtLTP1 luciferase expression during carrot somatic embryogenesis. *Plant J.* **12**: 1213-1221.
- Tordai, H., Banyai, L., et Patthy, L. (1999). The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins. *FEBS Lett.* **461**: 63-67.
- Torii, K. U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R. F., et Komeda, Y. (**1996**). The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* **8**: 735-746.
- Trotochaud, A. E., Hao, T., Wu, G., Yang, Z., et Clark, S. E. (**1999**). The CLAVATA1 receptor-like kinase requires CLAVATA3 for its assembly into a signaling complex that includes KAPP and a Rho-related protein. *Plant Cell* **11**: 393-406.
- Trotochaud, A. E., Jeong, S., et Clark, S. E. (2000). CLAVATA3, a multimeric ligand for the CLAVATA1 receptor-kinase. *Science* 289: 613-617.

- Tucker, M. R., Araujo, A. C., Paech, N. A., Hecht, V., Schmidt, E. D., Rossell, J. B., de Vries, S. C., et Koltunow, A. M. (2003). Sexual and apomictic reproduction in Hieracium subgenus pilosella are closely interrelated developmental pathways. *Plant Cell* 15: 1524-1537.
- Tutin, T. G., Heywood, V. H., Burges, N. A., Valentine, D. H., Walter, S. M., and Webb, D. A. (1964) Flora Europaea. Cambridge University Press, New York.
- Ulrich, T. U., Wurtele, E. S., et Nikolau, B. J. (**1990**). Sequence of EMB-1, an mRNA accumulating specifically in embryos of carrot. *Nucleic Acids Res.* **18**: 2826-
- Valon, C., Smalle, J., Goodman, H. M., et Giraudat, J. (1993). Characterization of an Arabidopsis thaliana gene (TMKL1) encoding a putative transmembrane protein with an unusual kinase-like domain. Plant Mol. Biol. 23: 415-421.
- van der Knaap, E., Song, W. Y., Ruan, D. L., Sauter, M., Ronald, P. C., et Kende, H. (**1999**). Expression of a gibberellin-induced leucine-rich repeat receptor-like protein kinase in deepwater rice and its interaction with kinase-associated protein phosphatase. *Plant Physiology* **120**: 559-570.
- van Engelen, F. A. et de Vries, S. C. (1992). Extracellular proteins in plant embryogenesis. *Trends Genet.* 8: 66-70.
- van Engelen, F. A., de Jong, A. J., Meijer, E. A., Kuil, C. W., Meyboom, J. K., Dirkse, W. G., Booij, H., Hartog, M. V., Vandekerckhove, J., et de Vries, S. C. (1995). Purification, immunological characterization and cDNA cloning of a 47 kDa glycoprotein secreted by carrot suspension cells. *Plant Mol. Biol.* 27: 901-910.
- van Hengel, A. J., Guzzo, F., van Kammen, A., et de Vries, S. C. (1998). Expression pattern of the carrot EP3 endochitinase genes in suspension cultures and in developing seeds. *Plant Physiology* 117: 43-53.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3: research0034.1–research0034.11
- Vanoosthuyse, V., Tichtinsky, G., Dumas, C., Gaude, T., et Cock, J. M. (2003). Interaction of calmodulin, a sorting nexin and kinase-associated protein phosphatase with the *Brassica oleracea* S locus receptor kinase. *Plant Physiology* 133: 919-929.
- Vasil, I. K., Hildebrandt, A. C., et Riker, A. J. (**1964**). Endive plantlets from freely suspended cells and cell groups grown *in vitro*. *Science* **146** : 76-77.
- Vasil, I. K. et Hildebrandt, A. C. (1966). Variation of morphogenetic behaviour in plant tissue culture. I. *Cichorium endivia. Am. J. Bot.* 53: 869-

- Vijayaraghvan, M. R. and Prabhakar K. (**1984**) The endosperm. Dans, *Embryology of Angiosperm* (pp 319-376), par Johri, B. M. Editions Springer-Verlag, Berlin.
- Wagner, T. A. et Kohorn, B. D. (2001). Wall-associated kinases are expressed throughout plant development and are required for cell expansion. *Plant Cell* 13: 303-318.
- Walker, J. C. (**1993**). Receptor-like protein kinase genes of *Arabidopsis thaliana*. *The Plant Journal* **3**: 451-456.
- Wang, Z. Y., Seto, H., Fujioka, S., Yoshida, S., et Chory, J. (2001). BRI1 is a critical component of a plasma-membrane receptor for plant steroids. *Nature* 410: 380-383.
- Willats, W. G. et Knox, J. P. (**1996**). A role for arabinogalactan-proteins in plant cell expansion: evidence from studies on the interaction of beta-glucosyl Yariv reagent with seedlings of *Arabidopsis thaliana*. *Plant J.* **9**: 919-925.
- Williams, R. W., Wilson, J. M., et Meyerowitz, E. M. (1997). A possible role for kinase-associated protein phosphatase in the *Arabidopsis* CLAVATA1 signaling pathway. *Proc. Natl. Acad. Sci. U. S. A* 94: 10467-10472.
- Wiweger, M., Farbos, I., Ingouff, M., Lagercrantz, U., et von Arnold, S. (2003). Expression of Chia4-Pa chitinase genes during somatic and zygotic embryo development in Norway spruce (*Picea abies*): similarities and differences between gymnosperm and angiosperm class IV chitinases. *J. Exp. Bot.* 54: 2691-2699.
- Wolf, S., Grsic-Rausch, S., Rausch, T., et Greiner, S. (2003). Identification of pollen-expressed pectin methylesterase inhibitors in *Arabidopsis. FEBS Lett.* 555: 551-555.
- Xu, N. et Bewley, J. D. (1992). Contrasting pattern of somatic and zygotic embryo development in alfalfa (*Medicago sativa* L.) as revealed by scanning electron microscopy. *Plant Cell Rep.* 11: 279-284.
- Yamamuro, C., Ihara, Y., Wu, X., Noguchi, T., Fujioka, S., Takatsuto, S., Ashikari, M., Kitano, H., et Matsuoka, M. (2000). Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and bending of the lamina joint. *Plant Cell* 12: 1591-1606.
- Yarbrough, J. A. (**1932**). Anatomical and developmental studies of the folair embryos of *Bryophyllum calycinuum*. *Am. J. Bot.* **19**: 443-453.
- Yasuda, H., Nakajima, M., Ito, T., Ohwada, T., et Masuda, H. (**2001**). Partial characterization of genes whose transcripts accumulate preferentially in cell clusters at the earliest stage of carrot somatic embryogenesis. *Plant Molecular Biology* **45**: 705-712.

- Yazawa, K., Takahata, K., et Kamada, H. (2004). Isolation of the gene encoding Carrot leafy cotyledon1 and expression analysis during somatic and zygotic embryogenesis. *Plant Physiol. Biochem.* 42: 215-223.
- Yeung, E. C. (1995) Structural and developmental patterns in somatic embryogenesis. Dans, *Embryogenesis in plant* (pp 205-248), par Thorpe, T. A. Editions Kluwer Academic Publishers, Dordrecht.
- Yin, Y., Wang, Z. Y., Mora-Garcia, S., Li, J., Yoshida, S., Asami, T., et Chory, J. (2002). BES1 accumulates in the nucleus in response to brassinosteroids to regulate gene expression and promote stem elongation. *Cell* 109: 181-191.
- Yokoyama, R., Takahashi, T., Kato, A., Torii, K. U., et Komeda, Y. (**1998**). The *Arabidopsis* ERECTA gene is expressed in the shoot apical meristem and organ primordia. *Plant J.* **15**: 301-310.
- Yu, L. P., Miller, A. K., et Clark, S. E. (2003). POLTERGEIST encodes a protein phosphatase 2C that regulates CLAVATA pathways controlling stem cell identity at *Arabidopsis* shoot and flower meristems. *Curr. Biol.* 13: 179-188.
- Yu, L. P., Simon, E. J., Trotochaud, A. E., et Clark, S. E. (2000). POLTERGEIST functions to regulate meristem development downstream of the CLAVATA loci. *Development* 127: 1661-1670.
- Zhang, R. et Walker, J. C. (1993). Structure and expression of the S locus-related genes of maize. *Plant Mol. Biol.* 21: 1171-1174.
- Zhao, D. Z., Wang, G. F., Speal, B., et Ma, H. (**2002a**). The excess microsporocytes1 gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. *Genes Dev.* **16**: 2021-2031.
- Zhao, J., Peng, P., Schmitz, R. J., Decker, A. D., Tax, F. E., et Li, J. (**2002b**). Two putative BIN2 substrates are nuclear components of brassinosteroid signaling. *Plant Physiology* **130**: 1221-1229.
- Zhu, H., Klemic, J. F., Chang, S., Bertone, P., Casamayor, A., Klemic, K. G., Smith, D., Gerstein, M., Reed, M. A., et Snyder, M. (2000). Analysis of yeast protein kinases using protein chips. *Nat. Genet.* 26: 283-289.
- Zimmerman, J. L. (**1993**). Somatic Embryogenesis: A Model for Early Development in Higher Plants. *Plant Cell* **5**: 1411-1423.
- Zuo, J., Niu, Q. W., Frugis, G., et Chua, N. H. (2002). The WUSCHEL gene promotes vegetative-toembryonic transition in *Arabidopsis*. *Plant J.* **30**: 349-359.

Annexes

Annexe 1 : Séquence d'ADNc pleine longueur et protéique déduite de CiSERK1474

A partir d'une séquence connue correspondant au domaine kinase partiel (KD partiel), 3 PCR inverses successives (a, b, c) ont été réalisées à partir de la même matrice d'ADNc autoliguée. Les fragments de séquences obtenus après la première PCR inverse (a) ont permis de définir les amorces nécessaires à la seconde PCR inverse (b) dont produits ont permis de définir les amorces utilisées pour la troisième PCR inverse (c). Chaque fragment est un consensus d'environ 8 séquences. La présence des régions chevauchantes a permis de créer une séquence continue déduite (continuum).

iPCRc iPCRb iPCRa KD partiel iPCRa/b/c	 AAGTTATA	10 AATTCTGGAT	20 	30 ATGGCCGGTTTC	40 	50 . GTTCAAAGCT	60 . ?AGCGTTTGG	70 GTYTTAGGTY	80 . YYTGAATCT	90 CTTGATCTGO	100 . GTTGAGTAATC	110 GATGATGAA	120 . .YYYAAAGCTYTT	130 FAGCTTGACT	140 rcgagttts:	150 PGGGAA
continuum	AAGTTAT	AATTCTGGAT	GTTATGGGT	ATGGCCGGTTT(3CTTGAGCTGT	GTTCAAAGCI	AGCGTTTGG	GTYTTAGGTY	YYTGAATCT	CTTGATCTG	3TTGAGTAATC	GATGATGAA	YYYAAAGCTYT	TAGCTTGAC	rcgagttts:	rgggaa
iPCRc iPCRb	ATGGATCO	160 GAACAATCTC		180 CATTTGGCTTA	190 PTTTCGGGTTC	200 . AATTATGTAT	210 .	220 ATGGAAATGC	230 . CGAAGGTGA	240 CGCATTGAA	250 . IGCCCTAAAGA	260 СССААТТАА	270 . CCGATCCGAATA CCGATCCGAATA T D P N	280 AATATTTTGO AATATTTTGO N I L	290 CAA CAAAGTTGGC Q S W	GATCCG
iPCRa KD partiel iPCRa/b/c																
continuum	AT GG ATCO M D F	GAACAATCTO R T I S	TCTAATCGC LIA	CATTTGGCTTA I W L I	FTTTC GGG TTC. I F G F	AATTATGTAT NYV	CAAAGGTTT SKV	ATGGAAATGC Y G N A	CGAAGGTGA E G D	CGCATTGAAI A L N	PGCCCTAAAGA A L K	CCCAATTAA T Q L	CCGATCCGAAT/ T D P N	ATATTTTG N I L	CAAAGTTGG(Q S W	JATCCG D ₽
iPCRb iPCRa KD partiel iPCRa/b/c	ACCCTGGT	310 TTAATCCATO	320 	330 FCATATTACTTO	340 	350 . AATAGCGTTI	360 . . CTAGACTTG	370 ATCTTGGAAA	380 . TGCAGGCTT 	390 GTCTGGTCA2	400 . ACTGGTTCCAG	410 AGCTTGGTC	420 . AACTITICAAAGI	430 . TTACAGTATO	440 CTGGAGCTT	450 FATGGG
continuum	ACCCTGGI T L \	ITAATCCATC V N ₽ C	TACATGGTT T W F	FCATATTACTTO H I T (GTGACCCTGGG. D P G	AATAGCGTTI N S V	CTAGACTTG S R L	ATCTTGGAAA D L G N	T GCAGGCTT A G L	STCTGGTCAR SGQ	ACTGGTTCCAG	AGCTTGGTC E L G	AACTTTCAAAG Q L S K	TTACAGTATO L Q Y	CTGGAGCTTT L E L	PAT GGG Y G
iPCRb iPCRa KD partiel iPCRa/b/c	 AACAACAT	460 . FCACTGGAAA	470 . .AATTCCTGA.	480 ACAGCTTGGAAA	490 	500 . FTAATGAGTT	510 . TGGATCTTT	520 	530 . ATTAGAAGG	540 . IGACATTCCA	550 . AGACACATTGG	560 GGAACCTTC 	570 . AAGAACTCCGT	580 . TTCCTTCGTC	590 CTCAACAACA	600 \ACACT
continuum	AACAACAT N N I	PCACTGGAAA I T G K	AATTCCTGA TPE	ACAGCTTGGAAA Q L G M	$\mathbf{ACTTAACACAA}^{ACTTAACACAA}$	P TAATGAG TT L M S	TGGATCTTT L D L	ACCTAAACAA Y L N K	ATTAGAAGG L E G	IGACATTCCA D I ₽	AGACACATTGG D T L	GGAACCTTC G N L	AAGAACTCCGT Q E L R	FLR	TCAACAACA L N N	ACACT N T

iPCRb iPCRa KD partiel	610 . TTGACAGGAAC	 TATTCCA	620 	630 PAACTACA	64(ATCGGATC	0 CGCTCCA	650 AGPTCTTC	660 GATCTAT) .	670 \ATAATT	. TAAGTO	580 GGATTGO	690 2TCCAAT'	7 FAATGGG	00 TCATTT	710 PCACTCT	 TCACAC	720 . CGATCA	7 GTTAT	30 GATGGT	740 AACGTCAA	750 . .CTTGCAA
iPCRa/b/c continuum	TTGACAGGAAC	TATTCC7	AATTCATT	ГААСТАСА	ATCGGAT(CGCTCCA	AGTTCTT	GATCTAI	CGCACA	ATAATI	TAAGT	GGATTG	CTCCAAT	TAATGGG	TCATTT	rcactct	TCACAC	CGATCA	GTTAT	GATGGT	AACGTCAA	
iPCRb iPCRa KD partiel iPCRa/b/c	760 . ATTCAAGCTCC	TCCTCC	770 	780 2TGTTTCA(79(. CCTGATTC	0 CTCAATC	800 TCCTTCA(810 . GTCAGCA) AACAGTG	820 	L S	330 ATAGCCGC	840 GAGGAGT [*]	N G 8 PGCCGCT	50 GGTGCA(GCA(860 GCTCTAC GCTCTAC	r i TTTTTG	870 . CAGGT CAGGTC	8 CGGCCG	80 ATTGCA	890 ATTGCTTG	900 • • • • GTATCGG
continuum	ATTCAAGCTCC I Q A P	TCCTCC7 F P	ACAATCACC Q S T	CTGTTTCA(V S	CCTGATTO	CTCAATC S Q S	TCCTTCA PS	GTCAGCA V S	ACAGTO N S	CCACCO A T	GAGCTA	TAGCCGO I A (GAGGAGT" G V	IGCCGCT A A	GGTGCAG G A	GCTCTAC A L	TTTTTG L F	CAGG TC A G	CGGCG	ATTGCA I A	ATTGCTTG I A W	GTATCGG Y R
iPCRa KD partiel iPCRa/b/c	910 . CGCAGAAAACC.	 ACAAGAT	920 . . CATTTCTT	930 TTGATGTAC	94(. . 	0 AAGAGGA'	950 TCCAGAA(960 . . GTCCACC) CTAGGAC	970 CAACTAA	. LAGAGA	980 . TCTCTCT	990 FACGCGA	1 ACTACAA	000 GTCGCAJ	1010 ACCGATA		1020 . GCAACA	1 ATAAC	.030 PATTCTT	1040 GGCCGAGG	1050 . TGGATTT
continuum	CGCAGAAAACC R R K P	ACAAGAI Q D	CATTTCTI H F F	rtgat g ta(PDV	CCCGCCGA	AAGAGGA [.] E E D	TCCAGAAG P E	GTCCACC V H	CTAGGAC L G	CAACTAZ Q L	AGAGA K R	TCTCTCT F S I	PACGCGA	ACTACAA L Q	GTCGCAJ V A	ACCGATA T D	CCTTTA T F	GCAACA S N	ATAAC N N	ATTCTT	GGCCGAGG G R G	/TGGATTT G F
	106 .	0 	1070 . .	1080	109 .	90 	1100 	111 .	.0 	1120	:	130	1140 	1 • • • • • •	150 	1160 		1170 .	1	180 	1190 	1200 .
iPCRa KD partiel iPCRa/b/c	GGTAAAGTTTA	CAAAGGO	CGATTAGC	CCGATGGT	FCTCTGG	FGGCGGT	TAAAAGA(TAAAAGA(CTCAAAG	BAGGAAC BAGGAAC	GTAGTO GTAGTO	CAAGGCO	GCGAATI	rgCAGTT' rgCAGTt'	TCAAACG TCAAACG	GAGGTC(GAGGTT(GAGATGA GAGATGA	TCAGTA TCAGTA	TGGCGG	FTGCAC	CGGAAT CGGAAC	CTTCTCCG	GCTAAGA GCTAAGA
continuum	GGTAAAG TTTA G K V Y	CAAAGGC K G	CGATTAGO R L 2	CCGATGGTI A D G	FCTCTGGI SL	TGGCGGT V A V	TAAAAGA(K R	CTCAAAG L K	AGGAAC E E	CGTAGTO R S	CAAGGCC ହୁତ୍ର	GCGAATI G E I	FGCAGTT L Q F	PCAAACG Q T	GAGGTCO E V	GAGATGA E M	TCAGTA I S	T GGCGG M A	TGCAC V H	CGGAAT R N	CTTCTCCG L L R	GCTAAGA L R
iPCRa KD partiel iPCRa/b/c	121 . GGATTTT GGATTTTGCATC	0 GACACC7	1220 . .	1230 	124 • • • • • GTTTATCC	40 •••• ••	1250 	126 . GGCAGCO	50 STCGCAI	1270 	- Tacgao	280 ATCGACC	1290 CTGATAC	1 • • • • • • • • • • • • • • • • • • •	300 CCGCTTC	1310 GATTGGC	 С G атаа	1320 • • • • • GGAAAC	1 	.330 :GCTTTG	1340 GGAGCCGC	1350 • • • • ACGTGGC GGC
continuum	GGATTTTGCAT	GACACCA T P	ACCGAACG	TCTACTCC R L L	GTTTATCC V Y I	CCTACATO	GGCCAAT(A N	GCAGCO G S	TCGCAT V A	CATGTI S C	TACGAC	ATCGACO	CTGATAC	ACAACCT Q P	CCGCTTO	GATTGGC D ₩	CGATAA P I	GGAAAC R K	GAATC R I	GCTTTG A L	GGAGCCGC G A A	ACGTGGC R G
	1360 1370 1380 1390 1400 1410 1420 1430 1450 1460 1470 1480 1490 1500																					
-------------------------	--																					
KD partiel iPCRa/b/c	CTTGCTŁATŤTACAĊGATCATTGTĠACCCĠAAAAŤTATAĊACCGŤGATGŤGAAAĠĊŁGCÀAATAŤACTAŤTGGAŤGAAGÅGTŁTĠAAGCÅGTTGŤGGGAĠATTTĊGGGTŤGGCTÁAACTĊATGGÀCTATÀAAGAŤACTCÀTGTCÀCGACĂ CTTGCTTATTTACACGATCATTGTGACCCGAAAATTATACACCGTGATGTGAAAGCTGCAAATATACTATTGGATGAAGAGTTTGAAGCAGTTGTGGGAGATTTCGGGTTGGCTAAACTCATGGACTATAAAGAŤACTCATGTCACGACA																					
continuum	CTTGCTLATTTACACGATCATTGTGACCCGAAAATTATACACCGTGATGTGAAAGCLGCAAATATACTATTGGATGAAGAGTLTGAAGCAGTTGTGGGAGATTTCGGGTTGGCTAAACTCATGGACTATAAAGATACTCATGTCACGACA L A Y L H D H C D P K I I H R D V K A A N I L L D E E F E A V V G D F G L A K L M D Y K D T H V T T																					
KD partiel iPCRa/b/c	1510 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630 1640 1650 </th																					
continuum	$ \begin{array}{l} GCTGTACGTGGTACAATTGGACACATAGCACCCGAGTATCTATC$																					
iPCRa/b/c	1660 1670 1680 1690 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800																					
continuum	GACGAGGTCATGTTGCTTGATTGGGTGAAAGGACTATTGACGGATAAGAAGCTTGAGATGTTGGTTG																					
iPCRa/b/c	1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950																					
continuum	ATGGAGCGGCCGAAGATGTCAGAAGTTGTTAGAATGCTTGAAGGGGATGGTTTAGCGGAGAGGTGGGAGGAGGTGGCAGAAGAGGAGATTTCCGGCAAGATTACAACCAGGCACAATCATAATATTGATTG																					
iPCRa/b/c	1960 1970 1980 1990 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100																					
continuum	ACTTATAACCTTAGCAACGAGCAATTGTCGGGTCCTAGATGATTGGATTGGATTCATGTGGGGTGGGGTGTGTGT																					
iPCRa/b/c continuum	2110 2120 2130 2140 2150 2160 2170 2180 2190 2200 2210 2220 																					

Annexe 2 : Séquence d'ADNc pleine longueur et protéique déduite de CiSERK2474

A partir d'une séquence connue correspondant au domaine kinase partiel (KD partiel), 2 PCR inverse successives (a, b) ont été réalisées à partir de la même matrice d'ADNc autoliguée. Les fragments de séquences obtenus après la première PCR inverse (a) ont permis de définir les amorces utilisées pour la seconde PCR inverse (b) dont les produits ont permis de définir les amorces utilisées pour l'obtention de la région 5' par le kit generacer (gr) (Invitrogen). Chaque fragment est un consensus d'environ 8 séquences. La présence des régions chevauchantes a permis de créer une séquence continue déduite (continuum).

gr iPCRb iPCRa KD partiel iPCRa/b	. CATATO	10 		20 . AAGTGG	30 . TTTTCTI) CTTTGG	40 	50 TTAGATTI	GTGGA	60 CGATATA(70 CATAGATT	 CTGCTC	80 . GTTTGAG	90 CTGTCTC	 TAAACC	100 . TAGGGTI	110 TGGTGTT	 TGAGGI	120 . TCTGGG1	130		140 	150 3AATGA
continuum	CATATC	CAAGGTI	PTCTTC	AAGTGG	TTTTCTI	CTTTGC	GTGGTGGA	TTAGATTI	GTGGA	CGATATAC	CATAGATI	CTGCTC	GTTTGAG	CTGTCTC	TAAACC	TAGGGTI	TGGTGTT	TGAGGI	TCTGGGI	PACTTCGA	ATTTC	PTGATCTO	3AAT G A
gr iPCRb iPCRa KD partiel iPCRa/b	. GTAATC	160 GATGAGA	 AAATCT	170 . AGGTTT	180 . TTAGCTI) TACCGA	190 	200 . CCGGATTI	 'Gaaaa <i>i</i>	210 	220 . GCCGAGTC	 TCTCTG	230 . ATTTCAG	240 CTTCTGT'	 TCTTGT	250 . CTGGTTC	260 ATTTTGG	 STATTCA	270 . ACAACTO	280 CGTCAAAA		290 CGGGAATC	300 3CCGAA
continuum	GTAATC	GATGAGA	\AATCI	AGGTTT	TTAGCTI	TACCGA	CCCGAAC	CCGGATTI	GAAAA	ATGGATCO M D H	GCCGAGTC R R V	S L	ATTTCAG I S	CTTCTGT A S V	TCTTGT L V	CTGGTTC W F	ATTTTGG I L	TATTCA V F	ACAACTO	CGTCAAAA S S K	AGTCCAC V H	CGGGAAT(G N	CCGAA A E
gr iPCRb iPCRa KD partiel	. GGCGAT'	310 GCATTGA		320 . TTAAAG	330 . ACTCAAT) ••••••••••••••••••••••••••••••••••••	340 SATCCGAA	350 TAATGTTC	 :TTCAA/	360 AGTTGGG/	370 \TGCGACT		380 . AACCCGT	390 GTACTTG		400 . TGTTACG	410 TGCAATA		420 . MATAGCG1	430	AGTTGA	440 rcttggaa	450 \ACGCG
continuum	GGCGAT G D	GCATTGA A L	AATGCC N A	TTAAAG L K	ACTCAAT T Q	TAGGTO L G	ATCCGAA D P N	FAATGTTC N V	CTTCAA) L Q	AGTTGGG/ S % I	A TGCG ACT D A T	CTTGTA L V	AACCCGT N	GTACTTG C T W	GTTTCA F H	TGTTACC V T	TGCAATA C N	ACGAGA N E	ATAGCGI N S N	TACAAGI / T R	AGTTGA V D	ICTTGGAJ L G	ACGCG N A
gr iPCRb iPCRa KD partiel iPCRa/b	. AAT'TTG 	460 TCAGGTC		470 . GTTGCA	480 . CAACTCG	GTGACC	490 TTCCCTAA	500 . TTTACAGI	 PATTTGO	510 	520 . ACAGTAA1	 AACATT	530 . ACTGGAA	540 AAATTCC		550 . ACTTGGA	560 .AACTTAA	 Caaatti	570 . TAGTGAC	580 . GTTTGGAT	CTTTA	590 rctcaaca	600 \AATTA
continuum	AATTTG N L	T CAGG TC S G	CAACTC Q L	GTTGCA V A	CAACTCG Q L	GTGACC G D	TCCCTAA L P N	rttacagt L Q	ATTIGO Y L	GAACTTTA EL	ACAGTAAT (SN)	AACATT N I	ACTGGAA T G	AAATTCC K I ₽	TGATGC. D A	ACTTGGA L G	AACTTAA N L	CAAATI T N	TAGTGAC L V S	FTTTGGAT 3 L D	CTTTA L Y	PCTCAACA L N	AATTA K L

gr iPCRb iPCRa KD partiel	610 620 630 640 650 660 670 680 690 700 710 720 730 740 750
iPCRa/b	
continuum	GATGGTGGCATTCCTGAAACATTGGGCAAGCTTCAAAAGCTTCGTTTTCTTCGTCTCAATAACAACACGTTAACAGGAACGATTCCGATTTCATTGACTACTACTACACAAGTTCTTGATCTGTCAAAACAACAACAAGGAGGA D G G I F E T L G K L Q K L R F L R L N N N T L T G T I F I S L T T I T S L Q V L D L S N N N L R G
gr iPCRb iPCRa KD partiel iPCRa/b	760 770 780 790 800 810 820 830 840 850 860 870 880 890 900
continuum	GACGTCCCGGTCAATGGTTCCTTTTCGCTTTTCACACCTATAAGTTTTGCCAATAATCCCCAGTTGAAAGCCCCTGCGGTTTCTCCTCAAGCTCCGGCCAACAATTCCCCATCTTCTTCCGTCGGCAACAGCGCGACGGGGGCAATC D V P V N G S F S L F T P I S F A N N P Q L K A P A V S P Q A F A P P N S F S S V G N S A T G A I
gr iPCRb iPCRa KD partiel iPCRa/b	910 920 930 940 950 960 970 980 990 1000 1010 1020 1030 1040 1050
continuum	GCGGGAGGAGTAGCCGCCGGTGCCGCCCCCTCTATTCGCCGGTCGGCGCGCGC
iPCRb iPCRa KD partiel iPCRa/b	TCTCTTCGCGAATTACAAGTGGCAACCGACAA TCTCTTCGCGAATTACAAGTGGCAACCGACAATCTCAACAACCGCCACATTCTCGGGCGGG
continuum	TCTCTTCGCGAATTACAAGTGGCAACCGACAATCTCAACAACCGCCACATTCTCGGGGGGGG
iPCRa KD partiel iPCRa/b	1210 1220 1230 1240 1260 1270 1280 1290 1310 1320 1330 1340 1350
continuum	GAGCTGCAGTTTCAGACGGAGGTGGAGATGATCAGTATGGCGGTTCACCGGAATTTACTCCGGTTGAGAGGGTTTTTGCATGACACCGAACGGTTACTTGTTTATCCATATGGCTAATGGTAGTGTTGCTTCGtGCTTAAGAGAG E L O F O T E V E M I S M A V H R N L L R L R G F C M T P T E R L L V Y P Y M A N G S V A S C L R E

-

KD partiel iPCRa/b	1360 1370 1380 1390 1400 1420 1430 1440 1450 1460 1470 1480 1490 1500
continuum	AGACCCGATACACGAGAACCGCTTGATTGGCCAATACGTAAGAGAATTGCACTGGGGTCCGCAAGAGGACTTGCGLATTTACATGACCATTGTGACCCGAAAATCATTCACCGTGATGTAAAAGCTGCGAATATATTACTAGATGAAGAA R P D T R E P L D W P I R K R I A L G S A R G L A Y L H D H C D P K I I H R D V K A A N I L L D E E
KD partiel iPCRa/b	1510 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630 1640 1650
continuum	TTCGAAGCAGTTGTGGGAGATTTCGGTTTGGCTAAACTTATGGATTATAAAGATACACATGTCACGACAGCtGTCCGTGGTACAATTGGCCATATTGCCCCCGAGTATTTATCAACCGGGAAATCATCGGAAAAAACCGATGTTTTTGGG F E A V V G D F G L A K L M D Y K D T H V T T A V R G T I G H I A P E Y L S T G K S S E K T D V F G
iPCRa/b	1660 1670 1680 1690 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
continuum	TACGGTGTCATGCTTCTAGAACTTATCACGGGACAACGCGCTTTCGATCTTGCAAGACTTGCCAATGATGATGATGTGTCATGTTGCGGGGAAAGGGCCTTTTGAGGGAAAAGAAATTGGAGACATTAGTGGATGCGGATTTGAAA Y G V M L L E L I T G Q R A F D L A R L A N D D D V M L L D W V K G L L R E K K L E T L V D A D L K
iPCRa/b	1810 1820 1830 1840 1850 1870 1880 1890 1910 1920 1930 1940 1950
continuum	GGTAATTATATTGATGAAGTTGAACAACTTATACAAGTGGCACTTTTATGCACACAAGGGACGCCATTAGAAAGACCAAAAATGTCGGAAGTTGTGAGAATGTTGGAAGGTGGGTTAGCGGAAAGGTGGGAAGAATGGCAAAGA G N Y I D D E V E Q L I Q V A L L C T Q G T P L E R P K M S E V V R M L E G D G L A E R W E E W Q R
iPCRa/b	1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
continuum	GAAGAAATGTTTAGACAAGAATTCAATACGACACATAATCCAAATACTGATTGGAGTTATTGCAGACTCAACGTATAACCTTCGTCCTGATGAATTATCTGGTCCCAGATGATTTGGGGTTTAGGGTTTTGGGGTTTTTTTAATGGTTATT E E M F R Q E F N T T H N P N T D W I I A D S T Y N L R P D E L S G P R *
iPCRa/b continuum	2110 2120 2130 2140 2150 2160 2170

í.

Séquence nucléotidique partielle codante du gène d'*actine* utilisé comme contrôle pour les expériences de PCR en temps réel. Les séquences correspondant aux amorces sont figurées par des flèches.

СіАст	ATTCTCCGTCTCGACCTCGCCGTGGCCGTGACCTCACCGACTCCCTAATGAAAATCCTAACC
CiAct	GAAAGAGGTTACATGTTCACCACCACAGCCGAACGGGAAATTGTCCGT GACATGAAAGAG
CiAct	AAACTCGCCTACGTGGCACTCGACTTCGAACAAGAGCTCGAAACCGCCAAATCCAGCTCA
СіАст СіАст СіАст СіАст СіАст	TCAGTCGAAAAGAACTACGAATTACCCGACGGACAGG TCATCACCATTGGAGCCGAAAG A TTCCGGTGCCCGGAGGTCCTCTTCCAGCCGTCTCTCATCGGAATGGAAGCTGCCGGAATT CACGAGACTACTTATAACTCCATCATGAAGTGTGACGTTGATATCAGGAAGGA

Séquence nucléotidique partielle d'ADN ribosomique *18s* de tournesol (N°Acc AF107577) et position des amorces utilisées pour les amplfications par PCR en temps réel. La séquence (figurée par des flèches) de tournesol a été choisie en raison de sa proximité botanique avec la chicorée (Astéracées). Les séquences correspondant aux amorces sont invariables quelle que soit l'espèce d'Astéracées considérée.

TCATATGCTTGTCTCAAAGATTAAGCCATGCATGTGTAAGTATGAACAAATTCAGACTGTGAAACTGCGAATG GCTCATTAAATCAGTTATAGTTTGTTGATGGTATCTTGCTACTCGGATAACCGTAGTAATTCTAGAGCTAATA CGTGCAACAAACCCCCGACTTCTGGAAGGGATGCATTTATTAGATAAAGGTCGACGCGGGCTTTGCCCGTTG CTGCGATGATTCATGATAACTCGACGGATCGCACGGCCCTTGTGCCGGCGACGCATCATTCAAATTTCTGCC CTATCAACTTTCGATGGTAGGATAGTGGCCTACTATGGTGGTGACGGGTGACGGAGAATTAGGGTTCGATTC ACACGGGGGGGGGGTAGTGACAATAACAATAACAATACCGGGCTCATACGAGTCTGGTAATTGGAATGAGTACAATC TAAATCCCTTAACGAGGATCCATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAG CGTATATTTAAGTTGTTGCAGTTAAAAAGCTCGTAGTTGGACTTTGGGTTGGGTCGGCCGGTCCGCCATCAG GTGTGCACCGGTTTACTCGTCCCTTCTGTCGGCGATGCGCTCCTACCCTTAACTGGGCGGGTCGTGCCTCC TCATAGGATTTCGGTCCTATTACGTTGGCCTTCGGGATCGGAGTAATGATTAACAGGGACAGTCGGGGGGCAT TCGTATTTCATAGTCAGAGGTGAAATTCTTGGATTTATGAAAGACGAACAACTGCGAAAGCATTTGCCAAGGA TGTTTTCATTAATCAAGAACGAAAGTTGGGGGGCTCGAAGACGATCAGATACCGTCCTAGTCTCAACCATAAAC GATGCCGACCAGGGATCAGCGGATGTTGCTTTTAGGACTCCGCTGGCACCTTATGAGAAATCAAAGTTTTTG GGTTCCGGGGGGGAGTATGGTCGCAAGGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGG AGCCTGCGGCTTAATTTGACTCAACACGGGGGAAACTTACCAGGTCCAGACATAGTAAGGATTGACAGACTGA GAGCTCTTTCTTGATTCTATGGGTGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATT CCGTTAACGAACGAGACCTCAGCCTGCTAACTAGCTATGTGGAGGTATCCCTCCATGGCCAGCTTCTTAGAG GGACTATGGCCTTTTAGGCCACGGAAGTTTGAGGCAATAACAGGTCTGTGATGCCCTTAGATGTTCTGGGCC GCACGCGCGCTACACTGATGTATTCAACGAGTATATAGCCTTGGCCGACAGGCCCGGGAAATCTTTGAAATT TCATCGTGATGGGGATAGATCATTGCAATTGTTGGTCTTAAACGAGGAATTCCTAGTAAGCGCGAGTCATCAG CTCGCGTTGACTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTCCTACCGATTGAATGGTCCGGTGAAG TGTTAGGATCGTGGCGACRTGGGCGGTTCGCTGCCCGCGACGTCGCGAGAATTCCACTGAACCTTATCATTT AGAGGAAGGAGATCG

Séquence nucléotidique codante et protéique déduite de l'ADNc d'hémoglobine non symbiotique de chicorée (Hendriks *et al.*, 1998)(N°Acc AJ007507) et position des amorces (figurées par des flèches) utilisées en PCR en temps réel.

CHI7507	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
CHI7507	708090100110120GACATTCCTGCTCTTAGTCTCTATCTTTACGCGATGATACTGGAGATAGCCCCGGAAGCADIPALSLYAMILEIAPEA	
CHI7507	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	0
CHI7507	19020021022023024	0
CHI7507	250 260 270 280 290 30 AAAGGTGAGGTTGTGGGTTCTGGGTTCTACACTCAAGTATTTGGGATCCGTTCATCTTGAG K G E V V V S G S T L K Y L G S V H L E	0
CHI7507	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	0
CHI7507	370 380 390 400 410 42	0
CHI7507	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0
CHI7507	490 500 510 520 530 54	0
CHI7507	550 560 570 580 590 60	0

Numéro d'accession GenBank des séquences de type SERK chez différentes espèces.

Les séquences d'ADN, d'ADNc et de protéines chez *Arabidopsis* ont été identifiées au moyen de leur nom de locus disponibles sur le site de "The *Arabidopsis* Information Ressource (TAIR) " (http://www.arabidopsis.org/index.jsp)

- <u> </u>	Gène	ADNc	Protéine	Locus
DcSERK	A67796	U93048		
AtSERK1	A67827			At1g71830
AtSERK2	AC007454			At1g34210
AtSERK3	AL035678			At4g33430
AtSERK4	AC006436			At2g13790
AtSERK5	AC006436			At2g13800
MtSERK1	AY162177	AY162176	AAN64293	
ZmSERK1	AJ400868	AJ277702	CAC37638	
ZmSERK2	AJ400869	AJ277703	CAC37639	