



Université de la Nouvelle-Calédonie  
Ecole doctorale du Pacifique (ED 469)  
Laboratoire Insulaire du Vivant et de l'Environnement

**Doctorat en Physiologie et Biologie des Organismes –  
Populations – Interactions**

Présenté par

**Matthieu VILLEGENTE**

**Caractérisation Biochimique et Moléculaire de Mécanismes de la  
Germination d'Espèces Endémiques de Nouvelle-Calédonie**

Thèse dirigée par **Bruno FOGLIANI**

**Valérie BURTET-SARRAMEGNA**

Thèse soutenue publiquement le 20 novembre 2013 devant le jury composé de :

Françoise CORBINEAU,	Professeur, Université Pierre et Marie Curie	Rapporteur
Ian SMALL,	Professeur, University of Western Australia	Rapporteur
Dominique JOB,	Directeur de Recherche émérite au CNRS Laboratoire mixte CNRS/Bayer CropScience	Examineur
Hamid AMIR,	Professeur, Université de la Nouvelle-Calédonie	Examineur
Bruno FOGLIANI,	Chercheur HDR, Institut Agronomique néo-Calédonien	Directeur de thèse
Valérie BURTET-SARRAMEGNA,	MCF, Université de la Nouvelle-Calédonie	Co-encadrante de thèse







# Remerciements

Ce travail de thèse est le résultat de nombreuses rencontres et collaborations. Je tiens à remercier toutes les personnes rencontrées sur ce chemin et tous les amis qui m'ont aidé à maintenir le cap.

Un grand merci à mes deux encadrants, Bruno Fogliani et Valérie Sarramégna. Merci pour les moments que vous avez pris pour les nombreuses discussions, les conseils, la recherche de financements multiples pour mes missions à l'étranger, l'établissement de collaborations, dans le but de nous permettre d'aboutir au mieux ce projet de thèse. Ces quatre années passées en votre compagnie auront aussi été l'occasion de développer une relation d'amitié dont la valeur m'est inestimable. Merci.

Si je suis dans le domaine du végétal, c'est en partie grâce (à cause) des chercheurs et enseignants passionnés que j'ai rencontrés, notamment lors de mon master à l'Université Pierre et Marie Curie et qui m'ont ainsi « dérouté » de la physiologie de la nutrition animale vers le monde fabuleux du végétal. Merci à Françoise Corbineau, Jean-Pierre Bouly et Eric Ruelland pour cette modification de parcours que vous m'avez inspirée et qui m'a mené jusqu'ici.

Un grand merci à Claudette Job et Dominique Job qui m'ont accueilli comme leur thésard. Merci Claudette pour les heures passées au laboratoire, le goût de la paillasse que tu prends plaisir à transmettre et tout ton investissement même quand j'étais à 22 000 km de Lyon. Un grand merci Dominique, pour ton encadrement, les heures passées sur Skype, le temps consacré à transmettre ta passion pour la recherche, et la rigueur que tu as su m'apprendre.

Je tiens à remercier le département Bayer CropScience pour l'accueil, le soutien technique et financier pour l'aspect de protéomique. Merci à Nathalie Poussereau et toute son équipe pour m'avoir accueilli au sein du laboratoire mixte CNRS/Bayer. Merci à Vincent Girard et Cindy Dieryckx.

Merci également à François Bernier pour les anticorps anti-DING et les conseils et discussion échangés concernant les protéines DING.

Merci à Charlie Scutt et Loïc Rajjou , membres de mon comité de thèse, pour le temps consacré à mes travaux de recherche et les conseils avisés.

Ces années de thèse furent l'occasion de nombreuses missions à l'étranger. Notamment à Versailles où j'ai eu la chance de faire deux passages à l'IJPB, et à Lyon (UMR5240, laboratoire mixte CNRS/Bayer CropScience) où j'ai effectué plusieurs séjours de longue durée.

Un grand merci à Loïc Rajjou pour m'avoir accueilli au sein de son équipe pour des échanges aussi riches en Sciences qu'en amitié. J'espère que vous viendrez un jour sur notre petit bout de France de l'autre côté du monde. A Erwann pour les conseils sur la 2D, l'aide aux manips et les séances microscopie du samedi et du dimanche, merci. Merci à Gwendal pour ton aide technique, les moments chivas/chocolats, le spotting de gel et toute l'aide aux manips que tu as réalisé quand j'étais à Nouméa. A Béatrice et Alima pour toute l'aide en cytologie dans la bonne humeur. Merci à Imen pour les séances coaching « on va y arriver », les tours de voiture jusqu'au casino de Saint-Cyr, les pauses café et les moments de rigolade, mais aussi les échanges de publis et conseils de dernière minute, même si 22 000 km nous séparent, j'ai l'impression qu'on les a un peu rédigées ensemble nos thèses ! Merci également à Marc Galland, post-doc, ami, mentor de la cytologie, on se retrouve en Australie ?

Merci à tous, les meilleurs moments restent sans doute ceux passés à la Pirogue nous permettant de voguer de la Science à autre.

Merci à Maya Belghazi, Jean-Marc Strub et Alain Van Dorsselaer pour la rapidité et la qualité des résultats de spectrométrie de masse. Merci également à tous les membres de la plateforme PAPPSO de l'INRA du Moulon.

Merci à Jim Leebens-Mack et Claude de Pamphilis pour m'avoir permis de venir dans leurs laboratoires et qui m'ont accueilli à bras ouvert. Cette opportunité m'a permis de voir la recherche à la mode US et de nouer des relations de travail et d'amitié qui je l'espère, dureront. Ainsi, merci à Joshua Der pour l'accueil chaleureux dans sa famille, au laboratoire et pour la formation accélérée en bio-info 1.1. Merci à Loren et Susan pour l'intégration à la vie sociale de State College et les discussions Sciences autour d'un bon hamburger.

Merci à Kingsley Dixon pour m'avoir permis de participer au colloque SERA 2012 à Perth, pour les conseils, les idées foisonnantes et la mise en relation du plus petit poil de mouche à l'écologie à grande échelle.

A tous les membres du LIVE, ce fut une réelle chance pour moi d'arriver dans ce laboratoire aux thèmes de recherche aussi divers que les personnalités s'y trouvant. Merci à Hamid Amir pour la direction du laboratoire et les efforts consentis à la réalisation de ma thèse. Merci à Yvon et Yannick pour les discussions/conseils. Merci à Valérie, Véronique et Stéphane pour les moments partagés autour d'un bon café. Merci à tous les autres pour la bonne humeur que vous participez à mettre, notamment Sarah, Audrey, Claire, et tous les stagiaires.

Une mention particulière à mes collègues thésards Laury Dijoux, Christelle Paillon, Marine Briand, Marine Toussirot, Josina Tiavouane et aux anciens, Adrien Wullf et Charly Zongo. Simon et Laurent, des mots ne suffiront pas, mais un grand merci à tous les deux pour votre soutien dans tous les moments, les rigolades et les mardis des thésards. Bientôt votre tour, gardez le cap, ça en vaut la peine.

Je tiens également à remercier toute l'équipe administrative et la présidence de l'UNC qui régit avec efficacité le fonctionnement de cette université.

Merci à Laurent, Gildas, Isabelle, Nicolas & Nicolas, Marion, Romain, Casi et tous les collègues de l'Institut Agronomique néo-Calédonien pour l'aide aux récoltes, les conseils, l'aide aux manip et les bons moments passés en votre compagnie.

Même si ces résultats n'apparaîtront pas dans ce manuscrit, ces années de thèse m'auront permis de travailler pour la société Koniambo Nickel SAS. J'ai ainsi pu jouir d'un soutien financier et d'une accessibilité à la transition laboratoire-terrain extraordinaire. Merci à Antoine Leveau pour m'avoir laissé intégrer le service REVEG de KNS. Un grand merci à tous les acteurs de la revégétalisation de KNS, Laurent, Patrick, Mado, Milo, Joan, Sika. Une mention toute particulière à Lyz pour tout ce temps passé sur le massif, les discussions dans la voiture et ta passion pour la préservation de notre patrimoine naturel.

Merci à mes supports financiers, la Province Sud de Nouvelle-Calédonie et la société Koniambo Nickel SAS.

Je tiens à remercier Françoise Corbineau et Ian Small qui ont accepté d'être les rapporteurs de mon manuscrit.

Je remercie mes amis. A vous qui m'avez écouté, remotivé, boosté,... . Sophie, Julie, Amandine, Guillaume, Esmé, Josine, Kelly, on va enfin pouvoir reprendre nos soirées pizza/champagne ! La prochaine est pour moi.

Un énorme merci à ma famille. Je ne pourrais pas tous vous citer ici, mais vous m'avez toujours soutenu, laisser libre de mes choix et encourager.

Merci à mes parents, Françoise et Jean-Jacques, pour m'avoir porté à bout de bras de la petite école de Boulari jusqu'à ma chère rue Mouffetard. Merci d'avoir cru en moi, si j'en suis là, c'est grâce à vous, continuons de regarder vers l'avenir avec Amour.

Olivier, Marie-Eve, Julie, Romance et Esmé, votre soutien m'a été précieux, merci. Merci également à mon grand-père et à ma grand-mère.

Je terminerai par remercier la personne avec qui j'ai la chance de partager ma vie, ma douce Gaëlle. Tu as vécu ces années de thèse aussi intensément que moi, tout en me laissant quitter la maison de longs mois. Ton soutien constant, tes encouragements, ta présence chaleureuse et aimante m'auront permis de rester sur le chemin. Merci, et maintenant, au jardin...



# Abréviations, sigles et symboles

A	Ampère
ABA	Abscissic Acid (acide abscissique)
ADN	Acide DésoxyriboNucléique
ANITA	familles du clade des plantes basales ANITA ( <i>Amborella</i> , <i>Nymphaeales</i> et <i>Illiciales</i> , <i>Trimeniaceae-Austrobaileya</i> )
APS	Ammonium PerSulfate
ARN	Acide RiboNucléique
BLAST	Basic Local Aligement Search Tool (outil de recherche par alignement local basique)
CHAPS	3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonate
CNRS	Centre National de la Recherche Scientifique
D	Dormant
Da	Dalton, unité de masse des atomes
DNase	Désoxyribonucléase
DTT	dithiothréitol, une petite molécule redox, également connue sous le nom de réactif de Cleland
EDTA	Ethylène Diamine TétraAcétique
ELISA	Enzyme-Linked Immunosorbent Assay (dosage d'immunoabsorption par enzyme liée)
ER	Endoplasmic Reticulum (réticulum endoplasmique)
FAO	Forme Activée de l'Oxygène
GA	Gibberellic acid (acide gibbérellique)
GPS	Global Positioning System
GSS	Germination <i>Sensu Stricto</i>
HAI	Heure Après Imbibition
HPLC	High Performance Liquid Chromatography (chromatographie liquide à haute performance)
HSP	Heat Shock Protein (protéine de choc thermique)
Hz	hertz, unité dérivée de fréquence du système international équivalente à une oscillation par seconde

IAC	Institut Agronomique néo-Calédonien
IAM	IodoAcétaMide
INRA	Institut National de la Recherche Agronomique
ISEE	Institut de la Statistique et des Etudes Economiques
LEA	Late Embryogenesis Abundant (protéine abondante en fin d'embryogenèse)
LIVE	Laboratoire Insulaire du Vivant et de l'Environnement
MAFFT	Multiple Alignment using Fast Fourier Transform ; a multiple alignment program for amino acid or nucleotide sequences
MS	Mass Spectrometry
Ma	Millions d'années
MD	Morphological Dormancy (dormance morphologique)
MF	Matière Fraîche
MS	Matière Sèche
ND	Non Dormant
Ni	Nickel
PBS	Phosphate Buffured Saline (tampon phosphate salin)
PD	Physiological Dormancy (dormance physiologique)
PDA	Piperazine-di-Acrylamide
PIB	Produit Intérieur Brut
PIMT	Protéine-L-isoaspartate- <i>O</i> -méthyltransférase
PPM	Partie Par Million
PY	PhYsical dormancy (dormance physique)
QSP	Quantité Suffisante Pour
QTL	Quantitative Trait Loci (locus de caractères quantitatifs)
RNase	Ribonucléase
ROS	Reactive Oxygen Species (espèces réactives dérivées de l'oxygène)
S	Svedberg, de symbole S, une unité de mesure du taux de sédimentation d'une macromolécule
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulfate – PolyAcrylamide Gel Electrophoresis (électrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium)
TAG	Triglycérides
TAIR	The Arabidopsis Information Resource

TBS	Tris-Buffered Saline
TCA	Trichloroacetic acid (acide trichloroacétique)
TEMED	Tétraméthyléthylènediamine
TRIS	Trishydroxyméthylaminométhane
UNC	Université de la Nouvelle-Calédonie
V	Volt
W	Watt



# Sommaire

<b>Préface</b> .....	<b>1</b>
<b>Chapitre I. Introduction générale</b> .....	<b>3</b>
<b>1 Introduction</b> .....	<b>5</b>
1.1 Situation générale.....	5
1.2 L'environnement néo-calédonien .....	6
1.2.1 Histoire géologique .....	6
1.2.2 La flore de la Nouvelle-Calédonie .....	7
1.2.2.1 Les formations végétales .....	7
1.2.2.2 Origine de la flore .....	9
1.2.3 Diversité et endémisme.....	10
1.3 La multiplication sexuée des plantes.....	11
1.3.1 La graine, une exceptionnelle innovation .....	11
1.3.2 Les types de graines .....	14
1.3.3 Le développement des graines.....	15
1.3.3.1 Généralités.....	15
1.3.3.2 Les réserves accumulées.....	18
Les protéines .....	18
Les lipides .....	19
L'amidon.....	20
1.3.4 La tolérance à la dessiccation .....	20
1.3.5 La dormance.....	21
1.3.6 La germination .....	26
1.4 La protéomique .....	33
Axes de recherche et objectifs de la thèse.....	37
<b>Chapitre II. Matériels et méthodes</b> .....	<b>39</b>
<b>1. Matériel végétal</b> .....	<b>41</b>
1.1 <i>Amborella trichopoda</i> .....	41
1.1.1 Récolte des semences.....	41
1.1.2 Préparation des semences .....	41
1.1.3 Préparation des embryons.....	42
1.1.4 Tests de germination.....	42
1.1.5 Obtention des différents stades de développement des graines d' <i>Amborella</i> suite à l'imbibition et la germination <i>sensu stricto</i> .....	43
1.2 Genre <i>Psychotria</i> .....	43
1.2.1 Récoltes utilisées pour les analyses protéomiques .....	44
1.2.2 Récoltes utilisées pour l'approche transcriptomique.....	45
<b>2. Flore endophyte de <i>Psychotria gabriellae</i></b> .....	<b>46</b>
2.1 Obtention et mise en cultures des bactéries.....	46
2.2 Caractérisation des bactéries isolées.....	47
<b>3. Composition biochimique des réserves séminales des graines</b> .....	<b>47</b>
3.1 Glucides.....	47
3.2 Protéines.....	47
3.3 Taux d'humidité.....	48
<b>4. Cytologie</b> .....	<b>48</b>
4.1 Inclusion en historésine .....	48
4.2 Inclusion dans l'agarose .....	48
4.3 Coloration au rouge Nil .....	49
4.4 Double coloration au bleu noir naphtol et acide périodique-réactif de Schiff.....	49

<b>5. Etude protéomique .....</b>	<b>50</b>
5.1 Approche fondée sur l'électrophorèse bidimensionnelle des protéines .....	50
5.1.1 Extraction des protéines de graines selon Osborne (1924) .....	50
5.1.2 Extraction des protéines totales de graines d' <i>Amborella</i> .....	51
5.1.3 Dosage des protéines.....	52
5.1.4 Electrophorèse SDS-PAGE.....	52
5.1.4.1 Electrophorèse monodimensionnelle.....	52
5.1.4.2 Electrophorèse bidimensionnelle.....	53
5.1.4.3 Révélation des gels à l'argent.....	55
5.1.4.4 Révélation des gels à l'argent modifié (méthode de Schevchenko).....	56
5.1.4.5 Identification des protéines par spectrométrie de masse .....	57
5.2 Approche de protéomique en vrac (shotgun).....	58
5.3 Construction de la base de données protéomiques des graines d' <i>Amborella</i> .....	58
<b>6. Immunodetection des protéines .....</b>	<b>59</b>
6.1 Obtention des anticorps.....	59
6.1.1 Genre <i>Psychotria</i> .....	59
6.1.2 <i>Amborella trichopoda</i> .....	60
6.2 Test ELISA.....	60
6.3 Western blot.....	61
<b>7. Etude transcriptomique .....</b>	<b>62</b>
7.1 Protocole d'extraction.....	63
7.2 Séquençage.....	63
7.3 Analyse des données .....	64
<b>8 Etude phylogénétique .....</b>	<b>64</b>
<b>Chapitre III. Physiologie de la graine d'<i>Amborella trichopoda</i>.....</b>	<b>65</b>
<b>1. Introduction.....</b>	<b>67</b>
<b>2. Analyse des graines matures sèches d'<i>Amborella</i>.....</b>	<b>72</b>
2.1 Résultats .....	72
2.1.1 Observations cytologiques.....	72
2.1.2 Etude du protéome de la graine mature sèche d' <i>Amborella</i> .....	73
2.1.2.1 Etude des protéines totales de la graine mature sèche par protéomique shotgun .....	74
2.1.2.2 Analyse des fractions albumines et globulines de la graine mature sèche d' <i>Amborella</i> par électrophorèse bidimensionnelle .....	78
2.1.2.2.1 Albumines .....	79
2.1.2.2.2 Globulines.....	80
2.1.2.2.2.1 Identification des gènes codant les globulines 11S d' <i>Amborella</i> .....	85
2.1.2.2.2.2 Classes ontologiques des globulines.....	94
2.2 Discussion .....	95
2.2.1 Caractérisation cellulaire et protéomique de la graine mature sèche d' <i>Amborella</i> .....	96
2.2.1.1 Caractérisation cytologique des graines matures sèches d' <i>Amborella</i> .....	96
2.2.1.2 Analyses par protéomique shotgun des protéines totales de l'embryon et de l'albumen isolés de la graine mature sèche d' <i>Amborella</i> .....	97
2.2.1.3 Caractérisation des fractions albumines et globulines de la graine mature sèche entière d' <i>Amborella</i> .....	112
<b>3. Protéome de la graine d'<i>Amborella</i> au cours de l'imbibition et de la germination .....</b>	<b>117</b>
3.1 Introduction .....	117
3.2 Résultats .....	117
3.2.1 Etude du protéome de la fraction albumines extraite de la graine d' <i>Amborella</i> au cours de l'imbibition et de la germination sensu stricto.....	118
3.2.2 Evolution des globulines de la graine d' <i>Amborella</i> au cours de l'imbibition et de la germination sensu stricto.....	122

3.3	Discussion .....	124
3.3.1	Evolution de la fraction des albumines au cours de la germination .....	124
3.3.1.1	Protéines s'accumulant progressivement tout au long de la germination .....	124
3.3.1.2	Protéines spécifiques de la percée radiculaire.....	127
3.3.1.3	Les protéines spécifiques du stade intermédiaire de la germination.....	128
3.3.1.4	Protéines diminuant progressivement lors de la germination.....	129
3.3.2	Evolution de la fraction des globulines au cours de la germination .....	130
	<b>Conclusion générale sur la caractérisation de la graine mature sèche d'<i>Amborella</i> et lors de sa germination.....</b>	<b>133</b>
	<b>Chapitre IV. Adaptations au stress nickélicifère : <i>Psychotria gabriellae</i>.....</b>	<b>135</b>
1	<b>Introduction.....</b>	<b>137</b>
1.1	Le milieu ultramafique .....	137
1.1.1	Origine et évolution.....	137
1.1.2	Caractéristiques édaphiques.....	140
1.2	Le nickel dans la plante .....	141
1.3	Etat des recherches au L.I.V.E.....	148
1.4	Axes de recherche.....	154
2	<b>Résultats.....</b>	<b>154</b>
2.1	Confirmation de la présence des protéines de type DING dans différents échantillons de graines de <i>P. gabriellae</i> .....	154
2.2	Nouvelles données de séquence .....	158
2.3	Caractérisation de la flore endophyte de <i>P. gabriellae</i> .....	160
2.4	Résultats préliminaires de transcriptomique .....	162
3	<b>Discussion.....</b>	<b>164</b>
3.1	Les protéines de type DING chez <i>P. gabriellae</i> .....	164
3.2	Les données préliminaires de transcriptomique.....	166
4	<b>Conclusion et perspectives.....</b>	<b>167</b>
	<b>Conclusion générale et perspectives .....</b>	<b>171</b>
1.	<i>Amborella trichopoda</i> .....	173
2.	Adaptations au stress nickélicifère .....	176
	<b>Références bibliographiques .....</b>	<b>179</b>
	<b>Annexes.....</b>	<b>207</b>
	<b>Tableaux supplémentaires.....</b>	<b>243</b>





# Liste des figures

Figure 1 : la Nouvelle-Calédonie et son positionnement géographique dans le Pacifique Sud.	5
Figure 2 : carte géologique des roches de la Nouvelle-Calédonie (Grande-Terre).....	7
Figure 3 : principales formations végétales autochtones de la Nouvelle-Calédonie.....	8
Figure 4 : photos (A) d'un plant d' <i>Amborella</i> dans son milieu naturel (Dogny-Sarraméa) et (B) d'un plant de <i>Psychotria gabriellae</i> .....	10
Figure 5 : les types de graines d'angiospermes .....	13
Figure 6 : les différentes étapes du développement des graines .....	17
Figure 7 : les différentes phases de la germination .....	27
Figure 8 : importance des ARNm et des protéines stockés et synthétisés <i>de novo</i> dans le contrôle de la germination des graines .....	30
Figure 9 : photographie des fruits d' <i>Amborella</i> (A). Carte de la répartition des différentes populations d' <i>Amborella</i> sur la Grande-Terre (B) (points rouges).....	41
Figure 10 : plant de <i>P. gabriellae</i> marqué pour le suivi mensuel (A). Plants de <i>P. gabriellae</i> (en bleu) et de <i>P. baillonii</i> (en jaune) en conditions naturelles.....	43
Figure 11 : répartition des population de <i>P. gabriellae</i> étudiées sur la Grande-Terre.....	44
Figure 12 : stades de maturité des fruits de <i>P. gabriellae</i> utilisés lors de l'approche transcriptomique.....	45
Figure 13 : extraction fractionnée des protéines de graines, d'après Osborne (1924).....	50
Figure 14 : photo d'une plaque de test ELISA .....	61
Figure 15 : photographies d'un plant d' <i>Amborella</i> parmi la végétation naturelle du plateau de Dogny (A) ainsi que ses fruits matures (B et C).....	67
Figure 16 : arbre phylogénétique des plantes supérieures, d'après <i>Amborella</i> genome project (2013).....	68
Figure 17 : liste des familles pour lesquelles au moins une espèce possède un embryon rudimentaire ou linéaire .....	70
Figure 18 : reconstruction de la séquence chronologique de la germination de graines d' <i>Amborella</i> .....	71
Figure 19 : taux de germination cumulé en fonction du temps des semences (○) ou des graines (▲) d' <i>Amborella</i> mises à germer en boîtes de Petri à 25°C ± 1°C, lumière continue. Les semences ont été scarifiées (■) ou pas (◆) à l'acide et semées en serre (25 ± 4°C, haute humidité) (Fogliani et al. 2013)..	71
Figure 20 : observations au microscope optique de coupes de graines matures sèches d' <i>Amborella</i> après coloration au bleu noir naphthol combinée à la coloration à l'acide périodique/Schiff.....	73
Figure 21 : représentation de l'importance des classes ontologiques (Bevan et al. 1998) des protéines totales extraites d'embryon (Embryo) et de l'albumen (Endosperm) de graines <i>stricto sensu</i> d' <i>Amborella</i> .....	75
Figure 22 : diagramme de Venn et signatures ontologiques (Bevan et al. 1998) des protéines totales de l'embryon et de l'albumen de graines <i>stricto sensu</i> d' <i>Amborella</i> identifiées par protéomique <i>shotgun</i> .....	78
Figure 23 : gels bidimensionnels colorés à l'argent des fractions albumines (A) et globulines (B) des graines matures sèches d' <i>Amborella</i> .....	79
Figure 24 : représentation des classes ontologiques (Bevan et al. 1998) de la fraction albumines des graines matures sèches d' <i>Amborella</i> .....	80
Figure 25 : les étapes de la biosynthèse des globulines de type 11S dans les cotylédons de soja. ....	83

Figure 26 : utilisation des protéines de réserve de type légumineuses (globulines 11S) .....	84
Figure 27 : recouvrement des séquences nucléotidiques (séquences codantes) et protéiques des trois gènes de globulines 11S d' <i>Amborella</i> .....	85
Figure 28 : recouvrement des séquences nucléotidiques (séquences codantes) et protéiques des trois gènes de globulines 11S identifiés à partir des données peptidiques expérimentales obtenues par protéomique <i>shotgun</i> .....	86
Figure 29 : arbre phylogénétique construit à partir des séquences des globulines 11S de diverses espèces.....	87
Figure 30 : alignement des séquences peptidiques de globulines 11S d' <i>Amborella</i> , <i>Glycine max</i> , <i>Ginkgo biloba</i> et <i>Arabidopsis thaliana</i> .....	90
Figure 31 : représentation schématique de la structure indiquant les exons (traits épais) et les introns (traits fins) des gènes codant les globulines 11S pour <i>Amborella</i> , <i>G. max</i> , <i>G. biloba</i> et <i>A. thaliana</i> .....	92
Figure 32 : prédictions structurales des globulines 11S d' <i>Amborella</i> , <i>G. max</i> , <i>G. biloba</i> et <i>A. thaliana</i> déduites des séquences des gènes codant les globulines 11S chez ces espèces. 93	
Figure 33 : représentation des classes ontologiques (Bevan et al. 1998) de la fraction globulines des graines matures sèches d' <i>Amborella</i> .....	94
Figure 34 : séquences des oléosines d' <i>Amborella</i> identifiées par approche protéomique et extraites du génome.....	100
Figure 35 : métabolisme du soufre dans le contrôle de la germination des graines.....	105
Figure 36 : mécanisme spontané de formation de L-Isoaspartyl et action enzymatique de réparation de la PIMT. ....	106
Figure 37 : arbre phylogénétique obtenu pour le gène unique d' <i>Amborella</i> codant la protéine isoaspartyl méthyltransférase.....	108
Figure 38 : arbre phylogénétique obtenu pour le gène d' <i>Amborella</i> codant la protéine Mother of FT and TFL1 .....	111
Figure 39 : germination <i>sensu stricto</i> des graines matures sèches d' <i>Amborella</i> , dans l'eau, à 25°C et lumière continue.....	117
Figure 40 : coupes longitudinales des stades graine mature (A), ¼ (B) et graine germée (C). L'embryon est matérialisé par le cercle rouge. ....	118
Figure 41 : gels d'électrophorèse bidimensionnelle colorés à l'argent des fractions albumines de la graine mature sèche d' <i>Amborella</i> (A), du stade ¼ (B) et du stade graine germée (GSS) (C).....	119
Figure 42 : classes ontologiques (Bevan et al. 1998) des quatre groupes de protéines de la fraction albumines de la graine d' <i>Amborella</i> dont l'abondance varie pendant l'imbibition et la germination.....	120
Figure 43 : classes ontologiques et fonctions des protéines de la fraction albumines des graines d' <i>Amborella</i> spécifiquement accumulées au moment de la percée radiculaire GSS (II) et du stade ¼ (III) .....	121
Figure 44 : (A) gel d'électrophorèse monodimensionnelle coloré au bleu colloïdal correspondant à la fraction globulines extraite d'une quantité fixe de graines pour les stades graine mature sèche (GMS, MS), ¼ et germination <i>sensu stricto</i> (GSS, GS). (B) Abondances relatives des globulines estimées par test ELISA pour les stades GMS (MS), ¼ et GSS (GS).....	122
Figure 45 : gel d'électrophorèse monodimensionnelle correspondant aux globulines extraites d'albumen isolé de graines matures sèches (t0), ou d'albumen isolé des graines matures sèches puis imbibé 17 jours avec de l'eau (eau) ou une solution de gibbérellines (GA <sub>3</sub> ) .....	123
Figure 46 : photographie d'un sol ultramafique et de sa végétation arbustive en haut du Massif du Koniambo (Koné).....	138

Figure 47 : schéma du profil d'altération des roches péridotidiques de Nouvelle-Calédonie	139
Figure 48 : stratégies de réponse des plantes exposées à une concentration croissante de métal dans le sol.	145
Figure 49 : mécanismes moléculaires proposés pour être impliqués dans l'hyperaccumulation de métaux par les plantes	146
Figure 50 : cartographies brutes aux rayons X de la répartition des éléments nickel (Ni), calcium (Ca) et phosphore (P) au sein de l'endocarpe (à gauche), de l'albumen (à gauche et à droite) et de l'embryon (à droite) de <i>P. gabriellae</i> (d'après Zongo 2010).	149
Figure 51 : cartographie au rayons X de la répartition des éléments calcium (Ca) et phosphore (P) au sein de l'albumen et de l'embryon de <i>P. baillonii</i> (d'après Zongo 2010).	150
Figure 52 : signatures ontologiques (Bevan et al. 1998) des protéomes des graines matures de <i>Psychotria gabriellae</i> , <i>Beta vulgaris</i> et <i>Arabidopsis thaliana</i> (Zongo 2010).	151
Figure 53 : séquence de la protéine <i>Human Phosphate Binding Protein</i> (HPBP) (P85173) (Morales et al. 2006).	152
Figure 54 : gels d'électrophorèse bidimensionnelle des fractions albumines des graines matures sèches de <i>P. gabriellae</i> (A) et <i>P. baillonii</i> (B) colorés à l'argent et après western blot avec l'anticorps anti-DING de tabac (Perera et al. 2008) pour <i>P. gabriellae</i> (C) et <i>P. baillonii</i> (D).	156
Figure 55 : gels d'électrophorèse bidimensionnelle des fractions albumines des graines matures de <i>P. gabriellae</i> de Poya (A), des Monts Koghi (B) et de la forêt Desmazures (C) et western blots correspondants révélés avec l'anticorps anti-DING de tabac (Perera et al. 2008). (D) carte de la répartition des populations sur la Grande-Terre.	157
Figure 56 : gels d'électrophorèse bidimensionnelle de graines de <i>P. gabriellae</i> aux stades de maturités II (A), III (B) et IV (C) colorés à l'argent.	158
Figure 57 : séquence peptidique de la protéine DING humaine (HPBP) (P85173) et alignement de différentes protéines de type DING (Berna et al. 2008).	159
Figure 58 : gel d'électrophorèse bidimensionnelle (A) de l'extrait des protéines totales de <i>Methylobacterium populi</i> et western blot correspondant (B) révélé en utilisant l'anticorps anti-DING de tabac (Perera et al. 2008).	161



# Liste des tableaux

Tableau 1 : positionnement des populations de <i>P. gabriellae</i> étudiées. ....	44
Tableau 2 : volumes des réactifs utilisés pour la réalisation d'un gel monodimensionnel de polyacrylamide. ....	53
Tableau 3 : quantités des réactifs intervenant dans la révélation à l'argent des gels. ....	55
Tableau 4 : quantités des réactifs intervenant dans la révélation à l'argent modifié (Schevchenko). ....	56
Tableau 5 : composition des différentes classes de protéines de la graine mature sèche d' <i>Amborella</i> . Les résultats sont exprimés en pourcentage de la matière fraîche (trois réplicats). ....	72
Tableau 6 : taux de recouvrement des séquences peptidiques prédites des trois gènes codant les globulines 11S d' <i>Amborella</i> par les peptides obtenus expérimentalement à partir de l'approche 2-DE ou celle de protéomique <i>shotgun</i> . ....	86
Tableau 7 : nombre de plantes hyperaccumulant différents métaux (Damien Lee Callahan 2007). ....	141
Tableau 8 : gènes (23) dont les transcrits issus d'un mélange de fruits de <i>P. gabriellae</i> ont été retrouvés plus de dix fois lors du séquençage du transcriptome. ....	163



# Liste des Annexes

<b>Annexe 1 : carte de référence de la fraction albumines de la graine d'<i>Amborella trichopoda</i></b> .....	<b>209</b>
<b>Annexe 2 : carte de référence de la fraction globulines de la graine d'<i>Amborella</i></b> .....	<b>217</b>
<b>Annexe 3 : analyse phylogénétique des globulines 11S d'<i>Amborella</i></b> .....	<b>221</b>
Participation au corps de texte (article principal) : .....	223
Participation en données supplémentaire : .....	224
<b>Annexe 4 : observations cytologiques de la graine mature sèche d'<i>Amborella</i> après coloration au rouge Nil</b> .....	<b>237</b>
<b>Annexe 5 : alignement des peptides expérimentaux pour la PIMT et MOTHER of FT and TFL1</b> .....	<b>239</b>
<b>Annexe 6 : résultats de la caractérisation de la flore endophyte de la graine mature sèche de <i>Psychotria gabriellae</i> réalisée à Angers par Marie-Agnès Jacques (INRA Angers)</b> .....	<b>241</b>
Séquence de l'amplification de l'ADNr 16S : .....	241
Résultat de l'analyse par BLAST : .....	241





## **ATTENTION**

Le contenu de cette thèse fait actuellement l'objet de rédaction de publications, les résultats présentés font l'objet d'une certaine confidentialité.



# Préface

L'étude de la biodiversité mondiale et de ses richesses fait l'objet d'un intérêt croissant dans la communauté scientifique. La Nouvelle-Calédonie est un archipel français du Pacifique Sud. Son histoire géologique est fascinante et a abouti à son immersion il y a environ 37 millions d'années la plaçant dans sa position actuelle et en faisant un des complexes ophiolitiques majeur dans le monde. Ainsi, il y a 37 millions d'années, la Grande-Terre de la Nouvelle-Calédonie représentait un refuge climatique stable propice au développement et à la conservation de la flore des autres terres émergées de la région, et notamment de l'Australie. La Grande-Terre héritât ainsi d'une flore de type gondwanienne qui fut cependant soumise à une forte pression de sélection. En effet, suite à la formation du complexe ophiolitique, un tiers de la Grande-Terre fut recouvert d'un sol ultramafique, pauvre en éléments nutritifs et riche en éléments traces potentiellement phytotoxiques. La Nouvelle-Calédonie contient ainsi, avec Cuba, le plus grand nombre d'espèces végétales hyperaccumulatrices de métaux. Cette combinaison complexe multifactorielle a abouti à la flore actuelle du Territoire, riche, unique et diverse. Le taux d'endémisme de la flore y est de 74,7%, faisant de la Nouvelle-Calédonie un laboratoire d'étude de la biodiversité à ciel ouvert. Les originalités de sa flore y sont nombreuses. Afin de préserver cette richesse, il convient de caractériser les espèces et mettre en évidence des mécanismes adaptatifs encore inconnus.

La domination mondiale des végétaux supérieurs est en partie due à l'extraordinaire innovation que représente la graine. Élément issu de la reproduction sexuée, elle assure la colonisation des milieux et possède de remarquables capacités d'adaptation. Les graines représentent la seule phase mobile dans le cycle de développement d'un végétal.

Ainsi, le Laboratoire Insulaire du Vivant et de l'Environnement (LIVE) de l'Université de Nouvelle-Calédonie développe des programmes de recherche, en collaboration avec ses partenaires scientifiques locaux (Institut Agronomique néo-calédonien (IAC) et Institut de Recherche pour le Développement (IRD)), visant à caractériser la biodiversité du Territoire. Les originalités foisonnent en Nouvelle-Calédonie, permettant ainsi d'aborder un spectre d'études très large allant de l'étude des communautés de microorganismes associées aux végétaux, à celle liée à l'intérêt des substances naturelles ou encore aux études concernant les réseaux trophiques des espèces animales du lagon. De par la place des graines dans le cycle de développement des végétaux, cet organe a une place de choix dans les études. Sur

ces bases, au cours de ce travail de thèse, nous nous sommes intéressés à étudier deux cas particuliers de la flore de la Nouvelle-Calédonie. Il s'agit plus particulièrement de la caractérisation des graines de i) d'une plante à fleur basale, *Amborella trichopoda*, et ii) d'une des plantes accumulant le plus de nickel au monde dans ses feuilles, *Psychotria gabriellae* (anciennement *P. douarrei*).

## **Axes du travail de recherche et principaux attendus de la thèse**

*Amborella trichopoda* est la sœur de toutes les plantes à fleurs. Elle bénéficie ainsi d'un très grand intérêt de la part de la communauté scientifique, comme en témoigne le très récent achèvement du séquençage de son génome (*Amborella* Genome project 2013). Dans ce cadre, le LIVE, l'IAC et la fondation Paul Ars Cuttoli de la Fondation de France (<http://www.fondationdefrance.org>) ont entrepris de développer un programme de recherche visant à caractériser la biologie de la graine de cette espèce. Ce travail de thèse a le double but i) de caractériser, au plan biochimique et structural, les graines matures sèches de cette espèce et ii) d'apporter une vision à l'échelle moléculaire des mécanismes régissant le contrôle de la germination de cette graine en rapport avec la place de cette plante dans la phylogénie.

*Psychotria gabriellae* est quant à elle une des plantes connues contenant naturellement le plus de nickel dans ses feuilles. La nécessité croissante de trouver de nouvelles alternatives aux techniques polluantes d'extraction minière, mais surtout à la décontamination des sols pollués par les activités anthropiques requiert une caractérisation des plantes pouvant naturellement séquestrer les métaux lourds du sol. C'est ainsi que le LIVE a développé un programme de recherche ayant pour but de caractériser les mécanismes adaptatifs permettant l'hyperaccumulation du nickel. Dans ce cadre, des travaux très prometteurs ont d'ores et déjà été obtenus dans le cadre d'une thèse précédente au laboratoire (Zongo 2010). Ce travail de thèse a eu pour but de valider les données précédemment obtenues et de consolider la dynamique de recherche autour de l'étude moléculaire de ces plantes hyperaccumulatrices, notamment en affinant la compréhension des mécanismes à l'échelle des gènes.

L'ensemble de ce travail a été favorisé par la mise en place ou le renforcement de collaborations avec des laboratoires en France Métropolitaine (approches protéomiques) et aux USA (approches génomiques). La mise en place et la continuité de telles collaborations sont un atout majeur pour le rayonnement scientifique de l'Université de Nouvelle-Calédonie et plus généralement de la recherche néo-calédonienne.

# **Chapitre I. Introduction générale**



# 1 Introduction

## 1.1 Situation générale

La Nouvelle-Calédonie est un archipel français d'Océanie, situé dans le Pacifique Sud. Elle est située à 1 500 km à l'Est de l'Australie et à 2 000 km au Nord de la Nouvelle-Zélande (17-23°S 163-167°E) (Figure 1). La plus grande île de l'archipel, la Grande-Terre, est d'allure étirée, longue de 400 km et large de 50 km. Elle est entourée de nombreuses îles, notamment au Nord les îles Belep, au Sud l'île des Pins et le long de la côte Est les îles Loyauté (Ouvéa, Maré et Lifou).



Figure 1 : la Nouvelle-Calédonie et son positionnement géographique dans le Pacifique Sud.

L'ensemble de l'archipel présente un climat tropical, tempéré par l'influence océanique et soumis périodiquement aux phénomènes « El Niño » et « La Niña ». Le caractère très montagneux de la Grande Terre (80%), avec sa chaîne centrale, permet la présence d'une variété de climats propices au développement de différents écosystèmes.

La Nouvelle-Calédonie représente environ 19 100 km<sup>2</sup> de terres émergées et une superficie de 1 740 000 km<sup>2</sup> en incluant les aires maritimes, ce qui en fait la troisième île du Pacifique. En 2009, le nombre d'habitants estimé était de 245 580, soit 13 habitants/km<sup>2</sup>. Le

PIB de 6,437 milliards d'euros (2007) fait de la Nouvelle-Calédonie le pays d'Outre-mer avec le plus haut niveau de vie (ISEE 2013). Ceci est notamment dû à la filière très développée de l'exploitation minière, principalement l'exploitation du nickel (Ni). En effet, la Grande-Terre contiendrait jusqu'à 40% des réserves mondiales de Ni. De fait, la Nouvelle-Calédonie est le cinquième producteur mondial de minerai de Ni et l'industrie minière représente jusqu'à 25% du PIB et 90% des exportations commerciales (L'Huillier et al. 2010).

## **1.2 L'environnement néo-calédonien**

### **1.2.1 Histoire géologique**

La Grande-Terre s'est détachée du Gondwana il y a environ 83 millions d'années (Ma) suite à l'ouverture de la Mer de Tasman. Ce mouvement, lié à la tectonique des plaques et qui a abouti à la position actuelle de l'archipel a été marqué par des événements géologiques prédominants. Au cours de cette dérive, la Grande-Terre a subi une période d'immersion d'environ 20 Ma (65-45 Ma), provoquant l'accumulation de nombreux sédiments (Grandcolas et al. 2008) notamment pélagiques, ce qui témoigne des fortes profondeurs de l'immersion (Cluzel et al. 2012). La subduction de la Grande-Terre sous la plaque Pacifique a ensuite été associée à un événement majeur d'obduction. L'orogénèse a donc commencé par le charriage d'une partie de la lithosphère océanique sur la Grande-Terre. La Nouvelle-Calédonie est ainsi un des complexes ophiolitiques les plus importants au monde (Pelletier 2007). La mise en place de ce complexe est à l'origine des substrats dits ultramafiques. Finalement, l'émersion de la Grande-Terre, la situant dans sa position actuelle, est estimée à 37 Ma (Pelletier 2007). L'érosion post-émersion a permis la réapparition de la multitude des sols de l'île ainsi que son relief. Les roches ultramafiques recouvrent aujourd'hui un tiers de la surface de la Grande-Terre, principalement sous forme de massifs, dont le plus important est le Massif du Grand Sud (Figure 2). Les îles Loyauté sont quant à elles apparues plus tardivement (2 Ma), suite au soulèvement de la ride des Loyauté (Murienne et al. 2005). Elles sont principalement formées de calcaire récifal formé à faible profondeur. Les roches de la Nouvelle-Calédonie sont donc d'origines très diverses (Figure 2), incluant principalement des péridotites, des granites, des basaltes et des gabbros (Cornuet et al. 1987).



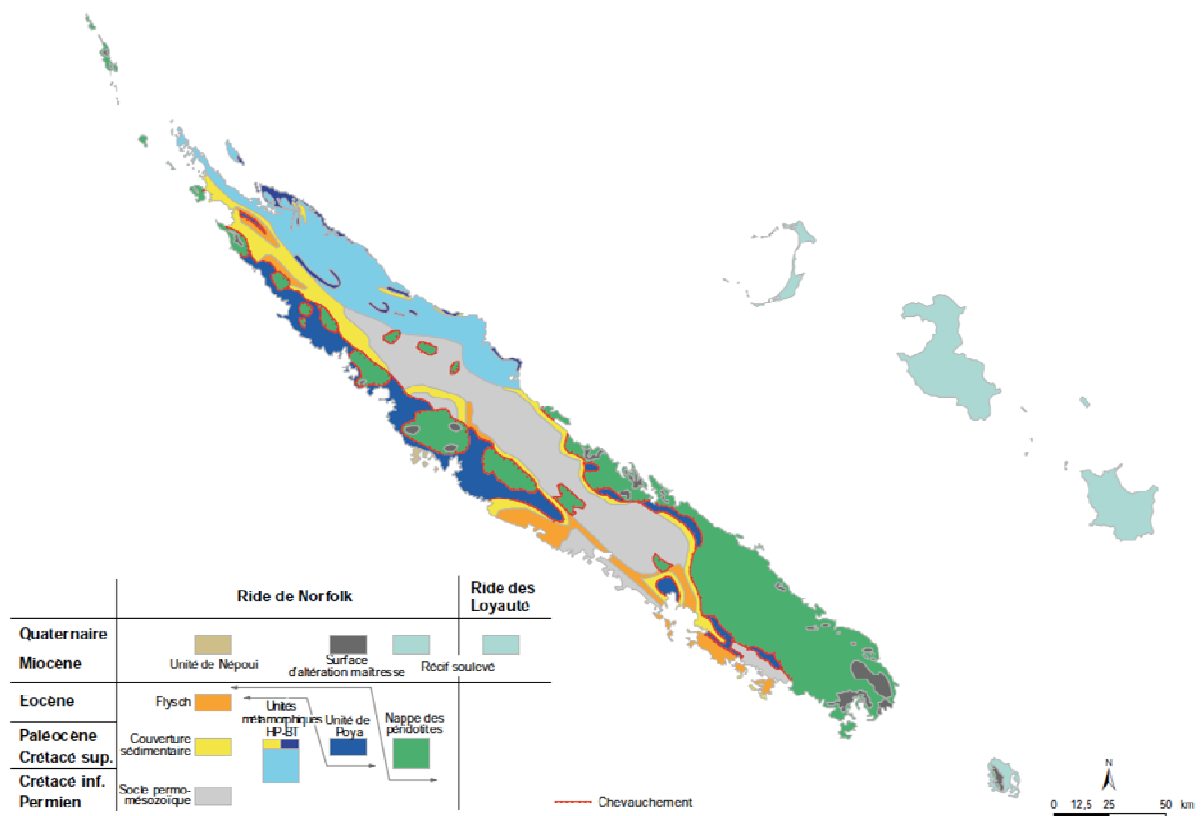


Figure 2 : carte géologique des roches de la Nouvelle-Calédonie (Grande-Terre) ([www.dimenc.gouv.nc](http://www.dimenc.gouv.nc)).

## 1.2.2 La flore de la Nouvelle-Calédonie

### *1.2.2.1 Les formations végétales*

La Grande-Terre présente cinq formations végétales autochtones principales (Figure 3).

Les **forêts humides** qui se retrouvent sur tous types de sols, des basses aux hautes altitudes, présentant ainsi des compositions spécifiques variables. Les arbres peuvent y atteindre 20 mètres de hauteur. Ces forêts couvrent aujourd'hui environ 22% de la surface et se concentrent surtout sur la côte Est et la chaîne centrale.

Les **forêts sclérophylles** (aussi appelée forêts sèches) qui se retrouvent uniquement sur les sols sédimentaires et à une altitude maximale de 300 mètres, principalement sur la côte Ouest. Il s'agit d'une végétation arbustive avec des arbres allant jusqu'à 10-12 mètres de hauteur. Cette formation est la plus menacée de toutes, du fait de nombreux facteurs tels que l'activité de l'Homme pour la création de nouveaux pâturages, les incendies, ou le défrichage par une population envahissante de cerfs. On estime que sa surface actuelle ne représente guère plus que 2,2% de sa surface originelle.

Les **maquis** qui sont constitués par une végétation sclérophylle arbustive. Parmi eux, les maquis miniers majoritaires sont inféodés aux sols ultramafiques des basses aux hautes altitudes. Ils couvrent environ 30% de la surface du territoire et ont un taux d'endémisme pouvant atteindre près de 90%. Les maquis sur roches acides, couvrant une très faible surface, sont quant à eux inféodés aux sols volcano-sédimentaires.

La **mangrove** qui est présente sur près de 200 km<sup>2</sup> de zone côtière composée d'estuaires, de cordons de littoraux ou de fonds de baies.

La **savane** dite à niaouli qui se trouve principalement sur la côte Ouest. Sa surface a augmenté suite à la dégradation des autres milieux par les activités humaines.

Du fait de l'anthropisation du territoire, 50% de la surface actuelle de la Nouvelle-Calédonie est composée de végétation secondaire correspondant notamment à de la savane ou à des zones de végétation primaire très dégradée.

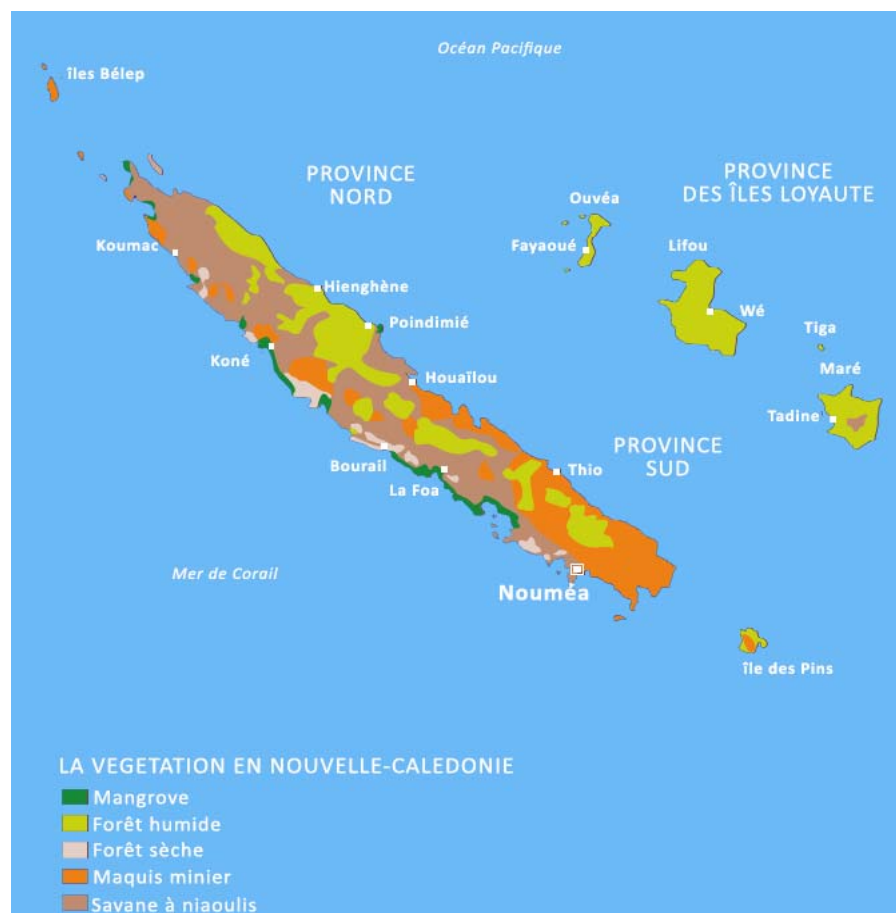


Figure 3 : principales formations végétales autochtones de la Nouvelle-Calédonie (Violette 2009).

### 1.2.2.2 Origine de la flore

La flore néo-calédonienne contient de nombreuses espèces basales. La plus connue d'entre toutes est *Amborella trichopoda* (Figure 4), aujourd'hui reconnue comme la sœur de toutes les plantes à fleurs (Qiu et al. 1999 ; Soltis & Soltis 2004 ; Soltis et al. 2008). Toutefois cette flore contient également d'autres espèces basales de la famille des *Trimeniaceae*. *Amborella trichopoda* étant la seule espèce de son genre, elle sera appelée *Amborella* tout au long de ce manuscrit.

La présence de ces espèces, combinée à la surreprésentation de certaines lignées dites gondwaniennes (*Proteaceae*, Gymnospermes) (Morat et al. 2012) avait conduit à la conclusion que la flore était d'origine gondwanienne (Morat et al. 1986). Cependant, compte tenu de l'histoire géologique de la Grande-Terre, la colonisation ayant donné lieu à la flore actuelle aurait commencé il y a 37 Ma (Grandcolas et al. 2008). Celle-ci serait le résultat de dispersions à plus ou moins longues distances grâce à la présence de ponts insulaires (Cruaud et al. 2012 ; Pillon 2012). La flore résulterait de la dispersion de graines anémochores, zoochores ou possédant des structures (endocarpe, téguments) très épaisses ayant permis un transport par la mer (Howarth et al. 2003). La recolonisation de la Grande-Terre assurée par les terres émergées voisines, combinée à la stabilité du climat de type tropical de la Nouvelle-Calédonie, aurait permis de préserver cette flore « d'origine », alors que celle-ci disparaissait notamment en Australie, du fait de changements climatiques (notamment la sécheresse) (Kershaw 1984).

La mise en place du complexe ophiolitique a engendré la formation de sols ultramafiques a priori peu propices au développement des plantes. En effet, les teneurs en azote, phosphore et potassium y sont basses, alors les teneurs en nickel, chrome, cobalt, manganèse ... y sont élevées et la biodisponibilité de l'eau est faible (L'Huillier et al. 2010).

La présence de ces sols semble cependant avoir joué un rôle primordial dans l'évolution de la flore de la Grande-Terre. D'une part, les espèces australasiennes présentes rapidement après l'émersion ont pu s'adapter à ces sols, leur créant ainsi une niche écologique et les séparant des autres espèces tropicales plus compétitives mais non adaptées. D'autre part, l'induction d'une forte pression de sélection a abouti à la richesse actuelle (Jaffré et al. 1994).

La Nouvelle-Calédonie possède ainsi, avec Cuba, la plus grande concentration au monde d'espèces hyperaccumulatrices de métaux lourds, notamment de nickel (Brooks et al. 1998). Les plantes hyperaccumulatrices sont capables de tolérer des doses intrinsèques en métaux lourds supérieures à 1000 ppm, doses normalement létales pour les espèces non

adaptées. Les deux cas les plus connus sont *Pycnantha acuminata* (*Sapotaceae*) dont le latex de couleur bleu-vert peut contenir jusqu'à 25% de son poids en nickel (Perrier et al. 2004) et *Psychotria gabriellae* (*Rubiaceae*) (Figure 4) dont les feuilles peuvent contenir jusqu'à 4,7 % (rapport à la matière sèche) de nickel (Jaffré & Schmid 1974). Les stratégies adaptatives mises en place par ces végétaux hyperaccumulateurs sont présentées dans le chapitre sur *Psychotria*.



Figure 4 : photos (A) d'un plant d'*Amborella* dans son milieu naturel (Dogny-Sarraméa) et (B) d'un plant de *Psychotria gabriellae* (ex *P. douarrei*) aux Monts Koghi.

### **1.2.3 Diversité et endémisme**

L'isolement géographique, la mosaïque des sols et les conditions climatiques et altitudinales variées ont permis la mise en place de biotopes très divers. Ainsi, la flore vasculaire indigène est représentée par 3 371 espèces, dont 3 099 pour la flore phanérogame. Les taux d'endémisme sont de 74,7% pour la totalité de la flore, et de 77,8% pour la seule flore phanérogamique (Morat et al. 2012). La Nouvelle-Calédonie fait ainsi partie des dix plus importants « hot-spots » mondiaux en terme de biodiversité (Myers et al. 2000 ; Mittermeier et al. 2004) et sa protection et sa restauration (Fogliani et al 2011 ; Villegente, et al 2012 ;

Zongo et al 2012) sont de fait une priorité, d'autant que 5 à 10% des espèces de plantes vasculaires présentes sur le territoire resteraient à décrire (Morat et al. 2012).

La flore des gymnospermes présente une grande originalité et diversité. En effet, sur les 46 espèces présentes, une seule, *Cycas seemanii*, n'est pas endémique. Le taux d'endémisme est ainsi de 97,8%. A elles seules, les gymnospermes néo-calédoniennes représentent 7% des espèces mondiales.

Les monocotylédones sont représentées par 560 espèces regroupées en 30 familles et 199 genres, leur taux d'endémisme étant de 47,1%. Les dicotylédones quant à elles sont représentées par 2 491 espèces regroupées en 126 familles et 490 genres, avec un taux d'endémisme encore plus élevé que chez les monocotylédones, atteignant 84,5%. Les trois plus grandes familles d'angiospermes sont les Myrtacées (257 espèces), les Orchidacées (237 espèces) et les Rubiacées (232 espèces) (Morat et al. 2012).

Ainsi, avec 77,8% d'endémisme pour la flore phanérogamique indigène, la Nouvelle-Calédonie se trouve en troisième position mondiale en terme d'endémicité, après Hawaï (89%) et la Nouvelle-Zélande (81,9%).

Il est donc clair que la flore de Nouvelle-Calédonie est un réservoir principal de biodiversité. Elle constitue un outil extraordinaire mis à la disposition des scientifiques pour mieux appréhender les mécanismes évolutifs et adaptatifs des végétaux. La domination mondiale des plantes phanérogames est notamment attribuée à leur mode de reproduction et de dissémination. Ainsi, la biologie des semences est un domaine de recherche incontournable en raison de l'importance de la graine dans le cycle de développement des plantes et de par le rôle joué par cet organe dans la dissémination des espèces végétales.

## **1.3 La multiplication sexuée des plantes**

### **1.3.1 La graine, une exceptionnelle innovation**

La dissémination des plantes peut être assurée par les deux voies principales que sont la multiplication asexuée tel que marcottage, bouturage, multiplication par bulbes ou stolons, ou la multiplication sexuée, c'est-à-dire, par dissémination de graines.

Au sens botanique du terme, une graine est l'organe issu de la fécondation de l'ovule par le grain de pollen. C'est elle qui assure la reproduction des végétaux supérieurs (les gymnospermes et angiospermes). Elle est une entité nouvelle et unique, résultat du brassage du patrimoine génétique des deux parents. L'incroyable variabilité génétique ainsi créée

permet d'accroître considérablement les capacités d'adaptation des plantes aux conditions variables de leur environnement. Ceci est d'une importance capitale étant donné le caractère sessile des végétaux qui, fixés au sol, ne peuvent fuir les stress environnants. Les plantes à graines ont donc un avantage adaptatif considérable leur permettant d'évoluer dans une large gamme d'écosystèmes. La graine peut ainsi être considérée comme une des adaptations les plus remarquables des végétaux supérieurs, leur permettant notamment de dominer la flore mondiale (Rajjou et al. 2012). Elle correspond à la phase mobile dans le cycle de développement d'une plante.

La double fécondation est caractéristique des angiospermes. Lorsque le pollen (gamète mâle) germe sur le stigmate, il se forme un tube pollinique qui va libérer deux noyaux spermatiques au sein de l'ovule. C'est alors que se produit une double fécondation, l'une donnant l'embryon ( $2n$ ), future plantule, et l'autre, le tissu de réserve, l'albumen, souvent triploïde, suite à la fusion du noyau spermatique avec les deux noyaux polaires issus du gamétophyte femelle. C'est suite à la double fécondation qu'il y a, d'une part, embryogenèse, puis d'autre part, accumulation de réserves lipidiques, glucidiques et protéiques qui sont mobilisées pendant la germination et le développement de la plantule. Puis vient l'acquisition de la tolérance à la dessiccation et la mise en place d'une dormance éventuelle (Bewley & Black 1994).

Les réserves peuvent être stockées dans l'embryon lui-même (au niveau des cotylédons), dans l'albumen (tissu triploïde issu de la double fécondation chez les angiospermes, alors que les gymnospermes contiennent un tissu de réserve haploïde). Dans certains cas, le nucelle maternel peut évoluer, se maintenir, et former le périsperme (cas de la betterave à sucre par exemple). Ainsi, on distingue trois grandes classes de graines en fonction de la nature de leur tissu de réserve (Figure 5), les graines exalbuminées, les graines à albumen et les graines à périsperme.

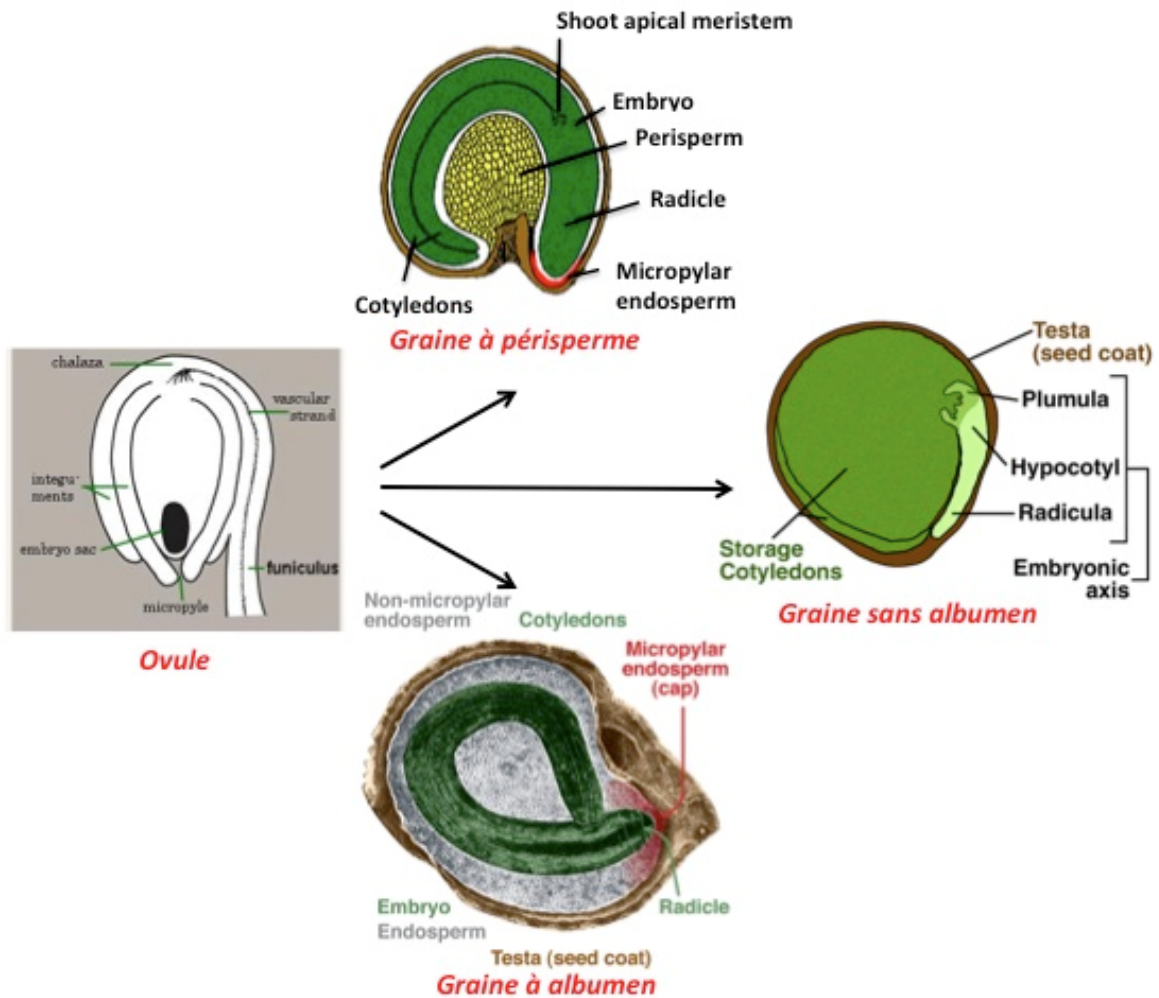


Figure 5 : les types de graines d'angiospermes (d'après The Seed Biology Place ; <http://www.seedbiology.de/index.html>).

L'embryon et le (ou les) tissu(s) de réserve sont entourés par les téguments de la graine. C'est cet ensemble qui constitue la graine *sensu stricto*, bien plus résistant aux stress environnementaux que les propagules végétatives. Cependant, dans certains cas, l'unité de dissémination est une graine entourée de parties du fruit. Dans de tels cas, l'usage du terme de semence est plus approprié. Il est d'une importance capitale de bien différencier l'unité de dissémination (semence, graine) car la présence de certaines structures du fruit lors de la conservation et de la germination peut entraîner des complications (pourrissement, retard à la germination, ...).

La capacité de conservation des graines est un de leurs avantages adaptatifs majeurs (Dickie & Pritchard 2002). Ainsi, des graines de lotus sacré ont été datées de plus de 1 000 ans (Shen-Miller et al. 1995 ; Shen-Miller 2002 ; Shen-Miller et al. 2013). Ceci est notamment possible grâce au fait i) que la plupart des graines tolère une déshydratation

intense pouvant aller jusqu'à 5% de leur teneur en eau, et ii) que cette dessiccation intense est associée à un ralentissement massif du métabolisme cellulaire. Au moment de leur dispersion, ces graines, dites orthodoxes, peuvent ainsi être considérées comme une unité de dissémination en quiescence, prêtes à germer dès que les conditions extérieures (eau, lumière, température, ...) sont favorables. La graine est un arrêt réversible de la croissance au cours du cycle du développement des plantes. Ainsi, pendant sa conservation, sa vigueur et son aptitude à germer ne doivent pas être perdues. Ceci implique la mise en place de mécanismes particuliers d'un point de vue cellulaire et moléculaire qui concourent d'une part, à assurer la tolérance à la dessiccation et gérer tous les stress associés à une perte en eau tout en maintenant la vigueur germinative, et, d'autre part, à favoriser la reprise métabolique nécessaire à l'accomplissement de la germination et l'implantation d'une plantule vigoureuse (Rajjou et al. 2012).

### **1.3.2 Les types de graines**

Les graines peuvent être classées au regard de leurs compositions en réserves séminales, ou selon leur teneur en eau. La composition en réserves séminales des graines est importante d'un point de vue nutritionnel, mais également pour leur conservation. En effet, certains composés comme les acides gras peuvent être très sensibles à la dégradation (oxydation) lors du stockage des graines et entraîner une perte de vigueur, voire la mort des graines par l'induction de chaînes radicalaires générant des espèces toxiques dérivées de l'oxygène.

On distingue trois grands groupes de graines en fonction de la nature de leurs réserves séminales.

- **Les graines à réserves amylacées** telles que le blé, qui contiennent beaucoup d'amidon (75 à 80%), des protéines (15 à 20%) et peu de lipides (1 à 2%).
- Les **graines à réserves lipidiques** qui se distinguent en deux sous-groupes :
  - celles contenant plus de protéines que de lipides, c'est notamment le cas du soja dont les graines renferment 20-30% de lipides, plus de 40% de protéines et très peu d'amidon ;
  - celles contenant plus de lipides que de protéines, c'est le cas du colza dont les graines renferment 40-50% de lipides et environ 20% de protéines.
- Les **graines à réserves protéiques (protéagineuses)** qui contiennent de 40 à 45% d'amidon, mais également des protéines et peu de lipides.



Cependant, d'un point de vue physiologique et en lien avec la conservation des graines (stockage à sec notamment), une autre classification des graines prend en compte leur teneur en eau. Trois grands types des graines sont ainsi définis : les orthodoxes, les intermédiaires et les récalcitrantes (Côme & Corbineau 2000).

- Les **graines orthodoxes** supportent une très forte déshydratation et peuvent présenter, lors de leur dissémination, une teneur en eau de seulement 5% sans perte de viabilité lors de leur conservation à sec. Elles représentent la majorité des graines et sont qualifiées de tolérantes à la dessiccation.
- Les **graines intermédiaires** sont des graines orthodoxes mais dont la vigueur chute drastiquement lorsque leur teneur en eau devient inférieure à 8%.
- Les **graines récalcitrantes** sont riches en eau à maturité et qui contiennent 40 à 50% d'eau. Elles ne tolèrent pas la dessiccation, se conservent mal, et on les retrouve principalement dans les forêts tropicales et subtropicales.

### **1.3.3 Le développement des graines**

#### ***1.3.3.1 Généralités***

De nombreux gènes sont exprimés pendant la formation de la graine (Bewley & Black 1994). Un contrôle très fin de leur expression est donc indispensable, ce, notamment grâce aux signaux hormonaux et aux facteurs de transcription (Wobus & Weber 1999 ; Vicente-Carbajosa & Carbonero 2005).

L'étude des phases du développement d'une graine prend en compte trois paramètres prédominants que sont : la masse de matière fraîche (MF), la masse de matière sèche (MS) et la quantité d'eau présente. A partir de ces paramètres, trois grandes phases se distinguent (Figure 6) :

- L'**embryogenèse**, au cours de laquelle la quantité de matière fraîche augmente de manière conséquente. Lors de cette phase, la teneur en eau initiale de 90 à 95% chute à une valeur de 70 à 75%. Il s'agit d'une période de divisions cellulaires intenses qui aboutit à la formation de l'embryon selon un plan de divisions cellulaires successives bien défini. Chez les dicotylédones, l'embryon possède deux cotylédons, la gemmule, la radicule, alors que chez les monocotylédones l'embryon possède un seul cotylédon, la gemmule et la radicule ;

- **L'accumulation des réserves** qui correspond, au début de la maturation, au remplissage de la graine par des réserves lipidiques, protéiques, glucidiques. La quantité d'eau est stable (plateau hydrique) mais la teneur en eau totale diminue. A la fin de cette phase, elle est d'environ 40 à 50%. Ces processus d'accumulation des réserves sont principalement sous le contrôle de l'acide abscissique (ABA), une phytohormone, mais également de la teneur en eau. Ainsi, de nombreux gènes codant des protéines de réserve sont activés, alors que les gènes anabolisants (eg, impliqués dans la mobilisation de ces réserves) sont inhibés. C'est également le moment de l'expression ABA dépendante de protéines chaperons, telles que les protéines LEA (*Late Embryogenesis Abundant*) ou les protéines de choc thermique de faible masse moléculaire (*small Heat Shock Proteins*, sHSPs) indispensables à la survie de la graine à l'état sec (Wehmeyer & Vierling 2000 ; Hundertmark & Hincha 2008) ;
- La **maturation tardive**, qui concerne les graines orthodoxes. Cette phase est capitale car elle détermine l'aptitude de la graine à la conservation, c'est-à-dire sa longévité (Rajjou & Debeaujon 2008). C'est à ce moment que les ARN messagers (ARNm) survivant à la dessiccation sont synthétisés (Rajjou et al. 2004) pour une utilisation suivant l'imbibition des graines mures sèches. La synthèse des protéines chaperons d'expression liée à la teneur en eau se poursuit. C'est le moment clé dans l'acquisition ou pas d'une dormance et de la tolérance à la dessiccation.

Trois régulateurs principaux de la maturation sont aujourd'hui connus. Il s'agit des protéines codées par les gènes *ABA-INSENSITIVE3 (ABI3)*, *FUSCA3 (FUS3)* et *LEAFY COTYLEDON1 (LEC1)* (Wobus & Weber 1999 ; Vicente-Carbajosa & Carbonero 2005). A la fin de la maturation, l'albumen est au point final de son développement alors que l'embryon est dans un état quiescent et retient sa capacité de régénération jusqu'à la germination. Cet état est possible grâce à l'accumulation de métabolites et de protéines conférant la tolérance à la dessiccation et à la conservation. Les réserves qui sont accumulées serviront principalement au développement de la plantule avant que cette dernière ait acquis ses capacités d'autotrophie.

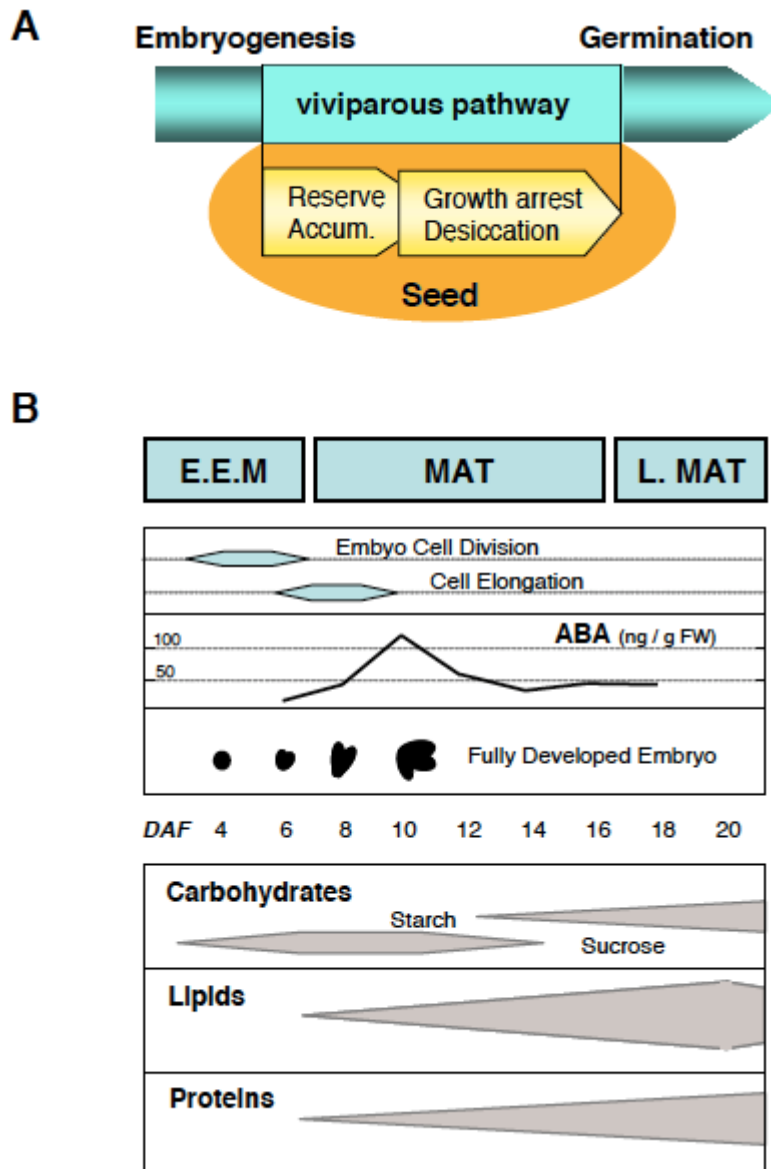


Figure 6 : les différentes étapes du développement des graines (Vicente-Carbajosa & Carbonero 2005). (A) Représentation du développement d'un embryon dans la graine avec la synthèse et l'accumulation des réserves, l'arrêt de la croissance, l'acquisition de la tolérance à la dessiccation et l'entrée en quiescence. (B) Etape du développement d'*Arabidopsis thaliana* indiquant les principales étapes durant l'embryogénèse (E.E.M), la maturation (MAT) et la maturation tardive (L.MAT). Les phases de développement de l'embryon (globulaire, cœur, torpille et mature) sont représentées en parallèle des niveaux d'ABA et de l'abondance des réserves séminales.

### **1.3.3.2 Les réserves accumulées**

#### **Les protéines**

Les principales protéines de la graine peuvent être répertoriées en deux grandes catégories. Premièrement les enzymes, qui sont responsables de la reprise de l'activité métabolique au moment de l'imbibition de la graine, les protéines régulatrices et les protéines structurales. Deuxièmement les protéines de réserve, accumulées lors de la formation de la graine et qui servent notamment de sources de carbone et d'azote lors de la germination et du développement de la plantule (Shewry et al. 1995). En 1924, Osborne a établi une définition fonctionnelle et représentative des protéines présentes chez les graines en les classant selon leur solubilité dans différentes solutions. Cette classification repose sur une extraction successive des protéines à partir d'un échantillon donné de graines. Les quatre catégories de protéines répertoriées sont les suivantes :

- Les **albumines**, principalement des enzymes, solubles dans l'eau à pH neutre ou légèrement acide ;
- Les **globulines**, solubles dans les solutions salines ;
- Les **prolamines**, solubles dans l'éthanol à 90% ;
- Les **glutélines** solubles dans des solutions très acides ou très basiques.

Les globulines, les prolamines et les glutélines sont les protéines de réserve. Leur rôle est d'assurer la fourniture en acides aminés issus de la mobilisation des réserves protéiques (Bewley 1997). Cette source en acides aminés sert à alimenter la synthèse protéique, mais conduit aussi à la production d'énergie et de pouvoir réducteur. Les prolamines sont les protéines de stockage majeures chez les herbacées et les céréales, à l'exception du riz et de l'orge où elles sont minoritaires. Les glutélines, décrites chez le blé, le maïs et d'autres céréales, ont une structure proche de celles des prolamines, mais sont insolubles dans l'éthanol en raison de la présence de ponts disulfures inter-chaînes (Shewry et al. 1995 ; Shewry & Halford 2002). Les globulines sont des protéines présentes en grande abondance chez les graines de la plupart des plantes dicotylédones. Il en est de même chez certaines monocotylédones telles que l'avoine et le riz, même si les globulines de ces deux espèces sont traditionnellement classées comme des glutélines en raison de leur insolubilité dans les sels, contrairement aux globulines canoniques. Les globulines sont classées en deux groupes selon leur coefficient de sédimentation ( $S_{20,w}$ ) : les vicilines 7-8S et les légumineuses 11-12S.

Les vicilines sont synthétisées sous forme de polypeptides de 45-50 kDa codés par des gènes différents. Les globulines 11-12S sont quant à elles synthétisées sous la forme de

précurseurs de 60-65 kDa. Après élimination d'un peptide signal, ces précurseurs forment des trimères dans le lumen du réticulum endoplasmique. Ils y subissent des modifications par glycosylation de certains résidus d'acides aminés dans l'appareil de Golgi et sont ensuite exportés dans des vacuoles de stockage (Jolliffe et al. 2005), grâce à deux courtes séquences d'adressage, l'une en C-terminal (KISIA) et l'autre interne (GNIFRGF) (Petrucci et al. 2007), où ils s'assemblent en hexamères. Ces vacuoles de stockage, également nommées corps protéiques, se forment par fragmentation de la grande vacuole dans le cas des globulines ou par réticulation localisée du réticulum endoplasmique granulaire dans le cas des prolamines (Buchanan & Jones 2000). Les corps protéiques sont localisés dans tout l'embryon, l'albumen et également au sein la couche à aleurone, couche spécialisée entourant l'albumen chez les céréales et l'embryon zygotique chez *Arabidopsis thaliana* (Bethke et al. 2007).

Dans cette thèse, une attention toute particulière est portée aux globulines et aux albumines et les données bibliographiques sont développées plus avant dans le chapitre sur les graines d'*Amborella*.

### ***Les lipides***

La grande majorité des lipides est stockée sous forme de triglycérides. Ils sont stockés dans des corps lipidiques de 0,5 à 2 µm de diamètre. Les corps lipidiques sont formés durant la maturation par bourgeonnement du réticulum endoplasmique. Ils sont constitués d'une matrice de triglycérides entourée par une monocouche de phospholipides contenant différentes protéines structurantes (Tzen et al. 1993). Parmi les protéines enchâssées dans la membrane monocouche, on peut citer :

- Les **oléosines** qui sont essentielles pour éviter la coalescence des corps lipidiques pendant la dessiccation et l'imbibition (Leprince et al. 1998) ;
- Les **caléosines** qui sont synthétisées et accumulées du milieu jusqu'à la fin de la phase de maturation. Elles servent principalement au recrutement des corps lipidiques lors de la germination en les associant aux peroxysomes et en facilitant l'interaction avec les lipases responsables de la mobilisation des triglycérides (Naested et al. 2000 ; Poxleitner et al. 2006) ;
- Le **stéroléosines** qui s'accumulent pendant les phases tardives de la maturation. Le rôle de ces protéines est encore mal connu, mais elles seraient impliquées dans la transmission du signal pour la croissance notamment via la

voie des brassinostéroïdes, et pourraient être impliquées dans la levée de dormance et l'utilisation des réserves lipidiques pendant la germination (Li et al. 2007 ; Baud et al. 2009).

Une fois mures, les corps lipidiques se retrouvent dans le cytosol des cellules.

### ***L'amidon***

L'amidon est la forme la plus répandue de stockage des glucides chez les graines, mais on peut aussi retrouver de l'hémicellulose ou du raffinose (Bewley & Black 1994). L'amidon peut être stocké sous forme d'amylose ou d'amylopectine dans des amyloplastes de taille variant de 2 à 100 µm de diamètre. A maturité, les membranes des amyloplastes se désagrègent exposant les granules directement dans le cytoplasme qui pourront ainsi être dégradés lors de la germination.

### **1.3.4 La tolérance à la dessiccation**

Comme nous l'avons vu précédemment, deux types de graines se distinguent d'un point de vue de leur teneur en eau lors de leur dispersion. Les graines récalcitrantes ont une teneur en eau forte (autour de 50%) à maturité et se conservent très difficilement. En revanche, les graines orthodoxes tolèrent la dessiccation et se conservent très bien de nombreuses années, voire des siècles. De fait, de nombreuses espèces cultivées sont orthodoxes, bien que certaines soient récalcitrantes, eg. celles d'avocatier (*Persea americana*) ou de caféier (*Coffea L.*).

Si, de prime abord, le fait de pouvoir conserver une graine peut paraître une évidence, il s'agit du reflet d'une extraordinaire organisation structurale et moléculaire.

La diminution de la teneur en eau dans la graine orthodoxe induit des modifications physiques. Il s'agit notamment de la minimisation de la vacuole, la protection de l'intégrité nucléaire et une réorganisation du cytosquelette. S'ajoute également une dé-différentiation des mitochondries associée à une « extinction » du métabolisme cellulaire. Toutes ces modifications sont associées à un haut niveau de stress oxydatif qui est géré par l'activation d'une multitude d'enzymes de détoxification ainsi que via la synthèse de molécules piège des espèces activées de l'oxygène (ROS) (Bailly 2004). Il a été montré que l'intensité du métabolisme était corrélée à la tolérance à la dessiccation chez des graines de pois et de concombre (Leprince et al. 2000). La limitation du métabolisme pourrait contribuer à la réduction de la production des ROS. L'activité métabolique étant faible, la transition vers un métabolisme fermentaire toxique serait ainsi évitée. Le cytoplasme des cellules

embryonnaires devient vitreux au moment de la dessiccation. Cet état vitreux est nécessaire à la survie à l'état sec mais ne constitue pas l'évènement déclencheur de la tolérance à la dessiccation (Buitink & Leprince 2008).

La dégradation des protéines due à la perte en eau est prévenue grâce à l'action de protéines chaperons. Ainsi, les protéines LEA jouent un rôle de protection cellulaire et pourraient également agir comme des « réservoirs » d'eau grâce à leur propriété d'hydrophilicité (Tunnacliffe & Wise 2007 ; Berjak & Pammenter 2008). Ces protéines LEA ont été appelées ainsi car elles furent tout d'abord identifiées comme étant abondantes au stade tardif de la maturation, au moment où la teneur en eau diminue et la tolérance à la dessiccation est acquise. Elles se retrouvent en fait dans d'autres organes végétaux et sont groupées en cinq classes selon leurs homologues en acides aminés (Wise 2003).

Les petites protéines de choc thermique (sHSPs) jouent également un rôle essentiel dans la tolérance à la dessiccation des graines orthodoxes. Ces petites protéines (12-42 kDa) constituent des oligomères qui forment un complexe autour de leurs substrats qui sont des protéines ou des complexes protéiques à protéger. Les substrats des sHSPs sont ainsi maintenus dans un état stable au cours de la déshydratation des graines et peuvent être rendus fonctionnels très rapidement grâce à l'action d'autres protéines de choc thermique (HSPs) (Friedrich et al. 2004). De nombreuses sHSPs ont été décrites dans les graines où leur action est indispensable (Lee et al. 1997 ; Wehmeyer & Vierling 2000 ; Jaya et al. 2009 ; Al-Wahaibi 2011).

L'ABA est un sesquiterpénoïde composé de 15 carbones ( $C_{15}H_{20}O_4$ ) résultant du clivage des caroténoïdes et qui contrôle le stockage et l'accumulation des réserves de la graine ainsi que la tolérance à la dessiccation des graines orthodoxes (Nambara et al. 2010). Cette phytohormone induit notamment l'expression des protéines LEA (Tunnacliffe & Wise 2007). De plus, au cours de la maturation des graines, l'ABA exerce un effet inhibiteur sur les mécanismes de déclenchement de la germination précoce, un phénomène délétère pouvant se manifester selon les conditions environnementales lors du développement des graines sur la plante mère (on parle de germination sur pied), permettant ainsi la poursuite du processus de maturation.

### 1.3.5 La dormance

Le passage de la fin de la maturation à l'état de vie ralentie de la graine puis à la germination est régi par de nombreux mécanismes moléculaires et biochimiques complexes encore mal compris (Nonogaki et al. 2010). De plus, dans certains cas, alors que toutes les conditions extérieures semblent favorables à la germination, les graines mises à imbiber refusent de germer. Cet empêchement de la germination est appelé dormance. Une graine viable est qualifiée de dormante lorsqu'elle ne germe pas dans des conditions environnementales a priori favorables. On parle de dormance lorsque le temps de latence avant toute germination visible est supérieur à 30 jours et/ou si la germination est étalée dans le temps jusqu'à plus de deux mois pour un même lot (Baskin & Baskin 2001).

Lors de son développement sur le pied mère, une graine orthodoxe acquiert la capacité à germer et peut en parallèle développer un blocage transitoire de la germination (Finch-Savage & Leubner-Metzger 2006 ; Hilhorst et al. 2010). Ce retard à la germination provient donc de la semence elle-même, ce qui différencie la dormance de la quiescence (germination dès que les conditions extérieures sont optimales). Si le phénomène de dormance peut être gênant en agriculture (hétérogénéité des levées, repousses indésirables), il est d'une importance capitale en écologie. En effet, la dormance est un mécanisme extraordinaire développé par les graines pour retenir le déclenchement de la germination jusqu'à ce que les conditions extérieures soient perçues comme optimales pour la germination et l'installation des plantules. Les chances de survie de l'espèce sont ainsi considérablement augmentées. Par ailleurs, les différents degrés de profondeur de dormance au sein d'un même lot procurent un moyen d'étaler la germination au cours du temps (Baskin & Baskin 2001), et donc une meilleure adaptation aux variations fluctuantes du climat (Donohue et al. 2010).

Dans le but de caractériser les gènes impliqués dans la mise en place et la régulation de la dormance, une approche sur différents écotypes d'*Arabidopsis thaliana* a permis d'isoler des QTL (Quantitative Trait Loci ; locus de caractères quantitatifs) impliqués dans le retard à la germination et appelés *Delay Of Germination (DOG)* (Bentsink & Koornneef 2008). L'un des gènes sous-tendant les QTL *DOG* a été identifié et dénommé *DOG1* (Bentsink et al. 2006). Ce gène appartient à une famille de gènes de cinq membres de fonction moléculaire encore inconnue chez *A. thaliana* et son expression est spécifique de la graine (Bentsink et al. 2006). *DOG1* est exprimé pendant le développement de la graine, et ses transcrits perdurent jusqu'au moment de l'imbibition (Bentsink & Koornneef 2008). Son expression chez *A. thaliana* (écotype Cvi) serait dépendante de la voie de signalisation liée à l'ABA et aux sucres (Teng et al. 2008). Si le mode d'action de ce gène est encore mal compris, il est établi que son niveau



d'expression module la dormance de la graine (Nguyen et al. 2012). L'activité du promoteur se situe au niveau des tissus vasculaires de l'embryon et la protéine codée par ce gène semble localisée dans le noyau. L'accumulation de la protéine DOG1 semble directement liée à l'importance de la dormance pour des graines fraîchement récoltées. En revanche, au cours du stockage à sec (*after ripening*), cette dernière semble subir des modifications, probablement oxydatives, diminuant ainsi la quantité de protéine DOG1 active pendant l'imbibition de la graine. DOG1 agirait comme un inhibiteur de la germination au moment de l'imbibition probablement au niveau de la traduction, tout en agissant également pendant la maturation au cours de laquelle DOG1 influence le niveau d'ABA (Nakabayashi et al. 2012). Par ailleurs, DOG1 semble être un facteur pléiotrope qui est capable d'agir non seulement sur la germination mais aussi sur la floraison. Ainsi, un même gène agissant sur le contrôle de la germination s'avère également capable d'influencer tout le cycle de vie de la plante en contrôlant le temps nécessaire à l'induction de la floraison (Chiang et al. 2013). Dans ce cadre, il est intéressant de noter qu'un des gènes majeurs de la régulation de la floraison, *Flowering Locus C*, est également capable d'influencer la germination. Si l'expression du gène codant ce régulateur est élevée au moment de la maturation des graines, le profil d'expression hormonal sera modifié au moment de la germination en faveur d'un ratio ABA/GA faible (Chiang et al. 2009). Ce mécanisme semble également dépendre d'autres gènes de la régulation de la floraison tels que *Flowering Locus T (FT)* et *Suppressor Of Overexpression of Constans1 (SOC1)* (Chiang et al. 2009).

La période de dormance est nécessaire au déroulement de processus biochimiques encore peu connus, au sein de la graine, et se manifestant avant le début de la germination. La dormance est sous le contrôle de nombreux facteurs génétiques, aussi bien dans sa mise en place (Koornneef et al. 2002) que dans sa levée (Hilhorst et al. 2010). L'hormone principale responsable de sa mise en place et de son maintien est l'ABA, qui agit notamment comme un suppresseur des activateurs de la germination (Nambara et al. 2010). La dormance mise en place sur le pied mère pendant la formation de la graine est qualifiée de dormance primaire. Elle est levée naturellement au cours du stockage à sec de la graine (état de quiescence). Si les conditions extérieures ne sont pas adéquates, la graine peut alors induire une dormance secondaire.

On distingue deux grands types de dormances primaires :

- **Endogènes**, ou embryonnaires, qui sont liées à la dormance de l'embryon lui-même ;
- **Exogènes**, engendrées par les structures entourant l'embryon telles que les téguments, l'albumen, ou même l'endocarpe (dans le cas des semences).

Les connaissances sur la dormance embryonnaire sont peu nombreuses, probablement en raison du très grand nombre d'études réalisées sur le modèle *A. thaliana* qui manifeste des dormances tégumentaires mais pas de dormance embryonnaire (Debeaujon et al. 2000). Toutefois les travaux de Françoise Corbineau et de son équipe à l'Université Pierre et Marie Curie ont fourni nombre d'informations importantes sur la dormance embryonnaire chez le tournesol (*Helianthus annuus* L.), en relation notamment avec les formes actives de l'oxygène (FAO) et l'éthylène (Corbineau et al. 1990 ; Oracz et al. 2007 ; Bailly et al. 2008 ; Oracz et al. 2009).

La compréhension des mécanismes physiologiques permettant la levée de la dormance fait l'objet de nombreuses études. De manière assez surprenante, plusieurs études témoignent d'une activité d'expression des gènes et des protéines chez les graines dormantes durant l'*after ripening*, c'est-à-dire le stockage des graines matures au sec (Hilhorst et al. 2010). Cette propriété d'expression de gènes à l'état sec serait expliquée par l'existence de poches d'eau séquestrées au sein des tissus de graines, pourtant considérées comme sèches (Leubner-Metzger 2005). Les gènes alors exprimés semblent principalement liés à la réponse au stress et au maintien de la dormance dû à l'action de l'ABA. Ce qui fait sens puisque le but de la dormance est, en partie, de permettre la conservation de graines viables.

Au cours de cette conservation, les protéines de la graine peuvent également subir des modifications post-traductionnelles impliquées dans la levée de la dormance. Ceci est notamment lié à l'action oxydante des FAO générées dans la graine et dont les protéines sont une cible principale. Ainsi, la levée de la dormance du tournesol a été en partie attribuée à la carbonylation de protéines spécifiques de l'embryon (Oracz et al. 2007). Ces travaux mettent en lumière de nouveaux mécanismes de levée de dormance, faisant appel à des modifications post-traductionnelles affectant spécifiquement certaines protéines de la graine (Arc et al. 2011). Ceci est notamment possible grâce à une homéostasie très fine des signaux et des dégradations provoqués par les formes activées de l'oxygène (Bailly 2004 ; Ogé et al. 2008).

Les mécanismes de contrôle de la dormance des graines d'*A. thaliana* ont été caractérisés par approche protéomique en utilisant l'accession dormante Cvi originaire des îles du Cap-Vert (Chibani et al. 2006). Des études comparatives réalisées avec des graines

dormantes fraîchement récoltées (D) et dont la dormance a été levée par stockage à sec (*after ripening*) (ND) ont révélé une accumulation différentielle spécifique de 32 protéines. Les données obtenues montrent que les protéines associées aux fonctions métaboliques potentiellement impliquées dans la germination peuvent s'accumuler pendant la levée de dormance par stockage des graines à l'état sec, indiquant une action très tardive des GAs lors de la germination *sensu stricto*, coïncidant avec l'émergence racinaire (Chibani et al. 2006). L'application exogène d'ABA aux graines ND entrave fortement leur germination, ce qui, physiologiquement, paraît mimer le comportement des graines dormantes imbibées. Pourtant, contrairement à cette attente, les données protéomiques démontrent que les mécanismes de blocage de la germination des graines non dormantes par l'application d'ABA sont différents de ceux empêchant la germination des graines dormantes imbibées en absence de la phytohormone (Chibani et al. 2006).

En plus de ces travaux montrant l'importance des protéines et leurs modifications dans le contrôle de la dormance, il convient de noter une publication récente illustrant que, chez le tournesol, la réduction de la dormance au cours de l'*after ripening* est associée à l'oxydation spécifique d'ARNm particuliers stockés dans la graine mature sèche et que cette oxydation est empêchée lorsque les graines sont maintenues en dormance. Cette étude a permis d'identifier 24 ARNm stockés qui se retrouvent très fortement oxydés au cours de l'*after ripening*. Ils correspondent principalement à des gènes impliqués dans la réponse au stress et dans la signalisation cellulaire. Les auteurs ont proposé que l'oxydation ciblée de ces ARNm puisse constituer un nouveau mécanisme engageant la signalisation cellulaire vers la germination dans les premières étapes de l'imbibition (Bazin et al. 2011 ; El-Maarouf-Bouteau et al. 2013).

De nombreux auteurs se sont aujourd'hui accordés pour distinguer cinq classes de dormance (Baskin & Baskin 2007) :

- **Physiologique** (PD). Cette dormance caractérise une inhibition physiologique de l'embryon empêchant l'émergence de la racine. Il s'agit de la forme de dormance la plus commune. Elle comprend trois niveaux : profond, intermédiaire et non profond. Ce type de dormance est très souvent associé à un ratio ABA/GA<sub>3</sub> élevé qu'il conviendra d'inverser par différents traitements pour permettre sa levée (Finch-Savage & Leubner-Metzger 2006) ;
- **Morphologique** (MD). Cette dormance est caractérisée par la présence d'un embryon immature. Ce dernier devra ainsi se développer dans la graine avant

que cette dernière puisse germer. Ce type particulier de dormance est considéré comme le plus primitif (Forbis et al. 2002). Il se retrouve dans la plupart des familles du clade des plantes basales ANITA (*Amborella*, *Nymphaeales* and *Illiciales*, *Trimeniaceae-Austrobaileya*) à l'exception de l'ordre des *Nymphaeales*, mais également chez certaines espèces « modernes » de monocotylédones et eudicotylédones (*Aracaceae* ou *Rubiaceae*) (Baskin & Baskin 2001) ;

- **Morpho-physiologique** (MPD). Certaines espèces sont soumises à la combinaison d'une dormance morphologique et physiologique. La levée de cette dormance nécessite souvent un traitement par stratification (froid ou chaud), avant ou après la maturation de l'embryon (Baskin et al. 2005) ;
- **Physique** (PY). Cette forme de dormance est induite par la présence d'un tégument qui imperméabilise la graine vis-à-vis du monde extérieur. L'embryon ne peut ni être hydraté ni avoir les ressources en oxygène nécessaires à la germination. En conditions naturelles, cette dormance est levée par l'action de microorganismes, par l'alternance de périodes humides et sèches, ou par l'ingestion par un animal. Le but est de fragiliser le tégument ou les structures du fruit responsable de l'imperméabilisation ;
- **Physique et physiologique** (PY + PD). Dans ce cas, le tégument de la graine (ou le péricarpe) est imperméable et l'embryon est aussi physiologiquement dormant.

Cette classification des dormances, principalement écologique, est le sujet d'un vif débat, notamment si l'on considère que, dans certains cas, une graine dormante peut s'avérer physiologiquement très active, notamment dans le cas des dormances morphologiques. En outre, il convient de préciser que, la plupart du temps, cette classification prend en compte l'unité de dispersion de la plante, qui peut être soit la graine, soit la semence (graine + structure du fruit). Le type de dormance peut ainsi varier selon les structures considérées.

### **1.3.6 La germination**

La germination est le processus physiologique et développemental qui permet de passer du stade de la graine quiescente à la plantule (Rajjou et al. 2012). La graine est une unité qui dispose, au moment de la séparation du pied mère, de toute la machinerie nécessaire à la germination (Rajjou et al. 2004). Elle est classiquement décrite en trois phases correspondant

aux phases de prise en eau de la graine suite à son imbibition (Bewley & Black 1994 ; Fait et al. 2006 ; Nonogaki et al. 2010). Ces trois phases sont les suivantes (Figure 7).

1. L'**imbibition** est la phase pendant laquelle la graine absorbe rapidement une grande quantité d'eau aboutissant à la réhydratation de tous les tissus. Les activités métaboliques peuvent ainsi reprendre et la respiration est très importante.
2. La **germination *sensu stricto*** qui est caractérisée par une stabilisation de la prise en eau par la graine. Au cours de cette phase, la graine peut être réversiblement déshydratée sans dommage pour sa viabilité. Cette phase se termine par l'émergence de la radicule (ou d'une partie de l'embryon) au travers des téguments.
3. La **croissance**, manifestée par une reprise de l'absorption d'eau. Cette phase correspond à l'installation et au développement de la plantule.

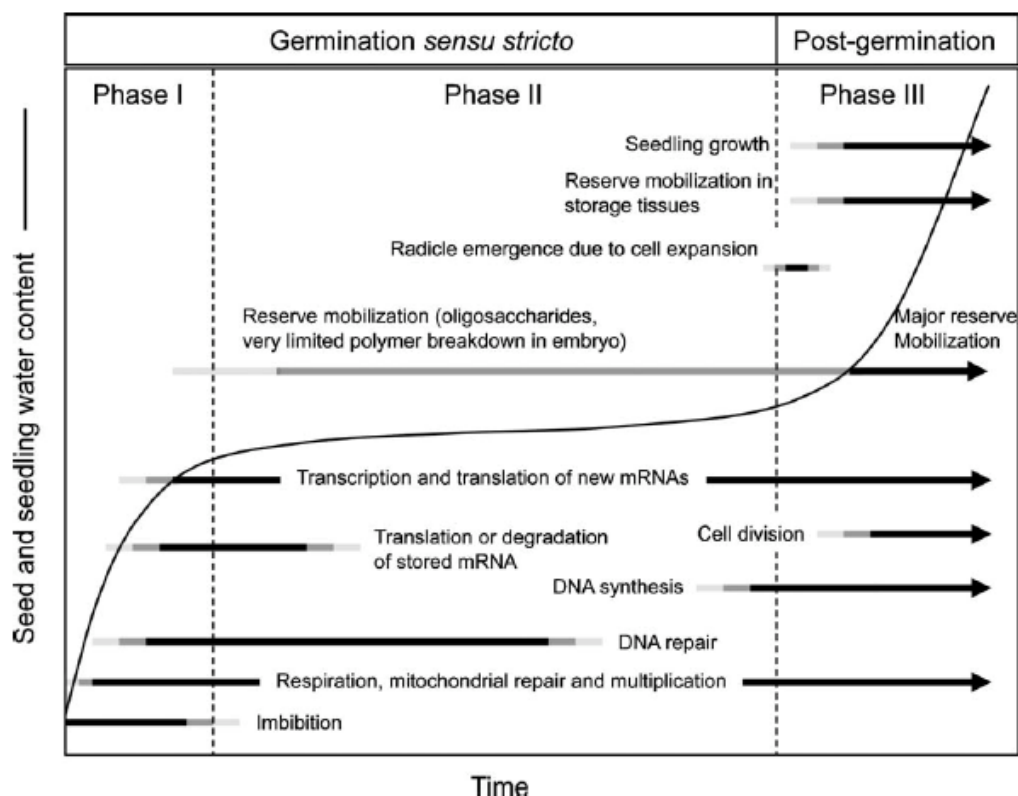


Figure 7 : les différentes phases de la germination (Nonogaki et al. 2010). Modifications physiques et métaboliques se manifestant lors de la germination (phase I et II) et du développement précoce de la plantule (phase III). Le temps nécessaire pour l'accomplissement de ces étapes est spécifique à chaque espèce.

Des travaux récents illustrent que la phase 1, classiquement considérée comme une imbibition physique des graines matures sèches, est associée avec des modifications dans l'expression de nombreux gènes, pouvant être détectées chez les graines d'*A. thaliana* dès 15 min après l'imbibition (Preston et al. 2009). De même, des changements très rapides de l'accumulation de certains métabolites sont décelables 1 h après imbibition (HAI) des graines de riz (Howell et al. 2009).

Plusieurs études ont établi que le développement d'un embryon en germination peut, avant émergence racinaire, être stoppé dans des conditions de déficit hydrique. Cet arrêt de la croissance correspond à un point de contrôle ABA-dépendant lors de la germination impliquant le facteur de transcription ABI5 (Lopez-Molina et al. 2001 ; Lopez-Molina et al. 2002). Des analyses génétiques ont révélé que le facteur embryonnaire ABI3, qui agit en amont de ABI5 et dont l'action est essentielle pour l'expression de ABI5, est également nécessaire pour promouvoir cet arrêt de la croissance. Puisque l'activité d'ABI3 est essentielle à la fin de la phase de maturation des graines, son accumulation durant les premiers stades de la germination en présence de l'ABA ou lors de l'imposition d'un stress osmotique reflète une réexpression du programme de développement se déroulant normalement à la fin de maturation des graines. En d'autres termes, cette première phase de la germination est un processus réversible dans lequel le programme de maturation peut être récapitulé (Figure 8). Considérant la fonction connue des gènes contrôlés par ABI3 durant la maturation des graines, notamment celle des protéines LEA, Lopez-Molina et al. (2001) ont proposé que la formation d'embryons arrêtés dans leur progression vers la germination corresponde à un mécanisme d'adaptation permettant d'augmenter la survie des graines dans des conditions de stress hydrique pouvant être naturellement rencontrées dans le sol. Des études protéomiques ont confirmé que les graines d'*A. thaliana* imbibées pouvaient réinitialiser le programme de maturation (Rajjou et al. 2004 ; Rajjou et al. 2006). La possibilité de récapituler le programme de maturation tardive à deux niveaux : la transcription (Lopez-Molina et al. 2002) et la traduction (Rajjou et al. 2004 ; Rajjou et al. 2006) pourrait permettre un ajustement rapide de la réponse des graines imbibées confrontés à des fluctuations rapides des conditions environnementales, par exemple, la disponibilité de l'eau dans le sol.

Des travaux récents se sont intéressés au rôle des différents compartiments de la graine (enveloppe, albumen l'embryon) au cours de la germination (Dekkers et al. 2013). Cette analyse de l'ensemble des données relatives à l'expression temporelle et spatiale du génome d'*A. thaliana* lors de la germination (vseed.nottingham.ac.uk) illustre le potentiel de ces

données à haute résolution pour la construction de réseaux de co-expression de gènes qui permettent, en particulier, de mieux appréhender les mécanismes du contrôle génétique de la germination. L'ensemble des données révèle deux phases de la transcription au cours de la germination qui sont séparées par la rupture de l'enveloppe de la graine lors de l'émergence radriculaire. La première phase est marquée par d'importants changements du transcriptome lorsque la graine passe d'un état quiescent (graine mature sèche) à un état hydraté et actif. A l'issue de cette première phase, le nombre de gènes différentiellement exprimés chute. Ce nombre augmente à nouveau lors de la rupture de l'enveloppe, marquant le début d'une deuxième phase d'activité transcriptionnelle. Ces données documentent avec précision les destins différents de l'albumen (sénescence) et de la radicule (croissance) lors de la germination (Dekkers et al. 2013).

Pour germer avec succès suite à l'imbibition et afin de contrer l'effet inhibiteur de l'ABA sur la germination, la graine dormante doit (a) établir un catabolisme spécifique, diminuer sa sensibilité, et inhiber la biosynthèse de cette molécule afin d'en réduire le niveau d'accumulation (Ali-Rachedi et al. 2004), et (b) synthétiser une autre classe d'hormones représentées par une grande famille de diterpènes tétracycliques, les gibbérellines (GAs), qui sont des activateurs essentiels de la germination (Koornneef & van der Veen 1980 ; Groot & Karssen 1987 ; Sun & Gubler 2004) (Figure 8). Les GAs régulent négativement l'accumulation de protéines se comportant comme des répresseurs de la germination, tels que les régulateurs de transcription GRAS impliqués dans la différenciation cellulaire (Zhang et al. 2011). En accord avec l'idée que le rapport ABA / GA régule la transition métabolique nécessaire à la germination (Yamaguchi 2008 ; Nambara et al. 2010), les graines dormantes imbibées conservent des niveaux élevés d'ABA (Ali-Rachedi et al. 2004).

D'autres hormones, telles que l'éthylène (Kepczynski & Kepczynski 1997 ; Linkies et al. 2009 ; Subbiah & Reddy 2010), les brassinostéroïdes (Clouse & Sasse 1998 ; Kucera et al. 2005), l'acide salicylique (Rajjou et al. 2006 ; Lee et al. 2010), les cytokinines (Mok & Mok 2001 ; Chiwocha et al. 2005 ; Subbiah & Reddy 2010), l'auxine (Kucera et al. 2005 ; Belin et al. 2009), l'acide jasmonique (Preston et al. 2009) et les oxylipines (Dave et al. 2011) influencent aussi la germination. Toutes ces phytohormones forment un réseau de signalisation très intriqué qui finalement régule très finement la germination, en particulier en réponse à des contraintes environnementales.

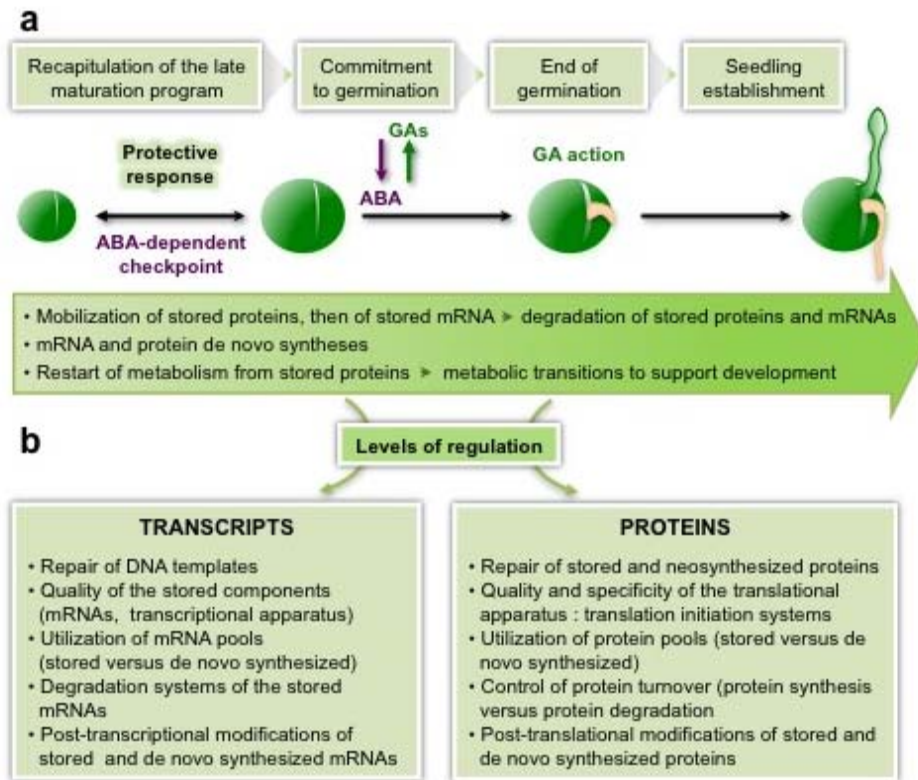


Figure 8 : importance des ARNm et des protéines stockés et synthétisés *de novo* dans le contrôle de la germination des graines (Rajjou et al. 2012). (a) Phénoménologie de la germination des graines. (b) Niveaux de régulation de l'accumulation des transcrits et des protéines pendant la germination.

Contre toute attente, une étude protéomique des graines du mutant GA-déficient *gal* d'*A. thaliana* a montré une implication très tardive des GAs dans la germination, survenant à un stade coïncidant avec l'émergence de la racicule (Figure 8) (Gallardo, Job, Steven P C Groot, et al. 2002). Il apparaît que les GAs, bien que nécessaires à l'achèvement de la germination, ne sont pas directement impliquées dans de nombreux processus qui se déroulent pendant la germination, comme la mobilisation initiale des protéines et des lipides de stockage des graines. Ainsi, sur les 46 modifications de protéines détectées pendant l'imbibition (24 HAI), une seule, correspondant à une  $\alpha$ -2,4-tubuline, semble dépendre de l'action des GAs. Cela implique que les graines imbibées du mutant *gal* sont métaboliquement actives, même si elles ne peuvent pas conclure la germination, c'est-à-dire l'émergence de la racicule (Gallardo, Job, Steven P C Groot, et al. 2002). Ce comportement a été confirmé par une étude de la germination des graines d'*A. thaliana* combinant la mesure



des niveaux endogènes des GAs par spectrométrie de masse en phase gazeuse et des analyses transcriptomiques sur les gènes régulés par les GAs (Ogawa et al. 2003).

Le processus de traduction joue un rôle essentiel au cours de la germination des graines (Rajjou et al. 2012). Ainsi, des plantes mutantes d'*A. thaliana*, correspondant à des doubles mutants (*i4g1/i4g2*) affectés dans l'expression de gènes essentiels pour l'initiation de la traduction présentent des taux réduits de germination et une réduction de la viabilité des graines (Lellis et al. 2010). Par ailleurs, l'utilisation d'inhibiteurs métaboliques a permis de montrer que la transcription *de novo* n'était pas nécessaire pour l'achèvement de la germination chez les graines d'*A. thaliana* (Rajjou et al. 2004). En effet, l'émergence de la radicule à travers les téguments est observable suite à l'imbibition des graines en présence d' $\alpha$ -amanitine, un inhibiteur très puissant et spécifique de l'ARN polymérase II, enzyme responsable de la synthèse des ARNm. Fait intéressant, une accumulation des sous-unités des globulines 12S et des membres de la famille des déhydrines (protéines LEA du groupe 2) a été observée chez les graines germées en présence d' $\alpha$ -amanitine, indiquant que, dans de telles conditions, les graines ont ré-initié leur programme de maturation (Rajjou et al. 2004). En revanche, la germination des graines d'*A. thaliana* est totalement bloquée en présence de cycloheximide, un inhibiteur de la traduction (Rajjou et al. 2004). Une étude de la germination des graines de riz a conduit aux mêmes conclusions (He et al. 2011). Ces résultats montrent ainsi que la synthèse des protéines est requise pour l'achèvement de la germination, en particulier à partir des ARNm stockés et dont la qualité conditionne la vigueur des graines (Sen & Osborne 1977 ; Smith & Bray 1982). Ces résultats mettent en évidence que chez *A. thaliana* et le riz, le potentiel de germination des graines est largement programmé au cours du processus de maturation. Le rôle crucial des mitochondries dans la germination des graines d'*A. thaliana* a été récemment démontré, notamment dans l'initiation de la traduction des protéines au sein de cet organite cellulaire (Law et al. 2012).

Chez *A. thaliana*, des variations dans l'abondance des ARNm sont initiées après seulement une à deux heures d'imbibition et une telle induction précoce n'est pas modifiée par le cycloheximide, bien que cette molécule bloque la germination à des stades ultérieurs (émergence radriculaire) (Kimura & Nambara 2010). Ainsi, contrairement à ce qui se passe en fin de germination (Rajjou et al. 2004), la synthèse protéique *de novo* n'est pas nécessaire pour l'expression génique pendant la phase d'imbibition initiale (phase 1 de la germination) et les protéines stockées dans les graines matures sont suffisantes à ce stade (Kimura &

Nambara 2010) (Figure 8). Une comparaison des résultats de Rajjou et al. (2004) (émergence racinaire) avec ceux de Kimura & Nambara (2010) (imbibition rapide, phase 1 de la germination) révèle l'importance des pools de protéines et d'ARNm stockés dans les graines matures sèches à des étapes successives du processus de germination. Il semble que très tôt, le pool de protéines stockées soit utilisé pour relancer l'activité cellulaire, suivi par la traduction des ARNm stockés (Figure 8). Ce rôle des protéines stockées renforce l'idée que la germination est préparée au cours de la maturation des graines. Des résultats similaires ont été rapportés pour la germination des graines d'orge (Sreenivasulu et al. 2008).

Des études protéomiques ont mis en évidence que les protéines des graines sont sujettes à un grand nombre de modifications post-traductionnelles (MPT), qui peuvent affecter leur fonction, leur localisation, leur interaction avec d'autres protéines, leur stabilité et leur activité (pour une revue, voir Arc et al. (2011) (Figure 8).

Ainsi, la germination des graines de céréales se caractérise par une profonde modification de l'état redox du protéome. Les protéines sont présentes principalement sous deux formes, une forme oxydée (SS) dans les graines sèches et une forme réduite (-SH) lors de l'imbibition (Buchanan & Balmer 2005). Ces études suggèrent un mécanisme dans lequel la forme oxydée du protéome chez les graines matures sèches rend compte de la quiescence métabolique, alors que le protéome est activé pendant la germination grâce à une réduction des ponts disulfures des protéines sous l'action de thiorédoxines (Buchanan & Balmer 2005). Une autre modification oxydative des protéines résulte de leur carbonylation par les ROS, un processus connu pour contribuer chez l'homme à diverses maladies (vieillesse, maladies neurodégénératives). De nombreuses protéines carbonylées s'accumulent au cours de l'imbibition des graines d'*A.thaliana* (Job et al. 2005). Cette accumulation vise des enzymes métaboliques, des facteurs de traduction et plusieurs chaperons moléculaires. Bien que l'accumulation de protéines carbonylées soit généralement considérée dans le contexte du vieillissement chez la plupart des systèmes étudiés, ce ne semble pas être le cas chez les graines. En effet, en dépit de l'accumulation de fortes quantités de protéines carbonylées suite à l'imbibition, les graines germent à un rythme élevé et conduisent à l'établissement de plantules vigoureuses (Job et al. 2005). Il semble que les changements spécifiques observés dans la carbonylation des protéines sont nécessaires pour contrecarrer la production de ROS causée par la reprise de l'activité métabolique dans les graines imbibés (Kranter, Roach, et al. 2010). Notamment, la carbonylation observée des abondantes protéines de stockage des

graines (globulines 11S) pourrait refléter leur rôle dans le piégeage des ROS produits au cours de l'imbibition des graines (Job et al. 2005).

Plusieurs protéine phosphatases et protéine kinases jouent un rôle éminent dans le contrôle de la germination à travers la modulation de la signalisation ABA (Brock et al. 2010 ; Hubbard et al. 2010). En outre, il a été proposé que la phosphorylation d'enzymes et de protéines impliquées dans la traduction des protéines (facteurs d'initiation et protéines ribosomiques) sont cruciales pour permettre la germination des graines (Le et al. 1998 ; Montoya-García et al. 2002).

Le radical NO<sup>o</sup> est un médiateur important et polyvalent dans les systèmes biologiques. Il peut en particulier causer la nitrosylation de la cystéine ou la nitration de la tyrosine des protéines de graines (Moreau et al. 2010 ; Arc et al. 2011). En faveur de son implication dans la germination, l'apparition transitoire de ce radical a été observée pendant les premières heures de l'imbibition des graines d'*A. thaliana* (Liu et al. 2010). Par ailleurs, une augmentation de plusieurs protéines nitrées a été observée dans les axes embryonnaires de sorgho après 24 h d'imbibition (Jasid et al. 2008).

Enfin, au-delà des modifications post-traductionnelles mentionnés ci-dessus, la contribution de modifications des protéines via leur biotinylation, glycosylation, ubiquitination, farnesylation ou acétylation a également été documentée lors de la germination (Arc et al. 2011).

Du fait de cet apport important de l'approche protéomique dans les travaux récents sur la physiologie des graines, nous précisons, dans le paragraphe suivant, les principales caractéristiques de ces techniques, que nous avons utilisées tout au long de cette thèse.

## **1.4 La protéomique**

Le terme PROTEOME désigne l'ensemble des PROTéines exprimées par le génOME d'une cellule ou d'un tissu, à un moment et dans un environnement donnés (stress, température, prolifération, différenciation...). Le concept est apparu en 1995 (Wasinger et al. 1995) pour désigner l'ensemble des produits des gènes fonctionnels d'un organisme vivant. L'analyse du protéome (protéomique) consiste en la caractérisation de l'ensemble des protéines et de leurs modifications s'exprimant dans un tissu ou un organisme donné à un moment donné ainsi que leur niveau d'accumulation. C'est donc un outil puissant pour déterminer le rôle fonctionnel des gènes et de leurs produits, complémentaire de l'analyse des transcrits. Cependant, cet outil présente quelques limitations, dont l'une réside dans le fait que

les protéines ne peuvent faire l'objet d'amplification quantitative, contrairement aux acides nucléiques (réactions de polymérisation en chaîne, PCR) qui représente une étape importante de l'analyse du transcriptome. Toutefois, les progrès technologiques en termes d'extraction, de séparation et d'identification des peptides et des protéines permettent aujourd'hui de caractériser de nombreuses protéines, même celles qui sont très faiblement représentées.

Avec l'achèvement de la séquence génomique d'un grand nombre d'organismes et le développement de méthodes analytiques pour la caractérisation des protéines, la protéomique est devenue un outil indispensable de génomique fonctionnelle (Bradshaw & Burlingame 2005 ; Rabilloud et al. 2010 ; Walther & Mann 2010) notamment chez les plantes (Agrawal et al. 2011). En particulier, plusieurs études protéomiques ont été réalisées récemment afin de caractériser le développement (maturation, dormance) et la germination des graines (cf. ci-dessus, pour des revues, voir Arc et al. 2011 et Rajjou et al. 2012). De plus, l'étude des voies métaboliques responsables de la synthèse et de l'accumulation des différentes réserves des graines est d'importance primordiale pour au moins deux raisons : (i) ces réserves conditionnent la vigueur des plantules après l'étape de germination et (ii) ces réserves sont largement utilisées en alimentation animale et humaine.

Le protéome est extrêmement complexe à plusieurs titres (cf. <http://ead.univ-angers.fr/~jaspard/Page2/COURS/9ModulGenFoncVeg/6Proteomique/1Proteomiq.htm>) :

- compte-tenu de l'épissage alternatif des transcrits primaires (plusieurs ARNm pour un gène) et compte-tenu des modifications post-traductionnelles des protéines, on peut estimer à plusieurs centaines de milliers les formes des protéines synthétisées dans les différents tissus humains par exemple ;
- Pour chaque condition environnementale (condition physiologique normale vs. conditions de stress) une cellule est caractérisée par un protéome adapté à cette condition alors qu'elle a toujours le même génome. Le cas des plantes est un exemple flagrant compte-tenu de leur nécessité de s'adapter tant aux variations de la lumière qu'aux effets de stress biotiques ou abiotiques ;
- Outre les modifications post-traductionnelles, les protéines subissent des transformations une fois synthétisées : clivage du peptide signal d'adressage, activation de la forme native à partir d'un précurseur (zymogène), assemblage en complexes oligomériques, association à des cofacteurs ;

- Il existe une grande dynamique de la synthèse des protéines : le rapport entre les protéines les moins abondantes et les plus abondantes dans une cellule peut être très variable ;
- Les protéines ont des temps de demi-vie très variables, s'étalant de quelques minutes à plusieurs jours.

Schématiquement, les étapes de l'analyse protéomique sont les suivantes :

- extraction des protéines,
- séparation des protéines par électrophorèse sur gel bidimensionnel,
- révélation des protéines dans les gels puis l'analyse d'image des gels,
- récupération des spots de protéines et digestion par des protéases (généralement la trypsine),
- obtention de cartes peptidiques massiques par des techniques de spectrométrie de masse,
- détermination de la séquence complète des protéines par des techniques de spectrométrie de masse dites en tandem,
- analyse bioinformatique (identification des protéines, annotation des protéines et des gènes, recherche de motifs structuraux).

La spectrométrie de masse est basée sur l'ionisation électronique. L'échantillon est introduit dans une enceinte sous vide où il est vaporisé puis soumis au bombardement d'un canon à électrons de grande énergie. Un électron est arraché aux molécules et on obtient une espèce qui est à la fois un cation (ion positif) et un radical libre (nombre impair d'électrons), que l'on appelle ion moléculaire. L'énergie du faisceau ionisant fragmente l'ion moléculaire par rupture des liaisons les plus faibles avant les liaisons les plus fortes et donne naissance à des ions positifs de masse plus faible, qui pourront être fragmentés à nouveau (eg. spectrométrie de masse dite en tandem - MS/MS). Ces ions sont ensuite accélérés dans un champ électrique et/ou magnétique puis dirigés entre les pôles d'un aimant selon une trajectoire circulaire qui dépend de leur rapport masse/charge  $[m/z]$ . En faisant varier le champ électrique, on fait varier la vitesse des ions moléculaires et on peut les faire ainsi parvenir au détecteur par ordre croissant de rapport  $[m/z]$ .

Les spectromètres de masse actuels permettent d'enregistrer automatiquement plusieurs milliers de spectres MS/MS en quelques heures. L'interprétation de ces spectres est le facteur limitant.

Parmi les techniques de protéomique quantitative, on peut citer celle dénommée protéomique en vrac ("*Shotgun Proteomics*") - "*Multidimensional Protein Identification Technology - MudPIT*" (Washburn et al. 2001). Son principe est le suivant :

- Un mélange de protéines est digéré par la trypsine.
- Le mélange de peptides qui en résulte est séparé par un système composé de deux gels de chromatographie successifs (en tandem), relié à un spectromètre de masse (MS/MS) en tandem : le premier gel est un échangeur de cations, le deuxième gel est une phase reverse.
- Une concentration de sel donnée permet l'élution de certains peptides de l'échangeur de cations vers la phase reverse.
- Les peptides qui se fixent sur la phase reverse sont élués par un gradient d'éthanol et analysés par spectrométrie de masse en tandem.
- Une fois l'élution de la phase reverse terminée, une concentration de sel plus importante est appliquée à l'échangeur de cations pour éluer d'autres peptides vers la phase reverse.
- Une vingtaine de ces cycles peuvent ainsi être réalisés.
- L'assignation des peptides à des protéines est effectuée par des algorithmes spécifiques comme SEQUEST (<http://en.wikipedia.org/wiki/SEQUEST>).

## Axes de recherche et objectifs de la thèse

Nous l'avons vu, la flore néo-calédonienne est riche d'une biodiversité extraordinaire. Son positionnement parmi les « hot-spots » de la biodiversité mondiale fait de la Nouvelle-Calédonie un lieu à préserver. L'intérêt pour la conservation des espèces a pris une ampleur considérable au cours des dernières années, comme en témoigne le « Millenium Seed Bank Project ». Ce programme initié en 2000 a pour objectif de collecter et de préserver 10% de la flore mondiale terrestre. De fait, l'Homme a à sa disposition un organe végétal extraordinaire capable de se conserver des dizaines, des centaines, voire des milliers d'années : la graine.

Cependant, afin de mieux protéger et conserver les espèces, il convient de les étudier pour mieux les connaître et comprendre leurs spécificités.

Les axes de recherche de cette thèse ont ainsi été orientés par les originalités de la flore néo-calédonienne et par l'importance générale des graines dans le cycle de développement des végétaux.

Ainsi, une première partie expérimentale de cette thèse est consacrée à l'étude de la graine d'*Amborella trichopoda* Baill.. Cette espèce, sœur de toutes les plantes à fleurs, est aujourd'hui endémique à la Nouvelle-Calédonie. Le séquençage récent de son génome (*Amborella* genome project 2013) témoigne de l'intérêt qui lui est accordé par la communauté scientifique. Nous avons ainsi profité de l'émulation scientifique autour de cette plante pour caractériser sa graine. Le but de cette étude débutée en octobre 2010 est de i) caractériser les processus contrôlant la germination de la sœur de toutes les plantes à fleurs, de ii) comprendre les mécanismes moléculaires responsables de la levée de dormance morphologique et de iii) replacer nos données dans un contexte phylogénétique.

Par ailleurs, du fait de la nature particulière de son sol, la Nouvelle-Calédonie est un des pays contenant le plus d'espèces hyperaccumulatrices de métaux lourds, notamment de nickel. Globalement, il existe un intérêt croissant pour la phytoremédiation des sols qui sont de plus en plus pollués. La caractérisation et l'utilisation des végétaux concentrant les métaux lourds nécessitent le développement de programmes de recherche afin de mieux comprendre les processus physiologiques responsables de l'accumulation. Fort de la présence de nombreuses espèces hyperaccumulatrices En Nouvelle-Calédonie, le Laboratoire Insulaire du Vivant et de l'Environnement (LIVE ; <http://www.univ-nc.nc/recherche/equipes-de-recherche#titre4>), en collaboration avec ses partenaires scientifiques locaux (IAC et IRD), a

développé un modèle d'étude comparative du genre *Psychotria* au sein duquel se trouvent une espèce hyperaccumulatrice, *Psychotria gabriellae* (Baill.) Guillaumin, et une espèce non-hyperaccumulatrice, *Psychotria baillonii* Schltr.. L'étude des graines de ces deux espèces a précédemment fait l'objet de la thèse de Charly Zongo au laboratoire (soutenance juillet 2010 ; [http://live.univ-nc.nc/index.php?option=com\\_content&task=view&id=30&Itemid=30](http://live.univ-nc.nc/index.php?option=com_content&task=view&id=30&Itemid=30)). Au cours du présent travail, nous nous sommes attachés à i) confirmer des résultats originaux précédemment obtenus et ii) créer une base de données des ARNm exprimés lors de la formation des fruits de *Psychotria gabriellae*. Cette dernière fera partie d'un nouveau programme de recherche plus large récemment financé par l'Agence Nationale pour la Recherche (ANR EvoMetoNicks).



## **Chapitre II. Matériels et méthodes**



# 1. Matériel végétal

## 1.1 Amborella trichopoda

### 1.1.1 Récolte des semences

Les graines d'*Amborella* ont été récoltées sur dix individus de la population du plateau de Dogny (long. 165.874 ; lat. -21.623) (Figure 9). Les fruits matures ont été prélevés et placés dans des sachets en papier le temps du transport jusqu'au laboratoire. Au laboratoire, les fruits ont ensuite été dépulpés sous l'eau courante par frottement sur un tamis (Fischer Scientific, maille 1 mm). Une fois la pulpe enlevée, les graines ont été séchées à l'air libre pendant 3 à 4 h.

La tolérance à la dessiccation des graines d'*Amborella* et leur aptitude à la conservation au sec n'étant pas encore clairement établies, toutes les manipulations sur les fruits ont été réalisées le plus rapidement possible après la récolte. Pendant ce court laps de temps, les fruits ont été conservés à l'obscurité dans des tubes hermétiques en plastique.

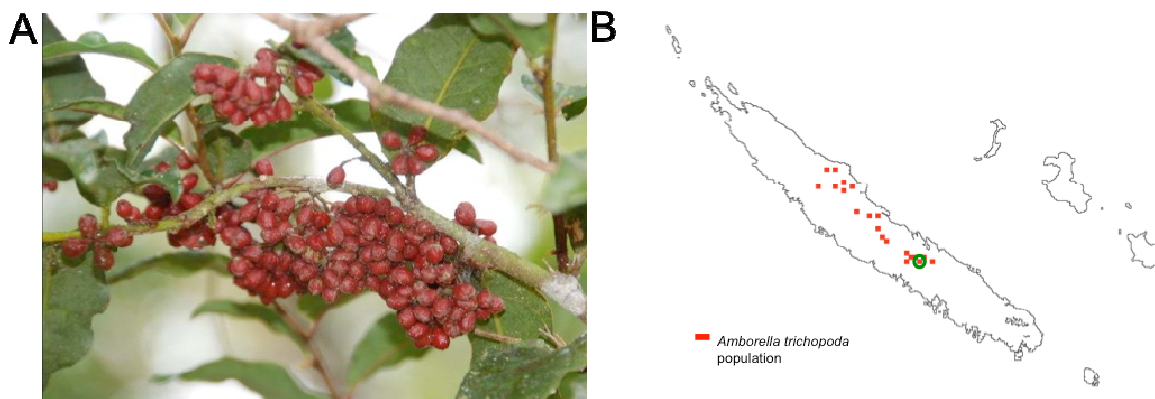


Figure 9 : photographie des fruits d'*Amborella* (A). Carte de la répartition des différentes populations d'*Amborella* sur la Grande-Terre (B) (points rouges). La population d'étude principale est signalée par un cercle vert.

### 1.1.2 Préparation des semences

Afin de limiter la présence de bactéries naturellement présentes à la surface des semences, celles-ci ont subi une désinfection de surface. Pour ce faire, les semences ont été placées pendant 5 min dans une solution d'eau distillée stérile contenant 2% d'hypochlorite de sodium et deux gouttes de Tween 20 (Prolabo 28829.296), sous agitation légère. Elles ont ensuite été rincées trois fois pendant 5 min dans des bains d'eau distillée stérile. Les graines ont ensuite été isolées à l'aide d'une pince et d'une lame de scalpel stériles. Une attention

particulière a été portée à ne pas toucher les téguments de la graine au moment de l'ouverture du méso-endocarpe sclérifié. En effet, si ce dernier est très dur, la graine est très fragile et tendre. En cas de doute sur l'intégrité des téguments de la graine, celle-ci a été éliminée.

### **1.1.3 Préparation des embryons**

Les embryons d'*Amborella* sont très petits et, à l'état sec, il est très difficile de les différencier de l'albumen. Afin de les extraire, les graines ont été coupées longitudinalement à l'aide d'une lame de rasoir stérile. Sous la loupe binoculaire, une goutte d'eau ultra pure a été déposée sur chacune des deux parties de la graine, au niveau de la zone où se trouve l'embryon. Après 3 à 4 s, l'embryon se distingue de l'albumen. Il est alors délicatement prélevé à l'aide d'une aiguille très fine fixée sur un porte aiguille. Aussitôt extrait, l'embryon a été placé dans un tube Eppendorf et plongé dans l'azote liquide.

### **1.1.4 Tests de germination**

Pour s'affranchir de la présence de bactéries et de champignons contaminant naturellement la surface des graines, ces dernières ont été placées pendant 5 min dans une solution d'eau distillée stérile contenant 2% d'hypochlorite de sodium et deux gouttes de Tween 20, sous agitation légère. Elles ont ensuite été rincées trois fois pendant 5 min dans des bains d'eau distillée stérile.

Les graines ont ensuite été mises à germer dans des boîtes de Pétri (diamètre 8 cm) sur deux épaisseurs de papier filtre imbibé d'eau distillée stérile ou d'une solution de GA<sub>3</sub> (Sigma – G1025) à 50 ppm (140 µM). Les boîtes de Pétri, fermées par leurs couvercles, ont ensuite été incubées à 25°C en conditions de lumière continue (tube néon Mazdafluor -18W - lumière du jour). Les papiers filtres ont été arrosés régulièrement afin qu'ils ne sèchent pas, toujours sous hotte à flux laminaire.

La germination de la semence d'*Amborella* a été réalisée dans les mêmes conditions (endocarpe + graine). La scarification de l'endocarpe a été réalisée en incubant les semences pendant 90 minutes en présence d'acide sulfurique 95% (Prolabo 20690.293) puis les semences ont été rincées abondamment à l'eau courante puis à l'eau distillé stérile avant d'être mises à germer comme les graines *sensu stricto*.

La germination a été suivie régulièrement et a été considérée comme accomplie lorsque la radicule est apparue au travers des téguments. Pour chaque expérience, 100 graines du même lot ont été mises à germer simultanément.

### 1.1.5 Obtention des différents stades de développement des graines d'Amborella suite à l'imbibition et la germination sensu stricto

Un lot de 200 graines a été mis à germer comme décrit ci-dessus. Après la germination de 20 d'entre elles, toutes les graines non germées restantes ont été découpées longitudinalement et sélectionnées en fonction de leur ratio entre longueur de l'embryon et longueur de la graine (E : S). Les graines dont le ratio était d'environ 0,25 ont été rassemblées pour créer le stade d'étude dénommé  $\frac{1}{4}$  (cf. Figure 40 du chapitre 3).

## 1.2 Genre *Psychotria*

Deux espèces ont été choisies dans le cadre de la construction d'un modèle d'étude comparatif entre une espèce hyperaccumulatrice de nickel et une autre espèce non hyperaccumulatrice. *Psychotria gabriellae* est hyperaccumulatrice, alors que *Psychotria baillonii* ne l'est pas, bien qu'étant tolérante au nickel. En plus d'être du même genre, ces plantes se développent au même endroit (Figure 10), ce qui permet de s'affranchir des effets dus à des variations environnementales (notamment les différences de composition du sol et les variations climatiques). Le choix de ce modèle comparatif a ainsi été fait au LIVE dans le cadre d'une approche globale visant à améliorer la compréhension des mécanismes de tolérance aux métaux lourds.

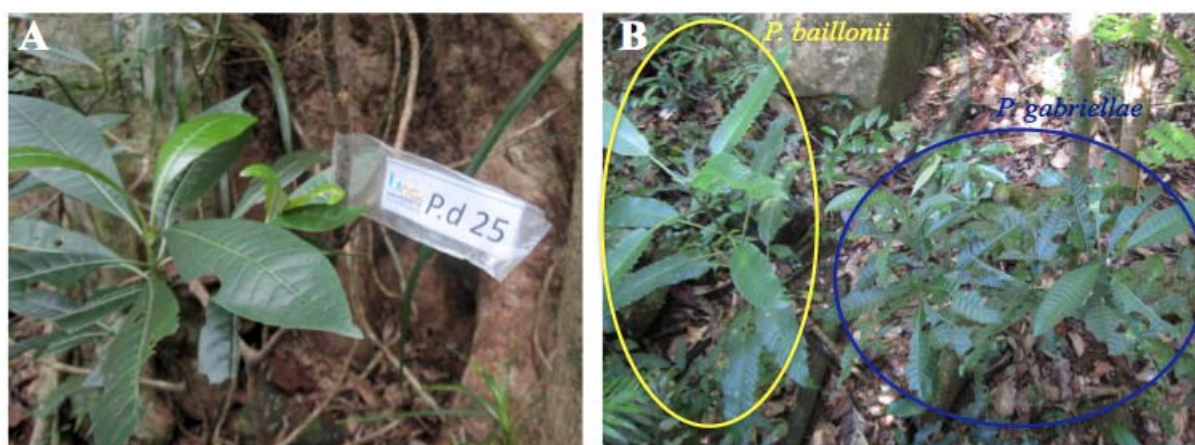


Figure 10 : plant de *P. gabriellae* marqué pour le suivi mensuel (A). Plants de *P. gabriellae* (en bleu) et de *P. baillonii* (en jaune) en conditions naturelles aux Monts Koghi.

Pour les travaux menés au cours de cette thèse, le site d'étude principal des collectes a été la forêt humide des Monts Koghi, où ces plantes se développent sur un sol ultramafique. Un suivi mensuel des plants de *P. gabriellae* et de *P. baillonii* a été réalisé afin de suivre le

développement des fruits pour assurer une récolte des graines au meilleur moment, c'est-à-dire lorsqu'elles atteignent leur stade final de maturité. Deux autres populations ont également été étudiées pour *P. gabriellae* (Poya et Mont-Dore). Les données GPS de ces populations sont présentées dans le Tableau 1 et sur la Figure 11.

Tableau 1 : positionnement des populations de *P. gabriellae* étudiées.

<b>Population</b>	<b>Longitude</b>	<b>Latitude</b>
Monts Koghi – Dumbéa	166.508	-22.176
Forêt Desmazures - Mont-Dore	166.624	-22.192
Nétéa – Poya	165.200	-21.231



Figure 11 : répartition des population de *P. gabriellae* étudiées sur la Grande-Terre.

### **1.2.1 Récoltes utilisées pour les analyses protéomiques**

Les fruits ont été récoltés et placés dans des sachets en papier le temps du transport jusqu'au laboratoire. Ils ont ensuite été dépulpés manuellement et les semences obtenues ont été séchées à l'air ambiant et à la lumière pendant 3 à 4 h, le temps que l'endocarpe redevienne sec. Les semences ont ensuite été conservées à 4°C dans des tubes hermétiques en plastique (Greiner Bio-One – CellStar – 227.261).

### 1.2.2 Récoltes utilisées pour l'approche transcriptomique

Les ARN messagers sont parmi les molécules biologiques les plus sensibles à la dégradation. Afin de préserver leur intégrité, une technique d'échantillonnage particulière a été mise en place sur le terrain. Les fruits à prélever ont été choisis à différents stades de maturité (Figure 12). Encore attachés au pied mère, ces derniers ont rapidement été aspergés d'eau savonneuse puis abondamment rincés à l'eau distillée stérile. Cette étape avait pour but de limiter au maximum la présence en surface des fruits de microorganismes et de RNAses.

La récolte des fruits a été faite ensuite sans les toucher à main nue (port de gants stériles), en les séparant du pied mère à l'aide d'une paire de ciseaux stériles, et en les stockant dans des tubes en plastique de 50 mL qui ont ensuite été plongés dans l'azote liquide.

Le stade I correspond au moment où les téguments sont encore photosynthétiquement actifs, permettant l'approvisionnement en énergie de l'embryon, notamment en assurant une source en oxygène pour les activités mitochondriales (Nakajima et al. 2012). L'ouverture de fruits à ce stade montre des graines très molles et difficilement distinguables des structures du fruit. Le stade II contient des graines dures et séparables du fruit et les chlorophylles commencent à être dégradées (Johnson-Flanagan & McLachlan 1990). Les stades III et IV contiennent tous deux des graines capables de germer. Certains fruits n'atteignent jamais le stade IV dont la pulpe est riche en pigments de type caroténoïdes ou flavonoïdes impliqués dans l'aptitude à la conservation au niveau de la graine (Bharti & Khurana 2003 ; Lepiniec et al. 2006).



Figure 12 : stades de maturité des fruits de *P. gabriellae* utilisés lors de l'approche transcriptomique.

## 2. Flore endophyte de *Psychotria gabriellae*

Les bactéries naturellement présentes dans les graines matures de *P. gabriellae* ont été caractérisées. Afin d'éliminer les bactéries non endophytes mais présentes à la surface des graines, celles-ci ont subi une désinfection par incubation pendant 5 min dans une solution d'eau distillée stérile contenant 5% d'hypochlorite de sodium sous agitation. Les graines ont ensuite été rincées trois fois pendant 5 min dans des bains d'eau distillée stérile.

### **2.1 Obtention et mise en cultures des bactéries**

Les graines désinfectées ont ensuite été broyées à l'aide d'un pilon et d'un mortier stériles, sous hôte à flux laminaire vertical. Le broyat de cinq graines ainsi obtenu a alors été remis en suspension dans 10 mL d'eau distillée stérile puis le mélange a été agité avec un agitateur mécanique à bras (Bioblock Scientific, Agitest 86212) pendant 5 min à 7 Hz (solution I). Une dilution au dixième du mélange obtenu (solution I) a ensuite été réalisée dans de l'eau distillée stérile (solution II). Au final, deux solutions de concentrations différentes ont été testées.

Les bactéries endophytes sont des organismes très sensibles et très exigeants en terme de conditions de culture. De fait, des milieux riches ont été choisis pour isoler de tels microorganismes. Il s'agit des milieux « viande-foie » (Bio-Rad – 64564) et « malt glucosé » (Panreac - 403690.1210) pH 6.0. La préparation a été effectuée selon les indications du fournisseur avec une concentration finale en agar de 20 g/L.

Des fractions de 200 µL de chaque solution (I et II) (trois répétitions par milieu de culture et par solution) ont été étalées sur des boîtes de Pétri (diamètre 8 cm) remplies de milieu de culture puis ont été incubées à 28°C à l'obscurité. L'apparition de colonies a été observée quotidiennement.

Les souches pures de bactéries endophytes ont été cultivées en milieu liquide afin de disposer d'assez de matériel pour pouvoir procéder à l'identification et la conservation des souches. Les colonies obtenues ont été prélevées à l'aide d'une anse et déposées dans 250 mL de milieu liquide « cœur-cervelle » (Fier – 413772.1210) à 37 g/L avec ou sans sulfate de nickel (100 ppm). Les suspensions ont été placées sous agitation à 37°C pendant 48 h.

Les souches bactériennes ont ensuite été conservées à -80°C dans des tubes Microbank™ (Pro-Lab diagnostic – PL.170/M).



## **2.2 Caractérisation des bactéries isolées**

Le séquençage des souches bactériennes isolées a été réalisé par l'entreprise MacroGen en Corée (<http://www.macrogen.com/eng/>). Les amorces utilisées visent la sous-unité 16S de l'ARN ribosomal, en utilisant le couple 518F (5'-CCAGCAGCCGCGGTAATACG-3') et 800R (5'-TACCAGGGTATCTAATCC-3'). Les séquences obtenues ont ensuite été identifiées par alignement sur BLAST (<http://blast.ncbi.nlm.nih.gov>).

## **3. Composition biochimique des réserves séminales des graines**

### **3.1 Glucides**

La méthode employée est adaptée de celle décrite par Roe (Roe 1995). Quatre répétitions de 20 mg de graines ont été finement broyées dans un broyeur à billes (Retsch) en présence de deux billes en inox pendant une minute avec une puissance maximum. Le carter cylindrique en inox dans lequel a été effectué le broyage des graines ainsi que les deux billes en inox ont été refroidis à -80°C préalablement au broyage. De l'eau distillée (1 mL) a été ajoutée à la poudre et le tout a été incubé à 45°C pendant 1 h. Les échantillons ont ensuite été centrifugés à 2 000 g pendant 15 min à 25°C. Le surnageant a alors été mis de côté avec précaution. De l'eau distillée (100 µL) a été mélangée avec 100 µL d'extrait aqueux de graines. Du réactif à l'antrone (2 mL) (Trevelyan et al. 1952) a été ajouté et l'ensemble a été homogénéisé à l'aide d'un vortex. Tous les échantillons ont été placés dans un bain-marie à 95°C pendant 10 min puis 10 min dans de l'eau glacée et enfin 10 min à température ambiante (20-25°C). Enfin, 4 mL de H<sub>2</sub>SO<sub>4</sub> (70%) ont été ajoutés dans chaque tube. Les glucides hydrosolubles ont alors été mesurés spectrophotométriquement par détermination de l'absorbance à 635 nm. Les résultats ont été déduits d'une courbe standard réalisée à partir de D<sup>+</sup>-glucose.

### **3.2 Protéines**

Les protocoles utilisés pour l'extraction des fractions protéiques décrites par Osborne (1924), albumines, globulines, prolamines et glutélines, ou de protéines totales à partir des graines sont détaillés dans le paragraphe 5.1.1.

### **3.3 Taux d'humidité**

Immédiatement après avoir extrait les graines des structures du fruit, quatre lots de 25 graines d'aspect fertile ont été triés. La teneur en eau de ces dernières a ensuite été mesurée suivant le protocole décrit par MacKay et al. (MacKay et al. 2002) et inspiré des procédures développées par l'International Seed Testing Association (ISTA 1999). Les graines ont été pesées avant puis après un traitement à  $103^{\circ}\text{C} \pm 2^{\circ}\text{C}$  pendant  $17\text{ h} \pm 1\text{ h}$  et le taux d'humidité a été déterminé.

## **4. Cytologie**

### **4.1 Inclusion en historésine**

Les graines ont été placées dans un tampon de fixation (paraformaldéhyde 2% (p/v) / glutaraldéhyde 1% (v/v) / caféine 1% (w/v), 0.05% Triton X-100 (v/v) dans un tampon de phosphate de sodium pH 7.2 (quatre infiltrations successives à l'aide d'une pompe à vide) et incubées toute la nuit à  $4^{\circ}\text{C}$  sous agitation circulaire lente. Après fixation, les graines ont été déshydratées progressivement sous agitation circulaire lente et sous filtration avec une série de solutions correspondant à 50% éthanol (EtOH) pendant 30 min, 70% EtOH pendant 30 min, 70% EtOH pendant 1 h, 95% EtOH deux fois pendant 30 min, 100% EtOH pendant 1 h, EtOH/Butan-1-ol (1:1 v:v) pendant 1 h, 100% Butan-1-ol toute la nuit à  $4^{\circ}\text{C}$ . Après déshydratation, les graines ont été imprégnées (toutes les étapes ont été réalisées sous agitation circulaire) dans du Butan-1-ol / Historésine Technovit (3v:1v) pendant 24 h à  $4^{\circ}\text{C}$ , Butan-1-ol / Historésine Technovit (2v:2v) pendant 24 h à  $4^{\circ}\text{C}$ , Butan-1-ol / Historésine Technovit (1v:3v) pendant 24 h à  $4^{\circ}\text{C}$ , Historésine Technovit 100%. Les graines ont alors été incluses dans 1 mL de solution d'inclusion d'historésine (Historésine/Durcissant 15:1) à température ambiante pendant 1 h puis à  $37^{\circ}\text{C}$  pendant au moins deux jours. Les graines ont ensuite été coupées à l'aide de microtome réglé à  $5\ \mu\text{m}$  d'épaisseur.

### **4.2 Inclusion dans l'agarose**

Les graines ont été fixées dans de l'agarose à faible point de gélification (Sigma – A9414) lorsque la température de la solution était inférieure à  $40^{\circ}\text{C}$ . La solidification a été obtenue à température ambiante pendant 4 h.

### **4.3 Coloration au rouge Nil**

Le rouge Nil permet la coloration spécifique des corps lipidiques (Greenspan et al. 1985). Les graines incluses dans l'agarose ont été coupées longitudinalement à l'aide d'un microtome à lame vibrante (Leica, VT1000S), réglé à une épaisseur de 30 µm.

La solution de travail de rouge Nil a été préparée extemporanément, en diluant la solution mère (1 mg/mL de rouge Nil dans de l'acétone 100%) au centième dans du glycérol 50% (v/v). Les coupes y ont été incubées 30 min à l'obscurité puis directement placées entre lame et lamelle dans une goutte d'eau.

L'observation a été réalisée à l'aide d'un microscope confocal (Leica, TCS SP2, Leica Microsystems SAS, France) avec une excitation à 485 nm et une émission à 525 nm.

### **4.4 Double coloration au bleu noir naphthol et acide périodique-réactif de Schiff**

Cette double coloration permet la visualisation simultanée des réseaux de polysaccharides (acide périodique-réactif de Schiff) (Puchtler et al. 1975) et des corps protéiques et nucléoles (bleu noir naphthol, aussi appelé noir amide 10B) (Mathe & Vieillescazes 2002).

Les coupes obtenues à partir des graines fixées dans la résine ont été hydrolysées 5 min dans de l'acide périodique (1%) puis rincées à l'eau distillé. La première coloration au réactif de Schiff a été réalisée pendant 10 min à l'obscurité. Les coupes ont ensuite été rincées deux fois dans l'acide sulfurique (0,25% de métabisulfite de sodium, 0,05 M d'acide chlorhydrique) puis à l'eau courante et à l'eau distillée jusqu'à ce que l'eau de lavage soit limpide.

Les coupes ont ensuite été traitées pendant 5 min avec la solution de bleu noir naphthol préalablement chauffée à 60°C (1% bleu noir naphthol Sigma – N3005 dans de l'acide acétique 7%) puis rapidement rincées à l'eau courante. La régression de la coloration a été obtenue avec de l'acide acétique (7%) sous microscope (microscope optique Leica Zeiss Axioplan, Leica Microsystems, Wetzlar, Allemagne).

## 5. Etude protéomique

Deux approches différentes ont été mises en œuvre, l'une basée sur la séparation des protéines par électrophorèse bidimensionnelle et l'autre, plus globale, nommée protéomique en vrac (*shotgun*).

### 5.1 Approche fondée sur l'électrophorèse bidimensionnelle des protéines

#### 5.1.1 Extraction des protéines de graines selon Osborne (1924)

L'extraction selon Osborne (1924) est la méthode de référence concernant l'extraction des protéines des graines. Cette méthode permet d'extraire successivement quatre grandes classes de protéines selon leur solubilité dans différents solvants. Ainsi, on obtient les albumines dans l'eau, les globulines dans les sels, les prolamines dans l'alcool et les glutélines dans la soude. Les globulines, prolamines et glutélines sont des protéines de réserve de la graine utilisées pour le démarrage du métabolisme lors des premières phases de croissance de la plantule (Bewley & Black 1994). Les albumines correspondent à l'ensemble des protéines hydrosolubles de la cellule et, en particulier, aux enzymes du métabolisme (Figure 13). Toutes les manipulations ont été réalisées à 4°C.

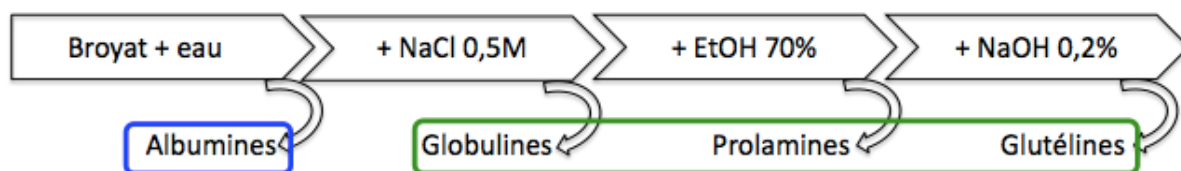


Figure 13 : extraction fractionnée des protéines de graines, d'après Osborne (1924).

Dans le cas de *P. gabriellae*, l'endocarpe des graines (environ 20 graines, trois réplicats) a été ôté et les graines ont été placées à -80°C pendant 20 min avant un broyage fin par agitation dans un broyeur à billes (chaque tube de broyage contient deux billes en acier inox). L'ensemble est agité par un broyeur à billes (Retsch, MM 301) 2 min à pleine puissance.

Dans le cas d'*Amborella*, les graines (environ 25, trois réplicats) ont été broyées dans l'azote liquide à l'aide d'un pilon et d'un mortier. Le broyat obtenu a ensuite été lyophilisé toute la nuit, placé dans des tubes scellés, pesé puis mélangé avec de l'eau MilliQ contenant

des inhibiteurs de protéases (Roche complete – 1 pastille pour 10 mL) à raison de 1 mL pour 250 mg de broyat (0,25 p:v). Le tout a été agité pendant 15 min en chambre froide. La suspension a été centrifugée (Sigma type 3K30) pendant 15 min à 15 000 g à 4°C. Le surnageant contenant les albumines a été récupéré. Trois lavages du précipité ont été effectués en utilisant de l'eau MilliQ et des inhibiteurs de protéases. Le culot a alors été de nouveau solubilisé avec du NaCl 0,5 M à raison de 1 mL de solution pour 250 mg de broyat sous agitation pendant 15 min en chambre froide. Après centrifugation (15 000 g ; 15 min ; 4°C), le surnageant contenant les globulines a été recueilli et le culot lavé de la même manière que précédemment, mais avec du NaCl 0,5 M. La troisième extraction a été réalisée comme les deux précédentes mais avec un mélange éthanol : eau (70:30). Le surnageant obtenu contenant les prolamines a été récupéré, le culot lavé trois fois avec de l'éthanol 70% puis solubilisé dans une solution de NaOH à 0,2% (p:v). Après centrifugation (15 000 g ; 15 min ; 4°C), le surnageant contenant les glutélines a été récupéré.

Pour concentrer les protéines et éliminer des contaminants pouvant interférer avec l'électrophorèse (e.g., polyphénols, lipides) les quatre fractions protéiques obtenues ont été précipitées pendant la nuit à -20°C après ajout d'un volume d'un mélange contenant acide trichloracétique (TCA) 20% / acétone 80% contenant du 1,4-dithiothréitol (DTT) à 1 mM final. Après précipitation, les échantillons ont été centrifugés (15 000 g ; 15 min ; 4°C) puis les culots ont été lavés deux fois avec 500 µL d'une solution d'acétone contenant 1 mM de DTT. Les culots ainsi lavés ont été séchés sous hotte puis repris dans un volume de tampon de réhydratation permettant d'obtenir une concentration finale en protéine d'environ 5 µg/µL. Le tampon de réhydratation est constitué de 25 mL d'eau MilliQ, 21 g d'urée (7 M), 7,6 g de thio-urée (2 M), 2 g de CHAPS (3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonate) à 4% (p :v), 0,5 mL d'ampholytes 1% (p :v), 1 mL de Triton 100X et de quelques gouttes de bleu de bromophénol. Le DTT a été rajouté extemporanément à raison de 1 g pour 50 mL de tampon.

### **5.1.2 Extraction des protéines totales de graines**

#### **d'Amborella**

Les graines ont été broyées dans l'azote liquide à l'aide d'un pilon et d'un mortier jusqu'à l'obtention d'une poudre fine. Puis, pour 100 mg de graines, 640 µL de tampon d'extraction et 107 µL d'inhibiteur de protéases (Roche, Complete mini protease inhibitor – 1 pastille dans 1,5 mL d'eau MilliQ) ont été ajoutés, et le tout a été de nouveau broyé comme précédemment.

Le tampon d'extraction est constitué de Trizma-HCl (18 mM), Trizma-base (14 mM), Triton 100X (0,2 %), urée (7 M), thio-urée (2 M), CHAPS (4%) (Gallardo et al. 2001).

La suspension ainsi obtenue a été récupérée et additionnée de 15  $\mu$ L de DTT 1 M, 4  $\mu$ L de DNase (10 u/ $\mu$ L) et 24  $\mu$ L de RNase (100 u/mg). L'ensemble a été mélangé au vortex puis placé sous agitation douce pendant 1 h en chambre froide. La suspension a ensuite été centrifugée à 35 000 g pendant 20 min à 4°C puis le surnageant contenant les protéines a été récupéré. S'il contenait des particules en suspension, une centrifugation de clairance a été réalisée (35 000 g ; 10 min ; 4°C). Une précipitation au TCA a été réalisée en fonction de la présence de polyphénols et/ou de lipides, comme décrit ci-dessus. Les culots ont ensuite été repris dans un volume de tampon de réhydratation permettant d'obtenir une concentration finale en protéines d'environ 5  $\mu$ g/ $\mu$ L.

### **5.1.3 Dosage des protéines**

Le dosage des protéines des échantillons extraits a été effectué grâce à une réaction colorimétrique avec un réactif Biorad (Total Protein Assay) dilué au 1/5<sup>ème</sup> dans de l'eau ultra pure. Une gamme étalon d'albumine de sérum bovin (BSA) a d'abord été préparée à partir d'une solution mère de concentration connue à 40 mg/mL et la densité optique de chaque dilution a été mesurée à 595 nm, permettant d'établir, par régression linéaire, la correspondance entre l'absorbance mesurée à 595 nm et la concentration de BSA. Ensuite, l'absorbance des différents extraits a été mesurée afin d'en déduire leur concentration en protéine. Deux dosages distincts ont été effectués, l'un avant précipitation au TCA et l'autre déterminant la concentration finale réelle en protéines.

### **5.1.4 Electrophorèse SDS-PAGE**

#### ***5.1.4.1 Electrophorèse monodimensionnelle***

Les échantillons, de concentrations protéiques connues, ont été dilués dans du tampon de charge afin d'obtenir une concentration finale de 1 $\mu$ g/ $\mu$ L. Dix  $\mu$ L (soit 10  $\mu$ g de protéines) de chaque échantillon et 5  $\mu$ L de marqueur de taille (Biorad – 161-0373) ont été déposés dans les puits du gel dont la composition est décrite dans le Tableau 2. La solution de « *stacking* » est composée de 0,17 mL d'un mélange acrylamide/PDA (piperazine-di-acrylamide) (100 mL/1,33 g) ; 0,25 mL de Tris (trishydroxyméthylaminométhane)-HCl pH 6,8 ; 0,58 mL d'eau et 10  $\mu$ L de SDS 10%. La migration a été effectuée dans un tampon Taurine 1X contenant 0,1% de SDS (Catusse, Job, et al. 2008).

Tableau 2 : volumes des réactifs utilisés pour la réalisation d'un gel monodimensionnel de polyacrylamide.

	<b>Acrylamide 10%</b>	<b>Solution de stacking</b>	<b>APS 25%</b>	<b>TEMED</b>
<b>Gel de migration</b>	4 mL		8 $\mu$ L	4 $\mu$ L
<b>Polymérisation avec de l'isopropanol en surface</b>				
<b>Gel de concentration</b>		1 mL	4 $\mu$ L	2 $\mu$ L

Les gels ont été soumis à un courant continu à ampérage constant, une première phase à 20 mA permettant la pénétration des échantillons dans le gel et la traversée de la zone de concentration. Une fois que les échantillons ont atteint la zone de migration, l'ampérage a été augmenté à 40 mA. Les gels ont été lavés cinq fois dans de l'eau MilliQ pendant 5 min puis mis en présence d'une solution de GelCode<sup>®</sup> (Thermo Scientific - 24596) pendant une heure sous agitation douce. Les gels ont alors été mis à décolorer dans une solution de glycérol à 5% jusqu'à ce que le fond s'éclaircisse. Ils ont ensuite été séchés entre deux feuilles de cellophane tendues sur un cadre en plexiglass.

#### **5.1.4.2 Electrophorèse bidimensionnelle**

La première dimension (isoélectrofocalisation) a été réalisée à l'aide de bandelettes (ou « *strips* ») déshydratées de gel de polyacrylamide contenant un gradient de pH immobilisé (3-10 non linéaire) sur une longueur de 11 ou 18 cm. Une étape de réhydratation a été réalisée. Les bandelettes, placées dans une cassette de réhydratation, ont été mises en contact, gel dessous, de 210  $\mu$ L (strips de 11 cm) ou 340  $\mu$ L (strips de 18 cm) de la solution contenant l'échantillon et le tampon de réhydratation. Les bandelettes ont ensuite été recouvertes d'huile minérale et la cassette a été placée à 22°C, la réhydratation se déroulant sur toute la nuit (un minimum de 12 h est nécessaire) (Catusse et al. 2008).

Dans le cas de strips de 11 cm (équipement Biorad), les bandelettes ont été déposées dans une cassette après élimination de l'huile minérale par simple égouttement. Chaque extrémité des bandelettes a été posée, face gel au dessous, sur un morceau de papier Wattman préalablement imbibé d'eau MilliQ pour assurer un meilleur contact avec les électrodes. Les bandelettes ont ensuite été recouvertes d'huile minérale et la cassette a été positionnée dans l'appareil d'isoélectrofocalisation (Biorad – Protean IEF cell 165-4000). Le programme de migration suivant a été utilisé : une première phase avec un gradient rapide à 100 V pendant 30 min, une seconde avec un gradient linéaire à 250 V pendant 30 min, une troisième avec un gradient linéaire à 1 000 V pendant 30 min, une quatrième avec un gradient linéaire à 7 000 V

pendant 3 h, une cinquième avec un gradient linéaire à 7 000 V jusqu'à ce que le voltage atteint soit de 32 000 V puis une sixième phase avec un gradient rapide à 500 V, ce qui permet de maintenir les protéines à leur pI jusqu'à l'arrêt de l'appareil par l'expérimentateur. L'ampérage a été limité à 50  $\mu$ A par bandelette et l'isoélectrofocalisation a été effectuée à une température de 20°C.

Dans le cas des strips de 18 cm (équipement Pharmacia), les bandelettes ont été déposées dans une cassette d'isoélectrofocalisation après élimination de l'huile avec la face gel au dessus. Du papier filtre épais abondamment imbibé d'eau MilliQ a été déposé à chaque extrémité des bandelettes puis les électrodes ont été disposées par dessus et les bandelettes ont été recouvertes d'huile minérale. Avant le dépôt de la cassette dans l'appareil, une goutte de kérosène a été ajoutée pour assurer un contact thermique entre la cassette et la plaque à température régulée. Le programme de migration a alors été lancé avec une première étape d'une heure à 300V (2 mA et 5 W) suivie par une migration à 3 500 V (2 mA et 5 W) pendant 7 h.

Une solution d'équilibration a été préparée avec 37,5 mL de tampon de « stacking 8X », 90 mL de glycérol, 50 mL d'eau MilliQ, 75 mL de SDS 10% et 109,8 g d'urée, sous agitation. Le tampon de « stacking » correspond à 251 g de bis-Tris dissous dans 800 mL d'HCl molaire et de l'eau MilliQ (qsp 1 L). L'équilibration a débuté par ajout d'une solution de 100 mL de tampon d'équilibration avec 1 g de DTT aux bandelettes pendant 20 min sous agitation. Les bandelettes ont ensuite été lavées deux fois avec du tampon d'équilibration pendant au moins 1 min avec au minimum 50 mL de tampon. Au final, 100 mL de tampon contenant 4 g d'iodoacétamide (IAM) ont été ajoutés pendant 20 min sous agitation. Les bandelettes ont finalement été relavées pendant 1 min avec le tampon de migration. Après cette étape, les strips peuvent être conservées à -80°C ou être déposées en seconde dimension.

Pour la migration en gels bidimensionnels, 25 gels SDS-PAGE peuvent être préparés simultanément, ce qui conduit à préparer une solution contenant 1 L d'eau MilliQ dégazée avec 400 mL d'acrylamide/PDA (13,3 g de PDA dans 1 L de solution stock d'acrylamide 40%) puis 266 mL de tampon Trizma-HCl pH 8,3 (130 g Trizma, 600 mL d'HCl molaire, eau MilliQ qsp 1 L). La polymérisation est catalysée par l'APS (1,4 g dans 10 mL d'eau MilliQ) et du TEMED (665  $\mu$ L) rajoutés extemporanément. Les gels ont été coulés et ont été polymérisés en 1 h.



La migration en gels 2D a été effectuée dans 20 L de tampon taurine contenant 500 g de taurine, 60 g de trizma dilué dans un litre d'eau MilliQ, 200 mL de SDS 10% et de l'eau MilliQ qsp 20 L. Avant le dépôt des bandelettes sur le haut du gel, de l'agarose a été ajouté. Pour 200 mL de cette solution, 25 mL de tampon de « stacking 8X », 8 mL de SDS 10% et 2 g d'agarose « Low Melting Point » ont été mis sous agitation et chauffage pour dissoudre l'agarose. Les bandelettes ont été mises en place une fois l'agarose refroidi mais non polymérisé. Une fois l'agarose polymérisé, les plaques ont été disposées dans la cuve et le voltage a été ajusté en fonction du nombre de gels et du temps de migration désiré (habituellement sur la nuit).

#### 5.1.4.3 Révélation des gels à l'argent

La fixation des protéines dans les gels a été effectuée pendant 1 h dans une solution contenant de l'éthanol et de l'acide acétique (600 mL et 100 mL, H<sub>2</sub>O qsp 2 L). Les gels sont ensuite transférés dans un automate pour coloration (Hoefer Automated Gel stainer apparatus, Amersham Pharmacia biotech.) (2 gels/cuve). La fixation a été suivie par un lavage à l'eau MilliQ de 5 min puis de trois lavages de 10 min. Les gels ont ensuite été sensibilisés avec le thiosulfate de sodium (Tableau 3) pendant 1 min 30 s puis lavés deux fois à l'eau MilliQ pendant 1 min. Les gels ont ensuite été mis en présence de nitrate d'argent (Tableau 3) pendant 10 min et lavés deux fois à l'eau MilliQ pendant 1 min. La coloration a été révélée par ajout d'une solution de révélation (Tableau 3) pendant 20 min et stoppée par transvasement dans un bain d'EDTA (éthylène diamine tétraacétique) (30 g d'EDTA pour 2 L d'eau MilliQ).

Tableau 3 : quantités des réactifs intervenant dans la révélation à l'argent des gels.

	Sensibilisation	Coloration		Révélation		
	Thiosulfate de sodium 15%	Nitrate d'argent	Formaldéhyde	Carbonate de potassium	Formaldéhyde	Thiosulfate de sodium 15%
400 mL (2 gels)	0,8 mL	4,8 µL	120 µL	12 g	120 µL	250 µL

Les gels ont finalement été lavés deux fois dans de l'eau MilliQ pendant 5 min avant d'être placés dans la solution de conservation (« preserving ») contenant de l'éthanol et du glycérol (600 mL EtOH, 100 mL glycérol, H<sub>2</sub>O qsp 2 L) pour une durée de 45 min. Ils ont

ensuite été séchés durant 24 h entre deux feuilles de papier cellophane poreux tendues sur un cadre en plexiglass avant d'être scannés et analysés par le logiciel SameSpots (Progenesis).

#### **5.1.4.4 Révélation des gels à l'argent modifié (méthode de Schevchenko)**

La fixation des protéines dans les gels a été effectuée pendant 1 h dans une solution contenant de l'éthanol et de l'acide acétique (600 mL et 100 mL dans 2 L final).

Les gels ont ensuite été transférés dans un automate pour coloration (Hoefler Automated Gel stainer apparatus, Amersham Pharmacia biotech.) (2 gels/cuve). La fixation et a été suivie de deux lavages à l'eau MilliQ de 10 min. Les gels ont été sensibilisés avec le thiosulfate de sodium pendant (Tableau 4) 2 min puis lavés deux fois dans de l'eau MilliQ pendant 1 min. Puis les gels ont ensuite été mis en présence de nitrate d'argent (Tableau 4) pendant 20 min et lavés deux fois à l'eau MilliQ pendant 1 min. La coloration a été révélée par ajout d'une solution de révélation (Tableau 4) pendant 5 min et stoppée par transvasement dans un bain d'acide acétique à 5% (Shevchenko et al. 1996).

Les gels ont finalement été lavés deux fois à l'eau MilliQ pendant 1 min avant d'être placés dans la solution de stockage contenant de l'acide acétique 1%. Les gels ont été conservés à 4°C en chambre froide avant prélèvement des spots d'intérêt. Cette méthode est compatible avec la spectrométrie de masse car elle permet de diminuer la quantité de réactifs et surtout la quantité en formaldéhyde. Sa faible présence a pour conséquence une coloration qui reste en surface des gels. Ainsi l'analyse par spectrométrie de masse pourra être réalisée sur la partie non colorée du spot, enfouie dans le gel.

Tableau 4 : quantités des réactifs intervenant dans la révélation à l'argent modifié (Shevchenko et al 1996).

	Sensibilisation	Coloration		Révélation	
	Thiosulfate de sodium 20%	Nitrate d'argent	Formaldéhyde	Carbonate de sodium	Formaldéhyde
400 mL (2 gels)	0,533 mL	2,4 µL	432 µL	8 g	432 µL

#### ***5.1.4.5 Identification des protéines par spectrométrie de masse***

Les spots d'intérêt ont été excisés à partir de gels 2D colorés à l'argent modifié (coloration de Schevchenko) avec des embouts de pipette stériles sous une hotte à flux laminaire horizontal et placés dans des microplaques stériles. La digestion en gel des protéines a ensuite été effectuée avec le système Progest (Solution génomique) en utilisant la trypsine. La coloration argentique a été éliminée selon Dunn (1996) (Dunn 1996). Les morceaux de gel ont été lavés deux fois par bains successifs de 10% d'acide acétique, d'éthanol à 40%, et d'acétonitrile (ACN). Ils ont ensuite été lavés deux fois avec des bains successifs de 25 mM de  $\text{NH}_4\text{CO}_3$  et d'ACN. La digestion a ensuite été effectuée pendant 6 h à 37 ° C avec 125 ng de trypsine (Promega) dissous dans 20% de méthanol et 20 mM de  $\text{NH}_4\text{CO}_3$ . Les peptides ainsi générés ont été extraits successivement avec de l'acide trifluoroacétique à 2% (TFA) et 50% d'ACN puis avec de l'ACN pur. Les extraits peptidiques ont été séchés dans une centrifugeuse sous vide et en suspension dans 20  $\mu\text{L}$  de TFA à 0,05%, 0,05%  $\text{HCOOH}$ , et 2% d'ACN. La séparation des peptides par HPLC a été réalisée sur un système NanoLC-Ultra (Eksigent). Un échantillon de 4  $\mu\text{L}$  a été chargé à 7,5  $\mu\text{L}/\text{min}^{-1}$  sur une cartouche en pré-colonne (phase stationnaire : Biosphere C18, 5  $\mu\text{m}$  ; colonne: 100  $\mu\text{m}$  de diamètre interne) et dessalé avec 0,1% de  $\text{HCOOH}$ . Après 3 min, la cartouche en pré-colonne a été reliée à la colonne de séparation C18 PepMap (phase stationnaire : Biosphere C18, 3  $\mu\text{m}$  ; colonne: diamètre interne de 75  $\mu\text{m}$ , 150 mm: Nanoseparations). Les tampons utilisés étaient  $\text{HCOOH}$  (0,1%) dans l'eau (A) et  $\text{HCOOH}$  (0,1%) dans l'ACN (B). La séparation des peptides a été réalisée avec un gradient linéaire de 5 à 30% de B pendant 11 min à un débit de 300  $\text{nL}/\text{min}^{-1}$ . Avec l'étape de régénération à 95% de B et l'étape de remise en équilibre à 95% de A, la totalité de la séparation des peptides a été effectuée sur un laps de temps de 25 min. Les peptides élués ont été analysés en ligne avec une trappe ionique LTQ XL (Thermo Electron) à l'aide d'une interface nanoelectrospray.

Une recherche des masses peptidiques sur la base de données génomiques chez *Amborella* a été réalisée avec XTandem (version 2010.12.01.1) (<http://www.thegpm.org/TANDEM/>). Le clivage enzymatique a été déclaré comme résultant d'une digestion par la trypsine en considérant une seule possibilité, au plus, de mauvais clivages possibles. Les modifications de carboxyamidométhylation des résidus Cys et l'oxydation de la Met ont été déclarées comme modifications statiques et possibles, respectivement. La tolérance de la mesure de la masse de l'ion précurseur et des fragments ont été de 2,0 et 0,5, respectivement. Une recherche de raffinement a été ajoutée avec des

paramètres similaires, sauf que le peptide semi-trypsique et les modifications d'acétylation N-terminale des peptides ont été recherchés. Seuls les peptides avec une valeur d'E inférieure à 0,1 ont été consignés. Les protéines identifiées ont été filtrées et regroupées en utilisant le pipeline XTandem (<http://pappso.inra.fr/bioinfo/xtandempipeline/>) en tenant compte de : (1) la détection d'un minimum de deux peptides différents avec une valeur E inférieure à 0,03, (2) d'une valeur de E en protéines (calculée comme le produit des valeurs de E peptidiques uniques) inférieure à  $10^{-3}$ . En cas d'identification avec seulement deux ou trois spectres MS/MS, la similitude entre les spectres MS/MS expérimentaux et théoriques a été vérifiée visuellement.

## **5.2 Approche de protéomique en vrac (shotgun)**

Dans le cas des échantillons de graines d'*Amborella*, l'extraction des protéines totales a été réalisée à partir de trois répétitions de 100 embryons isolés et de 20 portions d'albumen (débarassées des embryons). Ces matériels ont été broyés dans l'azote liquide avec un mortier et un pilon. Ensuite, les protéines totales ont été extraites à 25°C dans un tampon de lyse composé de thiourée / urée 400 (7 M d'urée, 2 M thiourée, Tris-HCl pH 6,4, 2 mM Trizma<sup>®</sup> base (Sigma-Aldrich, Lyon, France)), 4% (p / v) CHAPS) supplémenté avec 50 µL du cocktail inhibiteurs de protéases Mini Complete (Roche Diagnostics France, Meylan, France). Ensuite, 15 µL de DTT (1 M), 2 µL de DNase I (Roche Diagnostics) et 5 µL de RNase A (Sigma-Aldrich) ont été ajoutés à l'échantillon. Enfin, les extraits protéiques ont été agités pendant 2 h à 4°C. Tous les échantillons ont ensuite été centrifugés à 20 000 g à 4°C pendant 15 min, et les surnageants résultants ont été soumis à une deuxième centrifugation (20 000 g à 4°C pendant 15 min). Les surnageants ont été conservés à -80°C et les concentrations de protéines dans les différents extraits ont été mesurées selon Bradford (1976) (Bradford 1976) en utilisant l'albumine de sérum bovin comme standard. Ils ont ensuite été soumis à une analyse en spectrométrie de masse comme décrit ci-dessus (paragraphe 5.1.4.4).

## **5.3 Construction de la base de données protéomiques des graines d'*Amborella***

Une fois les protéines mises en évidence par spectrométrie de masse, les résultats d'identification contre la toute récente base de donnée d'*Amborella* (*Amborella* genome project 2013) étaient de la forme : « evm\_27.model.AmTr\_v1.0\_scaffold00004.99 ». L'ensemble des séquences peptidiques prédites à partir du génome d'*Amborella* ont alors été

extraites puis analysées par alignement en BLAST sur l'espèce végétale modèle *Arabidopsis thaliana*. Lorsque cet alignement ne donnait aucun résultat, la séquence peptidique prédite d'*Amborella* a alors été analysée par alignement sur BLAST (<http://blast.ncbi.nlm.nih.gov>) contre toutes les bases de données protéiques disponibles.

Une fois les protéines des graines d'*Amborella* identifiées, leur rôle dans la physiologie de la graine a été exploré, notamment grâce à la bibliographie disponible sur PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>), mais également en utilisant un site plus spécialisé dans le référencement de l'implication des protéines dans les voies métaboliques spécifiques, notamment du métabolisme des plantes (<http://metacyc.org>).

L'ensemble de ces analyses nous a permis d'assigner une fonction à la majorité des protéines identifiées. Les catégories et sous-catégories fonctionnelles sont celles définies par la classification de Bevan et al. (Bevan et al. 1998).

Pour certaines protéines d'*Amborella* (eg. globulines 11S) nous avons conduit des analyses phylogéniques. Dans ce cas, les arbres phylogénétiques ont été construits à partir des séquences protéines d'intérêt de diverses espèces. Les alignement ont été réalisés avec le logiciel MAFFT (*Multiple Alignment using Fast Fourier Transform* ; <http://www.ebi.ac.uk/Tools/msa/mafft/>) et les arbre construits avec la méthode du « maximum likelihood » à partir des séquences disponibles. Les analyses de bootstrap ont été fixées à 100 pseudo-réplicats.

## **6. Immunodetection des protéines**

Afin de rapidement, et spécifiquement, identifier certaines protéines présentes dans les extraits de graines, les échantillons de graines de *P. gabriellae* et/ou d'*Amborella* ont été soumis à extraction, migration en gels puis à des test ELISA après électroélution de protéines d'intérêt (*Amborella*) ou étude par Western Blot (genre *Psychotria*).

### **6.1 Obtention des anticorps**

#### **6.1.1 Genre *Psychotria***

Des anticorps dirigés contre une protéine DING de tabac (*Nicotiana tabacum*) nous ont très généreusement été donnés par le Dr François Bernier, DR CNRS, (CNRS, Strasbourg) (Perera et al. 2008). Lors des expériences de Western blot, ces anticorps ont été utilisés dilués au cinq millièmes.

### **6.1.2 *Amborella trichopoda***

Dans le cadre de cette thèse, un anticorps polyclonal dirigé contre la sous-unité B (22 kDa) des globulines 11S d'*Amborella* a été obtenu comme décrit par Job et al. (1997) dans le cas de la préparation d'un anticorps reconnaissant spécifiquement la sous-unité B des globulines 11S de graines de betterave à sucre. Brièvement, des extraits de la fraction globulines des graines matures sèches d'*Amborella* ont été résolus en gels 2D. Après coloration des spots au bleu de Coomassie, les spots contenant des protéines attribuables à la sous-unité B des globulines 11S, ainsi que déterminé par séquençage MS/MS, ont été prélevés des gels. Après avoir rassemblé plusieurs de ces spots, la sous-unité B des globulines 11S d'*Amborella* a été électroéluée des spots, les électroéluates étant concentrés puis injectés à des cobayes ainsi que décrit par Job et al. (1997). La spécificité de ces anticorps a été testée en Western blot.

## **6.2 Test ELISA**

Une série de dilutions des extraits protéiques a été préparée comme suit (Job et al. 1997). L'échantillon le plus concentré contenait l'équivalent d'un dixième de la quantité en globulines d'une graine, dilué dans 250  $\mu$ L de tampon PBS (phosphate buffered saline) (150 mL NaCl/10 mM Na<sub>2</sub>PO<sub>4</sub>). Cet extrait a ensuite subi 11 dilutions au demi successives dans du tampon PBS.

Cent  $\mu$ L de ces dilutions ont été transférés dans une plaque ELISA microtitre de 96 puits (F ELISA plates, Greiner), le dernier puits de chaque ligne (n°12) ne recevant que du PBS et servant de témoin. La plaque a ensuite été incubée 20 min à 35°C puis rincée quatre fois avec du PBS contenant 0,05% (v:v) de Tween 20 (Sigma). Cent  $\mu$ L de la solution contenant les anticorps primaire dilués au 1/7500<sup>ème</sup> dans du PBS ont alors été placés dans chaque puits. Les puits ont ensuite été rincés quatre fois avec du PBS-Tween 20 0,05% puis les plaques ont été incubées 20 min à 35°C. La solution contenant l'anticorps secondaire (peroxidase-labelled goat anti-guinea pig antibodies, Sigma) diluée au 1/10000 dans du PBS-Tween 0,05 % a été ajoutée dans chaque puits à raison de 100  $\mu$ L/puits. L'incubation s'est déroulée à 35°C pendant 20 min puis les puits ont été rincés quatre fois avec du PBS-Tween.

La révélation a été ensuite réalisée par l'ajout de substrats de la peroxydase à 0,05% [1.8 mM 2,2'-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) diammoniums salt (Aldrich), et 3 X 10,3% (p/p) H<sub>2</sub>O<sub>2</sub> (Sigma)] dans du tampon citrate-phosphate 0,1 M, pH 4,0 (Figure 14). Le développement de la couleur a été suivi à 405 nm avec un lecteur de plaques (EL340 Bio-Teck Instruments) commandé par un Macintosh IICx équipé du logiciel DeltaSoft (version 4.1, BioMetallics, Inc.). Le logiciel KaleidaGraph™ (Abelbeck Software) a été utilisé pour

interpréter les données. La quantité relative de globuline 11S par graine a été calculée par une analyse en régression linéaire à partir des valeurs d'absorbance obtenues pour chacun des puits à la longueur d'onde de 405 nm.



Figure 14 : photo d'une plaque de test ELISA après révélation de la sous-unité B des globulines 11S d'*Amborella*. Pour chaque ligne A, B, C, D, E, correspond à un extrait particulier, le puits N° 1 contient la plus grande quantité d'extrait. Les puits N° 2 à N° 11 sont des dilutions successives au demi. Les puits N° 12 ne reçoivent que du tampon et servent de témoins.

### **6.3 Western blot**

Les protéines des échantillons de graines ont été séparées par migration SDS-PAGE en gels mono- ou bidimensionnels comme décrit dans le paragraphe « 5.1.4 Electrophorèse SDS-PAGE ». Après migration, les protéines des gels ont ensuite été transférées sur une membrane de nitrocellulose par un transfert semi-sec grâce à un courant transmis par deux électrodes entre lesquelles se trouvent, un papier filtre, la membrane de nitrocellulose, le gel et un papier filtre (Towbin et al. 1979). Le tout a été humidifié avec du tampon de transfert. Celui-ci est constitué de 500 mL du tampon Tris-Taurine utilisé pour la migration en deuxième dimension des gels de 18 cm, 200 mL de méthanol et 200 mL d'eau MilliQ.

Une intensité de  $0,8 \text{ mA/cm}^2$  a ensuite été appliquée pendant une heure. La membrane a alors été rapidement colorée au rouge Ponceau qui permet de s'assurer de la présence des protéines sur la membrane (<http://www8.umoncton.ca/umcm->

gauthier\_didier/siitub/elecubuard.html). A cette fin, les membranes de nitrocellulose ont été recouvertes de solution de rouge Ponceau pendant une à deux min puis elle a été abondamment rincées à l'eau MilliQ jusqu'à disparition de la coloration.

La révélation des Western Blot a été réalisée à l'aide du Kit « BM Chemiluminescence Blotting Substrate Roche » selon les instructions du fabricant (Roche Applied Science - 11500708001). Les sites réactifs, non spécifiques, de la membrane de nitrocellulose ont été bloqués pendant une heure par imprégnation de la membrane dans une solution de « Blocking » (fournie par le Kit), diluée au 1/5 dans du TBS (Tris-buffered saline). Puis la membrane a été incubée pendant 30 min à 1 h avec l'anticorps primaire dilué dans la solution de « Blocking » au dixième. Après cinq lavages de 5 min dans du tampon TBS-Tween, la membrane a ensuite été incubée 30 min à 1 h avec l'anticorps secondaire (cochon d'Inde anti cobaye couplé à la peroxydase) dilué au 1/10,000 dans la solution de « Blocking » au dixième. Puis la membrane a été lavée cinq fois 5 min dans le tampon TBS-Tween.

La révélation a ensuite été réalisée en chambre noire en plaçant la membrane dans la solution de révélation additionnée du substrat de la peroxydase pendant 1 min. L'émission de la fluorescence a été enregistrée sur film photographique.

## **7. Etude transcriptomique**

L'étude des transcrits accumulés au cours de la formation des fruits et des graines de *P. gabriellae* a été réalisée dans le cadre d'un projet global de compréhension des mécanismes de tolérance au nickel à l'échelle de la plante. Cette analyse s'intègre dans un projet récemment financé par l'Agence Nationale pour la Recherche (ANR EvoMetoNicks).

L'extraction des transcrits a été réalisée sur un mélange de fruits comme indiqué dans le paragraphe « 1.2.2 Récoltes utilisées pour l'approche de transcriptomique ». L'extraction a été effectuée à Nouméa puis l'analyse a été réalisée dans le cadre du programme « One thousand plant transcriptomes » (<http://www.onekp.com>). Ce programme est une initiative issue d'un partenariat public-privé. Il a pour but de permettre un séquençage à grande échelle d'ARNm de 1 000 espèces de plantes différentes. Les sponsors majeurs en sont « Alberta's Department of Advanced Education and Technology (AET), Silicon Valley based Musea Ventures, Beijing Genomics Institute in Shenzhen, University of Alberta, and Alberta's Informatics Circle of Research Excellence (iCORE) ».



## **7.1 Protocole d'extraction**

Le protocole utilisé a été décrit pour l'extraction d'ARN de graines riches en amidon (Li & Trick 2005). Cette expérimentation doit être réalisée dans des conditions de propreté extrême, afin de limiter tout risque de contaminations de l'extrait par des RNases naturellement présentes dans notre environnement qui détruiraient les ARN extraits.

Les graines ont été broyées dans de l'azote liquide à l'aide d'un pilon et d'un mortier (environ 50 à 100 mg de matériel végétal). Le broyat a été disposé dans des tubes Eppendorf RNase-free avec 400 µL de tampon d'extraction I (100 mM Tris, pH 8.0, 150 mM LiCl, 50 mM EDTA, 1.5% SDS, 1.5% 2-mercaptoethanol). Le tout a été agité au vortex pendant 30 s. La suspension a été additionnée de 250 µL de phénol-chloroforme (1:1, pH 4,7) et mélangée par retournement plusieurs fois puis centrifugée à 13 000 g pendant 15 min à 4°C. Le surnageant (phase aqueuse) a été transféré dans un nouveau tube et mélangé avec 250 µL du tampon II (70% sulfate de guanidinium (p/ v), 0.75 M de citrate de sodium, 10% lauryl-sarcosine, 2 M d'acétate de sodium, pH 4.0). Le tout a été mélangé par retournement plusieurs fois puis incubé à température ambiante (25°C) pendant 10 min sous agitation. Deux cents µL de chloroforme-isoamyl alcool (24:1) ont ensuite été ajoutés et mélangés par retournement avant une centrifugation à 13 000 g pendant 15 min à 4°C. Le surnageant (phase aqueuse) a été transféré dans un autre tube et mélangé à 300 µL d'isopropanol et 250 µL de NaCl 1,2 M. L'ensemble a été laissé sur glace pendant 15 min puis mélangé par retournement plusieurs fois avant d'être centrifugé comme précédemment. Le surnageant a alors été éliminé et le culot d'ARN lavé avec de l'éthanol 70%. Le culot a ensuite été séché puis suspendu dans 50 µL d'eau « RNase free » avant d'être stocké à -20°C.

## **7.2 Séquençage**

Le séquençage a été réalisé par Beijing Genomics Institute in Shenzhen (<http://www.genomics.cn>). Les transcrits ont été séquencés par séquençage haut débit Illumina et les données seront mise en ligne, notamment sur GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>). Le but du programme est d'obtenir au minimum 1 Gb de données par espèce.

### **7.3 Analyse des données**

L'analyse des données sur *Amborella* a été faite au laboratoire du professeur Claude de Pamphilis en collaboration avec le Dr Joshua Der (Penn State University, USA).

Les séquences de basse qualité en fin des contigs ont été éliminées. Les paramètres utilisés étaient : cutoff (20) ; badfraction (0,1) ; lengthfraction (0,5) et minlength (0). L'assemblage des contigs a ensuite été réalisé en utilisant le logiciel Trinity (Grabherr et al. 2011). Après l'assemblage, les séquences ont été analysées avec ESTScan (<http://myhits.isb-sib.ch/cgi-bin/estscan>) afin de déterminer le potentiel de chacune à coder des protéines et pour éliminer les séquences présentant des erreurs d'assemblage dans les cassettes de lecture. Les séquences obtenues ont ensuite été sélectionnées en fonction de leur aptitude ou non à être identifiées par BLASTx sur Swissprot, TAIR (The Arabidopsis Information Resource) et trEMBL avec une e-value de  $10^{-5}$  et un minimum de 50 identifications (hits). Pour chaque séquence, l'identification avec le plus de force statistique a été conservée. La description du gène est associée au résultat donné par TAIR (<http://www.arabidopsis.org>).

## **8 Etude phylogénétique**

Une approche phylogénique a été réalisée sur la base des séquences peptidiques des globulines 11S d'*Amborella*. Les séquences d'*Amborella* et de 65 autres (issues de 55 espèces) ont été alignées avec MAFFT (Multiple alignment program for amino acid or nucleotide sequences ; séquences disponibles en annexe 3) puis un arbre phylogénétique a été construit avec MEGA5.1 en mode « maximum likelihood » avec un bootstrap à 100 pseudo-réplicats.

**Chapitre III.    Physiologie de la  
graine d'*Amborella trichopoda***



# 1. Introduction

*Amborella trichopoda* Baill. est la seule espèce connue de sa famille (*Amborellaceae*), elle sera donc citée « *Amborella* » tout le long du manuscrit. C'est un arbuste dioïque, sarmenteux haut de 2 à 6 m. L'espèce pousse à altitude moyenne (600 à 800 m) en sous-bois de forêt humide sur substrats schisteux. Les feuilles sont alternes, pétiolées, arrondies ou aiguës au sommet et la marge est ondulée voire dentée. Les fleurs, inodores, de couleur crème sont portées sur des grappes de 2 à 30 fleurs situées à l'aisselle des feuilles ou sur les tiges dénudées (Figure 15). L'espèce est répandue à travers la partie centrale de la Grande-Terre : du plateau de Dogny jusqu'au Mont Tonine à Touho (Figure 16).



Figure 15 : photographies d'un plant d'*Amborella* parmi la végétation naturelle du plateau de Dogny (A) ainsi que ses fruits matures (B et C).

Bien que sa position phylogénétique ait fait l'objet de discussions (Goremykin et al. 2003 ; Soltis & Soltis 2004), il est maintenant admis qu'*Amborella* se situe à la base de l'arbre phylogénétique des angiospermes (Mathews & Donoghue 1999 ; Parkinson et al. 1999 ; Qiu et al. 1999 ; Soltis et al. 1999 ; Savolainen et al. 2000) (Figure 16). Elle est ainsi aujourd'hui considérée comme la sœur de toutes les plantes à fleurs et serait apparue il y a environ 135 Ma.

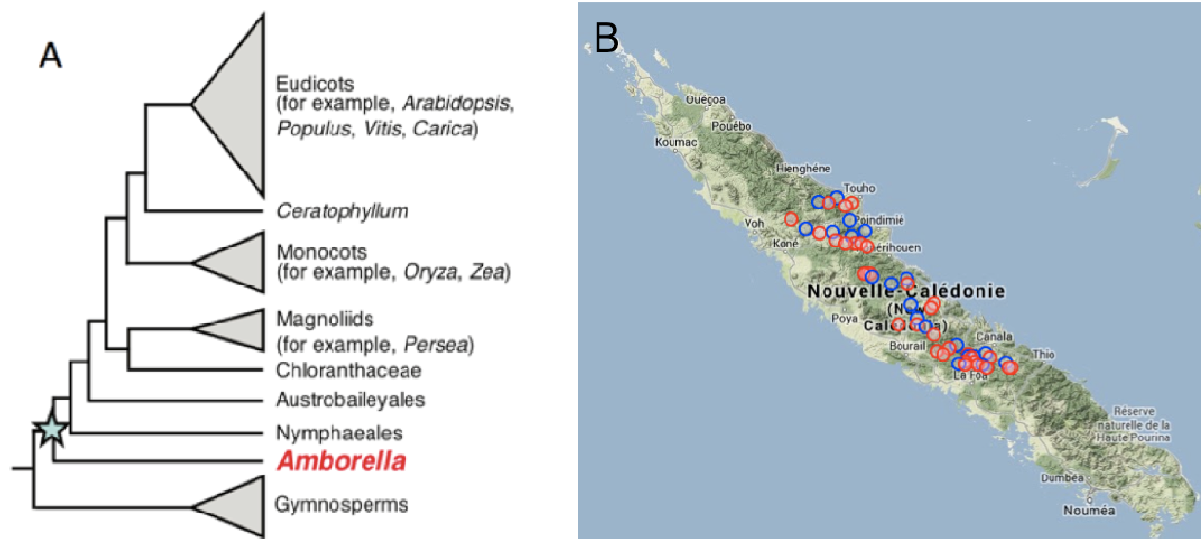


Figure 16 : arbre phylogénétique des plantes supérieures, d’après *Amborella* genome project. (2013) (A) et localisation des populations d’*Amborella* le long de la Grande-Terre (B) ; ronds rouges : données Herbar IRD Nouméa, ronds bleus : données Endémia ; carte extraite sur [www.endemia.nc](http://www.endemia.nc)).

Tenant compte de l’histoire géologique de la Grande-Terre, *Amborella* serait arrivée au plus tôt il y a 37 Ma en Nouvelle-Calédonie. Aucun fossile, ni pollen de cette espèce n’a encore été découvert ailleurs. Bien qu’*Amborella* ait pu poursuivre son évolution depuis son arrivée en Nouvelle-Calédonie, elle pourrait être considérée comme la meilleure fenêtre à disposition des scientifiques, de ce qui a été à la base des mono- et des dicotylédones.

De fait, de nombreuses études ont déjà été réalisées sur cette espèce afin de mettre en évidence, ou pas, des caractères plésiomorphes (primitifs), notamment au travers de l’étude de son anatomie végétative (Carlquist & Schneider 2001), de ses fruits (Bobrov et al. 2005) ou encore de la structure très particulière de son système vasculaire qui ne possède pas de vaisseaux xylémiens (Feild et al. 2000). De même, le séquençage récent du génome d’*Amborella* (*Amborella* Genome Project 2013) témoigne de l’importance de cette espèce pour la communauté scientifique.

Le mode de reproduction des angiospermes serait un de leurs avantages adaptatifs majeurs. C’est pourquoi cet aspect a été abordé avec beaucoup d’attention chez *Amborella*. Ainsi, l’étude du contrôle de la mise en place des carpelles (structures primordiales de la fleur femelle protégeant les ovules) a montré que les principaux gènes *CRABS CLAW* et *TOUSLED* contrôlant le développement des carpelles sont conservés au sein des angiospermes (Fourquin et al. 2005). Aussi, le développement embryonnaire présente de nombreux caractères

plésiomorphes (Tobe et al. 2000). C'est également le cas de l'albumen qui subit une première division cellulaire oblique, caractéristique des plantes basales (Floyd & Friedman 2001). Par ailleurs, l'étude de la structure du sac embryonnaire d'*Amborella* a permis de mettre en évidence une nouvelle organisation de ce dernier, qui contient huit cellules et neuf noyaux, différant ainsi du modèle classique (sept cellules, huit noyaux ; Friedman 2006). Cette découverte est en accord avec l'idée que les espèces basales reflètent le haut niveau d'expérimentation et de labilité liés à l'évolution, mais également un possible lien avec la structure du gamétophyte femelle des gymnospermes (Friedman 2006).

L'étude réalisée par Tobe et al. (2000), pour la première fois, sur des graines d'*Amborella* a permis la mise en évidence de la présence d'un embryon très petit, sous-développé, en forme de cœur dans la graine mature. Ces observations sont en accord avec l'idée qu'un embryon sous-développé est la forme la plus primitive chez les graines, bien que retrouvé chez d'autres espèces non basales (céleri, tomate,...) (Jacobsen & Pressman 1979 ; Hopher & Roberts 1985a ; Sansberro et al. 1998 ; Forbis et al. 2002 ; Vandeloos et al. 2007 ; Alves-Da-Silva et al. 2011 ; Chien et al. 2011). En effet, le ratio (E:S) entre la longueur de l'embryon (E) et celle de la graine (S) augmente au cours de l'évolution (Forbis et al. 2002). Les graines avec un embryon sous-développé, immature, présentent une dormance morphologique qui est considérée comme la dormance la plus primitive (Baskin & Baskin 2007). Celles-ci peuvent germer uniquement après poursuite du développement de l'embryon, suite à l'imbibition de la graine. La dispersion de graines matures possédant un petit embryon qui doit se développer avant de germer serait une stratégie ancienne qui permettrait d'étaler le déroulement de la germination dans le temps. Ce type de dormance, qui pose la question de la définition de la germination chez ces espèces, est rencontré dans toutes les familles du groupe des ANITA à l'exception des *Nymphaeales*, mais également chez certaines espèces plus modernes de mono- et eudicotylédones (Figure 17).

De nombreuses études se sont intéressées aux graines à dormance morpho- ou morphophysiological afin de mettre en évidence les moyens de lever de telles dormances (Steinbauer 1937 ; Hidayati et al. 2000 ; Baskin & Baskin 2001 ; Vandeloos et al. 2007 ; Alves-Da-Silva et al. 2011 ; Chien et al. 2011). En revanche, il existe peu d'études ayant caractérisé les mécanismes moléculaires qui contrôlent la levée de ces types de dormance (Hopher & Roberts 1985a ; Hopher & Roberts 1985b ; Sansberro et al. 1998 ; Homrichhausen et al. 2003).

Une étude menée au laboratoire LIVE (<http://live.univ-nc.nc>) en collaboration avec l'IAC (Institut Agronomique néo-Calédonien ; <http://www.iac.nc>) a confirmé que la semence

d'*Amborella* présente une dormance morphophysiological alors que la graine présente une dormance morphologique uniquement (Fogliani et al. 2013). Ces données (Figure 18 et Figure 19) confirment que :

- l'embryon sous-développé est en forme de cœur et entouré par un albumen volumineux,
- l'embryon se développe dans la graine avant germination *sensu stricto*,
- le choix du modèle d'étude est très important : la semence ou la graine.

Family	Primary region of geographical distribution <sup>b</sup>	Family	Primary region of geographical distribution <sup>b</sup>
Amaryllidaceae	Tropical or subtropical	Hydrophyllaceae	Cosmopolitan, except Australia
<b>Amborellaceae</b>	<b>Tropical</b>	Illiciaceae	Tropical
Annonaceae	Tropical (especially Old World)	Iridaceae	Tropical and temperate
Apiaceae	Northern temperate	Lactoridaceae	Tropical
Aquifoliaceae	Tropical and temperate	Lardizabalaceae	Temperate
Araceae	Tropical and temperate	Loranthaceae	Tropical and temperate
Araliaceae	Tropical	Liliaceae	Warm temperate and tropical
Arecaceae	Tropical and subtropical	Magnoliaceae	Temperate and tropical
Aristolochiaceae	Tropical and warm temperate	Melanthiaceae	Warm temperate and tropical
Berberidaceae	Northern temperate and tropical mountains	Menyanthaceae	Temperate and boreal
Buxaceae	Tropical and temperate	Monimiaceae	Southern tropical
Canellaceae	Tropical	Myristicaceae	Tropical
Cannaceae	Tropical	Nandinaceae	Northern temperate
Caprifoliaceae	Northern temperate and tropical mountains	Oleaceae	Temperate and tropical
Chloranthaceae	Tropical and subtropical	Paeoniaceae	Northern temperate
Convallariaceae	Temperate	Papaveraceae	Northern temperate
Cycadaceae	Tropical	Piperaceae	Tropical
Daphniphyllaceae	Tropical	Pittosporaceae	Tropical
Degeneriaceae	Tropical	Podocarpaceae	Southern temperate
Dilleniaceae	Tropical and subtropical	Ranunculaceae	Northern temperate
Escalloniaceae	Southern temperate	Santalaceae	Tropical and temperate
Eupomatiaceae	Tropical	Sarraceniaceae	Temperate and tropical
Fumariaceae	Northern temperate	Schisandraceae	Temperate and tropical
Garryaceae	Warm temperate	Smilacaceae	Tropical and temperate
Ginkgoaceae	Temperate	Taxaceae	Temperate and tropical
Grossulariaceae	Temperate	Trochodendraceae	Northern temperate
Haemodoraceae	Temperate and tropical	Winteraceae	Tropical

Figure 17 : liste des familles pour lesquelles au moins une espèce possède un embryon rudimentaire ou linéaire permettant de définir une dormance morphologique (Baskin & Baskin 2001).



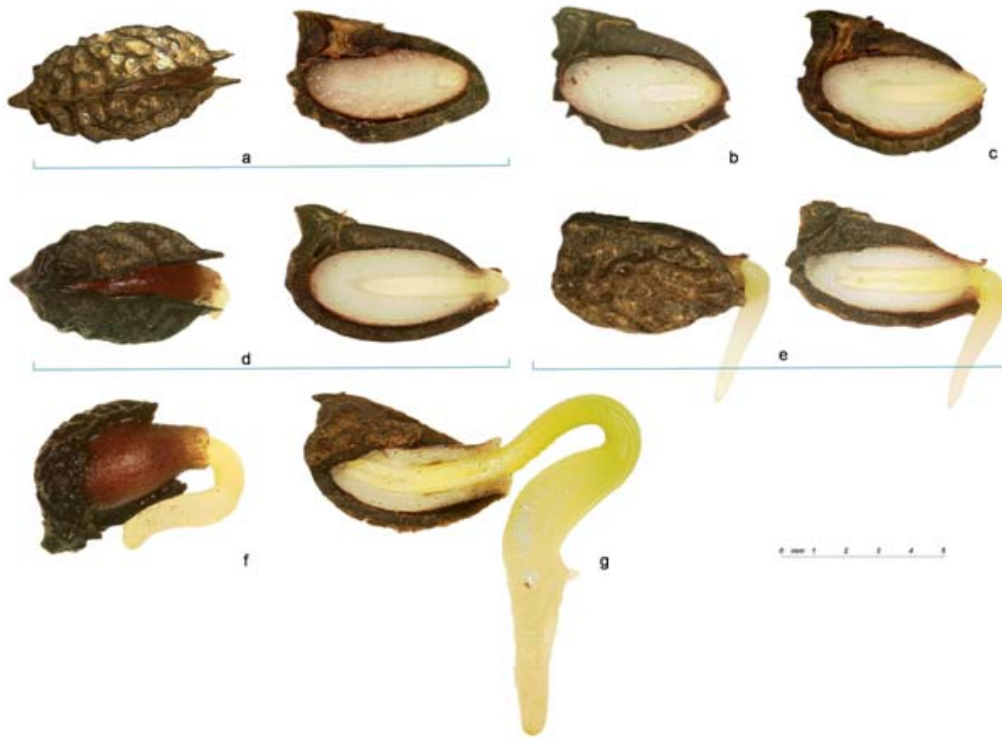


Figure 18 : reconstruction de la séquence chronologique de la germination de graines d'*Amborella* (Fogliani et al. 2013).

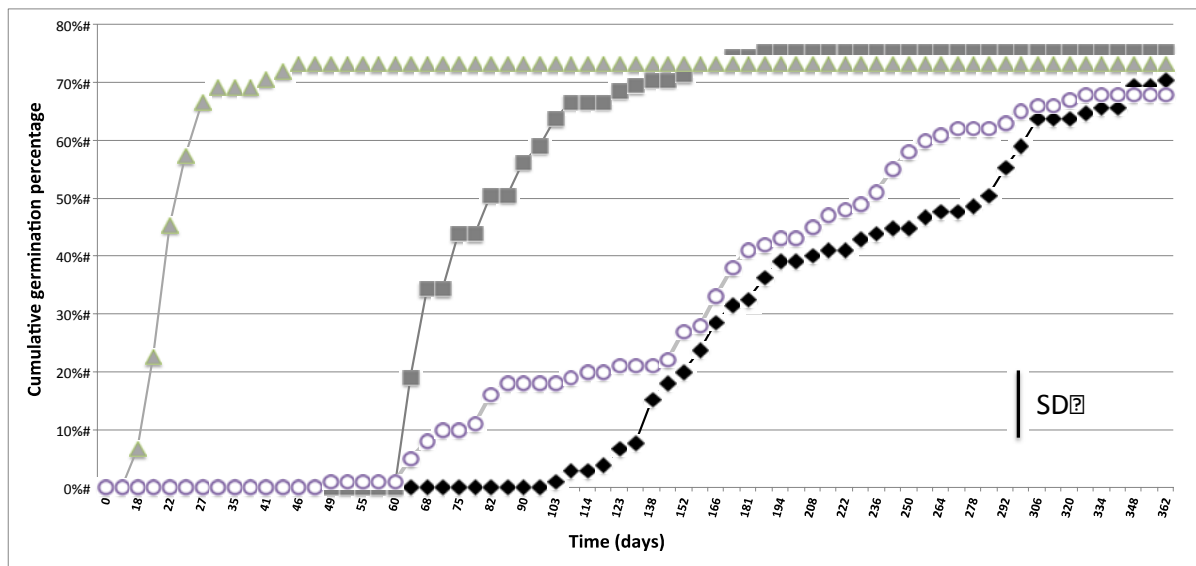


Figure 19 : taux de germination cumulé en fonction du temps des semences (○) ou des grains (▲) d'*Amborella* mises à germer en boîtes de Petri à  $25^{\circ}\text{C} \pm 1^{\circ}\text{C}$ , lumière continue. Les semences ont été scarifiées (■) ou pas (◆) à l'acide et semées en serre ( $25 \pm 4^{\circ}\text{C}$ , haute humidité). La barre verticale (SD) indique l'écart-type le plus important (Fogliani et al. 2013).

## 2. Analyse des graines matures sèches d'*Amborella*

Les résultats obtenus précédemment indiquent qu'au moment de sa dispersion (séparation du pied mère) la teneur en eau de la graine est de 8,6% (Fogliani et al. 2013). Ainsi, le terme de « graine mature sèche » sera utilisé pour décrire la graine au moment de sa dispersion (état mature) et de type orthodoxe (teneur en eau de 8,6%).

### 2.1 Résultats

Une l'analyse des classes protéiques décrites par Osborne (1924) montre que les deux principales classes de protéines de la graine mature sèche d'*Amborella* sont les albumines et les globulines, qui représentent respectivement 4,4% et 13,9% de la matière fraîche (MF). Les prolamines et les glutélines (Osborne 1924) ne sont quant à elles présentes respectivement qu'à hauteur de 0,02% et 0,36% de la MF (Tableau 5).

Tableau 5 : composition des différentes classes de protéines de la graine mature sèche d'*Amborella*. Les résultats sont exprimés en pourcentage de la matière fraîche (trois réplicats).

<b>Fractions protéiques des graines matures sèches d'<i>Amborella</i> selon Osborne (1924)</b>			
Albumines	Globulines	Prolamines	Glutélines
4,35 (±0,26)	13,91 (±0,2)	0,02 (±0,01)	0,36 (±0,09)

#### 2.1.1 Observations cytologiques

Les graines matures sèches d'*Amborella* sont composées d'un albumen volumineux, d'une enveloppe et d'un embryon rudimentaire de petite taille, excentré et en forme de cœur (ratio longueur embryon/graine = 0,08). Des coupes de graines ont été colorées au bleu noir naphtol, afin de visualiser les protéines (bleu) et les noyaux (bleu foncé). Cette coloration contenait également de l'acide périodique/réactif de Schiff afin de mettre en évidence les polysaccharides (rose) (Figure 20).

Les cellules embryonnaires présentent un fond bleu clair. Elles contiennent un noyau (flèche Figure 5B) et sont délimitées par une très fine paroi (coloration rose). On ne distingue pas de corps protéiques, le cytosol étant uniformément coloré en bleu pâle. L'ensemble de

l'embryon est entouré d'un épais réseau de polysaccharides (coloré en rose) dans lequel sont imbriquées des cellules mortes (Figure 20). Les cellules de l'albumen ont des parois plus visibles que celles de l'embryon et sont plus grosses (environ 70  $\mu\text{m}$  de long ; maximum 50  $\mu\text{m}$  pour l'embryon) (Figure 20). Elles contiennent un noyau et de nombreux corps protéiques (flèche, Figure 20A).

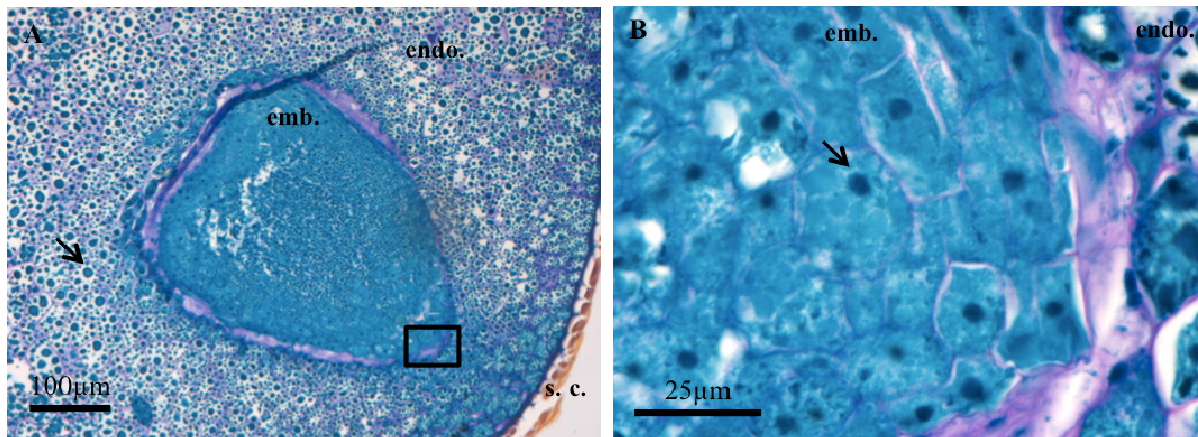


Figure 20 : observations au microscope optique de coupes de graines matures sèches d'*Amborella* après coloration au bleu noir naphthol combinée à la coloration à l'acide périodique/Schiff. (A) embryon (emb.), albumen (endo.), enveloppe (s.c.). (B) Agrandissement de la zone du rectangle noir en (A). Les protéines et corps protéiques apparaissent en bleu (flèche A), les nucléoles en bleu foncé (flèche B) et les polysaccharides en rose.

Ces observations indiquent une différence histologique marquée entre l'albumen et l'embryon tous deux étant séparés par un réseau de polysaccharides et de cellules mortes. Elles posent diverses questions quant à l'état de maturité de l'embryon et de l'albumen ainsi que la nature des protéines présentes dans la graine mature sèche et ses deux constituants principaux, l'albumen et l'embryon. Ces questions ont été abordées par approche protéomique.

### 2.1.2 Etude du protéome de la graine mature sèche d'*Amborella*

Deux types d'approches ont été utilisés. La première vise à caractériser les protéines totales de l'embryon et de l'albumen isolés de graines matures sèches d'*Amborella* grâce à la technique de protéomique en vrac (*shotgun*) dont le principe est décrit dans les chapitres

« Introduction Générale » et « Matériels et Méthodes ». L'autre a pour but d'étudier séparément le protéome des globulines et des albumines extraites de graines matures sèches d'*Amborella* en mettant en œuvre une approche basée sur l'électrophorèse 2D des protéines (cf. les chapitres « Introduction Générale » et « Matériels et Méthodes »).

### ***2.1.2.1 Etude des protéines totales de la graine mature sèche par protéomique shotgun***

Cette analyse a été réalisée à partir de i) 300 embryons isolés de graines matures sèches et ii) d'albumen dépourvu d'embryons (1,2 g) également isolé de graines matures sèches d'*Amborella* (Figure 18 et Figure 20) (cf. « Matériels et Méthodes »). Une approche de protéomique *shotgun* a été privilégiée car cette technique sensible est particulièrement adaptée pour des échantillons disponibles en très petite quantité, comme c'est le cas pour les embryons d'*Amborella* (Figure 18 et Figure 20).

L'extraction des protéines totales a permis d'identifier 415 protéines de l'embryon isolé (tableau supplémentaire 1) et 69 protéines de l'albumen débarrassé d'embryons (tableau supplémentaire 2). Les protéines identifiées ont ensuite été regroupées en fonction de leur classe ontologique et de leur description (Bevan et al. 1998).

La représentation de chaque catégorie et de chaque fonction a été exprimée en pourcentage du nombre total de protéines identifiées, ainsi que par rapport à la quantité relative de chaque identification (abondance). Cette estimation a été réalisée à partir du nombre de peptides correspondant à chaque identification (Figure 21).

Les 69 protéines de l'albumen correspondent à 22 fonctions uniques (Bevan et al. 1998). Elles se regroupent en quatre catégories majeures (Figure 21) au regard de leur abondance relative (tableau supplémentaire 2).

La catégorie « ***destination and storage*** » contient 19 protéines (27,5% de l'ensemble des protéines) qui représentent 85,1% de l'abondance de l'ensemble des protéines. Huit protéines sont de la fonction « ***storage protein*** » (11,6% de l'ensemble des protéines) et représentent à elles seules 82,6% en abondance. La fonction « ***folding and stability*** » contient sept protéines (10,1% de l'ensemble des protéines) qui représentent 1,8% en abondance, et la fonction « ***proteolysis*** » contient quatre protéines (5,8% de l'ensemble des protéines) qui représentent 0,7% en abondance.

Quinze protéines (21,7% de l'ensemble des protéines) sont de la catégorie « *disease/defence* » et représentent 5,1% en abondance. Les protéines de la fonction « *stress response* », au nombre de cinq (7,2% de l'ensemble des protéines), représentent 2,3% en abondance. Les protéines de la fonction « *detoxification* » sont au nombre de six (8,7% de l'ensemble des protéines) et représentent 2,2% en abondance.

Les 17 protéines (24,6%) de la catégorie « *energy* » représentent 3,6% en abondance. Elles sont principalement impliquées dans trois fonctions, la glycolyse, le cycle des acides tricarboxyliques (*TCA pathway*) et la fermentation et représentent 0,8% d'abondance relative à elles trois.

Les protéines impliquées dans la catégorie « *metabolism* » sont au nombre de quatre (5,8% de l'ensemble des protéines) et représentent 0,9% d'abondance.

La catégorie « *cell structure* » contient huit protéines (11,6% de l'ensemble des protéines) qui représentent 1,9% d'abondance.

Ces résultats illustrent clairement que l'albumen, constituant majoritaire de la graine entière, est principalement un tissu de stockage des réserves séminales. Ainsi, la quantité relative de protéines de réserve y est de 82,6%.

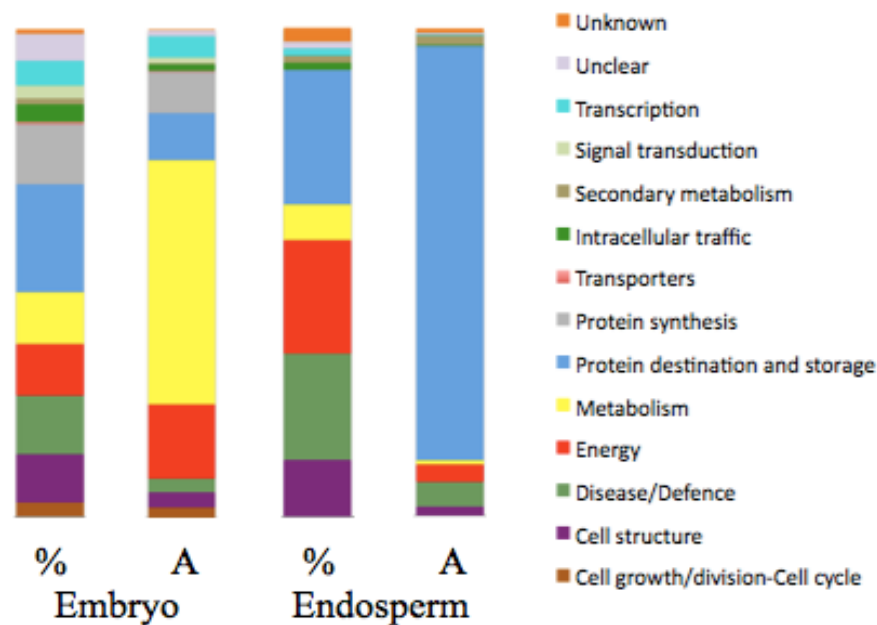


Figure 21 : représentation de l'importance des classes ontologiques (Bevan et al. 1998) des protéines totales extraites d'embryon (Embryo) et de l'albumen (Endosperm) de graines *stricto sensu* d'*Amborella*. Les résultats sont exprimés en fonction du nombre de protéines identifiées par classe par rapport au nombre total de protéines identifiées (%) ou en fonction de leurs quantités relatives parmi l'ensemble des protéines identifiées (A).

La classification par ontologies (Bevan et al. 1998) des 415 protéines identifiées dans l'embryon par approche *shotgun* est radicalement différente (Figure 21). Ces 415 protéines correspondent à 52 fonctions uniques. Ainsi, les protéines de la catégorie « **destination and storage** » qui sont au nombre de 91 (21,9% de l'ensemble des protéines totales identifiées) ne représentent que 9,5% en abondance (soit près de dix fois moins que dans l'albumen).

La classification ontologique des protéines de l'embryon a par ailleurs permis d'identifier plusieurs catégories majeures (Figure 21) (tableau supplémentaire 1).

La catégorie « **metabolism** » contient 43 protéines (10,4% de l'ensemble des protéines) et représentent 49,6% en abondance. Dix-sept protéines (4,1% de l'ensemble des protéines) sont impliquées dans la fonction « **amino acid** » et représentent à elles seules 37,1% d'abondance. Quatorze protéines (3,4% de l'ensemble des protéines) sont impliquées dans la fonction « **sugars and polysaccharides** » et représentent 7,3% des protéines en abondance. Six protéines (1,4% de l'ensemble des protéines) interviennent dans la fonction « **lipid and sterol** » et représentent 1,8% en abondance, quatre protéines sont impliquées dans la fonction « **nitrogen and sulphur** » et représentent 2,8% des protéines en abondance, deux protéines sont dans la fonction « **cofactors** » et représentent 0,6% en abondance.

La catégorie fonctionnelle « **energy** » contient 44 protéines (10,6% de l'ensemble des protéines) et représente 14,8% en abondance. Notamment, la fonction « **glycolysis** » contient 17 protéines (4,1% de l'ensemble des protéines) représentant 6,3% en abondance. La fonction « **TCA pathway** » est représentée par 14 protéines (3,4% de l'ensemble des protéines) qui constituent 4,5% des protéines en abondance. Les fonctions « **pentose phosphate** » et « **fermentation** » sont représentées chacune par quatre protéines (1% de l'ensemble des protéines) qui correspondent une abondance de 1,4% et 1,2% respectivement.

Les protéines de la catégorie « **protein synthesis** » sont au nombre de 51 (12,3% de l'ensemble des protéines) et représentent 8,4% des protéines de l'embryon en terme d'abondance relative. Quarante-deux d'entre elles (10,1% de l'ensemble des protéines) sont de la fonction « **ribosomal proteins** » et représentent 7,1% d'abondance relative, et neuf protéines (2,2% de l'ensemble des protéines) sont dans la catégorie « **translation factors** » et représentent 1,3% d'abondance relative.

Ainsi, les protéines de l'embryon sont majoritairement associées aux mécanismes cellulaires plutôt qu'à un rôle de stockage, contrairement à l'albumen (Figure 21).

Une représentation en diagramme de Venn (Figure 22) a permis de mettre en évidence les protéines spécifiques et communes à l'albumen et l'embryon. Ainsi, douze protéines sont

spécifiques de l'albumen, 358 sont spécifiques de l'embryon et 57 sont communes aux deux tissus.

Les protéines spécifiques de l'embryon sont principalement de la catégorie « *metabolism* » qui représente 33,3% des protéines (en abondance). Notamment, 24,8% sont de la fonction « *amino acid* », 4,9% de la fonction « *sugars and polysaccharides* », 2,0% de la fonction « *nitrogen and sulfur* », 1,2% de la fonction « *lipid and stérol* » et 0,4% de la fonction « *cofactors* ». La catégorie « *energy* » représente 7,0% des protéines spécifiques de l'embryon. La fonction « *glycolysis* » représente 1,6%, la fonction « *pentose phosphate* » 1,0% et la fonction « *TCA pathway* » 2,7%. Il apparaît que les protéines de la catégorie « *cell growth/division – cell cycle* » et « *protein synthesis* » sont spécifiques de l'embryon. La catégorie « *cell growth/division – cell cycle* » contient 12 protéines (2,8%) et représente 2% d'abondance. Quatre protéines (0,9%) sont de la fonction « *cell cycle* » qui représente 0,7% d'abondance. Trois protéines (0,7%) sont de la fonction « *DNA synthesis/replication* » et représentent 0,5% d'abondance. La catégorie « *protein synthesis* » contient 51 protéines (11,9%) représentant 6,0% d'abondance. Quarante deux d'entre elles (9,8%) sont de la fonction « *ribosomal proteins* » et représentent 5,0% d'abondance. Les neuf autres protéines (2,1%) sont de la fonction « *translation factors* » et représentent 0,9% d'abondance.

Sur les douze protéines spécifiques de l'albumen, cinq sont de la catégorie « *cell structure* » et représentent 0,32% des protéines (abondance). Trois d'entre elles sont de la fonction « *cytoskeleton* », les fonctions « *cell wall* » et « *ER/Golgi* » contiennent une protéine chacune. Les fonctions « *lipid and sterol* », « *mRNA synthesis* », « *folding and stability* », « *defence related* » et « *stress response* » contiennent chacune une protéine. La fonction « *TCA pathway* » contient elle deux protéines.

La grande majorité des protéines communes est de la catégorie « *protein destination and storage* » et vient du fait que les protéines de réserves ont été identifiées dans les deux compartiments, bien qu'en abondances très différentes (Figure 21 et Figure 22).

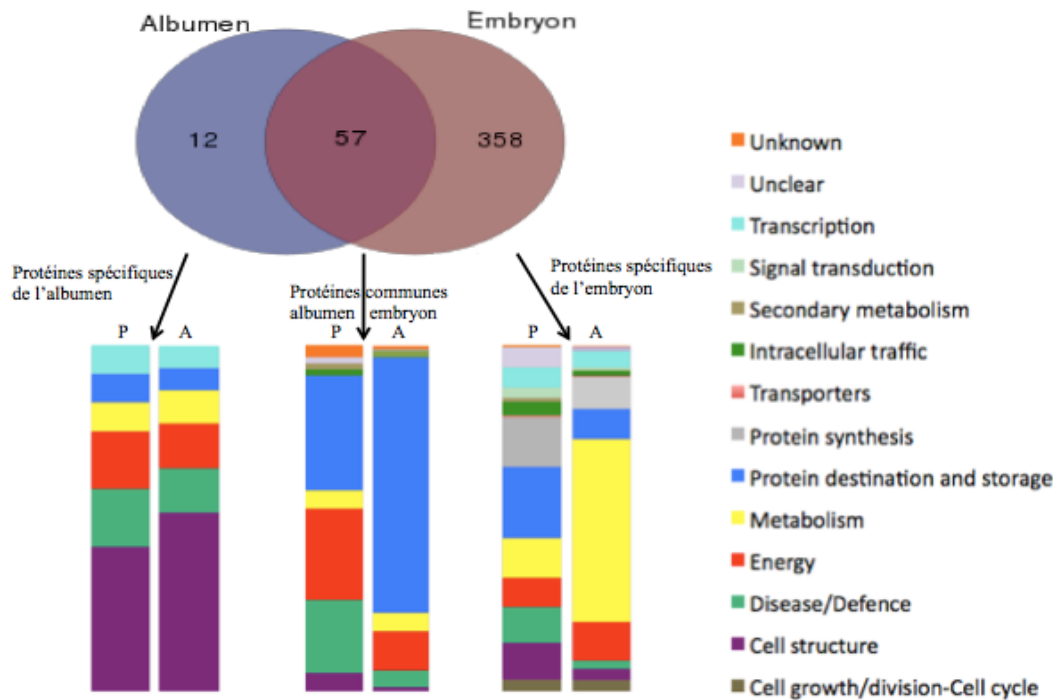


Figure 22 : diagramme de Venn et signatures ontologiques (Bevan et al. 1998) des protéines totales de l'embryon et de l'albumen de graines *stricto sensu* d'*Amborella* identifiées par protéomique *shotgun*. "P" : répartition des classes ontologiques en fonction du nombre total de protéines identifiées. "A" abondance relative des classes ontologiques identifiées.

### 2.1.2.2 Analyse des fractions albumines et globulines de la graine mature sèche d'*Amborella* par électrophorèse bidimensionnelle

L'analyse par protéomique *shotgun* a montré que les protéines totales de l'albumen sont principalement dominées par les protéines de réserve (82,6% de l'ensemble des protéines identifiées (Figure 21), en accord avec l'observation de la présence de nombreux corps protéiques dans les cellules de l'albumen, contrairement à l'embryon (Figure 20). Par contre, le fait que l'embryon contienne beaucoup moins de protéines de réserve (Figure 20 ; Figure 21) a permis l'identification d'un plus grand nombre de protéines (415) que dans l'albumen (Figure 21). Afin d'analyser plus en détail les fractions albumines et globulines de la graine mature sèche, ces deux fractions ont été séparées selon le protocole d'Osborne (1924) et soumises séparément à une analyse protéomique en 2D-PAGE.



### 2.1.2.2.1 Albumines

L'extrait de la fraction albumines des protéines de la graine mature sèche d'*Amborella* a été soumis à une migration en 2-DE (Figure 23). Après coloration au nitrate d'argent, le profil protéique révèle 212 spots majoritaires qui ont été prélevés des gels 2D et analysés par spectrométrie de masse en tandem. Les cartes de référence ont été établies (annexe 1). Parmi ces 212 spots, 50 contenaient une seule protéine, 71 spots contenaient deux protéines, 42 spots contenaient trois protéines différentes et 49 spots quatre ou plus (tableau supplémentaire 3). Au total, 561 protéines ont ainsi été identifiées (tableau supplémentaire 3). Elles correspondent à 123 gènes selon l'analyse du génome d'*Amborella* récemment établi (*Amborella* Genome Project 2013 ; <http://amborella.org>). La méthode utilisée pour l'extraction protéique a donc bien permis d'enrichir la fraction albumines.

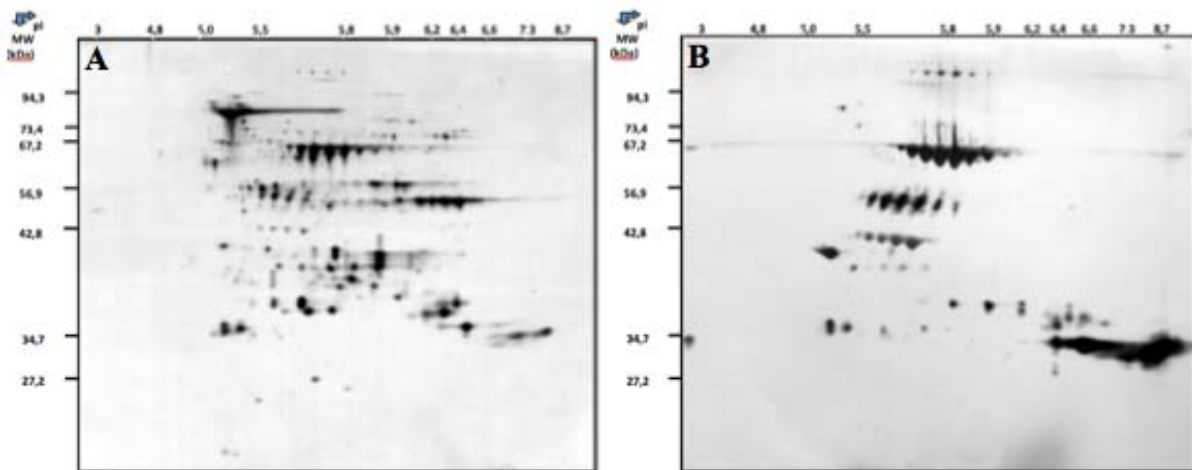


Figure 23 : gels bidimensionnels colorés à l'argent des fractions albumines (A) et globulines (B) des graines matures sèches d'*Amborella*.

Les protéines de la fraction albumines identifiées ont été classées en fonction de leurs catégories et fonctions ontologiques (Bevan et al. 1998). La proportion de chaque classe (ou fonction) a été exprimée en pourcentage du nombre de protéines par classe par rapport au nombre total de protéines (Figure 24).

La catégorie la plus importante est celle des « *protein destination and storage* » qui contient 35,9% des protéines identifiées, 3,9% d'entre elles appartenant à la fonction « *proteolysis* », 8,4% d'entre elles à la catégorie « *folding and stability* » et 23,7% à la catégorie « *storage proteins* ». Les protéines de la catégorie « *energy* » représentent 23,8% des protéines identifiées. La fonction « *glycolysis* » contient 17,8% des protéines identifiées, la fonction « *TCA pathway* » 3,2% des protéines identifiées, la fonction « *fermentation* » 1,6% des protéines identifiées et la fonction « *pentose phosphate* » 0,5% des protéines

identifiées. La catégorie « *disease/defence* » contient 13,4% des protéines identifiées, avec notamment les fonctions « *stress response* », « *detoxification* » et « *defence-related* » qui contiennent respectivement 7,7%, 5,2% et 0,5% de l'ensemble des protéines identifiées. La catégorie « *metabolism* » contient 11,7% des protéines identifiées, se répartissant dans les fonctions suivantes : 3,9% pour la fonction « *amino acid* », 3,7% pour la fonction « *sugars and polysaccharides* », 2,3% pour la fonction « *lipid and sterol* », 1,3% pour la fonction « *nucleotides* », 0,7% pour la fonction « *cofactors* » et 0,2% pour la fonction « *nitrogen and sulphur* ». La catégorie « *transcription* » contient 5,5% des protéines identifiées, avec notamment 0,4% des protéines dans la fonction « *chromatin modification* » et 5,1% dans la fonction « *mRNA processing* ».

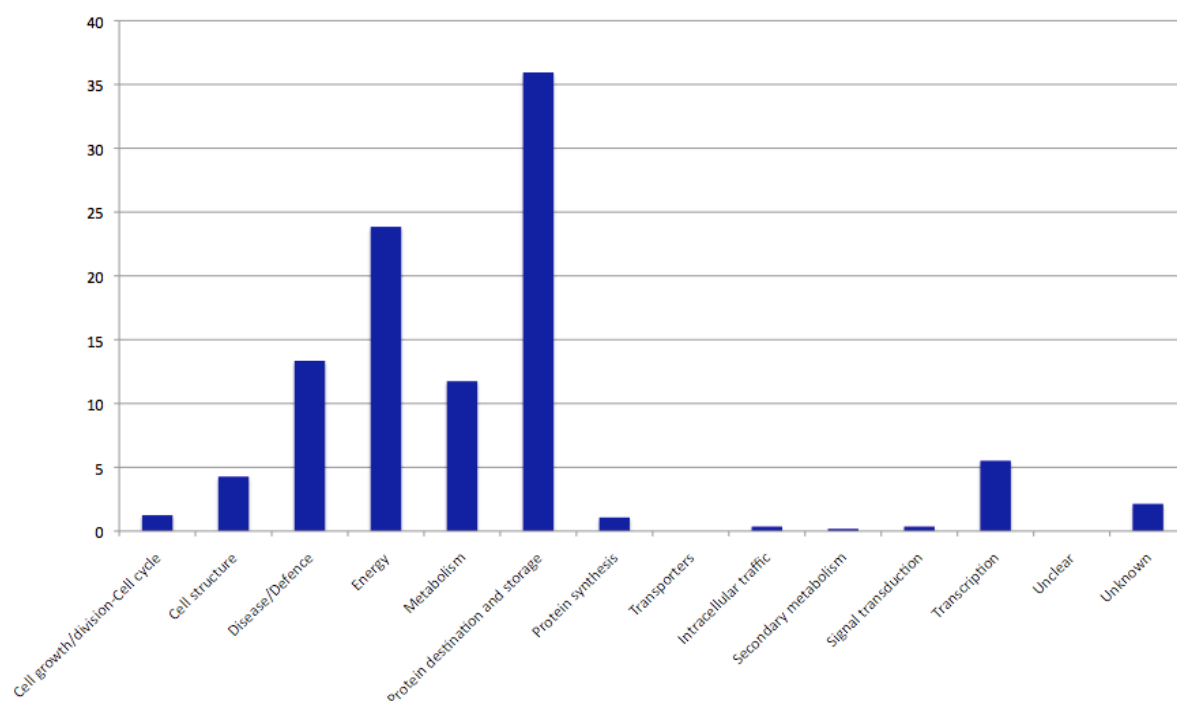


Figure 24 : représentation des classes ontologiques (Bevan et al. 1998) de la fraction albumines des graines matures sèches d'*Amborella*.

#### 2.1.2.2.2 Globulines

De la même manière, l'extrait de la fraction globulines des protéines de la graine mature sèche d'*Amborella* a été soumis à une migration par électrophorèse en gels 2-DE (Figure 23) puis les 133 spots majoritaires ont été prélevés des gels 2D et analysés par spectrométrie de masse. Les cartes de référence ont été établies (annexe 2). Parmi ces 133 spots, 26 contenaient une seule protéine, 49 spots contenaient deux protéines différentes, 39 spots contenaient trois protéines différentes et 19 spots quatre ou plus. Au total, 305 protéines ont été identifiées

(tableau supplémentaire 4) correspondant à seulement 20 gènes uniques. Parmi les 309 protéines identifiées, 99 correspondent aux globulines 11S, formes majoritaires des protéines de réserve chez de nombreuses graines (Shewry et al. 1995 ; Tandang-Silvas et al. 2010), qui ont fait l'objet dans cette thèse d'une analyse moléculaire, biochimique et phylogénétique.

A l'époque où de plus en plus de génomes sont séquencés, il persiste des difficultés dans leur annotation. En effet, l'identification des régions codantes est réalisée par comparaison avec des séquences similaires ou par une approche de prédiction non comparative (*ab initio*). Ces prédictions sont très souvent couplées à une comparaison avec des données de transcriptomique afin de valider la région codante et la séquence protéique correspondante. Cependant, cette méthode ne permet pas d'assurer que la région codante prédite est effectivement traduite. Avec l'avènement des possibilités de séquençage protéique à haut débit, de plus en plus d'études combinent : i) une prédiction par comparaison et *ab initio*, ii) une confirmation par utilisation des données de transcriptomique, et iii) une confirmation par utilisation des données de protéomique (Ansong et al. 2008). Ce type d'approche a ainsi permis de découvrir 778 nouveaux gènes codants chez *A. thaliana* et d'affiner l'annotation de 695 autres (Castellana et al. 2008).

Les protéines sont une des formes de réserve les plus présentes des graines, le groupe le plus répandu de ces protéines de réserve correspondant aux globulines. Elles appartiennent généralement à la famille des protéines à motif *cupin* (tonneau ; [http://en.wikipedia.org/wiki/Cupin\\_superfamily](http://en.wikipedia.org/wiki/Cupin_superfamily)) et ont évolué à partir d'enzymes bactériennes (Dunwell et al. 2001). Ces protéines forment une superfamille caractérisée par un feuillet  $\beta$  conservé. Leurs rôles sont divers : isomérase ou épimérase de modification des parois bactériennes, facteur de transcription multi domaine ou encore protéines de réserve des graines (Dunwell et al. 2004). Les globulines peuvent être classées en deux sous-groupes selon leur coefficient de sédimentation, les globulines de type viciline 7-8S et les globulines de type légumine 11-12S. Ces protéines sont présentes chez les mono- et les dicotylédones. Les globulines de type 11S sont codées par une même famille de gènes chez les mono- et les dicotylédones. En revanche, les gènes codant les globulines de type 7S ne sont pas communs à ces deux classes de plantes (Shewry & Halford 2002). Les protéines de réserve de type globulines 11-12S (légumine) sont les plus abondantes dans les graines matures (Bewley & Black 1994). Elles sont codées par trois gènes chez *A. thaliana* (At1g03880, At4g28520, At5g44120) et s'accumulent pendant les phases tardives de la maturation (Gruis et al. 2002 ; Q. Li et al. 2007). L'expression de ces trois gènes aboutit à la biosynthèse d'un polypeptide précurseur unique, la préprotéine contenant un peptide

signal qui est ensuite clivé (Figure 25). Les pro-globulines (50-60 kDa) ainsi formées s'assemblent en trimères et sont transportées dans la vacuole de stockage. Une endopeptidase agit ensuite au niveau d'un site consensus très conservé pour donner les sous-unités acides ( $\alpha$  ou A), d'environ 35 kDa, et les sous-unités basiques ( $\beta$  ou B), d'environ 25 kDa. Ces deux sous-unités restent liées par un pont disulfure grâce à la présence de résidus cystéines très conservés (Scott et al. 1992 ; Adachi et al. 2003 ; Tandang-Silvas et al. 2010 ; Li et al. 2012) (Figure 25). Cette coupure permet l'assemblage et le stockage sous forme d'hexamères alors que la forme non clivée ne forme que des trimères. Il s'agit ainsi d'un mécanisme d'évolution qui a permis d'assurer un compactage maximum des globulines 11S dans un minimum d'espace (Shutov et al. 2003). Les six paires de sous-unités sont associées de manière non covalente. Les sous-unités A se trouvent à l'extérieur de l'hexamère et les sous-unités B à l'intérieur (Adachi et al. 2003 ; Maltman et al. 2007) (Figure 25).

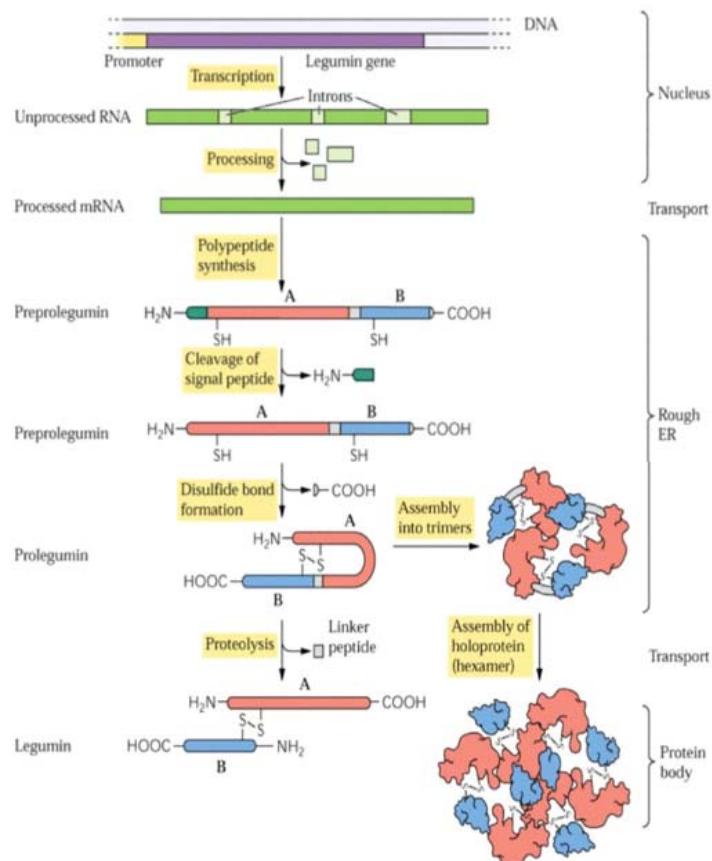


Figure 25 : les étapes de la biosynthèse des globulines de type 11S dans les cotylédons de soja (Catusse 2007). L'ARNm, une fois mature, sort du noyau pour aller dans le réticulum endoplasmique granulaire (Rough ER) où il est traduit. La protéine synthétisée est adressée au lumen du RE après clivage du peptide signal. La protéine est constituée des sous-unités A et B encore liées par une séquence conservée et reconnue par une protéase. Les protéines s'associent en trimère dans le RE puis entrent dans les corps protéiques. C'est alors que les sous-unités A et B sont séparées par protéolyse demeurant ainsi liées uniquement par le pont disulfure. Les formes matures de globulines 11S peuvent ainsi se compacter plus encore en formant des hexamères.

Les mécanismes de synthèse, d'accumulation, de stockage puis d'utilisation des globulines représentent une innovation extraordinaire. En effet, la graine doit différencier temporellement mais non spatialement la synthèse, le stockage puis l'utilisation de ces protéines (Figure 26). De plus, il existe des protéases synthétisées pendant la maturation de la graine et qui sont stockées dans les corps protéiques eux-mêmes. L'ajustement du pH dans les organites de réserve serait un des mécanismes permettant de contrôler l'enclenchement de leur activité (Müntz et al. 2001 ; Shutov et al. 2003 ; Li et al. 2007 ; Wang et al. 2007).

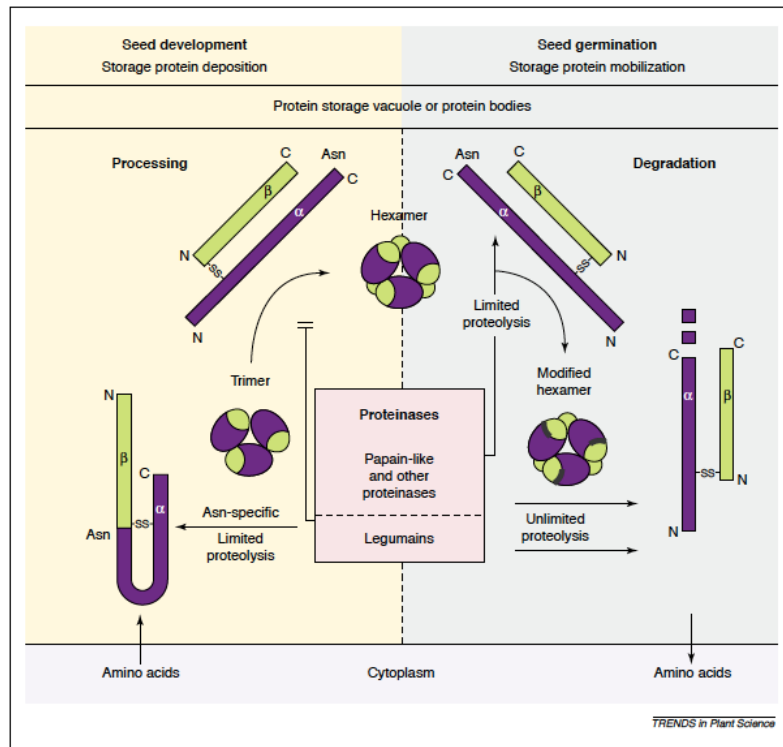


Figure 26 : utilisation des protéines de réserve de type légumineuses (globulines 11S) par les légumineuses (Müntz & Shutov 2002) au cours du développement de la graine puis lors de la germination et du développement de la plantule. Les modifications structurales des sous-unités A et B sont représentées en mauve et vert respectivement et associées aux modifications structurales des oligomères. L'asparagine (Asn) indique le site de coupure du précurseur protéique. « SS » indique le pont disulfure liant les sous-unités. Le cigne « = » indique qu'il n'y a pas de protéolyse des globulines 11S.

L'existence de propriétés structurales conservées des globulines 11S est largement documentée (Scott et al. 1992 ; Adachi et al. 2003 ; Tandang-Silvas et al. 2010 ; Li et al. 2012). Nous avons ainsi entrepris d'isoler dans le génome d'*Amborella* les gènes codant les globulines de type 11S à partir de nos identifications protéiques dans le but de i) valider les prédictions génomiques et ii) mettre en évidence des particularités structurales des globulines 11S. Ces résultats ont fait l'objet d'une contribution à la publication du génome d'*Amborella* récemment soumis dans *Science* (*Amborella* Genome Project 2013) et disponible en annexe 3. Ils sont présentés ci-dessous.

### 2.1.2.2.1 Identification des gènes codant les globulines 11S d'*Amborella*

L'analyse par électrophorèse bidimensionnelle de la fraction globuline de la graine mature sèche d'*Amborella* nous a permis de mettre en évidence la présence de 99 spots contenant des protéines globulines 11S codées par trois modèles de gènes différents chez *Amborella* : evm\_27.model.AmTr\_v1.0\_scaffold00067.7, evm\_27.model.AmTr\_v1.0\_scaffold00067.11 et evm\_27.model.AmTr\_v1.0\_scaffold00067.12. Les séquences peptidiques prédites à partir du génome ont ensuite été extraites et utilisées pour y aligner les peptides identifiés à partir de l'analyse 2-DE (Figure 27).

#### evm\_27.model.AmTr\_v1.0\_scaffold00067.7

MAHTKSIILTFVLLSLLVCAFADRRQSRRLSDAQQCRMNRIISGIRPTRVIRSEGGITELWDEDRDEFQACAGVAATRNTLNPNS  
LYLPSFSSAPQIVYIERGRGIIGLSAPGCSESFHSGESGSIQHRKIRQGGFRDQHQKQVQRIERGVDVIAIPPGITHWCYNDDNS  
EELVAFVTVDVTSDYNQLDTKQRQFFIAGGQPRGQRKQEGEGERGQKGRQEGEYGEEOQRGEQKKEKLIQTILPQIDTRFLAEA  
LDIPIELAQKIQREDEGIIKVEKEGLRILSPEGEEREERERETGPRANVIGVGERYCNAKIRQNIESLREADIYSRHGGH  
LKTINRRNLPILDILMSAAKVTLYSDAIILAPHWSINAHTIAYITRGEQIQIIGTNGQKVMDDRVROGDVIVPQFFTSMCK  
AGSQGIEWIAIKTSDLPMNSPLVGYTSAIKGMPIEVLTNAYRISNQQAQDIKYNREDQIMIFPSSSRSSASS

#### evm\_27.model.AmTr\_v1.0\_scaffold00067.12

MAFTKSLVSVLFLSFLVSAFSLSTETRSQESLREARQCRIDRIPTSRPARRIQSEGGHTEIWDEYEDQFLCAGVAAIRNTIHP  
NSLSLPNFEPAPRLVYIQKGRGIISIVFPGCTESFQSQGYIRTRVGGEGQQQQGIKDRHQKVQTIKQGDAAIIPAGAAHWCYN  
DGSEDLIAFSITDVTNDANQIEQSLKSYLAGGQPRRGQEEGQQEQEQEQEQERQNSVNVINELNAEQLAFAFVPIELIRSL  
QKPDREGWIVRVEKESLGVIRPDEEEEEERERYSERSNQYEEREENARIRQNIIDNPRRADIYSRQAGHIQIVNRQTLPILSILD  
MSIEKGLHPNALYAPHWTINAHTIVLITRGEQNIQVIGTNGRQKVMDDRVHEGDVFIQYFTAMSKAGNEGLEWVAIKTSDL  
PMKSPILGHASAIKGIPIEVLKNAIKITTTQEARDIKLNKRDQYMLLPPRSTSSRP

#### evm\_27.model.AmTr\_v1.0\_scaffold00067.11

MALISKSLVSLVSLFSLVCAFVSYRGLQRSLDDARQCRMSKISTSQASRIIRSEGGITELWDEYEDQFQACAGVAAMRNTIQPN  
SLSLPNFSPAPRVVYIEQGSGLIGLSAPGCSESFHSGERRSFQRGGQTRERRLRDQHQKQVQRIIRSGDIIGIPPGVTHWCYND  
NSQELIAFSVIDLTNKNQLDPKLRIDTRFLAESFGIPDELAQRIQREDEGLIVRVEKQGMRLVPIPEEEEEEREREREGAS  
VNGLEELYCNVKIRQNIIDTLRDIDIYSRQGGRLKTIINRKNLPILEILMSAQKVGLYSNAIFAPHWSTNAHTIAYITRGEQIQI  
QIIGTNGQKVMDDRVROGDVIVPQFFSSMCKAGSQGIEWIAIKTSDLPIHTPLVGYTSAIKGMPIEVLTNAYRISNQQAQDI  
KYNREEQLMIFPSSSRSDSS

Figure 27 : recouvrement des séquences nucléotidiques (séquences codantes) et protéiques des trois gènes de globulines 11S d'*Amborella* identifiés à partir des données peptidiques expérimentales. Les acides aminés surlignés en jaune représentent les régions couvertes par les peptides obtenus expérimentalement à partir des analyses protéomiques en 2-DE des globulines de la graine mature sèche d'*Amborella*.

Pour les gènes 67.7 et 67.12, les taux de recouvrement sont très élevés, de 84% et 73% respectivement (Figure 27). On note qu'expérimentalement, les séquences des peptides obtenues se chevauchent souvent permettant ainsi de conforter plus encore l'identification des gènes (tableau supplémentaire 4). De même, ces forts taux de recouvrement attestent de la qualité du séquençage du génome, ainsi que de la prédiction des introns et des exons. Par ailleurs, le faible taux de recouvrement du gène 67.11, de seulement 4,1%, pourrait indiquer que la protéine correspondante s'accumule moins que les deux autres. Suite à la soumission de cet article (*Amborella* Genome Project 2013), nous avons obtenu d'autres résultats sur le

séquençage des protéines totales de l'albumen d'*Amborella* par protéomique *shotgun*. La même approche du niveau de recouvrement gène/protéine a été réalisée sur la base des peptides expérimentaux obtenus (Figure 28). Là encore, les séquences des peptides obtenus se chevauchent et confortent nettement nos résultats d'identification (tableau supplémentaire 4), tant du point de vue de l'identification des trois gènes codant les globulines 11S d'*Amborella* que de la qualité du séquençage et de l'annotation du génome de cette plante. Par ailleurs, les peptides obtenus à partir de l'analyse protéomique *shotgun* ont permis d'augmenter le taux de couverture du gène 67.11, de 4,1% avec l'approche 2-DE à 61% avec l'approche *shotgun* (

Tableau 6). Ce taux est cependant plus faible que ceux des gènes 67.7 et 67.12, confortant l'hypothèse que le gène 67.11 soit moins exprimé que les deux autres.

```

evm 27.model.AmTr v1.0 scaffold00067.7
MAHTKSIILTFVLLSLLVCAFADRRQSQRRLSDAQQCRMNRISGIRPTFVIRSEGGITELWDEDRDEFQACAGVAATRNTLNPNS
LYLPSFPSSAPQIVYIERGRGIIGLSAPGCSESFHSGESGSIQHRKIRGQGFDRDQHQKQVRI ERGDVIAI PPGITHWCYNDNS
EELVAFTVTDVTSYDYNQDLTKORQFFIAGGQPRGQRKQGEGERGQKGRQEGEYGEERQGEQKKEKLIQTILPQIDTRFLAEB
LDIPIELAQKIQREDEIRGIIKVEKEGLRILSPEGEEREEERERETGPRAVIGVGEFYCNAKIRQNIESLREADIYSRHGGH
LKTINRRNLPILDILDMSAAKVTLYSDAI LAPHWS INAHTIAYITRQEGQIQIIGTNGQKVMDDRVROGDVIVPQPFQTSMCK
AGSQGIEWIAIKTSDLPMN SPLVGYTSAIKGMPIEVL TNAYRISNQQAQDIKYNREDQIMIFPSSSE SASS

```

```

evm 27.model.AmTr v1.0 scaffold00067.12
MAFTKSLVSVFLVFLVSAFSLSTETRSQESLRERQCRIDRIPTSRPARRIQSEGGHTBIWDEYEDQFLCAGVAATRNTIHP
NSLSLPNFEPAPRVYIQKGRGIIISIVFPGCTESFQSGYIRTRVGGEGQQQGIKDRHQKVQTIKQGDIAIPAGAAHWQYN
DGSEDLIAFSITDVTNDANQIEQSLKSFYLAGGQPRRQEQEQEQEQEQEQEQERQNSVNVINELNAEQLAFAFAVPIELIRSL
QKPDBERGWIVRVEKESLGVIRPDEEEERERYSERSRYEEREENARIQONIDNPRRADIYSRQAGHIQIVNRQTLPILSILD
MSIEKGHLHPNALYAPHWTINAHTI VLI TRGEGNIQVIGTNGRKVMDDRVHEGDVFIPOYFTAMSKAGNEGLEWVAIKTSDI
PMKSPILGHASAIKGIPIEVLKNAKYITTOEARDIKLNRKDOYMLLPPHSTSSRP

```

```

evm 27.model.AmTr v1.0 scaffold00067.11
MALISKSLVSLVSLVCAFVSRGKLRSLDDARQCRMSKISTSQASRIIRSEGGITELWDEYEDQFQACAGVAAMRNTIQPN
SLSLPNFSPAPRVYIQKGRGIIISIVFPGCTESFQSGYIRTRVGGEGQQQGIKDRHQKVQTIKQGDIAIPAGAAHWQYN
NSQELIAFSVIDLTNRINQLDPKLRIDTRFLAESFGIPDELAQRIQREDEGLIVRVEKQGMRVLIPEEEEEEREREREGAS
VHGLBELYCNVKIRQNIIDTLRDIDIYSRQGGRLKTTNRKNLPILEILDMSAQKVGLYSNAIFAPHWSTINAHTIAYITRQEGQI
QIIGTNGQKVMDDRVROGDVIVPQPFQTSMCKAGSQGIEWIAIKTSDLPVHTPLVGYTSAIKGMPIEVL TNAYRISNQQAQDI
KYNREBQIMIFPSSSE SDSS

```

Figure 28 : recouvrement des séquences nucléotidiques (séquences codantes) et protéiques des trois gènes de globulines 11S identifiés à partir des données peptidiques expérimentales obtenues par protéomique *shotgun*. Les acides aminés surlignés en bleu représentent les régions couvertes par les peptides obtenus expérimentalement à partir des analyses de protéomique *shotgun*.

Ces résultats illustrent clairement l'intérêt de combiner les différentes approches de protéomique pour la caractérisation du contenu protéique d'un organe.

Tableau 6 : taux de recouvrement des séquences peptidiques prédites des trois gènes codant les globulines 11S d'*Amborella* par les peptides obtenus expérimentalement à partir de l'approche 2-DE ou celle de protéomique *shotgun*.

Gène	67.7	67.12	67.11
2-DE	84%	73%	4,10%
<b>Shotgun</b>	82,30%	88,90%	61%



## Analyse phylogénétique des gènes de globulines 11S d'*Amborella*

L'utilisation des globulines 11S comme marqueurs de l'évolution des espèces végétales a été décrite et utilisée de nombreuses fois (Li et al. 2012). Ainsi, après avoir conforté le niveau d'identification des gènes codant les globulines 11S d'*Amborella*, un arbre phylogénétique a été construit en utilisant 65 séquences de globulines 11S provenant de 55 espèces différentes (Figure 29). Les séquences utilisées sont disponibles dans la publication située en annexe 3.

L'arbre obtenu montre des regroupements en clades similaires à ceux décrits dans la version actualisée de la classification des végétaux supérieurs (Bremer et al. 2009). En effet, les gymnospermes sont regroupées à la base de l'arbre et bien séparées des angiospermes et d'*Amborella* (Figure 29). Cette analyse conforte ainsi le positionnement basal d'*Amborella*.

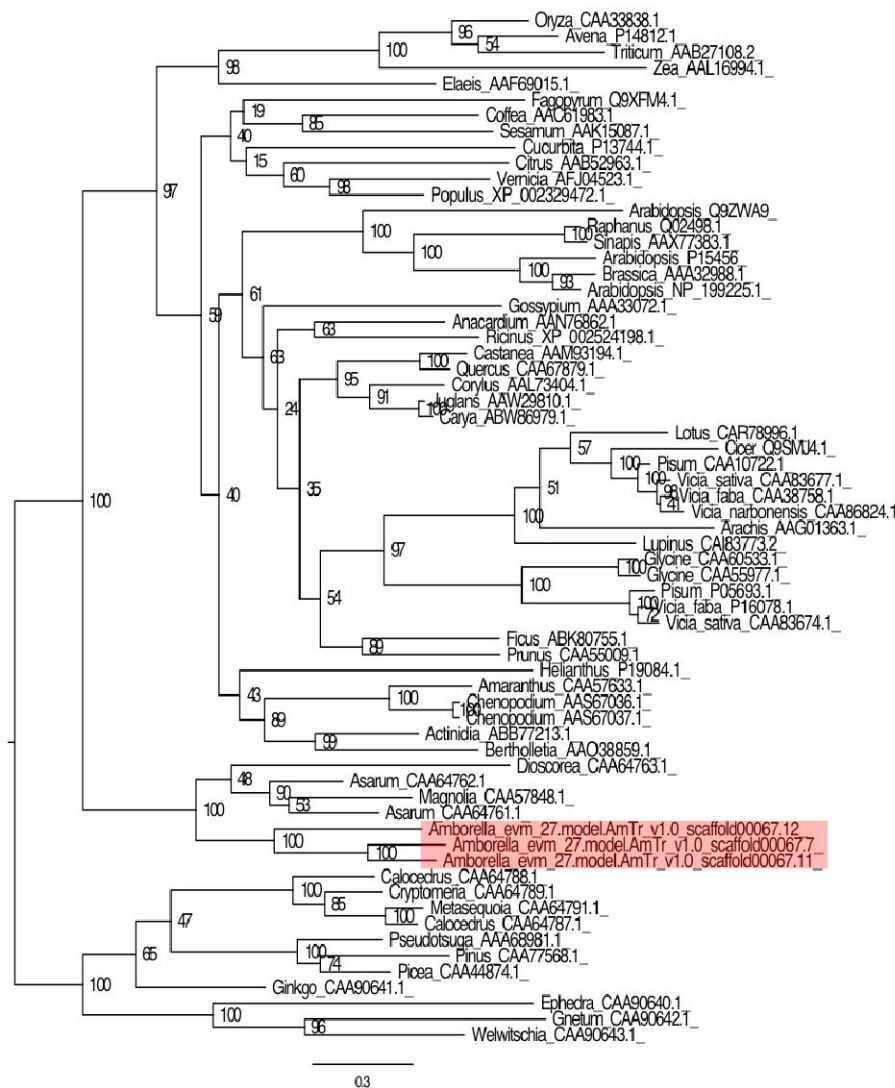


Figure 29 : arbre phylogénétique construit à partir des séquences des globulines 11S de diverses espèces. L'alignement a été réalisé avec le logiciel MAFT et l'arbre construit avec la méthode du « maximum likelihood » à partir de 65 séquences disponibles (55 espèces

différentes). L'analyse de bootstrap a été fixée à 100 pseudo-réplicats. Les trois gènes codant les globulines 11S d'*Amborella* sont surlignés en rouge.

Par ailleurs, les globulines 11S de monocotylédones auraient également une origine monophylétique alors que les globulines 11S des dicotylédones seraient polyphylétiques (Li et al. 2012). Il est donc important de noter que les trois gènes d'*Amborella* semblent être d'origine monophylétique. De plus, ils sont proches d'espèces décrites comme archaïques telles que *Asarum europeum* et *Magnolia salicifolia* et à la jonction entre les gymnospermes et les angiospermes, ce qui est en accord avec la vision globale de la phylogénie (Soltis et al. 2011). De même, les monocotylédones et les dicotylédones sont en groupes distincts. Ces résultats confirment ainsi l'aptitude des globulines 11S à être utilisées pour des études phylogénétiques.

#### **Alignement des séquences génomiques et protéiques des globulines 11S d'*Amborella trichopoda*, *Glycine max*, *Ginkgo biloba* et *Arabidopsis thaliana***

Nous avons aligné les séquences peptidiques prédites des trois gènes codant les globulines 11S d'*Amborella* avec les séquences des globulines 11S de soja (*Glycine max*), d'*Arabidopsis thaliana* (la plante modèle des dicotylédones) et de *Ginkgo biloba* (gymnosperme) dont les séquences sont connues, voire même les structures tridimensionnelles (Adachi et al. 2003). Ces résultats sont présentés dans la Figure 30.

L'alignement des séquences peptidiques de globulines 11S d'*Amborella* avec les séquences connues montre clairement la conservation des deux cystéines impliquées dans la formation du pont disulfure qui permet de maintenir ensemble les sous-unités A et B après le clivage de la pro-globuline pendant sa maturation (rectangles rouges, Figure 30). L'alignement révèle également la conservation d'acides aminés impliqués dans la formation et la stabilisation du trimère (en jaune, Figure 30) mais également d'autres acides aminés impliqués dans l'assemblage globulaire correct des globulines 11S (en bleu, Figure 30) (Tandang-Silvas et al. 2010). Par ailleurs, une région de 52 acides aminés est présente chez *G. max* (Figure 30) mais absente des globulines 11S codées par les gènes d'*Amborella* et de *G. biloba*. Cette région, qui joue un rôle prépondérant dans l'assemblage des formes hexamériques des globulines 11S matures (Tandang-Silvas et al. 2010), serait donc apparue après la divergence entre *Amborella* et les autres angiospermes. Il serait intéressant par la suite d'étudier plus finement les mécanismes d'assemblage des globulines 11S d'*Amborella*

par établissement de leur structure tridimensionnelle, comme cela a été décrit pour d'autres espèces (eg. Adachi et al. 2003 dans le cas des globulines 11S de soja).

Plus encore, on note la conservation du site de coupure par l'endopeptidase vacuolaire/asparaginyl endopeptidase (VPE/Asn-endopeptidase) chez *Amborella* (flèche rouge, Figure 30).



(Uniprot : P15455). La flèche rouge indique le site de coupure par l'endopeptidase vacuolaire qui libère les sous-unités A (flèche bleu) et B (flèche verte) pendant la maturation. Les deux rectangles rouges représentent les résidus cystéine conservés impliqués dans la formation du pont disulfure liant les deux sous-unités. Les acides aminés surlignés en jaune sont impliqués dans la formation et la stabilisation du trimère (formes précurseurs) et les acides aminés surlignés en bleu sont impliqués dans l'assemblage globulaire correct des globulines 11S (Tandang-Silvas et al. 2010). « Box a » et « Box b » correspondent aux homologues des motifs consensus PX14G et SX3G respectivement (Häger et al. 1996).

Ce site de coupure est crucial puisqu'il permet le clivage de la forme précurseur en sous-unités A et B pendant la maturation des globulines, assurant ainsi la compaction sous forme hexamérique des globulines dans les corps protéiques. Cette protéolyse spécifique et limitée se réalise entre une asparagine et une glycine, résultant ainsi en la formation d'une extrémité glycylo au niveau N-terminal de la sous-unité B. Ce site de coupure Asn-Gly (N-G) est très conservé chez les globulines de type 11S. Chez *A. thaliana* ce site est reconnu et clivé par le complexe VPE codé par les quatre gènes  $\alpha VPE$ ,  $\beta VPE$ ,  $\gamma VPE$  et  $\delta VPE$  (Gruis et al. 2002). Une recherche par homologie de séquence nous a permis de retrouver ces quatre gènes chez *Amborella*. Il s'agit des gènes *evm\_27.model.AmTr\_v1.0\_scaffold00036.100*, *scaffold00002.262*, *scaffold00002.263* et *scaffold00002.265*.

Nous pouvons conclure que les deux globulines 11S d'*Amborella* codées par les gènes 67.11 et 67.12, et qui contiennent le site de coupure N-G (Figure 30), sont susceptibles d'être reconnues et hydrolysées par le complexe VPE. En revanche, la séquence peptidique codée par le gène 67.7 ne possède pas la séquence consensus N-G qui est remplacée par une séquence N-V-I (Figure 30). Ce type d'altération est retrouvé dans les séquences de globulines 11S de gymnospermes, comme pour *Ginkgo biloba* et *Metasequoia glyptostroboides* (Figure 30) (Häger & Wind 1997), mettant ainsi en évidence le caractère ancestral du gène 67.7 d'*Amborella*.

Les séquences nucléiques des trois gènes codant les globulines 11S d'*Amborella* ont été alignées avec celles d'*A. thaliana*, *Glycine max* et *Ginkgo biloba* afin d'établir la position des introns et des exons (Figure 31) et la structure des protéines (présence de zones ordonnées et désordonnées) (Figure 32).

La sous-unité B des gymnospermes présente une organisation intron/exon très conservée et différente de celle des angiospermes (Häger et al. 1996). Plus particulièrement, les gènes de globulines 11S de gymnospermes ont un intron supplémentaire (intron IV) qui interrompt la séquence codante C-terminale de la sous-unité B. Cet intron se retrouve chez *G. biloba* mais pas pour les trois gènes d'*Amborella* ni chez les autres homologues angiospermes (*G. max* et *A. thaliana*) (Figure 31). Ce résultat suggère donc que cet intron a été perdu chez les angiospermes lors de l'évolution. Par ailleurs, on note la présence de motifs homologues à PX14G et SX3G pour les trois gènes d'*Amborella* (Box a et Box b, Figure 30). Ces motifs ont été décrits comme étant adjacents aux introns III et IV respectivement, et sont présents chez les gymnospermes et les angiospermes, indépendamment de la présence de l'intron IV (Häger et al. 1996).

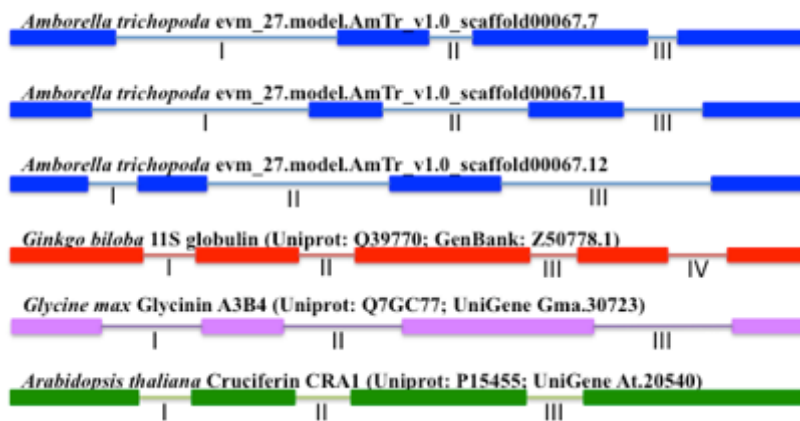


Figure 31 : représentation schématique de la structure indiquant les exons (traits épais) et les introns (traits fins) des gènes codant les globulines 11S pour *Amborella*, *G. max*, *G. biloba* et *A. thaliana*. Les introns sont numérotés en chiffres romains.

Nous avons également étudié, par prédiction de structure, la répartition, le long des séquences protéiques des zones structurées et déstructurées des globulines 11S à partir des séquences génomiques (Kelley & Sternberg 2009) (Figure 32). Il apparaît que la structure prédite de la globuline codée par le gène 67.11 est très différente des deux autres (Figure 32).

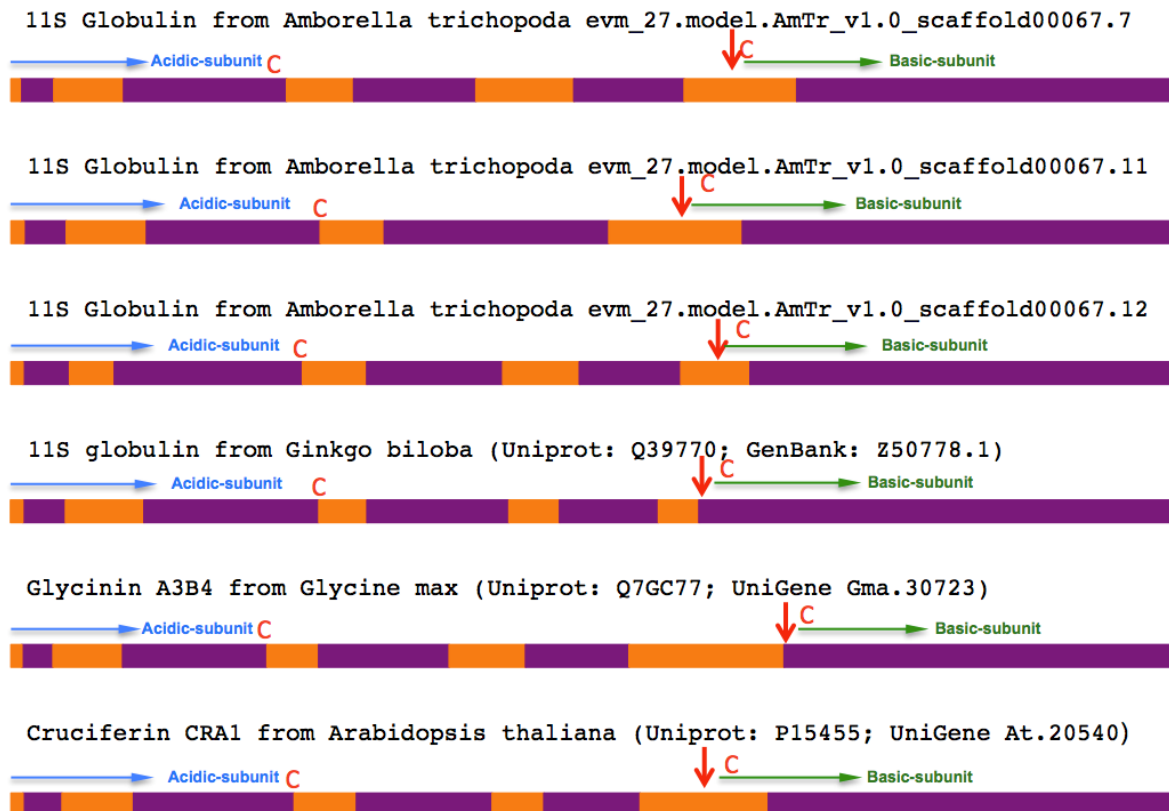


Figure 32 : prédictions structurales des globulines 11S d'*Amborella*, *G. max*, *G. biloba* et *A. thaliana* déduites des séquences des gènes codant les globulines 11S chez ces espèces grâce au serveur phyre2 (Protein Homology/AnalogY Recognition Engine) (<http://www.sbg.bio.ic.ac.uk/phyre2>) (Kelley & Sternberg 2009). Les rectangles mauves représentent les régions ordonnées, les rectangles oranges les régions désordonnées. Les « C » en rouge correspondent aux résidus cystéine conservés. La flèche rouge correspond au site de coupure par le système VPE.

En effet, dans la séquence prédite par ce gène, il manque une zone désordonnée dans la sous-unité A. Au cours de la germination, la mobilisation des globulines 11S commence préférentiellement par la protéolyse de la sous-unité A (Scott et al. 1992 ; Gallardo et al. 2001). Ainsi, l'absence de cette région désordonnée pourrait perturber la compaction de la protéine pendant sa maturation ou sa mobilisation lors de l'imbibition (Adachi et al. 2003). L'étude de l'origine des sous-unités A et B a révélé qu'il s'agit en fait d'un élément de duplication d'un gène ancestral unique (Li et al. 2012). Ainsi, l'absence de la région désordonnée du gène 67.11 pourrait être liée à une fonction divergente aux deux autres protéines.

#### 2.1.2.2.2 Classes ontologiques des globulines

L'analyse de la fraction globulines a également révélé d'autres protéines que les globulines 11S (tableau supplémentaire 4). Ces protéines ont été classées en fonction de leurs catégories et fonctions ontologiques (Bevan et al. 1998). La proportion de chaque classe (ou fonction) est exprimée en pourcentage du nombre de protéines par classe par rapport au nombre total de protéines identifiées (Figure 33).

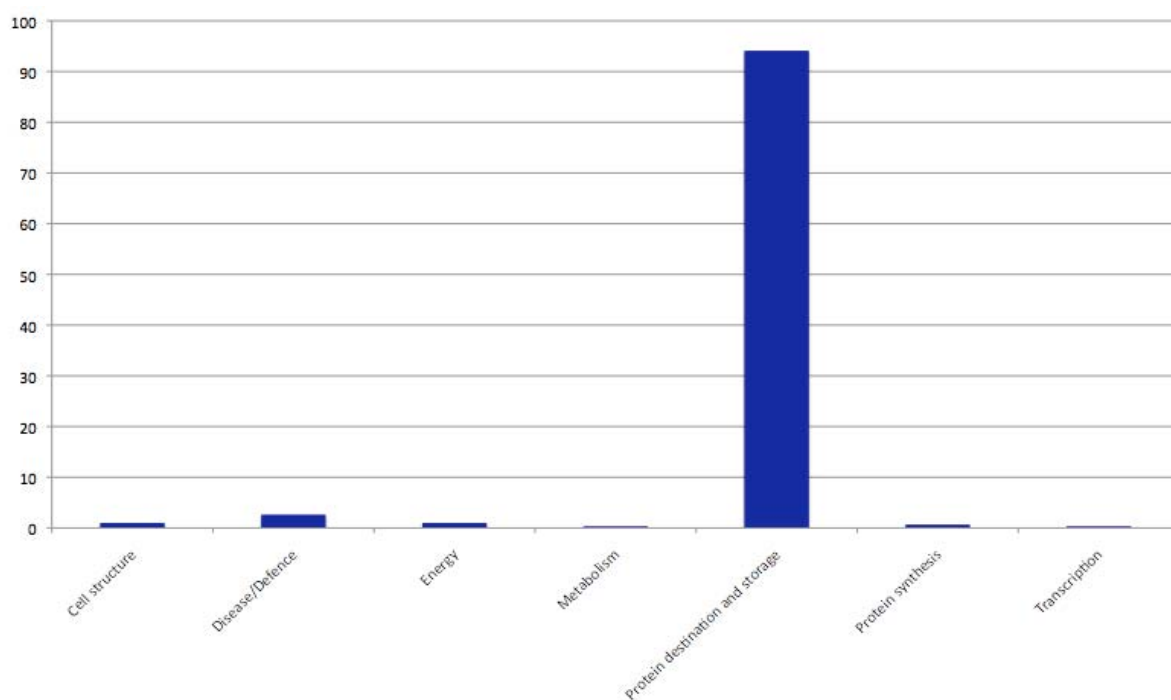


Figure 33 : représentation des classes ontologiques (Bevan et al. 1998) de la fraction globulines des graines matures sèches d'*Amborella*.

La catégorie « *protein destination and storage* » représente à elle seule 94% des protéines identifiées, dont 91% pour la fonction « *storage proteins* » (qui contient les globulines 11S présentées ci-dessus).



## **2.2 Discussion**

La graine est l'organe principal de la dissémination des plantes, un processus d'importance centrale tant du point de vue écologique (conservation/survie des espèces végétales, donc de la biodiversité végétale) qu'agronomique (nécessité de maintenir le rendement des cultures dans un contexte de très forte augmentation démographique). La germination est le premier stade du développement d'une plante, correspondant à une phase cruciale puisque de son succès dépend l'installation d'une plantule vigoureuse (Rajjou et al. 2012). Au cours de cette phase, l'embryon doit passer d'un état quiescent (graine mature sèche) à un état métaboliquement actif (plantule), via une transition associée à de nombreux changements biochimiques et physiologiques complexes (Rajjou et al. 2012). Notre étude a pour but de caractériser les modifications métaboliques permettant la germination de graines telles que celles d'*Amborella* possédant un embryon de très petite taille. De nombreuses études protéomiques existent pour les graines à gros embryons (pour une revue voir Rajjou et al. 2012). Par contre nous ne disposons pas de données équivalentes concernant les graines à petits embryons. Notre ambition est ici de combler cette lacune.

Au plan physiologique, certains auteurs avancent que ces embryons de petite taille enfouis au sein d'un volumineux tissu nourricier seraient immatures (expliquant leur très faible taille) et l'imbibition des graines correspondantes entraînerait une reprise de leur maturation, se manifestant par une croissance au dépend des tissus nourriciers (Baskin & Baskin 2001 ; Koornneef et al. 2002 ; Hilhorst et al. 2010). La fin d'une telle « maturation » aboutirait à l'émergence d'une partie de l'embryon au travers les enveloppes de la graine. Selon cette hypothèse, la germination *sensu stricto* de ces graines à petits embryons serait ainsi très similaire de celle communément admise pour les graines à gros embryons, lesquels atteignent leur maturité maximale lors des phases finales de la maturation des graines sur la plante mère, avant dispersion dans l'environnement.

Dans une autre hypothèse, cette longue période de temps nécessaire à la germination *sensu stricto* des graines à petits embryons serait une caractéristique de l'existence d'une dormance morpho- ou morphophysiologique de ces graines (Baskin & Baskin 2001 ; Finch-Savage & Leubner-Metzger 2006 ; Baskin & Baskin 2007).

Une dernière hypothèse, qui à notre connaissance n'a pas encore été envisagée, serait que cette croissance interne de l'embryon induite par l'imbibition de la graine corresponde en fait à une germination interne, ce qui amènerait ainsi à revoir la définition de la germination *sensu stricto* pour ce type de graines.

Une motivation importante de notre recherche était de déterminer si les mécanismes fondamentaux de la germination d'une graine d'espèce ancestrale (*Amborella*) sont similaires ou différents de ceux abondamment étudiés par diverses approches omiques (transcriptomique, protéomique, métabolomique) chez des espèces modèles (eg., *Arabidopsis*, riz) ou d'intérêt agronomique façonnées au cours d'étapes de domestication/ sélection des plantes (Skylas et al. 2000 ; Gallardo et al. 2001 ; Agrawal & Rakwal 2006 ; Catusse et al. 2008 ; Luo et al. 2011 ; Rajjou et al. 2012 ; Deng et al. 2013).

Pour répondre à ces questions, nous avons, dans un premier temps, étudié et caractérisé le protéome de la graine mature sèche d'*Amborella*, angiosperme la plus basale, et suivi, dans un deuxième temps, l'évolution de ce protéome au cours de la germination, plus précisément suite à l'imbibition de la graine mature sèche. Nos résultats montrent des originalités structurales et protéomiques qui soulèvent des questions particulières discutées ici.

## **2.2.1 Caractérisation cellulaire et protéomique de la graine mature sèche d'*Amborella***

### **2.2.1.1 Caractérisation cytologique des graines matures sèches d'*Amborella***

Nos résultats montrent que l'embryon d'*Amborella* est entouré d'un albumen volumineux (Figure 20), en accord avec des observations précédentes (Tobe et al. 2000 ; Fogliani et al. 2013). La graine contient donc essentiellement deux compartiments distincts, l'embryon d'une part, et l'albumen d'autre part, l'albumen étant le compartiment le plus volumineux, classant ainsi *Amborella* parmi les graines à très petit embryon, caractéristique des graines des plantes basales (Forbis et al. 2002 ; Baskin & Baskin 2007). De plus, nous montrons pour la première fois, que l'embryon est entouré par un réseau de polysaccharides entremêlé à un réseau de cellules vides, mortes (Figure 20). Un tel réseau de polysaccharides n'a pas été observé chez les graines d'autres espèces basales, notamment *Trimenia austinensis*, *Trithuria submersa*, *Trithuria cowieana*, *Trithuria lanterna*, *Nymphaea lotus* ou encore *Hydatella inconspicua* (Friedman 2008 ; Rudall et al. 2009 ; Tuckett et al. 2010). En revanche, la présence de cellules mortes entourant l'embryon a été décrite pour la graine de céleri qui contient également un embryon de très peu petite taille au sein d'un albumen volumineux (Jacobsen & Pressman 1979).

La germination des graines de *T. submersa* et *T. austinensis* est très particulière dans le sens que l'embryon rompt les téguments de la graine avant de se différencier (Tuckett et al.

2010). Cette stratégie expose donc l'embryon au milieu extérieur avant sa maturité, alors qu'il est très vulnérable. En revanche, l'embryon d'*Amborella* se développe et se différencie dans la graine avant germination *sensu stricto* tout comme pour le céleri (Jacobsen & Pressman 1979 ; Fogliani et al. 2013). Cette stratégie permet ainsi de libérer une plantule vigoureuse qui aura plus de chances de réussir son installation (par analogie avec la stratégie de la poche du kangourou ou de l'hippocampe). Dans cette optique, la présence du réseau de polysaccharides pourrait représenter une évolution structurale permettant de protéger l'embryon de la pression mécanique de l'albumen et de lui permettre de se différencier dans la graine.

Par ailleurs, le fond bleuté des cellules de l'embryon (après coloration au bleu noir naphthol) contraste avec le fond clair des cellules de l'albumen qui sont parsemées de corps protéiques, lieux de l'accumulation des protéines de réserve chez les graines (Figure 20). Nous observons que les parois des cellules de l'embryon sont beaucoup plus fines que celles de l'albumen (Figure 20). Cet aspect est typique d'un organe en pleine différenciation dont les cellules subissent de nombreuses divisions (Daher et al. 2010). Notre observation pose la question de l'état de maturité de l'embryon au moment de la dispersion de la graine. Afin d'étudier plus avant le niveau de maturité de l'embryon et les spécificités et points communs aux deux compartiments de la graine d'*Amborella*, nous avons mis en œuvre une étude protéomique de ces derniers compartiments.

#### ***2.2.1.2 Analyses par protéomique shotgun des protéines totales de l'embryon et de l'albumen isolés de la graine mature sèche d'Amborella***

L'organisation structurale très différenciée entre l'albumen et l'embryon nous a incités à étudier les protéomes de chacun des deux tissus afin de mettre en évidence des similitudes et/ou des différences entre eux. De par la très petite taille de l'embryon d'*Amborella* et la grande difficulté à l'extraire de la graine mature sèche, nous avons choisi d'utiliser l'approche protéomique en vrac (*shotgun*). En effet, cette technique très sensible requiert beaucoup moins de matériel végétal que l'approche basée sur la caractérisation des protéines par électrophorèse 2D-PAGE. Malgré cette difficulté, notre approche globale a été couronnée de succès. En effet, 415 protéines ont pu être extraites de l'embryon, et 69 de l'albumen (tableaux supplémentaires 1 et 2). Il s'agit, à notre connaissance, de la première étude ayant caractérisé le protéome total d'embryons de si faible taille.

L'approche de protéomique *shotgun* des protéines totales de l'albumen révèle que 82,6% (en abondance) d'entre elles sont des protéines de réserve alors que ces mêmes

protéines de réserve ne représentent que 0,7% (en abondance) des protéines identifiées à partir de l'embryon isolé (Figure 21). Ce résultat est en accord avec les observations cytologiques (Figure 20) qui montrent la présence de nombreux corps protéiques dans l'albumen, mais non pas dans l'embryon. Sachant que les protéines de réserve s'accumulent au cours de la maturation des graines (Bewley & Black 1994), cette différence d'accumulation des protéines de réserve entre l'embryon et l'albumen soulève de nouveau la question du stade de maturité de l'embryon au sein de la graine mature sèche d'*Amborella*.

Nous avons analysé plus avant cette question par plusieurs approches. Nous avons ainsi réalisé des observations sur les réserves lipidiques accumulées dans la graine mature sèche d'*Amborella*. Les graines d'*Amborella* étant en effet très riches en réserves lipidiques (tableau supplémentaire 5), nous avons cherché à visualiser la présence éventuelle de corps lipidiques (organes de stockage des triglycérides chez toutes les graines) (Frandsen et al. 2001 ; Poxleitner et al. 2006) au sein des graines matures sèches d'*Amborella*. L'observation au microscope confocal de coupes de graines d'*Amborella* colorées au rouge Nil nous a permis de visualiser les corps lipidiques (annexe 4). Il est clair que ces derniers sont présents dans l'embryon et dans l'albumen, à l'inverse des corps protéique qui, eux, ne sont décelables que dans l'albumen (Figure 20).

Les corps lipidiques sont composés d'une matrice de triglycérides (TAGs) entourée d'une couche de phospholipides dans laquelle est imbriquée un petit nombre de protéines transmembranaires. Ces dernières permettent de stabiliser la structure des corps lipidiques pendant la dessiccation et participent de plus au recrutement des corps lipidiques à la germination (Frandsen et al. 2001). Notre approche a permis de mettre en évidence que la graine mature sèche d'*Amborella* possède toutes les protéines nécessaires à la stabilisation puis la mobilisation des corps lipidiques, ainsi qu'à l'utilisation des TAGs comme source de carbone. En effet, nous avons pu identifier :

- Cinq oléosines dans l'embryon et une présente à la fois dans l'albumen et l'embryon (tableaux supplémentaires 1 et 2) (Figure 34). Ces protéines stabilisent les corps lipidiques et sont retrouvées chez les angiospermes et les gymnospermes (Lee et al. 1994). Leur rôle principal est d'empêcher la coalescence des corps lipidiques au moment de la dessiccation, mais aussi pendant la phase très rapide d'imbibition (Leprince et al. 1998). Il est intéressant de noter que les oléosines sont présentes en quantité beaucoup plus faible dans les graines récalcitrantes, intolérantes à la dessiccation, en comparaison des graines orthodoxes, tolérantes à la dessiccation (Leprince et al. 1998 ; Guilloteau et al. 2003 ; Lee et al. 2012). L'identification de

ces protéines en quantité et nombre importants chez les graines d'*Amborella*, est donc un bon indicateur d'acquisition de l'aptitude de la tolérance à la dessiccation de ces graines (Lee et al. 2012).

- Deux stéréoléosines dans l'embryon (Q9T0G0, Q9LUF1) et une dans l'albumen (Q9T0G0) (tableaux supplémentaires 1 et 2), accumulées pendant les phases tardives de la maturation. Le rôle de ces protéines est encore peu connu, mais elles seraient impliquées dans la transmission du signal pour la croissance notamment par la voie des brassinostéroïdes, et pourraient être impliquées dans la levée de dormance et l'utilisation des réserves lipidiques pendant la germination (F. Li et al. 2007 ; Baud et al. 2009).
- Une caléosine (O81270) (tableau supplémentaire 2), accumulée tardivement pendant la maturation et dont le rôle serait d'associer les peroxyosomes (dans lesquels se manifeste la mobilisation des TAGs chez les graines, (Eastmond 2007)) et les corps lipidiques au moment de la mobilisation des réserves ou de faciliter l'interaction entre les lipases et les corps lipidiques lors de la germination (Naested et al. 2000 ; Poxleitner et al. 2006).

Un examen approfondi des séquences (extraites du génome) des cinq oléosines décrites ci-dessus montre que les oléosines d'*Amborella* présentent le motif *proline-knot* présent chez toutes les oléosines (résidus en jaune dans la Figure 34) et qui est impliqué dans la stabilisation de ces dernières (Abell et al. 1997).

La séquence correspondant à l'oléosine Q39165 (scaffold *Amborella* 00014.132) présente à la fois dans l'embryon et l'albumen possède de plus une séquence nommée *H-form insertion*, correspondant à une insertion de 18 acides aminés en position C-terminale chez les oléosines de type H (Figure 34) (Tai et al. 2002 ; Hyun et al. 2013). Il est intéressant de noter que les gymnospermes n'ont que des oléosines du type L (c'est-à-dire dépourvues de la séquence H), alors que chez les angiospermes on trouve les deux types, H et L, d'oléosines (Wu et al. 1999 ; Liu et al. 2012 ; Hyun et al. 2013). De fait, la forme L correspondrait à la forme primitive des oléosines. Puisqu'une forme H est présente à la fois dans l'embryon et l'albumen, nous pouvons conclure qu'*Amborella*, du point de vue de la structure de ses oléosines, présente les caractéristiques des angiospermes. De plus, les peptides expérimentaux confirment l'identification des acides aminés impliqués dans la forme H (Figure 34).

Oléosine d'*Amborella* (Q39165) (scaffold *Amborella* 00014.132) accumulée dans l'embryon et l'albumen

MAEQOYHSVMDLIHDKAPSTSQVLAVVTLFPIGGILLTLSGITLAGTLACLAVVTPFLFVLFSPVLVPAIGILLLLA  
VSGFLASGAFGLTGVSSIGWIVKSVRGRGDVREQMEHAKRKMHESSGGHMGOOVGORAKDVGQGMCSKAGDGGRT

Oléosine d'*Amborella* (P29525) (scaffold *Amborella* 00056.169) accumulée dans l'embryon

MAEQPKPVTKLEAEKPPPTSHQAVKFSVAATAGAVLLLLLAGLTLTGTVIVLVLATPVLVIFSPILVPAIGIAIFL  
AVTGFLFSGGFGVAALSALAWIYNYVTGKHPVGS DKIDRARMTLASKAKDMKEKAKOYGOYVQOKAQEAST

Oléosine d'*Amborella* (Q9SS98) (scaffold *Amborella* 00061.105) accumulée dans l'embryon

MADRKVQEMKNKTS GSSVLSSSMAALVIGGPLLGM SFGSFLASMTILIISSPLFLIFSPLILVPAFVVLGCMVAFA  
AAGLMGMAGFCAITCVFRSVRGQOIGRIPAOVVDTAQKAKEHGKELGAHLOORVOIPIENKTRG

Oléosine d'*Amborella* (Q9LII2) (scaffold *Amborella* 00147.30) accumulée dans l'embryon

MADQRQARAASRPGPTPPIGGGATTRE RGFIO RLQEHAPNSTQVIGLLTLVIVSGAILLLLLTGLTLAATTVGLIVI  
TFLVLLSSPVWIPVLAFLGVAASGFVAVTGFLAVLGGLTWMYRYIKGRHPVGS DRVDYARMRLTDTASQMKDYA  
RDYRGYLQ GKVKDAAPGA

Oléosine d'*Amborella* (P29525) (scaffold *Amborella* 00111.94) accumulées dans l'embryon

MADRRSOLALSRSDQDITPKSHKVVKGLTAATLGGGSLVASGLTLTGTIIAITIATPLLIVFSPILVPAAGALLL  
ATAGFVTSGGFVAVAIIAFSWMYGYVMGKHPGAKTLDQAKAMVASKAODLKEQYASHTQSQ

Figure 34 : séquences des oléosines d'*Amborella* identifiées par approche protéomique et extraites du génome. En jaune le motif *proline-knot* conservé chez toutes les oléosines. En vert, la séquence H caractérisant les oléosines d'angiospermes. Les peptides identifiés expérimentalement sont soulignés.

En conclusion, nos données sur les corps lipidiques et leur détection non ambiguë au sein des embryons suggèrent très fortement d'une part que les cellules embryonnaires ont atteint un état de maturation avancé. D'autre part, leur présence révèle une utilisation préférentielle des lipides de réserve comme source d'énergie au moment de la reprise métabolique accompagnant l'imbibition des graines d'*Amborella*. Il est établi en effet que, durant la germination, les TAGs sont dégradés en acides gras qui sont ensuite utilisés pour la  $\beta$ -oxydation puis au sein du cycle du glyoxylate (Bewley & Black 1994). Chez *A. thaliana* cette voie est privilégiée au moment de la germination (Penfield et al. 2005). Il en est de même chez la betterave à sucre (Catusse, Strub, et al. 2008). Dans ce contexte, nous notons que la plupart des enzymes de la  $\beta$ -oxydation (à l'exception de l'acyl-CoA oxydase) et du cycle du glyoxylate (à l'exception de l'isocitrate lyase) ont été identifiées dans l'embryon d'*Amborella* (tableau supplémentaire 1). De fait, l'embryon possède la machinerie nécessaire à l'utilisation des lipides de réserve, la forme majeure de stockage chez la graine d'*Amborella*.

En revanche, l'absence de corps protéiques au niveau des cellules embryonnaires (Figure 20) signifierait que ces cellules ne nécessitent pas l'hydrolyse de telles réserves suite à l'imbibition des graines, qui sont massivement accumulées dans l'albumen (Figure 20 et Figure 21). Il s'ensuivrait que l'hydrolyse des protéines de réserve de l'albumen serait

suffisante pour la croissance interne de l'embryon, suggérant un transport actif des métabolites issus de cette mobilisation de l'albumen vers l'embryon.

D'autres données protéomiques, acquises au cours de cette thèse, sont compatibles avec l'atteinte d'un état avancé de maturité de l'embryon d'*Amborella*. Par exemple, l'embryon d'*Amborella* accumule tout un arsenal de protéines chaperons, dont la fonction est d'assister d'autres protéines dans leur maturation, leur assurant un repliement tridimensionnel adéquat. Il s'agit notamment des protéines nommées *Late Embryogenesis Abundant* (LEA) et des protéines de choc thermique (*Heat Shock Protein*, HSP) qui protègent les complexes macromoléculaires des stress tels que dessiccation, conservation au sec et imbibition (tableau supplémentaire 1). Il est bien documenté que ces protéines s'accumulent lors des phases tardives de la maturation, notamment lors de la phase de dessiccation. La présence de ces protéines chez l'embryon d'*Amborella* est donc révélatrice d'un état de maturité assez avancé, permettant la survie à sec de l'embryon. Leur expression est en partie sous le contrôle de l'ABA (Kotak et al. 2007 ; Tunnacliffe & Wise 2007). Un des stress majeurs que doit affronter une graine est la déshydratation intense qu'elle subit à la fin de la maturation. En effet, la perte en eau (de 90% à 5% chez les graines orthodoxes) peut causer la déstructuration et l'agrégation des protéines, voire leur dénaturation sous l'effet des espèces réactives de l'oxygène générées lors de la perte d'eau des graines (Bailly 2004). Plus encore, des protéines déstructurées peuvent interagir avec des protéines fonctionnelles les inactivant à leur tour.

Les protéines LEA ont été identifiées lors des phases tardives de la maturation (Galau et al. 1986). Leur implication dans la réponse au stress, notamment hydrique est sans équivoque (Battaglia et al. 2008 ; Hanin et al. 2011). Dans la graine, elles participeraient notamment à la stabilisation des phospholipides membranaires, l'atteinte et le maintien de l'état vitreux du cytoplasme ou encore remplaceraient l'eau perdue par leurs groupements polaires (Kalemba & Pukacka 2007 ; Tunnacliffe & Wise 2007). *Amborella* contient plusieurs protéines LEA dans l'embryon (treize identifications) et dans l'albumen (trois identifications) (tableau supplémentaires 1 et 2). Notamment, la protéine rd29B (Q04980), dont l'expression est sous la dépendance de l'ABA suite à un stress de dessiccation, est présente dans l'embryon (Yamaguchi-Shinozaki & Shinozaki 1993) (tableau supplémentaire 1). Par ailleurs, la protéine Em-like protein GEA6 (Q02973), qui est retrouvée dans l'embryon (tableau supplémentaire 1), est exprimée dans les phases tardives de la maturation des graines (Gaubier et al. 1993 ; Manfre et al. 2006).

Les protéines HSPs sont des protéines chaperons induites par de nombreux stress. Elles forment des complexes avec les protéines déstructurées et permettent leur réassemblage correct par utilisation d'ATP (Al-Whaibi 2011). Les petites HSPs (sHSPs) ont quant à elles une action indépendante de l'ATP. Elles forment des complexes avec des protéines partiellement structurées ou déstructurées et empêchent leur dénaturation complète (Friedrich et al. 2004). Des HSPs et des sHSPs sont présentes dans l'embryon (37 identifications) et dans l'albumen (huit identifications).

Deux sHSPs (HSP17.4 et HSP17.6) (P19036 et P29830) (tableau supplémentaire 1), décelables uniquement pendant la phase tardive de la maturation et au début de la germination des graines d'*A. thaliana* (Takahashi & Komeda 1989 ; Sun et al. 2001 ; Kotak et al. 2007) ont été identifiées dans l'embryon d'*Amborella* (tableau supplémentaire 1). HSP18.1 (P19037), décrite chez le pois comme une protéine chaperon conditionnant la stabilité de protéines cruciales du métabolisme, notamment la malate déshydrogénase ou la glycéraldéhyde-3-phosphate déshydrogénase (Stengel et al. 2010), a également été retrouvée dans le protéome de l'embryon d'*Amborella* (tableau supplémentaire 1). De même, HSP26.5 (Q9SSQ8), qui est accumulée pendant la maturation tardive chez le pois et qui serait impliquée dans le maintien de l'intégrité du complexe I mitochondrial (Grelet et al. 2005 ; Stupnikova et al. 2006), a été retrouvée au niveau du protéome de l'embryon d'*Amborella* (tableau supplémentaire 1).

Outre ces petites HSPs, l'embryon et l'albumen d'*Amborella* contiennent d'autres HSPs de plus haut poids moléculaire impliquées dans la structuration des protéines que sont notamment HSP70 (Q9STW6, Q9LKR3, F4HQD4) et HSP101 (P42730) (Su & Li 2008 ; Al-Whaibi 2011). Beaucoup de protéines chaperons sont des protéines de choc thermique (Heat shock proteins, HSP), c'est-à-dire des protéines exprimées en réponse à des variations de température, ou d'autres types de stress cellulaire. Par ailleurs, l'expression de HSP101 (P42730) est sous le contrôle du facteur de transcription HsFA9 pendant la phase tardive de la maturation des graines d'*A. thaliana* (Kotak et al. 2007). Sans stress, l'absence de HSP101 (par invalidation du gène codant cette protéine) n'est pas létale (Hong & Vierling 2001). Sa présence dans l'embryon d'*Amborella* laisse en revanche penser que celui-ci est « prêt » à affronter un stress. CPN60 (P29197) agit lors de la restructuration des complexes (Weiss et al. 2009) associés à son co-chaperon CPN21 (O65282) et ces protéines sont toutes deux retrouvées dans l'embryon d'*Amborella* (tableau supplémentaire 1). Fait très intéressant, CPN21 serait impliquée dans l'activation de la superoxyde dismutase à fer des plastides, une enzyme capitale dans les processus de détoxification cellulaire (Kuo et al. 2013).



Ainsi, notre étude démontre que la graine d'*Amborella* possède un grand nombre de protéines impliquées dans la tolérance à la dessiccation, une caractéristique acquise en fin de maturation des graines orthodoxes. Ces résultats nous conduisent à proposer que l'embryon soit mature au moment de sa dispersion.

En plus de ces protéines de réponse au stress, une grande partie de la machinerie de détoxification des cellules a été identifiée dans l'embryon (21 identifications) et dans l'albumen (six identifications) (tableaux supplémentaires 1 et 2) de la graine d'*Amborella*. Il s'agit notamment d'une catalase (P25819) et de protéines du système de transfert d'électrons peroxisomal dépendant de l'ascorbate telles que l'ascorbate peroxydase (Q1PER6) ou la monodéhydroascorbate réductase (Q9LFA3) (tableaux supplémentaires 1 et 2). Il est établi que lors de la germination de grandes quantités de peroxyde d'hydrogène sont produites, notamment par les péroxisomes lors de la  $\beta$ -oxydation des lipides de réserve (Eastmond 2007). La détoxification de ces composés est cruciale pour protéger les corps lipidiques de leur oxydation mais également des enzymes telles que l'aconitase (une enzyme des cycles des acides tricarboxyliques et du glyoxylate), très sensible au peroxyde d'hydrogène (Verniquet et al. 1991). Outre leurs effets délétères, le peroxyde d'hydrogène et les autres formes activées de l'oxygène jouent également un rôle important comme molécule signal et pour le relâchement des parois (Bailly 2004 ; Oracz et al. 2007).

Ainsi, d'un point de vue biochimique et moléculaire, l'embryon et l'albumen d'*Amborella* possèdent tous les mécanismes nécessaires à l'acquisition de la tolérance au stress de dessiccation se manifestant lors des phases ultimes de la maturation des graines orthodoxes. Ces données sont importantes puisque l'origine d'une telle tolérance au cours de l'évolution est sujette à discussions. En effet, cette tolérance apparaît correspondre à un caractère complexe, nécessitant l'interaction de nombreux facteurs génétiques (Berjak & Pammenter 2008). Une première étude sur 45 espèces avait abouti à la conclusion que les graines récalcitrantes (non tolérantes à la dessiccation) étaient associées à des ovules de type ancestral. De fait, l'orthodoxie fut donc considérée comme correspondant au caractère évolué (von Teichman & van Wyk 1994). Cependant, des études plus récentes portant sur un plus grand nombre d'espèces ont abouti à une conclusion inverse (Farnsworth 2000 ; Dickie & Pritchard 2002) et ce, notamment, pour des espèces de la famille des *Hydatellaceae* (plantes herbacées de Nouvelle-Zélande et de l'ouest de l'Australie, annuelles, à rosette, aquatiques avec des feuilles submergées et émergentes, autogames ou à pollinisation aquatique) qui sont

considérées comme l'une des plus anciennes lignées de plantes à fleurs (Tuckett et al. 2010). Ces observations sont en accord avec la loi de Dollo (un paléontologiste belge, connu pour ses travaux sur les dinosaures et pour son hypothèse sur la nature irréversible de l'évolution), qui suggère que l'apparition répétée d'un caractère complexe au cours de l'évolution est très improbable alors que l'inverse est très facile (Tweddle et al. 2003). De même, la graine la plus ancienne et encore capable de germer a été datée de 1 300 ans chez le lotus sacré (*Nelumbo nucifera*), qui est une espèce relativement basale (Shen-Miller 2002 ; Shen-Miller et al. 2013). Nos observations concernant la détection de protéines HSPs et LEA laissent penser que la graine d'*Amborella* est effectivement tolérante à la dessiccation et que ce trait serait le caractère ancestral. Cependant, des études physiologiques complémentaires, par exemple portant sur la conservation à différentes températures, avec différentes humidités relatives et pendant des temps différents seraient nécessaires afin d'étayer plus avant nos hypothèses.

Notre analyse des protéines totales par protéomique *shotgun* démontre que si l'albumen est un véritable tissu de stockage, l'embryon l'est nettement moins. En effet, 49,6% des protéines de l'embryon sont impliquées dans la catégorie fonctionnelle « *metabolism* ». Plus encore, 37,1% de ses protéines sont spécifiques du métabolisme des acides aminés, et en particulier du métabolisme de la méthionine et des acides aminés branchés. Cette observation remet en cause l'hypothèse longtemps admise que la source principale de carbone et d'azote dans les graines provient de l'utilisation des protéines de réserve.

Le métabolisme de la méthionine est crucial pour tous les organismes vivants. En effet, la méthionine entre non seulement dans la composition des protéines, mais elle supporte des fonctions métaboliques vitales comme la méthylation des protéines et d'une myriade de métabolites. Chez les plantes, la méthionine est également à l'origine de la synthèse d'éthylène, une phytohormone, et de la synthèse de biotine, une vitamine cofacteur de plusieurs carboxylases cellulaires (Ravanel et al. 1998). Il est établi que le métabolisme du soufre joue un rôle prépondérant dans la germination des graines, notamment en lien avec les voies de signalisation via les principales phytohormones (Figure 35 ; Rajjou et al. 2012).

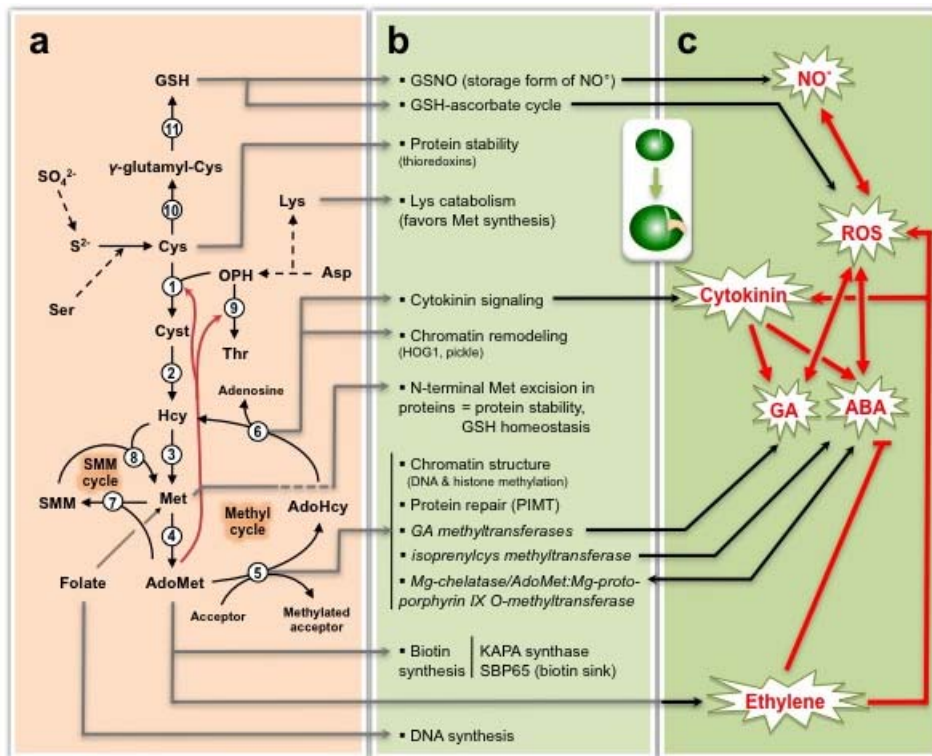


Figure 35 : métabolisme du soufre dans le contrôle de la germination des graines (Rajjou et al. 2012). (a) Métabolisme des acides aminés soufrés. L'*O*-phosphohomoserine (OPH) est un point de branchement entre les voies de synthèse Met/*S*-adénosylmethionine (AdoMet) et Thr/Ileu. AdoHcy, *S*-adénosylhomocystéine ; Cyst, cystathionine ; GSH, glutathion ; Hcy, homocystéine ; SMM, *S*-méthylméthionine. Enzymes : 1, Cyst  $\gamma$ -synthase ; 2, Cyst  $\gamma$ -lyase ; 3, Met synthase ; 4, AdoMet synthétase ; 5, transméthylases dépendant de l'AdoMet ; 6, AdoHcy hydrolase ; 7, AdoMet:Met *S*-méthyltransférase ; 8, SMM:Hcy *S*-méthyltransférase ; 9, Thr synthase ; 10,  $\gamma$ -glutamyl-Cys synthase ; 11, GSH synthase. Les flèches en pointillé indiquent des réactions multiples. Les deux flèches rouges indiquent que l'AdoMet est a) un régulateur allostérique de la Thr synthase (enzyme 9) permettant de contrôler le partage des flux d'OPH entre les voies de biosynthèse de Met and Thr, et b) un régulateur de l'expression du gène codant la Cyst  $\gamma$ -synthase (enzyme 1). (b) Processus régulés. KAPA, acide 7-keto-8-aminopelargonique ; PIMT, L-isoaspartyl-*O*-méthyltransférase. (c) Liens avec les hormones et les molécules de signalisation.

Dans le cadre de ce métabolisme du soufre, il est très intéressant de noter qu'une enzyme nommée protéine-L-isoaspartate-*O*-méthyltransférase (PIMT) a été identifiée uniquement dans l'embryon d'*Amborella* (Q42539) (tableau supplémentaire 1). C'est à notre connaissance la première identification par approche protéomique de cette enzyme chez les

plantes, suggérant qu'elle s'accumule en quantité substantielle dans les graines d'*Amborella*. Cette protéine est d'une importance capitale d'un point de vue de l'intégrité du protéome cellulaire. Elle est en effet présente dans tous les règnes (Kagan et al. 1997 ; Mudgett et al. 1997 ; Thapar et al. 2001) et son rôle se situe plus précisément au niveau du vieillissement, plus particulièrement dans la réparation des dommages subis par les protéines lors de ce vieillissement. L'isomérisation des résidus protéiques L-asparagine et L-aspartate en leurs formes L-iso-Asp ou D-Asp se manifeste spontanément au cours du vieillissement et est augmentée en cas de stress. Elle aboutit à une altération des protéines impactant ainsi leurs rôles et activités biologiques. La PIMT est une méthyltransférase capable de catalyser la conversion des résidus anormaux L-iso-Asp en leur forme non délétère L-Asp, en utilisant de la *S*-adénosyl-méthionine comme co-substrat (Figure 36) (Ogé et al. 2008).

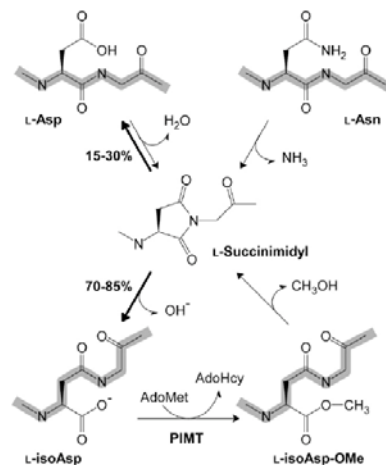


Figure 36 : mécanisme spontané de formation de L-Isoaspartyl et action enzymatique de réparation de la PIMT (Ogé et al. 2008). La déamination de l'asparagine (haut droit) et l'isomérisation de l'aspartate (haut gauche) provoquent la formation d'un cycle succinimidyle instable qui est hydrolysé spontanément pour donner un mélange d'Asp (15 à 30%) et isoAsp anormal (70 à 85%) aux pH physiologiques. Le squelette du peptide est souligné en gris pour illustrer l'effet de la présence d'un résidu isoAsp sur la structure normalement linéaire. La PIMT catalyse la première étape de la conversion de l'isoAsp en Asp. La forme L-isoAsp-OMe aboutit spontanément à la formation d'un résidu Asp (15 à 30%) et isoAsp (70 à 85%). De fait, plusieurs cycles d'action enzymatique sont nécessaires pour la réparation complète des protéines. AdoHcy : *S*-adenosyl homocystéine, L-Asp : L-aspartate, L-Asn : L-asparagine, L-isoAsp : L-isoaspartate, L-isoAsp-OMe : L-isoaspartate-*O*-méthylester.

Il a été démontré que cette enzyme était nécessaire au bon fonctionnement du système nerveux central de la souris où des protéines substrats ont été identifiées (Kim et al. 1997 ; Vigneswara et al. 2006 ; Zhu et al. 2006). Elle influe également la tolérance au stress chez le nématode *Caenorhabditis elegans* (Khare et al. 2011) et explique que la bactérie thermophile *Thermotoga maritima* supporte des températures de 100°C (Ichikawa & Clarke 1998). Elle est également suspectée jouer un rôle prépondérant dans les maladies neurodégénératives chez l'homme (Desrosiers & Fanélus 2011). Chez les plantes, l'enzyme est directement liée à la conservation des graines (après dispersion et/ou récolte) et leur vigueur germinative (Ogé et al. 2008). Ainsi, la surexpression de l'enzyme chez les graines d'*A. thaliana* augmente leur vigueur (viabilité), alors que l'inverse est observé par sous-expression (Ogé et al. 2008). Il a d'ailleurs été proposé que la longévité exceptionnelle de la graine de lotus sacré (*Nelumbo nucifera*) serait en partie liée à l'accumulation extraordinaire de la PIMT au sein de ses tissus, puisqu'il s'agit de la plus forte accumulation observée dans le règne vivant (Shen-Miller et al. 1995 ; Shen-Miller et al. 2013). L'identification de la PIMT dans l'embryon d'*Amborella* suggère ainsi que la graine est capable d'assurer la réparation de son protéome endommagé lors du stockage à sec, qui induit le vieillissement des graines chez de nombreuses espèces (Job et al. 2005 ; Rajjou et al. 2008, 2012).

Par ailleurs, 17 substrats protéiques de la PIMT sont connus chez *A. thaliana* (Chen et al. 2010). L'un d'entre eux est présent dans le protéome d'*Amborella*, correspondant à une sHSP (HSP17.4) (P19036 – AT3G46230). Quatre autres sont similaires à ceux identifiés chez *A. thaliana*. Il s'agit chez *Amborella* de la xylose synthase (AT1G08200), de la protéine Mother of FT and TFL1 (AT1G18100), de protéines ribosomales 60S (multiples identifiants) et de la cinnamoyl-CoA réductase (AT5G58490). Toutes ces protéines pourraient donc être des substrats potentiels de la PIMT d'*Amborella*.

Il est particulièrement intéressant de noter que la PIMT, dont l'expression peut ou non être influencée par l'ABA, est codée par deux gènes chez *A. thaliana*. Une forme est cytosolique (PIMT1), s'accumulant spécifiquement dans les graines et une deuxième forme dont l'ARNm est sujet à modifications post-transcriptionnelles est nucléaire (PIMT2), s'accumulant dans les graines mais également d'autres tissus (eg., feuilles) (Xu et al. 2004). D'autres plantes, par exemple le pois chiche, expriment également deux formes de la PIMT (Verma et al. 2013). Une recherche par BLAST contre le génome d'*Amborella* nous a montré qu'il n'existe qu'une forme de PIMT chez cette espèce ([www.amborella.org](http://www.amborella.org)) (scaffold00166.14) (tableau supplémentaire 1). Des analyses de séquences par BLAST

(<http://blast.ncbi.nlm.nih.gov>) indiquent que le gène d'*Amborella* coderait une protéine plus proche de la PIMT1 (E value de  $2^e-118$ ) que de la PIMT2 (E value de  $2^e-113$ ) d'*A. thaliana*.

Par ailleurs, une approche préliminaire par comparaison de la séquence de la PIMT d'*Amborella* contre toutes les séquences de PIMT contenues dans les bases de données chez les plantes à graines place cette protéine proche des gymnospermes, ce qui suggère la conservation d'un caractère ancestral, vital pour la conservation des espèces (Figure 37). La conservation d'un gène codant une enzyme de type PIMT1, plutôt que de type PIMT2 (par référence à *A. thaliana*) souligne l'importance de la survie des plantules résultant de la germination chez les plantes basales.

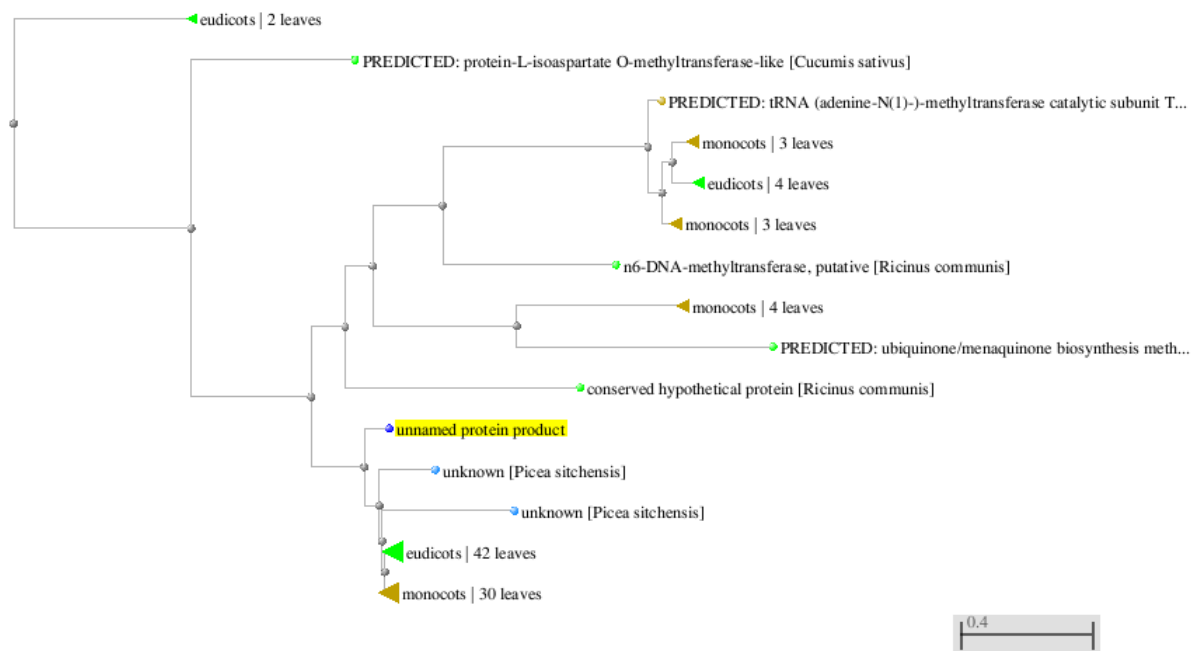


Figure 37 : arbre phylogénétique obtenu pour le gène unique d'*Amborella* codant la protéine isoaspartyl méthyltransférase (PIMT, unnamed protein product surligné en jaune), enzyme impliquée dans le vieillissement et la survie des graines. Cette analyse a été effectuée à l'aide du logiciel disponible sur le site <http://www.ncbi.nlm.nih.gov/gate1.inist.fr/tools/cobalt/cobalt.cgi> considérant toutes les séquences de PIMT répertoriées chez les plantes à graines.

De concert avec la biosynthèse des acides aminés, la synthèse *de novo* des protéines semble également en bon accord avec des observations chez la graine de betterave à sucre (Catusse et al. 2008) et d'*A. thaliana* (Fait et al. 2006). L'accumulation de protéines impliquées dans cette fonction dans la graine mature sèche d'*Amborella* suggère l'importance de la synthèse protéique *de novo* lors de l'imbibition, une caractéristique qui a été bien étayée

chez *A. thaliana* (Rajjou et al. 2004), la betterave à sucre (Catusse, Strub, et al. 2008) et le riz (Deng et al. 2013). Dans notre cas, une telle importance de la synthèse protéique *de novo* est également corroborée par la présence d'une grande quantité de protéines ribosomales qui représentent 7,1% des protéines de l'embryon (tableau supplémentaire 1). De plus, l'embryon contient de nombreux facteurs de traduction, tels que eEF-1 $\alpha$ 4, eEF-1B $\gamma$ 1 et eEF-1B $\beta$ 2 (tableau supplémentaire 1). Ainsi, la synthèse *de novo* des protéines serait l'une des fonctions importantes mises en œuvre dans l'embryon d'*Amborella* lors de l'imbibition de la graine mature sèche.

Par ailleurs, l'embryon possède 7,3% de protéines impliquées dans le métabolisme des sucres et polysaccharides (tableau supplémentaire 1). On note l'absence d'amylase ce qui concorde avec l'observation que la graine ne contient quasiment pas d'amidon (Floyd & Friedman 2001). La majorité d'entre elles sont impliquées dans la synthèse ou le remodelage de la paroi cellulaire (alpha-mannosidase, UDP-glucose pyrophosphorylase, glucan endo 1,3 beta-glucosidase, UDP-glucose 4 épimérase) (P94078, P57751, Q9M088, Q9T0A7) (tableau supplémentaire 1). La présence de ces protéines dans l'embryon d'*Amborella* n'est pas surprenante. En effet, comme cité et observé précédemment (Figure 20), l'embryon semble en plein état de différenciation. En revanche, cette présence pourrait contraster avec l'idée que la germination est associée à une déstructuration des parois cellulaires pour permettre la poussée et la sortie radicaire (Ogawa et al. 2003 ; Müller et al. 2009).

La fonction « **lipid and sterol** » représente 1,8% des protéines identifiées à partir de l'embryon. Il est intéressant de noter qu'ont été identifiées deux lipases (enzymes clés de la mobilisation des lipides) (O65660, Q9LY84) (tableau supplémentaire 1) présentes uniquement dans l'embryon. Ainsi, il semble que du point de vue de l'utilisation des réserves lipidiques, l'albumen n'est pas autonome et requiert les enzymes présentes dans l'embryon. Pour tester cette hypothèse, il conviendra d'effectuer une analyse plus approfondie de la fraction albumines de l'albumen. Par contre, nos résultats suggèrent à nouveau l'importance de la mobilisation des lipides de réserve au niveau de l'embryon, probablement en vue de l'établissement d'une transition métabolique favorable à la germination.

L'embryon d'*Amborella* possède, pour assurer son développement et la préparation à l'installation d'une plantule lors de l'imbibition de la graine mature sèche, de nombreuses protéines impliquées dans diverses voies métaboliques permettant notamment l'utilisation des réserves ainsi que la néo-synthèse protéique (tableau supplémentaire 1). Cependant, des sources énergétiques sont nécessaires pour la mise en place de ces transitions développementales. Suite à l'imbibition, la graine doit en effet subvenir très rapidement aux

besoins énergétiques de l'embryon, à l'origine de la plantule. L'autotrophie n'est pas encore en place, mais les réserves accumulées pendant la maturation de la graine doivent assurer le développement hétérotrophe. Nous l'avons vu précédemment, *Amborella* possède les protéines permettant la mobilisation puis l'utilisation des acides gras libérés par les corps lipidiques alimentant ainsi le cycle du glyoxylate (tableau supplémentaire 1). Ainsi, la catégorie « **energy** » contient 14,8% des protéines de l'embryon. Notamment, les enzymes de la glycolyse représentent à elles seules 6,3% des protéines, soit près de la moitié des protéines de la catégorie. Il est remarquable de constater qu'une seule enzyme de la glycolyse n'a pas été identifiée dans l'embryon (c'est-à-dire la phosphoglycérate mutase) (tableau supplémentaire 1). De manière intéressante, cette même enzyme est absente lors du développement des graines de blé et de ricin (Plaxton 1996). Sa synthèse plus tardive lors de la germination pourrait ainsi représenter un facteur de régulation de la glycolyse. De même, il ne manque que deux enzymes pour le cycle des acides tricarboxyliques (succinate déshydrogénase et fumarase). Ces enzymes sont également absentes de l'albumen.

L'embryon possède des protéines impliquées dans la catégorie « **secondary metabolism** ». Il s'agit notamment d'une isochorismatase (Q93Z51) (tableau supplémentaire 1) qui hydrolyse l'isochorismate (précurseur de l'acide salicylique). L'identification de cette enzyme pourrait donc suggérer une inhibition potentielle de cette voie de biosynthèse chez l'embryon d'*Amborella*. Par ailleurs, l'identification d'une tocophérol *O*-méthyltransférase (Q9ZSK1) (tableau supplémentaire 1) révèle l'aptitude de la graine à utiliser les acides gras présents pour synthétiser une puissante molécule piégeant les espèces oxydantes (ROS) : l' $\alpha$ -tocophérol (Shintani & DellaPenna 1998 ; Bailly 2004).

Une représentation en diagramme de Venn des protéines totales de l'embryon et de l'albumen nous a permis de mettre en évidence les spécificités des profils protéiques de chacun des deux compartiments de la graine (Figure 22). Ainsi, la catégorie « **synthèse protéique** » est spécifique de l'embryon et concerne principalement les protéines ribosomales comme nous l'avons vu précédemment. Les protéines de la catégorie « **croissance cellulaire et division** » sont également spécifiques de l'embryon. En particulier la protéine Mother of FT and TFL1 (Q9XFK7) (annexe 5), un homologue de FLOWERING LOCUS (FT) et TERMINAL FLOWER correspondant à deux régulateurs jouant des rôles opposés dans le contrôle de la floraison (Bradley et al. 1997 ; Kobayashi et al. 1999), a été identifiée dans l'embryon d'*Amborella* (tableau supplémentaire 1). Des données récentes indiquent clairement que cette protéine joue un rôle crucial dans la germination en exerçant un rétrocontrôle négatif sur la signalisation par l'ABA (un inhibiteur de germination) (Xi et al.



2010). La fonction de cette protéine semble très conservée chez les plantes, y compris chez les bryophytes, ancêtres de toutes les plantes terrestres (Hedman et al. 2009). Nos résultats chez *Amborella* sont donc en parfait accord avec ces données. Ainsi, l'identification de cette protéine indiquerait que les mécanismes de la régulation de la maturation/germination contrôlés par le ratio ABA/GA<sub>3</sub> sont présents chez *Amborella*. Des résultats préliminaires permettent de construire un arbre phylogénétique à partir de cette protéine d'*Amborella* (Figure 38). Cette analyse place la séquence d'*Amborella* proche de celles répertoriées chez les gymnospermes, indiquant ainsi la conservation de caractères ancestraux au niveau des mécanismes régulant la germination chez *Amborella*.

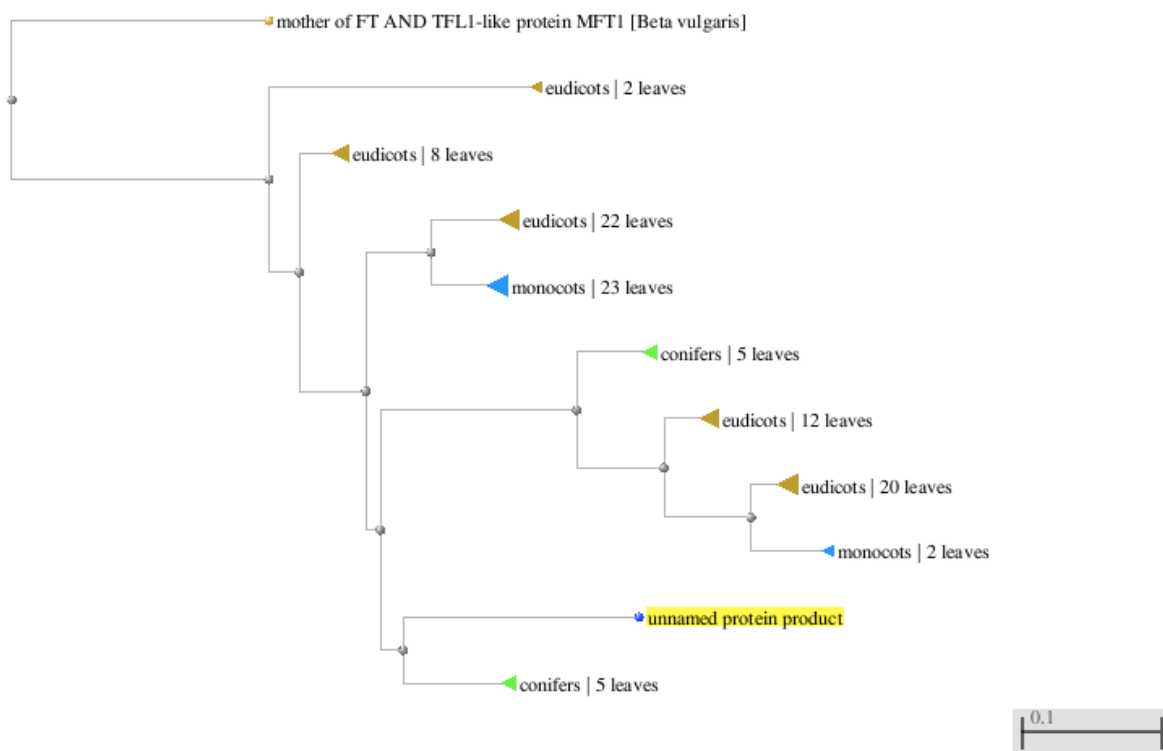


Figure 38 : arbre phylogénétique obtenu pour le gène d'*Amborella* codant la protéine Mother of FT and TFL1 (unnamed protein product surligné en jaune), protéine impliquée dans le contrôle de la floraison mais aussi dans celui de la germination des graines. Cette analyse a été effectuée à l'aide du logiciel disponible sur le site <http://www.ncbi.nlm.nih.gov/gate1.inist.fr/tools/cobalt/cobalt.cgi> considérant toutes les séquences de PIMT répertoriées chez les plantes à graines.

### **2.2.1.3 Caractérisation des fractions albumines et globulines de la graine mature sèche entière d'*Amborella***

Du fait de l'apparente pauvreté en classes ontologiques des protéines totales autres que les protéines de réserve de l'albumen en comparaison de celles de l'embryon, nous avons cherché à augmenter la finesse de représentation des protéines albumines, tout d'abord dans la graine entière, en procédant à une séparation des albumines et des globulines selon Osborne (1924). Ce protocole de solubilisation sélective des différentes classes des protéines des graines d'*Amborella* permettait donc de s'affranchir des protéines globulines majoritaires dans la fraction albumines, tout en obtenant une fraction globulines débarrassée des albumines. Les fractions albumines et globulines ainsi obtenues ont alors été soumises à électrophorèse 2D et identification des spots visibles par LC-MS/MS.

Cette approche révèle que la catégorie « *protein destination and storage* » de la fraction globulines contient à elle seule 94% des protéines extraites (Figure 33). Cette extraction a donc permis d'obtenir majoritairement les protéines de la fraction globulines (protéines de réserve). D'autres protéines sont cependant présentes dans la fraction globulines. Il s'agit notamment de protéines de la catégorie « *disease/defence* » (2,6% des protéines). On note ainsi l'identification d'une thaumatine (Q9ZV34) (tableau supplémentaire 4). Ces protéines sont souvent associées à une infection de la plante par un pathogène (Wang et al. 2010). Sa présence pourrait donc indiquer que la graine se prépare à devoir affronter des stress biotiques lors de la germination. Deux autres protéines ont été identifiées dans la fraction globulines uniquement. Il s'agit de la thiamine thiazole synthase (Q38814) (tableau supplémentaire 4). Cette enzyme est impliquée dans la biosynthèse de la thiamine (vitamine B1), essentielle pour le métabolisme, notamment dans le catabolisme des glucides, la synthèse d'ATP, de NADPH et la formation des acides nucléiques. L'activation de sa voie de biosynthèse est également associée à de nombreux stress, notamment la sécheresse (Rapala-Kozik et al. 2012). La seconde protéine retrouvée uniquement dans la fraction globulines est une tripeptidyl peptidase (F4JVN6). Sa présence dans la fraction globulines pourrait suggérer une interaction forte avec les protéines de réserve qu'elle devra dégrader et/ou une structure particulière expliquant sa solubilisation dans les sels. Les autres protéines identifiées dans la fraction globulines se retrouvent majoritairement dans la fraction albumines. Leur présence parmi les globulines pourrait indiquer une contamination au moment de l'extraction. Une autre explication serait que la structure de ces protéines permet de les solubiliser dans les deux conditions.

Nous pouvons conclure que l'application du protocole d'extraction fractionnée des protéines de la graine (Osborne 1924) a donc permis d'enrichir la fraction des globulines en protéines de réserves (Figure 33).

Comme nous l'avons vu, le compartiment majoritaire de la graine est l'albumen (Figure 20) qui est le principal tissu de stockage majoritaire composé de protéines de réserve (Figure 20 et Figure 21). Afin de s'affranchir de la forte quantité de protéines de réserve dans la graine entière, la fraction des albumines de la graine entière a été extraite selon le protocole d'Osborne (1924), c'est-à-dire dans l'eau (cf. le chapitre « Matériels et Méthodes »). Les résultats obtenus montrent que l'objectif d'enrichissement de la fraction des albumines a été parfaitement atteint. En effet, 561 protéines ont été identifiées (tableau supplémentaire 3) qui correspondent à 31 fonctions différentes (Figure 24).

De manière surprenante, la catégorie la plus importante est « *protein destination and storage* » qui contient 35,9% des protéines identifiées. Notamment, la fonction « *storage protein* » contient 23,7% des protéines, représentée majoritairement par une protéine à motif cupin (nutrient réservoir protein – Q9LUJ7) (tableau supplémentaire 3). Il convient d'observer que l'analyse des spots contenant cette protéine montre une très grande diversité de poids moléculaire et de point isoélectrique (de 5 à 7,5 pour le pI et de 95 à 27 kDa pour le poids moléculaire) (carte de référence en annexe 2). Notre protocole d'extraction des albumines, impliquant notamment de nombreux lavages de cette fraction protéique préalablement à l'analyse par protéomique 2D, n'est pas en faveur de l'hypothèse d'une contamination des albumines par la fraction globulines. Nous proposons au contraire qu'il s'agirait de fragments des différentes formes des globulines rendus solubles dans l'eau suite à des protéolyses discrètes de ces différentes formes survenant durant la maturation des graines d'*Amborella*. Ce phénomène a été très largement décrit chez les graines matures sèches de betterave à sucre (Job et al. 1997 ; Catusse, Strub, et al. 2008) et d'*A. thaliana* (Gallardo et al. 2001). Ainsi, chez toutes les graines, y compris celles d'*Amborella*, une pré-mobilisation des globulines se manifesterait dès la maturation, ainsi préparant et facilitant une hydrolyse plus complète de ces réserves lors de la germination et de l'installation de la plantule. Outre les protéines de réserve dans cette catégorie, nous pouvons noter la présence de nombreuses HSP (notamment HSP 70 et CPN60) (tableau supplémentaire 3) tout comme pour l'approche en protéomique *shotgun*. Les deux types d'approches protéomiques ont par ailleurs conduit à l'identification de sHSPs (HSP17.4 et HSP18.1) (tableau supplémentaire 3).

La catégorie « **energy** » contient, tout comme observé lors de l'approche protéomique *shotgun*, de nombreuses enzymes de la glycolyse (17,8%) et du cycle des acides tricarboxyliques (3,2%). Il est intéressant de noter que la phosphoglycérate mutase (Q9M9K1) absente par l'approche *shotgun* est retrouvée ici (tableau supplémentaire 3). Toutes les enzymes de la glycolyse ont donc été finalement identifiées dans la graine. Ce résultat montre la complémentarité des deux méthodes d'approche protéomique.

La catégorie « **metabolism** » représente là aussi une part importante des identifications (11,7%) (Figure 24). Tout comme dans le cas de l'approche par protéomique *shotgun*, il y a une forte représentation des enzymes des voies de biosynthèse des acides aminés, y compris de la méthionine (tableau supplémentaire 3). Nous notons également des enzymes impliquées dans la formation d'intermédiaires de biosynthèse des parois ou encore dans l'utilisation des lipides.

De manière corollaire à l'importance du métabolisme des acides aminés, nous pouvons noter la présence de protéines et enzymes impliquées dans la traduction des protéines. (Figure 24). Il est très intéressant d'observer dans ce contexte qu'une proportion notable des protéines identifiées (5,2%) correspond à une enzyme impliquée dans la mise en place de la coiffe des ARN messagers, la *S*-adenosyl-L-méthionine-dépendent-méthyltransférase (Q9S9P4) (tableau supplémentaire 3). Cette étape est cruciale puisqu'en son absence, les ARNm ne peuvent pas être traduits (Kapp & Lorsch 2004). Ces résultats sont en parfait accord avec ceux obtenus par l'approche de protéomique *shotgun* qui montrent l'importance de la synthèse *de novo* dans la graine d'*Amborella*.

La catégorie « **cell structure** » contient de nombreuses enzymes impliquées dans la biosynthèse des composés des parois cellulaires, tout comme celles que nous avons identifiées par l'approche protéomique *shotgun*. La catégorie « **cell growth/division – cell cycle** » contient notamment une monooxygénase (Q9SU40) également identifiée par l'approche protéomique *shotgun*. Cette enzyme est impliquée dans la croissance en intervenant sur le relâchement des parois (Sedbrook et al. 2002). Sa présence est donc révélatrice des processus de « croissance interne » de l'embryon.

Aucun transporteur n'a été identifié. En revanche, on note la présence d'un « vacuolar sorting receptor » (P93026) (tableau supplémentaire 3) retrouvé uniquement par l'approche 2D-PAGE. Cette protéine assure le transport des globulines du réticulum endoplasmique aux vacuoles de stockage en se fixant à la partie C-terminale des pro-globulines de manière calcium dépendante (Shimada et al. 2003). Son identification indique donc que les processus

de transport/maturation des globulines chez *A. thaliana* se retrouvent également chez *Amborella*.

L'obtention de la fraction albumines de la graine entière mature sèche d'*Amborella* a permis un net enrichissement en protéines du métabolisme. La signature ontologique (Bevan et al. 1998) de cette fraction est similaire à celle du protéome total de l'embryon obtenue avec l'approche protéomique *shotgun* à l'exception de la forte représentation de protéines de réserve dont l'origine est discutée plus haut (Figure 21 et Figure 24). Les classes majoritairement identifiées confirment l'aptitude de l'embryon à utiliser ses ressources énergétiques afin d'alimenter le fonctionnement de son métabolisme et, notamment, la synthèse *de novo* des protéines et la biosynthèse des composés des parois cellulaires.

En conclusion, notre caractérisation des protéines de la graine mature sèche d'*Amborella* a permis, pour la première fois, d'obtenir le protéome d'un embryon indifférencié dans une graine mature. Nos données indiquent que, malgré son apparence rudimentaire, cet embryon renferme des protéines habituellement associées à un stade de développement avancé chez les graines orthodoxes notamment i) l'aptitude à utiliser les réserves stockées mais également ii) l'aptitude à tolérer la dessiccation (capacité apparaissant tardivement lors de la maturation des graines orthodoxes).

Nos données indiquent également une compartimentation bien marquée entre l'embryon et l'albumen. S'agissant des réserves protéiques, l'embryon apparaît comme une structure quasiment dépourvue de telles réserves et semble contraint d'utiliser les ressources stockées dans l'albumen.

La caractérisation des protéines de la fraction globulines nous a permis de réaliser une étude protéogénomique sur la base des séquences des globulines 11S, permettant ainsi de i) valider l'annotation du tout récent génome d'*Amborella*, mais également de ii) mettre en évidence des caractères ancestraux et évolués des globulines 11S d'*Amborella*. L'approche protéomique menée sur la graine mature sèche d'*Amborella* a ainsi permis de mettre en évidence les trois gènes codant les globulines 11S chez cette espèce, annotés respectivement 67.7, 67.11 et 67.12. De même que chez de nombreuses espèces, ce type de globulines est la forme la plus répandue de protéines de réserve chez les graines d'*Amborella*. Concernant ces protéines, nos données ont démontré :

- la conservation des cystéines impliquées dans la formation du pont disulfure liant les sous-unités A et B des globulines 11S codées par les trois gènes,

- la conservation du motif consensus de coupure entre les deux sous-unités A et B pour deux des trois gènes (67.11 et 67.12),
- la présence d'un motif de coupure de type gymnosperme pour le gène 67.7, témoin du caractère ancestral de ce gène,
- l'absence de l'intron IV caractéristique des gymnospermes, témoignant de l'évolution des trois gènes d'*Amborella*,
- l'absence d'une région de 52 acides aminés importante pour la formation de l'hexamère au sein des corps protéiques et qui serait apparue après la divergence entre *Amborella* et les autres angiospermes.

La suite de cette étude sur la graine d'*Amborella* a pour but d'étudier l'évolution des deux fractions protéiques, albumines et globulines, au cours de l'imbibition/germination afin de mieux appréhender les mécanismes contrôlant la « croissance interne » de l'embryon.

### 3. Protéome de la graine d'*Amborella* au cours de l'imbibition et de la germination

#### 3.1 Introduction

Les résultats obtenus dans le chapitre précédent montrent que la graine mature sèche d'*Amborella* est constituée d'un très petit embryon, contenant très peu de réserves protéiques, au sein d'un albumen volumineux riche en lipides et protéines. Les données de Fogliani et al. (2013) indiquent une croissance interne de l'embryon suite à l'imbibition des graines matures sèches d'*Amborella* (Fogliani et al. 2013 ; Figure 18). Afin de caractériser les modifications métaboliques et moléculaires associées à une telle croissance, nous avons étudié par analyse protéomique les étapes de l'imbibition/germination de la graine mature sèche d'*Amborella*.

#### 3.2 Résultats

Des graines d'*Amborella* ont été mises à germer en boîtes de Petri en présence d'eau distillée (Figure 39) (cf. « Matériels et Méthodes »).

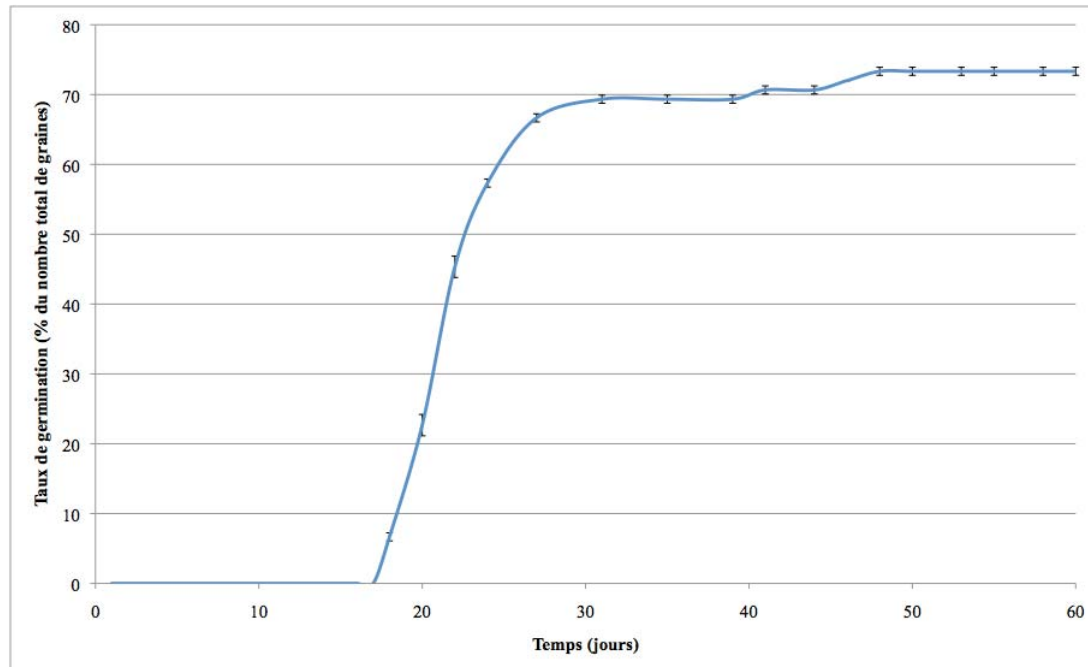


Figure 39 : germination *sensu stricto* des graines matures sèches d'*Amborella*, dans l'eau, à 25°C et lumière continue.

Le temps de latence du lot utilisé est de 17 jours, 73% des graines ont germé en 48 jours. Dans cette description, nous utilisons les termes de « germination *sensu stricto* » pour désigner les graines imbibées pour lesquelles une partie de l'embryon (ici la radicule) émerge au travers des enveloppes, en accord avec la définition communément admise de la germination (Côme 1982 ; Bewley & Black 1994).

Afin de caractériser les processus biochimiques et moléculaires contrôlant le développement interne de l'embryon (Fogliani et al. 2013), nous avons défini trois stades d'étude pour notre analyse protéomique (Figure 40).

- La graine mature, désignée par GMS (graine mature sèche) ou MS pour mature seed,
- la graine germée correspondant au stade d'atteinte de la germination *sensu stricto* (GSS),
- un stade intermédiaire (¼, quart ou quarter), correspondant à une croissance interne de l'embryon représentant de l'ordre d'un quart du volume total de la graine mature sèche.



Figure 40 : coupes longitudinales des stades graine mature (A), ¼ (B) et graine germée (C). L'embryon est matérialisé par le cercle rouge.

### **3.2.1 Etude du protéome de la fraction albumines extraite de la graine d'Amborella au cours de l'imbibition et de la germination *sensu stricto***

Les fractions albumines des trois stades de développement, GMS, ¼ et GSS, ont été extraites. Les extraits ont ensuite été soumis à une migration bidimensionnelle puis les gels obtenus ont été analysés par le logiciel d'acquisition et d'analyse d'images SameSpot (Progenesis) (cf. « Matériels et Méthodes »). Il a ainsi été possible d'identifier les spots significativement reproductibles et dont le volume varie au cours de l'imbibition (stade ¼) et de la germination *sensu stricto*. Des exemples de l'évolution du volume de tels spots sont présentés dans la Figure 41.



Parmi les 296 spots détectés dans la fraction albumines de la graine mature sèche, le volume de 125 d'entre eux varie significativement considérant un ratio minimum de 1,2 des volumes par comparaison avec la graine mature sèche et une p value < 0,02. Cent six spots répondant à ces critères ont conduit à une identification de protéines par spectrométrie de masse (tableau supplémentaire 6).

Les protéines ainsi identifiées (519) ont été classées en quatre groupes en fonction de l'évolution de leur profil d'accumulation au cours de la germination (Figure 42 et Figure 43) :

- **groupe I** : protéines dont l'accumulation est progressive de la GMS à la GSS (313 protéines),
- **groupe II** : protéines s'accumulant uniquement au stade GSS (58 protéines),
- **groupe III** : protéines s'accumulant uniquement au stade ¼ (35 protéines),
- **groupe IV** : protéines dont la quantité diminue progressivement de la GMS à la GSS (113 protéines).

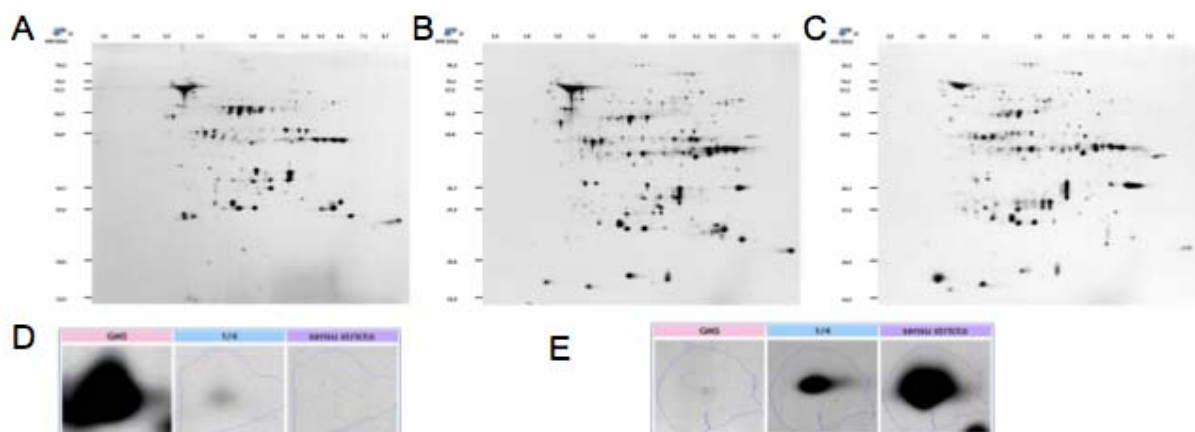


Figure 41 : gels d'électrophorèse bidimensionnelle colorés à l'argent des fractions albumines de la graine mature sèche d'*Amborella* (A), du stade ¼ (B) et du stade graine germée (GSS) (C) (295 protéines détectées). Cent µg d'extrait protéique ont été déposés sur chaque gel. D et E représentent des spots de protéines dont le volume diminue (D) ou augmente (E) au cours de l'imbibition (stade ¼) et de la germination *sensu stricto*. (D) identifiant SameSpot No. 877 (tableau supplémentaire 6) contenant une protéine de réserve à motif cupin. (E) identifiant SameSpot No. 1066 (tableau supplémentaire 6) contenant majoritairement une protéine Major Latex Protein-like.

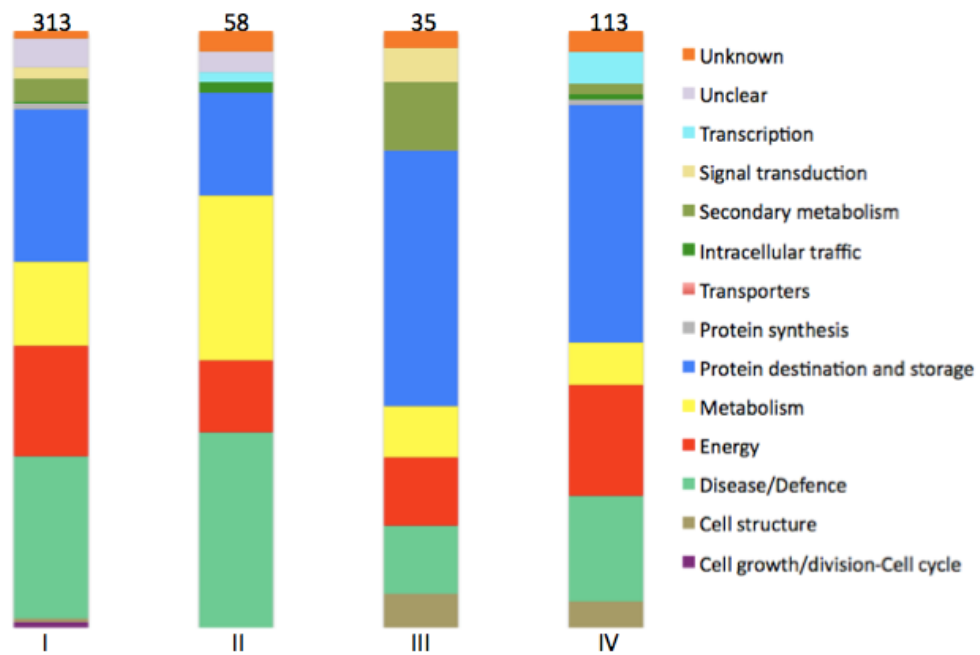


Figure 42 : classes ontologiques (Bevan et al. 1998) des quatre groupes de protéines de la fraction albumines de la graine d'*Amborella* dont l'abondance varie pendant l'imbibition et la germination. I : protéines s'accumulent progressivement de la graine mature sèche (GMS) à la germination *sensu stricto* (GSS). II : protéines s'accumulent uniquement à la GSS. III : protéines s'accumulent uniquement au stade ¼ (stade interne de la croissance de l'embryon cf. Figure 24). IV : protéines dont la quantité diminue progressivement entre les stades GMS et GSS.

Trente cinq protéines sont spécifiquement accumulées au stade ¼ et représentent 6,7% des protéines totales identifiées (Figure 27). Il s'agit notamment des protéines de la catégorie « *protein destination and storage* » dont 2,9% sont spécifiquement accumulées à ce stade, avec 2,3% des protéines dans la fonction « *storage protein* ». Les protéines de la catégorie « *energy* » accumulées au stade ¼ représentent 0,4% des protéines identifiées, et se décomposent en parts égales entre les fonctions « *pentose phosphate* » et « *TCA pathway* » (0,2% chacune). Les protéines impliquées dans la catégorie « *secondary metabolism* » représentent 0,8% des protéines identifiées, avec notamment les fonctions « *terpenoids* » et « *phenylpropanoids/phenolics* » (0,6% et 0,2% des protéines identifiées respectivement). Cinquante huit protéines sont spécifiquement accumulées au stade GSS et représentent 11,2% des protéines identifiées (Figure 27). Notamment, 3,7% des protéines identifiées spécifiquement accumulées à ce stade concernent la catégorie « *disease/defence* », avec 2,7% des protéines dans la fonction « *detoxification* ». La catégorie « *metabolism* » spécifiquement

accumulée au stade GSS contient 3,1% des protéines identifiées, 1,2% d'entre elles étant impliquées dans la fonction « *lipid and sterol* », 1% dans la fonction « *amino acid* » et 0,4% dans les fonctions « *nucleotides* » et « *sugars and polysaccharides* ».

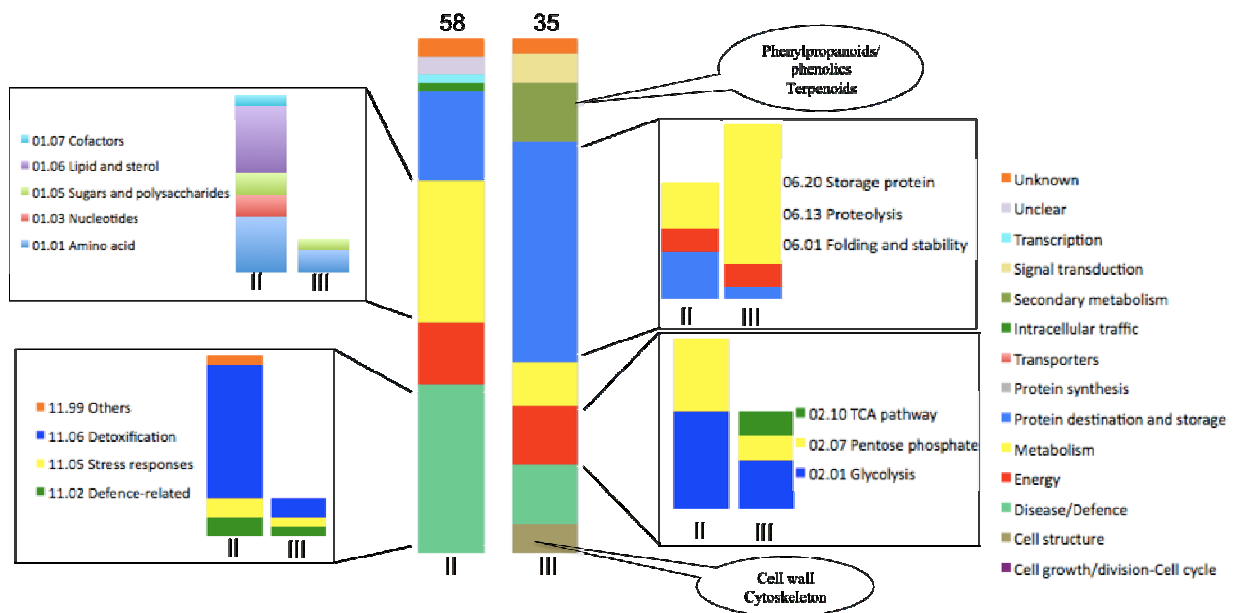


Figure 43 : classes ontologiques et fonctions des protéines de la fraction albumines des graines d'*Amborella* spécifiquement accumulées au moment de la percée radriculaire GSS (II) et du stade 1/4 (III) définis dans la Figure 40.

### 3.2.2 Evolution des globulines de la graine d'*Amborella* au cours de l'imbibition et de la germination *sensu stricto*

Les fractions globulines des trois stades sélectionnés ont été extraites et analysées par électrophorèse monodimensionnelle (Figure 44).

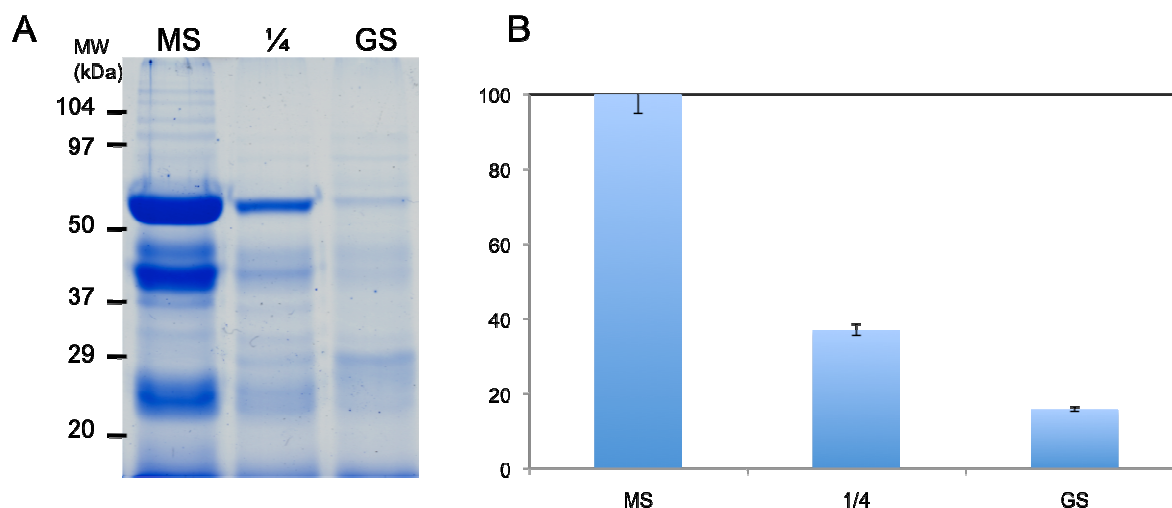


Figure 44 : (A) gel d'électrophorèse monodimensionnelle coloré au bleu colloïdal correspondant à la fraction globulines extraite d'une quantité fixe de graines (1/10ème de graine) pour les stades graine mature sèche (GMS, MS), 1/4 et germination *sensu stricto* (GSS, GS). (B) Abondances relatives des globulines estimées par test ELISA (anticorps reconnaissant la sous-unité B des globulines 11S d'*Amborella* ; cf. « Matériels et Méthodes ») pour les stades GMS (MS), 1/4 et GSS (GS).

Chaque dépôt correspond aux globulines extraites à partir d'une quantité fixe de graines (un dixième de graine), aux trois stades définis (GMS, 1/4, GSS). Une nette diminution de la quantité des globulines est observable (Figure 44A) pendant la croissance interne de l'embryon (stade 1/4), avant la germination *sensu stricto*. Cette observation est confirmée et quantifiée par les résultats de tests ELISA (Figure 44 B) mettant en oeuvre un anticorps spécifique reconnaissant la sous-unité B des globulines 11S d'*Amborella*, préparé ainsi que décrit dans le chapitre Matériels et Méthodes. Il apparaît que la quantité de globuline 11S chute considérablement, de plus de 60% dès le stade 1/4 et de plus de 80% au stade GSS, en parfait accord avec l'analyse électrophorétique de la Figure 44 A. Ainsi, nous pouvons conclure qu'il y a une forte mobilisation des protéines de réserve (globulines 11S) avant même la fin de la germination *sensu stricto* (émergence de la racicule au travers l'enveloppe).

Afin de clarifier la question de l'autonomie de l'albumen dans cette hydrolyse des protéines de réserve lors de l'imbibition, des graines d'*Amborella* dépourvues d'embryon ont été mises à imbiber en présence d'eau. Un temps d'imbibition de 17 jours a été choisi car correspondant au temps nécessaire à la première percée radiculaire lors des germinations des graines entières (Figure 39). Les fractions globulines des différentes conditions ont été extraites et analysées par électrophorèse monodimensionnelle (Figure 45).

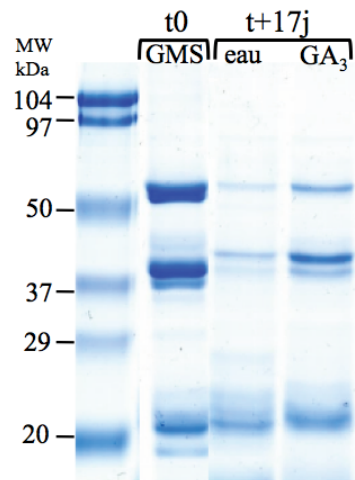


Figure 45 : gel d'électrophorèse monodimensionnelle coloré au bleu colloïdal correspondant aux globulines extraites d'albumen isolé de graines matures sèches (t0), ou d'albumen isolé des graines matures sèches puis imbibé 17 jours avec de l'eau (eau) ou une solution de gibbérellines (GA<sub>3</sub>) et une quantité de tissu correspondant à un dixième de graine.

Dans la graine mature sèche (t0), les bandes protéiques majoritaires correspondant aux différentes sous-unités des globulines sont très abondantes, en accord avec les résultats de la Figure 44. L'intensité de ces bandes chute très fortement suite à l'imbibition de l'albumen seul (sans embryon) en présence d'eau seulement (Figure 45), de manière similaire à ce qui est observé avec la graine entière (Figure 44).

Les mêmes expériences ont été réalisées en imbibant l'albumen isolé des graines matures sèches d'*Amborella* en présence d'acide gibbérélique (GA<sub>3</sub>), toujours à partir d'une quantité fixe de graines (un dixième de graine), aux trois stades définis (GMS, 1/4, GSS). Le profil protéique est très similaire de celui obtenu par imbibition dans de l'eau seulement (Figure 45). Ainsi, il apparaît que la mobilisation des protéines de réserve dans l'albumen est un phénomène qui se manifeste lors de l'imbibition indépendamment de la présence de l'embryon et de l'apport exogène de gibbérélines (GA<sub>3</sub>). Ces résultats contrastent avec les

données accumulées chez les graines, notamment de céréales (Berger 2003 ; Finnie et al. 2011).

### **3.3 Discussion**

La caractérisation des variations protéiques au cours de la germination d'espèces modèles ou d'intérêt agronomique est bien documentée (Kriz et al. 1990 ; Gallardo et al. 2001 ; Gallardo et al. 2002 ; Rajjou et al. 2004 ; Agrawal & Rakwal 2006 ; Ahsan et al. 2007 ; Yang et al. 2007 ; Catusse et al. 2008 ; Yacoubi et al. 2010). En revanche, à notre connaissance, il n'existe aucune étude documentant cette étape pour les graines à petits embryons. Notre étude est ainsi la première à avoir caractérisé aussi finement les variations du protéome de la fraction albumines au cours de l'imbibition et de la germination *sensu stricto*, qui plus est, sur la graine de la sœur de toutes les plantes à fleurs.

Les protéines contenues dans des spots (protéomique 2D) dont le volume varie suite à l'imbibition de la graine mature sèche d'*Amborella* ont été identifiées par spectrométrie de masse. Elles ont été classées en quatre groupes en fonction de leur profil d'accumulation durant l'imbibition et la germination *sensu stricto* (Figure 42). Du fait de l'identification de plusieurs protéines dans un même spot (cas des spots multiples), seules les protéines présentant un index PAI (Protein Abundance Index ; [http://pappso.inra.fr/documents/bioinformatique/Xtandem\\_parser.pdf](http://pappso.inra.fr/documents/bioinformatique/Xtandem_parser.pdf)) supérieur à 0,2 sont ici discutées.

#### **3.3.1 Evolution de la fraction des albumines au cours de la germination**

##### ***3.3.1.1 Protéines s'accumulant progressivement tout au long de la germination***

Trois cent treize protéines s'accumulent progressivement au cours de la croissance interne (stade ¼) de l'embryon jusqu'à émergence de la radicule (stade GSS) (Figure 42) (tableau supplémentaire 6).

Il s'agit notamment d'une augmentation globale des protéines du métabolisme et de la production d'énergie. Ainsi, sur les 44 protéines de la catégorie « *metabolism* », 54,5% sont associées à la fonction « *lipid and stérol* » et 29,5% à la fonction « *amino acid* ». De même, 58 protéines correspondent à la catégorie « *energy* », 53,4% de ces protéines étant associées à la fonction « *glycolysis* » et 34,5% au cycle des acides tricarboxyliques « *TCA pathway* ».

Ainsi, nos résultats illustrent la reprise globale d'une activité métabolique et énergétique décrite chez les graines lors de la germination (Bewley & Black 1994).

L'accumulation des protéines HSP de type HSP70 de la catégorie « *protein folding and stability* » augmente tout au long de l'imbibition (stade ¼) et de la germination (stade GSS) (tableau supplémentaire 6). Ces protéines sont exprimées dans tout organe qui synthétise des protéines puisqu'elles sont responsables de repliement des protéines suite à la traduction des ARNm. Ainsi, le développement de l'embryon semble associé à une forte synthèse protéique. De plus, on note la présence d'une rotamase (Q38931) (tableau supplémentaire 6). Cette protéine est un co-chaperon des HSPs de haut poids moléculaire. Elle est notamment associée à l'augmentation de la tolérance à la chaleur et faciliterait l'action des HSPs de haut poids moléculaire dans ces conditions (Aviezer-Hagai et al. 2007).

De même, nous notons une forte augmentation de l'accumulation des protéines de la catégorie « *disease/defence* », notamment l'apparition d'une endochitinase (P19171) (tableau supplémentaire 6), une protéine induite en réponse à un stress biotique par l'acide jasmonique (une phytohormone dérivée de cyclopentanones issus de la voie de biosynthèse des acides gras), indépendamment de l'acide salicylique (Thomma et al. 1998). Toutes les enzymes impliquées dans la détoxification des composés oxydants sont fortement accumulées. Par ailleurs, nous notons la présence d'une spermidine synthase (Q9Z9B3) (tableau supplémentaire 6). Il est intéressant d'observer que l'application de spermidine (une polyamine synthétisée par les plantes) sur des plantules de tomate augmente considérablement l'expression d'enzymes de détoxification (Hu et al. 2009). Dans le présent contexte, il semble que la graine en germination présente un fort besoin de détoxification. Ainsi, la présence de spermidine pourrait être un mécanisme important déclenchant et amplifiant l'expression de tels gènes de défense. De plus, l'implication des formes activées de l'oxygène dans le relâchement des parois cellulaires permettant la percée racinaire est bien décrit (Müller et al. 2009 ; Iglesias-Fernández & Matilla 2010). Leur présence est donc nécessaire et fait l'objet d'une régulation très fine (Bailly 2004 ; Oracz et al. 2007). L'augmentation de l'accumulation des enzymes de détoxification pourrait donc correspondre à un marqueur de l'émergence racinaire.

Dans le cadre du métabolisme du soufre, il est très intéressant de noter l'augmentation du volume du spot contenant la PIMT au cours de l'imbibition et de la germination *sensu stricto* (tableau supplémentaire 6) (annexe 5). La protéine a été identifiée dans un spot multiple (SameSpot No. 808, tableau supplémentaire 6). Il convient donc de rester prudent quant à la réelle contribution de la PIMT à l'augmentation du volume de ce spot (tableau

supplémentaire 6). Néanmoins, le volume du spot concerné augmente d'un facteur trois entre la graine mature sèche et le stade ¼ pour finalement être quatre fois plus important à la germination *sensu stricto* en comparaison de la graine mature sèche, ce qui nous encourage à discuter cette observation en relation avec le rôle de la PIMT chez les graines. Des travaux chez *A. thaliana* ont démontré que cette enzyme jouait un rôle primordial dans la vigueur des graines et la germination (Ogé et al. 2008). Le fonctionnement de la PIMT implique la synthèse de son substrat, la *S*-adénosyl-méthionine synthétase (Mudgett et al. 1997). De plus, suite à l'utilisation par la PIMT de la *S*-adénosyl-méthionine il y a formation de *S*-adénosyl-homocystéine (Ogé et al. 2008). Or, ce produit de la réaction est un inhibiteur très puissant des méthyltransférases utilisant la *S*-adénosyl-méthionine comme co-substrat (incluant la PIMT) (Ravanel et al. 1998). L'hydrolyse rapide de ce métabolite est donc primordiale pour un fonctionnement efficace du cycle des méthylations. Cette étape d'hydrolyse est catalysée par la *S*-adénosyl-homocystéine hydrolase, identifiée chez *Amborella* (O23255) (tableaux supplémentaires 1 et 6) et dont l'accumulation augmente au cours de l'imbibition et de la germination *sensu stricto* (tableau supplémentaire 6).

Ainsi, il est tout à fait raisonnable de proposer que l'activité de la PIMT soit importante lors de l'imbibition/germination de la graine d'*Amborella*, permettant la réparation de protéines endommagées lors de la dessiccation des graines, et que son fonctionnement soit assuré par la présence de la *S*-adénosyl-homocystéine hydrolase. Par ailleurs, nous avons également identifié la méthionine synthase (O50008) (tableaux supplémentaires 1 et 6), autre enzyme indispensable au cycle des méthyles très important dans la physiologie de la graine (Gallardo, Job, Steven P. C. Groot, et al. 2002). En outre, la phosphoglycérate déshydrogénase (O49485) (tableaux supplémentaires 1 et 3) (synthèse de la sérine) et la cystéine synthase (P47998) (tableaux supplémentaires 1 et 3) ont été identifiées. Bien que ces protéines ne soient pas directement impliquées dans le cycle des méthyles, elles y jouent néanmoins un rôle important car la cystéine (qui dérive de la sérine) est le précurseur de l'homocystéine qui donnera la méthionine par action de la méthionine synthase (Ravanel et al. 1998).

Chez la graine de tomate (*Lycopersicon esculentum* Mill.) en germination, l'activité de la PIMT reste constante pendant 48 heures après l'imbibition puis diminue. Ces données suggèrent que l'accumulation de l'enzyme est contrôlée par le stade développemental et dégradée ou inactivée à la percée radriculaire (Kester et al. 1997). L'étude de l'accumulation des transcrits codant la PIMT chez le blé (*Triticum aestivum*) a également conduit à la conclusion que son niveau d'expression était influencé par le statut développemental et



l'ABA. L'activité de l'enzyme est décelable principalement pendant la phase tardive de maturation et pendant les vingt-quatre premières heures d'imbibition des grains de blé (Mudgett & Clarke 1994). En revanche, l'activité de la PIMT persiste lors de la germination des graines de lotus sacré (Shen-Miller et al. 1995). Ainsi, le profil obtenu chez *Amborella* laisse penser que le maintien (voire l'augmentation) de l'activité de cette enzyme est un processus ancien de réparation du protéome permettant de conserver une vigueur germinative appropriée à l'installation d'une plantule viable. L'augmentation ou la reprise de l'activité de cette enzyme suite à l'imbibition de la graine d'*Amborella* concorde avec l'ensemble des données acquises témoignant d'un rôle majeur de cette enzyme pour le succès de la germination et l'installation de la plantule (Rajjou et al. 2012). L'arbre phylogénétique de la Figure 37 montre que l'unique PIMT d'*Amborella* (qui contraste avec l'existence de deux gènes distincts chez les angiospermes) est proche d'enzymes équivalentes de gymnospermes, indiquant la préservation d'un mécanisme très ancien de survie des graines suite à leur dispersion chez les plantes.

Une  $\beta$ -mannanase (Q9FJZ3) (tableau supplémentaire 6) a été identifiée dans notre protéome au cours de la germination des graines d'*Amborella*. Cette enzyme est impliquée dans la dégradation du squelette des mannanes afin de faciliter l'émergence radiculaire (Yuan et al. 2007).

### ***3.3.1.2 Protéines spécifiques de la percée radiculaire***

Cinquante huit protéines sont spécifiquement accumulées au moment de la percée radiculaire (Figure 43). Une part très importante de ces protéines correspond à diverses enzymes de détoxification. Comme nous venons de le voir, la percée radiculaire nécessite un relâchement/ramollissement des tissus environnants. Cette étape clé de la germination est réalisée grâce à l'action concordante d'enzymes de relâchement des parois, ainsi que des formes actives de l'oxygène (Bailly 2004 ; Yuan et al. 2007 ; Müller et al. 2009). Ainsi, la forte représentativité des enzymes de détoxification concorde avec le stade développemental étudié.

L'identification d'une chorismate mutase (P42738) (tableau supplémentaire 6) avec une prédiction de localisation chloroplastique indiquerait la mise en place des chloroplastes pour préparer l'autotrophie de la plantule. Cette enzyme catalyse la synthèse de chorismate à partir du shikimate (Herrmann & Weaver 1999). De plus, deux enzymes de la voie de biosynthèse

du shikimate ont été identifiées. Il s'agit de la 3-déhydroquinate déshydratase/shikimate déshydrogenase (Q9SQT8) bifonctionnelle et la 3-phosphoshikimate 1-carboxyvinyltransférase (P05466) (tableau supplémentaire 6), toutes deux augmentant de la graine mature sèche à la germination *sensu stricto*. Ainsi, il semble que la graine ait le potentiel d'alimenter les voies de biosynthèse des acides aminés aromatiques, ce qui corrobore nos observations précédentes concernant l'importance de la synthèse protéique *de novo*. La production de métabolites secondaires dérivés du shikimate est également associée avec l'établissement des mécanismes de défense et la biosynthèse de l'auxine, l'acide indolyl-3-acétique, évènements nécessaires au déclenchement de la germination (Fait et al. 2006).

### ***3.3.1.3 Les protéines spécifiques du stade intermédiaire de la germination***

Trente cinq protéines spécifiquement accumulées au stade  $\frac{1}{4}$  ont été identifiées (Figure 43). Il s'agit notamment des protéines de réserve qui représentent 34,3% de l'ensemble des protéines identifiées dans ce groupe. Leur abondance dans la fraction albumines, qui tend à diminuer tout au long de la germination témoigne du recrutement des protéines de réserve qui deviennent alors solubles, puis de leur utilisation pour la développement et l'accomplissement de la germination (Bewley & Black 1994 ; Gallardo et al. 2001 ; Shewry & Halford 2002).

L'identification d'une bétaine déshydrogénase (Q9STS1) (tableau supplémentaire 6) dont l'accumulation est spécifique du stade  $\frac{1}{4}$  est très intéressante. Cette enzyme est impliquée dans la biosynthèse de la glycine bétaine, un osmolyte formé à partir de la choline (Russell et al. 1998 ; Catusse, Strub, et al. 2008). L'identification de cette enzyme et la position d'*Amborella* dans la phylogénie suggèrent qu'il s'agit d'un mécanisme ancien et qu'il permet à la graine de tolérer le stress hydrique lors de l'imbibition.

Les protéines de la fonction « *cell structure* » spécifique de ce stade sont une villine-2 (O81644) et une  $\alpha$ -L-arabinofuranosidase (Q9SG80) (tableau supplémentaire 6). Les villines sont des protéines interagissant avec l'actine. Chez les plantes, il semble qu'elles soient impliquées dans la croissance cellulaire et la différenciation (Klahre et al. 2000). De même, l' $\alpha$ -L-arabinofuranosidase serait impliquée dans la modification des parois (Minic et al. 2006 ; Chávez Montes et al. 2008). L'identification de ces deux enzymes au stade  $\frac{1}{4}$  confirme que l'embryon est en pleine expansion.

#### **3.3.1.4 Protéines diminuant progressivement lors de la germination**

Les niveaux d'accumulation de 113 protéines diminuent progressivement au cours de la germination (Figure 42 ; tableau supplémentaire 6). Il s'agit notamment de la catégorie « *protein destination and storage* » qui représente 39,8% des protéines diminuant constamment (Figure 42 ; tableau supplémentaire 6). La présence de telles protéines dans la fraction albumines des graines s'explique par leur mobilisation lors de la germination, ce qui les rend solubles dans l'eau (Gallardo et al. 2001). Elles sont ensuite utilisées notamment comme source d'azote et de carbone pour permettre le développement de l'embryon, expliquant leur diminution.

Par ailleurs, une protéine (Q9S9P4) correspondant à une *S*-adénosyl-L-méthionine-dépendante-méthyltransférase (tableau supplémentaire 6), plus précisément correspondant à une triméthylguanosine synthase, représente à elle seule 5,3% des protéines diminuant constamment pendant la germination. Cette enzyme est impliquée dans la mise en place de la coiffe des ARN messagers, accompagnant la synthèse *de novo* de ces ARNm. Cette observation suggère une diminution du besoin de transcription au cours de la germination *sensu stricto*, en parfait accord avec les résultats de Rajjou et al. (2004) documentant l'importance de la synthèse des protéines à partir des ARN stockés lors de la maturation de la graine, en préparation à la germination. En effet, l'utilisation dans des expériences de germination de graines d'*A. thaliana* de l' $\alpha$ -amanitine, un inhibiteur spécifique et très puissant de l'ARN polymérase II (de Mercoyrol et al. 1989), a démontré que la transcription *de novo* n'était pas requise pour l'accomplissement de la germination *sensu stricto* (Rajjou et al. 2004). Des résultats analogues ont été observés lors de la germination des graines de riz (Sano et al. 2012). Ce mécanisme d'utilisation des ARNm stockés pour le succès de la germination devra être plus approfondi par des expériences de séquençage d'ARN à haut débit, établissant le transcriptome des graines matures sèches d'*Amborella*, et l'évolution de ce dernier lors de l'imbibition.

### 3.3.2 Evolution de la fraction des globulines au cours de la germination

Les fractions globulines des trois stades de développement (Figure 40) ont été extraites puis analysées par gel monodimensionnel (Figure 44 A). Afin de pouvoir visualiser l'évolution du contenu en globulines dans chaque graine, le dépôt protéique de chaque puits correspondait à une unité de graine constante (0,1 graine). Les résultats, très surprenants, montrent une diminution des globulines très importante dès le stade  $\frac{1}{4}$  et quasi complète au stade correspondant à la germination *sensu stricto*. Afin de confirmer cette observation, un test ELISA a été réalisé sur les extraits globulines des trois stades en utilisant un anticorps anti-sous-unité-B spécifiquement développé à partir de l'antigène d'*Amborella* (Figure 44 B). La forte diminution des globulines a ainsi été confirmée et nos résultats indiquent qu'au moment de la percée radriculaire, il ne reste que 20% de la quantité initiale de globuline 11S présente dans la graine mature sèche.

Ce résultat est très original. En effet, il est couramment admis que l'utilisation massive des protéines de réserve est un processus post-germinatif, bien qu'un début de mobilisation soit visible avant la fin de la maturation et pendant la germination (Bewley & Black 1994 ; Job et al. 1997 ; Nonogaki et al. 2010 ; Rajjou et al. 2012). Cependant, les études ont été menées sur des espèces sans dormance morphologique.

Très peu d'études sont disponibles concernant l'évolution des protéines de réserve lors de la germination. Le céleri et *Trollius ledebouri* sont deux espèces à dormance morphologique. Ces études menées au début des années 1980 avaient déjà mis en évidence des modifications du contenu des albumens des deux espèces, sans pour autant avoir pu attribuer clairement l'origine de ces dernières (Jacobsen & Pressman 1979 ; Hopher & Roberts 1985b).

Le rôle de l'acide gibbérellique comme médiateur de l'embryon vers l'albumen est un mécanisme très documenté (Bewley & Black 1994). L'exemple le plus connu est l'effet des GAs sur les cellules de la couche à aleurone de l'orge (Finnie et al. 2011). Les GAs sont un facteur diffusible de l'embryon qui va stimuler la production d' $\alpha$ -amylase dans les cellules de la couche à aleurone, qui permettra ainsi la mobilisation des réserves de l'albumen (Varner & Chandra 1964). L'effet des GAs sur la germination provient également d'un ramollissement de l'albumen facilitant l'émergence radriculaire (Groot & Karssen 1987). Chez *Arabidopsis*, les GAs ont une action relativement tardive à la germination, et les événements sous leur

contrôle commencent réellement peu avant la percée radiculaire (Gallardo, Job, Steven P C Groot, et al. 2002).

Nos résultats témoignent d'une large autonomie de l'albumen vis-à-vis de l'embryon et des gibbérellines (GA<sub>3</sub>) (Figure 45) pour la mobilisation des réserves protéiques. Ainsi, cette mobilisation semble se manifester dès lors que l'albumen est mis à imbiber avec de l'eau et l'ajout de GA<sub>3</sub> ne modifie que très peu le profil protéique des globulines stockées. Nos données contrastent de ce point de vue avec les résultats sur le céleri et *T. ledebouri* pour lesquelles l'acide gibbérellique a un effet sur la structure et le contenu des réserves de l'albumen (Jacobsen & Pressman 1979 ; Hepher & Roberts 1985a ; Hepher & Roberts 1985b). Ainsi, l'albumen d'*Amborella* pourrait avoir un rôle essentiel sur le contrôle du développement de l'embryon. En effet, il pourrait être à lui seul responsable de la dégradation et de la mise à disposition des protéines de réserve sans lesquelles l'embryon ne pourrait pas se développer, sous la perception d'un signal environnemental indiquant une germination favorable. L'albumen de cette espèce basale serait donc un tissu nourricier apte à réguler le développement de l'embryon. Il serait ainsi bien plus qu'un simple tissu de stockage quasi mort, comme certains auteurs le considèrent. En ce sens, il serait un véritable gardien de l'embryon, détectant les conditions favorables à une future germination et lui procurant les réserves énergétiques nécessaires à l'atteinte de ce développement.

Nos observations montrent une reprise de l'activité métabolique tout au long de l'imbibition jusqu'à la percée radiculaire hors des téguments. Cet événement est associé à une augmentation de tous les moyens de défense contre les formes actives de l'oxygène et autres radicaux oxydants, notamment au moment de la percée radiculaire.

Par ailleurs, l'ampleur de l'utilisation des protéines de réserves avant la percée radiculaire est un processus qui n'a encore jamais été décrit à une telle échelle d'après nos connaissances. Ainsi, la définition classique de la germination de ce point de vue semble ne pas pouvoir s'appliquer pour la germination d'espèces à petits embryons, au moins à celle d'*Amborella*. Nous pensons que, dans ce cas précis, il conviendrait de qualifier de « croissance interne de l'embryon » les événements se produisant suite à l'imbibition de la graine, plutôt que de se référer aux définitions de dormance invoquées pour ces graines à petits embryons. En particulier, le profil de l'utilisation des protéines de réserve (globulines) pendant la croissance interne de l'embryon accompagnant l'imbibition serait à confirmer sur d'autres espèces à dormance morphologique. Néanmoins, une révision de la définition de la germination semble nécessaire à ce stade. En effet, nos données montrent clairement qu'il

existe une discordance entre les marqueurs moléculaires de fin de germination (utilisation massive des globulines) et le marqueur morphologique (percée d'une partie de l'axe embryonnaire au travers des téguments). Dans ce cadre, il est utile de rappeler les propositions de Bewley & Black (1994) concernant les stades marqueurs de fin de la germination *sensu stricto* des graines. Ainsi, selon ces auteurs, « *The only stage of germination that we can time fairly is its termination! Emergence of the axis (usually the radicle) from the seed normally enables us to recognize when germination has gone to completion, though in those cases, where the axis may grow before it penetrates through the surrounding tissues, the completion of germination can be determine as the time when a sustained rise in fresh weight begins* ». Ainsi, il serait très intéressant d'analyser le profil d'évolution des globulines à la fin du plateau d'imbibition caractéristique de la courbe de la prise en eau lors de la germination, pour voir si il y a concordance entre ce marqueur de fin de germination et le début de la mobilisation massive des globulines.

# Conclusion générale sur la caractérisation de la graine mature sèche d'*Amborella* et lors de sa germination

Notre étude est la première à avoir mis en évidence les originalités structurales et protéomiques de la graine mature d'*Amborella*, sœur de toutes les plantes à fleurs. Elle a notamment permis :

- la mise en évidence de la présence de deux compartiments séparés par un réseau de polysaccharides dans la graine mature,
- la première caractérisation des protéomes d'un embryon rudimentaire isolé et de l'albumen issu d'une graine à dormance morphologique,
- la caractérisation de l'évolution des fractions albumines et globulines au cours de l'imbibition/germination,
- de montrer un état de maturité moléculaire avancé de l'embryon alors que son aspect est rudimentaire, notamment par l'identification de chaperons (LEAs, HSPs) et de réserves lipidiques dans l'embryon,
- de montrer l'aptitude de l'embryon rudimentaire à utiliser les réserves lipidiques dans un premier temps, puis les réserves protéiques,
- la validation d'une partie du tout récent génome d'*Amborella* grâce à une approche de protéogénomique.

Ainsi, nos données indiquent que l'embryon rudimentaire au sein de la graine mature renferment plusieurs marqueurs moléculaires d'un embryon mature (réserves lipidiques, chaperons moléculaires), ce qui pourrait contraster avec l'apparence de son stade de développement morphologique.

Le suivi de l'évolution de son protéome au cours de l'imbibition/germination montre une reprise globale des activités métaboliques et énergétiques couramment observées lors de la germination d'autres espèces. Cependant, le profil d'utilisation des protéines de réserve (massif avant la percée radiculaire) indique que les marqueurs couramment utilisés pour suivre la germination ont un profil différent lors de la germination d'espèces à petits embryons. La définition de la germination dans ces cas bien précis devrait donc être précisée.

Comme nous l'avons vu, la signature ontologique des protéines totales de l'embryon d'*Amborella* ressemble fortement à celle de la fraction albumines de la graine entière. L'albumen est principalement constitué de protéines de réserve. La forte représentativité de ces dernières aurait pu masquer certaines autres protéines moins représentées mais toutes aussi importantes. Ainsi, afin de s'affranchir des globulines dans l'extrait total des protéines de l'albumen, un extrait de la fraction albumines de l'albumen sera réalisé et analysé dans les plus brefs délais par protéomique *shotgun*. Nous espérons ainsi pouvoir caractériser un plus grand nombre de protéines du métabolisme dans l'albumen isolé.

Nos données suggèrent également l'implication des phytohormones ABA et GA dans le contrôle de la germination (protéine Mother of TF and TFL1). Ainsi, dans un premier temps, une recherche dans le génome des enzymes des voies de biosynthèse de l'ABA et des GA devrait être réalisée, de même que la recherche des éléments de réponse spécifiques (éléments ABRE pour la réponse à l'ABA et éléments GARE pour les GAs). Puis le rôle de ces phytohormones devrait être étudié au cours de l'imbibition/germination. De même, nos données suggèrent un rôle important de l'albumen vis-à-vis de la mise à disposition des protéines de réserve pour l'embryon. L'autonomie au développement de l'embryon devrait être étudiée en réalisant des expériences de germination *in vitro* à partir d'embryons isolés de graines matures.

Afin de valider nos observations, l'étude de l'évolution du profil des globulines mériterait d'être réalisée sur d'autres graines à dormance morphologique, ainsi que sur d'autres espèces basales.

L'identification de différents marqueurs de maturité et de tolérance à la dessiccation (protéines chaperons LEA et HSPs) indique que l'embryon d'*Amborella* serait tolérant à la dessiccation. Cependant, il sera nécessaire d'approfondir cet aspect, notamment en étudiant l'évolution de la vigueur germinative suite à la conservation des graines à différentes humidités, température et pendant plusieurs périodes de temps.

Nos résultats d'approche phylogénétique sur les globulines 11S, les oléosines et la PIMT indiquent qu'*Amborella* présente à la fois des caractéristiques propres aux gymnospermes et d'autres propres aux angiospermes. Afin d'approfondir plus amplement cet aspect d'évolution, nous allons démarrer une collaboration avec l'équipe du professeur Claude de Pamphilis (Penn State University, USA ; *Amborella* Genome Project) afin de d'examiner à l'échelle de l'ensemble du protéome que nous avons caractérisé les traits ancestraux de l'expression de gènes d'*Amborella* s'exprimant dans les graines.



**Chapitre IV. Adaptations au stress  
nickélique : *Psychotria gabriellae***



# 1 Introduction

Comme nous l'avons vu, la flore néo-calédonienne est l'une des plus riches au monde en plantes hyperaccumulatrices de nickel (Ni). Ceci est dû à la formation d'un vaste complexe ophiolitique il a 37 Ma dont l'érosion a abouti à la formation de sols ultramafiques sur un tiers de la surface de la Grande-Terre. Ces sols sont pauvres en macroéléments essentiels au développement des végétaux (azote, carbone, phosphore, potassium et soufre) et riches en éléments traces métalliques (chrome, cobalt, manganèse et nickel). Les pressions de sélection et de spéciation y sont donc fortes. Les plantes qui s'y développent ont ainsi façonné des mécanismes spécifiques pour surmonter la toxicité de ces éléments métalliques.

Les travaux de ce chapitre de thèse font suite à des résultats obtenus précédemment au LIVE sur le genre *Psychotria*, et à la mise en place récente d'un projet d'approche transcriptomique chez *Psychotria gabriellae* (ex *P. douarrei*). Nos travaux concernent principalement l'adaptation des graines de cette plante au stress nickélicifère.

Un état de l'art sera tout d'abord présenté. Ce dernier inclura l'énoncé des résultats précédemment acquis au laboratoire et ayant abouti aux problématiques qui ont guidées notre démarche.

## **1.1 Le milieu ultramafique**

### **1.1.1 Origine et évolution**

Les roches ultramafiques (ou ultrabasiqes) proviennent de la partie supérieure du manteau terrestre, composé principalement de péridotites. Elles sont pauvres en silice (moins de 45%), riches en fer (Fe) et magnésium (Mn) (terme ma-fique : Mn et Fe). Plus de 70% de ces minéraux sont ferromagnésiens (Coleman & Jove 1992). Les éléments sidérophiles (fer, cobalt, chrome, nickel,...) y sont en proportions importantes (Brooks 1987). Les éléments tels que calcium, azote, phosphore et potassium, peu affins des minéraux ferromagnésiens, y sont présents en faibles quantités (Zongo 2010). La présence de fer oxydé donne une couleur noirâtre-rougeâtre à ces sols (Figure 46). L'albédo y est donc faible, et les températures de surface très élevées au cours de la journée.



Figure 46 : photographie d'un sol ultramafique et de sa végétation arbustive en haut du Massif du Koniambo (Koné).

Les roches ultramafiques proviennent de l'ophiolite mis en place il y a environ 37 Ma (cf. « Introduction générale »). Cette formation a subi une forte érosion, et les roches ultramafiques sont maintenant présentes sur un tiers de la Grande-Terre.

Les péridotites sont donc l'horizon le plus primitif des roches ultramafiques de Nouvelle-Calédonie. Les processus d'altération dus aux pluies, aux variations de pH du sol et aux modifications climatiques ont abouti à la formation d'un profil d'altération type (Figure 47). Bien que ce dernier puisse présenter des variations en fonction de la roche mère et du degré d'érosion, ce profil est la plupart du temps composé des différentes couches suivantes du bas vers le haut (Figure 47).

- La roche mère constituée des péridotites, dont les fractures peuvent favoriser une altération dite pénétrante.
- La saprolite qui est une couche conservant la structure des péridotites mais pour laquelle on observe une diminution de la teneur en magnésium, une accumulation du fer et un enrichissement en nickel.
- La latérite jaune qui résulte de l'altération avancée des saprolites. Le magnésium et la silice y sont très peu représentés, alors que la quantité de fer continue d'augmenter. Les teneurs en nickel y sont faibles.
- La latérite rouge, un horizon remanié où les structures de la roche mère sont effacées.

- La grenaille et la cuirasse qui correspondent à l'altération ultime du profil. La cuirasse est la partie la plus colonisée par les organismes vivants, et est souvent appelée « *top soil* ».

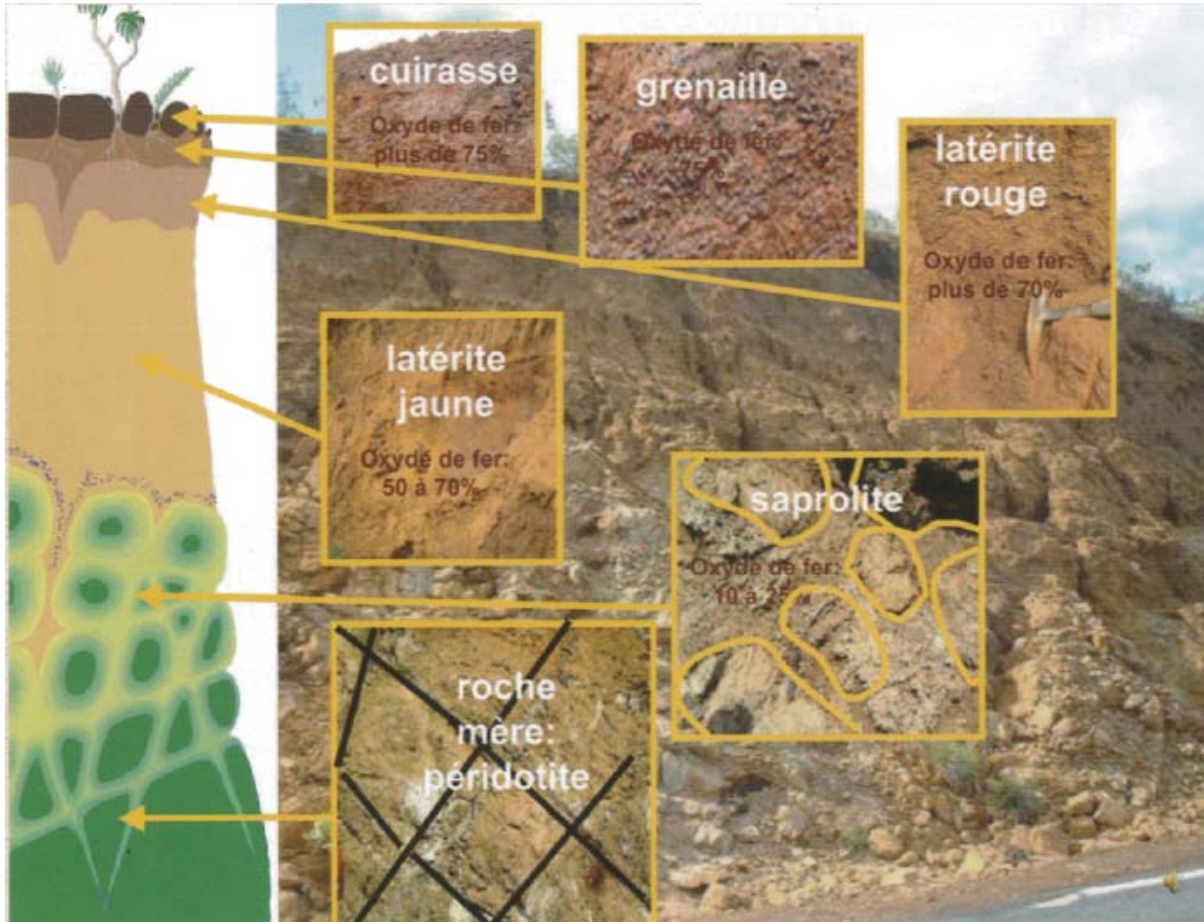


Figure 47 : schéma du profil d'altération des roches péridotidiques de Nouvelle-Calédonie (source : le maquis minier, Collection CIE - milieux naturels ; [www.cie.nc](http://www.cie.nc))

Les processus d'altération majeurs qui conduisent à la concentration des métaux sont de deux ordres.

- La serpentinisation, réalisée à haute température (100 à 400°C) sur des péridotites hydratées, un processus d'altération résultant principalement en un départ privilégié de certains éléments tels que la magnésium et l'enrichissement en nickel dans les horizons inférieurs (Pelletier 1982). L'érosion des péridotites par ce processus aboutit à la formation de serpentine. Ainsi le terme de sol serpentique désigne les écosystèmes développés sur roches ultramafiques altérées en serpentine.

- L'altération supergène qui intervient suite à la mise en place d'un affleurement de formation ultramafique. La pluie et les températures élevées modifient la géochimie du sol. La roche se dissout, le nickel est libéré en solution et circulera vers les strates inférieures qu'il enrichira.

### 1.1.2 Caractéristiques édaphiques

L'originalité de la composition des sols ultramafiques impacte les caractéristiques édaphiques de ces milieux. Elles sont principalement caractérisées par :

- Une faible teneur en macroéléments essentiels au développement des végétaux. Les teneurs en azote, phosphate, potassium, soufre et calcium y sont très faibles. La déplétion en calcium serait la plus limitante pour le développement des végétaux. De même, les pH basiques sont défavorables à l'absorption des éléments par les plantes (Brooks 1987).
- Une forte teneur en élément traces métalliques, principalement le chrome, le nickel, le cobalt et le manganèse. Bien que certains de ces éléments traces soient des microéléments indispensables au fonctionnement des organismes vivants, leur forte accumulation dans les cellules peut devenir toxique. Il n'existe malheureusement pas de valeur moyenne quant à la biodisponibilité de chacun de ces éléments dans les sols ultramafiques néo-calédoniens. En effet, leurs concentrations varient grandement selon les conditions, mais elles sont toujours nettement supérieures à celles retrouvées dans d'autres sols non ultramafiques.
- Un faible recyclage de la matière organique, lié d'une part à la présence d'une flore de petite taille générant une faible quantité de débris difficiles à dégrader, et d'autre part, à une faible teneur en eau ne favorisant pas la décomposition des débris végétaux. La couche de litière y est donc très mince, voire quasi inexistante.
- Une faible disponibilité de l'eau, qui, de par la structure compacte et craquelée de la cuirasse, n'est pas retenue.

## 1.2 Le nickel dans la plante

D'un point de vue global, les effets de fortes concentrations en éléments traces métalliques sur une plante sont similaires à ceux engendrés par la sécheresse (Barceló & Poschenrieder 1990). Il s'agit principalement d'une réduction de la capacité d'absorption de la solution du sol, de la capacité à transporter les solutés et de la transpiration, tout cela étant associé à un ralentissement de la croissance végétative.

La Nouvelle-Calédonie possède avec Cuba la plus grande représentation au monde d'espèces végétales hyperaccumulatrices de métaux lourds, notamment de nickel (Brooks et al. 1998). Ce dernier est d'ailleurs l'élément trace métallique le plus accumulé par les plantes (Tableau 7).

Tableau 7 : nombre de plantes hyperaccumulant différents métaux (Callahan 2007).

<b>Element</b>	<b>Normal plant mg/kg dried leaf matter</b>	<b>Hyperaccumulator threshold mg/kg dried leaf matter</b>	<b>Number of known hyperaccumulator species</b>
Mn	20-400	>10,000	9
Co	0.03-2	>1,000	28
<b>Ni</b>	<b>1-10</b>	<b>&gt;1,000</b>	<b>360</b>
Cu	5-25	>1,000	37
Zn	20-400	>10,000	11
Pb	0.1-5	>1,000	14

Les métaux de transition (Ni, Fe, Mn, Cu et Zn) sont des composants intrinsèques de certaines enzymes et protéines et sont donc essentiels à la survie des organismes vivants. Les caractéristiques de ces métaux de transition incluent i) une haute affinité de leurs ions pour les ligands contenant de l'azote, de l'oxygène et du soufre, ainsi que ii) différents statuts redox dans les conditions physiologiques. Néanmoins, ces propriétés font également des éléments de transition de redoutables ennemis moléculaires à forte concentration. Les métaux lourds se

fixent aux groupements carboxyles, imidazoles, sulfhydriles et aux amines aliphatiques. Cette association engendre des modifications de la structure tridimensionnelle des protéines pouvant altérer, voire inhiber, leur activité physiologique (Seregin & Kozhevnikova 2006). L'homéostasie fine de ces éléments est donc cruciale pour la survie des plantes.

Le nickel est le vingt-deuxième élément le plus abondant de la croûte terrestre. Son numéro atomique est 28, sa masse atomique est de 58,71 g/mol et sa forme la plus stable est  $Ni^{2+}$  dans une large gamme de conditions physiologiques (Yusuf et al. 2011). Sa concentration moyenne dans les sols est généralement inférieure à 100 ppm bien que les rejets liés à l'activité anthropique puissent aboutir à une concentration de 26 000 ppm dans des sols pollués (Wuana & Okieimen 2011). En effet, le nickel est un constituant des batteries, il est utilisé dans l'industrie des alliages pour ses propriétés antirouille, et se retrouve également dans les fertilisants du sol ou même les rejets ménagers.

Ses propriétés physico-chimiques sont relativement proches de celle du zinc ( $Zn^{2+}$ ). Ce dernier est le constituant de nombreuses enzymes et protéines (e.g. ARN polymérase, ....., protéines régulatrices à doigts de zinc). Ainsi, en cas de forte présence de nickel à l'état ionique dans la cellule, les deux éléments, zinc et nickel, peuvent rentrer en compétition provoquant l'inactivation des enzymes si le nickel « l'emporte » (Seregin & Kozhevnikova 2006). Le nickel peut se lier à de nombreux groupements, notamment aux groupements sulfhydriles (-SH), et aux atomes d'azote impliqués dans des cycles aromatiques (cas de l'histidine). Ainsi, une forte concentration, non régulée, de nickel dans la cellule provoquera un ralentissement du métabolisme par « simple » inactivation ou détérioration des activités enzymatiques des enzymes et des protéines avec lesquelles il aurait engagé des liaisons (Seregin & Kozhevnikova 2006 ; Yusuf et al. 2011 ; Sharma & Dhiman 2013). Dans ce cadre, il a été démontré que le nickel entraîne une diminution de l'activité d'enzymes telles que la Rubisco, la glycéraldéhyde 3-phosphate déshydrogénase, l'aldolase, ... probablement via une interaction avec les groupements -SH ou les histidines (Seregin & Kozhevnikova 2006). L'inactivation d'enzymes impliquées dans les chaînes de transfert d'électrons pourra notamment aboutir à la production d'espèces activées de l'oxygène (*reactive oxygen species*, ROS). En outre, comme tous les métaux de transition, le nickel peut participer à la production de  $OH^\circ$  via la réaction de Fenton/Haber-Weiss (Kehrer 2000), bien que ce dernier soit un mauvais candidat de par son potentiel redox élevé (Leonard et al. 2004). Ainsi, lors d'un stress nickélicifère, la plupart des enzymes de détoxification cellulaire sont activées. De manière surprenante, l'activation de ces enzymes observée suite à la présence de cadmium chez les espèces non hyperaccumulatrices n'est pas retrouvée chez les espèces hyperaccumulatrices de



nickel (Schickler & Caspi 1999). Ceci suggère la mise en place de mécanismes de tolérance spécifiques pour les espèces hyperaccumulatrices de différents métaux de transition.

Les membranes cellulaires forment la première barrière de la cellule. Une forte concentration en nickel a pour effet de diminuer la fluidité de celles-ci, notamment en interagissant avec les protéines transmembranaires (Yusuf et al. 2011).

Le nickel est cependant un microélément indispensable au développement des végétaux. Il peut se lier à diverses enzymes. Cependant, le seul cas aujourd'hui connu et clairement identifié nécessitant son rôle catalytique au sein d'une enzyme concerne son interaction avec l'uréase (Eskew et al. 1984 ; Seregin & Kozhevnikova 2006 ; Yusuf et al. 2011 ; Sharma & Dhiman 2013). Cette enzyme, présente dans des bactéries, des levures et certaines plantes, est indispensable à la dégradation de l'urée qui devient toxique si elle n'est pas éliminée. En particulier, une forme de l'uréase des plantes est exprimée spécifiquement et abondamment dans l'embryon (soja, *Arabidopsis*), alors qu'une autre forme, dite ubiquiste, s'accumule en quantités variables dans les tissus végétatifs de diverses plantes (Sirko & Brodzik 2000). Chaque sous-unité de l'enzyme contient deux atomes de nickel dont la stabilité est assurée par des protéines auxiliaires (Sirko & Brodzik 2000). Le nickel est également un inhibiteur des répresseurs de l'expression des gènes codant l'uréase. Ce microélément ne peut pas être remplacé (Eskew et al. 1984). Par ailleurs, une diminution de l'activité de l'hydrogénase a été observée en absence de nickel. Ce métal semble donc important pour l'activité de cette enzyme dont l'action est primordiale pour la fixation de l'azote au sein des nodules symbiotiques (Dalton et al. 1985).

D'autres observations indiquent que le nickel appliqué à faible dose stimule la germination du riz, du blé, ou du soja notamment en activant certaines enzymes. Cependant, les mécanismes moléculaires de cette régulation restent à définir (Seregin & Kozhevnikova 2006 ; Yusuf et al. 2011).

Certaines plantes, en plus d'avoir développé une capacité à tolérer de fortes concentrations en métaux dans les sols, s'avèrent capables de les séquestrer et de les stocker en quantités extraordinairement élevées dans leurs parties aériennes. Ces plantes sont qualifiées d'hyperaccumulatrices de métaux, à des seuils variables au regard de l'élément concerné (Tableau 7).

Sur les 45 espèces végétales hyperaccumulatrices de métaux répertoriées, 75% sont hyperaccumulatrices de nickel. L'existence de cette faculté d'hyperaccumulation est présente chez de nombreuses familles végétales, ce qui suggère une origine phylogénétique (Verbruggen et al. 2009). Au regard de l'élément nickel, une plante est qualifiée

d'*accumulatrice* lorsque la concentration de l'élément est supérieure à 100 ppm dans ses feuilles (100 mg/kg MS), et d'*hyperaccumulatrice* au delà de 1 000 ppm (1 000 mg/kg MS).

Les sources de nickel dans les milieux naturels sont les sols serpentiniques (ultramafiques) et les différents rejets issus de l'action de l'homme (fertilisants du sol, industrie minière) (Brooks et al. 1998 ; Wuana & Okieimen 2011). Il est difficile d'établir clairement la teneur en nickel total des sols ultramafiques qui peut varier de 80 à 7 100 µg Ni/g MS (Brooks 1987).

La première plante hyperaccumulante le nickel fut décrite en 1948. Il s'agit d'*Alyssum bertolonii* qui se développe sur un sol avec une teneur en NiO de 4 200 mg/kg et qui peut présenter des concentrations foliaires de nickel de l'ordre de 10 000 mg/kg (Minguzzi & Vergnano 1948). D'autres études ont par la suite étoffé les données sur ces plantes hyperaccumulatrices de nickel (Brooks et al. 1977 ; Lee et al. 1978 ; Jaffré et al. 1979 ; Kersten et al. 1980 ; Brooks 1987 ; Brooks et al. 1998). Une troisième classe fut en particulier créée pour les plantes contenant plus de 10 000 ppm (10 000 mg/kg MS) de nickel dans leurs feuilles. Dans ce cas, les plantes ont été qualifiées d'*hypernickelophores* (Jaffré & Schmid 1974 ; Sagner et al. 1998).

Une plante se développant sur un sol contenant une forte teneur en élément métallique potentiellement toxique peut présenter trois types de réponse possibles (Figure 48) : l'absorption non régulée, l'exclusion et l'(hyper)accumulation. L'absorption non régulée conduira généralement à la mort de l'individu si celui-ci n'a pas développé de mécanismes de régulation interne envers la toxicité du métal. La stratégie d'exclusion limite l'entrée de l'élément au niveau racinaire et sa translocation dans les parties aériennes, permettant ainsi le développement de la plante. La stratégie d'hyperaccumulation conduit à la concentration de l'élément dans les parties foliaires, ce dernier étant souvent plus concentré dans les feuilles que dans le sol lui-même, sans signe de toxicité pour la plante. Par ailleurs, une plante hyperaccumulatrice transplantée dans un sol contenant une faible teneur en élément hyperaccumulé aura un développement normal. Il n'existe donc pas de besoin physiologique à l'hyperaccumulation (Reeves & Baker 1984). En revanche, l'hyperaccumulation aurait une signification écologique importante. Elle permettrait de protéger la plante des herbivores et des attaques fongiques (Davis et al. 2001 ; Hanson et al. 2004), bien que certains insectes semblent avoir développé des stratégies de tolérance au nickel (Boyd et al. 2006). L'hyperaccumulation dans les feuilles pourrait également provoquer un enrichissement localisé de l'élément autour de la plante, lorsque les feuilles sénescents chutent au sol. La plante augmenterait ainsi la concentration de l'élément au niveau du sol, limitant la

germination et le développement d'autres espèces moins adaptées (Boyd 2004 ; L. Zhang et al. 2007).

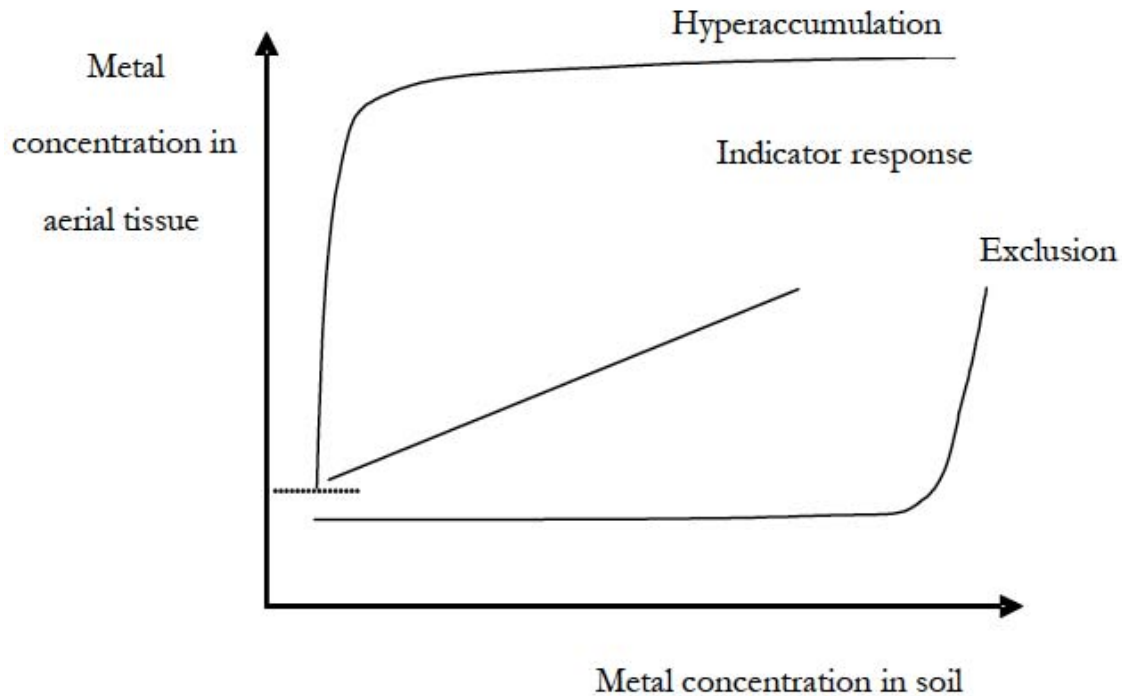


Figure 48 : stratégies de réponse des plantes exposées à une concentration croissante de métal dans le sol (Callahan 2007).

Les mécanismes physiologiques de l'hyperaccumulation sont encore mal compris. La plante doit être capable d'absorber l'ion, de le transporter et de le stocker sous une forme non toxique pour la cellule, généralement sous une forme complexée à un ligand. Les zones de stockage les plus probables sont la paroi végétale, le cytosol et la vacuole (Clemens et al. 2002) (Figure 49). L'idée d'adaptation admise pour le moment est que le mécanisme d'hyperaccumulation ne résulte pas de la création de nouvelles voies, mais plus d'une régulation différente des mécanismes de tolérance communs à toutes les espèces (Verbruggen et al. 2009) concernant l'absorption, le transport, la chélation et le processus de détoxification.

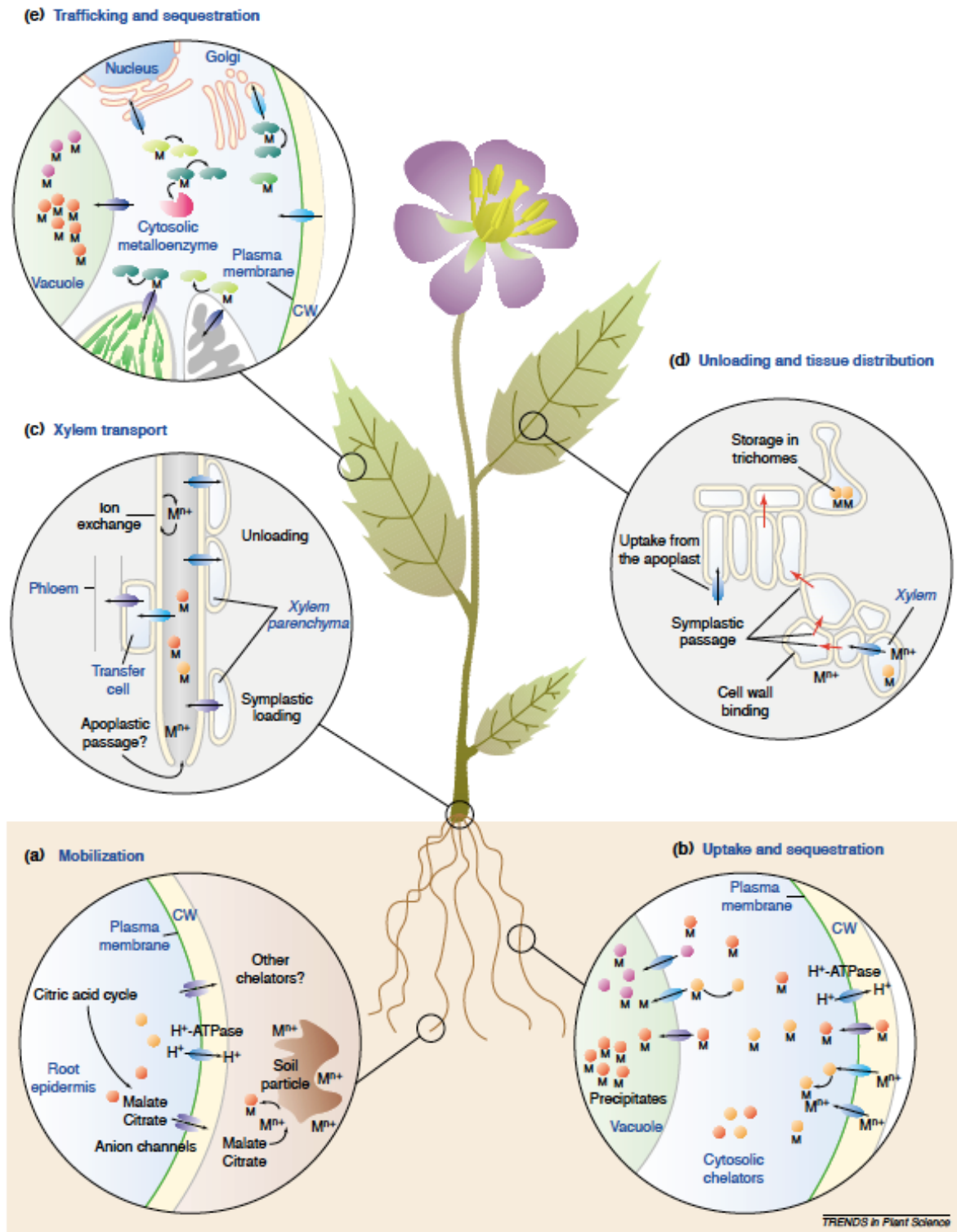


Figure 49 : mécanismes moléculaires proposés pour être impliqués dans l'hyperaccumulation de métaux par les plantes (Clemens et al. 2002). (a) Les ions sont mobilisés par la sécrétion de chélateurs et l'acidification de la rhizosphère. (b) Importation des ions ou du complexe ion-chélateur par différents systèmes membranaires. Une fois dans la cellule, les ions sont chélatés et/ou séquestrés dans la vacuole par différents systèmes de transporteurs

tonoplastiques. (c) Les ions sont transportés des racines vers les feuilles via le xylème. Le transport, des cellules racinaires vers les vaisseaux xylémiens, se ferait très vraisemblablement de manière symplastique car le transfert apoplastique n'existerait qu'au niveau de l'apex racinaire. Une fois dans le xylème, les ions sont transportés à l'état hydraté ou complexé. (d) Les métaux, une fois dans l'apoplaste foliaire, sont transloqués dans les cellules puis circulent par les plasmodesmes. Le stockage se fait principalement dans les trichomes. (e) La translocation des ions dans les cellules foliaires est réalisée par différents transporteurs (non représentés). La distribution intracellulaire est réalisée par l'action de métallochaperons et de transporteurs internes. CW : paroi cellulaire, M : métal, ronds pleins : chélateurs, ovales pleins : transporteurs, structure forme de haricots : métallochaperons.

Chez les plantes, la principale force motrice de l'absorption des éléments minéraux est le gradient de potentiel électrique membranaire (négatif) qui facilite l'entrée de cations. Tout d'abord, au niveau du sol, les métaux sont souvent sous forme insoluble, c'est notamment le cas du fer. Afin de solubiliser les métaux, la plante doit interagir avec la rhizosphère, notamment via des champignons mycorhiziens, induire un abaissement du pH du sol ou la synthèse d'exsudats (ligands des métaux) (Brown 1978 ; Whiting et al. 2001 ; Amir et al. 2007). Ces stratégies concernent principalement le fer, et aucune donnée ne permet de conclure sur leur implication s'agissant d'autres métaux lourds (Zhao et al. 2001).

Il existe des transporteurs impliqués dans le mouvement du nickel mais dont les mécanismes d'action restent à approfondir :

- P-type ATPase-HMA : ou « *heavy-metal transporting ATPases* » : ce type de transporteurs serait notamment responsable du chargement des métaux dans le xylème par utilisation d'ATP.
- OPT : la superfamille de transporteur d'oligopeptides, donc la sous-famille des *yellow-stripe 1-like* est impliquée dans l'influx et l'efflux de complexes métal-nicotianamine au niveau du système vasculaire.
- CDF : pour « *Cation Diffusion Facilitator* », aussi appelé « *Metal Transporter Protein* » (MTPs). Ces transporteurs seraient impliqués dans le passage des ions du cytoplasme vers les organelles.
- ABC : pour « *ATP-binding cassette* ». Ces transporteurs seraient impliqués dans la séquestration vacuolaire des métaux sous leur forme chélatée.

Aucun transporteur spécifique de nickel n'a encore été identifié chez les plantes. AtRT1 est un transporteur du fer sous sa forme  $Fe^{2+}$  chez *A. thaliana*, principalement au niveau des racines. Ce dernier est également impliqué dans le transport du nickel, notamment en cas de carence en fer. Le nickel et le fer seraient donc en compétition *in vivo* pour ce transporteur dont la spécificité semble faible (Vert et al. 2002 ; Nishida et al. 2011). AtIREG2 serait quant à lui exprimé au niveau du tonoplaste pour compenser la faible sélectivité de AtRT1 et ainsi permettre l'entrée de nickel dans la vacuole (Schaaf et al. 2006).

Une très faible proportion des métaux présents dans la plante se trouve sous forme d'ions libres (aqueux). La plupart des métaux sont liés à des ligands ou des protéines, ce qui permet de neutraliser leurs effets nocifs (Haydon & Cobbett 2007). Ces ligands sont variés, et il peut s'agir d'acides organiques, d'acides aminés, de peptides ou de protéines.

Au niveau de la graine, les métaux sont délivrés par le phloème dans l'apoplaste au niveau de la jonction entre les tissus maternels et les téguments de la graine. Les métaux ne peuvent pas passer la barrière apoplastique pour atteindre la graine à moins d'être pris en charge par des transporteurs spécifiques (Vogel-Mikus et al. 2007). La répartition des métaux dans les tissus de la graine est ensuite très variable selon l'espèce (Kranner & Colville 2011). Les graines comptent certainement parmi les organes les plus résistants aux stress, notamment dans le cas des graines orthodoxes (Kranner, Minibayeva, et al. 2010). Cependant, un excès de métaux dans les graines pourra provoquer la production de ROS entraînant la dégradation des protéines et des lipides par oxydation (Sharma & Dietz 2009).

### **1.3 Etat des recherches au L.I.V.E**

Afin d'étudier les mécanismes développés par les graines pour se prévenir des effets toxiques des métaux lourds et notamment du nickel, des recherches sont menées au sein de L.I.V.E sur l'espèce hyperaccumulatrice de nickel, *Psychotria gabriellae* (ex *P. douarrei*). Le modèle d'étude inclut également la comparaison avec une espèce non-hyperaccumulatrice du même genre, *Psychotria baillonii* (cf. chapitre « Matériel et Méthodes »). Des résultats importants ont été obtenus dans le cadre de la thèse de Charly Zongo (Zongo 2010) et sont présentés brièvement ci-après.

Une première étude avait conclu que 63% du nickel présent chez *P. gabriellae* était complexé avec du malate (Kersten et al. 1980). Plus récemment il a été proposé que le ligand principal du nickel chez *P. gabriellae* est l'acide malonique. Cependant, d'autres ligands

potentiels ont été retrouvés tels que l'acide salicylique, l'acide fumarique et la nicotianamine (Kersten et al. 1980 ; Callahan et al. 2012). Toutefois, ces données concernent les feuilles et aucune donnée n'est disponible concernant les ligands du nickel dans la graine.

La caractérisation minérale de plants de *P. gabriellae* se développant sur un sol contenant environ 4 000 ppm total de nickel indique que les racines, tiges et feuilles contiennent respectivement 3 000, 9 800 et 18 000 ppm de nickel, confirmant ainsi le caractère d'hypernickelophore de l'espèce (Zongo 2010). Par ailleurs, les graines contiennent 9 300 ppm de nickel (Zongo 2010). Pour comprendre une telle accumulation de nickel, une approche de microscopie électronique à balayage avec une analyse point par point a été réalisée. Fait remarquable, la cartographie brute aux rayons X (Figure 50) de la répartition des éléments nickel, calcium et phosphore a indiqué l'existence d'un gradient de nickel dans la graine. Ce gradient semble concentrer le métal dans les parties externes de la graine, le nickel étant très peu présent au voisinage immédiat de l'embryon. Autre élément notable, la répartition du phosphore est similaire à celle du nickel, à l'exception de la forte présence de phosphore dans l'embryon, très probablement en raison des réactions métaboliques s'y déroulant. En revanche, le calcium se comporte de manière opposée. Il est peu présent dans les parties externes et se concentre au voisinage de l'embryon. Ce gradient de nickel, d'origine inconnue, séquestre l'élément à l'extérieur de l'embryon. Il s'agit très vraisemblablement d'un mécanisme particulier qui tend à protéger l'embryon de la toxicité du métal.

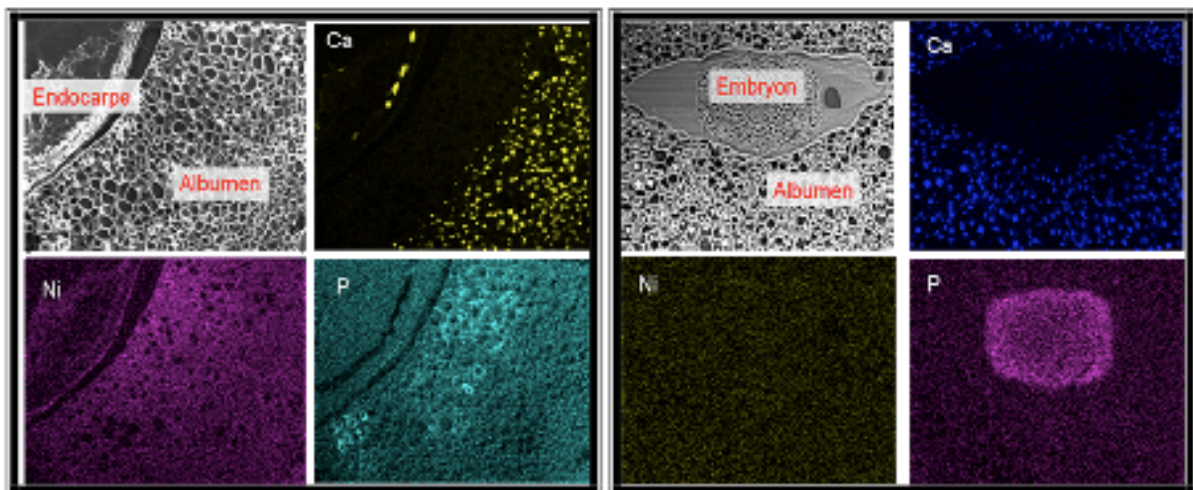


Figure 50 : cartographies brutes aux rayons X de la répartition des éléments nickel (Ni), calcium (Ca) et phosphore (P) au sein de l'endocarpe (à gauche), de l'albumen (à gauche et à droite) et de l'embryon (à droite) de *P. gabriellae* (d'après Zongo 2010).

La même étude a été réalisée chez *P. baillonii*, se développant sur le même sol que *P. gabriellae*. Les concentrations en nickel y sont de 430 ppm dans les racines, 230 ppm dans les tiges, 550 ppm dans les feuilles et 40 ppm dans les graines (Zongo 2010). L'observation au microscope électronique et l'établissement de la cartographie aux rayons X (Figure 51) ont permis d'observer que le calcium était également très présent au voisinage de l'embryon, et en quantité moindre dans les parties externes de la graine. Contrairement à *P. gabriellae*, le nickel n'était pas observable. De plus, le phosphore était visible dans l'embryon, et presque pas dans les parties externes de la graine. Les auteurs conclurent ainsi que le profil de répartition du calcium était indépendant du nickel, mais pas celui du phosphore (Zongo 2010). Ces données indiquent également que le nickel n'est pas présent dans la graine de *P. baillonii*. De fait, il n'est donc pas nécessaire de mettre en place chez les graines de *P. baillonii* un mécanisme spécifique de protection de l'embryon.

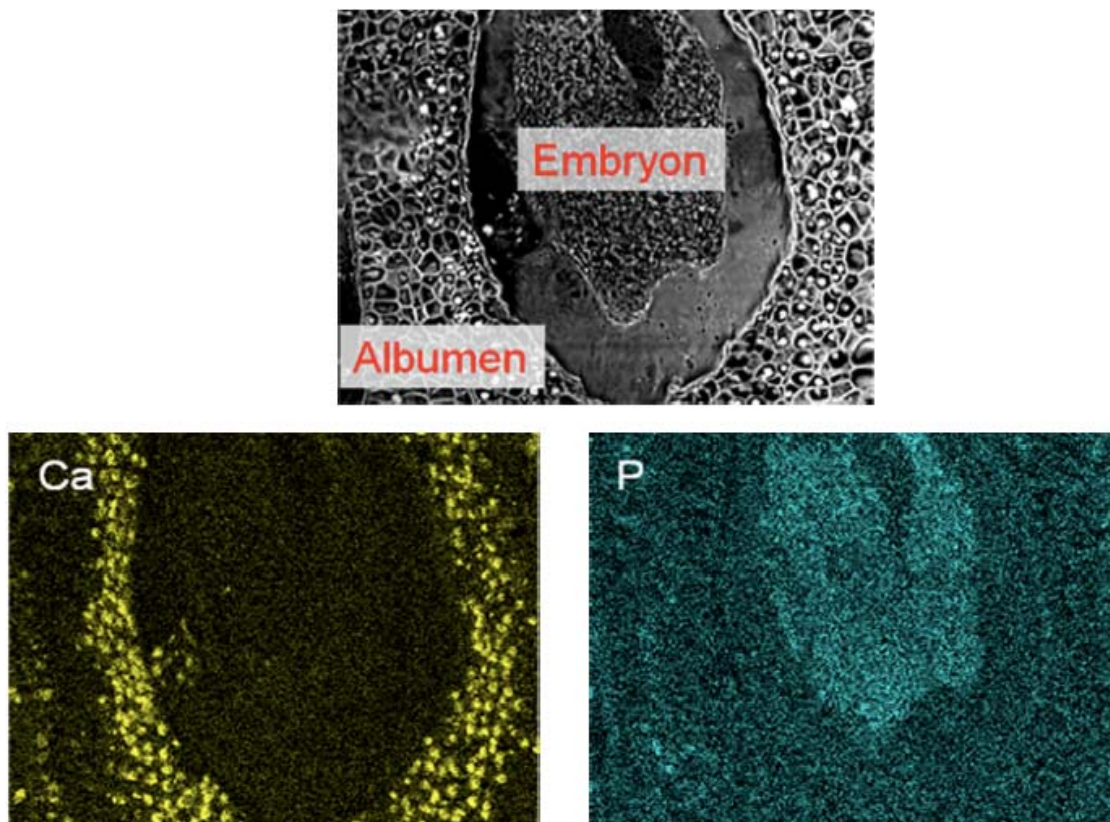


Figure 51 : cartographie au rayons X de la répartition des éléments calcium (Ca) et phosphore (P) au sein de l'albumen et de l'embryon de *P. baillonii* (d'après Zongo 2010).

Ces observations suggèrent donc l'existence d'un mécanisme de régulation de la répartition du nickel dans la graine de *P. gabriellae* impliquant l'élément phosphore. Une approche d'analyse protéomique a donc été développée afin de mettre en évidence des



protéines qui pourraient être potentiellement impliquées dans un tel mécanisme adaptatif. Dans ce but, les protéines des graines matures des deux espèces *P. gabriellae* et *P. baillonii*, ont été analysées par électrophorèse bidimensionnelle, et les protéines d'intérêt ont été identifiées par séquençage MS-MS. L'analyse des protéines et leur classification ontologique ont permis de conclure que la signature ontologique de la graine de *P. gabriellae* (Zongo 2010) était similaire à celle des graines de betterave (Catusse, Strub, et al. 2008) et d'*A. thaliana* (Rajjou et al. 2008), à l'exception notable d'une très forte représentativité de protéines impliquées dans le métabolisme du phosphate (Figure 52).

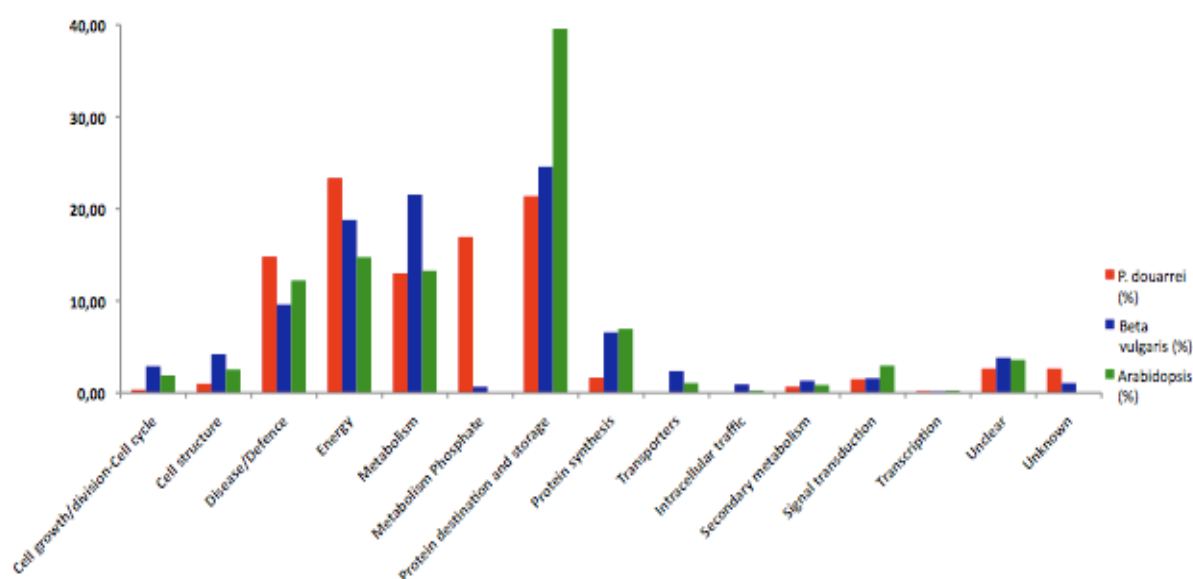


Figure 52 : signatures ontologiques (Bevan et al. 1998) des protéomes des graines matures de *Psychotria gabriellae*, *Beta vulgaris* et *Arabidopsis thaliana* (Zongo 2010).

Ainsi la fraction albumines de la graine de *P. gabriellae* a révélé 608 protéines, dont 103 (soit 17%) impliquées dans le métabolisme du phosphate. Une telle proportion est très inhabituelle chez les graines et ne se retrouve pas chez *P. baillonii*. Fait plus marquant encore, ces 103 protéines font toutes partie de la même famille de protéines, à savoir les protéines « DING », nommées ainsi en raison de leur séquence N-terminale. Les protéines de type DING constituent un ensemble de protéines assez conservées (environ 90% d'homologie dans les séquences) et sont caractérisées par un enchaînement spécifique en leur partie N-terminale de type DINGGG (Figure 53).

DINGGGATLTPQKLYLTPDVLTAGFAPYIGVGSKGKIAFLENKYNQFGTD'TTKNVHWAGSDS  
KLTATELATYAADKEPGWGKLIQVPSVATSVAIPFRKAGANAVDLSVKELCGVFSGRIADWS  
GITGAGRSGPIQVVYRAESSGTTELFTRFLNAKCTTEPGTFAVTTTTFANSYSLGLTPLAGAV  
AATGSDGVMALNDTTVAEGRI TYI SPDFAAPTLAGLDDATKVARVGKGVVNGVAVEGKSPA  
AANVSAAISVVPLPAAADRGNDVWVPVFGATTGGGVVAYPDSGYPI LGFTNLI FSQCYANA  
TQTGQVRDFFTKHYGTSANNDAIEANAFVPLPSNWKA AVRASFLTASNALSIGNTNVCNGK  
GRPQ

Figure 53 : séquence de la protéine *Human Phosphate Binding Protein* (HPBP) (P85173) (Morales et al. 2006). La séquence DINGGG en partie N-terminale est surlignée en jaune.

Les observations de protéines de type DING à partir d'organismes eucaryotes sont nombreuses. L'une d'entre elles a été retrouvée dans le liquide synovial de patient atteints d'arthrose (Berna et al. 2002), et d'autres dans le plasma sanguin (Morales et al. 2006), chez le millepertuis (Perera et al. 2008), la tomate, le tabac, la pomme de terre, la patate douce, le blé ou *A. thaliana* (Berna et al. 2002 ; Berna, Scott, et al. 2009 ; Berna, Bernier, et al. 2009). Cependant, les avis sont très partagés quant à l'appartenance réelle de ces protéines au règne animal ou végétal (Berna et al. 2008). En effet, aucun gène codant ce type de protéine n'a été identifié chez les plantes ou les animaux, en dépit du fait que les protéines y aient été retrouvées et que les séquences de nombreux génomes complets de différentes plantes soient disponibles (Berna et al. 2002 ; Berna, Scott, et al. 2009). En revanche, les séquences codantes de ces protéines DING ont été identifiées chez les procaryotes, notamment *Pseudomonas aeruginosa* et *Thermus thermophilus* (Lewis & Crowther 2005 ; Pantazaki et al. 2008).

L'analyse des séquences révèle que les protéines de type DING de *Pseudomonas* ont une grande similitude avec les protéines de type DING des eucaryotes allant jusqu'à 70 à 80% d'homologie. Elles sont en revanche très distantes des protéines pstS des procaryotes (25 à 30%) (Phosphate binding protein of bacterial ABC cassette phosphate-uptake) également impliquées dans l'homéostasie du phosphate. Une analyse de la séquence de ces différentes protéines indique l'existence d'acides aminés conservés (Berna et al. 2008). Il s'agirait notamment :

- de la partie N-terminale (DINGGG) très conservée chez les eucaryotes et nettement moins chez les procaryotes ; ce caractère est retrouvé par séquençage des protéines DING de *P. gabriellae* ;
- la présence de quatre cystéines conservées chez les eucaryotes, et qui sont responsables de la formation de deux ponts disulfures importants pour la structure

tridimensionnelle des protéines ; pour le moment une de ses cystéine a été identifiée avec certitude chez *P. gabriellae*.

Chez *Pseudomonas fluorescens*, le rôle des protéines de type DING est lié à la concentration et à la séquestration du phosphate (Scott & Wu 2005). Ainsi, les protéines de type DING semblent impliquées dans l'activation de la biosynthèse de la phosphatase alcaline. L'activation de cette dernière est régulée par la concentration en phosphate environnant et nécessite l'intervention d'un système de transporteur de type ABC spécifique du phosphate (Pantazaki et al. 2008). De fait, chez *P. fluorescens*, les protéines de type DING sont des transporteurs de type ABC (X.-X. Zhang et al. 2007).

Chez les organismes eucaryotes, les rôles des protéines de type DING semblent plus divers. Certaines sont capables de se lier à des enzymes. C'est le cas notamment de la « *germin-like-protein binding protein 1a* » capable de se lier à des germes chez *A. thaliana* (Membré et al. 2000), mais aussi à l'oxalate oxydase ou la superoxyde dismutase (Berna et al. 2002 ; Berna, Scott, et al. 2009). Les protéines de type DING peuvent également servir de ligands à d'autres protéines, notamment en agissant comme des chaperons ou co-transporteurs (Berna, Bernier, et al. 2009). Un des éléments les plus intéressants en relation avec le profil minéralogique observé chez la graine de *P. gabriellae* (Zongo 2010) est l'existence d'une protéine DING appelée « *crystal adhesion inhibitor* » chez l'homme (Kumar et al. 2004). Cette protéine empêche la fixation de cristaux d'oxalate de calcium à la surface des cellules épithéliales (Kumar et al. 2004). La protéine DING p27<sup>SJ</sup> du millepertuis (*Hypericum perforatum*) a été identifiée à partir d'une culture de cals (Darbinian-Sarkissian et al. 2006). Il est très intéressant d'observer que cette protéine est capable de supprimer la réplication du VIH-1 (Darbinian-Sarkissian et al. 2006 ; Darbinian et al. 2008) en interagissant avec le facteur de transcription C/EBP $\beta$  et l'activateur viral Tat. L'interaction avec le facteur de transcription C/EBP $\beta$  lui confère également une action de répression sur la réaction inflammatoire (Mukerjee et al. 2008). Par ailleurs, son activité de phosphatase altère l'état de phosphorylation de protéines impliquées dans la transduction du signal. C'est notamment le cas de Erk1/2 dont un niveau de phosphorylation dérégulé par p27<sup>SJ</sup> provoque une perturbation du cycle cellulaire (Darbinian et al. 2009). Les protéines de type DING ont donc un large spectre d'actions biologiques, suscitant des recherches et développements très actifs dans le secteur médical. De plus, chez *A. thaliana*, les protéines DING contenant suffisamment d'histidines pourraient lier les métaux (Membré et al. 2000).

Ainsi, l'identification de telles protéines en forte proportion dans le protéome de l'espèce hyperaccumulatrice *P. gabriellae* en lien avec les transporteurs de type ABC et

l'oxalate de calcium permettait d'envisager une implication de ces protéines dans la formation des profils de répartition minérale tout à fait atypiques des éléments nickel et phosphate observés chez les graines de cette espèce.

En revanche, l'absence de démonstration de la présence de gènes codant les protéines de type DING chez les eucaryotes pose problème. De fait, leur présence dans les graines de *P. gabriellae* pourrait être expliquée par l'existence d'une symbiose avec une/des bactérie(s) endophyte(s). Ces dernières sont naturellement présentes dans les plantes et les graines et notamment dans le genre *Psychotria* (Vega et al. 2005 ; Mendes et al. 2007 ; Ferreira et al. 2008 ; Darsonval et al. 2009 ; Lemaire et al. 2012).

## **1.4 Axes de recherche**

Sur ces bases, le but des travaux de recherche menés lors de la présente thèse était de valider la présence de ces protéines de type DING dans les graines de *P. gabriellae* sur une large période temporelle et à partir de populations différentes, mais également d'étudier et caractériser la flore endophyte de la graine de *P. gabriellae* (Villegente, Zongo et al. 2011).

Par ailleurs, aucune banque de données génomiques (de type EST) n'était disponible pour cette espèce. Afin de mettre en évidence des gènes impliqués dans le mécanisme d'hyperaccumulation de nickel, nous avons entrepris de séquencer les transcrits exprimés à différents stades de maturité des fruits de *P. gabriellae*. Cette approche préliminaire permettra par la suite de disposer d'un ensemble d'outils pour caractériser précisément les gènes impliqués dans le mécanisme de l'hyperaccumulation chez *P. gabriellae*. Elle s'inscrit dans un projet plus global récemment financé par l'agence nationale de la recherche (ANR) appelé « Diversité et génomique de l'hyperaccumulation du nickel chez les plantes ».

# **2 Résultats**

## **2.1 Confirmation de la présence des protéines de type DING dans différents échantillons de graines de *P.***

### ***gabriellae***

La présence des protéines DING chez *P. gabriellae* avait été observée avec un échantillon de graines récolté au Mont Koghi (Zongo 2010). Pour lever l'hypothèse que cette observation résultait d'un artefact, il était important de confirmer la présence de ces protéines DING dans

les graines de *P. gabriellae* en utilisant un plus grand nombre de graines provenant de lieux et de dates de récoltes différents.

Nous disposions au laboratoire de graines récoltées sur la population du Mont Koghi en 2006, 2007, 2009, 2010 et 2011, couvrant ainsi une période de 6 années avec une seule année (2008) manquante. Après récolte, ces graines ont été conservées à 4°C dans des tubes hermétiques. Par ailleurs, deux nouvelles populations ont été rajoutées à l'étude (Figure 55) :

- l'une récoltée au Nord de la Grande-Terre, la population de Nétéa-Poya,
- et une autre récoltée au sud de la Grande-Terre, la population de la forêt Desmazures.

Pour mener à bien nos études, nous avons disposé de deux moyens essentiellement. L'un repose sur l'identification de protéines par approche protéomique. L'autre est basé sur l'utilisation d'anticorps spécifiques. Nous avons ainsi pu bénéficier d'un anticorps spécifiquement développé contre une protéine DING du tabac (*Nicotiana tabacum*) (Perera et al. 2008) que nous avons testé sur les fractions albumines des graines des deux espèces de *Psychotria*, *P. gabriellae* et *P. baillonii* (Figure 54).

Les Western blots obtenus ont permis la révélation d'un groupe de spots (en haut et à droite des gels 2D) spécifiquement reconnus par l'anticorps anti-DING chez *P. gabriellae*, mais absent chez *P. baillonii* (Figure 54). Les spots protéiques correspondant à la zone reconnue par l'anticorps ont été prélevés dans les extraits préparés à partir de *P. gabriellae* et analysés par spectrométrie de masse. Les résultats obtenus ont confirmé l'identification de protéines DING (tableau supplémentaire 7), permettant ainsi d'utiliser cet anticorps pour le reste de nos expérimentations.

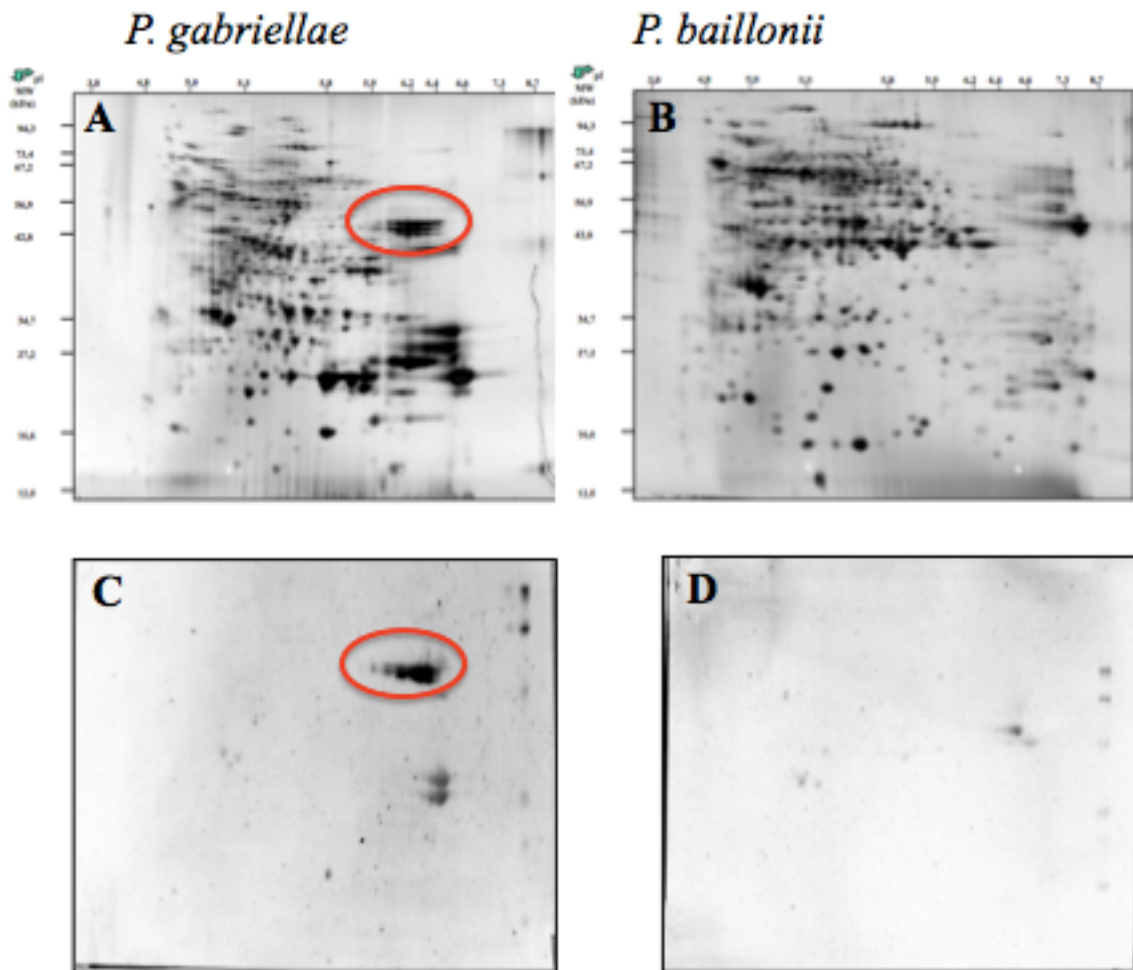


Figure 54 : gels d'électrophorèse bidimensionnelle des fractions albumines des graines matures sèches de *P. gabriellae* (A) et *P. baillonii* (B) colorés à l'argent et après western blot avec l'anticorps anti-DING de tabac (Perera et al. 2008) pour *P. gabriellae* (C) et *P. baillonii* (D). Les spots protéiques spécifiquement reconnus par l'anticorps anti-DING sont entourés par un ovale rouge.

Les extraits des fractions albumines des graines matures sèches de *P. gabriellae* provenant des trois populations de la Grande-Terre furent ainsi analysés avec cet anticorps, puis les protéines contenues dans ces spots furent identifiées par spectrométrie de masse (tableau supplémentaire 7). Pour les trois populations, les spots précédemment observés chez *P. gabriellae* furent de nouveaux reconnus par l'anticorps, et les données de la spectrométrie de masse ont permis de confirmer la présence de protéines DING (Figure 55 ; tableau supplémentaire 7).

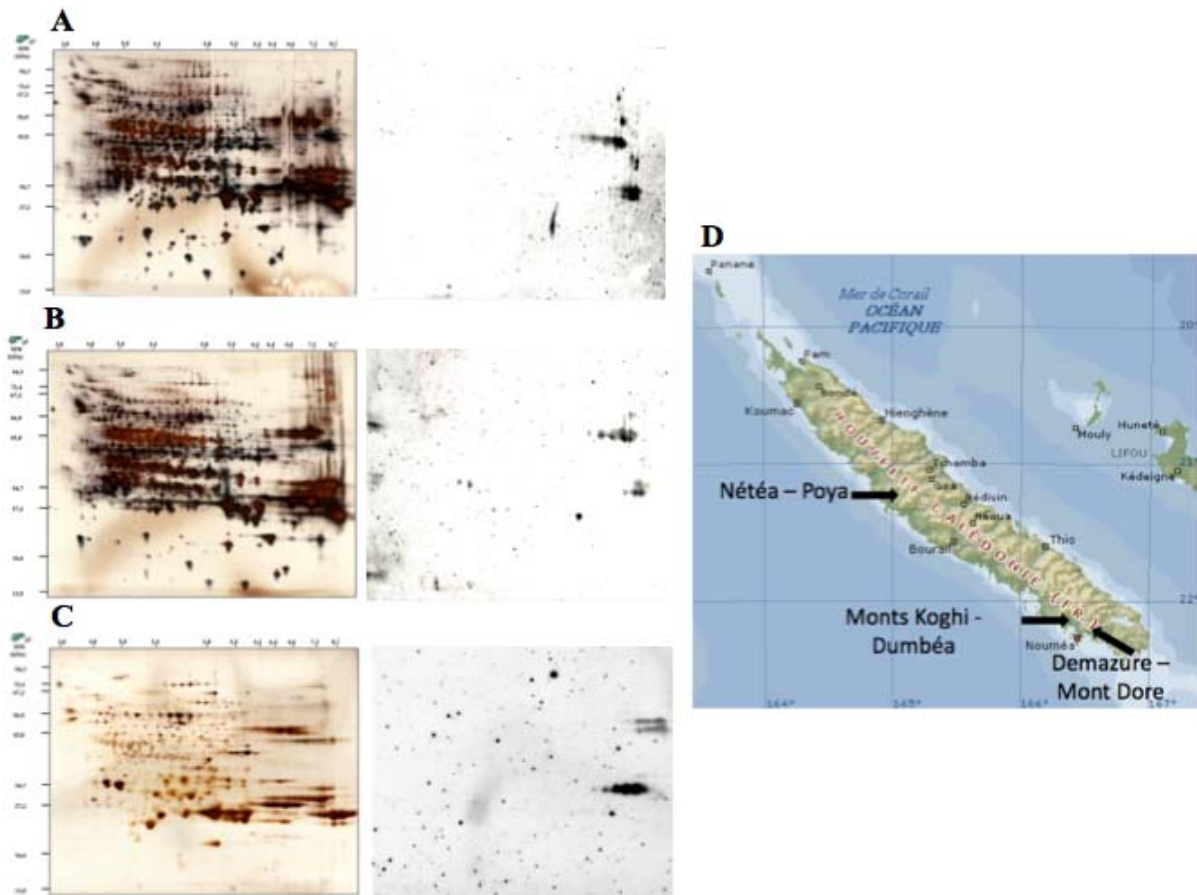


Figure 55 : gels d'électrophorèse bidimensionnelle des fractions albumines des graines matures de *P. gabriellae* de Poya (A), des Monts Koghi (B) et de la forêt Desmazes (C) et western blots correspondants révélés avec l'anticorps anti-DING de tabac (Perera et al. 2008). (D) carte de la répartition des populations sur la Grande-Terre.

Cette même approche a été réalisée pour étudier les graines de *P. gabriellae* récoltées en 2006, 2007, 2009, 2010 et 2011 et provenant de la population des Monts Koghi. Dans tous les cas, les protéines DING ont été retrouvées de 2006 à 2010 (tableau supplémentaire 7). Les peptides obtenus étaient similaires à ceux présents chez la protéine « *Chain A, High-Resolution Structure Of The Ding Protein From Pseudomonas Fluorescens* », ou encore chez la protéine « *GLP-binding protein 1a from Arabidopsis thaliana* ». Les échantillons de l'année 2011 ont été analysés par western blot. La spécificité de l'anticorps ayant été prouvée sur huit échantillons différents.

Nos résultats confirment donc bien la présence des protéines DING dans les graines matures de *P. gabriellae* provenant de trois populations différentes récoltées sur une période de 6 ans.

Afin de caractériser le profil d'accumulation de ces protéines au cours de la formation de la graine, trois stades de développement ont été définis (cf. Matériels et Méthodes). Les extraits des fractions albumines ont alors été analysés par électrophorèse bidimensionnelle (Figure 56). Pour les trois stades étudiés, nous avons constaté que les spots (en haut à droite) correspondant aux protéines DING reconnues par l'anticorps anti-DING de tabac sont toujours présents, et de même intensité, quelque soit le stade de développement. Un stade plus précoce (I) a été étudié mais malheureusement les résultats n'ont pas été exploitables. Il convient de noter qu'à ce stade (I), les graines sont indissociables du fruit. Un extrait des albumines sur le fruit total a été réalisée, mais les gels obtenus n'étaient pas satisfaisants. Nous ne pouvons donc pas à ce jour conclure sur le profil d'accumulation de ces protéines à un stade précoce du développement des graines de *P. gabriellae*.

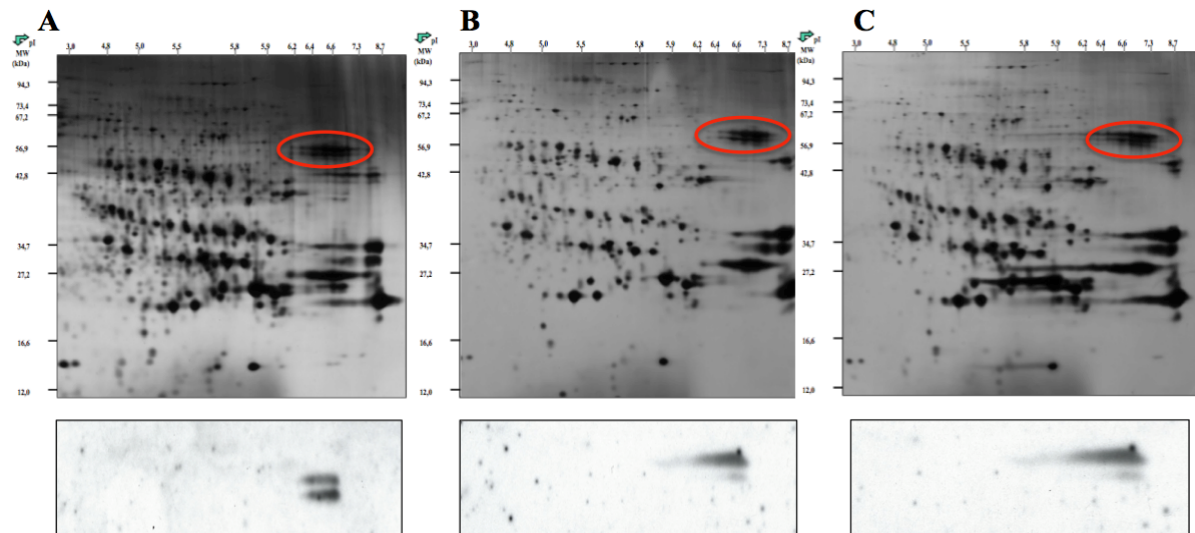


Figure 56 : gels d'électrophorèse bidimensionnelle de graines de *P. gabriellae* aux stades de maturités II (A), III (B) et IV (C) colorés à l'argent. Les spots reconnus par l'anticorps anti-DING de tabac (Perera et al. 2008) sont entourés d'un ovale rouge et la partie d'intérêt des westerns blots est présentée sous les photographies des gels.

## **2.2 Nouvelles données de séquence**

Afin de caractériser plus avant la protéine de type DING de *P. gabriellae*, la séquence de cette protéine a été réexaminée en détails par séquençage *de novo* (collaboration Jean-Marc Strub, CNRS Strasbourg et Maya Belghazi, CNRS Marseille). Les résultats obtenus nous ont permis de mettre en évidence 16 peptides (provenant des deux plateformes d'analyse) matchant avec les séquences de protéines DING connues (Figure 57).



**A**  
**Human Phosphate Binding Protein (HPBP – P85173)**

DINGGGATLPQKLYLTPDVLTAGFAPYIGVSGGKGIKAFLENKYNQFGTDTTKNVHWAGSDS  
 KLTATELATYAADKEPGWGKLIQVPSVATSVAIFFRKAGANAVDLSVKELCGVFSGRIADWS  
 GITGAGRSGPIQVVYRAESSGTTELFTRFLNAKCTTEPGTFAVTTTTFANSYSLGLTPLAGAV  
 AATGSDGVMMAALNDTTVAEGRITYISPDFAAPTLAGLDDATKVARVVGKGVVNGVAVEGKSPA  
 AANVSAATISVVPLPAAADRGNPDVWVPVFGATTGGGVVAYPDSGYPIILGFTNLIFSQCAYANA  
 TQTGQVRDFFTKHYGTSANNDAAIEANAFVPLPSNWKAAVRASFLTASNALSIGNTNVNCNGK  
 GRPQ

**B**

	8	32	63	114	141	159	300	363
PfluDING	DINGGGATLPQA...	VSGSN...	GSDSK...	ELCGV...	VVYRSESSGTTELFTRFLNAKCN	---	SQCYA...	NVCNG
SJW	DINGGGATLPQA...	VSGSN...	GSDSK...	ELCGV...	VVYRSESSGTTELFTRFLNAK	---	---	---
HPBP	DINGGGATLPQK...	VSGSK...	GSDSK...	ELCGV...	VVYRAESSGTTELFTRFLNAKCT	---	SECYA...	NVCNG
Pa14	DINGGGATLPQK...	VSGSN...	GSDSK...	TLCGV...	VVYRSESSGTTELFTRFLNACSS	---	SQCYA...	NVCGI
St	DINGGGATLPQK...	VSGSK...	GSDSK...	QLCGV...	VVYRNEASSGTTELFTRFLAAKCV	---	---	---
Lm	DINGGGATLPQP...	VSGSK...	GSDSK...	QLCGV...	VVYRNEASSGTTELFTRFLAAKCV	---	SQCYA...	---
Pf5/1	DINGGGATLPQP...	VSGSN...	GSDSK...	DLCGV...	VVYRNESSGTTELFTRFLNAKCA	---	SQCYA...	SVCNA
Pf5/2	DINGGGATLPQS...	SDLGD...	ATESK...	DVCGI...	VVYRQGGSGTTELFTRFLNAKCS	---	SQCYA...	NICNA
CAI	DINGGGATLPQP...	---	LTSTE...	QLCGV...	---	---	SQCYA...	---
Pa01/1	-VTGGGASLPAE...	TGSGT...	GSDSV...	QLCDA...	IVYRTGSSGTTELFTRHLNSICPT	---	GQCYK...	NVCNT
Pa01/2	-VTGGGASMPAK...	SGSGI...	GSDSI...	QLCDA...	IVYRGGSSGATELLTRHLNSVCPT	---	GQCYK...	SVCNG
Pf01	NVTGGGATLPAA...	TGSGT...	GSDSI...	QLCDA...	VVYRNVSSGTTEILTRHLNSICPT	---	GQCYK...	SVCNA
Rat	PVTGGGASLPPD...	TGSGT...	GSDSV...	QICGI...	VVYRGESSGTELLTRFLTSACQP	---	---	---
PpKT	-LTGSGASFPAP...	KSGSA...	ASDSA...	GKITQ...	VV-RADSSGTTAVFTKHLSAINA	---	RKDNQ...	---
	PP	P	P	C	P	PPP	C	C

Figure 57 : (A) séquence peptidique de la protéine de type DING humaine (HPBP) (P85173). Les peptides expérimentaux obtenus par Maya Belghazi (Marseille) et Jean-Marc Strub (Strasbourg) à partir du spot « DING » majoritaire reconnu par l'anticorps anti-DING de tabac chez *P. gabriellae* sont surlignés en vert. Les acides aminés conservés impliqués dans la liaison du phosphate sont en rouge, les cystéines impliquées dans la formations des deux ponts disulfures sont en bleu (Berna et al. 2008). (B) Figure extraite de Berna et al (2008) montrant l'alignement entre la protéine DING de *Pseudomonas* et les protéines DING caractérisées chez eucaryotes. Les acides aminés conservés impliqués dans la liaison du phosphate sont en rouge et signalés par la lettre « p ». Les cystéines conservées sont en vert et signalées par la lettre « c ». Les séquences bactériennes sont Pa01/1 : *P. aeruginosa* (PA0688). Pa01/2 : *P. aeruginosa* (PA0689). Pf01: *P. fluorescens* (ABA75910.1). PfluDING : DINGsequence from *P. fluorescens* SBW25 (GL019191). PA14 : *P. aeruginosa* (ZP\_00138283). Pf5/1 : *P. fluorescens* (AAY2032.1). Pf5/2 : *P. fluorescens* (AAY92033.1). PpKT : *P. putida* (PP2656). Les séquences de DING eucaryotes sont HPBP (Morales et al. 2006), CAI: partial sequence of crystal adhesion inhibitor (Kumar et al. 2004), SJW: *Hypericum perforatum* truncated DING protein (Darbinian-Sarkissian et al. 2006), Lm: translated partialDNAsequence from *Leishmania major* (Contig2206 and Contig2072;www.sanger.ac.uk), Rat: translated partial DNA sequence from *Rattus norvegicus* (GenBank AC135282.3), St: *Solanum tuberosum* sequence composed of two overlapping genomic DNA clones and the N-terminal protein sequence.

Les données de séquençage *de novo* ont permis de couvrir à 52% la séquence de la protéine DING humaine (HPBP) à partir des peptides obtenus par séquençage de la protéine DING de *P. gabriellae*. Ainsi, sept des huit acides aminés impliqués dans la liaison avec le phosphate ont été séquencés, une des quatre cystéines également et l'extrémité Nterminale contenant la séquence consensus DINGGG- a été mise en évidence.

Par ailleurs, ces résultats proviennent de deux plateformes d'analyses différentes et les analyses furent réalisées par deux expérimentateurs différents. Ces données permettent donc de conforter l'identification de la protéine DING chez *P. gabriellae* et d'éliminer la possibilité d'une contamination technique au moment du séquençage.

### **2.3 Caractérisation de la flore endophyte de *P. gabriellae***

L'origine des protéines de type DING chez les organismes eucaryotes fait l'objet d'une forte controverse. Il en est ainsi de la *Human Phosphate Binding Protein* (P85173) pourtant retrouvée en abondance dans le plasma sanguin humain (Morales et al. 2006). Son abondance y est d'ailleurs telle que la séquence protéique complète de la protéine a été obtenue par séquençage MS-MS et que la protéine a pu être cristallisée permettant d'établir sa structure tridimensionnelle (Diemer et al. 2008). En dépit de ces résultats, cette protéine est actuellement classée dans la base de données Uniprot ([www.uniprot.org](http://www.uniprot.org)) comme provenant d'une contamination par un organisme procaryote inconnu.

Ainsi, dans notre cas, l'hypothèse d'une origine procaryotique des protéines DING retrouvées chez *P. gabriellae* ne pouvait être écartée. Une telle origine pouvait refléter une contamination artéfactuelle mais également l'existence d'une coopération, voire d'une symbiose entre de telles bactéries et les cellules végétales, pouvant rendre compte de la tolérance au nickel de *P. gabriellae*. L'identification de bactéries endophytes (naturellement présentes dans les tissus végétaux) a donc été entreprise.

Le protocole que nous avons développé à Nouméa avait pour objectif d'extraire, de cultiver, d'identifier des bactéries endophytes des graines de *P. gabriellae* et de tester la présence des protéines de type DING chez de telles bactéries. A noter que ce protocole ne permettait d'obtenir que des bactéries endophytes cultivables. Trois souches bactériennes ont été isolées et caractérisées par séquençage de leur ARN 16S (cf. Matériels et Méthodes).

Ces trois souches sont :

- *Staphylococcus pasteurii*,
- *Methylobacterium populi*, et
- *Bacillus cereus*.

Ces bactéries ont été cultivées en milieu liquide en absence ou en présence de 100 ppm de sulfate de nickel puis les protéines totales solubles ont été extraites, analysées par électrophorèse bidimensionnelle et testées pour leur réactivité vis-à-vis de l'anticorps anti-DING de tabac reconnaissant les protéines de type DING de *P. gabraellae*. Un seul de ces résultats est montré en (Figure 58).

Par ailleurs, les bactéries se sont parfaitement développées dans le milieu de culture contenant 100 ppm de nickel. Ainsi, les bactéries semblent tolérer sans souci cette concentration de l'élément métallique.

L'anticorps n'a montré aucune réactivité spécifique dans la zone des poids moléculaires et points isoélectriques des protéines de type DING des graines de *P. gabraellae* (Figure 54). Toutefois, nous avons noté la présence d'un spot (en forme de doublet) de faible poids moléculaire dans les extraits réalisés à partir de *Methylobacterium populi* (Figure 58). Le spot correspondant à l'emplacement reconnu par l'anticorps a été prélevé dans un gel préparatif de protéines extraites de *Methylobacterium populi* et analysé par spectrométrie de masse. Les protéines identifiées sont une *formaldehyde activating enzyme* et une *YbaK/poly-tRNA synthetase associated protein* toutes deux de l'espèce *Methylobacterium populi*, validant ainsi au passage notre identification de ces bactéries par séquençage de l'ARN 16S. Ils indiquent également l'absence de protéines canoniques de type DING, pourtant reconnues par l'anticorps anti-DING de tabac.

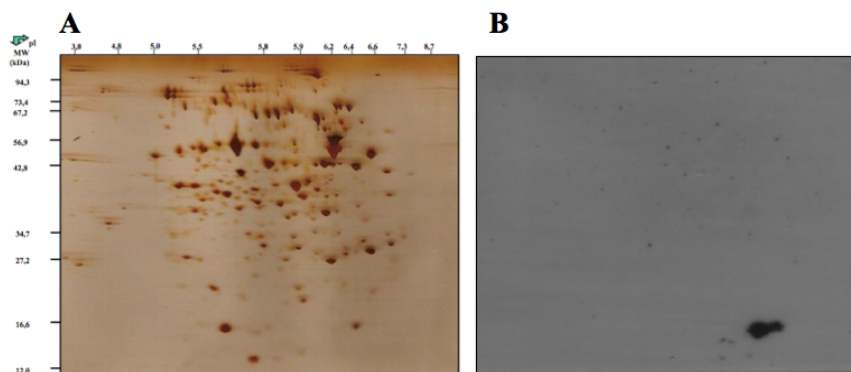


Figure 58 : gel d'électrophorèse bidimensionnelle (A) de l'extrait des protéines totales de *Methylobacterium populi* et western blot correspondant (B) révélé en utilisant l'anticorps anti-DING de tabac (Perera et al. 2008).

Par ailleurs, pour approfondir cette question sur la présence de bactéries au sein des graines de *P. gabriellae*, une collaboration a été initiée avec le Dr. Marie-Agnès Jacques de l'Institut de Recherche en Horticulture et Semences (INRA-Angers), équipe Emersys (Emergence, systématique et écologie des bactéries phytopathogènes). Ce laboratoire est notamment spécialisé dans l'extraction et l'identification de bactéries endophytes. Leur protocole consiste à extraire les bactéries en faisant macérer les tissus végétaux puis l'ARN 16S est séquencé afin de pouvoir identifier les bactéries présentes. Cette approche nous permettait ainsi d'isoler la présence éventuelle de bactéries non cultivables chez les graines de *P. gabriellae*, en complément des bactéries cultivables que nous avons isolées à Nouméa (cf. ci-dessus). A cette fin, des graines de *P. gabriellae* ont été envoyées au laboratoire d'Angers. Cette approche a ainsi permis l'identification d'un seul genre bactérien, correspondant à l'ordre des *Burkholderiales*. Les résultats de séquences et de BLAST correspondants sont présentés en annexe 6.

## **2.4 Résultats préliminaires de transcriptomique**

Les données acquises durant ce travail de thèse constituent les prémices d'un projet à grande échelle visant à caractériser le plus grand nombre de transcrits provenant des racines, des feuilles et des graines de *P. gabriellae*. Dans ce but, un mélange de fruits de tous les stades a été réalisé et les transcrits ont été extraits et analysés dans le cadre du projet 1kP (cf. « Matériels et Méthodes »).

Ces données ont été analysées par comparaison avec celles disponibles pour *A. thaliana*. Ainsi, 18 242 séquences de *P. gabriellae* ont pu être identifiées avec une e-value de  $1^{e-5}$  par comparaison avec des protéines d'*A. thaliana* (cf. Matériels et méthodes). Ces 18 242 correspondent en fait à 9 661 gènes uniques.

Les résultats ont été obtenus très récemment et il ne nous est pas encore possible de présenter ici une analyse exhaustive du transcriptome des fruits de *P. gabriellae*. Dans un premier temps, nous avons choisi d'étudier les transcrits retrouvés au minimum 10 fois lors du séquençage, correspondant donc aux transcrits les plus abondants. Ce seuil a permis d'identifier 23 gènes (Tableau 8).

Tous les gènes retrouvés plus de dix fois sont des gènes de ménage (*housekeeping genes*) impliqués notamment dans la dégradation des protéines par des protéases (subtilase) ou via le protéasome (*polyubiquitin*, *ubiquitin-protein ligase*), la traduction des ARN messagers (*elongation factors* 1 et 2) et l'assemblage correct des protéines (HSP 70 et HSP 81) ou encore le métabolisme des acide gras (lipoxygénase). Il est intéressant de noter

l'identification d'un transporteur de type ABC (AT1G15520) dont les transcrits ont été retrouvés 10 fois.

De même, une analyse par BLAST a été réalisée afin d'essayer d'identifier des transcrits codant les protéines de type DING de *P. gabriellae* sur la base des peptides identifiés par spectrométrie de masse. Cette recherche n'a donné aucun résultat positif.

A ce stade, nous ne pouvons conclure sur les gènes impliqués dans l'hyperaccumulation de nickel chez *P. gabriellae*. Il conviendra d'attendre les résultats d'analyses plus exhaustives des transcrits (gènes spécifiques, classification ontologiques des fonctions des gènes). Ces expériences sont en cours en collaboration avec le Dr. Sylvain Merlot (CNRS Gif-sur-Yvette) au sein du projet ANR EvoMetoNicks.

Tableau 8 : gènes (23) dont les transcrits issus d'un mélange de fruits de *P. gabriellae* ont été retrouvés plus de dix fois lors du séquençage du transcriptome.

Nombre de hits	Locus <i>A. thaliana</i>	Description
23	AT5G03240	similar to polyubiquitin (UBQ10) (SEN3)
21	ATMG01360	COX1   cytochrome c oxidase subunit 1
18	AT5G60390	elongation factor 1-alpha / EF-1-alpha
17	AT1G21680	Putative uncharacterized protein
13	AT2G05920	subtilase family protein (similarity to cucumisin-like serine protease)
13	AT3G12110	actin 11 (ACT11)
13	AT3G12500	basic endochitinase
13	AT3G51550	protein kinase family protein, contains protein kinase domain
13	AT5G52640	heat shock protein 81-1 (HSP81-1)
12	AT1G60710	aldo/keto reductase family protein
12	AT2G07727	cytochrome b (MTCYB) (COB) (CYTB)
12	AT3G12580	HSP70   heat shock protein 70, putative
12	ATCG00860	YCF2   hypothetical protein
11	AT1G15690	pyrophosphate-energized vacuolar membrane proton pump
11	AT1G55020	lipoxygenase (LOX1)
11	AT1G56070	elongation factor 2, putative / EF-2, putative
11	AT5G54160	quercetin 3-O-methyltransferase 1
10	AT1G15520	PDR12   ABC transporter family involved in resistant to lead
10	AT1G55860	ubiquitin-protein ligase 1 (UPL1)
10	AT1G60420	DC1 domain-containing protein
10	AT1G68570	proton-dependent oligopeptide transport (POT) family protein
10	AT1G77120	Symbol: ATADH Alcohol dehydrogenase class-P
10	AT3G25800	serine/threonine protein phosphatase 2A (PP2A)

## 3 Discussion

### 3.1 Les protéines de type DING chez *P. gabriellae*

Comme nous l'avons vu, l'origine des protéines de type DING fait l'objet d'une vive discussion dans la communauté scientifique (Lewis & Crowther 2005 ; Berna, Scott, et al. 2009). Alors que certains aspects structuraux semblent spécifiques des protéines DING de type eucaryote (Berna et al. 2008), aucune protéine de type DING n'est référencée dans les bases de données comme provenant de gènes d'organismes eucaryotes, et cela, en dépit du fait que des protéines de type DING abondantes aient été identifiées dans des tissus ou milieux de cultures *a priori* sains, tels que le plasma humain (Morales et al. 2006), un milieu de culture cellulaire (Perera et al. 2008) ou encore des souris totalement dépourvues de germes bactériens (lignée *germ free*) (Collombet et al. 2010).

Ainsi, afin de s'assurer de la reproductibilité des résultats obtenus précédemment au laboratoire chez *P. gabriellae* montrant que les graines récoltées au Mont Koghi en 2009 contenaient une grande quantité de protéines de type DING (Zongo 2010), plusieurs récoltes additionnelles ont ici été étudiées.

Nos résultats montrent sans ambiguïté la présence de ces protéines sur une période de six ans pour la population des Monts Koghi (tableau supplémentaire 7). Par ailleurs, la présence de ces protéines a été confirmée dans des graines provenant de deux autres populations du Nord et du Sud de la Grande-Terre (Figure 55 ; tableau supplémentaire 7), les récoltes ayant été effectuées par des personnes différentes.

Ainsi, si l'origine des protéines de type DING dans la graine de *P. gabriellae* n'est toujours pas définie, leur présence est clairement établie et semble corrélérer un mécanisme d'adaptation. Par ailleurs, nos résultats indiquent que des protéines de type DING sont présentes dès le plus jeune stade de développement étudié (Figure 56). Nous pouvons conclure que la détection des protéines de type DING à partir de la graine de *P. gabriellae* est reproductible et ne peut s'expliquer par une contamination artéfactuelle du lot de graines utilisé dans les expériences précédentes (Zongo 2010).

Bien que les gènes codant les protéines de type DING chez les eucaryotes n'aient pas encore été découverts, il est possible qu'ils existent malgré tout dans ces génomes. Le fait de ne pas encore avoir découvert ces gènes dans les génomes *a priori* complets d'organismes modèles pourrait s'expliquer par le fait que ces gènes sont i) très morcelés (fort morcellement en introns et exons), ou que ii) les transcrits aboutissant aux protéines de type DING résultent d'un mécanisme de transcription complexe (eg. transcription antisens) (Lee et al. 1999 ;

Beiter et al. 2009 ; Katayama et al. 2005). En ce sens, les données nouvelles obtenues par séquençage de novo (Figure 54) permettraient de faciliter la recherche de gènes codant ce type de protéines.

En outre, de plus en plus de gènes sont nouvellement annotés grâce aux informations apportées par les données de protéomique (Ansong et al. 2008 ; Castellana et al. 2008). Ainsi, il pourrait être envisageable que des gènes codant les protéines de type DING existent chez les eucaryotes mais ne soient pas encore correctement annotés.

Outre cette hypothèse, nous avons voulu étudier la présence de bactéries endophytes qui seraient naturellement présentes dans les graines de *P. gabriellae*. Trois bactéries endophytes cultivables ont été identifiées au sein de la graine mature sèche. Une recherche dans les banques de données disponibles sur internet ne nous a pas permis d'identifier une protéine de type DING chez ces bactéries. Cependant, les génomes complets de ces bactéries n'étant pas disponibles, nous ne pouvons pas conclure de ce point de vue à la présence ou pas de gènes codant ces protéines chez ces bactéries. Toutefois, nos analyses par approche protéomique couplée à l'utilisation d'un anticorps spécifique n'ont pas révélé la présence de protéines DING chez ces bactéries. Or, nous avons démontré que l'anticorps utilisé reconnaît spécifiquement les protéines de type DING de *P. gabriellae* (Figure 54). Ainsi, il semble que les bactéries endophytes cultivables que nous avons isolées n'expriment pas de protéines de type DING. Il convient toutefois de rester prudent car nos conditions de culture des bactéries n'étaient peut-être pas appropriées à la production de telles protéines.

L'équipe du Dr. Marie-Agnès Jacques de l'INRA d'Angers a réussi à identifier une seule souche bactérienne du genre des *Burkholderiales*, par une approche basée sur l'isolement de l'ensemble des acides nucléiques présents dans les graines mures sèches de *P. gabriellae*. De manière intéressante, ce genre bactérien est connu pour être endophyte des plantes, et notamment de *Psychotria* (Lemaire et al. 2012). Il s'agit toutefois de la première identification de *Burkholderiales* endophytes en Nouvelle-Calédonie. Deux espèces de *Burkholderiales* ont été décrites comme fixatrices d'azote, *B. vietnamiensis* et *B. tropica*, bien que d'autres espèces du genre en seraient capables (Santos et al. 2001 ; Reis et al. 2004). Ainsi, l'identification de ces bactéries dans les graines de *P. gabriellae* pourrait être un indice d'une symbiose avec la plante, permettant la fixation d'azote. Connaissant la pauvreté en macroéléments des sols ultramafiques, cette symbiose serait un avantage considérable au développement de la plante.

Nos résultats ont ainsi permis de confirmer clairement la présence de protéines de type DING chez les graines de *P. gabriellae* sans toutefois clairement statuer sur leur origine. Nos observations suggèrent l'existence d'une flore bactérienne intéressante d'un point de vue de la symbiose pour l'acquisition de l'azote.

### **3.2 Les données préliminaires de transcriptomique**

Le but essentiel de ces travaux était de procurer des données en masse permettant d'expliquer sans *a priori* l'étonnante accumulation de nickel chez *P. gabriellae*, notamment dans les graines. Cet objectif a été atteint puisqu'en effet notre protocole d'extraction des ARNm a permis d'identifier 18 242 transcrits par comparaison avec le génome modèle d'*Arabidopsis thaliana*. Ces résultats seront pleinement exploités dans le cadre du projet EvoMetoNicks récemment soutenu par l'ANR.

Il est toutefois intéressant de noter que parmi les transcrits retrouvés au minimum dix fois dans le transcriptome, un transporteur de type ABC a été identifié. AtPDR12 a tout d'abord été décrit comme étant impliqué dans la régulation de l'homéostasie du plomb chez *Arabidopsis* (Lee et al. 2005). Cette étude a mis en évidence un efflux de plomb associé à la présence du transporteur et dont l'expression semblait induite par le métal lui-même. Cependant, plus récemment, AtPDR12, renommé ABCG40, a été décrit comme transporteur transmembranaire de l'ABA (Kang et al. 2010 ; Kuromori & Shinozaki 2010). Cette étude a montré qu'il n'y avait pas de compétition entre l'ABA et le plomb au niveau de ce transporteur. Ainsi, le rôle dans la tolérance au plomb de ce transporteur ABC viendrait de son action sur la voie de signalisation de l'ABA. Un défaut dans son expression entraînerait notamment un ralentissement dans les mécanismes de réponse aux stress, notamment hydrique via la fermeture des stomates. La présence de ce transporteur a déjà été décrite dans des graines, bien qu'à plus faible intensité en comparaison des tissus végétatifs (Kang et al. 2010). Cette expression est en accord avec l'action prépondérante de l'ABA lors de la maturation des graines et du maintien de la dormance. Sa forte représentativité en terme d'abondance de transcrits dans la graine de *P. gabriellae* (retrouvé dix fois) pourrait résulter d'un rôle de l'ABA, peut être en relation avec la mise en place de mécanismes de tolérance aux métaux lourds. En effet, l'ABA régule le niveau d'expression d'une métallothionéine, impliquée dans la chélation des métaux (Chatthai et al. 1997). Cette hormone serait également impliquée dans l'expression de différents gènes pouvant contribuer à la mise en place de la



tolérance aux métaux lourds tels que le plomb et le cadmium (Talanova et al. 2000 ; Kang et al. 2010).

Par ailleurs, nos données montrent également l'identification d'un transcrit codant une protéine nommée « *High-affinity nickel-transport family protein, putative* » (G7JMC5) (données non montrées), de trois transcrits impliqués dans le transport de cuivre (Q10ST6), une métallothionéine (Q9FUJ8), d'un transporteur du zinc (C4LVL2), d'une perméase à fer/zinc (F6DER4), d'une « *Heavy metal transport/detoxification superfamily protein* » (O49613). Ainsi, ces résultats sont très prometteurs pour l'identification et la caractérisation fonctionnelle de gènes impliqués dans la tolérance aux métaux lourds.

Dans l'avenir, il conviendra de caractériser plus en détails les transporteurs de métaux lourds possiblement découverts, ainsi que tout le cortège de chélateurs possibles et de protéines de détoxification. Néanmoins, nos données illustrent la force d'une approche transcriptomique pour mettre en évidence sans *a priori* les mécanismes de tolérance aux métaux lourds développés par les plantes, et cela, même pour une espèce pour laquelle très peu de données génomiques sont encore disponibles.

Par ailleurs, une recherche préliminaire par BLAST entre les transcrits de *P. gabiellae* et les « peptides DING » identifiés par spectrométrie de masse a été réalisée. Aucune identification positive n'a été obtenue. Il convient dans ce cadre de rappeler que les transcrits séquencés ne sont que ceux d'origine eucaryote (sélection polyA). Ainsi, dans l'hypothèse d'une symbiose avec un organisme procaryote, les transcrits de ce dernier ont été perdus.

## 4 Conclusion et perspectives

Les mécanismes d'adaptation développés par les plantes pour résister à un stress en particulier sont très souvent le fruit d'une régulation complexe de différentes voies préexistantes et qui interagissent pour permettre à l'organisme de tolérer le stress sans que l'effort consacré à la survie soit trop pénalisant.

L'adaptation des plantes aux fortes concentrations en métaux lourds, et particulièrement au nickel, fait l'objet d'une attention grandissante pour faire face aux problèmes de pollution des sols et également aborder la question du « *phytomining* », c'est-à-dire l'extraction de métaux lourds des sols en vue de leur concentration et réutilisation.

La Nouvelle-Calédonie, qui possède une flore parmi les plus riches en plantes hyperaccumulatrices de nickel, est un site d'étude à privilégier pour mieux caractériser les mécanismes adaptatifs développés pour faire face au stress nickélicifère.

*P. gabriellae* stocke 18 000 ppm de nickel dans ses feuilles, et jusqu'à 9 000 ppm dans ses graines. Ces dernières semblent avoir mis au point un mécanisme spécifique de protection de l'embryon en concentrant le nickel dans les parties externes de la graine (Zongo 2010). Ainsi, ce mécanisme actif étudié par une approche protéomique en 2009 avait abouti à l'observation très originale d'une forte proportion de protéines impliquées dans le métabolisme du phosphate, les protéines de type DING.

L'identification de ces dernières et l'incertitude régnant quant à l'origine de leur appartenance (procaryote/eucaryote) nous ont conduit à confirmer la présence systématique de telles protéines en utilisant des graines de *P. gabriellae* récoltées sur une période de temps allant de 2006 à 2011 ainsi que dans trois populations différentes distantes d'environ 200 km. Par ailleurs, l'hypothèse d'une activité symbiotique avec des bactéries pour la production de ces protéines de type DING a été envisagée. Ainsi, nous avons caractérisé quatre genres bactériens endophytes de la graine de *P. gabriellae*. Aucun d'eux ne semble impliqué dans la biosynthèse de protéines DING. Toutefois, leur identification a permis la mise en évidence des bactéries susceptibles de tolérer de fortes concentrations en nickel. Cette observation a déjà été reportée pour d'autres espèces, et il conviendrait d'étudier plus en détail ces bactéries. Il s'agit également de la première observation du genre des *Burkholderiales* endophytes en Nouvelle-Calédonie.

Nos travaux ont permis de confirmer la présence des protéines de type DING dans la graine de *P. gabriellae*. Le rôle connu des protéines DING dans l'interaction avec les transporteurs de type ABC ainsi qu'avec les éléments minéraux, et notamment le phosphore reste intrigant, notamment, en lien avec l'aspect très particulier de la répartition des éléments minéraux dans la graine de *P. gabriellae*.

L'étude de stades de développement plus précoces que ceux que nous avons analysés sera nécessaire afin de mettre en évidence l'évolution du profil d'accumulation de ce type de protéines DING. Cet aspect serait combiné avec la caractérisation de l'évolution de la flore endophyte de la fécondation à la graine mature sèche. De même, un séquençage approfondi de ces protéines pourrait permettre de couvrir au maximum sa séquence peptidique permettant ainsi d'étudier les traits spécifiques aux « DING eucaryotes » et de les repérer plus facilement dans les génomes des plantes, notamment chez *P. gabriellae*.

L'étude de transcrits issus des fruits de *P. gabriellae* amène des données très encourageantes indiquant une forte représentation des transcrits impliqués dans l'homéostasie des métaux lourds, ainsi que dans leur chélation. Une analyse plus approfondie reste à faire, mais nous sommes confiants que la qualité du matériel que nous avons produit permettra d'avancer rapidement sur ces questions. Il s'agit là en particulier de la construction d'une base de données génomiques qui permettra également par la suite d'étudier le profil d'expression de gène d'intérêts. Cet outil permettra également d'envisager la possibilité de clonage de gènes d'intérêt afin de pouvoir caractériser leur rôle fonctionnel.



## **Conclusion générale et perspectives**



La Nouvelle-Calédonie est une source extraordinaire de biodiversité à ciel ouvert qu'il convient de protéger et d'étudier, afin de permettre sa préservation et l'utilisation de ses ressources. Ce travail de thèse s'est intégré aux travaux de recherches déjà en cours ou récemment initiés au sein de LIVE en collaboration avec l'IAC et a permis d'affiner la compréhension de certains mécanismes de germination et d'hyperaccumulation à l'échelle moléculaire.

## 1. *Amborella trichopoda*

La caractérisation du protéome de la graine d'*Amborella trichopoda* nous a permis de développer ou de renforcer des collaborations avec nos partenaires en France et aux Etats-Unis d'Amérique, notamment, de participer à l'important projet de séquençage du génome de cette espèce.

Tout d'abord, nos résultats nous ont permis de caractériser les gènes codant les globulines 11S d'*Amborella*. Ces gènes ont fait l'objet d'études approfondies chez d'autres espèces modèles et notre étude comparative a permis de mettre en évidence la présence de caractères ancestraux et de caractères plus évolués pour les globulines 11S d'*Amborella*. Nos données corroborent ainsi la place d'*Amborella* à la base de l'arbre phylogénétique des angiospermes, à la scission avec les gymnospermes.

Notre étude s'est également intéressée à caractériser le protéome de la graine mature sèche d'*Amborella*. En effet, en plus d'être la sœur de toutes les plantes à fleurs, la graine d'*Amborella* a la particularité de présenter une dormance morphologique. Ainsi, au moment de sa dispersion, l'embryon présent dans la graine d'*Amborella* est rudimentaire et ce caractère contraste avec l'aspect développé des embryons de la majorité des autres graines. Nombreuses sont les questions concernant les embryons rudimentaires, notamment : quel est leur stade de maturité ? quel(s) programme(s) gouverne(nt) l'aboutissement de leur développement permettant la germination ?

Tout d'abord, la caractérisation du contenu protéique des tissus a utilisé la technique de protéomique *shotgun*. Nous avons ainsi réussi à caractériser le premier protéome d'un embryon rudimentaire isolé de son albumen. Nos données ont permis de mettre en évidence la présence de tout un cortège de protéines chaperons (HSPs, LEA) ainsi que de toute la gamme enzymatique nécessaire à l'utilisation des réserves stockées au cours de la formation de la graine. Ces données suggèrent que malgré son aspect rudimentaire, l'embryon d'*Amborella* est mature d'un point de vue moléculaire au moment de la dispersion de la graine. La

présence de protéines chaperons suggère également que la graine d'*Amborella* est tolérante à la dessiccation.

Par ailleurs, l'étude des protéines structurant les corps lipidique, les oléosines, a encore une fois permis de mettre en évidence des caractères ancestraux et évolués, confortant la position phylogénétique d'*Amborella*.

De même, l'identification d'une protéine « Mother of TF and TFL1 » suggère la présence de mécanismes de régulation hormonale impliquant l'acide abscissique et l'acide gibbéréllique et largement documenté chez les espèces modèles.

La caractérisation du protéome de l'albumen de la graine nous a également permis d'illustrer avec force l'importance de ce tissu comme lieu de stockage des réserves, notamment lipidiques et protéiques dans le cadre d'*Amborella*, tout en suggérant une gamme protéique dynamique moins importante que dans l'embryon. Cependant, la forte représentativité des protéines de réserve a pu masquer certaines protéines du métabolisme dans l'albumen.

Nous nous sommes ensuite intéressés à caractériser l'évolution du protéome de la graine en cours de germination/imbibition afin de caractériser les processus contrôlant la « croissance interne » de l'embryon dans la graine. A cette fin, nous avons utilisé une approche d'analyse bidimensionnelle des fractions protéiques de la graine entière. Cette approche fractionnée nous a permis de nous affranchir de la forte représentativité des protéines de réserve dans la graine totale qui aurait pu masquer la dynamique fonctionnelle des protéines du métabolisme.

Nous avons ainsi pu mettre en évidence une réactivation globale du métabolisme cellulaire, comme observé couramment lors de la germination. Nos données confirment également l'importance du cycle des méthylations lors de la germination notamment en permettant la réparation du protéome, indispensable à la vigueur germinative.

En outre, nos données indiquent également une mobilisation massive des protéines de réserve avant la fin de la germination *sensu stricto* qui semble indépendante de l'embryon et de l'acide gibbéréllique. Cette mobilisation est habituellement observée après la germination une fois que la radicule (le plus souvent) a traversé les téguments. Ainsi, nos données remettent en cause l'applicabilité de la définition de germination pour les graines à petits embryons et soulèvent des questions quant aux dialogues moléculaires entre l'embryon et l'albumen au moment de l'imbibition.

En perspective de ces résultats, et dans un avenir proche, nous allons réaliser l'analyse du protéome de l'albumen débarrassé des protéines de réserve, dans le but de mettre en



évidence plus de protéines du métabolisme qui auraient pu être masquées lors de la précédente analyse. De même, il conviendra de caractériser de manière précise l'origine et la nature du réseau de polysaccharides entourant l'embryon, et ce notamment grâce à une approche par microscopie à effet tunnel au cours de la formation et de la germination de la graine.

A plus longue échelle, il conviendra de caractériser les mécanismes de régulation hormonaux de la germination. Le génome séquencé est un formidable outil à notre disposition, notamment pour réaliser une étude à grande échelle du positionnement phylogénétique de certaines familles protéiques (telles que les globulines 11S, la PIMT ou encore les oléosines). La recherche des gènes codant les voies de biosynthèse de l'acide abscissique et de l'acide gibbérellique est également possible et serait à réaliser. L'étude des profils d'accumulation de ces hormones devrait également être réalisée au cours de l'imbibition/germination de la graine.

Par ailleurs, le caractère autonome de l'albumen vis-à-vis de l'embryon, répondant simplement à l'imbibition pour enclencher le processus de mobilisation des réserves est intrigant. Cet aspect devra être validé et caractérisé plus finement car il suggère une dépendance de l'embryon envers l'albumen pour l'utilisation des protéines de réserve. Ainsi, l'autonomie de l'embryon pour son propre développement devrait aussi être caractérisé, notamment par des expériences de culture *in vitro* d'embryons isolés.

Nos données suggèrent également une discordance entre les marques moléculaires de la fin de la germination (utilisation massive des globulines 11S) et le marqueur visuel couramment utilisé (percée des enveloppes par une partie de l'embryon). Ainsi, il serait intéressant d'étudier le profil d'évolution des globulines 11S chez d'autres espèces basales et d'autres espèces à dormance morphologique, pour permettre de proposer une nouvelle définition ou un nouveau marqueur de germination pour ce type de graines.

Une partie de ces résultats a fait l'objet de la présentation d'un poster (Villegente, Burtet-Sarramegna et al 2011). Une publication sur la caractérisation du protéome de la graine mature et son évolution pendant la germination est en cours de rédaction et sera soumise prochainement.

Proteomics highlights specific features of ancestral *Amborella trichopoda* seeds

Matthieu Villegente<sup>1</sup>, Claudette Job<sup>2</sup>, Loïc Rajjou<sup>3,4</sup>, Dominique Job<sup>2,4,\*</sup>, Valérie Sarramegna<sup>1</sup>, Gwendal Cueff<sup>3,4</sup>, Benoît Valot<sup>5</sup>, Thierry Balliau<sup>5</sup>, Bruno Fogliani<sup>1,6\*</sup>,

<sup>1</sup>Laboratoire Insulaire du Vivant et de l'Environnement, University of New Caledonia, Noumea, NEW CALEDONIA

<sup>2</sup>CNRS-Université Claude Bernard Lyon-Institut National des Sciences Appliquées-Bayer CropScience Joint Laboratory (UMR5240), Bayer CropScience, F-69263 Lyon cedex 9, FRANCE

<sup>3</sup>INRA, Jean-Pierre Bourgin Institute (IJPB, UMR1318 INRA-AgroParisTech), Laboratory of Excellence "Saclay Plant Sciences" (LabEx SPS), F-78026 Versailles, France

<sup>4</sup>AgroParisTech, Chair of Plant Physiology, F-75231 Paris, FRANCE

<sup>5</sup>INRA, PAPPSO, Plateforme d'Analyse Protéomique de Paris Sud-Ouest, UMR0320 de Génétique Végétale, Gif-sur-Yvette, FRANCE

<sup>6</sup>Institut Agronomique néo-Calédonien, Diversités biologique et fonctionnelle des écosystèmes terrestres, BP 73, 98890 Païta, NEW CALEDONIA

\* Corresponding authors

## 2. Adaptations au stress nickélique

La nécessité de pouvoir dépolluer les sols souillés par les activités anthropiques (pollution) ou pour aider à l'extraction minière a conduit à un accroissement de l'intérêt pour les plantes concentrant naturellement les métaux dans leurs parties aériennes.

Les études menées au LIVE avaient conduit à l'identification, dans la graine de l'espèce hyperaccumulatrice *Psychotria gabriellae*, de protéines de type DING. Ces protéines dont l'appartenance au règne eucaryote fait polémique, sont impliquées dans la séquestration d'éléments minéraux, notamment du phosphate, ou associées à des transporteurs de type ABC. Cette observation tend à suggérer un possible rôle de ces protéines dans la création du gradient de nickel dans la graine de *P. gabriellae* qui tend à protéger l'embryon, quelle que soit leur origine (endosymbiotique ou de la plante elle-même).

Les travaux réalisés dans le cadre de cette thèse ont permis de confirmer la présence des protéines de type DING dans la graine de *P. gabriellae* provenant de trois populations différentes et répartie sur toute la côte Ouest de la Grande-Terre. Ces protéines ont également été retrouvées dans des lots de graines récoltés de 2006 jusqu'à 2011 et, ce, à tous les stades de maturités étudiés. Ces données nous permettent donc d'affirmer que les protéines de type DING sont présentes dans la graine et dès les stades précoces. Il serait intéressant d'étudier la

présence de protéines de type DING pour d'autres graines d'espèces hyperaccumulatrices de nickel (eg. *Geissois pruinosa*).

L'absence de gènes codant des protéines de type DING dans les génomes des eucaryotes suggère que de telles protéines de type DING observées expérimentalement proviendraient en fait de contaminations bactériennes. Dans notre cas, la présence des protéines de type DING dans la graine de *P. gabriellae* pourrait résulter d'une symbiose particulière régissant la tolérance au nickel dans la graine. Dans ce but, nous avons caractérisé la flore bactérienne endophyte des graines matures sèches de *P. gabriellae*. Quatre souches bactériennes ont été isolées et caractérisées. D'après les données génomiques disponibles actuellement et les résultats de notre approche protéomique, il semble que ces bactéries ne soient pas la source des protéines de type DING de la graine de *P. gabriellae*.

Afin d'affiner la compréhension des mécanismes de l'hyperaccumulation de nickel, un projet de séquençage des transcrits de la plante (racines, feuilles, graines) a été développé, EvoMetoNicks, récemment soutenu par l'ANR (Agence Nationale de la Recherche). L'identification de gènes importants dans la mise en place de ces mécanismes est une étape cruciale en vue de la compréhension et surtout pour l'utilisation par la suite par l'homme.

Ainsi, dans le cadre de ce projet et grâce à la participation au projet 1kP pour le séquençage, nous avons obtenu 18 242 séquences issus de transcrits extraits de fruits de *P. gabriellae*. Bien que l'analyse de ces données soit à un stade préliminaire, elles montrent cependant l'identification de nombreux transporteurs de métaux, dont un transporteur prenant en charge le nickel, ainsi que de protéines de neutralisation du nickel (métallothionéine). Ces résultats sont très encourageants et nous confortent dans l'approche choisie pour caractériser les mécanismes mis en place par la plante pour l'hyperaccumulation de ce métal.

Nos résultats confirment donc la présence des protéines de type DING sans pour autant expliquer leur origine et leur rôle. Avant toute chose, la caractérisation de leur profil d'accumulation serait à déterminer, notamment en lien avec la formation du gradient de nickel dans la graine, par exemple par des expériences d'immuno-localisation des protéines de type DING sur des coupes de graines de *P. gabriellae*.

Par ailleurs, il conviendra de caractériser en profondeur les transcrits issus de la graine de *P. gabriellae* notamment en réalisant une annotation fine des classes ontologiques des gènes identifiés. Nous pourrions ainsi obtenir une image des transcrits accumulés dans les fruits/graines.

Par la suite, la base de données génomiques ainsi construite procurera un outil puissant pour l'étude des profils d'accumulation de gènes cibles pour différentes conditions afin de mettre en évidence une régulation particulière des gènes. Cet outil permettrait également la caractérisation de transporteurs spécifiquement impliqués dans l'homéostasie du nickel.

La rédaction d'un article sur l'étude du protéome de *P. gabriellae* sera initiée prochainement s'appuyant tant sur les résultats accumulés par Zongo (2010) et au cours de cette thèse.

Proteomic analysis of seeds of *Psychotria* species endemic to New Caledonia in relation with nickel tolerance. Serendipitous discovery of DING-type proteins in the seeds of the Ni-hyperaccumulator *Psychotria gabriellae*.

Charly Zongo<sup>1,\$</sup>, Matthieu Villegente<sup>1,\$</sup>, Claudette Job<sup>2</sup>, Jean-Marc Strub<sup>3</sup>, Alain Van Dorsselaer<sup>3</sup>, Christine Schaeffer<sup>3</sup>, François Bernier<sup>4</sup>, Maya Belghazi<sup>5</sup>, Marie-Agnès Jacques<sup>6</sup>, Valérie Sarramegna<sup>1</sup>, Dominique Job<sup>2,\*</sup>, Bruno Fogliani<sup>1,7,\*</sup>

<sup>1</sup> Laboratoire Insulaire du Vivant et de l'Environnement (LIVE, Université de la Nouvelle-Calédonie)

<sup>2</sup> CNRS/Bayer CropScience Joint Laboratory (UMR5240), Lyon, France

<sup>3</sup> Université de Strasbourg, IPHC, CNRS, UMR, Strasbourg, France

<sup>4</sup> Institut de Biologie Moléculaire des Plantes (UPR CNRS 2357), Strasbourg, France

<sup>5</sup> Proteomics Analysis Center (CAPM), Faculty of Medicine, 13916 Marseilles, France

<sup>6</sup> IRHS, INRA-Angers

<sup>7</sup> Institut Agronomique néo-Calédonien (IAC), Nouméa, Nouvelle-Calédonie

\$ Equal contributions

\* Corresponding authors

# Références bibliographiques

- Abell, B. M., Holbrook, L. A., Abenes, M., Murphy, D. J., Hills, M. J., & Moloney, M. M. (1997). Role of the proline knot motif in oleosin endoplasmic reticulum topology and oil body targeting. *The Plant cell*, *9*, 1481–1493. doi:10.1105/tpc.9.8.1481
- Adachi, M., Kanamori, J., Masuda, T., Yagasaki, K., Kitamura, K., Mikami, B., & Utsumi, S. (2003). Crystal structure of soybean 11S globulin: glycinin A3B4 homo-hexamers. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(12), 7395–400. doi:10.1073/pnas.0832158100
- Agrawal, G. K., & Rakwal, R. (2006). Rice proteomics: a cornerstone for cereal food crop proteomes. *Mass spectrometry reviews*, *25*, 1–53. doi:10.1002/mas.20056
- Agrawal, G. K., Job, D., Zivy, M., Agrawal, V. P., Bradshaw, R. A., Dunn, M. J., Haynes, P. A., van Wijk, K. J., & Rakwal, R. (2011). Time to articulate a vision for the future of plant proteomics - A global perspective: An initiative for establishing the International Plant Proteomics Organization (INPPO). *Proteomics*, *11*, 1559–1568. doi:10.1002/pmic.201000608
- Ahsan, N., Lee, S.-H., Lee, D.-G., Lee, H., Lee, S. W., Bahk, J. D., & Lee, B.-H. (2007). Physiological and protein profiles alternation of germinating rice seedlings exposed to acute cadmium toxicity. *Comptes rendus biologiques*, *330*, 735–746. doi:10.1016/j.crv.2007.08.001
- Al-Whaibi, M. H. (2011). Plant heat-shock proteins: A mini review. *Journal of King Saud University - Science*, *23*(2), 139–150. doi:10.1016/j.jksus.2010.06.022
- Albertin, W., Marullo, P., Bely, M., Aigle, M., Bourgeois, A., Langella, O., Balliau, T., Chevret, D., Valot, B., da Silva, T., Dillmann, C., de Vienne, D., & Sicard, D. (2013). Linking post-translational modifications and variation of phenotypic traits. *Molecular & cellular proteomics: MCP*, *12*(3), 720–35. doi:10.1074/mcp.M112.024349
- Ali-Rachedi, S., Bouinot, D., Wagner, M.-H., Bonnet, M., Sotta, B., Grappin, P., & Jullien, M. (2004). Changes in endogenous abscisic acid levels during dormancy release and maintenance of mature seeds: studies with the Cape Verde Islands ecotype, the dormant model of *Arabidopsis thaliana*. *Planta*, *219*, 479–488. doi:10.1007/s00425-004-1251-4
- Alves-Da-Silva, D., Borghetti, F., Thompson, K., Pritchard, H., & Grime, J. P. (2011). Underdeveloped embryos and germination in *Aristolochia galeata* seeds. *Plant biology*, *13 Suppl 1*, 104–108. doi:10.1111/j.1438-8677.2009.00302.x
- Amborella Genome Project. (2013). The Amborella Genome and the Evolution of Flowering Plants. *submitted to Science*.
- Amir, H., Perrier, N., Rigault, F., & Jaffré, T. (2007). Relationships between Ni-hyperaccumulation and mycorrhizal status of different endemic plant species from New Caledonian ultramafic soils. *Plant and Soil*, *293*, 23–35. doi:10.1007/s11104-007-9238-0
- Ansong, C., Purvine, S. O., Adkins, J. N., Lipton, M. S., & Smith, R. D. (2008). Proteogenomics: needs and roles to be filled by proteomics in genome annotation. *Briefings in functional genomics & proteomics*, *7*(1), 50–62. doi:10.1093/bfgp/eln010
- Arc, E., Galland, M., Cuff, G., Godin, B., Lounifi, I., Job, D., & Rajjou, L. (2011). Reboot the system thanks to protein post-translational modifications and proteome diversity: How quiescent seeds restart their metabolism to prepare seedling establishment. *Proteomics*, *11*, 1606–1618. doi:10.1002/pmic.201000641

- Aviezer-Hagai, K., Skovorodnikova, J., Galigniana, M., Farchi-Pisanty, O., Maayan, E., Bocovza, S., Efrat, Y., von Koskull-Döring, P., Ohad, N., & Breiman, A.** (2007). *Arabidopsis* immunophilins ROF1 (AtFKBP62) and ROF2 (AtFKBP65) exhibit tissue specificity, are heat-stress induced, and bind HSP90. *Plant molecular biology*, *63*, 237–255. doi:10.1007/s11103-006-9085-z
- Bailly, C.** (2004). Active oxygen species and antioxidants in seed biology. *Seed Science Research*, *14*, 93–107. doi:10.1079/SSR2004159
- Bailly, C., El-Maarouf-Bouteau, H., & Corbineau, F.** (2008). From intracellular signaling networks to cell death: the dual role of reactive oxygen species in seed physiology. *Comptes rendus biologiques*, *331*, 806–814. doi:10.1016/j.crv.2008.07.022
- Barceló, J., & Poschenrieder, C.** (1990). Plant water relations as affected by heavy metal stress: A review. *Journal of Plant Nutrition*, *13*(1), 1–37. doi:10.1080/01904169009364057
- Baskin, C. C., & Baskin, J. M.** (2001). *Seeds - Ecology, Biogeography, and Evolution of Dormancy and Germination* (Academic P., p. 666). San Diego.
- Baskin, C. C., Baskin, J. M., & Yoshinaga, A.** (2005). Morphophysiological dormancy in seeds of six endemic lobelioid shrubs (Campanulaceae) from the montane zone in Hawaii. *Canadian Journal of Botany*, *83*(12), 1630–1637. doi:10.1139/b05-126
- Baskin, J. M., & Baskin, C. C.** (2007). A classification system for seed dormancy. *Seed Science Research*, *14*(01), 1–16. doi:10.1079/SSR2003150
- Battaglia, M., Olvera-Carrillo, Y., Garcarrubio, A., Campos, F., & Covarrubias, A. A.** (2008). The enigmatic LEA proteins and other hydrophilins. *Plant physiology*, *148*, 6–24. doi:10.1104/pp.108.120725
- Baud, S., Dichow, N. R., Kelemen, Z., D'Andréa, S., To, A., Berger, N., Canonge, M., Kronenberger, J., Viterbo, D., Dubreucq, B., Lepiniec, L., Chardot, T., & Miquel, M.** (2009). Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant & cell physiology*, *50*(8), 1463–1478. doi:10.1093/pcp/pcp092
- Bazin, J., Langlade, N., Vincourt, P., Arribat, S., Balzergue, S., El-Maarouf-Bouteau, H., & Bailly, C.** (2011). Targeted mRNA oxidation regulates sunflower seed dormancy alleviation during dry after-ripening. *The Plant cell*, *23*, 2196–2208. doi:10.1105/tpc.111.086694
- Beiter, T., Reich, E., Williams, R. W., & Simon, P.** (2009). Antisense transcription: a critical look in both directions. *Cellular and molecular life sciences: CMLS*, *66*(1), 94–112. doi:10.1007/s00018-008-8381-y
- Belin, C., Megies, C., Hauserová, E., & Lopez-Molina, L.** (2009). Abscisic acid represses growth of the *Arabidopsis* embryonic axis after germination by enhancing auxin signaling. *The Plant cell*, *21*, 2253–2268. doi:10.1105/tpc.109.067702
- Bentsink, L., Jowett, J., Hanhart, C. J., & Koornneef, M.** (2006). Cloning of DOG1, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(45), 17042–17047. doi:10.1073/pnas.0607877103
- Bentsink, L., & Koornneef, M.** (2008). Seed dormancy and germination. *The Arabidopsis book / American Society of Plant Biologists*, *6*(6), e0119. doi:10.1199/tab.0119
- Berger, F.** (2003). Endosperm: the crossroad of seed development. *Current Opinion in Plant Biology*, *6*, 42–50. doi:10.1016/S1369-5266(02)00004-3
- Berjak, P., & Pammenter, N. W.** (2008). From *Avicennia* to *Zizania*: seed recalcitrance in perspective. *Annals of botany*, *101*, 213–228. doi:10.1093/aob/mcm168
- Berna, A., Bernier, F., Scott, K., & Stuhlmüller, B.** (2002). Ring up the curtain on DING proteins. *FEBS letters*, *524*, 6–10. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12135732>

- Berna, A., Bernier, F., Chabrière, E., Perera, T., & Scott, K.** (2008). DING proteins ; novel members of a prokaryotic phosphate-binding protein superfamily which extends into the eukaryotic kingdom. *The international journal of biochemistry & cell biology*, *40*, 170–175. doi:10.1016/j.biocel.2007.02.004
- Berna, A., Bernier, F., Chabrière, E., Elias, M., Scott, K., & Suh, A.** (2009). For whom the bell tolls? DING proteins in health and disease. *Cellular and molecular life sciences*: *CMLS*, *66*, 2205–2218. doi:10.1007/s00018-009-0006-6
- Berna, A., Scott, K., Chabrière, E., & Bernier, F.** (2009). The DING family of proteins: ubiquitous in eukaryotes, but where are the genes? *BioEssays*, *31*, 570–580. doi:10.1002/bies.200800174
- Bethke, P. C., Libourel, I. G. L., Aoyama, N., Chung, Y.-Y., Still, D. W., & Jones, R. L.** (2007). The *Arabidopsis* aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. *Plant physiology*, *143*, 1173–1188. doi:10.1104/pp.106.093435
- Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., Drost, L., Ridley, P., Hudson, S.-A., Patel, L., Murphy, G., Piffanelli, P., Wedler, E., Wambutt, R., Weitzenegger, T., Pohl, T. M., Terryn, N., Gielen, J., Villarreal, R., De Clerk, R., Van Montagu, M., Lecharny, A., Auborg, S., Gy, I., Kreis, M., Lao, N., Kavanagh, T., Hempel, S., Kotter, P., Entian, K.-D., Rieger, M., Schaeffer, M., Funk, B., Mueller-Auer, S., Silvey, M., James, R., Montfort, A., Pons, A., Puigdomenech, P., Douka, A., Voukelatou, E., Milioni, D., Hatzopoulos, P., Piravandi, E., Obermaier, B., Hilbert, H., Düsterhöft, A., Moores, T., Jones, J. D. G., Eneva, T., Palme, K., Benes, V., Rechman, S., Ansorge, W., Cooke, R., Berger, C., Delseny, M., Voet, M., Volckaert, G., Mewes, H.-W., Klosterman, S., Schueller, C., & Chalwatzis, N.** (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature*, *391*, 485–488.
- Bewley, J. D., & Black, M.** (1994). *Seeds: Physiology of Development and Germination* (p. 445).
- Bewley, J. D.** (1997). Seed Germination and Dormancy. *The Plant cell*, *9*, 1055–1066. doi:10.1105/tpc.9.7.1055
- Bharti, A. K., & Khurana, J. P.** (2003). Molecular characterization of transparent testa (tt) mutants of *Arabidopsis thaliana* (ecotype Estland) impaired in flavonoid biosynthetic pathway. *Plant Science*, *165*(6), 1321–1332. doi:10.1016/S0168-9452(03)00344-3
- Blum, H., Beier, H., & Gross, H. J.** (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis*, *8*(2), 93–99. doi:10.1002/elps.1150080203
- Bobrov, A. V. F. C., Endress, P. K., Melikian, A. P., Romanov, M. S., Sorokin, A. N., & Bejerano, A. P.** (2005). Fruit structure of *Amborella trichopoda* (Amborellaceae). *Botanical Journal of the Linnean Society*, *148*, 265–274. doi:10.1111/j.1095-8339.2005.00422.x
- Boyd, R. S.** (2004). Ecology of metal hyperaccumulation. *New Phytologist*, *162*, 563–567. doi:10.1111/j.1469-8137.2004.01079.x
- Boyd, R. S., Wall, M. A., & Jaffre, T.** (2006). Nickel levels in arthropods associated with Ni hyperaccumulator plants from an ultramafic site in New Caledonia. *Insect Science*, *13*, 271–277. doi:10.1111/j.1744-7917.2006.00094.x
- Bradford, M. M.** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, *72*, 248–254.

- Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R., & Coen, E.** (1997). Inflorescence Commitment and Architecture in *Arabidopsis*. *Science*, 275, 80–83. doi:10.1126/science.275.5296.80
- Bradshaw, R. A., & Burlingame, A. L.** (2005). From proteins to proteomics. *IUBMB life*, 57(4-5), 267–272. doi:10.1080/15216540500091536
- Bremer, B., Bremer, K., Chase, M. W., Fay, M. F., Reveal, J. L., Soltis, D. E., Soltis P.S., & Stevens, P. F.** (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, 161, 105–121.
- Brock, A. K., Willmann, R., Kolb, D., Grefen, L., Lajunen, H. M., Bethke, G., Lee, J., Nürnberger, T., & Gust, A. A.** (2010). The *Arabidopsis* mitogen-activated protein kinase phosphatase PP2C5 affects seed germination, stomatal aperture, and abscisic acid-inducible gene expression. *Plant physiology*, 153, 1098–1111. doi:10.1104/pp.110.156109
- Brooks, R. R., Lee, J., Reeves, R. D., & Jaffré, T.** (1977). Detection of nickeliferous rocks by analysis of herbarium specimens of indicator plants. *Journal of Geochemical Exploration*, 7, 49–57. doi:10.1016/0375-6742(77)90074-7
- Brooks, R. R.** (1987). *Serpentine and its vegetation, a multidisciplinary approach*. (O. Portland, Ed.). Dioscorides Press.
- Brooks, R. R., Chambers, M. F., Nicks, L. J., & Robinson, B. H.** (1998). Phytomining. *Trends in Plant Science*, 3(9), 359–362.
- Brown, J. C.** (1978). Mechanism of iron uptake by plants. *Plant, Cell and Environment*, 1(4), 249–257. doi:10.1111/j.1365-3040.1978.tb02037.x
- Buchanan, B. B., & Jones, R. L.** (2000). *Biochemistry and molecular biology of plants*. (I. K. I. P. House, Ed.) (p. 1408). Rockville.
- Buchanan, B. B., & Balmer, Y.** (2005). Redox regulation: a broadening horizon. *Annual review of plant biology*, 56, 187–220. doi:10.1146/annurev.arplant.56.032604.144246
- Buitink, J., & Leprince, O.** (2008). Intracellular glasses and seed survival in the dry state. *Comptes rendus biologiques*, 331, 788–795. doi:10.1016/j.crv.2008.08.002
- Callahan, Damien Lee.** (2007). *The coordination of nickel in hyperaccumulating plants*. PhD thesis, University of Melbourne.
- Callahan, Damien L., Roessner, U., Dumontet, V., De Livera, A. M., Doronila, A., Baker, A. J. M., & Kolev, S. D.** (2012). Elemental and metabolite profiling of nickel hyperaccumulators from New Caledonia. *Phytochemistry*, 81, 80–9. doi:10.1016/j.phytochem.2012.06.010
- Carlquist, S., & Schneider, E. L.** (2001). Vegetative Anatomy of the New Caledonian Endemic *Amborella trichopoda*: Relationships with the Dliciales and Implications for Vessel Origin. *Pacific Science*, 55(3), 305–312.
- Castellana, N. E., Payne, S. H., Shen, Z., Stanke, M., Bafna, V., & Briggs, S. P.** (2008). Discovery and revision of *Arabidopsis* genes by proteogenomics. *Proceedings of the National Academy of Sciences of the United States of America*, 105(52), 21034–21038.
- Catusse, J.** (2007). *Utilisation de la protéomique pour la dissection des processus de germination et de vigueur germinative des graines de betterave à sucre*. Pau.
- Catusse, J., Job, C., & Job, D.** (2008). Transcriptome- and proteome-wide analyses of seed germination. *Comptes rendus biologiques*, 331(10), 815–22. doi:10.1016/j.crv.2008.07.023
- Catusse, J., Strub, J.-M., Job, C., Van Dorselaer, A., & Job, D.** (2008). Proteome-wide characterization of sugarbeet seed vigor and its tissue specific expression. *Proceedings of the National Academy of Sciences of the United States of America*, 105(29), 10262–10267. doi:10.1073/pnas.0800585105



- Chatthai, M., Kaukinen, K. H., Tranbarger, T. J., Gupta, P. K., & Misra, S.** (1997). The isolation of a novel metallothionein-related cDNA expressed in somatic and zygotic embryos of Douglas-fir: regulation by ABA, osmoticum, and metal ions. *Plant molecular biology*, 34, 243–254. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9207840>
- Chávez Montes, R. A., Ranocha, P., Martinez, Y., Minic, Z., Jouanin, L., Marquis, M., Saulnier, L., Fulton, L. M., Cobbett, C. S., Bitton, F., Renou, J-P., Jauneau, A., & Goffner, D.** (2008). Cell wall modifications in *Arabidopsis* plants with altered alpha-L-arabinofuranosidase activity. *Plant physiology*, 147, 63–77. doi:10.1104/pp.107.110023
- Chen, T., Nayak, N., Majee, S. M., Lowenson, J., Schäfermeyer, K. R., Eliopoulos, A. C., Lloyd, T. D., Dinkins, R., Perry, S. E., Forsthoefel, N. R., Clarke, S. G., Vernon, D. M., Zhou, Z. S., Rejtar, T., & Downie, A. B.** (2010). Substrates of the *Arabidopsis thaliana* protein isoaspartyl methyltransferase 1 identified using phage display and biopanning. *The Journal of biological chemistry*, 285, 37281–37292. doi:10.1074/jbc.M110.157008
- Chiang, G. C. K., Barua, D., Kramer, E. M., Amasino, R. M., & Donohue, K.** (2009). Major flowering time gene, flowering locus C, regulates seed germination in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, 106(28), 11661–11666. doi:10.1073/pnas.0901367106
- Chiang, G. C. K., Barua, D., Dittmar, E., Kramer, E. M., Casas, R. R. De, & Donohue, K.** (2013). Pleiotropy in the wild : the dormancy gene DOG1 exerts cascading control on life cycles. *Evolution*, 67(3), 883–893. doi:10.5061/dryad.g792f
- Chibani, K., Ali-Rachedi, S., Job, C., Job, D., Jullien, M., & Grappin, P.** (2006). Proteomic analysis of seed dormancy in *Arabidopsis*. *Plant physiology*, 142, 1493–1510. doi:10.1104/pp.106.087452
- Chien, C.-T., Chen, S.-Y., Tsai, C.-C., Baskin, J. M., Baskin, C. C., & Kuo-Huang, L.-L.** (2011). Deep simple epicotyl morphophysiological dormancy in seeds of two *Viburnum* species, with special reference to shoot growth and development inside the seed. *Annals of botany*, 108, 13–22. doi:10.1093/aob/mcr096
- Chiwocha, S. D. S., Cutler, A. J., Abrams, S. R., Ambrose, S. J., Yang, J., Ross, A. R. S., & Kermode, A. R.** (2005). The *etr1-2* mutation in *Arabidopsis thaliana* affects the abscisic acid, auxin, cytokinin and gibberellin metabolic pathways during maintenance of seed dormancy, moist-chilling and germination. *The Plant journal*, 42, 35–48. doi:10.1111/j.1365-313X.2005.02359.x
- Clemens, S., Palmgren, M. G., & Krämer, U.** (2002). A long way ahead: understanding and engineering plant metal accumulation. *Trends in plant science*, 7(7), 309–315. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12119168>
- Clouse, S. D., & Sasse, J. M.** (1998). BRASSINOSTEROIDS: Essential Regulators of Plant Growth and Development. *Annual review of plant physiology and plant molecular biology*, 49, 427–451. doi:10.1146/annurev.arplant.49.1.427
- Cluzel, D., Maurizot, P., Collot, J., & Sevin, B.** (2012). An outline of the Geology of New Caledonia; from Permian – Mesozoic Southeast Gondwanaland active margin to Cenozoic obduction and supergene evolution. *Episodes-News magazine of the International Union of Geological Sciences*, 35, 72–86.
- Coleman, R. G., & Jove, C.** (1992). Geological origin of serpentinites. In A. J. M. Baker, J. Proctor, & R. D. Reeves (Eds.), *The vegetation of ultramafic (serpentine) soils* (Intercept., pp. 1–17). Andover.

- Collombet, J.-M., Elias, M., Gotthard, G., Four, E., Renault, F., Joffre, A., Baubichon, D., Rochu, D., & Chabrière, E.** (2010). Eukaryotic DING proteins are endogenous: an immunohistological study in mouse tissues. *PloS one*, 5(2), e9099. doi:10.1371/journal.pone.0009099
- Côme, D.** (1982). Germination. In P. Mazliak (Ed.), *Physiologie végétale II* (Hermann., pp. 129–225). Paris.
- Côme, D., & Corbineau, F.** (2000). Storage of seeds. In F. Artes, M. I. Gill, & M. A. Conesa (Eds.), *Improving postharvest technologies of fruits, vegetables and ornamentals* (pp. 755–770).
- Corbineau, F., Bagniol, S., & Côme, D.** (1990). Sunflower (*Helianthus annuus* L.) seed dormancy and its regulation by ethylene. *Israel Journal of Botany*, 39(4–6), 313–325.
- Cornuet, V., Nicaise, T., Oliva, J., & Talon, M.** (1987). *Ecologie en Nouvelle-Calédonie* (C.T.R.D.P., p. 297). Nouméa.
- Cruaud, A., Jabbour-Zahab, R., Genson, G., Ungricht, S., & Rasplus, J.-Y.** (2012). Testing the emergence of New Caledonia: fig wasp mutualism as a case study and a review of evidence. *PloS one*, 7(2), e30941. doi:10.1371/journal.pone.0030941
- Daher, A., Adam, H., Chabrilange, N., Collin, M., Mohamed, N., Tregear, J. W., & Aberlenc-Bertossi, F.** (2010). Cell cycle arrest characterizes the transition from a bisexual floral bud to a unisexual flower in *Phoenix dactylifera*. *Annals of botany*, 106, 255–266. doi:10.1093/aob/mcq103
- Dalton, D. A., Evans, H. J., & Hanus, F. J.** (1985). Stimulation by nickel of soil microbial urease activity and urease and hydrogenase activities in soybeans grown in a low-nickel soil. *Plant and Soil*, 88, 245–258.
- Darbinian, N., Popov, Y., Khalili, K., & Amini, S.** (2008). Creation of a bi-directional protein transduction system for suppression of HIV-1 expression by p27SJ. *Antiviral research*, 79, 136–141. doi:10.1016/j.antiviral.2007.11.006
- Darbinian, N., Czernik, M., Darbinyan, A., Elias, M., Chabrière, E., Bonasu, S., Khalili, K., Amini, S.** (2009). Evidence for phosphatase activity of p27SJ and its impact on the cell cycle. *Journal of cellular biochemistry*, 107, 400–407. doi:10.1002/jcb.22135
- Darbinian-Sarkissian, N., Darbinyan, A., Otte, J., Radhakrishnan, S., Sawaya, B. E., Arzumanyan, A., Chipitsyna, G., Popov, Y., Rappaport, J., Amini, S., & Khalili, K.** (2006). p27SJ, a novel protein in St John's Wort, that suppresses expression of HIV-1 genome. *Gene therapy*, 13, 288–295. doi:10.1038/sj.gt.3302649
- Darsonval, A., Darrasse, A., Durand, K., Bureau, C., Cesbron, S., & Jacques, M.-A.** (2009). Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Molecular plant-microbe interactions*, 22(6), 747–757. doi:10.1094/MPMI-22-6-0747
- Dave, A., Hernández, M. L., He, Z., Andriotis, V. M. E., Vaistij, F. E., Larson, T. R., & Graham, I. A.** (2011). 12-oxo-phytodienoic acid accumulation during seed development represses seed germination in *Arabidopsis*. *The Plant cell*, 23, 583–599. doi:10.1105/tpc.110.081489
- Davis, M. A., Boyd, R. S., & Cane, J. H.** (2001). Host-switching does not circumvent the Ni-based defense of the Ni hyperaccumulator *Streptanthus polygaloides* (Brassicaceae). *South African Journal of Science*, 97, 554–557.
- De Mercoyrol, L., Job, C., & Job, D.** (1989). Studies on the inhibition by a-amanitin of single-step addition reactions and productive RNA synthesis catalysed by wheat-germ RNA polymerase II. *The Biochemical journal*, 258, 165–169.

- Debeaujon, I., Léon-Kloosterziel, K. M., & Koornneef, M.** (2000). Influence of the testa on seed dormancy, germination, and longevity in *Arabidopsis*. *Plant physiology*, *122*, 403–414. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58877&tool=pmcentrez&rendertype=abstract>
- Dekkers, B. J. W., Pearce, S., van Bolderen-Veldkamp, R. P., Marshall, A., Widera, P., Gilbert, J., Drost, H.-G., Bassel, G. W., Müller, K., King, J. R., Wood, A. T. A., Grosse, I., Quint, M., Krasnogor, N., Leubner-Metzger, G., Holdsworth, M. J., & Bentsink, L.** (2013). Transcriptional Dynamics of Two Seed Compartments with Opposing Roles in *Arabidopsis* Seed Germination. *Plant physiology*, *163*, 205–215. doi:10.1104/pp.113.223511
- Deng, Z. Y., Gong, C. Y., & Wang, T.** (2013). Use of proteomics to understand seed development in rice. *Proteomics*, *13*, 1784–800. doi:10.1002/pmic.201200389
- Desrosiers, R. R., & Fanéus, I.** (2011). Damaged Proteins Bearing L-Isoaspartyl Residues and Aging: A Dynamic Equilibrium Between Generation of Isomerized Forms and Repair by PIMT. *Current Aging Science*, *4*, 8–18.
- Dickie, J. B., & Pritchard, H. W.** (2002). Systematic and evolutionary aspects of desiccation tolerance in seeds. In *Desiccation and Plant Survival* (p. 416). CABI.
- Diemer, H., Elias, M., Renault, F., Rochu, D., Contreras-Martel, C., Schaeffer, C., Van Dorselaer, A., & Chabriere, E.** (2008). Tandem use of X-ray crystallography and mass spectrometry to obtain ab initio the complete and exact amino acids sequence of HPBP, a human 38-kDa apolipoprotein. *Proteins*, *71*, 1708–1720. doi:10.1002/prot.21866
- Donohue, K., Rubio de Casas, R., Burghardt, L., Kovach, K., & Willis, C. G.** (2010). Germination, Postgermination Adaptation, and Species Ecological Ranges. *Annual Review of Ecology, Evolution, and Systematics*, *41*(1), 293–319. doi:10.1146/annurev-ecolsys-102209-144715
- Dunn, M. J.** (1996). Detection of Proteins in Polyacrylamide Gels by Silver Staining. In J. M. Walker (Ed.), *The Protein Protocols Handbook* (Humana Pre., pp. 229–233).
- Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R., & Goodenough, P. W.** (2001). Evolution of functional diversity in the cupin superfamily. *Trends in biochemical sciences*, *26*(12), 740–746. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11738598>
- Dunwell, J. M., Purvis, A., & Khuri, S.** (2004). Cupins: the most functionally diverse protein superfamily? *Phytochemistry*, *65*(1), 7–17. doi:10.1016/j.phytochem.2003.08.016
- Eastmond, P. J.** (2007). MONODEHYDROASCORBATE REDUCTASE4 is required for seed storage oil hydrolysis and postgerminative growth in *Arabidopsis*. *The Plant cell*, *19*, 1376–1387. doi:10.1105/tpc.106.043992
- El-Maarouf-Bouteau, H., Meimoun, P., Job, C., Job, D., & Bailly, C.** (2013). Role of protein and mRNA oxidation in seed dormancy and germination. *Frontiers in plant science*, *4*, 77. doi:10.3389/fpls.2013.00077
- Eskew, D. L., Welch, R. M., & Norvell, W. A.** (1984). Nickel in Higher Plants. *Plant physiology*, *76*, 691–693.
- Fait, A., Angelovici, R., Less, H., Ohad, I., Urbanczyk-Wochniak, E., Fernie, A. R., & Galili, G.** (2006). *Arabidopsis* seed development and germination is associated with temporally distinct metabolic switches. *Plant physiology*, *142*(3), 839–854. doi:10.1104/pp.106.086694
- Farnsworth, E.** (2000). The ecology and physiology of viviparous and recalcitrant seeds. *Annual review of ecology, evolution and systematics*, *31*, 107–138.

- Feild, T. S., Zweiniecki, M. A., Brodribb, T., Jaffré, T., Donoghue, M. J., & Holbrook, M. N.** (2000). Structure and function of tracheary elements in *Amborella trichopoda*. *International journal of plant sciences*, *161*(5), 705–712.
- Ferreira, A., Qucine, M. C., Lacava, P. T., Oda, S., Azevedo, J. L., & Araújo, W. L.** (2008). Diversity of endophytic bacteria from *Eucalyptus* species seeds and colonization of seedlings by *Pantoea agglomerans*. *FEMS microbiology letters*, *287*, 8–14. doi:10.1111/j.1574-6968.2008.01258.x
- Finch-Savage, W. E., & Leubner-Metzger, G.** (2006). Seed dormancy and the control of germination. *The New phytologist*, *171*, 501–523. doi:10.1111/j.1469-8137.2006.01787.x
- Finnie, C., Andersen, B., Shahpiri, A., & Svensson, B.** (2011). Proteomes of the barley aleurone layer: A model system for plant signalling and protein secretion. *Proteomics*, *11*, 1595–1605. doi:10.1002/pmic.201000656
- Floyd, S. K., & Friedman, W. E.** (2001). Developmental evolution of endosperm in basal angiosperms: evidence from *Amborella* (Amborellaceae), *Nuphar* (Nymphaeaceae), and *Illicium* (Illiciaceae). *Plant Systematics and Evolution*, *228*, 153–169. doi:10.1007/s006060170026
- Fogliani, B., Zongo, C., L'Huillier, L., Villegente, M., Wulff, A.,** 2011. Stratégies de recolonisation développées par les espèces pionnières des milieux ultramafiques néo-calédoniens au travers de leurs graines. Colloque GRAINES 2011, Nantes, France, 27 & 28 Octobre 2011 (Communication orale)
- Fogliani, B., Gateblé, G., Villegente, M., Klein, N., Scutt, C., Fabre, I., & Baskin, C. C.** (2013). Seed dormancy and germination in the basal angiosperm *Amborella trichopoda*. *submitted*.
- Forbis, T. a, Floyd, S. K., & de Queiroz, A.** (2002). The evolution of embryo size in angiosperms and other seed plants: implications for the evolution of seed dormancy. *Evolution*, *56*(11), 2112–2125. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12487343>
- Fourquin, C., Vinauger-Douard, M., Fogliani, B., Dumas, C., & Scutt, C. P.** (2005). Evidence that CRABS CLAW and TOUSLED have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(12), 4649–4654. doi:10.1073/pnas.0409577102
- Frandsen, G. I., Mundy, J., & Tzen, J. T. C.** (2001). Oil bodies and their associated proteins, oleosin and caleosin. *Physiologia plantarum*, *112*, 301–307.
- Friedman, W. E.** (2006). Embryological evidence for developmental lability during early angiosperm evolution. *Nature*, *441*, 337–340. doi:10.1038/nature04690
- Friedman, W. E.** (2008). Hydatellaceae are water lilies with gymnospermous tendencies. *Nature*, *453*(7191), 94–97. doi:10.1038/nature06733
- Friedrich, K. L., Giese, K. C., Buan, N. R., & Vierling, E.** (2004). Interactions between small heat shock protein subunits and substrate in small heat shock protein-substrate complexes. *The Journal of biological chemistry*, *279*(2), 1080–1089. doi:10.1074/jbc.M311104200
- Galau, G. A., Hughes, W. D., & Dure, L.** (1986). Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. *Plant molecular biology*, *7*, 155–170.
- Gallardo, K., Job, C., Groot, S. P. C., Puype, M., Demol, H., Vandekerckhove, J., & Job, D.** (2001). Proteomic analysis of *Arabidopsis* seed germination and priming. *Plant physiology*, *126*, 835–848. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=111173&tool=pmcentrez&rendertype=abstract>

- Gallardo, K., Job, C., Groot, S. P. C., Puype, M., Demol, H., & Job, D.** (2002). Proteomics of *Arabidopsis* Seed Germination . A Comparative Study of Wild-Type and Gibberellin-Deficient Seeds. *Plant Physiology*, *129*, 823–837. doi:10.1104/pp.002816.gence
- Gallardo, K., Job, C., Groot, S. P. C., Puype, M., Demol, H., Vandekerckhove, J., & Job, D.** (2002). Importance of methionine biosynthesis for *Arabidopsis* seed germination and seedling growth. *Physiologia plantarum*, *116*, 238–247. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12354201>
- Gaubier, P., Raynal, M., Hull, G., Huestis, G. M., Grellet, F., Arenas, C., Pagès, M., & Delseny, M.** (1993). Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. *Molecular & general genetics*, *238*, 409–418. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8492809>
- Goremykin, V. V., Hirsch-Ernst, K. I., Wolf, S., & Hellwig, F. H.** (2003). Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Molecular biology and evolution*, *20*(9), 1499–1505. doi:10.1093/molbev/msg159
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A.** (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, *29*(7), 644–652. doi:10.1038/nbt.1883
- Grandcolas, P., Muriene, J., Robillard, T., Desutter-Grandcolas, L., Jourdan, H., Guilbert, E., & Deharveng, L.** (2008). New Caledonia: a very old Darwinian island? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1508), 3309–3317. doi:10.1098/rstb.2008.0122
- Greenspan, P., Mayer, E. P., & Fowler, S. D.** (1985). Nile red: a selective fluorescent stain for intracellular lipid droplets. *The Journal of cell biology*, *100*(3), 965–973. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2113505&tool=pmcentrez&endertype=abstract>
- Grelet, J., Benamar, A., Teyssier, E., Avelange-Macherel, M.-H., Grunwald, D., & Macherel, D.** (2005). Identification in Pea Seed Mitochondria of a Late-Embryogenesis Abundant Protein Able to Protect Enzymes from Drying. *Plant Physiology*, *137*, 157–167. doi:10.1104/pp.104.052480.1
- Groot, S. P. C., & Karssen, C. M.** (1987). Gibberellins regulate seed germination in tomato by endosperm weakening: a study with gibberellin-deficient mutants. *Planta*, *171*, 525–531. doi:10.1007/BF00392302
- Gruis, D. F., Selinger, D. A., Curran, J. M., & Jung, R.** (2002). Redundant Proteolytic Mechanisms Process Seed Storage Proteins in the Absence of Seed-Type Members of the Vacuolar Processing Enzyme Family of Cysteine Proteases. *The Plant cell*, *14*, 2863–2882. doi:10.1105/tpc.005009.eign
- Guilloteau, M., Laloi, M., Blais, D., Crouzillat, D., & Mc Carthy, J.** (2003). Oil bodies in *Theobroma cacao* seeds: cloning and characterization of cDNA encoding the 15.8 and 16.9 kDa oleosins. *Plant Science*, *164*, 597–606. doi:10.1016/S0168-9452(03)00011-6
- Häger, K. P., Braun, H., Czihal, A., Müller, B., & Bäumlein, H.** (1995). Evolution of seed storage protein genes: legumin genes of *Ginkgo biloba*. *Journal of molecular evolution*, *41*, 457–466. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7563133>

- Häger, K.-P., Müller, B., Wind, C., Erbach, S., & Fischer, H.** (1996). Evolution of legumin genes: loss of an ancestral intron at the beginning of angiosperm diversification. *FEBS letters*, *387*, 94–98. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8654576>
- Häger, Klaus-Peter, & Wind, C.** (1997). Two ways of legumin-precursor processing in conifers. Characterization and evolutionary relationships of *Metasequoia* cDNAs representing two divergent legumin gene subfamilies. *European journal of biochemistry*, *246*, 763–771. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9219537>
- Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S., & Masmoudi, K.** (2011). Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant signaling & behavior*, *6*(10), 1503–1509. doi:10.4161/psb.6.10.17088
- Hanson, B., Lindblom, S. D., Loeffler, M. L., & Pilon-Smits, E. A. H.** (2004). Selenium protects plants from phloem-feeding aphids due to both deterrence and toxicity. *New Phytologist*, *162*(3), 655–662. doi:10.1111/j.1469-8137.2004.01067.x
- Haydon, M. J., & Cobbett, C. S.** (2007). Transporters of ligands for essential metal ions in plants. *The New phytologist*, *174*, 499–506. doi:10.1111/j.1469-8137.2007.02051.x
- He, D., Han, C., Yao, J., Shen, S., & Yang, P.** (2011). Constructing the metabolic and regulatory pathways in germinating rice seeds through proteomic approach. *Proteomics*, *11*, 2693–2713. doi:10.1002/pmic.201000598
- Hedman, H., Källman, T., & Lagercrantz, U.** (2009). Early evolution of the MFT-like gene family in plants. *Plant molecular biology*, *70*, 359–369. doi:10.1007/s11103-009-9478-x
- Hepher, A., & Roberts, J. A.** (1985a). The control of seed germination in *Trollius ledebouri* A model of seed dormancy. *Planta*, *166*, 321–328. doi:10.1007/BF00401168
- Hepher, A., & Roberts, J. A.** (1985b). The control of seed germination in *Trollius ledebouri*: The breaking of dormancy. *Planta*, *166*, 314–320. doi:10.1007/BF00401167
- Herrmann, K. M., & Weaver, L. M.** (1999). The shikimate pathway. *Annual review of plant physiology and plant molecular biology*, *50*, 473–503.
- Hidayati, S. N., Baskin, J. M., & Baskin, C. C.** (2000). Morphophysiological dormancy in seeds of two north american and one eurasian species of *Sambucus* (Caprifoliaceae) with underdeveloped spatulate embryos. *American Journal of Botany*, *87*(11), 1669–1678.
- Hilhorst, H. W. M., Finch-savage, W. E., Buitink, J., Bolingue, W., & Leubner-Metzger, G.** (2010). Dormancy in plant seeds. In E. Lubzens, J. Cerda, & M. Clark (Eds.), *Dormancy and Resistance in Harsh Environments* (Vol. 21, pp. 43–67). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-12422-8
- Homrichhausen, T. M., Hewitt, J. R., & Nonogaki, H.** (2003). Endo- $\beta$ -mannanase activity is associated with the completion of embryogenesis in imbibed carrot (*Daucus carota* L.) seeds. *Seed Science Research*, *13*(3), 219–227. doi:10.1079/SSR2003139
- Hong, S.-W., & Vierling, E.** (2001). Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. *The Plant journal*, *27*(1), 25–35. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11489180>
- Howarth, D. G., Gustafsson, M. H. G., Baum, D. A., & Motley, T. J.** (2003). Phylogenetics of the genus *Scaevola* (Goodeniaceae): implication for dispersal patterns across the Pacific Basin and colonization of the Hawaii Islands. *American journal of botany*, *90*(6), 915–923.
- Howell, K. A., Narsai, R., Carroll, A., Ivanova, A., Lohse, M., Usadel, B., Millar, A. H., & Whelan, J.** (2009). Mapping Metabolic and Transcript Temporal Switches during Germination in Rice Highlights Specific Transcription Factors and the Role of RNA Instability in the Germination Process. *Plant physiology*, *149*, 961–980. doi:10.1104/pp.108.129874

- Hu, X.-H., Wang, S.-P., & Qu, B.** (2009). Effects of spermidine on seed germination and seedling antioxidant system of tomato under NaCl stress. *Ying yong sheng tai xue bao*, 20(2), 446–450. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19459389>
- Hubbard, K. E., Nishimura, N., Hitomi, K., Getzoff, E. D., & Schroeder, J. I.** (2010). Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes & development*, 24, 1695–1708. doi:10.1101/gad.1953910
- Hundertmark, M., & Hinch, D. K.** (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC genomics*, 9, 118. doi:10.1186/1471-2164-9-118
- Hyun, T. K., Kumar, D., Cho, Y.-Y., Hyun, H.-N., & Kim, J.-S.** (2013). Computational identification and phylogenetic analysis of the oil-body structural proteins, oleosin and caleosin, in castor bean and flax. *Gene*, 515, 454–460. doi:10.1016/j.gene.2012.11.065
- Ichikawa, J. K., & Clarke, S.** (1998). A highly active protein repair enzyme from an extreme thermophile: the L-isoaspartyl methyltransferase from *Thermotoga maritima*. *Archives of biochemistry and biophysics*, 358(2), 222–231. doi:10.1006/abbi.1998.0830
- Iglesias-Fernández, R., & Matilla, A. J.** (2010). Genes involved in ethylene and gibberellins metabolism are required for endosperm-limited germination of *Sisymbrium officinale* L. seeds: germination in *Sisymbrium officinale* L. seeds. *Planta*, 231, 653–664. doi:10.1007/s00425-009-1073-5
- ISEE.** (2013). Les chiffres clé. Site consulté en juin 2013. Institut de la statistique et des études économiques.
- ISTA.** (1999). International rules for seed testing. *Seed Science and Technology*, 27 (suppl.).
- Jacobsen, J. V., & Pressman, E.** (1979). A structural study of germination in celery (*Apium graveolens* L.) seed with emphasis on endosperm breakdown. *Planta*, 144, 241–248.
- Jaffré, T., & Schmid, M.** (1974). Accumulation du nickel par une rubiacée de Nouvelle-Calédonie, *Psychotria douarrei* (G. Beauvisage). *Comptes rendus de l'académie des sciences de Paris*, 278, 1727–1730.
- Jaffré, T., Brooks, R. R., & Trow, J. M.** (1979). Hyperaccumulation of Nickel by *Geissois* species. *Plant and Soil*, 51, 157–162. doi:10.1111/j.1467-9647.2008.00448.x
- Jaffré, T., Morat, P., & Veillon, J.-M.** (1994). Caractéristiques et composition floristique des principales formations végétales. *Bois et forêts des tropiques*, 242, 7–30.
- Jasid, S., Simontacchi, M., & Puntarulo, S.** (2008). Exposure to nitric oxide protects against oxidative damage but increases the labile iron pool in sorghum embryonic axes. *Journal of experimental botany*, 59(14), 3953–3962. doi:10.1093/jxb/ern235
- Jaya, N., Garcia, V., & Vierling, E.** (2009). Substrate binding site flexibility of the small heat shock protein molecular chaperones. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), 15604–15609. doi:10.1073/pnas.0902177106
- Job, C., Kersulec, A., Ravasio, L., Chareyre, S., Pepin, R., & Job, D.** (1997). The solubilization of the basic subunit of sugarbeet seed 11-S globulin during priming and early germination. *Seed Science Research*, 225–243.
- Job, C., Rajjou, L., Lovigny, Y., Belghazi, M., & Job, D.** (2005). Patterns of Protein Oxidation in *Arabidopsis* Seeds and during Germination. *Plant physiology*, 138, 790–802. doi:10.1104/pp.105.062778.ing
- Johnson-Flanagan, A. M., & McLachlan, G.** (1990). The role of chlorophyllase in degreening canola (*Brassica napus*) seeds and its activation by sublethal freezing. *Physiologia Plantarum*, 80(3), 460–466. doi:10.1111/j.1399-3054.1990.tb00068.x

- Jolliffe, N. A., Craddock, C. P., & Frigerio, L.** (2005). Pathways for protein transport to seed storage vacuoles. *Biochemical Society transactions*, 33, 1016–1018. doi:10.1042/BST20051016
- Kagan, R. M., McFadden, H. J., McFadden, P. N., O'Connor, C., & Clarke, S.** (1997). Molecular phylogenetics of a protein repair methyltransferase. *Comparative biochemistry and physiology*, 117B(3), 379–385. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9253175>
- Kalemba, E. M., & Pukacka, S.** (2007). Possible roles of LEA proteins and sHSPs in seed protection: a short review. *Biological letter*, 44(1), 3–16.
- Kang, J., Hwang, J.-U., Lee, M., Kim, Y.-Y., Assmann, S. M., Martinoia, E., & Lee, Y.** (2010). PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proceedings of the National Academy of Sciences of the United States of America*, 107(5), 2355–2360. doi:10.1073/pnas.0909222107
- Kapp, L. D., & Lorsch, J. R.** (2004). The molecular mechanics of eukaryotic translation. *Annual review of biochemistry*, 73, 657–704. doi:10.1146/annurev.biochem.73.030403.080419
- Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engström, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B., & Wahlestedt, C.** (2005). Antisense transcription in the mammalian transcriptome. *Science*, 309(5740), 1564–1566. doi:10.1126/science.1112009
- Kehrer, J. P.** (2000). The Haber-Weiss reaction and mechanisms of toxicity. *Toxicology*, 149(1), 43–50. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10963860>
- Kelley, L. a, & Sternberg, M. J. E.** (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols*, 4(3), 363–371. doi:10.1038/nprot.2009.2
- Kepczynski, J., & Kepczynski, E.** (1997). Ethylene in seed dormancy and germination. *Physiologia plantarum*, 101, 720–726.
- Kershaw, A. P.** (1984). Late Cenozoic plant extinctions in Australia. In P. Martin & R. Klein (Eds.), *Quaternary extinctions: a prehistoric revolution* (p. 705). Tucson: Univ. Arizona Press.
- Kersten, W. J., Brooks, R. R., Reeves, R. D., & Jaffré, T.** (1980). Nature of nickel complexes in *Psychotria douarrei* and other nickel-accumulating plants. *Phytochemistry*, 19, 1963–1965. doi:10.1016/0031-9422(80)83013-5
- Kester, S. T., Geneve, R. L., & Houtz, R. L.** (1997). Priming and accelerated ageing affect L-isoaspartyl methyltransferase activity in tomato (*Lycopersicon esculentum* Mill.) seed. *Journal of experimental botany*, 48(309), 943–949.
- Khare, S., Linster, C. L., & Clarke, S. G.** (2011). The interplay between protein L-isoaspartyl methyltransferase activity and insulin-like signaling to extend lifespan in *Caenorhabditis elegans*. *PloS one*, 6(6), e20850. doi:10.1371/journal.pone.0020850
- Kim, E., Lowenson, J. D., MacLaren, D. C., Clarke, S., & Young, S. G.** (1997). Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 6132–6137. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=21014&tool=pmcentrez&rendertype=abstract>



- Kimura, M., & Nambara, E.** (2010). Stored and neosynthesized mRNA in *Arabidopsis* seeds: effects of cycloheximide and controlled deterioration treatment on the resumption of transcription during imbibition. *Plant molecular biology*, *73*, 119–129. doi:10.1007/s11103-010-9603-x
- Klahre, U., Friederich, E., Kost, B., Louvard, D., & Chua, N.-H.** (2000). Villin-like actin-binding proteins are expressed ubiquitously in *Arabidopsis*. *Plant physiology*, *122*, 35–48. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58842&tool=pmcentrez&rendertype=abstract>
- Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., & Araki, T.** (1999). A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. *Science*, *286*, 1960–1962. doi:10.1126/science.286.5446.1960
- Koornneef, M., & van der Veen, J. . H.** (1980). Induction and Analysis of Gibberelin Sensitive Mutants in *Arabidopsis thaliana* ( L .) Heynh. *Theoretical and Applied Genetics*, *263*, 257–263.
- Koornneef, Maarten, Bentsink, L., & Hilhorst, H.** (2002). Seed dormancy and germination. *Current Opinion in Plant Biology*, *5*, 33–36.
- Kotak, S., Vierling, E., Bäumlein, H., & von Koskull-Döring, P.** (2007). A novel transcriptional cascade regulating expression of heat stress proteins during seed development of *Arabidopsis*. *The Plant cell*, *19*, 182–195. doi:10.1105/tpc.106.048165
- Kranner, I., Minibayeva, F. V., Beckett, R. P., & Seal, C. E.** (2010). What is stress? Concepts, definitions and applications in seed science. *The New phytologist*, *188*, 655–673. doi:10.1111/j.1469-8137.2010.03461.x
- Kranner, I., Roach, T., Beckett, R. P., Whitaker, C., & Minibayeva, F. V.** (2010). Extracellular production of reactive oxygen species during seed germination and early seedling growth in *Pisum sativum*. *Journal of plant physiology*, *167*, 805–811. doi:10.1016/j.jplph.2010.01.019
- Kranner, I., & Colville, L.** (2011). Metals and seeds: Biochemical and molecular implications and their significance for seed germination. *Environmental and Experimental Botany*, *72*(1), 93–105. doi:10.1016/j.envexpbot.2010.05.005
- Kriz, A. R., Wallace, M. S., & Paiva, R.** (1990). Globulin Gene Expression in Embryos of Maize viviparous mutants. *Plant Physiology*, *92*, 538–542.
- Kucera, B., Cohn, M. A., & Leubner-Metzger, G.** (2005). Plant hormone interactions during seed dormancy release and germination. *Seed Science Research*, *15*, 281–307. doi:10.1079/SSR2005218
- Kumar, V., Yu, S., Farell, G., Toback, F. G., & Lieske, J. C.** (2004). Renal epithelial cells constitutively produce a protein that blocks adhesion of crystals to their surface. *American journal of physiology. Renal physiology*, *287*(3), F373–83. doi:10.1152/ajprenal.00418.2003
- Kuo, W. Y., Huang, C. H., Liu, A. C., Cheng, C. P., Li, S. H., Chang, W. C., Weiss, C., Azem, A., & Jinn, T. L.** (2013). CHAPERONIN 20 mediates iron superoxide dismutase (FeSOD) activity independent of its co-chaperonin role in *Arabidopsis* chloroplasts. *New phytologist*, *197*, 99–110. doi:10.1111/j.1469-8137.2012.04369.x
- Kuromori, T., & Shinozaki, K.** (2010). ABA transport factors found in *Arabidopsis* ABC transporters. *Plant Signaling & Behavior*, *5*(9), 1124–1126. doi:10.4161/psb.5.9.12566
- L’Huillier, L., Jaffré, T., & Wulff, A.** (2010). *Mines et environnement en Nouvelle-Calédonie: les milieux sur substrats ultramafiques et leur restauration.* (IAC., p. 412). Nouméa.

- Law, S. R., Narsai, R., Taylor, N. L., Delannoy, E., Carrie, C., Giraud, E., Millar, A. H., Small, I., & Whelan, J.** (2012). Nucleotide and RNA metabolism prime translational initiation in the earliest events of mitochondrial biogenesis during *Arabidopsis* germination. *Plant physiology*, *158*, 1610–1627. doi:10.1104/pp.111.192351
- Le, H., Browning, K. S., & Gallie, D. R.** (1998). The Phosphorylation State of the Wheat Translation Initiation Factors eIF4B, eIF4A, and eIF2 Is Differentially Regulated during Seed Development and Germination. *Journal of Biological Chemistry*, *273*(32), 20084–20089. doi:10.1074/jbc.273.32.20084
- Lee, A. K., Slovin, J. P., & Suh, J. K.** (2012). Dehydration intolerant seeds of *Ardisia* species accumulate storage and stress proteins during development. *Horticulture, Environment, and Biotechnology*, *53*(6), 530–538. doi:10.1007/s13580-012-0073-8
- Lee, G. J., Roseman, A. M., Saibil, H. R., & Vierling, E.** (1997). A small heat shock protein stably binds heat-denatured model substrates and can maintain a substrate in a folding-competent state. *The EMBO journal*, *16*(3), 659–671. doi:10.1093/emboj/16.3.659
- Lee, J., Reeves, R. D., Brooks, R. R., & Jaffré, T.** (1978). The relation between Nickel and citric acid in some Nickel-accumulating plants. *Phytochemistry*, *17*, 1033–1035.
- Lee, J. T., Davidow, L. S., & Warshawsky, D.** (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature genetics*, *21*(4), 400–404. doi:10.1038/7734
- Lee, K., Bih, F. Y., Learn, G. H., Ting, J. ulie T. L., Sellers, C., & Huang, A. H. C.** (1994). Oleosins in the gametophytes of *Pinus* and *Brassica* and their phylogenetic relationship with those in the sporophytes of various species. *Planta*, *193*, 461–469. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7764875>
- Lee, M., Lee, K., Lee, J., Noh, E. W., & Lee, Y.** (2005). AtPDR12 Contributes to Lead Resistance in *Arabidopsis*. *Plant physiology*, *138*, 827–836. doi:10.1104/pp.104.058107.2000
- Lee, S., Kim, S.-G., & Park, C.-M.** (2010). Salicylic acid promotes seed germination under high salinity by modulating antioxidant activity in *Arabidopsis*. *The New phytologist*, *188*, 626–637. doi:10.1111/j.1469-8137.2010.03378.x
- Lellis, A. D., Allen, M. L., Aertker, A. W., Tran, J. K., Hillis, D. M., Harbin, C. R., Caldwell, C., Gallie, D. R., & Browning, K. S.** (2010). Deletion of the eIFiso4G subunit of the *Arabidopsis* eIFiso4F translation initiation complex impairs health and viability. *Plant molecular biology*, *74*, 249–263. doi:10.1007/s11103-010-9670-z
- Lemaire, B., Lachenaud, O., Persson, C., Smets, E., & Dessein, S.** (2012). Screening for leaf-associated endophytes in the genus *Psychotria* (Rubiaceae). *FEMS microbiology ecology*, 1–9. doi:10.1111/j.1574-6941.2012.01356.x
- Leonard, S. S., Harris, G. K., & Shi, X.** (2004). Serial Review: Redox-Active Metal Ions , Reactive Oxygen Species , and Apoptosis. *Free radical biology & medicine*, *37*(12), 1921–1942. doi:10.1016/j.freeradbiomed.2004.09.010
- Lepiniec, L., Debeaujon, I., Routaboul, J.-M., Baudry, A., Pourcel, L., Nesi, N., & Caboche, M.** (2006). Genetics and biochemistry of seed flavonoids. *Annual review of plant biology*, *57*, 405–430. doi:10.1146/annurev.arplant.57.032905.105252
- Leprince, O., Aelst, A. C. van, Pritchard, H. W., & Murphy, D. J.** (1998). Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and -sensitive oilseeds. *Planta*, *204*, 109–119.
- Leprince, Olivier, Harren, F. J. M., Buitink, J., Alberda, M., & Hoekstra, F. A.** (2000). Metabolic Dysfunction and Unabated Respiration Precede the Loss of Membrane Integrity during Dehydration of Germinating Radicles. *Plant physiology*, *122*, 597–608. Retrieved from

<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58896&tool=pmcentrez&rendertype=abstract>

- Leubner-Metzger, G.** (2005). beta-1,3-Glucanase gene expression in low-hydrated seeds as a mechanism for dormancy release during tobacco after-ripening. *The Plant journal*, *41*, 133–145. doi:10.1111/j.1365-313X.2004.02284.x
- Lewis, A. P., & Crowther, D.** (2005). DING proteins are from *Pseudomonas*. *FEMS microbiology letters*, *252*, 215–222. doi:10.1016/j.femsle.2005.08.047
- Li, C., Li, M., Dunwell, J. M., & Zhang, Y.-M.** (2012). Gene duplication and an accelerated evolutionary rate in 11S globulin genes are associated with higher protein synthesis in dicots as compared to monocots. *BMC evolutionary biology*, *12*, 15. doi:10.1186/1471-2148-12-15
- Li, F., Asami, T., Wu, X., Tsang, E. W. T., & Cutler, A. J.** (2007). A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant physiology*, *145*, 87–97. doi:10.1104/pp.107.100560
- Li, Q., Wang, B.-C., Xu, Y., & Zhu, Y.-X.** (2007). Systematic Studies of 12S Seed Storage Protein Accumulation and Degradation Patterns during *Arabidopsis* Seed Maturation and Early Seedling Germination Stages. *Journal of Biochemistry and Molecular Biology*, *40*, 373–381.
- Li, Z., & Trick, H. N.** (2005). Rapid method for high-quality RNA isolation from seed endosperm containing high levels of starch. *BioTechniques*, *38*, 872–876. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16018547>
- Linkies, A., Müller, K., Morris, K., Turecková, V., Wenk, M., Cadman, C. S. C., Corbineau, F., Strnad, M., Lynn, J. R., Finch-Savage, W. E., & Leubner-Metzger, G.** (2009). Ethylene interacts with abscisic acid to regulate endosperm rupture during germination: a comparative approach using *Lepidium sativum* and *Arabidopsis thaliana*. *The Plant cell*, *21*, 3803–3822. doi:10.1105/tpc.109.070201
- Liu, Q., Sun, Y., Su, W., Yang, J., Liu, X., Wang, Y., Wang, F., Li, H., & Li, X.** (2012). Species-specific size expansion and molecular evolution of the oleosins in angiosperms. *Gene*, *509*, 247–257. doi:10.1016/j.gene.2012.08.014
- Liu, Y., Ye, N., Liu, R., Chen, M., & Zhang, J.** (2010). H<sub>2</sub>O<sub>2</sub> mediates the regulation of ABA catabolism and GA biosynthesis in *Arabidopsis* seed dormancy and germination. *Journal of experimental botany*, *61*(11), 2979–2990. doi:10.1093/jxb/erq125
- Lopez-Molina, L., Mongrand, S., & Chua, N.-H.** (2001). A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(8), 4782–4787. doi:10.1073/pnas.081594298
- Lopez-Molina, L., Mongrand, S., McLachlin, D. T., Chait, B. T., & Chua, N.-H.** (2002). ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *The Plant Journal*, *32*, 317–328. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12410810>
- Luo, M., Taylor, J. M., Spriggs, A., Zhang, H., Wu, X., Russell, S., Singh, M., & Koltunow, A.** (2011). A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS genetics*, *7*(6), e1002125. doi:10.1371/journal.pgen.1002125
- MacKay, A. C., McGill, C. R., Fountain, D. W., & Southward, R. C.** (2002). Seed dormancy and germination of a panel of New Zealand plants suitable for re-vegetation. *New Zealand Journal of Botany*, *40*(3), 373–382. doi:10.1080/0028825X.2002.9512798

- Maltman, D. J., Gadd, S. M., Simon, W. J., & Slabas, A. R.** (2007). Differential proteomic analysis of the endoplasmic reticulum from developing and germinating seeds of castor (*Ricinus communis*) identifies seed protein precursors as significant components of the endoplasmic reticulum. *Proteomics*, *7*, 1513–1528. doi:10.1002/pmic.200600694
- Manfre, A. J., Lanni, L. M., & Marcotte, W. R. J.** (2006). The *Arabidopsis* Group 1 LATE EMBRYOGENESIS ABUNDANT Protein ATEM6 Is Required for Normal Seed Development. *Plant Physiology*, *140*, 140–149. doi:10.1104/pp.105.072967.development.
- Mathe, C., & Vieillescazes, C.** (2002). Compréhension des mécanismes de coloration des liants protéiques picturaux à l'aide du Noir Amide 10B. *l'actualité chimique*, 11–14.
- Mathews, S., & Donoghue, M. J.** (1999). The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science*, *286*, 947–950. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10542147>
- Membré, N., Bernier, F., Staiger, D., & Berna, A.** (2000). *Arabidopsis thaliana* germin-like proteins: common and specific features point to a variety of functions. *Planta*, *211*(3), 345–354. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10987552>
- Mendes, R., Pizzirani-Kleiner, A. A., Araujo, W. L., & Raaijmakers, J. M.** (2007). Diversity of cultivated endophytic bacteria from sugarcane: Genetic and Biochemical characterization of *Burkholderia cepacia* complex isolates. *Applied and environmental microbiology*, *73*(22), 7259–7267. doi:10.1128/AEM.01222-07
- Minguzzi, C., & Vergnano, O.** (1948). Nickel content of the ash of *Alyssum bertolonii*. *Atti soc. toscana sci. nat.*, *55*, 49–74.
- Minic, Z., Do, C.-T., Rihouey, C., Morin, H., Lerouge, P., & Jouanin, L.** (2006). Purification, functional characterization, cloning, and identification of mutants of a seed-specific arabinan hydrolase in *Arabidopsis*. *Journal of experimental botany*, *57*(10), 2339–2351. doi:10.1093/jxb/erj205
- Mittermeier, R. A., Gil, P. R., Hoffman, M., Pilgrim, J., Brooks, T., Mittermeier, C. G., Lamoreux, J., da Fonseca, G. A. B., Seligmann, P. A., & Ford, H.** (2004). *Hotspots Revisited*. (CEMEX, Ed.) (p. 390). Mexico: Conservation International.
- Mok, D. W. S., & Mok, M. C.** (2001). Cytokinin metabolism and action. *Annual review of plant physiology and plant molecular biology*, *52*, 89–118.
- Montoya-García, L., Muñoz-Ocoter, V., Aguilar, R., & Sánchez de Jiménez, E.** (2002). Regulation of acidic ribosomal protein expression and phosphorylation in maize. *Biochemistry*, *41*, 10166–10172. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12146982>
- Morales, R., Berna, A., Carpentier, P., Contreras-Martel, C., Renault, F., Nicodeme, M., Chesne-Seck, M.-L., Bernier, F., Dupuy, J., Schaeffer, C., Diemer, H., Van-Dorselaer, A., Fontecilla-Camps, J. C., Masson, P., Rochu, D., & Chabriere, E.** (2006). Serendipitous discovery and X-ray structure of a human phosphate binding apolipoprotein. *Structure*, *14*, 601–609. doi:10.1016/j.str.2005.12.012
- Morat, P., Jaffré, T., & Veillon, J.-M.** (1986). Grande Terre, New Caledonia, France. In S. D. Davis, V. H. Heywood, & A. C. Amilton (Eds.), (Centres of., pp. 529–537). Gland.
- Morat, P., Jaffré, Tronchet, F., T., Muzinger, J., Pillon, Y., Veillon, J.-M., Chapolin, M., Birnbaum Ph, Rigault F, Dagostini G, Tinel J, & Lowry P.** (2012). Le référentiel taxonomique Florical et les caractéristiques de la flore vasculaire indigène de la Nouvelle-Calédonie. *Adansonia*, *34*, 179–221.
- Moreau, M., Lindermayr, C., Durner, J., & Klessig, D. F.** (2010). NO synthesis and signaling in plants--where do we stand? *Physiologia plantarum*, *138*, 372–383. doi:10.1111/j.1399-3054.2009.01308.x

- Mudgett, M. B., & Clarke, S.** (1994). Hormonal and environmental responsiveness of a developmentally regulated protein repair L-isoaspartyl methyltransferase in wheat. *The Journal of biological chemistry*, 269(41), 25605–25612. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7929264>
- Mudgett, M. B., Lowenson, J. D., & Clarke, S.** (1997). Protein Repair L-Isoaspartyl Methyltransferase in Plants. *Plant Physiology*, 115, 1481–1489.
- Mukerjee, R., Deshmane, S. L., Darbinian, N., Czernik, M., Khalili, K., Amini, S., & Sawaya, B. E.** (2008). St. John's Wort protein, p27SJ, regulates the MCP-1 promoter. *Molecular immunology*, 45, 4028–4035. doi:10.1016/j.molimm.2008.06.001
- Müller, K., Linkies, A., Vreeburg, R. A. M., Fry, S. C., Krieger-Liszkay, A., & Leubner-Metzger, G.** (2009). In vivo cell wall loosening by hydroxyl radicals during cress seed germination and elongation growth. *Plant physiology*, 150, 1855–1865. doi:10.1104/pp.109.139204
- Müntz, K., Belozersky, M. a, Dunaevsky, Y. E., Schlereth, A., & Tiedemann, J.** (2001). Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *Journal of experimental botany*, 52(362), 1741–1752. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11520862>
- Müntz, K., & Shutov, A. D.** (2002). Legumains and their functions in plants. *Trends in Plant Science*, 7, 340–344. doi:10.1016/S1360-1385(02)02298-7
- Murienne, J., Grandcolas, P., Piulachs, M. D., Bellés, X., D'Haese, C., Legendre, F., Pellens, R., & Guilbert, E.** (2005). Cladistics Evolution on a shaky piece of Gondwana□: is local endemism recent in New Caledonia ? *Cladistics*, 21, 2–7.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J.** (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858. doi:10.1038/35002501
- Naested, H., Frandsen, G. I., Jauh, G.-Y., Hernandez-Pinzon, I., Nielsen, H. B., Murphy, D. J., Rogers, J. C., & Mundy, J.** (2000). Caleosins: Ca<sup>2+</sup>-binding proteins associated with lipid bodies. *Plant molecular biology*, 44, 463–476. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11197322>
- Nakabayashi, K., Bartsch, M., Xiang, Y., Miatton, E., Pellengahr, S., Yano, R., Seo, M., & Soppe, W. J. J.** (2012). The time required for dormancy release in *Arabidopsis* is determined by DELAY OF GERMINATION1 protein levels in freshly harvested seeds. *The Plant cell*, 24(7), 2826–2838. doi:10.1105/tpc.112.100214
- Nakajima, S., Ito, H., Tanaka, R., & Tanaka, A.** (2012). Chlorophyll b reductase plays an essential role in maturation and storability of *Arabidopsis* seeds. *Plant physiology*, 160, 261–273. doi:10.1104/pp.112.196881
- Nambara, E., Okamoto, M., Tatematsu, K., Yano, R., Seo, M., & Kamiya, Y.** (2010). Abscisic acid and the control of seed dormancy and germination. *Seed Science Research*, 20(02), 55. doi:10.1017/S0960258510000012
- Nguyen, T.-P., Keizer, P., van Eeuwijk, F., Smeekens, S., & Bentsink, L.** (2012). Natural variation for seed longevity and seed dormancy are negatively correlated in *Arabidopsis*. *Plant physiology*, 160, 2083–2092. doi:10.1104/pp.112.206649
- Nishida, S., Tsuzuki, C., Kato, A., Aisu, A., Yoshida, J., & Mizuno, T.** (2011). AtIRT1 , the Primary Iron Uptake Transporter in the Root , Mediates Excess Nickel Accumulation in *Arabidopsis thaliana*, 52(8), 1433–1442. doi:10.1093/pcp/pcr089
- Nonogaki, H., Bassel, G. W., & Bewley, J. D.** (2010). Germination—Still a mystery. *Plant Science*, 179, 574–581. doi:10.1016/j.plantsci.2010.02.010
- Ogawa, M., Hanada, A., Yamauchi, Y., Kuwahara, A., Kamiya, Y., & Yamaguchi, S.** (2003). Gibberellin Biosynthesis and Response during *Arabidopsis* Seed Germination. *The Plant cell*, 15, 1591–1604. doi:10.1105/tpc.011650.ble

- Ogé, L., Bourdais, G., Bove, J., Collet, B., Godin, B., Granier, F., Boutin, J-P., Job, D., Jullien, M., & Grappin, P.** (2008). Protein repair L-isoaspartyl methyltransferase 1 is involved in both seed longevity and germination vigor in *Arabidopsis*. *The Plant cell*, *20*, 3022–3037. doi:10.1105/tpc.108.058479
- Oracz, K., El-Maarouf Bouteau, H., Farrant, J. M., Cooper, K., Belghazi, M., Job, C., Job, D., Corbineau, F., & Bailly, C.** (2007). ROS production and protein oxidation as a novel mechanism for seed dormancy alleviation. *The Plant journal*, *50*, 452–465. doi:10.1111/j.1365-313X.2007.03063.x
- Oracz, K., El-Maarouf-Bouteau, H., Kranner, I., Bogatek, R., Corbineau, F., & Bailly, C.** (2009). The mechanisms involved in seed dormancy alleviation by hydrogen cyanide unravel the role of reactive oxygen species as key factors of cellular signaling during germination. *Plant physiology*, *150*, 494–505. doi:10.1104/pp.109.138107
- Osborne, T. B.** (1924). *The vegetable proteins* (2nd editio., p. 125). London Longmans, Green and Co. Retrieved from [http://books.google.com/books/about/The\\_vegetable\\_proteins.html?id=MVFKAAAAMAAJ&pgis=1](http://books.google.com/books/about/The_vegetable_proteins.html?id=MVFKAAAAMAAJ&pgis=1)
- Pantazaki, A. A., Tsolkas, G. P., & Kyriakidis, D. A.** (2008). A DING phosphatase in *Thermus thermophilus*. *Amino acids*, *34*, 437–448. doi:10.1007/s00726-007-0549-5
- Parkinson, C. L., Adams, K. L., & Palmer, J. D.** (1999). Multigene analyses identify the three earliest lineages of extant flowering plants. *Current biology*, *9*(24), 1485–1488. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10607592>
- Pelletier, B.** (1982). *Origine et nature des minerais de nickel de Nouvelle-Calédonie* (SLN, Servi., pp. 21–47).
- Pelletier, B.** (2007). Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. In C. E. Payri & B. Richer de Forges (Eds.), *Compendium of marine species of New Caledonia* (IRD Nouméa., pp. 19–32). Nouméa.
- Penfield, S., Graham, S., & Graham, I. A.** (2005). Storage reserve mobilization in germinating oilseeds: *Arabidopsis* as a model system. *Biochemical Society transactions*, *33*, 380–383. doi:10.1042/BST0330380
- Perera, T., Berna, A., Scott, K., Lemaitre-Guillier, C., & Bernier, F.** (2008). Proteins related to St. John's Wort p27SJ, a suppressor of HIV-1 expression, are ubiquitous in plants. *Phytochemistry*, *69*, 865–872. doi:10.1016/j.phytochem.2007.10.001
- Perrier, N., Colin, F., Jaffré, T., Ambrosi, J.-P., Rose, J., & Bottero, J.-Y.** (2004). Nickel speciation in *Sebertia acuminata*, a plant growing on a lateritic soil of New Caledonie. *Comptes Rendus Geosciences*, *336*, 567–577. doi:10.1016/j.crte.2003.12.014
- Petrucelli, S., Molina, M. I., Lareu, F. J., & Circosta, A.** (2007). Two short sequences from amaranth 11S globulin are sufficient to target green fluorescent protein and beta-glucuronidase to vacuoles in *Arabidopsis* cells. *Plant Physiology and Biochemistry*, *45*, 400–409. doi:10.1016/j.plaphy.2007.02.008
- Pillon, Y.** (2012). Time and tempo of diversification in the flora of New Caledonia. *Botanical Journal of the Linnean Society*, *170*, 288–298.
- Plaxton, W. C.** (1996). The Organization and Regulation of Plant Glycolysis. *Annual review of plant physiology and plant molecular biology*, *47*, 185–214. doi:10.1146/annurev.arplant.47.1.185
- Poxleitner, M., Rogers, S. W., Lacey, S. A., Browse, J., & Rogers, J. C.** (2006). A role for caleosin in degradation of oil-body storage lipid during seed germination. *The Plant journal*, *47*, 917–933. doi:10.1111/j.1365-313X.2006.02845.x

- Preston, J., Tatematsu, K., Kanno, Y., Hobo, T., Kimura, M., Jikumaru, Y., Yano, R., Kamiya, Y., & Nambara, E.** (2009). Temporal expression patterns of hormone metabolism genes during imbibition of *Arabidopsis thaliana* seeds: a comparative study on dormant and non-dormant accessions. *Plant & cell physiology*, *50*(10), 1786–1800. doi:10.1093/pcp/pcp121
- Puchtler, H., Meloan, S. N., & Brewton, B. R.** (1975). On the history of basic fuchsin and aldehyde-Schiff reactions from 1862 to 1935. *Histochemistry*, *41*(3), 185–94. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1090565>
- Qiu, Y.-L., Lee, J., Bernasconi-Quadroni, F., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A., Chen, Z., Savolainen, V., & Chase, M. W.** (1999). The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature*, *402*(6760), 404–407. doi:10.1038/46536
- Rabilloud, T., Chevallet, M., Luche, S., & Lelong, C.** (2010). Two-dimensional gel electrophoresis in proteomics: Past, present and future. *Journal of proteomics*, *73*, 2064–2077. doi:10.1016/j.jprot.2010.05.016
- Rajjou, L., Gallardo, K., Debeaujon, I., Vandekerckhove, J., Job, C., & Job, D.** (2004). The Effect of a-Amanitin on the *Arabidopsis* Seed Proteome Highlights the Distinct Roles of Stored and Neosynthesized mRNAs during Germination 1. *Plant Physiology*, *134*, 1598–1613. doi:10.1104/pp.103.036293.1598
- Rajjou, L., Belghazi, M., Huguet, R., Robin, C., Moreau, A., Job, C., & Job, D.** (2006). Proteomic Investigation of the Effect of Salicylic Acid on *Arabidopsis* Seed Germination and Establishment of Early Defense Mechanisms. *Plant Physiology*, *141*, 910–923. doi:10.1104/pp.106.082057.1994
- Rajjou, L., & Debeaujon, I.** (2008). Seed longevity: survival and maintenance of high germination ability of dry seeds. *Comptes rendus biologiques*, *331*, 796–805. doi:10.1016/j.crv.2008.07.021
- Rajjou, L., Lovigny, Y., Groot, S. P. C., Belghazi, M., Job, C., & Job, D.** (2008). Proteome-wide characterization of seed aging in *Arabidopsis*: a comparison between artificial and natural aging protocols. *Plant physiology*, *148*, 620–641. doi:10.1104/pp.108.123141
- Rajjou, L., Duval, M., Gallardo, K., Catusse, J., Bally, J., Job, C., & Job, D.** (2012). Seed germination and vigor. *Annual review of plant biology*, *63*, 507–533. doi:10.1146/annurev-arplant-042811-105550
- Rapala-Kozik, M., Wolak, N., Kujda, M., & Banas, A. K.** (2012). The upregulation of thiamine (vitamin B1) biosynthesis in *Arabidopsis thaliana* seedlings under salt and osmotic stress conditions is mediated by abscisic acid at the early stages of this stress response. *BMC plant biology*, *12*, 2. doi:10.1186/1471-2229-12-2
- Ravanel, S., Gakière, B., Job, D., & Douce, R.** (1998). The specific features of methionine biosynthesis and metabolism in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 7805–7812. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=22764&tool=pmcentrez&rendertype=abstract>
- Reeves, R. D., & Baker, J. M.** (1984). Studies on metal uptake by plants from serpentine and non-serpentine populations of *Thlaspi goesingense* Halácsy (Cruciferae). *New Phytologist*, *98*, 191–204.
- Reis, V. M., Estrada-de los Santos, P., Tenorio-Salgado, S., Vogel, J., Stoffels, M., Guyon, S., Mavingui, P., Baldani, V. L. D., Schmid, M., Baldani, J. I., Balandreau, J., Hartmann, A., Caballero-Mellado, J.** (2004). *Burkholderia tropica* sp. nov., a novel nitrogen-fixing, plant-associated bacterium. *International journal of systematic and evolutionary microbiology*, *54*, 2155–2162. doi:10.1099/ijs.0.02879-0

- Roe, J. H.** (1995). The determination of sugar in blood and spinal fluid with anthrone reagent. *Journal of Biological Chemistry*, 212, 335–343.
- Rudall, P. J., Eldridge, T., Tratt, J., Ramsay, M. M., Tuckett, R. E., Smith, S. Y., Collinson, M. E., Remizowa, M. V., & Sokoloff, D. D.** (2009). Seed fertilization, development, and germination in Hydatellaceae (Nymphaeales): Implications for endosperm evolution in early angiosperms. *American journal of botany*, 96(9), 1581–1593. doi:10.3732/ajb.0900033
- Russell, B. L., Rathinasabapathi, B., & Hanson, A. D.** (1998). Osmotic stress induces expression of choline monooxygenase in sugar beet and Amaranth. *Plant physiology*, 116, 859–865. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=35146&tool=pmcentrez&rendertype=abstract>
- Sagner, S., Kneer, R., Wanner, G., Cosson, J.-P., Deus-Neumann, B., & Zenk, M. H.** (1998). Hyperaccumulation, complexation and distribution of Nickel in *Sebertia acuminata*. *Phytochemistry*, 47(3), 339–347.
- Sano, N., Permana, H., Kumada, R., Shinozaki, Y., Tanabata, T., Yamada, T., Hirasawa, T., & Kanekatsu, M.** (2012). Proteomic analysis of embryonic proteins synthesized from long-lived mRNAs during germination of rice seeds. *Plant & cell physiology*, 53(4), 687–698. doi:10.1093/pcp/pcs024
- Sansberro, P. A., Rey, H. Y., Mroginski, L. A., & Collavino, M. M.** (1998). In Vitro Culture of Rudimentary Embryos of *Ilex paraguariensis*: Responses to Exogenous Cytokinins. *Journal of Plant Growth Regulation*, 17, 101–105. doi:10.1007/PL00007014
- Santos, P. E. L., Bustillos-cristales, R., & Ecologi, P. De.** (2001). Burkholderia , a Genus Rich in Plant-Associated Nitrogen Fixers with Wide Environmental and Geographic Distribution. *Applied and environmental microbiology*, 67(6), 2790–2798. doi:10.1128/AEM.67.6.2790
- Savolainen, V., Chase, M. W., Hoot, S. B., Morton, C. M., Soltis, D. E., Bayer, C., Fay, M. F., de Bruijn, A. Y., Sullivan, S., Qiu, Y.-L.** (2000). Phylogenetics of flowering plants based on combined analysis of plastid atpB and rbcL gene sequences. *Systematic biology*, 49(2), 306–362. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12118410>
- Schaaf, G., Honsbein, A., Meda, A. R., Kirchner, S., Wipf, D., & von Wirén, N.** (2006). AtIREG2 encodes a tonoplast transport protein involved in iron-dependent nickel detoxification in *Arabidopsis thaliana* roots. *The Journal of biological chemistry*, 281(35), 25532–25540. doi:10.1074/jbc.M601062200
- Schickler, H., & Caspi, H.** (1999). Response of antioxidative enzymes to nickel and cadmium stress in hyperaccumulator plants of the genus *Alyssum*. *Physiologia plantarum*, 105, 39–44.
- Scott, K., & Wu, L.** (2005). Functional properties of a recombinant bacterial DING protein: comparison with a homologous human protein. *Biochimica et biophysica acta*, 1744(2), 234–244. doi:10.1016/j.bbamcr.2005.02.003
- Scott, M. P., Jung, R., Muntz, K., & Nielsen, N. C.** (1992). A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean. *Proceedings of the National Academy of Sciences of the United States of America*, 89(2), 658–662. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=48298&tool=pmcentrez&rendertype=abstract>



- Sedbrook, J. C., Carroll, K. L., Hung, K. F., Masson, P. H., & Somerville, C. R.** (2002). The *Arabidopsis* SKU5 Gene Encodes an Extracellular Glycosyl Phosphatidylinositol – Anchored Glycoprotein Involved in Directional Root Growth. *The Plant cell*, 14(July), 1635–1648. doi:10.1105/tpc.002360.1996
- Sen, S., & Osborne, D. J.** (1977). Decline in ribonucleic acid and protein synthesis with loss of viability during the early hours of imbibition of rye (*Secale cereale* L.) embryos. *The Biochemical journal*, 166, 33–38. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1164953&tool=pmcentrez&rendertype=abstract>
- Seregin, I. V., & Kozhevnikova, A. D.** (2006). Physiological role of nickel and its toxic effects on higher plants. *Russian Journal of Plant Physiology*, 53(2), 257–277. doi:10.1134/S1021443706020178
- Sharma, A., & Dhiman, A.** (2013). Nickel and Cadmium toxicity in plants. *Journal of Pharmaceutical and Scientific Innovation*, 2(2), 20–24. doi:10.7897/2277-4572.02213
- Sharma, S. S., & Dietz, K.-J.** (2009). The relationship between metal toxicity and cellular redox imbalance. *Trends in plant science*, 14(1), 43–50. doi:10.1016/j.tplants.2008.10.007
- Shen-Miller, J., Mudgett, M. B., Schopf, J. W., Clarke, S., & Berger, R.** (1995). Exceptional Seed Longevity and Robust Growth: Ancient Sacred Lotus from China. *American journal of botany*, 82(11), 1367–1380.
- Shen-Miller, J.** (2002). Sacred lotus, the long-living fruits of China Antique. *Seed Science Research*, 12(03), 131–143.
- Shen-Miller, J., Aung, L. H., Turek, J., Schopf, J. W., Tholandi, M., Yang, M., & Czaja, A.** (2013). Centuries-Old Viable Fruit of Sacred Lotus *Nelumbo nucifera* Gaertn var. China Antique. *Tropical Plant Biology*, August. doi:10.1007/s12042-013-9125-1
- Shevchenko, A., Wilm, M., Vorm, O., & Mann, M.** (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical chemistry*, 68(5), 850–858. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8779443>
- Shewry, P. R., Napier, J. A., & Tatham, A. S.** (1995). Seed storage proteins: structures and biosynthesis. *The Plant cell*, 7, 945–956. doi:10.1105/tpc.7.7.945
- Shewry, P. R., & Halford, N. G.** (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of experimental botany*, 53(370), 947–958. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11912237>
- Shimada, T., Fuji, K., Tamura, K., Kondo, M., Nishimura, M., & Hara-Nishimura, I.** (2003). Vacuolar sorting receptor for seed storage proteins in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26), 16095–16100. doi:10.1073/pnas.2530568100
- Shintani, D., & DellaPenna, D.** (1998). Elevating the Vitamin E Content of Plants Through Metabolic Engineering. *Science*, 282, 2098–2100. doi:10.1126/science.282.5396.2098
- Shutov, A. D., Bäumlein, H., Blattner, F. R., & Müntz, K.** (2003). Storage and mobilization as antagonistic functional constraints on seed storage globulin evolution. *Journal of experimental botany*, 54, 1645–1654. doi:10.1093/jxb/erg165
- Sirko, A., & Brodzik, R.** (2000). Plant ureases: roles and regulation. *Acta biochimica Polonica*, 47(4), 1189–1195. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11996109>
- Skylas, D. J., Mackintosh, J. A., Cordwell, S. J., Basseal, D. J., Walsh, B. J., Harry, J., Blumenthal, C., Copeland, L., Wrigley, C. W., & Rathmell, W.** (2000). Proteome Approach to the Characterisation of Protein Composition in the Developing and Mature Wheat-grain Endosperm. *Journal of Cereal Science*, 32, 169–188. doi:10.1006/jcrs.2000.0321

- Smith, C. A. D., & Bray, C. M.** (1982). Intracellular levels of polyadenylated RNA and loss of vigour in germinating wheat embryos. *Planta*, *156*, 413–420.
- Soltis, D. E., & Soltis, P. S.** (2004). *Amborella* not a “basal angiosperm”? Not so fast. *American Journal of Botany*, *91*(6), 997–1001.
- Soltis, D. E., Albert, V. a, Leebens-Mack, J., Palmer, J. D., Wing, R. a, DePamphilis, C. W., Ma, H., Carlson, J. E., Altman, N., Kim, S., Wall, P. K., Zuccolo, A., & Soltis, P. S.** (2008). The *Amborella* genome: an evolutionary reference for plant biology. *Genome biology*, *9*, 402. doi:10.1186/gb-2008-9-3-402
- Soltis, D. E., Smith, S. a, Cellinese, N., Wurdack, K. J., Tank, D. C., Brockington, S. F., Refulio-Rodriguez, N. F., Walker, J. B., Moore, M. J., Carlswald, B. S., Bell, C. D., Latvis, M., Crawley, S., Black, C., Diouf, D., Xi, Z., Rushworth, C. A., Gitzendanner, M. A., Sytsma, K. J., Qiu, Y-L., Hilu, K. W., Davis, C. C., Sanderson, M. J., Beaman, R. S., Olmstead, R. G., Judd, W. S., Donoghue, M. J., & Soltis, P. S.** (2011). Angiosperm phylogeny: 17 genes, 640 taxa. *American journal of botany*, *98*(4), 704–730. doi:10.3732/ajb.1000404
- Soltis, P. S., Soltis, D. E., & Chase, M. W.** (1999). Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature*, *402*, 402–404. doi:10.1038/46528
- Sreenivasulu, N., Usadel, B., Winter, A., Radchuk, V., Scholz, U., Stein, N., Weschke, W., Strickert, M., Close, T. J., Stitt, M., Graner, A., Wobus, U.** (2008). Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant physiology*, *146*, 1738–1758. doi:10.1104/pp.107.111781
- Steinbauer, G. P.** (1937). Dormancy and germination of *Fraxinus* seeds. *Plant physiology*, *12*, 813–824.
- Stengel, F., Baldwin, A. J., Painter, A. J., Jaya, N., Basha, E., Kay, L. E., Vierling, E., Robinson, C. V., & Benesch, J. L. P.** (2010). Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 2007–2012. doi:10.1073/pnas.0910126107
- Stupnikova, I., Benamar, A., Tolleter, D., Grelet, J., Borovskii, G., Dorne, A., & Macherel, D.** (2006). Pea Seed Mitochondria Are Endowed with a Remarkable Tolerance to Extreme Physiological Temperatures. *Plant Physiology*, *140*, 326–335. doi:10.1104/pp.105.073015.326
- Su, P.-H., & Li, H.-M.** (2008). *Arabidopsis* stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. *Plant physiology*, *146*, 1231–1241. doi:10.1104/pp.107.114496
- Subbiah, V., & Reddy, K. J.** (2010). Interactions between ethylene, abscisic acid and cytokinin during germination and seedling establishment in *Arabidopsis*. *Journal of Biosciences*, *35*, 451–458. doi:10.1007/s12038-010-0050-2
- Sun, T.-P., & Gubler, F.** (2004). Molecular mechanism of gibberellin signaling in plants. *Annual review of plant biology*, *55*, 197–223. doi:10.1146/annurev.arplant.55.031903.141753
- Sun, W., Bernard, C., van de Cotte, B., Van Montagu, M., & Verbruggen, N.** (2001). At-HSP17.6A, encoding a small heat-shock protein in *Arabidopsis*, can enhance osmotolerance upon overexpression. *The Plant journal*, *27*(5), 407–415. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11576425>

- Tai, S. S. K., Chen, M. C. M., Peng, C.-C., & Tzen, J. T. C.** (2002). Gene family of oleosin isoforms and their structural stabilization in sesame seed oil bodies. *Bioscience, biotechnology, and biochemistry*, 66(10), 2146–2153. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12450125>
- Takahashi, T., & Komeda, Y.** (1989). Characterization of two genes encoding small heat-shock proteins in *Arabidopsis thaliana*. *Molecular & general genetics*, 219, 365–372. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2482931>
- Talanova, V. V., Titov, A. F., & Boeva, N. P.** (2000). Effect of increasing concentrations of lead and cadmium on cucumber seedlings. *Biologia plantarum*, 43(3), 441–444.
- Tandang-Silvas, M. R. G., Fukuda, T., Fukuda, C., Prak, K., Cabanos, C., Kimura, A., Itoh, T., Mikami, B., Utsumi, S., & Maruyama, N.** (2010). Conservation and divergence on plant seed 11S globulins based on crystal structures. *Biochimica et biophysica acta*, 1804, 1432–1442. doi:10.1016/j.bbapap.2010.02.016
- Teng, S., Rognoni, S., Bentsink, L., & Smeekens, S.** (2008). The *Arabidopsis* GSQ5/DOG1 Cvi allele is induced by the ABA-mediated sugar signalling pathway, and enhances sugar sensitivity by stimulating ABI4 expression. *The Plant journal: for cell and molecular biology*, 55(3), 372–381. doi:10.1111/j.1365-3113X.2008.03515.x
- Thapar, N., Kim, A.-K., & Clarke, S.** (2001). Distinct patterns of expression but similar biochemical properties of protein L-isoaspartyl methyltransferase in higher plants. *Plant physiology*, 125, 1023–10235. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=64902&tool=pmcentrez&rendertype=abstract>
- Thomma, B. P. H. J., Eggermont, K., Penninckx, I. A. M. A., Mauch-Mani, B., Vogelsang, R., Cammue, B. P. A., & Broekaert, W. F.** (1998). Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 15107–15111. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=24583&tool=pmcentrez&rendertype=abstract>
- Tobe, H., Jaffré, T., & Raven, P. H.** (2000). Embryology of *Amborella* (Amborellaceae): Descriptions and polarity of character states. *Journal of plant research*, 113, 271–280.
- Towbin, H., Staehelin, T., & Gordon, J.** (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, 76(9), 4350–4354. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1422008>
- Trevelyan, W. E., Forrest, R. S., & Harrison, J. S.** (1952). Determination of Yeast Carbohydrates with the Anthrone Reagent. *Nature*, 170, 626–627. doi:10.1038/170626a0
- Tuckett, R. E., Merritt, D. J., Rudall, P. J., Hay, F., Hopper, S. D., Baskin, C. C., Baskin, J. M., Tratt, J., & Dixon, K. W.** (2010). A new type of specialized morphophysiological dormancy and seed storage behaviour in Hydatellaceae, an early-divergent angiosperm family. *Annals of botany*, 105, 1053–1061. doi:10.1093/aob/mcq062
- Tunnacliffe, A., & Wise, M. J.** (2007). The continuing conundrum of the LEA proteins. *Die Naturwissenschaften*, 94, 791–812. doi:10.1007/s00114-007-0254-y
- Tweddle, J. C., Dickie, J. B., Baskin, C. C., & Baskin, J. M.** (2003). Ecological aspects of seed desiccation sensitivity. *Journal of Ecology*, 91, 294–304. doi:10.1046/j.1365-2745.2003.00760.x

- Tzen, J. T. C., Cao, Y., Laurent, P., Ratnayake, Cc., & Huang, A. H. C.** (1993). Lipids, Proteins, and Structure of Seed Oil Bodies from Diverse Species. *Plant physiology*, *101*, 267–276. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=158673&tool=pmcentrez&rendertype=abstract>
- Valot, B., Langella, O., Nano, E., & Zivy, M.** (2011). MassChroQ: a versatile tool for mass spectrometry quantification. *Proteomics*, *11*(17), 3572–7. doi:10.1002/pmic.201100120
- Vandelook, F., Bolle, N., & Van Assche, J. A.** (2007). Seed dormancy and germination of the European *Chaerophyllum temulum* (Apiaceae), a member of a trans-Atlantic genus. *Annals of botany*, *100*, 233–239. doi:10.1093/aob/mcm090
- Varner, J. E., & Chandra, G. R.** (1964). Hormonal control of enzyme synthesis in Barley endosperm. *Proceedings of the National Academy of Sciences of the United States of America*, *52*, 100–106.
- Vega, F. E., Pava-Ripoll, M., Posada, F., & Buyer, J. S.** (2005). Endophytic bacteria in *Coffea arabica* L. *Journal of basic microbiology*, *45*, 371–380. doi:10.1002/jobm.200410551
- Verbruggen, N., Hermans, C., & Schat, H.** (2009). Molecular mechanisms of metal hyperaccumulation in plants. *The New phytologist*, *181*, 759–776. doi:10.1111/j.1469-8137.2008.02748.x
- Verma, P., Kaur, H., Petla, B. P., Rao, V., Saxena, S. C., & Majee, M.** (2013). PROTEIN L-ISOASPARTYL METHYLTRANSFERASE2 is differentially expressed in chickpea and enhances seed vigor and longevity by reducing abnormal isoaspartyl accumulation predominantly in seed nuclear proteins. *Plant physiology*, *161*, 1141–1157. doi:10.1104/pp.112.206243
- Verniquet, F., Gaillard, J., Neuburger, M., & Douce, R.** (1991). Rapid inactivation of plant aconitase by hydrogen peroxide. *The Biochemical journal*, *276*, 643–648. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1151053&tool=pmcentrez&rendertype=abstract>
- Vert, G., Grotz, N., Dédaldéchamp, F., Gaynard, F., Guerinot, L., Briat, J., & Curie, C.** (2002). IRT1, an *Arabidopsis* Transporter Essential for Iron Uptake from the Soil and for Plant Growth. *The Plant cell*, *14*, 1223–1233. doi:10.1105/tpc.001388.Nongrasses
- Vicente-Carbajosa, J., & Carbonero, P.** (2005). Seed maturation: developing an intrusive phase to accomplish a quiescent state. *The International journal of developmental biology*, *49*, 645–651. doi:10.1387/ijdb.052046jc
- Vigneswara, V., Lowenson, J. D., Powell, C. D., Thakur, M., Bailey, K., Clarke, S., Ray, D. E., Carter, W. G.** (2006). Proteomic identification of novel substrates of a protein isoaspartyl methyltransferase repair enzyme. *The Journal of biological chemistry*, *281*, 32619–32629. doi:10.1074/jbc.M605421200
- Villegente, M., Burtet-Sarramégn, V., Rajjou, L., Job, C., Job, D., & Fogliani, B.,** (2011). Caractérisation des processus biochimiques ayant cours lors de la germination d’*Amborella trichopoda*. Colloque GRAINES 2011, Nantes, France, 27 & 28 Octobre 2011 (Poster)
- Villegente, M., Zongo, C., Job, C., Job, D., Van Dorsselaer, A., Strub, J.-M., Burtet-Sarramegna, V., & Fogliani B.,** (2011). Dissecting the protective Strategy developed by the seed of the nickel hyperaccumulator shrub *Psychotria douarrei*. 10th International Conference on Seed Science, Salvador-Costa do Sauipe, Bahia, Brazil, 10-15 avril 2011

- Villegente, M., Boissard, Y., Carteron, A., Leveau, A., & Fogliani, B.,** (2012). On the study of New Caledonian dormant native seeds for ecological restoration after mining impact using hydroseeding on the Massif du Koniambo. Inaugural conference of the Society for Ecological Restoration Australasia, Perth, Australia, 27-30 Novembre 2012. (Communication orale)
- Violette, Z.** (2009). *La biodiversité de la Nouvelle-Calédonie*. (J. Viratelle, J. Muzinger, & H. Jourdan, Eds.) *Maison de la Nouvelle-Calédonie* (p. 42). Paris: Horizon Pacifique.
- Vogel-Mikus, K., Pongrac, P., Kump, P., Necemer, M., Simcic, J., Pelicon, P., Budnar, M., Povh, B., & Regvar, M.** (2007). Localisation and quantification of elements within seeds of Cd/Zn hyperaccumulator *Thlaspi praecox* by micro-PIXE. *Environmental pollution (Barking, Essex)*: 1987, 147, 50–59. doi:10.1016/j.envpol.2006.08.026
- Von Teichman, I., & van Wyk, A. E.** (1994). Structural aspects and trends in the evolution of recalcitrant seeds in dicotyledons. *Seed Science Research*, 4(02), 225–239. doi:10.1017/S096025850000221X
- Walther, T. C., & Mann, M.** (2010). Mass spectrometry-based proteomics in cell biology. *The Journal of cell biology*, 190, 491–500. doi:10.1083/jcb.201004052
- Wang, J., Li, Y., Lo, S. W., Hillmer, S., Sun, S. S. M., Robinson, D. G., & Jiang, L.** (2007). Protein mobilization in germinating mung bean seeds involves vacuolar sorting receptors and multivesicular bodies. *Plant physiology*, 143, 1628–1639. doi:10.1104/pp.107.096263
- Wang, X., Tang, C., Deng, L., Cai, G., Liu, X., Liu, B., Han, Q., Buchenauer, H., Wei, G., Han, D., Huang, L., & Kang, Z.** (2010). Characterization of a pathogenesis-related thaumatin-like protein gene TaPR5 from wheat induced by stripe rust fungus. *Physiologia plantarum*, 139, 27–38. doi:10.1111/j.1399-3054.2009.01338.x
- Washburn, M. P., Wolters, D., & Yates, J. R.** (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature biotechnology*, 19, 242–247. doi:10.1038/85686
- Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., & Humphery-Smith, I.** (1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*, 16, 1090–1094. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7498152>
- Wehmeyer, N., & Vierling, E.** (2000). The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance. *Plant physiology*, 122, 1099–1108. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58944&tool=pmcentrez&rendertype=abstract>
- Weiss, C., Bonshtien, A., Farchi-Pisanty, O., Vitlin, A., & Azem, A.** (2009). Cpn20: siamese twins of the chaperonin world. *Plant molecular biology*, 69, 227–238. doi:10.1007/s11103-008-9432-3
- Whiting, S. N., de Souza, M. P., & Terry, N.** (2001). Rhizosphere bacteria mobilize Zn for hyperaccumulation by *Thlaspi caerulescens*. *Environmental science & technology*, 35(15), 3144–3150. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11505990>
- Wise, M. J.** (2003). LEAPing to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC bioinformatics*, 4, 52. doi:10.1186/1471-2105-4-52
- Wobus, U., & Weber, H.** (1999). Seed maturation: genetic programmes and control signals. *Current opinion in plant biology*, 2, 33–38. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10047566>

- Wu, L. S. H., Hong, G. H. H., Hou, R. F., & Tzen, J. T. C.** (1999). Classification of the Single Oleosin Isoform and Characterization of Seed Oil Bodies in Gymnosperms. *Plant & cell physiology*, *40*(3), 326–334.
- Wuana, R. A., & Okieimen, F. E.** (2011). Heavy Metals in Contaminated Soils: A Review of Sources, Chemistry, Risks and Best Available Strategies for Remediation. *ISRN Ecology*, *2011*, 1–20. doi:10.5402/2011/402647
- Xi, W., Liu, C., Hou, X., & Yu, H.** (2010). MOTHER OF FT AND TFL1 regulates seed germination through a negative feedback loop modulating ABA signaling in *Arabidopsis*. *The Plant cell*, *22*, 1733–1748. doi:10.1105/tpc.109.073072
- Xu, Q., Belcastro, M. P., Villa, S. T., Dinkins, R. D., Clarke, S. G., & Downie, A. B.** (2004). A Second Protein L-Isoaspartyl Methyltransferase Gene in *Arabidopsis* Produces Two Transcripts Whose Products Are Sequestered in the Nucleus. *Plant Physiology*, *136*, 2652–2664. doi:10.1104/pp.104.046094.2652
- Yacoubi, R., Job, C., Belghazi, M., Chaibi, W., & Job, D.** (2010). Toward Characterizing Seed Vigor in Alfalfa Through Proteomic Analysis of Germination and Priming. *Journal of proteome research*.
- Yamaguchi, S.** (2008). Gibberellin metabolism and its regulation. *Annual review of plant biology*, *59*, 225–51. doi:10.1146/annurev.arplant.59.032607.092804
- Yamaguchi-Shinozaki, K., & Shinozaki, K.** (1993). *Arabidopsis* DNA encoding two desiccation-responsive rd29 genes. *Plant physiology*, *101*, 1119–1120. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=158736&tool=pmcentrez&rendertype=abstract>
- Yang, P., Li, X., Wang, X., Chen, H., Chen, F., & Shen, S.** (2007). Proteomic analysis of rice (*Oryza sativa*) seeds during germination. *Proteomics*, *7*, 3358–3368. doi:10.1002/pmic.200700207
- Yuan, J. S., Yang, X., Lai, J., Lin, H., Cheng, Z.-M., Nonogaki, H., & Chen, F.** (2007). The endo-beta-mannanase gene families in *Arabidopsis*, rice, and poplar. *Functional & integrative genomics*, *7*, 1–16. doi:10.1007/s10142-006-0034-3
- Yusuf, M., Fariduddin, Q., Hayat, S., & Ahmad, A.** (2011). Nickel: an overview of uptake, essentiality and toxicity in plants. *Bulletin of environmental contamination and toxicology*, *86*, 1–17. doi:10.1007/s00128-010-0171-1
- Zhang, L., Angle, J. S., Chaney, R. L., & Angle, S.** (2007). Do high-nickel leaves shed by the nickel hyperaccumulator *Alyssum murale* inhibit seed germination of competing plants? *New Phytologist*, *173*, 509–516.
- Zhang, X.-X., Scott, K., Meffin, R., & Rainey, P. B.** (2007). Genetic characterization of psp encoding the DING protein in *Pseudomonas fluorescens* SBW25. *BMC microbiology*, *7*, 114. doi:10.1186/1471-2180-7-114
- Zhang, Z., Ogawa, M., Fleet, C. M., Zentella, R., Hu, J., Heo, J.-O., Lim, J., Kamiya, Y., & Yamaguchi, S., Sun, T.** (2011). SCARECROW-LIKE 3 promotes gibberellin signaling by antagonizing master growth repressor DELLA in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(5), 2160–2165. doi:10.1073/pnas.1012232108/-DCSupplemental.[www.pnas.org/cgi/doi/10.1073/pnas.1012232108](http://www.pnas.org/cgi/doi/10.1073/pnas.1012232108)
- Zhao, F. J., Hamon, R. E., & McLaughlin, M. J.** (2001). Root exudates of the hyperaccumulator *Thlaspi caerulescens* do not enhance metal mobilization. *New Phytologist*, *151*, 613–620. doi:10.1046/j.0028-646x.2001.00213.x
- Zhu, J. X., Doyle, H. A., Mamula, M. J., & Aswad, D. W.** (2006). Protein repair in the brain, proteomic analysis of endogenous substrates for protein L-isoaspartyl methyltransferase in mouse brain. *The Journal of biological chemistry*, *281*, 33802–33813. doi:10.1074/jbc.M606958200

**Zongo, C.** (2010). *Données biologiques et physiologiques sur les semences de quelques espèces de milieux ultramafiques néo-calédoniens*. Thèse de doctorat, 228p. Université de la Nouvelle-Calédonie.

**Zongo, C., L'Huillier, L., Wulff, A., Villegente, M., Fogliani, B.,** 2012. Restoration in a south Pacific biodiversity hotspot, the case of New Caledonia. Inaugural conference of the Society for Ecological Restoration Australasia, Perth, Australia, 27-30 Novembre 2012. (Communication orale)





# **Annexes**



## **Annexe 1 : carte de référence de la fraction albumines de la graine d'*Amborella trichopoda***

Localisation des protéines albumines sur gel d'électrophorèse bidimensionnelle de référence de la graine totale mature sèche d'*Amborella*. Une quantité de 100 µg de protéines albumines a été déposée. Les protéines ont été séparées en première dimension par isoélectrofocalisation suivant un gradient de pH immobilisé 3-10 non linéaire puis en seconde dimension par électrophorèse SDS-PAGE sur des gels de polyacrylamide 10%. Les protéines ont été révélées par coloration au nitrate d'argent.

# Spots 1 à 46

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3

73,4

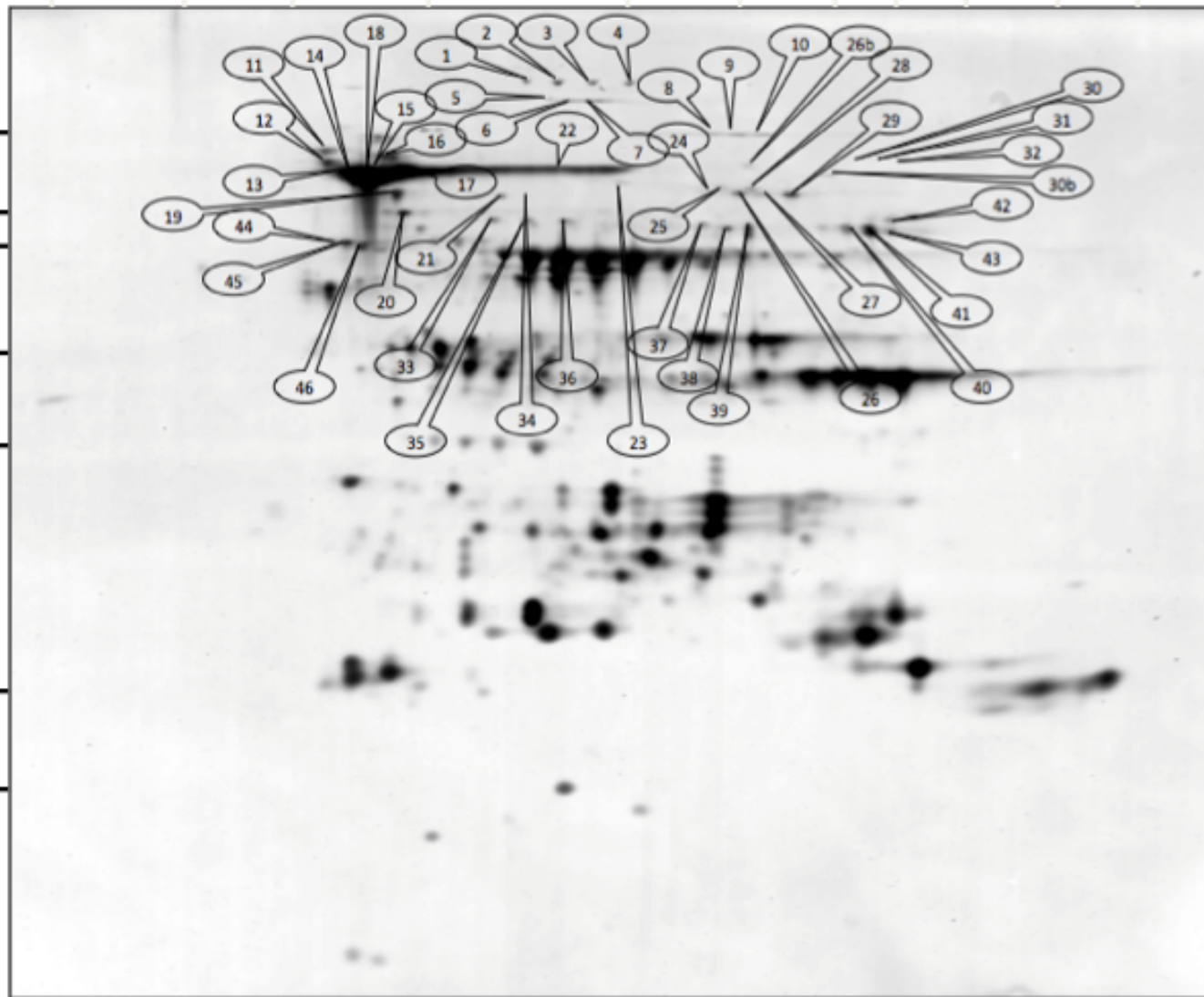
67,2

56,9

42,8

34,7

27,2

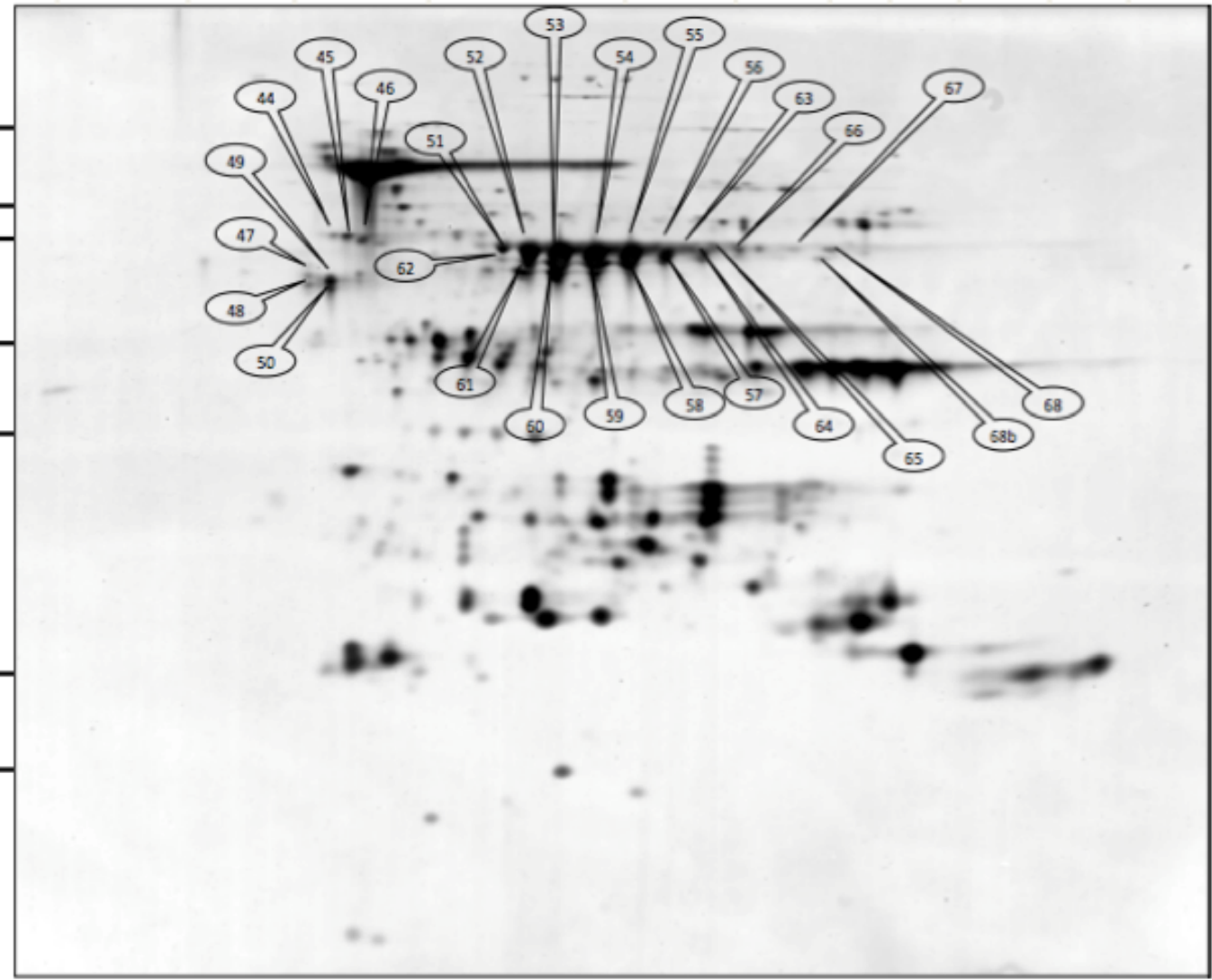


# Spots 47 à 68b

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3  
73,4  
67,2  
56,9  
42,8  
34,7  
27,2



# Spots 69 à 88

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3

73,4

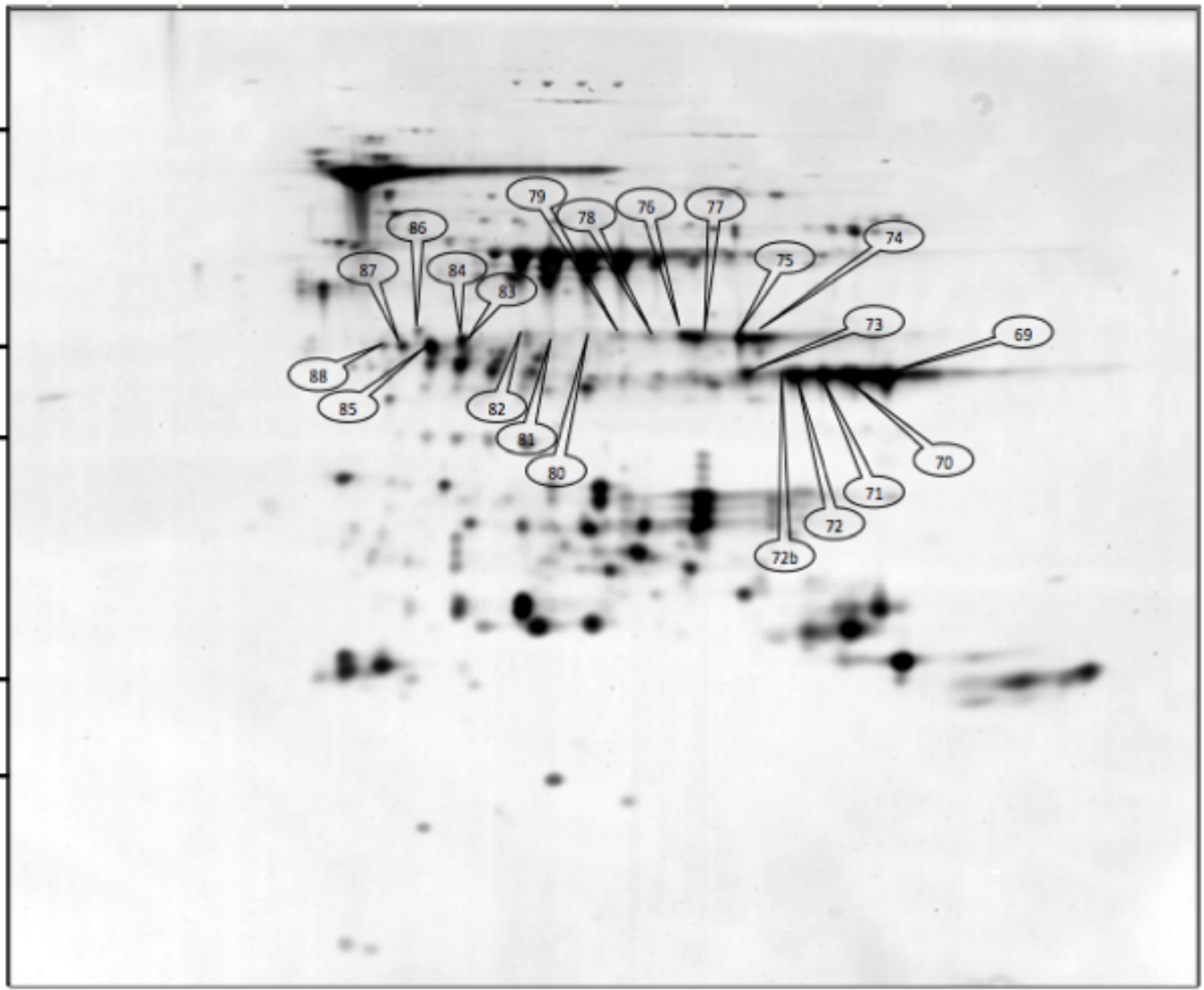
67,2

56,9

42,8

34,7

27,2



# Spots 89 à 105

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3

73,4

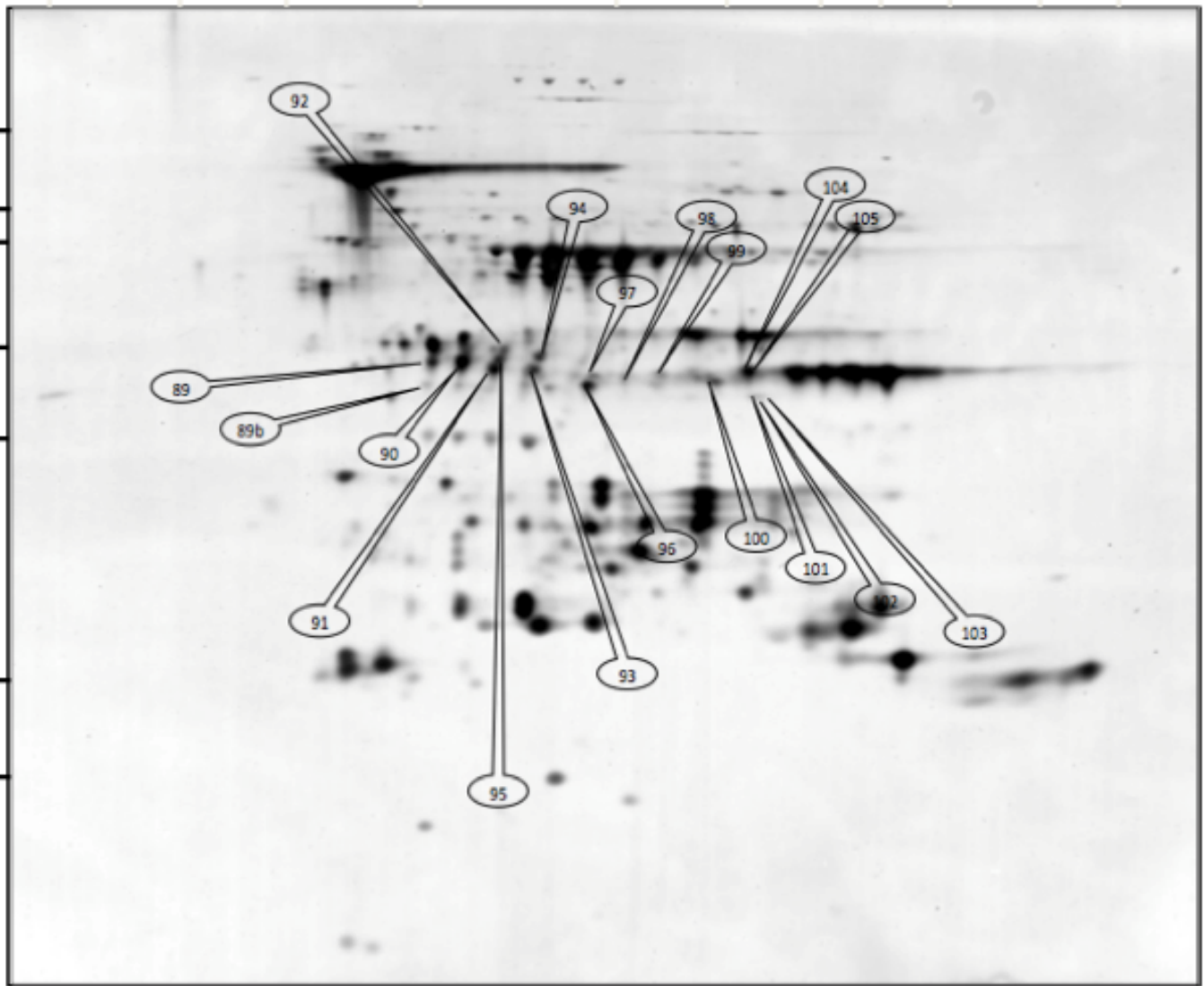
67,2

56,9

42,8

34,7

27,2



# Spots 106 à 128

pi  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3

73,4

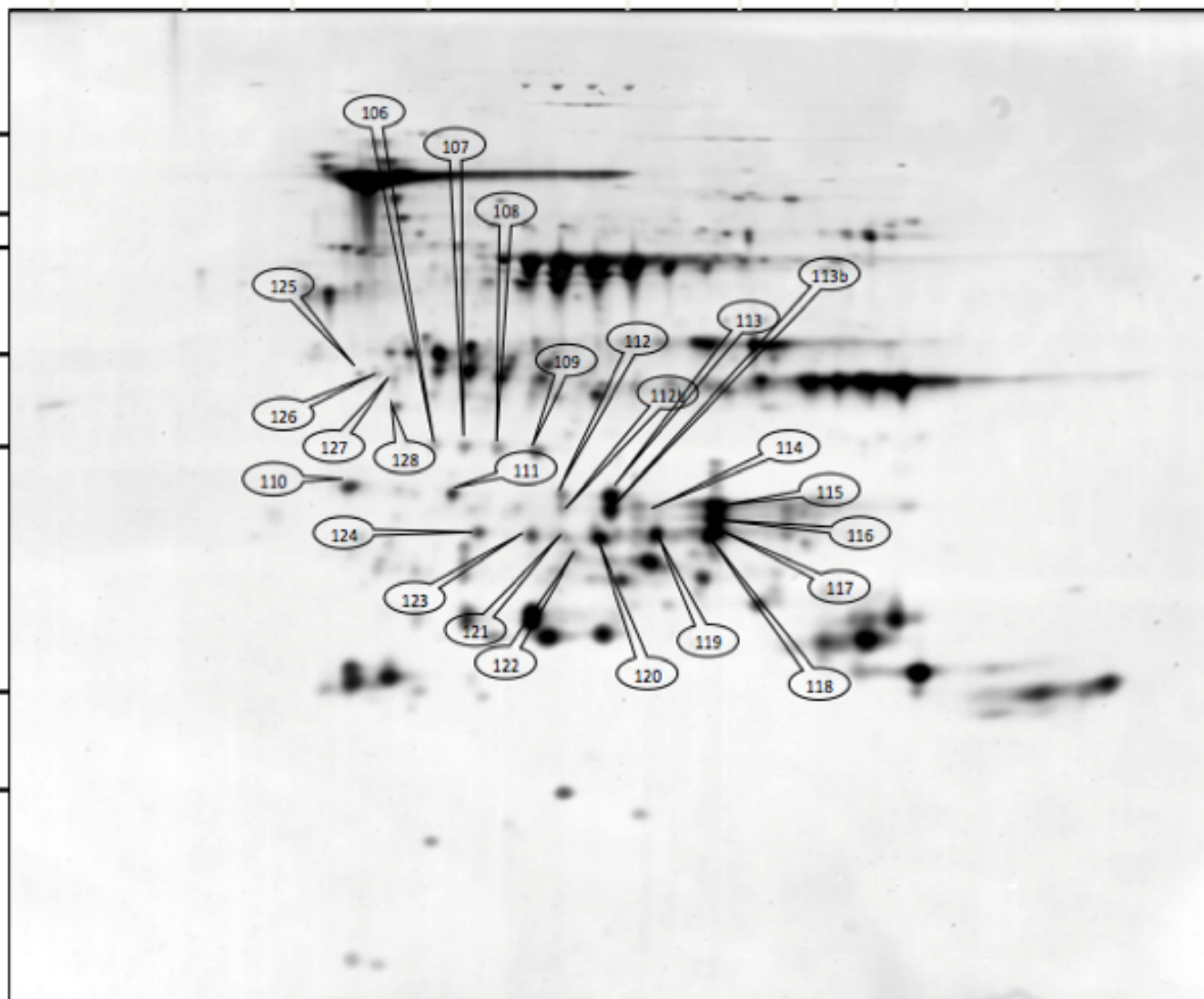
67,2

56,9

42,8

34,7

27,2



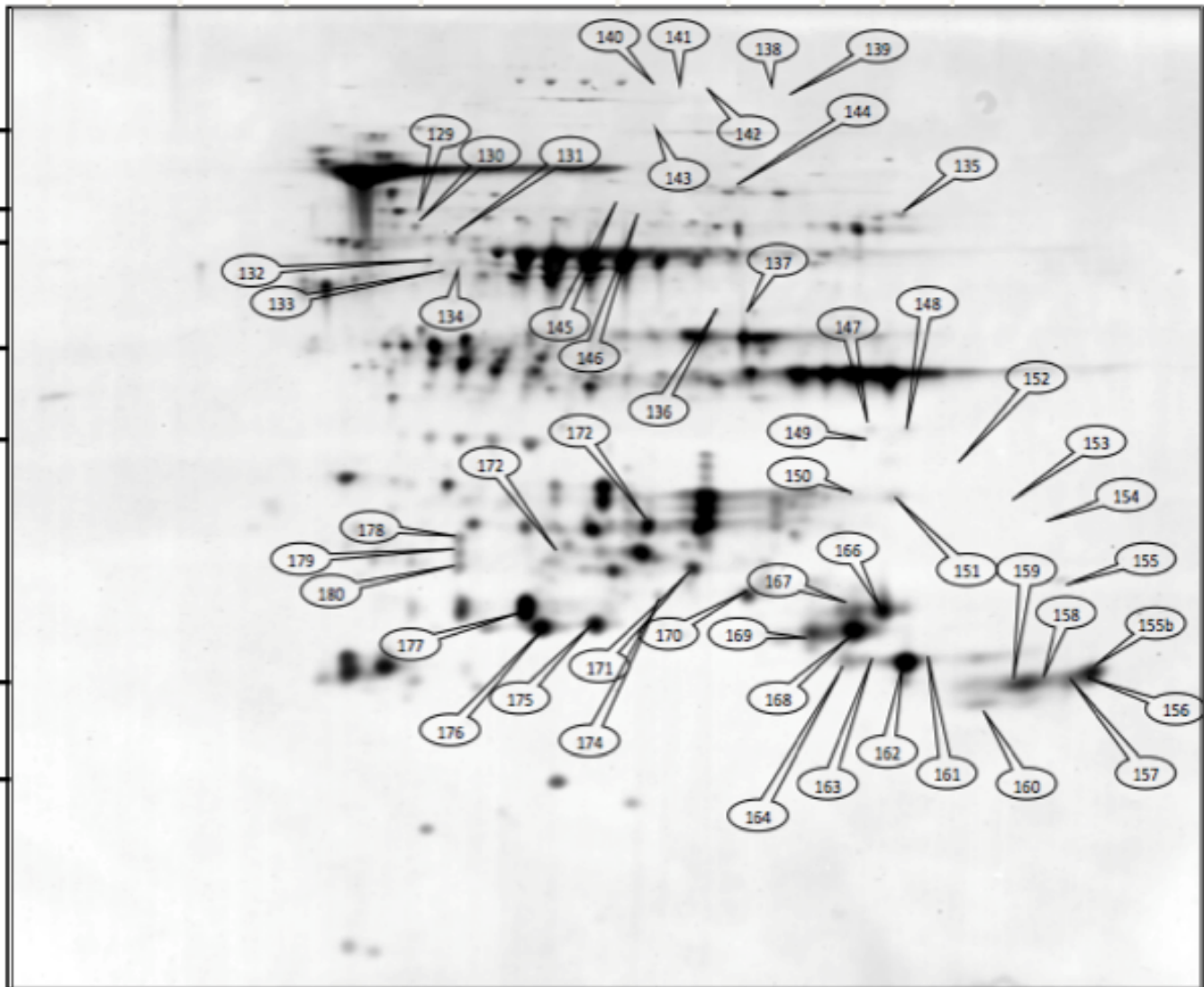


# Spots 129 à 180

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3  
73,4  
67,2  
56,9  
42,8  
34,7  
27,2



# Spots 181 à 220

pi  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3

73,4

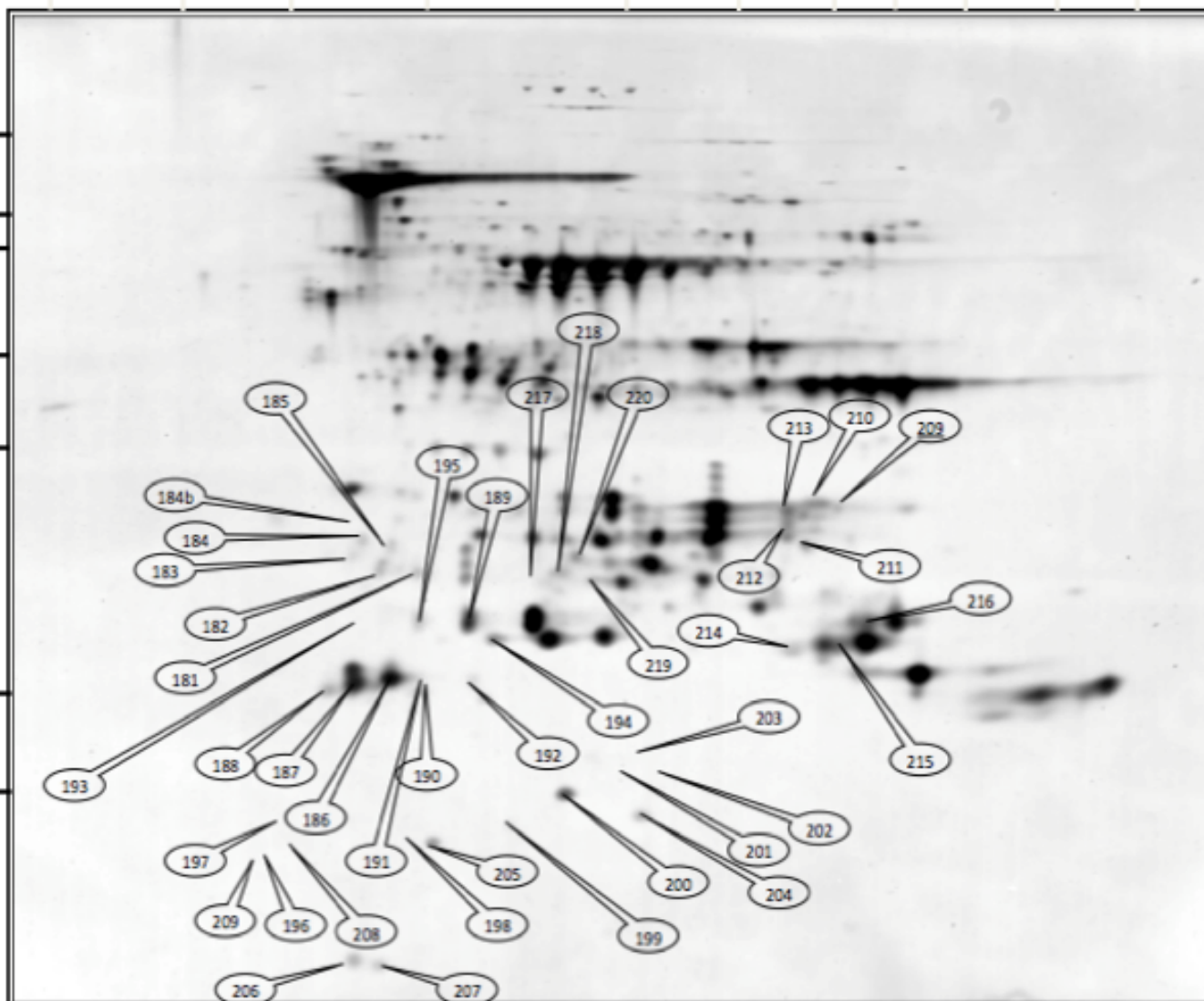
67,2

56,9

42,8

34,7

27,2



## **Annexe 2 : carte de référence de la fraction globulines de la graine d'*Amborella***

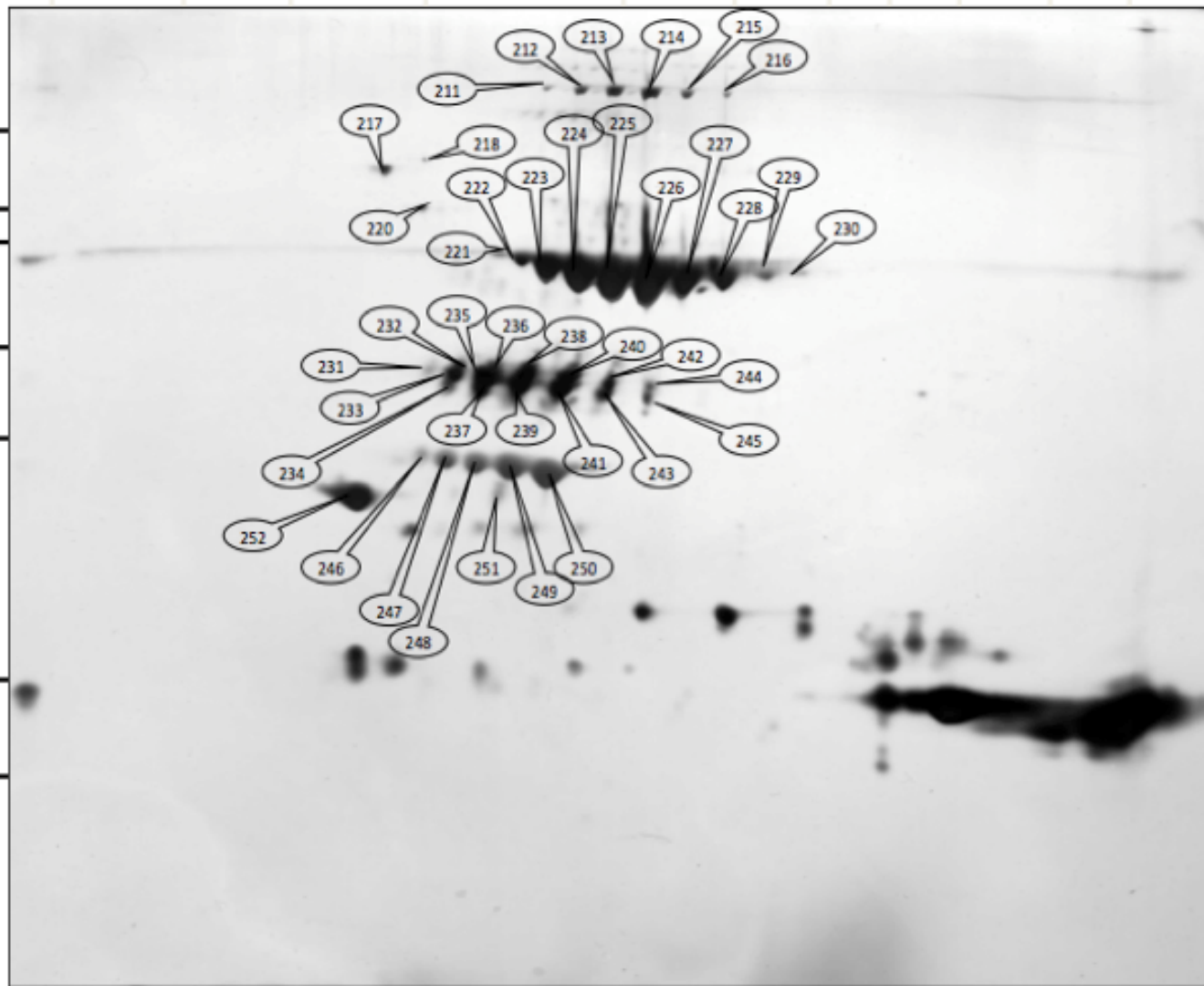
Localisation des protéines globulines sur gel d'électrophorèse bidimensionnelle de référence de la graine totale mature sèche d'*Amborella*. Une quantité de 100 µg de protéines globulines a été déposée. Les protéines ont été séparées en première dimension par isoélectrofocalisation suivant un gradient de pH immobilisé 3-10 non linéaire puis en seconde dimension par électrophorèse SDS-PAGE sur des gels de polyacrylamide 10%. Les protéines ont été révélées par coloration au nitrate d'argent.

# Spots 211 à 252

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3  
73,4  
67,2  
56,9  
42,8  
34,7  
27,2



# Spots 253 à 290

← pi  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3

73,4

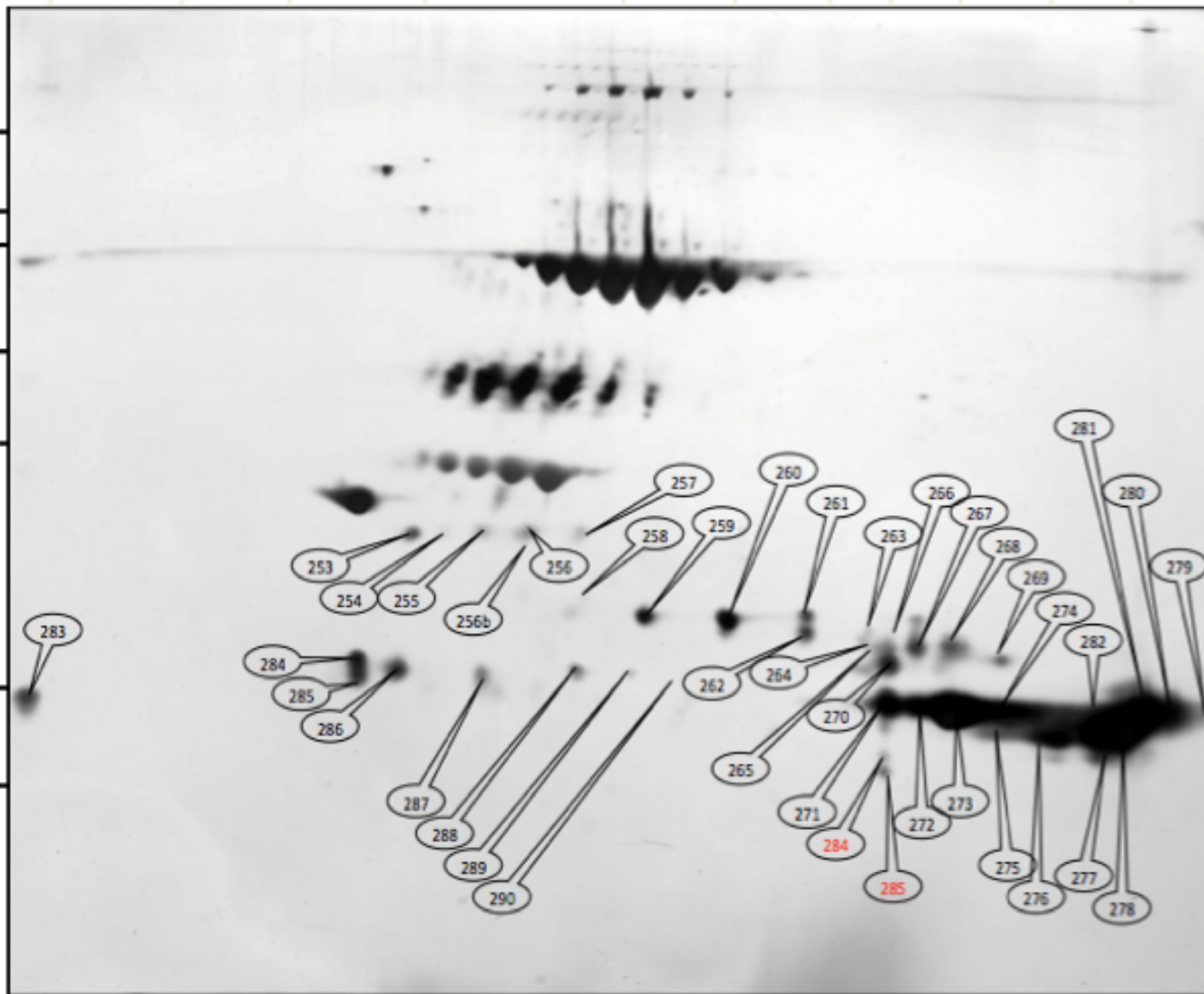
67,2

56,9

42,8

34,7

27,2

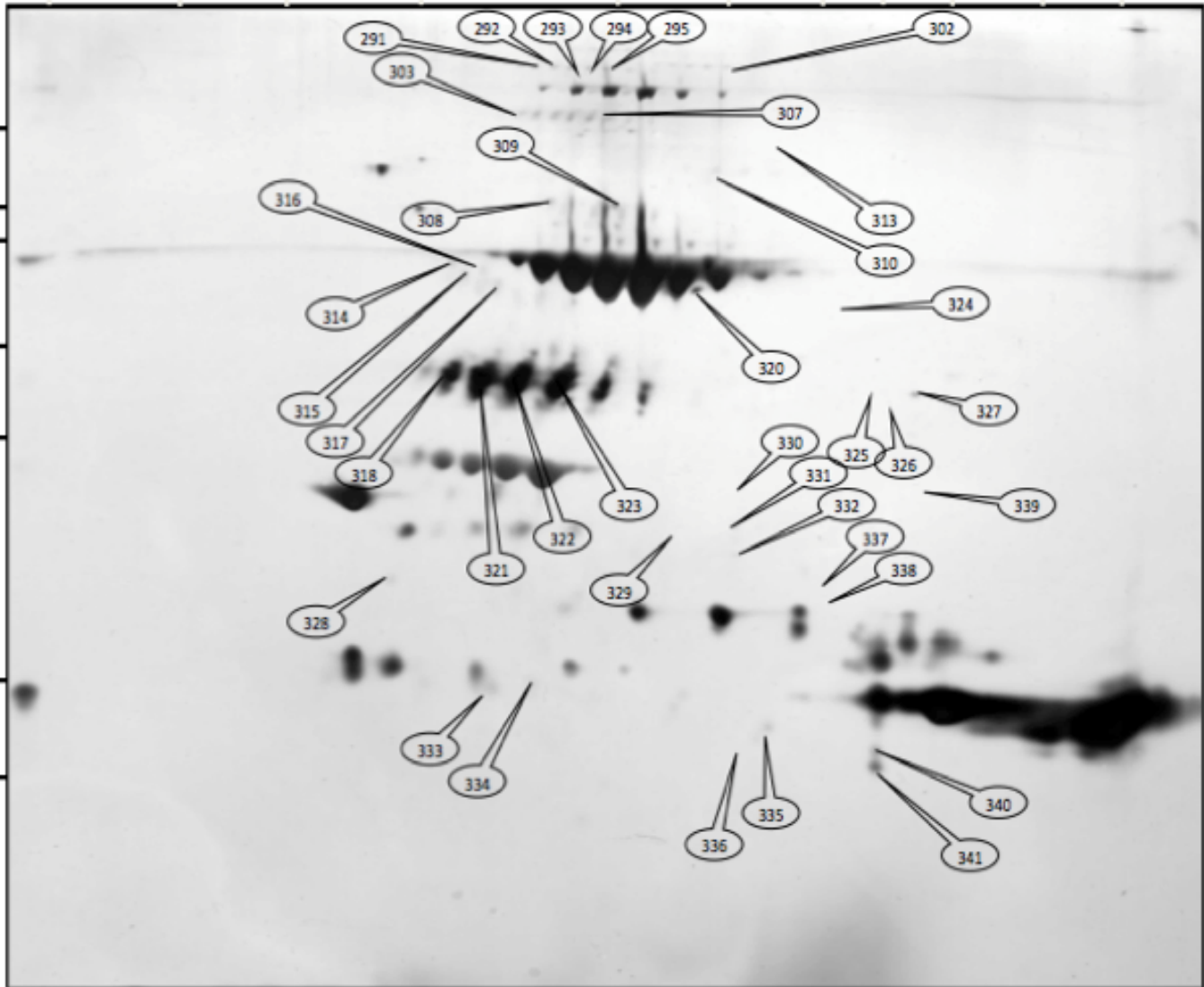


# Spots 291 à 341

pl  
MW (kDa)

3,0 4,8 5,0 5,5 5,8 5,9 6,2 6,4 6,6 7,3 8,7

94,3  
73,4  
67,2  
56,9  
42,8  
34,7  
27,2



# Annexe 3 : analyse phylogénétique des globulines 11S d'*Amborella*

Cette analyse fait partie de la publication « *The Amborella Genome and the Evolution of Flowering Plants* » soumise à la revue *Science*. La version présentée ici date du 25 juillet 2013.

## **The *Amborella* Genome and the Evolution of Flowering Plants**

### ***Amborella* Genome Project**‡\*

‡All authors with affiliations and contributions are listed at the end of the paper.

\*Address for general correspondence: [cwd3@psu.edu](mailto:cwd3@psu.edu). Contact information for working groups is provided in the authorship details.

### ***Amborella* Genome Project**

Authorship of this paper should be cited as “*Amborella* Genome Project.” Participants are arranged by working group, and then are listed in alphabetical order. Major contributions (†) and the author for correspondence (\*) are indicated within each working group. Joshua P. Der, Srikar Chamala, Andre S. Chanderbali, and James C. Estill made major and equal contributions to this project.

**Research leadership:** Victor A. Albert,<sup>1</sup> W. Bradley Barbazuk,<sup>2,3</sup> Claude W. dePamphilis<sup>4,5,6,\*</sup> ([cwd3@psu.edu](mailto:cwd3@psu.edu)), Joshua P. Der,<sup>4,6,†</sup> James Leebens-Mack,<sup>7</sup> Hong Ma,<sup>4,8</sup> Jeffrey D. Palmer,<sup>9</sup> Steve Rounsley,<sup>10,11</sup> David Sankoff,<sup>12</sup> Stephan C. Schuster,<sup>6,13,14</sup> Douglas E. Soltis,<sup>2,3,15</sup> Pamela S. Soltis,<sup>3,15</sup> Susan R. Wessler,<sup>16</sup> Rod A. Wing<sup>10,17</sup>

**Genome sequencing and assembly:** Victor A. Albert,<sup>1</sup> Jetty S.S. Ammiraju,<sup>10,17</sup> W. Bradley Barbazuk<sup>2,3,\*</sup> ([bbarbazuk@ufl.edu](mailto:bbarbazuk@ufl.edu)), Srikar Chamala,<sup>2,†</sup> Andre S. Chanderbali,<sup>2</sup> Claude W. dePamphilis,<sup>4,5,6</sup> Joshua P. Der,<sup>4,6</sup> Ronald Determann,<sup>18</sup> James Leebens-Mack,<sup>7</sup> Hong Ma,<sup>4,8</sup> Paula Ralph,<sup>4</sup> Steve Rounsley,<sup>10,11</sup> Stephan C. Schuster,<sup>6,13,14</sup> Douglas E. Soltis,<sup>2,3,15</sup> Pamela S. Soltis,<sup>3,15</sup> Jason Talag,<sup>10,17</sup> Lynn Tomsho,<sup>13</sup> Brandon Walts,<sup>2</sup> Stefan Wanke,<sup>19</sup> Rod A. Wing<sup>10,17</sup>

**Cytogenetics:** Victor A. Albert,<sup>1</sup> W. Bradley Barbazuk,<sup>2,3</sup> Srikar Chamala,<sup>2</sup> Andre S. Chanderbali,<sup>2,†</sup> Tien-Hao Chang,<sup>1</sup> Ronald Determann,<sup>18</sup> Tianying Lan,<sup>1</sup> Douglas E. Soltis<sup>2,3,15,\*</sup> ([dsoltis@ufl.edu](mailto:dsoltis@ufl.edu)), Pamela S. Soltis<sup>3,15</sup>

**Genome annotation and database development:** Siwaret Arikrit,<sup>20</sup> Michael Axtell,<sup>4,5</sup> Saravananaraj Ayyampalayam,<sup>7</sup> W. Bradley Barbazuk,<sup>2,3</sup> James M. Burnette, III,<sup>16</sup> Srikar Chamala,<sup>2</sup> Emanuele De Paoli,<sup>21</sup> Claude W. dePamphilis,<sup>4,5,6</sup> Joshua P. Der,<sup>4,6</sup> James C. Estill,<sup>7,†</sup> Nina P. Farrell,<sup>4</sup> Alex Harkess,<sup>7</sup> Yuannian Jiao,<sup>4,22</sup> James Leebens-Mack<sup>7,\*</sup> ([jleebensmack@plantbio.uga.edu](mailto:jleebensmack@plantbio.uga.edu)), Kun Liu,<sup>16</sup> Wenbin Mei,<sup>2</sup> Blake C. Meyers,<sup>20</sup> Saima Shahid,<sup>5</sup> Eric Wafula,<sup>4</sup> Brandon Walts,<sup>2</sup> Susan R. Wessler,<sup>16</sup> Jixian Zhai,<sup>20</sup> Xiaoyu Zhang<sup>7</sup>

**Synteny analysis:** Victor A. Albert<sup>1,\*</sup> ([vaalbert@buffalo.edu](mailto:vaalbert@buffalo.edu)), Lorenzo Carretero-Paulet,<sup>1</sup> Claude W. dePamphilis,<sup>4,5,6</sup> Joshua P. Der,<sup>4,6</sup> Yuannian Jiao,<sup>4,22</sup> James Leebens-Mack,<sup>7</sup> Eric Lyons,<sup>10,23</sup> David Sankoff,<sup>12,†</sup> Haibao Tang,<sup>24</sup> Eric Wafula,<sup>4</sup> Chunfang Zheng<sup>12</sup>

**Global gene family analysis:** Victor A. Albert,<sup>1</sup> Naomi S. Altman,<sup>25</sup> W. Bradley Barbazuk,<sup>2,3</sup> Claude W. dePamphilis<sup>4,5,6,\*</sup> ([cwg3@psu.edu](mailto:cwd3@psu.edu)), Joshua P. Der,<sup>4,6,†</sup> James C. Estill,<sup>7</sup> Yuannian Jiao,<sup>4,22,†</sup> James Leebens-Mack,<sup>7</sup> Kun Liu,<sup>16</sup> Wenbin Mei,<sup>2</sup> Eric Wafula<sup>4</sup>

**Targeted gene family curation and analysis:** Naomi S. Altman,<sup>25</sup> Siwaret Arikrit,<sup>20</sup> Michael Axtell,<sup>4,5</sup> Srikar Chamala,<sup>2</sup> Andre S. Chanderbali,<sup>2</sup> Feng Chen,<sup>26</sup> Jian-Qun Chen,<sup>27</sup> Vincent Chiang,<sup>28</sup> Emanuele De Paoli,<sup>21</sup> Claude W. dePamphilis,<sup>4,5,6</sup> Joshua P. Der<sup>4,6,\*</sup> ([jpd18@psu.edu](mailto:jpd18@psu.edu)), Ronald Determann,<sup>18</sup> Bruno Fogliani,<sup>29,30</sup> Chunce Guo,<sup>31</sup> Jesper Harholt,<sup>32</sup> Alex Harkess,<sup>7</sup> Claudette Job,<sup>33</sup> Dominique Job,<sup>33</sup> Sangtae Kim,<sup>34</sup> Hongzhi Kong,<sup>31</sup> James Leebens-Mack,<sup>7</sup> Guanglin Li,<sup>26</sup> Lin Li,<sup>31</sup> Jie Liu,<sup>28</sup> Hong Ma,<sup>4,8</sup> Blake C. Meyers,<sup>20</sup> Jongsun Park,<sup>34</sup> Xinshuai Qi,<sup>2,35</sup> Loïc Rajjou,<sup>36</sup> Valérie Sarramegna,<sup>29</sup> Ron Sederoff,<sup>28</sup> Saima Shahid,<sup>5</sup> Douglas E. Soltis,<sup>2,3,15</sup> Pamela S. Soltis,<sup>3,15</sup> Ying-Hsuan Sun,<sup>37</sup> Peter Ulvskov,<sup>32</sup> Matthieu Villegente,<sup>29</sup> Jia- Yu Xue,<sup>27</sup> Ting-Feng Yeh,<sup>38</sup> Xianxian Yu,<sup>31</sup> Jixian Zhai<sup>20</sup>

**Population genomics:** Juan J. Acosta,<sup>39</sup> Victor A. Albert,<sup>1</sup> W. Bradley Barbazuk,<sup>2,3</sup> Riva A. Bruenn,<sup>4,40</sup> Srikar Chamala,<sup>2</sup> Alexandre de Kochko,<sup>41</sup> Claude W. dePamphilis,<sup>4,5,6</sup> Joshua P. Der,<sup>4,6</sup> Luis R. Herrera-Estrella,<sup>42</sup> Enrique Ibarra-Laclette,<sup>42</sup> Matias Kirst,<sup>39,3</sup> James Leebens-Mack,<sup>7</sup> Solon P. Pissis,<sup>15,43</sup> Valérie Poncet,<sup>41</sup> Stephan C. Schuster,<sup>6,13,14</sup> Douglas E. Soltis,<sup>2,3,15</sup> Pamela S. Soltis<sup>3,15,\*</sup> ([psoltis@flmnh.ufl.edu](mailto:psoltis@flmnh.ufl.edu)), Lynn Tomsho<sup>13</sup>

#### **Affiliations:**

1Department of Biological Sciences, University at Buffalo, Buffalo, NY 14260, USA

2Department of Biology, University of Florida, Gainesville, FL 32611, USA

3University of Florida Genetics Institute, University of Florida, Gainesville, FL 32610, USA

4Department of Biology and Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA 16802, USA

5Intercollege Plant Biology Graduate Program, The Pennsylvania State University, University Park, PA 16802, USA

6Center for Comparative Genomics and Bioinformatics, The Pennsylvania State University, University Park, PA 16802 USA

7Department of Plant Biology, University of Georgia, Athens, GA 30602, USA

8State Key Laboratory of Genetic Engineering and Institute of Genetics, Institute of Plant Biology, Center for Evolutionary Biology, Institutes of Biomedical Sciences, School of Life Sciences Fudan University, Shanghai 200433, CHINA

9Department of Biology, Indiana University, Bloomington, IN 47405, USA

10School of Plant Science & BIO5 Institute for Collaborative Research, University of Arizona, Tucson, AZ 85721, USA

11Dow Agrosiences, Indianapolis, IN 46268, USA

12Department of Mathematics and Statistics, University of Ottawa, Ottawa, ONT, K1N 6N5, CANADA

13Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA

14Singapore Centre on Environmental Life Sciences Engineering, SINGAPORE

15Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

16Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA 92521, USA

17Arizona Genomics Institute, University of Arizona, Tucson, AZ 85721, USA

18Atlanta Botanic Garden, Atlanta, GA 30309, USA

19Technische Universität Dresden, Institut für Botanik, 01062 Dresden, GERMANY

20Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, USA



- 21Dipartimento di Scienze Agrarie ed Ambientali, Università degli Studi di Udine, via delle Scienze 206, 33100 Udine, ITALY
- 22Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602, USA
- 23iPlant Collaborative, University of Arizona, Tucson, AZ 85721, USA
- 24J. Craig Venter Institute, Rockville, MD 20850, USA
- 25Department of Statistics, The Pennsylvania State University, University Park, PA 16802, USA
- 26Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA
- 27School of Life Sciences, Nanjing University, Nanjing 210093, CHINA
- 28Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA
- 29Laboratoire Insulaire du Vivant et de l'Environnement, University of New Caledonia, Noumea, NEW CALEDONIA
- 30Institut Agronomique néo-Calédonien (IAC), Diversités biologique et fonctionnelle des écosystèmes terrestres, BP 73, 98890 Païta, NEW CALEDONIA
- 31State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, The Chinese Academy of Sciences, Beijing 100093, CHINA
- 32Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, DENMARK
- 33CNRS-Université Claude Bernard Lyon-Institut National des Sciences Appliquées-Bayer CropScience Joint Laboratory, Bayer CropScience, F-69263 Lyon cedex 9, FRANCE
- 34School of Biological Sciences and Chemistry, and Basic Science Research Institute, Sungshin Women's University, Seoul 142-732, REPUBLIC OF KOREA
- 35Key Laboratory of Conservation Biology for Endangered Wildlife of the Ministry of Education, and Laboratory of Systematic & Evolutionary Botany and Biodiversity, College of Life Sciences, Zhejiang University, Hangzhou 310058, CHINA
- 36INRA-AgroParisTech, Jean-Pierre Bourgin Institute (IJPB, UMR1318), Laboratory of Excellence "Saclay Plant Sciences" (LabEx SPS), F-78026 Versailles, FRANCE
- 37Department of Forestry, National Chung Hsing University, Taichung 40227, TAIWAN
- 38School of Forestry and Resource Conservation, National Taiwan University, Taipei 10617, TAIWAN
- 39School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611, USA
- 40Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA 94720, USA
- 41Institut de Recherche pour le Développement (IRD), Unité Mixte de Recherche–Diversité Adaptation et Développement des plantes (UMR DIADE), BP 64501, 34394 Montpellier Cedex 5, FRANCE
- 42Laboratorio Nacional de Genómica para la Biodiversidad, Irapuato, MEXICO
- 43Scientific Computing Group, Heidelberg Institute for Theoretical Studies, Heidelberg 69118, GERMANY

### **Participation au corps de texte (article principal) :**

*Seed storage globulins* – Seed storage proteins, including globulins, are critical for embryo and early seedling development in seed plants. These proteins are embedded in the very diverse cupin super-family, which is distributed across the tree of life (46). The 11S legumin-type globulins are particularly widespread across seed plant phylogeny (e.g. 47-49). Three distinct 11S legumin-type globulins have been identified in proteomic analyses of the globulin fraction in *Amborella* seeds

(Figs. S8-1 and S8-2). Comparisons of the *Amborella* globulin-coding gene sequences to other seed plants revealed that key cysteine residues contributing to disulfide bonding between subunits and the absence of Intron IV found in gymnosperms are conserved characteristics of angiosperm legumins (Fig. S8-4). In contrast, a conserved 52-residue region present in soybean and thought to be important for mature hexamer formation (50) was apparently derived after the divergence of *Amborella* from other angiosperms (Fig. S8-4). The present proteomic data provide support for validating the *A. trichopoda* genome annotation (Fig. S8-3). Also, they allow determining the timing of key changes in the genes that distinguish previously characterized angiosperm and gymnosperm members of the 11S-globulin gene family. It appears that structural features of *Amborella* 11S globulins are both reminiscent of modern angiosperms, yet still exhibit specific features of corresponding seed storage proteins in gymnosperms, thus highlighting an ancestral character of the *Amborella* 11S globulins (Figs. S8-4 and S8-5).

### **Participation en données supplémentaire :**

#### **8: Seed Proteomics and 11S Globulin Proteins: Understanding Seed Proteins in the Ancestral Angiosperms While Validating the *Amborella trichopoda* Genome Annotation**

Seed germination processes are diverse but can be characterized in part by studying seed storage proteins, because the remobilization of these proteins during germination and initial seedling growth is crucial for the establishment of vigorous seedlings (Rajjou et al. 2012). We used a proteomic approach to study the seed storage proteins of *Amborella*, with the goal of identifying proteome changes that could be associated with the origin and early diversification of angiosperms. We focused most of our attention on globulins. Whereas 7S vicilin type globulin genes sampled from monocots and eudicots cannot be aligned (Shewry & Halford 2002), 11S legumin-type globulins have been characterized and compared across seed plants in evolutionary analyses of these important protein-coding genes (Häger et al. 1995 ; Adachi et al. 2003 ; Li et al. 2012). Here we refine understanding of this gene family with the characterization of three *Amborella* 11S legumin type globulin homologs. These new data allow us to determine the timing of key changes in the genes that distinguish previously characterized angiosperm and gymnosperm members of this gene family.

**Preparation of Protein Extracts** - Fruits of *A. trichopoda* (Fig. S8-1) were collected at “plateau de Dogny-Sarraméa” (New Caledonia ; 21°37'0" N, 165°52'59" E ; [http://fr.getamap.net/cartes/new\\_caledonia\\_\[france\]/new\\_caledonia\\_\(general\)/dogny\\_plateaud\\_e/](http://fr.getamap.net/cartes/new_caledonia_[france]/new_caledonia_(general)/dogny_plateaud_e/)).



**Figure S8-1.** Left, *Amborella trichopoda* fruits on a tree. Right, transverse cut of the fruit without the fleshy exocarp, showing the localization of the endosperm (endo) and embryo (emb) in the seed within the sclerotized endocarp/mesocarp.

Protein extracts were prepared from 25 mature *Amborella* seeds (approximately 350 mg) after removing the dried pericarp tissue. Following grinding of the seeds in liquid nitrogen with a mortar and pestle, the protein characterization protocols of Osborne (Osborne 1924) were applied. Albumins were first extracted by placing the powder in 1.5 ml of distilled water containing a cocktail of protease inhibitors (Roche Molecular Biochemicals). After incubation for 20 min at 4°C, the suspension was centrifuged (15 min, 4°C, 40,000 g), and the supernatant containing albumins was recovered. Following three washes, the pellet was resuspended in 1.5 ml 0.8 M NaCl containing the cocktail of protease inhibitors for 20 min at 4°C. After centrifugation (4°C, 40,000 g, 15 min), the supernatant containing the globulins was collected. Following precipitation of the proteins in trichloroacetic (TCA) solution (10% TCA and 0.07% 2-mercaptoethanol in acetone) overnight at -20°C, the pellet was suspended in 1.5 ml of the buffer used for the isoelectrofocusing (IEF) step (see below) and containing 8 M urea, 2% [w/v] 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 0.5% [v/v] immobilized pH gradient (IPG) buffer 4-7, 20 mM dithiothreitol (DTT), and 0.01% [w/v] bromophenol blue

(Catusse, Strub, et al. 2008). Protein concentration in the various extracts was measured according to Bradford assay using bovine serum albumin as a standard.

**Two-Dimensional Gel Electrophoresis** - Isoelectric focusing (IEF) of protein extracts in IEF buffer was conducted using 18-cm (3-10 nonlinear) IPG strips on a Multiphor II system (Amersham-Pharmacia Biotech). Strips were rehydrated for 24 h at 22°C with the thiourea/urea lysis buffer containing 2% (v/v) Triton X-100, 20 mM DTT, and the protein extracts. IEF was performed at 22°C for 1 min at 300 V, 1 h 30 at 500 V and for 17 h at 3,500 V. IEF strips were equilibrated for 2 x 20 min in 2 x 100 ml of equilibration solution containing 6 M urea, 30% (v/v) glycerol, 2.5% (w/v) SDS, 0.15 M bis-Tris, and 0.1M HCl (9-10). DTT (50 mM) was added to the first equilibration solution, and iodoacetamide (4% [w/v]) was added to the second (Catusse, Strub, et al. 2008). Equilibrated gel strips were placed on top of vertical polyacrylamide gels (10% [v/v] acrylamide, 0.33% [w/v] piperazine diacrylamide, 0.18 M Trizma base, 0.166 M HCl, 0.07% [w/v] ammonium persulfate, and 0.035% [v/v] tetramethylethylenediamine [TEMED]). A denaturing solution (1% [w/v] low-melting agarose, 0.4% [w/v] SDS, 0.15 M bis-Tris, and 0.1 M HCl) was loaded on gel strips. After agarose solidification, electrophoresis was performed at 10°C in a buffer (pH 8.3) containing 25 mM Trizma Base, 200 mM taurine, and 0.1% (w/v) SDS, for 1 h at 35 V and 110 V overnight. Ten gels were run in parallel (Isodalt system from Amersham Pharmacia Biotech). All 2D gel assays were run at least in triplicate on three independent protein isolations.

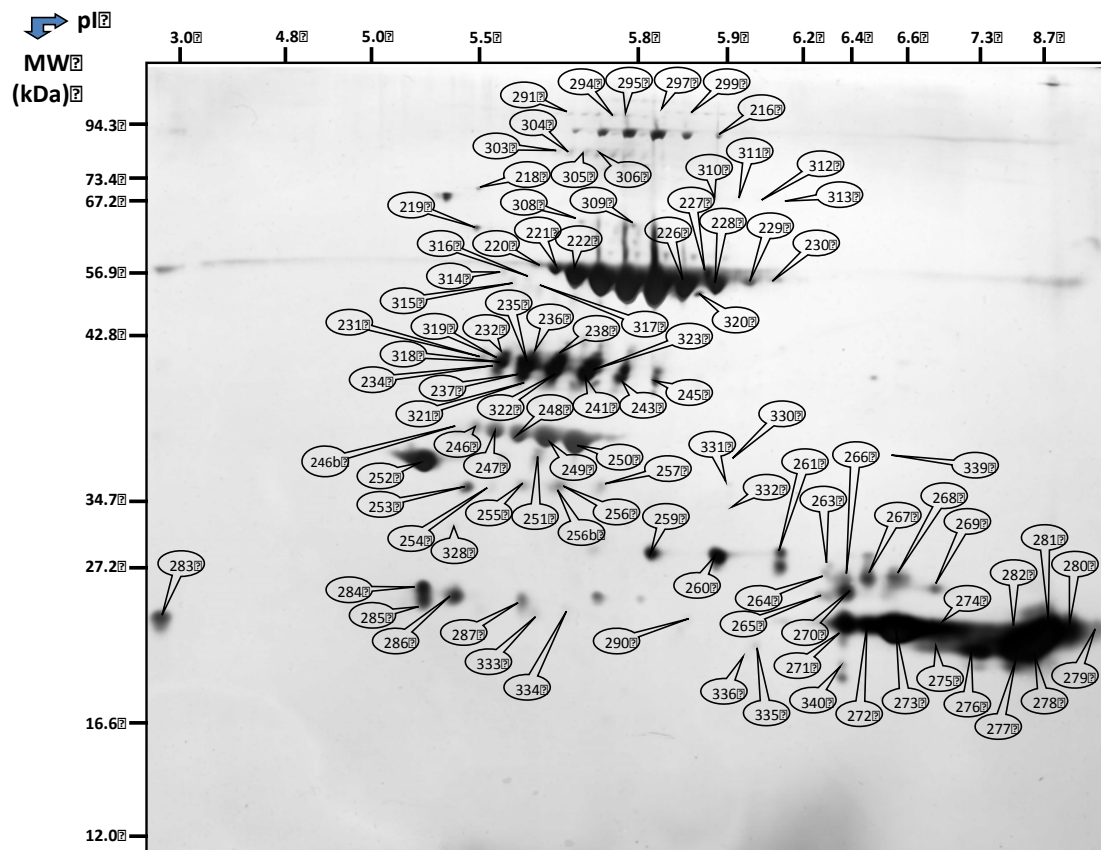
**Protein Staining** - Gels were stained with silver nitrate either as described by Blum *et al.* (Blum *et al.* 1987) for the analytical gels or by Shevchenko *et al.* (Shevchenko *et al.* 1996) for the preparative gels using the Hoefer Automated Gel stainer from Amersham Pharmacia Biotech. Silver-stained gels were scanned with the Sharp JX-330 scanner equipped with the Labscan version 3.00 from Amersham Pharmacia Biotech.

**Protein Identification Through Mass Spectrometry Analyses** - Protein spots were analyzed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) on the PAPPSO platform (B. Valot, T. Balliau, M. Zivy, INRA Moulon, France ; <http://pappso.inra.fr>). Based on the spectrum generated, proteins were identified using the X-Tandem software (Valot *et al.* 2011 ; Albertin *et al.* 2013). All peptides with an E-value match lower than 0.05 were parsed with our X-Tandem pipeline v. 3.1.2 (<http://pappso.inra.fr/bioinfo/xtandempipeline>). Proteins identified with at least two unique peptides and an E-value lower than 10E-8 were validated. Peptide sequences and protein identifications are listed in Table S8-1 (a separate file can be located at the Science

website: *Amborella* Supplementary Tables, Table S8-1. 11S globulins identified from 2D-electrophoresis (2DE) individual spots from *Amborella trichopoda* seed protein extracts).

**Phylogenetic Tree of 11S Globulin Proteins in Seed Plants** - Alignment was performed using MAFFT and a gene tree was estimated using maximum likelihood with available 11S globulin protein sequences (65 sequences) from 55 seed plant species (Table S8-2). Bootstrap analysis was performed with 100 pseudoreplicates.

**Reference Map and Peptides** - We found 210 albumin spots corresponding to 162 *Amborella* genes and 129 globulin spots corresponding to 24 *Amborella* genes, of which 99 globulin spots. Fig. S8-2 depicts the reference map for *Amborella* 11S globulins and shows the localization of the 99 protein spots identified by LC-MS/MS analysis.



**Figure S8-2.** Reference map for *Amborella* 11S globulins. Localization of the 99 protein spots identified by MS (labeled from 216 to 340) on a reference 2DE gel of 11S globulins from mature *Amborella* seeds. Peptide sequences are listed in Table S8-1 (a separate file can be located at the Science website: *Amborella* Supplementary Tables, Table S8-1. 11S globulins identified from 2D-electrophoresis (2DE) individual spots from *Amborella* seed protein extracts).

Peptide sequences for these spots are listed in Table S8-1 (a separate file can be located at the Science website: *Amborella* Supplementary Tables, Table S8-1. 11S globulins identified from 2D-electrophoresis (2DE) individual spots from *Amborella* seed protein extracts). These peptide sequences allowed revealing three distinct 11S legumin-type globulins from the *Amborella* genome sequence:

- i) evm\_27.model.AmTr\_v1.0\_scaffold00067.7,
- ii) evm\_27.model.AmTr\_v1.0\_scaffold00067.11,
- iii) and evm\_27.model.AmTr\_v1.0\_scaffold00067.12.

**Coverage of 11S Globulin Protein Sequences Encoded by the Three *Amborella* Gene Models by Experimentally Determined Peptide Sequences** - Peptides were mapped to the three *Amborella* 11S globulin gene models as follows (Fig. S8-3).

**evm\_27.model.AmTr\_v1.0\_scaffold00067.7**  
 MAHTKSIILTFVLLSLLVCAFADRRQSQRRLSDAQQCRMNRISGIRPTRVIRSEGGITELWDEDRDEFQCAGVAATRNTL  
 NPNSLYLPSFSSAPQIVYIERGRGIIGLSAPGCSESFHSGESGSIQHRKIRGQGF RDQHQKVQRIERG DVIAI PPGITH  
 WCYNDDNSEELVAFTVTDVTS DYNQLDTKQROFFIAGGQPRGORKQEGEGERGQKGRQEGEYGEERERERETGPRANVIGVGERYCNAKIRQN  
 LPQIDTRFLAEALDIPIELAQKI QREDE RGI I I KVEKEGLRILSPEGEEREERERERETGPRANVIGVGERYCNAKIRQN  
 IESLREADIYSRHGGHLKTINRRNLPILDILDM SAAKVTLYSDAI LAPHWSINAHTIAYITRGEQQIQIIGTNGQKVM D  
 DRVRQGDVIVPQFFTS MCKAGSQGIEWIAIKTSDLP MNSPLVGYTSAIKGMP IEVLTNAYRISNQQAQDIKYNREDQI  
 MIFPSSSRSSASS

**evm\_27.model.AmTr\_v1.0\_scaffold00067.12**  
 MAFTKSLVSFVLFSLFVSAFSLSTETRSQESLREARQCRIDRIPTSRPARRIQSEGGHTTEIWDEYEDQFLCAGVAAIRN  
 TIHPNSLSL PNFEPAPRLVYIQKGRGIISIVFP GCTESFQSQGYIRTRVGGEGQQQQGIKDRHQKVQTIKQGD AIAI PA  
 GAAHWCYNDGSEDLIAFSITDVTNDANQIEQSLKSFYLAGGQPRRGQEEGQQEQEQEQEQERQNSVNVINELNAEQ LAE  
 AFAPVPIELIRSLQKPD ERGWIVRVEKESLGVIRPDEEEEEERERYSERSNGYEERE CNARIRQNI DNPRRADIYSRQAGH  
 IQIVNRQTLPILSILDMSIEKGHLHPNLYAPHWTINAHTIVLITR GEGNIQVIGTNGRKMVDDRVVHEGDV FVVPQYFT  
 AMSKAGNEGLEWVAIKTSDLP MKSPILGHASAIKGIPIEVLKNAYKITTQEAR DIKLN RKDQYMLLPPRSTSSRP

**evm\_27.model.AmTr\_v1.0\_scaffold00067.11**  
 MALISKSLLVSLFLSLLVCAFVS YRGKLRSLDDARQCRMSKISTSQASRIIRSEGGITELWDEYEDQFQCAGVAAMRNT  
 IQPNLSL PNFSPAPRVVYIEQGS GILGLSAPGCSESF RSGERRSFQGGQTRERRLRDQHQQVQRI RSGDIIGIPPGV  
 THWCYNDDNSQELIAFSVIDLTNKNQLD PKLRIDTRFLAESFGIPDELAQRIQREDE RGLIVRVEKQGM RVLIP EEEEE  
 REEEEREREGASVNGLEELYCNVKIRQNI DTLRDI IYSRQGGRLKTINRKNLP ILEILDMSAQKVGLYSNAIFAPHWS  
 TNAHTIAYITRGEQQIQIIGTNGQKVMDDRVVRQGDVIVVPQFFSSMCKAGSQGIEWIAIKTSDLP IHTPLVGYTSAIKG  
 MPIEVL TNAYRISNQQAQDIKYNREEQLMIFPSSSRSDSS

**Figure S8-3.** Inferred protein sequences for the three *Amborella* gene models with 11S globulin peptides mapping them. Highlighted regions show peptide matches.

The peptide mapping coverage likely corresponds to the relative concentration of these three seed proteins. For the gene models evm\_27.model.AmTr\_v1.0\_scaffold00067.7 and evm\_27.model.AmTr\_v1.0\_scaffold00067.12, the protein coverage is extensive, 84% and 73%, respectively (Fig. S8-3). We note that experimentally determined peptide sequences frequently overlap (Table S8-1 ; a separate file can be located at the Science website: *Amborella* Supplementary Tables, Table S8-1. 11S globulins identified from 2D-electrophoresis (2DE) individual spots from *Amborella* seed protein extracts), thereby further increasing confidence in

validation of these two gene models. Furthermore, this high coverage validates the quality of genome sequencing data and of the exon/intron predictions.

The third gene model, *evm\_27.model.AmTr\_v1.0\_scaffold00067.11*, exhibited much lower coverage in the peptide data (4.1%, Fig. S8-3), presumably indicating that this 11S globulin accumulates in much lower concentration than the other two.

**Protein and Gene Sequence Alignments of 11S Globulin Family from *Amborella*, *Glycine max*, *Ginkgo biloba*, and *Arabidopsis thaliana*** - A multiple sequence alignment of the *Amborella* 11S globulin homologs to canonical seed plant 11S globulin genes revealed that two Cys residues involved in the disulfide bond linking the A- and B-subunits of 11S globulins are conserved in *Amborella* 11S globulins (red boxes in Fig. S8-4A).

Importantly, the conserved vacuolar processing enzyme/asparaginyl endopeptidase (VPE/Asn-endopeptidase) cut site is conserved in *Amborella* 11S globulins (Fig. S8-4A). Cutting at this site yields A- and B-subunits (red arrow in Fig. S8-4A, with the blue and green arrows designating the A and the B subunits, respectively). This limited proteolysis is regularly directed to an Asn-Gly (N-G) junction, resulting in an N-terminal glycyl residue of the B-subunit formed. This proteolytic processing triggers protein conformational changes that enable dense packaging and long-term stable storage of reserves within the protein storage vacuoles (PSV). After transport to the PSV, proteolytic processing at the conserved N-G peptide bond by the VPE/Asn-endopeptidase converts the pro-form of 11S globulins to the two disulfide-linked mature polypeptides A- and B-subunits, which is accompanied by further assembly of the trimer precursor-protein complexes into mature hexamers that sediment as a 11S complex (Scott et al. 1992 ; Gruis et al. 2002).

By conducting BLAST analyses we found that as in *Arabidopsis* (Scott et al. 1992 ; Gruis et al. 2002), the *Amborella* VPE family is composed of four genes (corresponding to the  $\alpha$ VPE,  $\beta$ VPE,  $\gamma$ VPE,  $\delta$ VPE subunits), namely: *evm\_27.model.AmTr\_v1.0\_scaffold00036.100*, *scaffold00002.262*, *scaffold00002.263*, and *scaffold00002.265*, thereby suggesting that at least the two *Amborella* 11S globulins encoded by the gene models *evm\_27.model.AmTr\_v1.0\_scaffold00067.11* and *evm\_27.model.AmTr\_v1.0\_scaffold00067.12*, which do contain the N-G cutting site (Fig. S8-4), are processed by an *Amborella* VPE system. In marked contrast, the *Amborella* 11S globulin encoded by the *evm\_27.model.AmTr\_v1.0\_scaffold00067.7* gene deviates conspicuously from the two above *Amborella* globulin precursors as the N-G junction is replaced by a N-V-I sequence. The alteration

of this N-G junction motif is found for 11S globulins from gymnosperms, as observed in *Gingko biloba* (sequence 4 in Fig. S8-4A) and *Metasequoia glyptostroboides* (Häger & Wind 1997), thus highlighting an ancestral character of this 67.7 11S *Amborella* globulin (Fig. S8-4A).

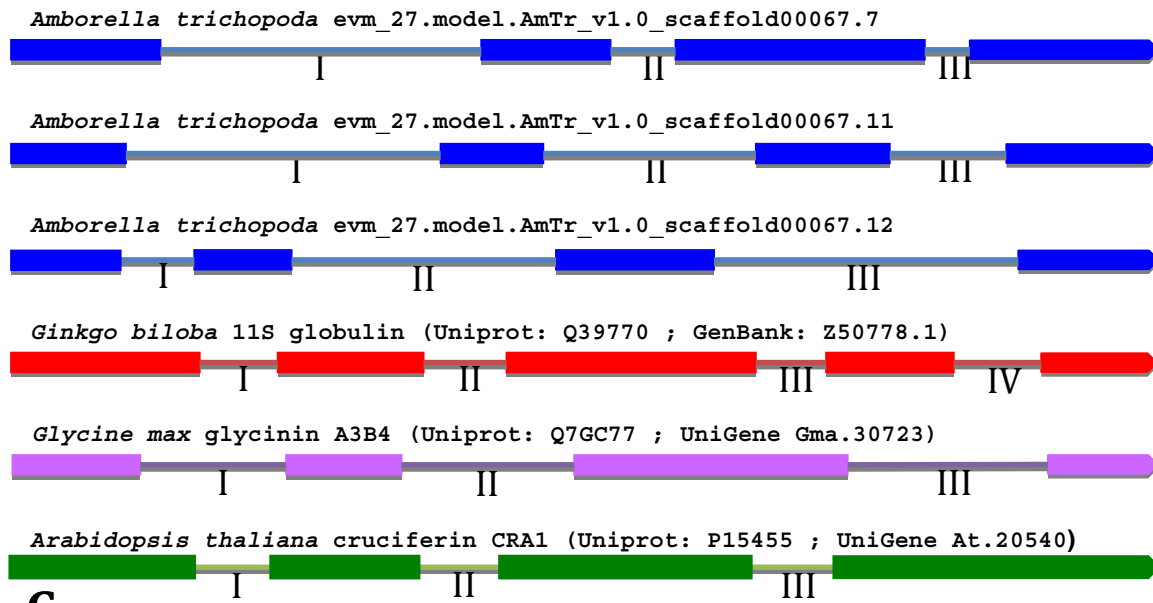
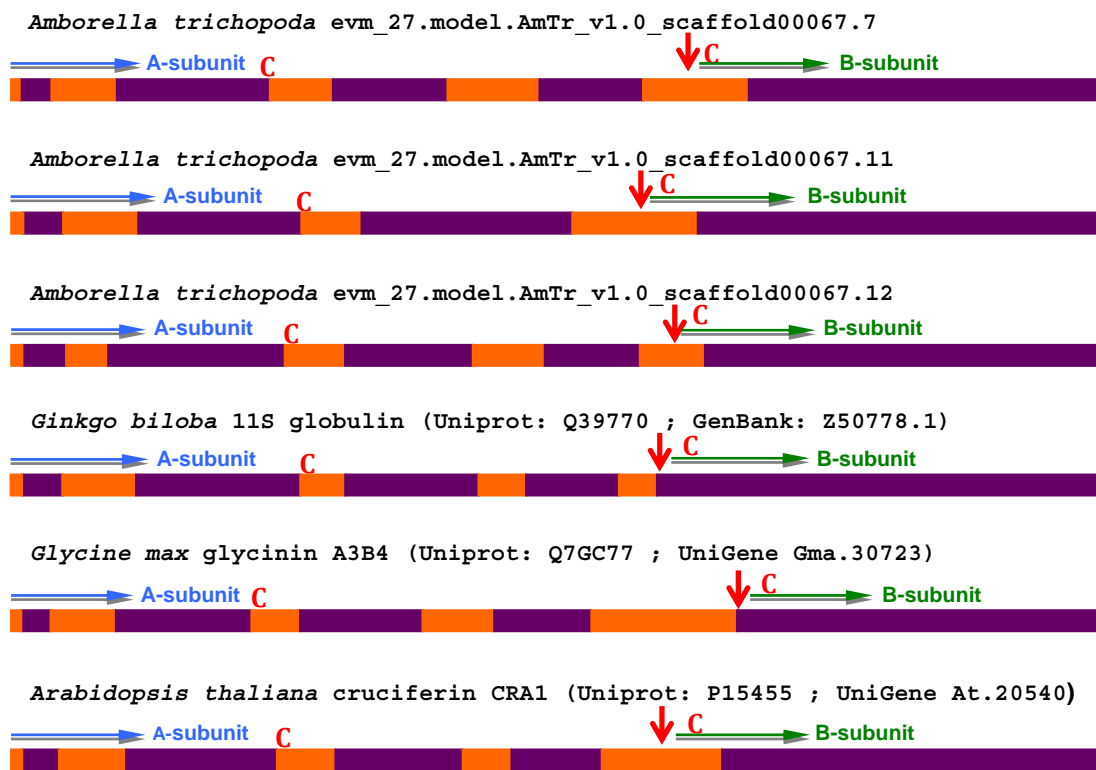
The sequence alignment also revealed additional conserved motifs, including the amino acid residues involved in trimer formation and stabilization of 11S globulins (Tandang-Silvas et al. 2010), and those involved in the correct globular folding of 11S globulins (Häger & Wind 1997).

Furthermore, this analysis revealed protein regions that are homologous to PX14G (box a) and SX3G (box b) motifs (Fig. S8-4A) adjacent to introns III and IV in 11S globulin genes, respectively (Häger et al. 1996). The B-subunits of gymnosperm globulins exhibit a highly conserved intron/exon structure differing from those of angiosperms (Häger et al. 1996). In particular, the 11S globulin genes from gymnosperms display an additional intron (referred to as intron IV) that uniformly interrupts the sequence coding for the C-terminal part of the basic-subunits of 11S globulins (Häger et al. 1996). In contrast to the *Gingko biloba* 11S globulin homolog, intron IV is absent in *Amborella* 11S-globulin-encoding genes and other angiosperm homologs (Fig. S8-4B), consistent with the finding that this intron was lost in the angiosperm stem group (Häger et al. 1996).

We investigated the distribution between ordered and disordered regions in 11S globulin globulin genes from *Amborella* (Fig. S8-4C). It appears that gene model 67.11 is very different from other 11S globulins as it lacks a disordered region in the A subunit (Fig. S8-4C). During seed germination the initial mobilization of 11S globulins preferentially begins with proteolysis of the A chains (Gallardo et al. 2001). It is therefore possible that lack of this disordered region somehow impedes processing of the globulin encoded by 67.11 during seed imbibition. The function of 67.11 may have diverged from the other 11S globulins given its divergent structure and low concentration relative to the other 11S globulin homologs.





**B****C****Figure****S8-4.** 11S globulin family of *Amborella trichopoda*.

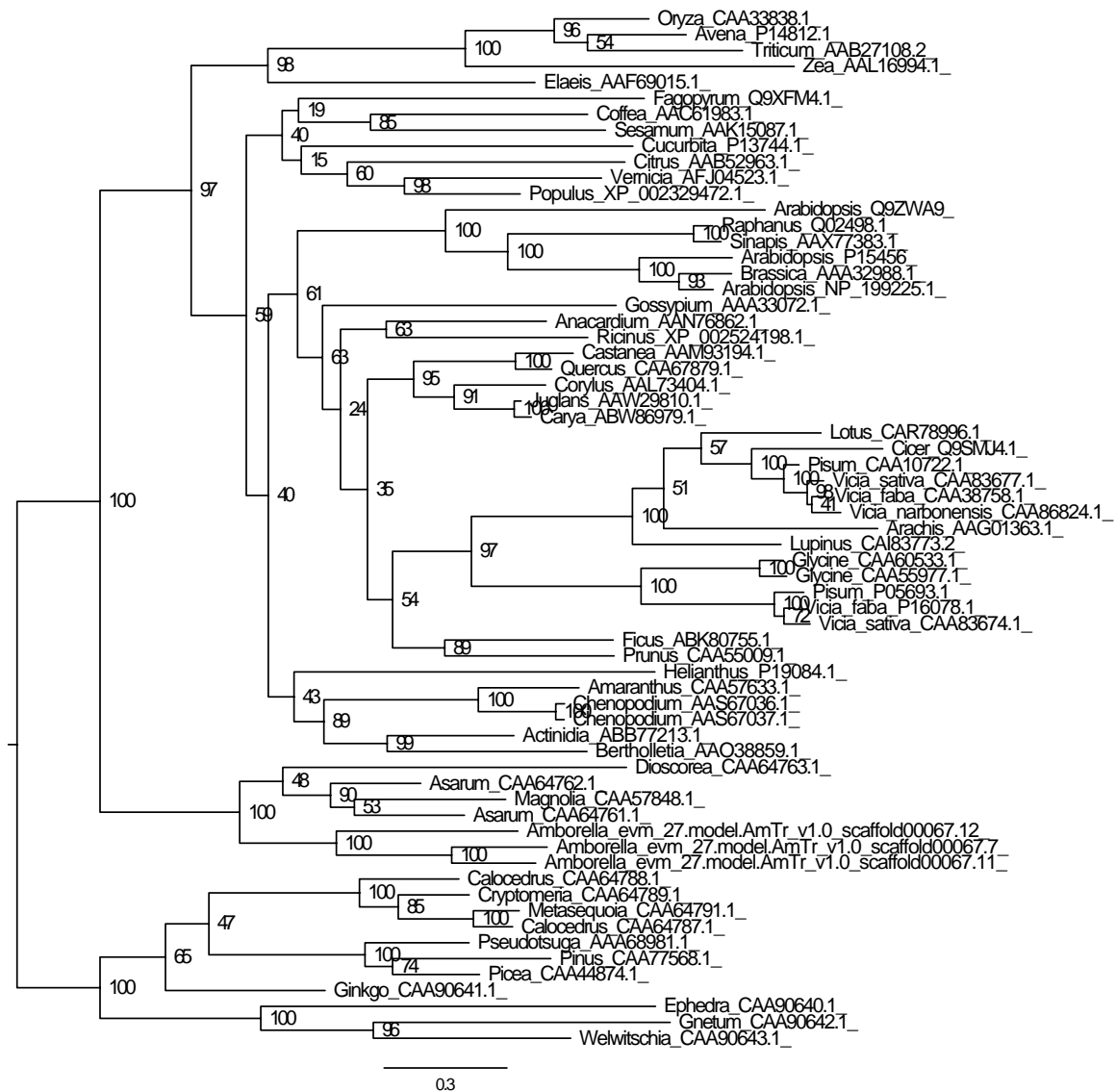
(A) Protein sequence alignment of 11S globulin family from *Amborella*, *Glycine max* (Uniprot: Q7GC77), *Ginkgo biloba* (Uniprot: Q39770) and *Arabidopsis thaliana* (Uniprot: P15455). Multiple sequence alignment was performed with ClustalW2 program ([http://www.ebi.ac.uk/Tools/services/web\\_clustalw2/toolform.ebi](http://www.ebi.ac.uk/Tools/services/web_clustalw2/toolform.ebi)). BLOSUM matrix was used to score the alignment. The Multiple Sequence alignment parameters were: Gap Open = 5 ; Gap

Extension = 0.05 ; Gap Distances = 1 ; No End Gaps = Yes. The red boxes show the two Cys involved in the conserved disulfide bond linking acidic (A)- and basic (B)- subunits of 11S globulins. The red arrow indicates the conserved cutting site by a vacuolar processing enzyme/asparaginyl endopeptidase (VPE/Asn-endopeptidase). The blue and green arrows point the A- and B-subunits respectively. The amino acids highlighted with gray shading indicate the conserved motif presumably involved in the recognition of the cutting site just behind the first Asn residue (N). The amino acids highlighted with yellow shading indicate the residues involved in trimer formation of the precursor forms and stabilization of 11S globulins. The amino acids highlighted with blue shading indicate the residues involved in the correct globular folding of 11S globulins. Black boxes a and b indicate regions that are homologous to PX14G (Box a) and SX3G (Box b) motifs adjacent to the position of introns III and IV respectively.

(B) Schematic structural description of 11S globulin gene models is displayed here indicating exons (boxes) and introns (lines) positions for *Amborella*, *Glycine max*, *Ginkgo biloba* and *Arabidopsis*. Introns are labeled with Roman numbers.

(C) 11S-Globulins predicted structural organization from *Amborella trichopoda*, *Glycine max*, *Ginkgo biloba* and *Arabidopsis thaliana*. Protein structure was predicted with Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2>). Orange boxes represent disordered regions. Purple boxes represent organized regions (including alpha helix and beta strand). The red arrow shows the conserved cleavage site by a vacuolar processing enzyme/asparaginyl endopeptidase (VPE/Asn-endopeptidase). The red C show the two Cys involved in the conserved disulfide bond linking the A- and B- subunits of 11S globulins. The blue and green arrows point the A- and B-subunits respectively.

**Phylogenetic Tree of 11S Globulins Based on Protein Sequences** - The phylogenetic tree constructed using 11S globulins from seed plants (Fig. S8-5) shows clades comparable to those described by Angiosperm Phylogeny Group III (Bremer et al. 2009), thus confirming the ability of 11S globulins to be used for phylogenetic studies (Li et al. 2012). The position of *Amborella* close to magnoliids (i.e. *Asarum europeum* (Piperales), *Magnolia salicifolia* (Magnoliales)) generally agrees with broad molecular phylogenetic analyses (Soltis et al. 2011), although the sister relationship of magnoliid and *Amborella* genes is not consistent with current understanding of species relationships (Fig. S8-1). As suggested by Li *et al.* (Li et al. 2012), most monocot 11S globulin homologs form a clade that is sister to a eudicot gene clade.



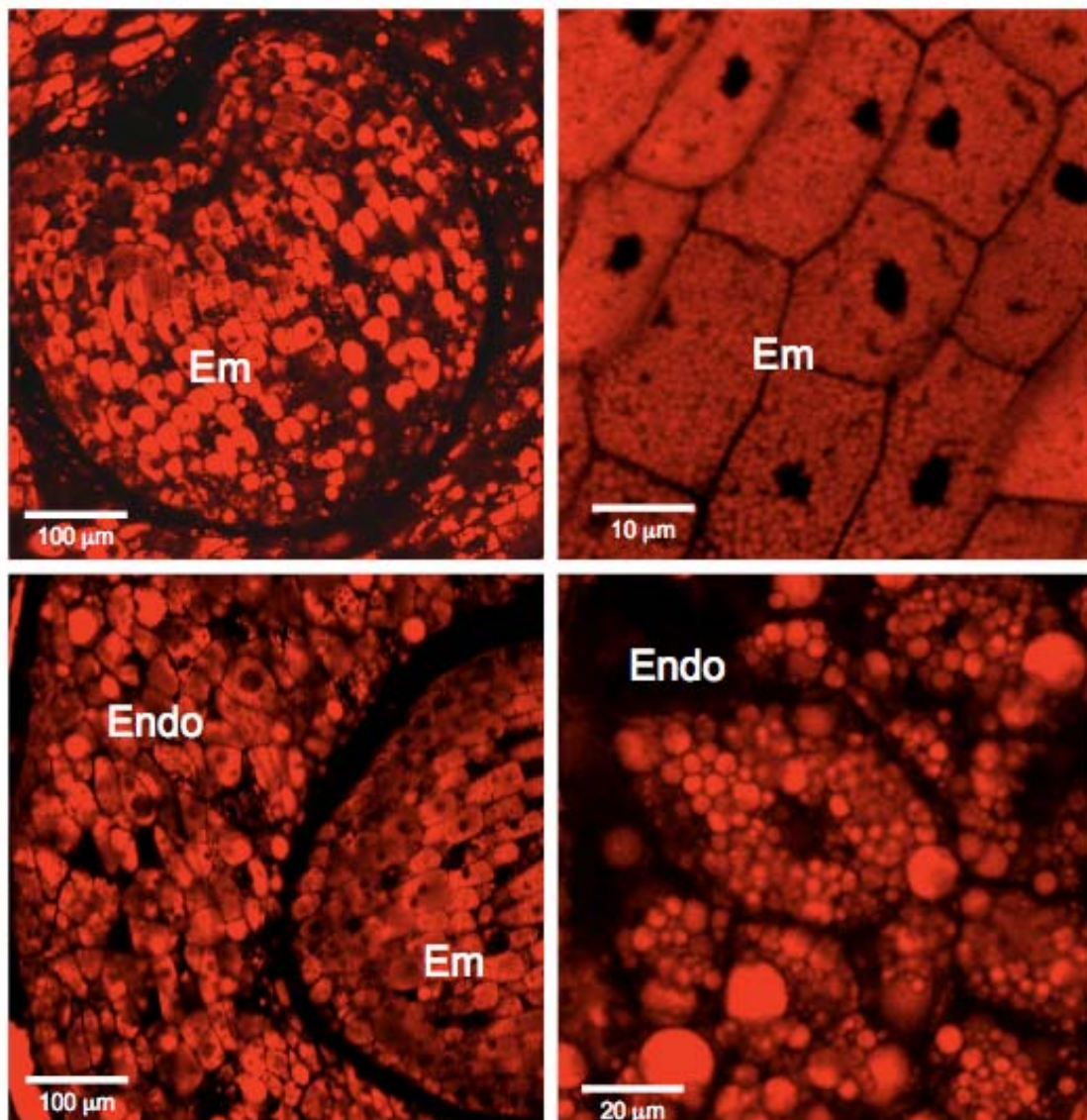
**Figure S8-5.** Phylogenetic tree of 11S globulin proteins in seed plants shows that angiosperm homologs form a well-supported clade. Estimated relationships among eudicots, monocots, magnoliids, and *Amborella* suggest clade-specific duplications within the *Amborella* and eudicot lineages. Seed plant 11S globulin genes included in the phylogenetic analysis are listed in Table S8-2.

**Table S8-2.** Seed plant 11S-globulin genes included in phylogenetic analysis of Fig. S8-5.

<b>Source Species</b>	<b>GenBank ID or <i>Amborella</i> Gene model ID</b>
<i>Actinidia chinensis</i>	ABB77213.1
<i>Amaranthus hypochondriacus</i>	CAA57633.1
<i>Amborella trichopoda</i>	evm_27.model.AmTr_v1.0_scaffold00067.11
<i>Amborella trichopoda</i>	evm_27.model.AmTr_v1.0_scaffold00067.12
<i>Amborella trichopoda</i>	evm_27.model.AmTr_v1.0_scaffold00067.7
<i>Anacardium occidentale</i>	AAN76862.1
<i>Arabidopsis thaliana</i>	NP_199225.1
<i>Arabidopsis thaliana</i>	P15456
<i>Arabidopsis thaliana</i>	Q9ZWA9
<i>Arachis hypogaea</i>	AAG01363.1
<i>Asarum europaeum</i>	CAA64761.1
<i>Asarum europaeum</i>	CAA64762.1
<i>Avena sativa</i>	P14812.1
<i>Bertholletia excelsa</i>	AAO38859.1
<i>Brassica napus</i>	AAA32988.1
<i>Calocedrus decurrens</i>	CAA64787.1
<i>Calocedrus decurrens</i>	CAA64788.1
<i>Carya illinoensis</i>	ABW86979.1
<i>Castanea crenata</i>	AAM93194.1
<i>Chenopodium quinoa</i>	AAS67036.1
<i>Chenopodium quinoa</i>	AAS67037.1
<i>Cicer arietinum</i>	Q9SMJ4.1
<i>Citrus sinensis</i>	AAB52963.1
<i>Coffea arabica</i>	AAC61983.1
<i>Corylus avellana</i>	AAL73404.1
<i>Cryptomeria japonica</i>	CAA64789.1
<i>Cucurbita maxima</i>	P13744.1
<i>Dioscorea caucasia</i>	CAA64763.1
<i>Elaeis guineensis</i>	AAF69015.1
<i>Ephedra gerardiana</i>	CAA90640.1
<i>Fagopyrum esculentum</i>	Q9XFM4.1
<i>Ficus pumila</i> var. <i>awkeotsang</i>	ABK80755.1
<i>Ginkgo biloba</i>	CAA90641.1
<i>Glycine max</i>	CAA55977.1
<i>Glycine max</i>	CAA60533.1
<i>Gnetum gnemon</i>	CAA90642.1
<i>Gossypium hirsutum</i>	AAA33072.1
<i>Helianthus annuus</i>	P19084.1
<i>Juglans regia</i>	AAW29810.1
<i>Lotus japonicus</i>	CAR78996.1
<i>Lupinus albus</i>	CAI83773.2
<i>Magnolia salicifolia</i>	CAA57848.1
<i>Metasequoia glyptostroboides</i>	CAA64791.1
<i>Oryza sativa</i>	CAA33838.1
<i>Picea glauca</i>	CAA44874.1
<i>Pinus strobus</i>	CAA77568.1
<i>Pisum sativum</i>	CAA10722.1

<i>Pisum sativum</i>	P05693.1
<i>Populus trichocarpa</i>	XP_002329472.1
<i>Prunus dulcis</i>	CAA55009.1
<i>Pseudotsuga menziesii</i>	AAA68981.1
<i>Quercus robur</i>	CAA67879.1
<i>Raphanus sativus</i>	Q02498.1
<i>Ricinus communis</i>	XP_002524198.1
<i>Sesamum indicum</i>	AAK15087.1
<i>Sinapis alba</i>	AAX77383.1
<i>Triticum aestivum</i>	AAB27108.2
<i>Vernicia fordii</i>	AFJ04523.1
<i>Vicia faba</i>	P16078.1
<i>Vicia faba var. minor</i>	CAA38758.1
<i>Vicia narbonensis</i>	CAA86824.1
<i>Vicia sativa</i>	CAA83674.1
<i>Vicia sativa</i>	CAA83677.1
<i>Welwitschia mirabilis</i>	CAA90643.1
<i>Zea mays</i>	AAL16994.1.

## Annexe 4 : observations cytologiques de la graine mature sèche d'*Amborella* après coloration au rouge Nil



Annexe 1 : observation au microscope confocal de coupes de graines d'*Amborella* colorées au rouge Nil montrant l'embryon (Em) et l'albumen (Endo).





## Annexe 5 : alignement des peptides expérimentaux pour la PIMT et MOTHER of FT and TFL1

### Protein-L-isoaspartyl O-methyltransferase (scaffold00166.14)

MVLVSKKNCFGFLIVLLVMRSAEDPALRAQFDSWMSCTQNLWTKGGMDRNKSLVEHLCLYGVIKS  
KRVAEVMETIDRGLFVVPDGNLAYADSPMPIGYNATISAPMHATCLELLLDEHLQPGMRALDVGSG  
TGYLTACFALMVGPQGRAVGVEHIPELVAMSIENIKSPAAPLLOEGSLQIHVGDGRQGWPEAP  
YDAIHVGAAASDIPAPLIEQLKPGGRMVIPVGNIFQDLKVVDKKADGSI SVRDETSVRYVPLTSR  
ESQLRDY

### Protein MOTHER of FT and TFL1 (scaffold00003.462)

MAALFPLPLKHEKSHPAITLPLSCFSGETILPNLTLAPRFISKETFLSPSMAR SVEPLAVGK VIG  
DVIDMFVPTVDVSVSYVSKQVANGCELKPSGTAERPFQLSARSLDDNLYTLVMTDPDAPSPSEP  
TMR EWLHWI VVDI PGGSDAMR GREIVSYMGPKPPTGIHRYVFSLFK QQGPLGSVAR APEARNNFS  
TRAF AAR LGLGLPVA AVYFNAQKEPASR KR

Alignement des peptides expérimentaux (en vert) sur les séquences correspondantes de la PIMT  
(Protein-L-isoaspartyl O-methyltransferase) et la Protein MOTHER of FT and TFL1



## **Annexe 6 : résultats de la caractérisation de la flore endophyte de la graine mature sèche de *Psychotria gabriellae* réalisée à Angers par Marie-Agnès Jacques (INRA Angers)**

Les graines de *P. gabriellae* ont été macérées pendant 12 h à 4°C sous agitation puis le macérât a été porté à ébullition pendant 10 min. L'amplification de l'ADNr 16S a ensuite été réalisée à l'aide d'amorces universelles puis séquencé.

### **Séquence de l'amplification de l'ADNr 16S :**

```
GCCCTGTAGTGGGGGATAACTAGTCGAAAGATTAGCTAATACCGCATACGACCTGAGGGTGAAAGTGGGGGACCGCAAGGCCT  
CATGCTATAGGAGCGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCCACGATCAGTAGCTGGTCTGAG  
AGGACGATCAGCCACACTGGGACTGAGACACGGCCACACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCA  
ACCCGTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTTGTAAAGCAC'TTTTGTCCGAAAGAAATCGCACTTACTA  
ATATTAGGTGTGGATGACGGTACCGGAAGAATAAGGACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTCCAAG  
CGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTGTGCAAGACCGATGTGAAATCCCCGGGCTTAACCTGGGAAT  
TGCATTGGTACTGCACGGCTAGAGTGTGTGAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAGATGTGGAGG  
AATACCGATGGCGAAGGCAGCCCCCTGGGATAACACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGGATTAGATACCC  
TGG
```

## Résultat de l'analyse par BLAST :

### Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

<u>Accession</u>	<u>Description</u>	<u>Max score</u>	<u>Total score</u>	<u>Query coverage</u>	<u>E value</u>	<u>Max ident</u>	<u>Links</u>
<a href="#">JF703496.1</a>	Uncultured Burkholderiales bacterium clone H3-88 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">HM780306.1</a>	Uncultured bacterium clone JFR0702_jaa37c12 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">HM780007.1</a>	Uncultured bacterium clone JFR0702_jaa39g02 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">HQ585154.1</a>	Uncultured bacterium clone G01YRMYR06SP38STGEOTHERMALENV 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">HQ585147.1</a>	Uncultured bacterium clone A08KNSYR05SP28STGEOTHERMALENV 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">HQ530524.1</a>	Uncultured Ralstonia sp. clone Atlantis_l_d 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">GQ487980.1</a>	Uncultured bacterium clone V20-13 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">FJ981669.1</a>	Uncultured bacterium clone CF6 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">GQ032544.1</a>	Uncultured bacterium clone nbw908c02c1 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">GQ032479.1</a>	Uncultured bacterium clone nbw908c09c1 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">GQ029348.1</a>	Uncultured bacterium clone nbw924f10c1 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">FN293248.1</a>	Ralstonia sp. A24MG105 partial 16S rRNA gene, strain A24MG105	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">AM936912.1</a>	Uncultured Burkholderia sp. partial 16S rRNA gene, clone CM14H4	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">EU263295.1</a>	Ralstonia sp. LR35 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">EU263294.1</a>	Ralstonia sp. LR33 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">AM990870.1</a>	Uncultured beta proteobacterium partial 16S rRNA gene, clone TH1-74	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">DQ676998.1</a>	Iron-reducing enrichment clone CI-A6 clone CI-A6 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">AY509958.1</a>	Ralstonia sp. 1F2 16S ribosomal RNA gene, partial sequence	<a href="#">1221</a>	1221	100%	0.0	99%	
<a href="#">HM811408.1</a>	Uncultured bacterium clone nby232e07c1 16S ribosomal RNA gene, partial sequence	<a href="#">1216</a>	1216	100%	0.0	99%	
<a href="#">GQ032568.1</a>	Uncultured bacterium clone nbw908e12c1 16S ribosomal RNA gene, partial sequence	<a href="#">1216</a>	1216	100%	0.0	99%	
<a href="#">GQ055776.1</a>	Uncultured bacterium clone nbw1119h07c1 16S ribosomal RNA gene, partial sequence	<a href="#">1216</a>	1216	100%	0.0	99%	
<a href="#">AM936909.1</a>	Uncultured Burkholderia sp. partial 16S rRNA gene, clone CM14G12	<a href="#">1216</a>	1216	100%	0.0	99%	
<a href="#">AM936971.1</a>	Uncultured Burkholderia sp. partial 16S rRNA gene, clone CM14C8	<a href="#">1216</a>	1216	100%	0.0	99%	

# Tableaux supplémentaires

Les tableaux supplémentaires sont disponibles sur demande en version informatique. Ils sont regroupés dans un fichier Excel composé de plusieurs onglets :

- Onglet 1 : liste des protéines identifiées à partir de l'extrait des protéines totales d'embryons isolés de graines matures d'*Amborella*
- Onglet 2 : liste des protéines identifiées à partir de l'extrait des protéines totales d'albumens isolé de graines matures d'*Amborella*
- Onglet 3 : liste des protéines albumines identifiées à partir de graines totales matures d'*Amborella*
- Onglet 4 : liste des protéines globulines identifiées à partir de graines totales matures d'*Amborella*
- Onglet 5 : composition biochimique de la graine mature d'*Amborella*
- Onglet 6 : liste des protéines albumines identifiées à partir de graines en imbibition/germination d'*Amborella*
- Onglet 7 : liste des protéines reconnues par l'anticorps anti-DING et identifiées chez *P. gabriellae*





## Résumé

La Nouvelle-Calédonie possède une flore riche, diverse et unique au monde. Sa forte endémicité (74,7%) résulte en partie de l'origine gondwanienne de la flore et de la forte pression de sélection exercée par les sols ultramafiques, riches en éléments traces métalliques, dont le nickel. Si cet élément fait la richesse du pays par son exploitation minière, cette dernière, ainsi que l'anthropisation du territoire engendrent une détérioration des écosystèmes. Dans le cadre de la dynamique mondiale de conservation, de protection et de restauration de la biodiversité, il convient de caractériser au mieux la flore néo-calédonienne foisonnante de plantes uniques. Les graines, innovation des végétaux supérieurs assurant leur dissémination, sont un des points clés du succès de la domination mondiale des plantes supérieures et un outil indispensable à la restauration écologique. Notre étude s'est attachée à caractériser par une approche biochimique la biologie des graines de deux plantes endémiques exceptionnelles, *Amborella trichopoda*, sœur de toutes les plantes à fleurs et *Psychotria gabriellae*, une des plantes contenant le plus de nickel dans ses feuilles au monde.

La caractérisation protéomique de la graine d'*Amborella trichopoda* a permis d'obtenir le premier protéome d'un embryon rudimentaire. L'étude de ce cortège protéique a apporté des éléments de réponse aux nombreuses questions que soulèvent les graines à dormance morphologique comme celles d'*A. trichopoda*. Notamment, nous avons montré que l'embryon rudimentaire a acquis un stade de maturité moléculaire (présence de protéines chaperons, de réserves lipidiques). L'étude phylogénique de ces protéines a permis de conforter la place basale d'*A. trichopoda*. La caractérisation de l'évolution du protéome au cours de la germination a quant à elle mise en évidence une utilisation massive des protéines de réserve avant la fin de la germination ce qui questionne la définition de la germination *sensu stricto* pour les espèces à embryon rudimentaire.

Une précédente étude du cortège protéique de *P. gabriellae* avait révélé la présence de protéines DING notamment impliquées dans l'homéostasie d'éléments minéraux via leur interaction avec des transporteurs de type ABC ou encore en séquestrant elles-mêmes le phosphore. Leur identification fut corrélée avec l'observation d'un gradient de nickel dans la graine visant à protéger l'embryon du caractère toxique de ce dernier. Les données recueillies au cours du présent travail ont permis de confirmer la présence de ces protéines dont l'origine eucaryotique fait cependant débat. Face à cette controverse, nous avons cherché à identifier la présence de bactéries chez la graine mature sèche. Quatre bactéries endophytes de graine ont été identifiées mais aucune ne semble produire de protéines DING. Le rôle de ces protéines dans la physiologie de la graine de *P. gabriellae* et dans l'adaptation au nickel restent à explorer.

Par ailleurs, cette approche protéomique a été complétée par l'obtention des transcrits exprimés au cours de la formation de la graine de *Psychotria gabriellae*. Cette base de données représente une source de données génomique utile pour approfondir les mécanismes impliqués dans la mise en place de l'hyperaccumulation de nickel chez les plantes, mécanismes qui pourront un jour s'avérer utiles pour répondre à des questions de phytoremédiation.

**Mots clés** : dormance, germination, protéomique, nickel, hyperaccumulation, *Psychotria gabriellae*, *Amborella trichopoda*, Nouvelle-Calédonie

## Abstract

New Caledonia possesses one of the world most rich, diverse and unique flora. Its high endemism (74,7%) is partly due to the gondwanian origin of its flora and to the high speciation induced by the ultramafic soils rich in heavy metals, including nickel. If this element is the source of the country richness, its mining exploitation and human colonization of the land induce ecosystems degradation. The study and comprehension of the new Caledonian flora is essential to be able to preserve, protect and restore its rich biodiversity. Preservation and restoration both depend on seeds. They are the unit of dispersal of higher plants, and responsible of their world domination on flora. We focused our study on the biochemical characterisation of seed biology of two extraordinary species, *Amborella trichopoda*, the sister to all extant flowering plants and *Psychotria gabriellae*, one of the world most nickel hyperaccumulating plant.

Proteomic characterisation of *A. trichopoda* seeds was the first study that documented a rudimentary embryo proteome. This approach provides a better understanding of the mechanisms involved in the control of dormancy and germination of seeds with morphological dormancy such as *A. trichopoda*. The results obtained allow us to highlight the molecular maturity of the rudimentary embryo, as well as confirming the basal position of *Amborella trichopoda* through phylogenetic analyses of selected protein families. The characterisation of the protein evolution during germination highlights massive mobilisation of storage proteins before the end of germination *sensu stricto*, and suggests a new definition of germination for seeds with rudimentary embryo.

Previous proteomics characterisation of *P. gabriellae* seeds revealed a high representation of DING proteins that are known to be involved with ABC type transporters or to bind phosphorus. This observation was associated with an observed gradient of nickel inside the seed presumably to protect the embryo from its toxicity. During this work, we confirmed the presence of this protein family in the seeds, from which the belonging to the eukaryotic kingdom remains a subject of debate. To answer about the origin of these proteins in seed, we tried to determine the presence or not of bacteria in the dry mature seed. Four endophytic bacteria were identified but none of them seems to produce such proteins. However, the physiological signification of these bacteria to account for physiological features of the *Psychotria gabriellae* seeds and their exceptional tolerance toward nickel toxicity remains to be established.

Beside this proteomics approach, we sequenced a large number of transcripts expressed during *Psychotria gabriellae* seed formation. This database will enrich the very limited genomic data available for this specie. It will allow a better understanding of the mechanisms involved in nickel hyperaccumulation, and may highlight novel tools for phytoremediation.

**Keywords**: dormancy, germination, proteomic, nickel, hyperaccumulation, *Psychotria gabriellae*, *Amborella trichopoda*, New Caledonia