

## **Habilitation à Diriger des Recherches**

**Université de Lille**

**Ecole Doctorale Sciences de la Matière, du Rayonnement et de l'Environnement**

**ED - 104**

# **Métabolisme et rôles des composés phénoliques chez les plantes**

par **David GAGNEUL**

**Maître de Conférences des Universités**

**CNU 66 - Physiologie**

soutenue le 8 avril 2021 devant la commission d'examen :

Sylvie BAUDINO, Professeur, Université de Saint-Etienne, Rapporteur

Emmanuel GAQUEREL, Professeur, Université de Strasbourg, Rapporteur

Nathalie GIGLIOLI-GUIVARC'H, Professeur, Université de Tours, Rapporteur

Alain BOUCHEREAU, Professeur, Université de Rennes 1, Examineur, Président du jury

Jean-Louis HILBERT, Professeur, Université de Lille, Garant







## **Habilitation à Diriger des Recherches**

**Université de Lille**

**Ecole Doctorale Sciences de la Matière, du Rayonnement et de l'Environnement**

**ED - 104**

# **Métabolisme et rôles des composés phénoliques chez les plantes**

par **David GAGNEUL**

**Maître de Conférences des Universités**

**CNU 66 - Physiologie**

soutenue le 8 avril 2021 devant la commission d'examen :

Sylvie BAUDINO, Professeur, Université de Saint-Etienne, Rapporteur

Emmanuel GAQUEREL, Professeur, Université de Strasbourg, Rapporteur

Nathalie GIGLIOLI-GUIVARC'H, Professeur, Université de Tours, Rapporteur

Alain BOUCHEREAU, Professeur, Université de Rennes 1, Examineur, Président du jury

Jean-Louis HILBERT, Professeur, Université de Lille, Garant

Institut Charles Viollette, UMRT BioEcoAgro, n°1158, Université de Lille

*"The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge" Stephen Hawking*

## Remerciements

Je tiens tout d'abord à remercier les membres du jury qui ont accepté d'évaluer ce travail : Sylvie Baudino, Emmanuel Gaquerel et Nathalie Giglioli-Guivarc'h en tant que rapporteurs et Alain Bouchereau en tant qu'examineur.

Mes remerciements vont ensuite à Jean-Louis Hilbert, garant de mon HDR, pour m'avoir accueilli dans son laboratoire il y a plus de 10 ans maintenant, pour m'avoir accordé sa confiance, pour avoir soutenu mes activités de recherche et m'avoir donné l'opportunité de co-encadrer avec lui plusieurs doctorants.

Je tiens à remercier les personnes qui m'ont fait confiance et qui m'ont permis d'être en mesure aujourd'hui de présenter mon HDR : Alain Bouchereau, mon directeur de thèse, ainsi que M. le Professeur Larher dont le soutien scientifique, la rigueur et la bienveillance au cours de mes années de thèse ont largement contribué à ma formation et Andreas Weber qui m'a accueilli dans son laboratoire aux Etats-Unis puis en Allemagne.

Un grand merci à tous les étudiants encadrés pendant ces années à l'Université de Lille et, en particulier, Guillaume Legrand, Marianne Delporte et Julie Buges qui ont accepté de réaliser une thèse à mes côtés. Merci également à Guillaume Bernard dont les travaux de thèse ont largement contribué à mes projets. Le travail présenté dans ce manuscrit est le fruit du travail de ces doctorants et de nombreux autres stagiaires de M2, M1, L3 et BTS.

Merci à Caroline Rambaud qui a mis à contribution son expertise pour faire avancer nos problématiques. Ton enthousiasme est sans limite.

Merci à Monika Mörchen et Amandine Watrin dont l'aide technique a été très déterminante pour l'accomplissement des projets présentés dans ce manuscrit.

Merci à tous nos collaborateurs pour leur aide précieuse dans la réalisation de ces travaux. Bien souvent ces collaborations ont permis de débloquer des situations ou de magnifier les résultats obtenus au laboratoire et également de faire des rencontres très enrichissantes scientifiquement et humainement.

Merci à toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Merci aux membres du laboratoire du SN2 (anciens ou actuels).

Merci à l'équipe pédagogique de physiologie végétale et de biologie végétale pour m'avoir intégré rapidement dans différents modules et m'avoir accordé leur confiance.

Merci à Yasmine pour son soutien sans faille. Merci à mes 2 enfants, Inès et Paul, sources intarissables d'inspiration... Merci à tous les 3 pour le bonheur que vous m'apportez au quotidien.





## SOMMAIRE

<b>I- Curriculum vitae</b> .....	<b>3</b>
<b>II- Liste de la production scientifique</b> .....	<b>9</b>
1. Articles dans des revues internationales avec comité de lecture (rang A) .....	9
2. Communications orales (*: orateur).....	11
3. Communications par affiches.....	11
4. Rayonnement de nos travaux .....	12
5. Brevet.....	13
<b>III- Activités scientifiques antérieures à mon recrutement</b> .....	<b>15</b>
1. Travaux de thèse .....	15
1.1. Caractérisation des modalités d'osmorégulation du contenu en osmolytes chez 2 espèces de <i>Limonium</i> soumises à différents traitements salins.....	18
1.2. Dynamique d'accumulation constitutive ou osmo-induite d'osmolytes chez <i>L. latifolium</i> au cours des premiers stades de la vie végétative en présence ou non de sel.....	20
1.3. Implication du catabolisme oxydatif des polyamines aliphatiques dans la production de $\beta$ -alanine, précurseur privilégié de la $\beta$ -AB.....	22
2. Travaux de post-doctorat.....	26
2.1. Caractérisation de transporteurs impliqués dans la photorespiration.....	26
2.2. Caractérisation des éléments moléculaires à l'origine de la transition d'une photosynthèse de type C3 vers un type C4 .....	35
2.3. Séquençage du premier génome d'algue rouge unicellulaire.....	38
<b>IV- Métabolisme et rôles des acides hydroxycinnamiques</b> .....	<b>41</b>
1. Introduction .....	41
2. Bilan des activités de recherche.....	59
2.1. Développement de l'outil qRT-PCR chez la chicorée.....	59
2.2. Edition du génome de chicorée par la technologie CRISPR/Cas9 .....	61
2.3. Identification et classement des membres de la famille BAHD chez la chicorée .....	62
2.4. Caractérisation de la voie de biosynthèse du CQA chez la chicorée .....	64
2.5. Caractérisation de 2 enzymes impliquées dans la synthèse de phénolamides originales accumulées dans le manteau pollinique des Astéracées.....	65
2.6. Caractérisation de la voie de biosynthèse du diCQA .....	69
3. Projets de recherche .....	73
3.1. Caractérisation fonctionnelle et biochimique des SHT-like .....	73
3.2. Métabolisme et rôle du diCQA chez la chicorée.....	79
3.3. Caractérisation des enzymes du sous-groupe spécifique du clade IIIa des BAHDs.....	80
3.4. Caractérisation du métabolisme du CTA et du diCTA.....	80
3.5. Analyse de la redondance fonctionnelle des HCTs et des HQTs.....	81
3.6. Caractérisation du métabolisme des flavones C-glycosides chez le lin.....	82
4. Conclusion .....	85
<b>V. Bibliographie</b> .....	<b>87</b>
<b>VI- Annexes</b> .....	<b>95</b>



## I- Curriculum vitae

### David Gagneul

Né le 02/09/1976 à Rennes (Ille-et-Vilaine)

Situation familiale : Pacsé, 2 enfants

Nationalité : Française

#### Adresse professionnelle :

UMRT BioEcoAgro n°1158

Pôle « Biomolécules d'origine végétale et microbienne : de l'identification à la bioproduction »

Equipe « Métabolites spécialisés d'origine végétale »

Institut Charles Viollette

Université de Lille

Cité Scientifique

Bât SN2, 3ème étage, Bur. 309

F-59655 Villeneuve d'Ascq Cédex

Téléphone : (+33) 320434021

david.gagneul@univ-lille.fr

---

### Situation actuelle

Maître de conférences, Classe Normale, Université de Lille, Section CNU 66

---

### Formations universitaires

- 2005** Thèse de Doctorat de l'Université de Rennes 1, Allocataire de Recherche du MENRT  
**Ecole Doctorale :** Vie, Agronomie et Santé  
**Laboratoire d'accueil :** Equipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026, Université de Rennes 1  
**Titre de la thèse :** Accumulation constitutive ou halo-induite de solutés compatibles chez deux espèces de *Limonium* (*Plumbaginaceae*) : Aspects quantitatifs et métaboliques  
**Directeur de thèse :** Pr. A. Bouchereau  
**Soutenance :** 2 décembre 2005  
**Mention :** Très honorable avec félicitations du jury  
**Composition du jury :**  
J-L. Prioul, Université de Paris-Sud, Rapporteur  
S. Lutts, Université Catholique de Louvain (Belgique), Rapporteur  
E. Deslandes, Université de Bretagne Occidentale, Examineur  
P. Potin, CNRS, Station Biologique de Roscoff, Examineur  
F.R. Larher, Université de Rennes 1, Examineur  
A. Bouchereau, Université de Rennes 1, Directeur de thèse
- 2001** **DEA** de Génétique, Adaptation et Production Végétales, option "Physiologie et Agrophysiologie. Université de Rennes 1. Mention bien.  
**Titre :** Accumulation halodépendante de proline et de bêtaïnes et activités d'enzymes impliquées dans la biosynthèse de la  $\beta$ -alanine chez *Limonium* sp. (Plombaginacées)
- 2000** **Maîtrise** de Biologie Cellulaire et Physiologie, option Physiologie Végétale. Université de Rennes 1. Mention Bien.
- 1999** **Licence** de Biologie, option Biologie Cellulaire et Physiologie. Université de Rennes 1.

1998 **DEUG** de Sciences, option Sciences de la Vie. Université de Rennes 1. Mention Bien.

### **Expériences professionnelles**

---

**Depuis septembre 2009 : Maître de Conférences CN (CNU 66), Université de Lille.**

**2007-2009 Chercheur post-doctoral**, Heinrich-Heine-University, Institute for Plant Biochemistry, Düsseldorf, Allemagne.

**2006-2007 Chercheur post-doctoral**, Michigan State University, Department of Plant Biology, East Lansing, USA.

**2004-2005 Attaché temporaire d'Enseignement et de Recherche**, demi-poste, section 66, au sein de l'équipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026, UFR SVE, Université de Rennes 1.

**2001-2004 Allocataire de recherche du MENRT**, au sein de l'équipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026, Université de Rennes 1.

**2001-2004 Moniteur (CIES)**. Equipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026 ICM, UFR SVE, Université de Rennes 1.

**2001 Stage de DEA** au sein de l'équipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026, Université de Rennes 1 (Directeur : Pr. Alain Bouchereau).

**2000 Stage de Maîtrise** réalisé à l'I.N.R.A. d'Avignon (Domaine Saint Paul, Montfavet) dans l'Unité de Recherche en Ecophysiologie et en Horticulture (Directeur : Pr. Loïc Pagès).

### **Activités d'encadrement**

---

- **Thèses**

**Octobre 2018 - septembre 2021** : Co-encadrement (50%) de la thèse de **Julie Buges**. « Métabolisme et rôles des phénolamides ». Directeur de thèse : Jean-Louis Hilbert.

**Octobre 2014 - septembre 2017** : Co-encadrement (50%) de la thèse de **Marianne Delporte**. « Néo-fonctionnalisation des acyltransférases de la famille BAHD : conséquences sur le métabolisme des acides hydroxycinnamiques des Astéracées ». Directeur de thèse : Jean-Louis Hilbert.

**Octobre 2011 - septembre 2015** : Co-encadrement (50%) de la thèse de **Guillaume Legrand**. « Contribution à la caractérisation du métabolisme des acides chlorogéniques chez la chicorée : approches biochimique et moléculaire ». Directeur de thèse : Jean-Louis Hilbert.

- **Stage de Master 2 ou équivalents**

**Février – juillet 2020** : Stage de master 2, **Mohamad Cheaib**. Caractérisation des métabolismes primaire et spécialisé de chevelus racinaires de chicorée.

**Mars - septembre 2017** : Stage de fin d'études d'Ingénieur (ISARA), **Emeric Morel**. Métabolisme et rôle des dérivés d'acide hydroxycinnamique chez les plantes.

**Janvier - juin 2015** : Stage de master 2, **Gianni Galati**. Caractérisation de la voie de biosynthèse de l'acide caftarique chez la chicorée *Cichorium intybus* L..

**Mars - septembre 2014** : Stage de fin d'études d'Ingénieur (ENSAIA), **Marianne Delporte**. Vers une meilleure compréhension du métabolisme des acides hydroxycinnamiques chez *Cichorium intybus* L. : analyses transcriptionnelle et biochimique.

**Janvier - juin 2013** : Stage de master 2, **Chahinez Khelifi**. Caractérisations fonctionnelle et biochimique d'enzymes impliquées dans la voie de biosynthèse de l'acide chlorogénique chez la chicorée.

**Janvier - juin 2011** : Stage de master 2, **Guillaume Legrand**. Métabolisme des acides hydroxycinnamiques et halodépendance de leur accumulation chez *Cichorium intybus* L..

- **Stage de Master 1**

**Juin - juillet 2017** : Stage de 2 mois non obligatoire, **Valene Khouri**. Isolation de plantes transgéniques.

**Stage de licence 3, obligatoire, 1 mois**

**Mai 2018** : **Damien Vicaire**. Analyse structurale de protéines de la sous famille SHT-like de la famille des acyltransférases BAHD.

**Mai 2017** : **Audrey Bloch**. Caractérisation biochimique d'enzymes impliquées dans la synthèse des dérivés d'acide chlorogénique.

**Mai 2016** : **Julie Buges**. Molecular cloning and heterologous expression of enzymes from chicory involved in phenylpropanoid metabolism.

**Mai 2014** : **Ingrid Verfaillie**. Contribution à la caractérisation biochimique et fonctionnelle de gènes de la famille des BAHD de la chicorée.

**Mai 2013** : **Jennifer Vanhove**. Répartition tissulaire des acides hydroxycinnamiques au cours du développement de *Cichorium intybus* L..

**Avril 2011** : **Antoine Clément**. Mise en évidence de précurseurs potentiels de la voie de biosynthèse de composés phénoliques chez *Cichorium intybus*.

- **Stage de BTS**

**Mai - juillet 2020 puis septembre – octobre 2020** : **Valaury Brogniez**. Métabolisme des acides hydroxycinnamiques.

**Mai - juillet 2019 puis septembre - octobre 2019** : **Bryan Garin**. Expression transitoire de protéines chez *Nicotiana benthamiana*.

**Mai - juillet 2019** : **Yann Degroise**. Production et purification de protéines recombinantes et mesure de leur activité enzymatique.

**Mai - juillet 2018 puis octobre - décembre 2018** : **Lucie Plaitin**. Préparation des outils moléculaires nécessaires à l'introduction de nouvelles voies métaboliques chez *Synechocystis* sp. PCC6803.

**Mai - juillet 2017 puis octobre - décembre 2017** : **Sophie Bayart**. Mise au point d'une méthode de dosage du CoASH.

**Mai - juillet 2016 puis octobre - décembre 2016** : **Amandine Watrin**. Production d'esters de CoA par voie enzymatique.

## **Rayonnement**

---

**Bibliométrie** : h-index : 14 ; 911 citations, *Source* : <https://www.scopus.com> (18 janvier 2021)

## **Responsabilités scientifiques**

Membre élu du conseil de laboratoire de l'Institut Charles Violette depuis 2019.

Responsable de la thématique « Phénolamides » de l'équipe mixte laboratoire-entreprise "CHIC41H" (Chicory for one health) dirigée par le Dr Caroline Rambaud (1 M€).

Coordinateur du projet DECRIP (Développement d'un pipeline pour le criblage de gènes impliqués dans la synthèse de molécules à haute valeur ajoutée) financé par la SFR Condorcet (2019). Ce projet impliquait l'équipe BIOPI (Université de Picardie Jules Verne, UPJV) et la plateforme PFA (UPJV). Il s'agissait d'un projet de 1 an (10 k€).

Participation au projet CRISTALIN (Contribution des flavones C-glucoside dans la tolérance au stress froid du lin ; dispositif STIMULE, région Hauts-de-France ; coordinateur Anthony Quéro, BIOPI, UPJV) (70 k€).

Référent pour les demandes d'utilisation d'OGM au laboratoire (rédaction et dépôt du dossier d'agrément).

Responsable d'une plateforme de profilage métabolique (Institute for Plant Biochemistry, Heinrich-Heine University, Düsseldorf, Germany) entre 2007 et 2009.

### **Implication dans des sociétés savantes**

Membre du conseil d'administration de la Société Française de Biologie Végétale (depuis 2017).

Membre du réseau Métabolisme Spécialisé (MétaSpé).

### **Membre élu du CNU section 66 – 2020-2024**

**Evaluation d'articles pour des revues de rang A** : Journal of Experimental Botany, Frontiers in Plant Science, Plant, Cell and Environment, Plant Biology, Botany, PeerJ, Microbial Cell Factories, Agriculture, Metabolites...

**Review editor** de la revue Frontiers in Plant Science, section Marine and Freshwater plants.

**Guest editor** de la revue Metabolites pour un numéro spécial (publication en mai 2021) intitulé « Advances in plant specialized metabolites: how to take full advantages of the properties of specialized metabolites? ».

### **Organisation de congrès nationaux et internationaux**

Membre du comité scientifique des 7ème Journée Scientifique du Réseau Français de Métabolomique et Fluxomique (RFMF). Amiens, 10-13 juin 2013.

Membre du comité scientifique et du comité d'organisation des 9ème Journée Scientifique du Réseau Français de Métabolomique et Fluxomique (RFMF). Lille, 9-11 juin 2015.

Membre du comité d'organisation du PSE Young Scientist's Meeting (Phytochemical Society of Europe). Villeneuve d'Ascq, 28 juin - 1 juillet 2017.

Modérateur (chairman) de la session "Allelopathy - Metabolomics - Genomics" lors du PSE young scientists's meeting, natural products in health, agro-food and cosmetics, Villeneuve d'Ascq, 28 juin-1 juillet 2017.

Membre du comité d'évaluation pour l'attribution de bourses de voyage pour le congrès Plant Biology Europe 2020, FESPB, Turin, Italie, 29 juin - 2 juillet 2020 (congrès annulé et reporté en 2021).

### **Participation à des comités de suivi de thèse**

Nicole Pontarin, EA3900, Unité de Recherche "Biologie des Plantes et Innovation", Université de Picardie Jules Verne, Amiens (thèse soutenue en 2019).

Guillaume Bernard, Institut Charles Viollette, EA7394, Université de Lille (thèse soutenue en 2019).

Justine Domont, UMRT BioEcoAgro n°1158, Université de Lille.

### **Participation à des jurys de thèse**

Guillaume Legrand (2015), Université de Lille1 ; Marianne Delporte (2017), Université de Lille1 ; Guillaume Bernard (2019), Université de Lille.

### **Autres activités**

Membre élu des comités de sélection (collège rang B - spécialité CNU66) en vue du

recrutement des enseignants-chercheurs.

Mai 2014 : Président de jury du comité de sélection poste de technicien en SVT, Université de Lille 2 et ULCO

Je participe tous les ans aux **jurys nationaux de BTS** (BioAnalyses et Contrôle) au Lycée Valentine Labbé (la Madeleine). Il s'agit de l'audition de candidats à l'issue de leur stage (soutenance du rapport de stage).

### **Activités d'enseignement**

---

- 2009-** Maître de Conférences, Université de Lille.
- 2007-2009** Poste d'assistant-professeur, Institute for Plant Biochemistry, Heinrich-Heine-University, Düsseldorf, Allemagne.
- 2004-2005** Attaché Temporaire d'Enseignement et de Recherche (1/2 poste). Equipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026 ICM, UFR SVE, Université de Rennes 1.
- 2001-2004** Moniteur (CIES). Equipe Osmoadaptation et Métabolismes de Stress, UMR CNRS 6026 ICM, UFR SVE, Université de Rennes 1.

### **Enseignement à l'Université de Rennes 1**

J'ai été recruté en tant que moniteur pendant ma thèse (3 années, 64 h/an). J'ai pendant ces 3 années assuré mes enseignements en Travaux Pratiques de Physiologie Végétale en DEUG "Sciences de la vie" 2ème année, en Maîtrise "Sciences de la terre et de l'univers - Biologie Générale et Géologie", en Maîtrise "Biologie Cellulaire et Physiologie" et en IUP "Génie biologique option production végétale" (3ème année).

Par la suite, j'ai bénéficié d'un poste de demi ATER (2004-2005, 96h). J'ai effectué mes enseignements en Travaux pratiques et Travaux dirigés dans des modules de Physiologie Végétale en Maîtrise "Sciences de la terre et de l'univers -Biologie Générale et Géologie", en IUP "Génie Biologique" 1ère année et en IUP "Génie biologique option production végétale" (3ème année).

### **Enseignements à l'Université Heinrich-Heine, Düsseldorf, Allemagne**

De 2007 à 2009, j'ai été recruté sur un poste d'assistant professeur. Ceci impliquait d'effectuer une charge d'enseignement (150h/an, 140 h TP et 10 h TD). Ces enseignements étaient réalisés en anglais et s'adressaient à des étudiants de niveau Master (1 et 2). Ces enseignements s'adressaient à environ 15 étudiants. J'ai mis en place des travaux pratiques dont la thématique était la caractérisation du métabolisme de l'amidon chez des mutants d'*Arabidopsis thaliana*. J'ai également participé à la mise en place d'un module de TDs dédié aux techniques dites à haut-débit (génomique, transcriptomique, métabolomique, protéomique).

### **Enseignement dans le département de biologie, Université de Lille**

J'effectue, depuis mon recrutement en 2009, l'intégralité de mon service statutaire (formation initiale, 192 h).

Je suis responsable d'un module de Licence (S6) proposé à la filière Biologie Cellulaire et Physiologie et à la filière Science de la vie et de la Terre intitulé "Physiologie du Développement des Plantes". Ce module s'adresse à environ 60 étudiants. J'ai notamment effectué la refonte de ce module existant lors de l'élaboration des nouvelles maquettes en 2014 puis en 2020. La thématique est la physiologie du transport des métabolites dans la plante, de la fructification et de la tubérisation.

Je suis coresponsable de 2 modules en Master : "Outils modernes de l'innovation et de la recherche" et "Les dogmes de la biochimie et de la physiologie remis en cause" dans le master agroalimentaire régional, Nutrition et Sciences des Aliments (NSA) spécialité Innovations en

Biotechnologies Végétales, Enzymatiques et Microbiennes (IBVEM). Ces modules s'adressent à environ 15 étudiants.

Je contribue également, de manière très active, à la formation des étudiants au travers des stages (BTS, Licence3, Master1, Master2).

J'ai participé en 2019/2020 à un projet d'innovation pédagogique nommé A3PV (Accompagnement des pratiques pédagogiques en physiologie végétale ; coordinateur : Godfrey Neutelings). A ce titre, j'ai contribué à la réalisation de supports audiovisuels pour les TP de Physiologie Végétale (première et deuxième année). Ces vidéos servent de support à la préparation des TP pour les étudiants.

Voici, la répartition la plus récente de mes enseignements :

Niveau	Intitulé du module	Niveau	Nature	Volume
Licence	Biologie végétale 1	L1	TD / TP	7.5 / 10
	Physiologie végétale	L1	CM / TD / TP	15 / 13.5 / 6
	Valorisation végétale	L2	CM / TP	1.5 / 3
	Ecophysiologie végétale	L2	CM / TP	1.5 / 6
	Grande fonctions des végétaux	L2	TD / TP	9 / 24
	Défense des organismes	L3	CM	3
	Biologie du développement	L3	TD / TP	3 / 18
	Physiologie du développement des plantes	L3	CM / TD / TP	6 / 2 / 15
	Perception et communication chez les plantes	L3	TD	2
Master	Métabolites spécialisés des plantes	M1	CM / TD / TP	6 / 4 / 16
	Ingénierie cellulaire	M2	CM	4
	Métabolisme spécialisé des plantes	M2	CM / TD	4 / 2
	Bioprocédés	M2	TP	5
	Ateliers Biotechnologiques	M2	TP	7

Ainsi, 25% de mes enseignements sont effectués en Master et 75% au niveau Licence. Environ 20% de mes enseignements sont des cours magistraux.



## II- Liste de la production scientifique

**25 articles** dans des revues internationales à comité de lecture dont 4 en dernier auteur et corresponding author, 2 en premier auteur, 5 en deuxième auteur et 2 en avant dernier auteur.

**Un brevet** a été déposé le 13 mars 2020.

h-index : 14 ; 911 citations (Source : <https://www.scopus.com> (18 janvier 2021))

### 1. Articles dans des revues internationales avec comité de lecture (rang A)

Les noms des étudiants encadrés sont soulignés.

25. Perrin J., Kulagina N., Unlubayir M., Munsch T., Carqueijeiro I., Dugé de Bernonville T., De Craene J.O., Clastre M., St-Pierre B., Giglioli-Guivarc'h N., **Gagneul D.**, Lanoue A., Courdavault V., Besseau S., 2021. Exploiting spermidine N-hydroxycinnamoyl transferase diversity and substrate promiscuity to produce various tri-hydroxycinnamoyl spermidines and analogs in engineered yeast. *ACS Synth. Biol.*, in press.
24. Roumani M., Besseau S., **Gagneul D.**, Robin C., Larbat R., 2020. Phenolamides in plants: an update on their function, regulation, and origin of their biosynthetic enzymes. *J. Exp. Bot.*, in press.
23. Miguel S., Legrand G., Duriot L., Delporte M., Menin B., Michel C., Olry A., Chataigné G., Salwinski A., Bygdell J., Vercaigne D., Wingsle G., Hilbert J.L., Bourgaud F., Hehn A., **Gagneul D.**, 2020. A GDSL lipase-like from *Ipomoea batatas* catalyzes efficient production of 3,5-diCQA when expressed in *Pichia pastoris*. *Comm. Biol.*, 3: 673.
22. Bernard G., Dos Santos H., Etienne A., Samaillie J., Neut C., Sahpaz S., Hilbert J.L., **Gagneul D.**, Jullian N., Tahrioui A., Chevalier S., Rivière C., Rambaud C., 2020. MeJA elicitation of chicory hairy roots promotes efficient increase of 3,5-diCQA accumulation, a potent antioxidant and antibacterial molecule. *Antibiotics*, 9: 659.
21. Pontarin N., Molinié R., Mathiron D., Tchoumtchoua J., Bassard S., **Gagneul D.**, Thiombiano B., Demailly H., Fontaine J.X., Guillot X., Sarazin V., Quéro A., Mesnard F., 2020. Age-dependent metabolic profiles unravel the metabolic relationships within and between flax leaves (*Linum usitatissimum*). *Metabolites*, 10: 218.
20. Tchoumtchoua J., Mathiron D., Pontarin N., **Gagneul D.**, van Bohemen A.I., Otogo N'ngang E., Mesnard F., Petit E., Fontaine J.X., Molinié R., Quéro A., 2019. Phenolic profiling of flax highlights contrasting patterns in winter and spring varieties. *Molecules*, 24: 1-14.
19. Bernard G., **Gagneul D.**, Alves Dos Santos H., Etienne A., Hilbert J.L., Rambaud C., 2019. Efficient genome editing using CRISPR/Cas9 technology in chicory. *Int. J. Mol. Sci.*, 20: 1155.
18. Bogdanovic M., Cankar K., Todorovic S., Dragicevic M., Simonovic A., van Houwelingen A., Schijlen E., Schipper B., **Gagneul D.**, Hendriks T., Quillet M.C., Bouwmeester H., Bosch D., Beekwilder J., 2019. Tissue specific expression and genomic organization of bitter sesquiterpene lactone biosynthesis in *Cichorium intybus* L. (Asteraceae). *Ind. Crops Prod.*, 129: 253-260.
17. Delporte M., Bernard G., Legrand G., Hielscher B., Lanoue A., Molinié R., Rambaud C., Mathiron D., Besseau S., Linka N., Hilbert J.L., **Gagneul D.**, 2018. A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in pollen coat of the Asteraceae family. *J. Exp. Bot.*, 69: 5365-5371.
16. Monné M., Daddabbo L., **Gagneul D.**, Obata T., Hielscher B., Palmieri L., Miniero D.V., Fernie A.R., Weber A.P.M., Palmieri F., 2018. Uncoupling proteins 1 and 2 (UCP1 and UCP2) from *Arabidopsis thaliana* are mitochondrial transporters of aspartate, glutamate and dicarboxylates. *J. Biol. Chem.*, 293, 4213-4227.

15. Legrand G., Delporte M., Khelifi C., Harant A., Vuylsteker C., Mörchen M., Hance P., Hilbert J.L., **Gagneul D.**, 2016 Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.*, 7, 741.
14. Delporte M., Legrand G., Hilbert J.L., **Gagneul D.**, 2015. Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. *Front. Plant Sci.*, 6, 651.
13. Schönknecht G., Chen W.H., Ternes C.M., Barbier G.G., Shrestha R.P., Stanke M., Bräutigam A., Baker B.J., Banfield J.F., Garavito R.M., Carr K., Wilkerson C., Rensing S.A., **Gagneul D.**, Dickenson N.E., Oesterhelt C., Lercher M.J., Weber A.P.M., 2013. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science*, 339, 1207-1210.
12. Eisenhut M., Planchais S., Cabassa C., Guivarc'h A., Justin A.M., Taconnat L., Renou J.P., Linka M., **Gagneul D.**, Timm S., Bauwe H., Carol P., Weber A.P.M., 2013. Arabidopsis A BOUT DE SOUFFLE is a putative mitochondrial transporter involved in photorespiratory metabolism and is required for meristem growth at ambient CO<sub>2</sub> levels. *Plant J.*, 73, 836-849.
11. Larher F.R., **Gagneul D.**, Niogret M.F., Emery N., Monnier C., Maerskalck V., Quillien G., Lefort C., Plasman M., Bouchard F. and Bouchereau A., 2013. Potentiation of proline accumulation in oilseed rape leaf discs exogenously supplied with combinations of PEG and cryoprotective agents is associated with overproduction of ABA, *Env. Exp. Bot.*, 87, 167-178.
10. Bräutigam A., Kajala K., Wullenweber J., Sommer M., **Gagneul D.**, Weber K.L., Carr K.M., Gowik U., Mass J., Lercher M.J., Westhoff P., Hibberd J.M. and Weber A.P.M., 2011. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species, *Plant Physiol.*, 155 : 142-156.
9. Larher F.R., Lukan R., **Gagneul D.**, Monnier C., Lespinasse Y. and Bouchereau A., 2009. A reassessment of the prevalent organic solutes constitutively accumulated in Pear leaves which can play a role in osmotic adjustment, *Env. Exp. Bot.*, 66, 230-241.
8. **Gagneul D.**, Aïnouche A., Duhazé C., Lukan R., Larher F.R. and Bouchereau A., 2007. A reassessment of the function of the so-called compatible solutes in the halophytic *Plumbaginaceae* *Limonium latifolium*, *Plant Physiol.*, 144, 1598-1611.
7. Larher F., **Gagneul D.**, Deleu C. and Bouchereau A., 2007. The physiological functions of nitrogenous solutes accumulated by higher plants subjected to environmental stress. *Acta Hort.*, 729, 33-41.
6. Petrivalsky M., Brauner F., Luhova L., **Gagneul D.** and Sebela M., 2007. Aminoaldehyde dehydrogenase activity during wound healing of mechanically injured pea seedlings, *J. Plant Physiol.*, 164, 1410-1418.
5. Lalouette L., Kostal V., Colinet H., **Gagneul D.** and Renault D., 2007. Metabolite profile during cold stress in the beetle *Alphitobius diaperinus* (Coleoptera: Tenebrionidae), *FEBS J.*, 274, 1759-1757.
4. **Gagneul D.**, Bräutigam A. and Weber A., 2007. High-throughput colorimetric method for the parallel assay of glyoxylic acid and ammonium in a single extract. *Anal. Biochem.*, 362, 151-153.
3. Tıprıdamaz R., **Gagneul D.**, Duhazé C., Aïnouche A., Monnier C., Özkum D. and Larher F.R., 2006. Clustering of halophytes from an inland salt marsh in Turkey according to their ability to accumulate sodium and nitrogenous osmolytes. *Env. Exp. Bot.*, 57, 139-153.
2. Duhazé C., **Gagneul D.**, Leport L., Larher F.R. and Bouchereau A., 2003. Uracil as one of the multiple sources of  $\beta$ -alanine in *Limonium latifolium*, a halotolerant  $\beta$ -alanine betaine accumulating *Plumbaginaceae*. *Plant Physiol. Biochem.*, 41, 993-998.
1. Duhazé C., Gouzerh G., **Gagneul D.**, Larher F.R. and Bouchereau A., 2002. The conversion of spermidine to putrescine and 1,3-diaminopropane in the roots of *Limonium tataricum*. *Plant Sci.*, 163, 639-646.

## 2. Communications orales (\*: orateur)

8. Delporte M., Bernard G., Legrand G., Hielscher B., Lanoue A., Molinié R., Rambaud C., Mathiron D., Besseau S., Linka N., Hilbert J.L., \***Gagneul D.**, 2019. Neofunctionalization of acyltransferases from the BAHD family results in the accumulation of unusual spermine-conjugated hydroxycinnamic acids in the pollen coat of the Asteraceae. *12<sup>th</sup> General Symposium of the French Society of Plant Physiology, Strasbourg, France.*
7. Delporte M., Bernard G., Legrand G., Hielscher B., Lanoue A., Molinié R., Rambaud C., Mathiron D., Besseau S., Linka N., Hilbert J.L., \***Gagneul D.**, 2018. Neofunctionalization of acyltransferases from the BAHD family results in the accumulation of a novel phenolamide in the pollen coat of the Asteraceae. *International SPS conference: Plants Sciences for the Future, Gif-sur-Yvette, France.*
6. \*Delporte M., Legrand G., Hilbert J.L., **Gagneul D.** 2016. Identification and characterization of BAHD acyltransferases involved in phenolic acid metabolism in chicory. *11<sup>th</sup> General Symposium of the French Society of Plant Physiology, Angers, France.*
5. \*Legrand G., Delporte M., Khelifi C., Mörchen M., Hilbert J.L., **Gagneul D.**, 2014. Identification and characterization of hydroxycinnamoyl transferases putatively involved in caffeic ester production in chicory. *8th congrès jeunes chercheurs de la SFBV, Tours, France.*
4. \*Hendricks T., Gonther L., Blassiau C., Legrand S., **Gagneul D.**, Huss B., Bergès H., Quillet M.C., 2014. Unravelling self-incompatibility in the Asteraceae: cloning of the S locus in chicory (*Chicorium intybus* L). *Plant & Animal Genome XXII, San Diego, USA.*
3. \***Gagneul D.**, Linka M., Linka N., Weber A., 2009. Towards the identification and functional characterization of mitochondrial transporters involved in photorespiration. *International Conference for Plant Mitochondrial Biology, Tahoe, USA.*
2. \***Gagneul D.**, Linka M., Linka N., Weber A., 2007. Towards the identification and characterization of transporters connecting the different compartments involved in photorespiration. *Botanical Congress, Hamburg, Germany.*
1. \*Weber A., Linka M., **Gagneul D.**, Braütigam A. and Linka N., 2007. Making the connections-the role of intracellular metabolite transport in photosynthesis and photorespiration. *Botanical Congress, Hamburg, Germany.*

## 3. Communications par affiches

18. Bernard G., **Gagneul D.**, Alves Dos Santos H., Etienne A., Hilbert J.L., Rambaud C., 2019. Efficient genome editing using CRISPR/Cas9 technology for basic research and chicory improvement. *12<sup>th</sup> General Symposium of the French Society of Plant Physiology, Strasbourg, France.*
17. Delporte M., Bernard G., Legrand G., Hielscher B., Lanoue A., Molinié R., Rambaud C., Mathiron D., Besseau S., Linka N., Hilbert J.L., **Gagneul D.**, 2018. Neofunctionalization of acyltransferases from the BAHD family results in the accumulation of a novel phenolamide in the pollen coat of the Asteraceae. *International SPS conference: Plants Sciences for the Future, Gif-sur-Yvette, France.*
16. Delporte M., Legrand G., Besseau S., Molinié R., Lanoue A., Mathiron D., Hilbert J.L., **Gagneul D.**, 2017. Metabolic diversification of pollen coat phenolamides in the Asteraceae family through neo-functionalization of a BAHD enzyme. *Plant Biology meeting, Hawaii, USA.*
15. Le Roy J., Huss B., Creach A., Bompard C., Delporte M., Mouajjah D., **Gagneul D.**, Hawkins S., Neutelings G., 2016. Is there a link between monolignol glycosylation and lignification in flax and Arabidopsis? *14<sup>th</sup> Cell Wall Meeting, Chania, Crete, Greece.*
14. Cankar K., Bogdanovic M., Van Houwelingen A., Schilijen E., Van de Geest H., Simonovic A., Todorovic S., Dragicevic M., Hendriks T., **Gagneul D.**, Quillet M.C., Bouwmeester H., Bosch D., Beekwilder J., 2015. Genomic organization and expression of sesquiterpene lactone biosynthetic genes in chicory. *12th TERPNET, Vancouver, Canada.*

13. Legrand G., Khelifi C., Mörchen M., Hilbert J.L., **Gagneul D.**, 2013. Molecular cloning and biochemical characterization of five hydroxycinnamoyl transferases putatively involved in chlorogenic acids accumulation in chicory. *10<sup>th</sup> International Congress of SFBV, Créteil, France.*
12. Legrand G., Mörchen M., Hance P., Vuylsteker C., Cadalen T., Hilbert J.L., **Gagneul D.**, 2012. Towards the identification of chlorogenic acid biosynthetic genes in *Cichorium intybus* L. *Colloque jeunes chercheurs, Grenoble, France.*
11. Harant A., Guitton Y., Hance P., Vanderrielle M., Mörchen M., Voedt N., **Gagneul D.**, Vuylsteker C., 2012. Production of 3,5-di-O-caffeoylquinic acid by *Cichorium intybus* cell suspension culture : metabolomic study by LC-DAD, LC-ESI-MSn and gene expression monitoring by RT-PCR. *6 JS RFMF, Nantes, France.*
10. Chrobok D., **Gagneul D.**, Marin K., Huhn S., Keech O., Obata T., Fernie A.R., Gardeström P., Krämer R., Weber A.P.M., 2012 Uncoupling proteins as putative pyruvate transporters in mitochondria. *PROMICS, Rostock, Germany.*
9. Quillet M.C., Blassiau C., Mörchen M., Habarugira I., Huss B., **Gagneul D.**, Gonthier L., Rambaud C., Cadalen T., Heuvelmans P., Van de Wal M., Prat E., Hilbert J.L., Hendriks T., 2011. Cloning and characterization of nuclear male sterility 1 (nms1) in chicory (*Cichorium intybus* L., Asteraceae). *Eucarpia Leafy Vegetables 2011, Lille, France.*
8. **Gagneul D.**, Sordet C., Duhazé C., Larher F.R. and Bouchereau A., 2006. Are the organic osmolytes that increase osmolarity of salt-grown seedlings of *Limonium latifolium* compatible solutes? *15<sup>ème</sup> congrès de la FESPB, Lyon, France.*
7. Larher F.R., Niogret M.-F., Monnier C., **Gagneul D.** and Bouchereau A., 2006. Agonistic effects between PEG 6000 and dimethylsulfoxide (DMSO) lead to overproduction of ABA and accumulation of proline (Pro) in rape leaf discs (RLDs). *15<sup>ème</sup> congrès de la FESPB, Lyon, France.*
6. Bouchereau A., **Gagneul D.**, Duhazé C., Le-Cahérec F., Leport L. and Larher F.R., 2005. Osmolyte accumulation in seedlings of the halophyte *Limonium latifolium* could arise without the inducing effect of sodium chloride. *6th General Symposium of the French Society of Plant Biology, Arcachon, France.*
5. **Gagneul D.**, Lugin R., Bouchereau A., Larher F., Leport L. and Niogret M.F., 2004. Metabolic snapshots of *Arabidopsis thaliana* exhibiting proline or glycine betaine accumulation when exposed or not to sodium chloride. *3<sup>rd</sup> International Congress on Plant Metabolomics, Ames, Iowa, USA.*
4. Larher F.R., **Gagneul D.**, Deleu C. and Bouchereau A., 2004. The physiological functions of nitrogenous solutes accumulated by higher plants subjected to environmental stress. *3<sup>rd</sup> BSVP; Bursa, Turkey.*
3. **Gagneul D.**, Duhazé C., Larher F.R. and Bouchereau A., 2003. Profiling and subcellular compartmentation of stress metabolites in the halotolerant species, *Limonium latifolium* (Plumbaginaceae). *5<sup>th</sup> General Symposium of the French Society of Plant Physiology, Plant Biology and the challenge of functional genomics, Orsay, France.*
2. Bouchereau A., **Gagneul D.**, Duhazé C., Moulin M., Leport L. and Larher F.R., 2003. Osmolyte profiling and related metabolic investigations in halotolerant *Limonium* species (Plumbaginaceae). *2<sup>nd</sup> International conference on Plant Metabolomics, Potsdam, Germany.*
1. Duhazé C., **Gagneul D.**, Moulin M., Larher F.R. and Bouchereau A., 2001. Coaccumulation and compartmentation of compatible solutes in halotolerant *Limonium* species in response to salt treatments. *International symposium, Plants under environmental stress, Moscow, Russia.*

#### 4. Rayonnement de nos travaux

1. Vogt T., 2018. Unusual spermine-conjugated hydroxycinnamic acids on pollen: function and evolutionary advantage. *J. Exp. Bot.*, 22: 5311-5318.

"Insight" soulignant l'originalité des résultats présentés dans la publication n°17.

## 5. Brevet

1. Miguel S., Bourgaud F., Salwinski A., Signe C., Hehn A., Hilbert J.L., **Gagneul D.**, Legrand G., déposé le 13 mars 2020. Procédé de production de l'acide 3,5-dicaféoylquinique (3,5-DCQ) par une levure (*Pichia pastoris*) exprimant une enzyme GDSL estérase/lipase recombinante d'*Ipomoea batatas*.



### III- Activités scientifiques antérieures à mon recrutement

#### 1. Travaux de thèse

Thèse soutenue le 2 décembre 2005

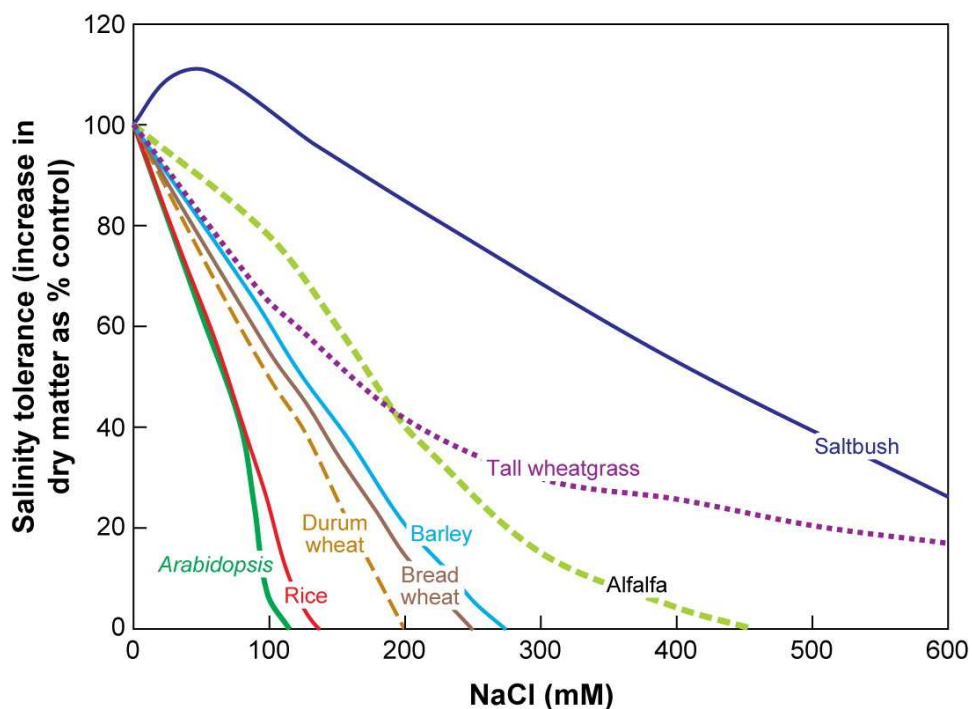
**Laboratoire d'accueil :** UMR CNRS 6026 ICM, Equipe Osmoadaptation et Métabolismes de Stress, Université de Rennes 1

**Directeur de thèse :** Prof Alain Bouchereau

Titre de la thèse : **Accumulation constitutive ou halo-induite de solutés compatibles chez deux espèces de *Limonium* (*Plumbaginaceae*) : aspects quantitatifs et métaboliques.**

Les plantes sont des organismes fixés qui ont la capacité à modifier leur morphologie, leur physiologie et leur métabolisme en réponse aux conditions environnementales. Cette flexibilité a permis aux plantes de coloniser presque toutes les niches écologiques de la planète et de survivre aux conditions les plus sévères. Aujourd'hui, la demande pour la nourriture est en augmentation permanente en lien avec une augmentation démographique mondiale et la diminution de la disponibilité des terres arables. Pour répondre aux besoins futurs des populations, une augmentation des rendements des plantes cultivées est donc nécessaire. La salinité des sols est l'un des facteurs abiotiques majeurs qui affecte la productivité des plantes cultivées. L'accumulation de sels sur les terres arables est principalement liée à l'irrigation avec de l'eau qui contient des traces de NaCl et à l'eau de mer<sup>1,2</sup>. La mise à jour des déterminants moléculaires à l'origine de la tolérance des plantes à la salinité constitue un enjeu majeur afin d'envisager des stratégies d'amélioration des résistances chez les espèces sensibles.

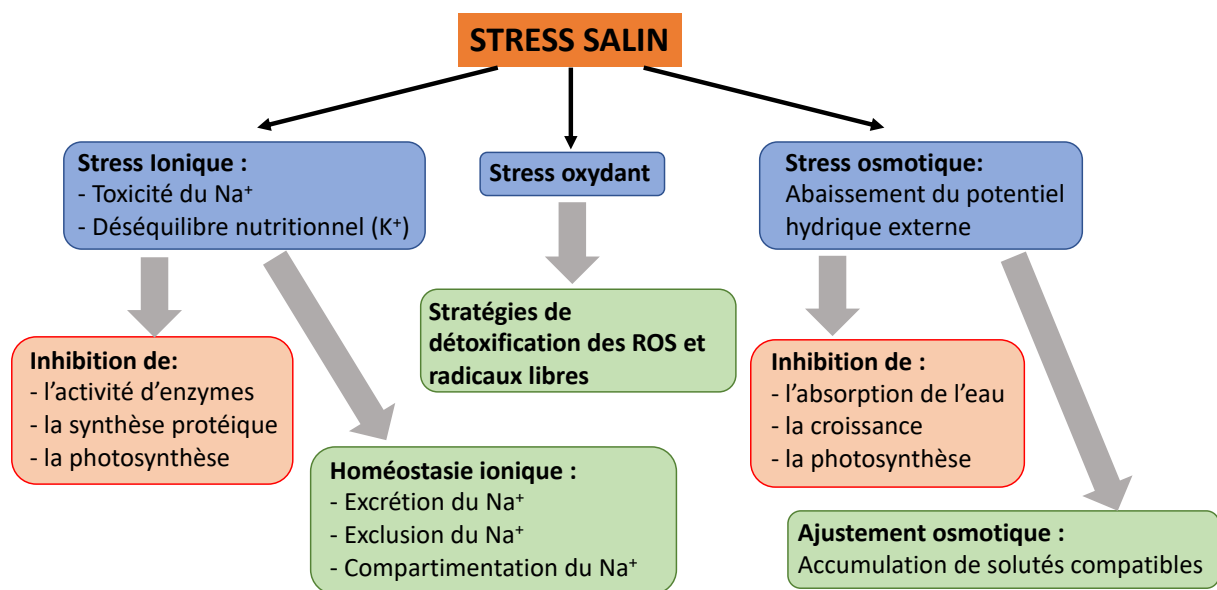
La tolérance des plantes à la salinité se définit comme leur capacité à croître et à effectuer leur cycle complet de développement sur un support contenant de fortes concentrations en sels (généralement le chlorure de sodium). Cette tolérance s'étale sur une large gamme (Fig. 1).



**Figure 1.** Degré de tolérance à la salinité de différentes espèces végétales. La tolérance est exprimée en augmentation de la masse sèche de la tige après au moins 3 semaines de croissance sur un milieu additionné de NaCl par rapport à la croissance en absence de sels (d'après Munns and Tester, 2008)<sup>3</sup>.

Ainsi, les halophytes (ou plantes halophiles) tolèrent des niveaux très élevés en sels tandis que les glycophytes sont peu tolérantes. Certaines halophytes nécessitent même une concentration en sels non nulle pour croître de manière optimale comme c'est le cas pour certaines espèces du genre *Atriplex* (Fig. 1).

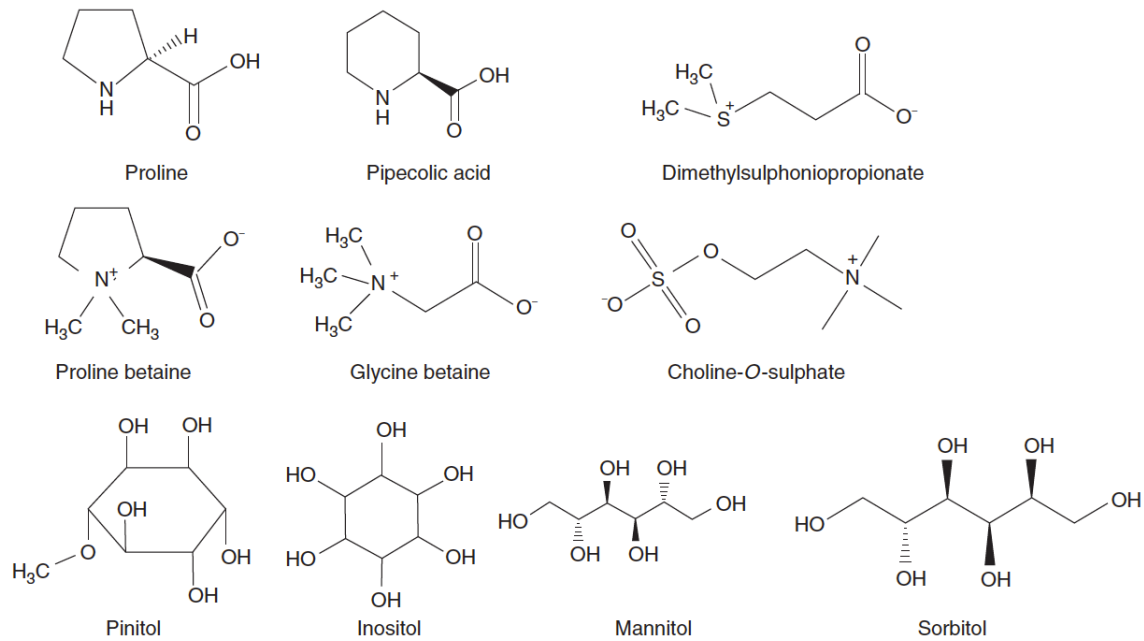
La discrimination des halophytes et des glycophytes se base sur la capacité de ces plantes à supporter tout au long de leur vie des concentrations plus ou moins importantes en NaCl. Ainsi, historiquement, cette limite était fixée à 86 mM de NaCl (0.5% NaCl)<sup>4</sup>. Plus récemment, une concentration de 200 mM a été suggérée afin de séparer les vrais halophytes des plantes relativement tolérantes<sup>5</sup>. De manière générale, les fortes concentrations en NaCl conduisent à différents stress chez les plantes : le stress ionique lié à l'absorption excessive d'ions indésirables et les stress nutritionnels associés, le stress osmotique lié à l'abaissement du potentiel hydrique externe et le stress oxydant qualifié de secondaire (Fig. 2). Les plantes utilisent différentes stratégies pour lutter contre les effets délétères du stress salin (Fig. 2).



**Figure 2.** Résumé synthétique des stress (en bleu) subis par la plante quand elle est soumise à de fortes concentrations en sels, les effets sur sa physiologie et son métabolisme (en rouge) et les réponses (en vert) mises en place pour survivre à leurs effets délétères (d'après Horie et al., 2012)<sup>6</sup>.

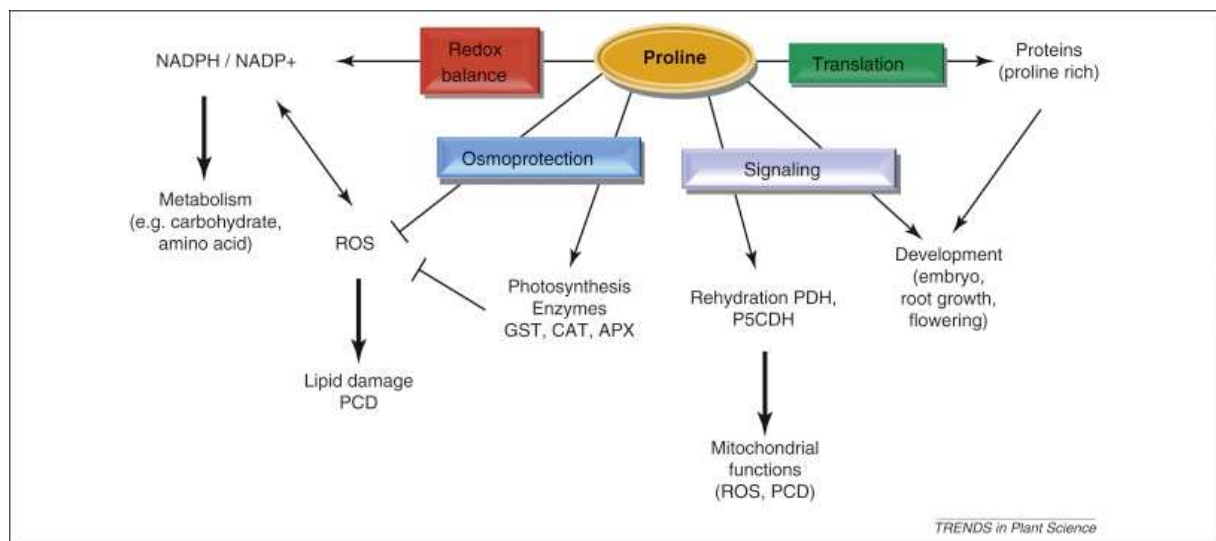
L'halophilie chez certaines plantes supérieures est basée sur des aptitudes particulières à accumuler les ions chlorure et sodium présents en excès dans le milieu extérieur et à les utiliser pour l'ajustement osmotique au niveau cellulaire. Ces ions indésirables sont séquestrés au niveau vacuolaire isolant ainsi le cytoplasme et ses organites de leurs effets néfastes. Le maintien de faibles concentrations en sodium dans le cytosol comparées à celles présentes dans la vacuole exige l'accumulation d'autres solutés dans le cytoplasme afin de réaliser l'ajustement osmotique des deux compartiments mitoyens. L'accumulation de solutés organiques non toxiques qualifiés de compatibles mis en évidence chez les halophytes et retrouvée chez les glycophytes répond à cette nécessité. Ces solutés organiques de faibles masses moléculaires, très solubles dans l'eau, neutres aux pH physiologiques peuvent s'accumuler à de fortes concentrations et ainsi participer à l'ajustement osmotique sans effet néfaste sur la physiologie cellulaire. Ces composés organiques osmoprotecteurs ou haloprotecteurs sont de natures chimiques très diverses (Fig. 3). Il peut s'agir d'acides aminés (proline, citrulline, acide pipécolique) et de leurs dérivés bétaïniques (glycine bétaïne,  $\beta$ -alanine bétaïne, choline-O-sulphate, proline bétaïne), de sucres solubles non structuraux (fructose, saccharose, fructanes) et de leurs dérivés polyhydroxylés (pinitol, sorbitol, *myo*-inositol, mannitol) et de composés à groupements sulfonium ternaire comme le diméthylsulfoniopropionate.





**Figure 3.** Structures chimiques des principaux solutés compatibles accumulés chez les plantes supérieures (d'après Abdelly et al. 2015)<sup>7</sup>.

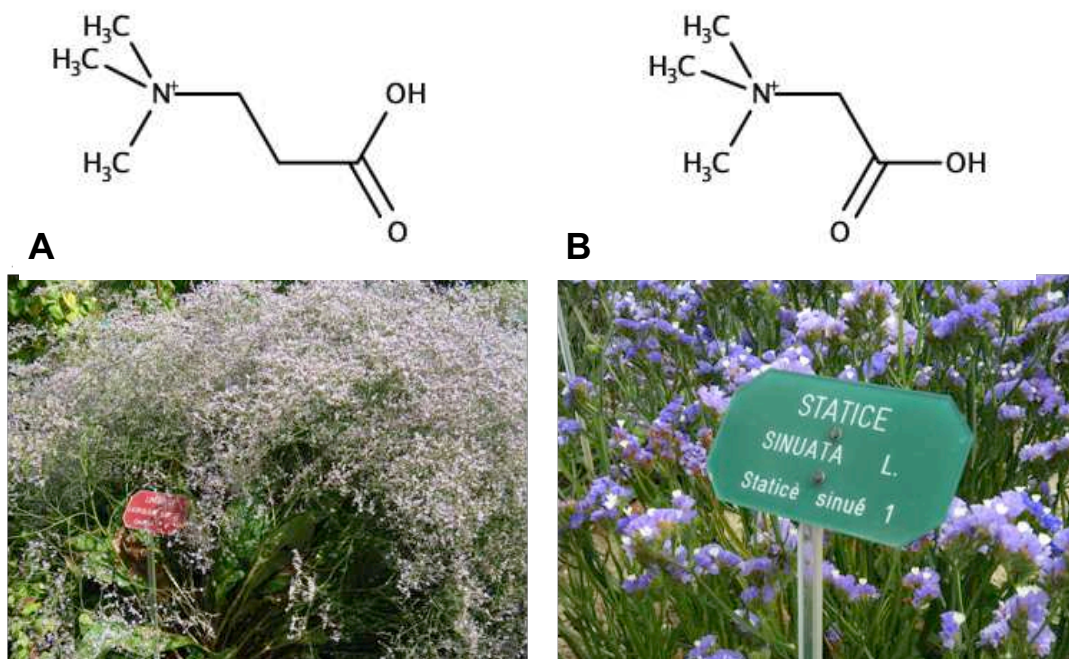
D'une manière générale, il est considéré par la plupart des auteurs que leur fonction principale est d'assurer l'ajustement osmotique au niveau cellulaire mais des études plus récentes suggèrent l'existence d'autres fonctions pour ces composés eux-mêmes et les voies métaboliques responsables de leur biosynthèse. Le cas le plus documenté est celui de la proline. Ainsi, le stress augmente sa synthèse alors que son catabolisme est stimulée pendant les phases de retour à des conditions plus clémentes<sup>8,9</sup>. Lors de la levée de la contrainte, la proline joue un rôle de molécule signal et serait impliquée dans la régulation de la prolifération cellulaire, de la mort cellulaire et de l'expression de gènes nécessaires au retour à ces conditions<sup>10</sup>. En plus de son rôle dans l'ajustement osmotique, la proline jouerait un rôle dans le piégeage des radicaux libres, dans l'homéostasie redox ou comme une chaperonne moléculaire stabilisant ainsi les protéines et les structures membranaires pendant les stress<sup>10</sup>(Fig. 4).



**Figure 4.** La proline a de nombreuses fonctions en plus de son rôle dans l'ajustement osmotique intracellulaire. La proline est utilisée pour la synthèse des protéines, a un rôle d'osmoprotecteur, contribue au maintien de l'homéostasie redox, régule le développement et est un acteur dans le réseau de signalisation qui contrôle les fonctions mitochondriales, le développement et les éléments nécessaires lors de la levée de la contrainte (d'après Szabados et Savouré, 2009)<sup>10</sup>.

Dans la mesure où les voies métaboliques concernées sont bien élucidées, notamment dans leur composante génique, les solutés compatibles constituent des cibles privilégiées pour le génie génétique et ses applications pour l'obtention de sources de tolérance au sel. Cependant, les expériences de transgénèse répertoriées mettent en évidence le fait qu'il s'agit tout d'abord de réaliser tous les approfondissements nécessaires avant d'exploiter leurs réelles potentialités. Les connaissances concernant la compartimentation des osmolytes, la description et la régulation des voies métaboliques et les flux métaboliques sous-tendus sont souvent superficielles et obèrent les potentialités des méthodes de transgénèse. Ceci justifie la poursuite de travaux de recherche visant à étudier les effets du sel tant sur des modèles glycophytes qu'halophytes sachant que les caractères à valeur adaptative seront plus facilement appréhendés sur les modèles halophytes.

Ma thèse avait pour objet l'étude des osmolytes synthétisés par les halophytes du genre *Limonium* en réponse à la salinité. Les espèces de ce genre coaccumulent des osmolytes de familles chimiques variées et en particulier des bêtaïnes réputées être des solutés compatibles très performants. Une étude par profilage métabolique des différentes classes de solutés en relation avec la salinité a été effectuée chez 2 espèces du genre *Limonium* (Fig. 5) considérées comme halophytes facultatives accumulant en complément d'autres solutés organiques comme la proline et les sucres polyhydroxylés soit la glycine bêtaïne (GB) (*L. sinuatum*), soit la  $\beta$ -alanine bêtaïne ( $\beta$ -AB) (*L. latifolium*). D'autre part, une étude de la compartimentation intracellulaire de ces solutés a été entreprise. Sur le plan métabolique, certaines voies possibles de biosynthèse de la  $\beta$ -alanine, le précurseur de la  $\beta$ -AB, ont été étudiées.

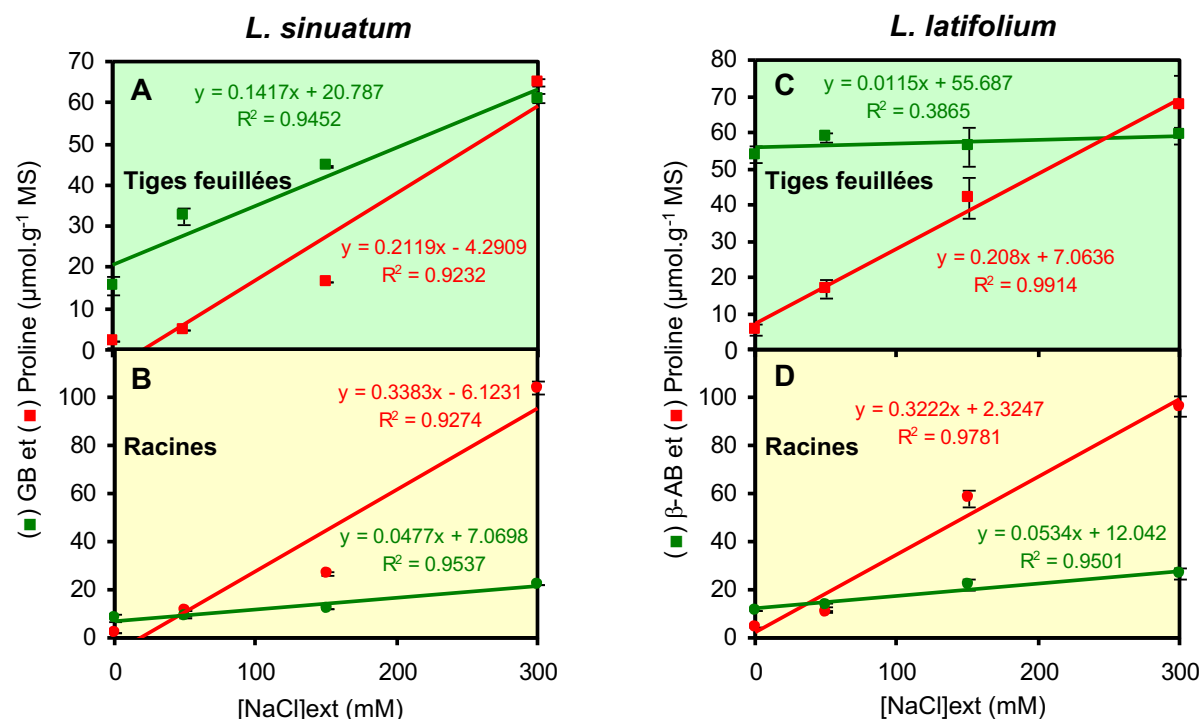


**Figure 5.** Plantes de *L. latifolium* (A) et de *L. sinuatum* (B) dans le jardin botanique du Thabor (Rennes, France) et la structure chimique de la bêtaïne accumulée par ces espèces,  $\beta$ -alanine bêtaïne et glycine bêtaïne, respectivement.

### 1.1. Caractérisation des modalités d'osmorégulation du contenu en osmolytes chez 2 espèces de *Limonium* soumises à différents traitements salins

Il est apparu que si les concentrations internes en proline et en GB coaccumulées chez *L. sinuatum* sont ajustées en relation avec la salinité du milieu extérieur, celle de la  $\beta$ -AB chez *L. latifolium* semble acquise de façon constitutive et indépendante des variations des contenus externes ou internes en sodium (Fig. 6). Les métabolismes prolinique et bêtaïnique, en termes de performance, chez les deux espèces à l'étude, semblent s'exclure mutuellement, les bêtaïnes étant plus abondamment accumulées dans les parties aériennes et la proline dans les parties racinaires. Ceci témoigne d'une sorte de gestion parcimonieuse par les 2 espèces

de leurs ressources et de leur énergie. Les mécanismes de régulation sous-jacents à ces options métaboliques coordonnées ne sont pas connus. Ceci suggère des actions spécifiques pour chacun de ces solutés dans les organes concernés et des modalités de régulations différentes des systèmes métaboliques impliqués.



**Figure 6.** Teneurs en proline et en bétaines (GB ou  $\beta$ -AB) des racines et des tiges feuillées de plantes de *L. sinuatum* (A et B) et de *L. latifolium* (C et D) ayant subi, pendant 10 jours, un traitement au NaCl à différentes concentrations dans les milieux de culture.

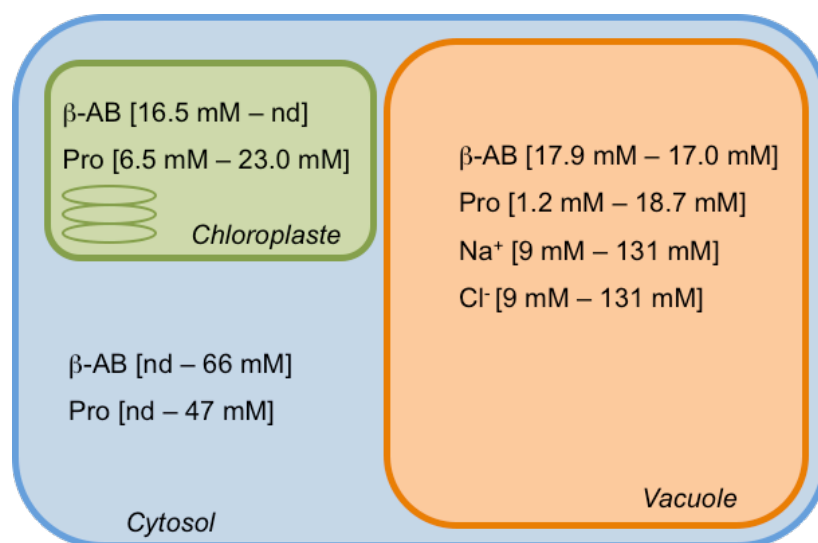
Les devenir métaboliques de ces solutés lors d'une séquence traitement hyperosmotique-traitement hypoosmotique ont également été appréhendés. Alors que le contenu en proline (accru en réponse au traitement salin, Fig. 6) décroît très rapidement après la levée de la contrainte (retour à un niveau identique à celui des plantes témoins 10 jours après la levée de la contrainte), ceux en GB (accru également en réponse à l'application de la contrainte) et en  $\beta$ -AB (élevé constitutivement mais stable en réponse à l'application du stress) restent inchangés au terme de la période d'étude. Ceci étaye l'hypothèse que ces bétaines constituent des "culs de sac" métaboliques alors que la proline accumulée en réponse au traitement hyperosmotique (ou hypersalin) se présente comme une réserve d'azote et de carbone recyclable rapidement lorsque les conditions redeviennent favorables. Ceci est en accord avec des études plus récentes menées chez *Arabidopsis*<sup>9</sup> et élargissant le rôle de la proline à d'autres fonctions que la simple régulation osmotique, en particulier lors de la levée des stress.

Malgré tout l'intérêt du couple "proline + bétaines", j'ai élargi la caractérisation, sur le plan structural, des solutés susceptibles de participer à l'ajustement osmotique aux acides aminés libres, aux cyclitols et aux sucres solubles non structuraux chez *L. latifolium*. Le contenu en sucres solubles non structuraux est apparu très important en l'absence de toute contrainte saline. En plus du saccharose et des hexoses habituels, *L. latifolium* s'est révélé être un accumulateur de *chiro*-inositol, ce cyclitol étant accompagné de *myo*-inositol ce qui constitue une situation métabolique rarement décrite chez les plantes. Cette étude cinétique et quantitative donnant lieu à un profilage métabolique ciblé combiné à une analyse en composantes principales (ACP) a permis de distinguer 3 catégories de solutés : (i) ceux dont le contenu interne est halorégulé et augmenté en présence de sel (proline, fructose, saccharose, glucose, *myo*-inositol et *chiro*-inositol), (ii) ceux dont le contenu est halorégulé et diminué en situation de traitement salin (glutamate, glutamine, aspartate) et (iii) ceux dont le

contenu est stable quel que soit le traitement appliqué ( $\beta$ -AB, choline, choline-O-sulfate). Il a été montré que les reconfigurations métaboliques induites par le sel sont réversibles lorsque le traitement est levé. L'accumulation d'autres solutés organiques, tels que les sucres solubles non structuraux et des cyclitols (*chiro*- et *myo*-inositol) en réponse à la contrainte saline suggère aussi leur implication dans l'ajustement osmotique ou l'osmoprotection.

La participation active de la  $\beta$ -AB à l'ajustement osmotique ayant été remise en cause (sur la base d'une étude quantitative), une étude de sa répartition tissulaire et intracellulaire a été entreprise. Les teneurs en  $\beta$ -AB décroissent selon un gradient s'étalant des feuilles jeunes vers les feuilles les plus âgées, ce qui pourrait correspondre à une accumulation privilégiée du soluté dans les organes les plus sensibles à l'invasion par le sodium. Cette allocation privilégiée de molécules protectrices dans les organes jeunes est conforme à la théorie dite théorie de la défense optimale (optimal defense theory)<sup>11,12</sup>.

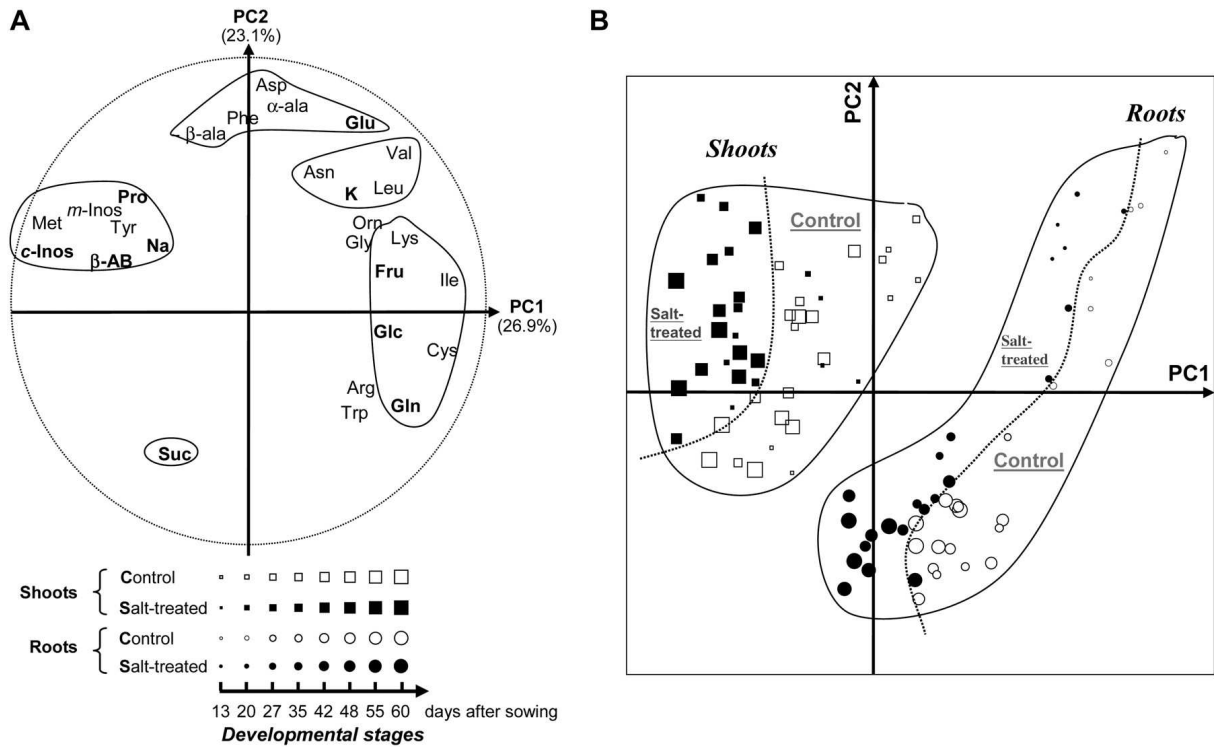
Pour vérifier l'hypothèse selon laquelle la  $\beta$ -AB, bien que non accumulée en réponse au sel, jouerait un rôle non négligeable d'osmorégulateur en migrant, *via* des processus halodépendants, d'un ou plusieurs compartiments de réserve vers des sites demandeurs en particules osmotiquement actives, une méthode de fractionnement subcellulaire en milieu non aqueux a été développée au laboratoire ceci afin de préciser la localisation subcellulaire de ce composé et de la proline et les variations susceptibles d'être déclenchées par le sel. Outre une répartition des solutés remettant en cause la définition même des solutés dits compatibles (quantité importante de solutés localisée dans la vacuole), un mécanisme original de localisation-relocalisation de la  $\beta$ -AB en réponse à la salinité a été mis en évidence (Fig. 7). La  $\beta$ -AB pourrait donc jouer un rôle osmotique puisqu'elle fait l'objet d'une redistribution interne au niveau cellulaire lorsque la concentration en sel du milieu extérieur varie. Ces résultats vont à l'encontre du dogme selon lequel à l'échelle cellulaire l'accumulation des solutés compatibles s'effectue dans le cytoplasme.



**Figure 7.** Modèle d'ajustement osmotique au niveau cellulaire lors de l'application d'une contrainte saline (300 mM NaCl pendant 10 jours). Les chiffres entre crochets indiquent la concentration en osmolytes dans les compartiments avant et après traitement, respectivement. Pro : proline,  $\beta$ -AB :  $\beta$ -alanine bêtaïne, nd : non détecté.

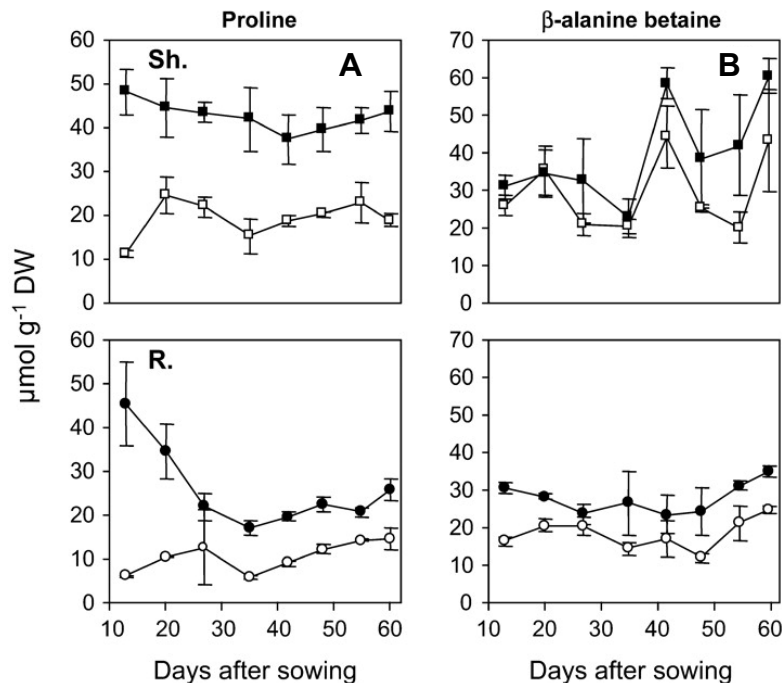
## 1.2. Dynamique d'accumulation constitutive ou osmo-induite d'osmolytes chez *L. latifolium* au cours des premiers stades de la vie végétative en présence ou non de sel

Tous les solutés détectés chez *L. latifolium* sont présents à des teneurs non négligeables dans les plantes soustraites à la contrainte saline. Une étude cinétique réalisée à partir de jeunes plantes révèle que l'apparition de ces composés (acides aminés, composés à groupement amine quaternaire, sucres solubles non structuraux, cyclitols) est plus fortement dépendante de leur stade de développement que des concentrations en sel du milieu extérieur (Fig. 8).



**Figure 8.** Analyse en Composantes Principales de l'évolution des teneurs en métabolites dans les tiges feuillées et les racines de plantules de *L. latifolium* soumises à un traitement salin (NaCl 100 mM) ou non traitées. Des plantules âgées de 6 jours ont été divisées en 2 lots et cultivées pendant 54 jours supplémentaires en présence de NaCl ou non. A : Représentation des variables sur les composantes principales 1 et 2. B : Projection des individus sur les composantes principales.

Seul le contenu en proline apparaissait véritablement halo-régulé alors que celui en  $\beta$ -AB ne semble pas significativement modulé par l'application du traitement salin (Fig. 9).



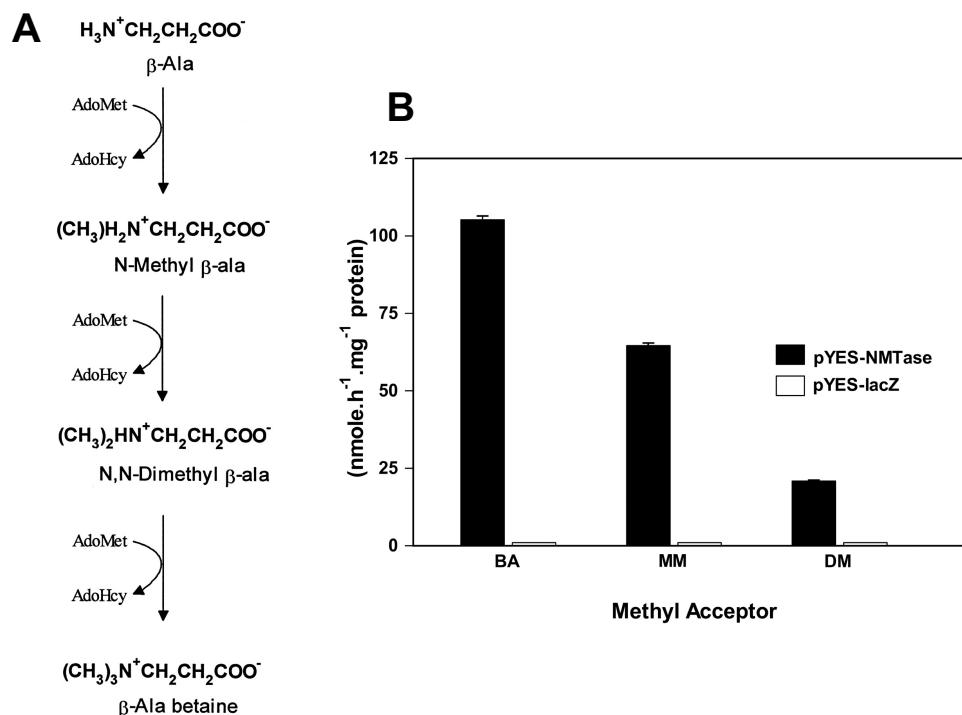
**Figure 9.** Changements des teneurs en proline (A) et en  $\beta$ -AB dans les racines (R.) et les feuilles (Sh.) de *L. latifolium* soumis (symboles noirs) ou non (symboles vides) à un traitement salin pendant 54 jours (voir Fig. 8 pour les détails de la manipulation).

La relative halophilie de cette espèce semble relever de son aptitude à coaccumuler constitutivement un nombre important de solutés de familles chimiques variées et ceci à des

concentrations suffisamment élevées pour provoquer un abaissement important du potentiel osmotique. Ce processus constitutif doit permettre à la plante de s'acclimater sans délai aux modifications soudaines du potentiel hydrique externe provoquées par la salinité. L'investissement important, tant sur le plan des ressources que sur le plan de l'énergie, pour la production de ces composés potentiellement impliqués dans la tolérance au stress est à mettre en relation avec la faible vitesse relative de croissance mise en évidence chez cette espèce. La diversité des solutés compatibles accumulés chez *L. latifolium* (choline-O-sulfate,  $\beta$ -AB, sucres solubles non structuraux, proline, *myo*-inositol, *chiro*-inositol) suggère que si leur fonction osmorégulatrice est incontestable en raison de leur teneur et de leur grande solubilité, chacun d'entre eux peut exercer par ailleurs des fonctions spécifiques d'osmoprotection contribuant elles-aussi à la tolérance au chlorure de sodium ce qu'il reste à identifier. Ces capacités spectaculaires d'accumulation de solutés organiques ajoutées aux mécanismes de relocalisation mis en évidence pour la  $\beta$ -AB, suggère l'existence chez *L. latifolium* d'adaptations métaboliques acquises associées à d'autres inductibles par le sel qui permettent de faire face, dans des conditions naturelles, aux variations rapides du potentiel osmotique externe au moindre coût énergétique avec en contrepartie une faible vitesse relative de croissance. Par ailleurs, *L. latifolium* dispose d'aptitudes métabolique et morphologique (glandes à sel) particulières pour gérer la compartimentation et la concentration interne du NaCl.

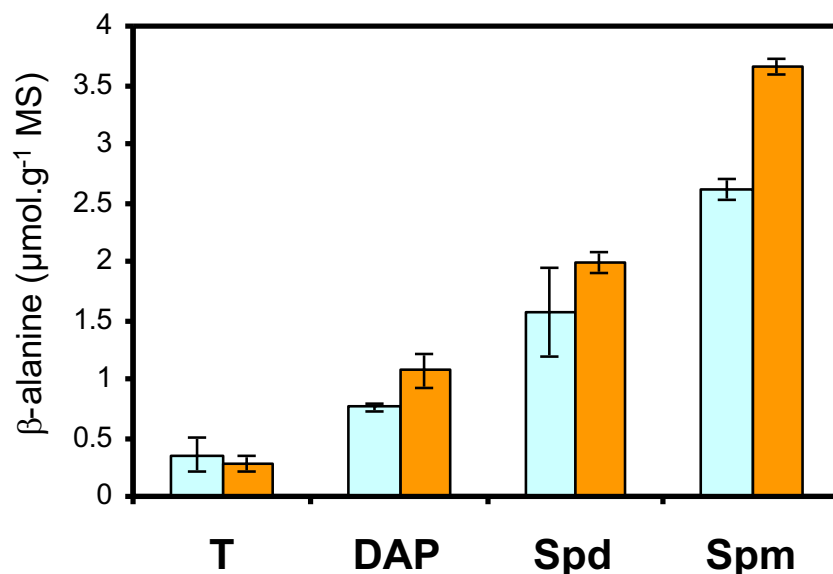
### 1.3. Implication du catabolisme oxydatif des polyamines aliphatiques dans la production de $\beta$ -alanine, précurseur privilégié de la $\beta$ -AB

La  $\beta$ -AB est issue des *N*-méthylations successives de la  $\beta$ -alanine et de ses dérivés mono- et diméthylés (Fig. 10). Ces réactions sont catalysées par une seule enzyme, la S-Adenosyl-L-méthionine-dépendant N-méthyltransférase (NMTase)<sup>13-15</sup>.



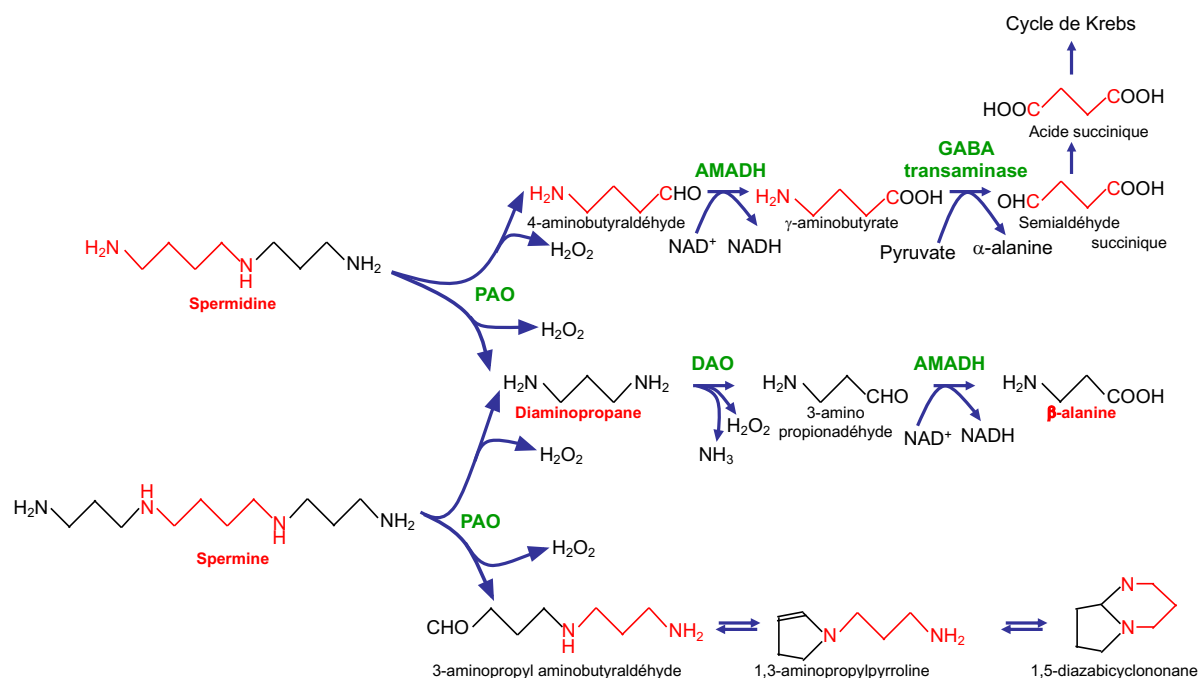
**Figure 10.** Voie de biosynthèse de la  $\beta$ -AB chez *L. Latifolium* (A) et caractérisation de l'unique enzyme responsable de ces méthylations successives, la NMTase dont le gène a été identifié chez *L. latifolium* (B). Le gène a été introduit chez *Saccharomyces cerevisiae* et les extraits protéiques obtenus ont été mis en présence de  $\beta$ -alanine (BA), de N-méthyl  $\beta$ -alanine (MM) ou de N,N-diméthyl  $\beta$ -alanine à la concentration de 10mM. En noir : levures transformées avec pYES-NMTase, en gris : levures contrôles contenant le vecteur vide pYES-lacZ. A l'issue de l'expérience, les produits méthylés ont été quantifiés par mesure de radioactivité (d'après Raman et Rathinasabapathi, 2003 et Rathinasabapathi et al., 2001)<sup>13,14</sup>.

La  $\beta$ -alanine participe également à la production du pantothénate, précurseur du coenzyme A et de l'ACP (Acyl Carrier Protein), à la production de l'homoglutathion ( $\gamma$ -glutamyl-cystéinyl- $\beta$ -alanine), homologue structural et fonctionnel du glutathion et, chez les animaux, à la formation de conjugués tels que la carnosine et l'ansérine. La  $\beta$ -alanine peut également être converti en semialdéhyde malonique lui-même pouvant être orienté vers le métabolisme des acides gras. Les voies métaboliques de synthèse de la  $\beta$ -alanine sont peu explorées chez les plantes. Les seules voies partiellement décrites conduisant à sa synthèse sont celle de l'oxydation des polyamines aliphatiques, celle de la dégradation de l'uracile et celle de la dégradation du propionate. Ces trois voies apparaissent fonctionnelles chez les espèces du genre *Limonium*. Les travaux que nous avons réalisés au laboratoire ont permis de mettre en évidence l'implication de la voie de dégradation de l'uracile chez *L. latifolium* et celle de la dégradation oxydatives des polyamines chez *L. tataricum*<sup>16,17</sup>. Ces premiers travaux m'ont incité à privilégier l'étude de l'implication de la dégradation oxydative des polyamines aliphatiques. L'approche privilégiée a été la fourniture exogène de substrats dont on veut tester l'effet sur les concentrations internes en  $\beta$ -alanine. L'alimentation en spermine, en spermidine et en 1,3-diaminopropane (DAP) de tissus foliaires isolés de *L. latifolium* se traduit en effet par une augmentation de la teneur relative en  $\beta$ -alanine, les effets les plus spectaculaires étant provoqués par la spermidine et la spermine (Fig. 11). Lorsque les précurseurs potentiels sont fournis en combinaison avec de l'aminooxyacétate, un inhibiteur de transaminase, les teneurs en  $\beta$ -alanine mesurées sont plus fortes que lorsque ceux-ci sont fournis seuls. Ceci indique qu'une partie du pool de  $\beta$ -alanine formée est transformée en semialdéhyde malonique qui peut être orienté vers le métabolisme des acides gras via le malonate et le malonyl-CoA.



**Figure 11.** Teneurs en  $\beta$ -alanine de disques foliaires de *L. latifolium* incubés 24h sur milieux de référence enrichis ou non de différents précurseurs (5 mM) et additionnés (■) ou non (□) d'aminooxyacétate (2 mM), un inhibiteur de transaminase. T : Témoin, DAP : Diaminopropane, Spd : Spermidine, Spm : Spermine.

Les amines testées peuvent donc être converties en  $\beta$ -alanine selon les voies de dégradation décrite dans la figure 12.



**Figure 12.** Dégradation oxydative de la spermidine, de la spermine et du diaminopropane conduisant potentiellement à la production de β-alanine. PAO : Polyamine oxydase, DAO : Diamine oxydase, AMADH : Aminoaldehyde déshydrogénase, GABA : γ-aminobutyrate.

La fourniture exogène de polyamines à différents supports expérimentaux (plantes entières, fragments foliaires et racinaires excisées) conduit à des désaminations oxydatives qui font intervenir des polyamine oxydases, des diamine oxydases et des déshydrogénases. Dans un tel schéma métabolique, les interventions successives de ces enzymes doivent permettre la conversion des polyamines en β-alanine et en γ-aminobutyrate (GABA). Le métabolisme des aminoaldéhydes n'a fait l'objet que de rares investigations. L'étude de l'activité aminoaldehyde déshydrogénase totale (ci-après notée AMADH) extraite de plantes de *L. latifolium* a permis de démontrer que cette enzyme permet la conversion de l'APAL (3-aminopropionaldéhyde, meilleur substrat de l'enzyme) en β-alanine et celle de l'ABAL (4-aminobutyraldéhyde) en GABA, ce qui suggère une coopération fonctionnelle de cette enzyme avec les amine oxydases (Tableau I).

**Tableau I.** Activité relative de l'AMADH en présence de différents aldéhydes dans le milieu réactionnel. Les aldéhydes sont fournis à la concentration de 20 mM. Les mesures sont réalisées à partir d'extraits protéiques partiellement purifiés obtenus de racines de *L. latifolium*.

Substrats	Structures	Activités relatives (%)
3-aminopropionaldéhyde	<chem>NCC=O</chem>	100
3-guanidinopropionaldéhyde	<chem>NC(=N)NCC=O</chem>	61
4-aminobutyraldéhyde	<chem>NCCC=O</chem>	45
4-guanidinobutyraldéhyde	<chem>NC(=N)NCCC=O</chem>	35
Triméthylaminoacétaldéhyde	<chem>CN(C)C(=O)C=O</chem>	0
Aminoacétaldéhyde	<chem>NC=O</chem>	0

De plus cette enzyme est également capable d'oxyder les dérivés guanidinés de ces aldéhydes. A l'inverse, elle semble inactive vis-à-vis de la glycine bêtaïne aldéhyde, l'aldéhyde



précurseur immédiat de la glycine bêtaïne. Le pH optimum de l'enzyme est proche de celui observé chez d'autres espèces végétales. Ce travail a permis de mettre en évidence une enzyme capable de produire de la  $\beta$ -alanine à partir d'APAL ce qui n'avait jamais été démontré chez le genre *Limonium*, l'APAL lui-même pouvant résulter de l'oxydation du DAP par les DAO ou de celle de la spermidine par des PAO. Ce système enzymatique catalyse aussi la conversion de l'ABAL en GABA. La voie métabolique privilégiée par de nombreux auteurs pour la synthèse de GABA s'intègre à une voie métabolique nommée le "shunt du GABA". Les résultats obtenus montrent que le GABA peut aussi être synthétisé à partir du ABAL dans un extrait protéique de *L. latifolium*.

En association avec les DAO et les PAO, qui sont localisées dans la paroi cellulaire, l'AMADH doit participer au catabolisme des polyamines. La localisation tissulaire de l'activité AMADH chez *L. latifolium* a été recherchée afin de savoir si une telle coopération fonctionnelle était basée ou non sur une colocalisation tissulaire. La colocalisation tissulaire de ces enzymes, démontrée par immunocytochimie, alimente l'hypothèse d'une possible coopération d'ordre physiologique entre ces enzymes, impliquées dans la remobilisation de l'azote et du carbone d'origine polyaminique et dans la production de solutés compatibles dont la participation à l'halotolérance est postulée.

L'ensemble des travaux décrits ci-dessus et d'autres travaux réalisés en collaboration pendant ma thèse ont permis la publication de 9 articles (cf liste de publications II. 1.) :

- **Métabolisme des polyamines** : articles 1, 2, 6
- **Plantes et stress salin** : articles 3, 7, 8, 9, 11
- **Insectes et stress froid** : article 5

## 2. Travaux de post-doctorat

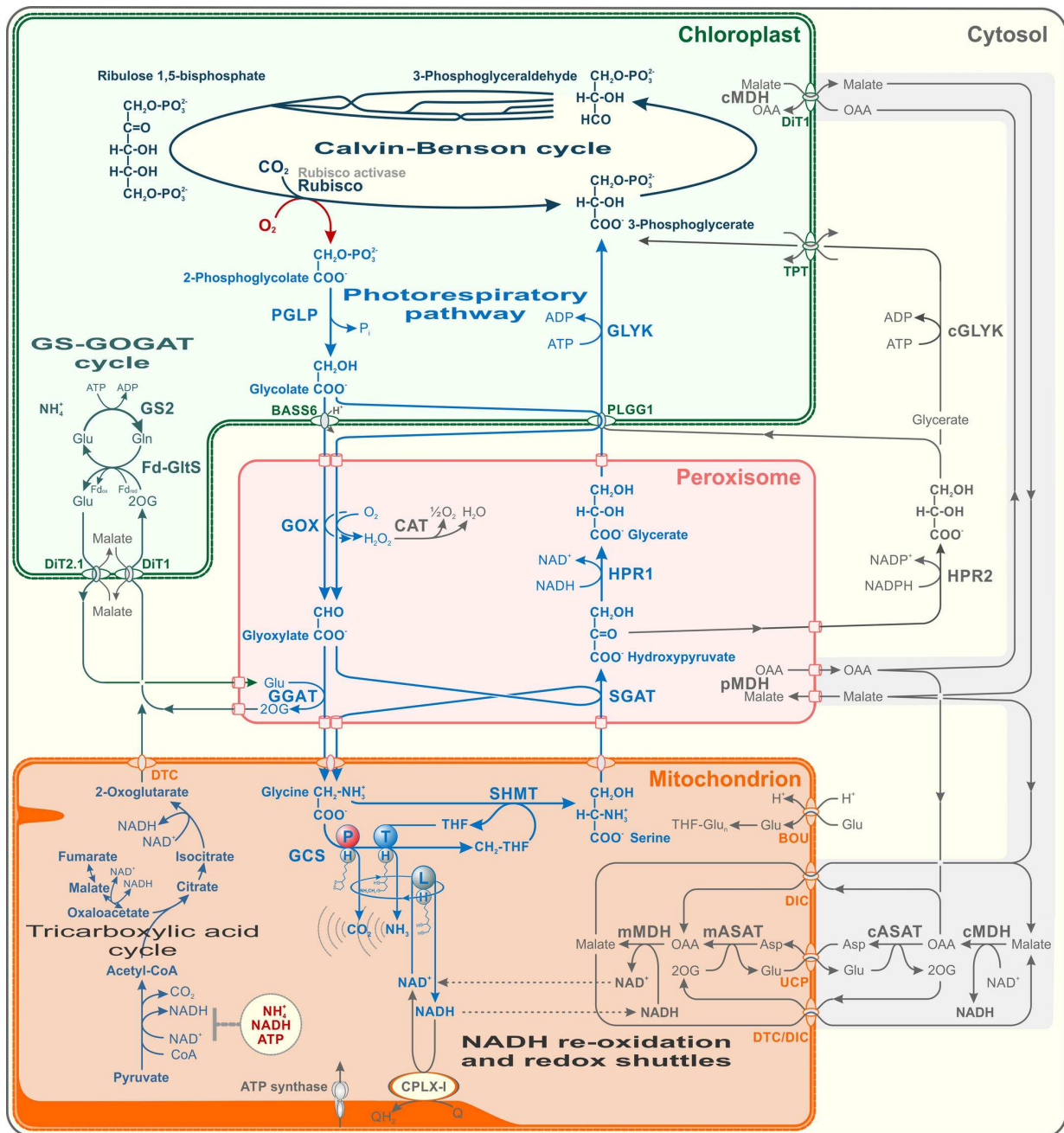
Mes 2 stages post-doctoraux ont été effectués sous la direction du Prof. Andreas Weber tout d'abord aux USA (Plant Biology Department, Michigan State University, East-Lansing, MI, USA, 2006-2007) puis en Allemagne (Plant Biochemistry Institut, Heinrich-Heine University, Düsseldorf, 2007-2009). J'ai participé, au cours de ces séjours post-doctoraux, à différents projets qui ont été finalisés parfois beaucoup plus tard.

### 2.1. Caractérisation de transporteurs impliqués dans la photorespiration

La cellule végétale est caractérisée par un très fort degré de compartimentation. En plus des organites présents dans toutes cellules eucaryotes, les cellules de plantes possèdent une vacuole, le plus grand compartiment cellulaire, et des plastes. La coordination du métabolisme repose sur un contrôle fin des flux entre ces différents compartiments. L'amélioration de la productivité des plantes reposent sur une compréhension la plus exhaustive possible des éléments moléculaires contrôlant ces flux. Les transporteurs font partie des éléments clés dont l'identification est essentielle afin de prédire le métabolisme.

La photorespiration est un métabolisme actif chez tous les organismes photosynthétiques. Chez les plantes, c'est un métabolisme hautement compartimenté impliquant les chloroplastes, les mitochondries, les peroxyosomes et le cytosol (Fig. 13). C'est un processus qui permet le recyclage du phosphoglycolate généré par l'activité oxygénase de la RubisCO (Ribulose-1,5-bisphosphate Carboxylase/Oxygenase), enzyme clé de l'activité photosynthétique. En effet, la RubisCO catalyse la carboxylation du ribulose-1,5-bisphosphate en présence de CO<sub>2</sub> pour générer 2 molécules de phosphoglycérate qui seront ensuite converties en triose-phosphate mais elle catalyse également l'oxygénation du ribulose-1,5-bisphosphate en présence d'O<sub>2</sub> ce qui génère une molécule de phosphoglycérate et une molécule de phosphoglycolate. L'intensité relative des réactions de carboxylation et d'oxygénation dépend du rapport CO<sub>2</sub>/O<sub>2</sub>. D'un point de vue évolutif, la RubisCO est apparu dans une atmosphère où le ratio CO<sub>2</sub>/O<sub>2</sub> était largement en faveur du CO<sub>2</sub>. Ceci explique pourquoi cette activité n'a pas été contre-sélectionnée. Dans notre atmosphère actuelle, cette activité est non négligeable et la photorespiration permet le recyclage du carbone perdu dans ces conditions et empêche l'accumulation d'intermédiaires toxiques dans la cellule végétale. Si pendant longtemps, on attribuait uniquement un rôle de recyclage à ce métabolisme, on sait aujourd'hui qu'il est indispensable pour le métabolisme des plantes. Les enzymes impliquées dans les différentes étapes de cette voie métabolique ont été identifiées et caractérisées (Fig. 13). Cependant, peu d'informations étaient disponibles au moment de ces travaux en ce qui concerne les transporteurs. Seul un transporteur, DiT2.1/DCT1 (Dicarboxylate Transporter) impliqué dans l'export de glutamate du chloroplaste avait été caractérisé par une approche de génétique classique<sup>18,19</sup>. Le glutamate en provenance du chloroplaste est ensuite importé dans le peroxyosome et est utilisé comme donneur de groupement amine par la glyoxylate : glutamate amino transférase (GGAT).

J'ai au cours de ce post-doctorat contribué à l'identification et à la caractérisation de 3 transporteurs mitochondriaux impliqués dans la photorespiration chez *Arabidopsis thaliana* : BOU (à bout de souffle), UCP1 et UCP2 (Uncoupling protein 1 and 2). Comme indiqué par la suite, l'implication de UCP1 et UCP2 dans la photorespiration n'a été démontré que tardivement et leur rôle hypothétique initial était une implication dans l'importation de pyruvate dans la mitochondrie, hypothèse réfutée par la suite.



**Figure 13.** La photorespiration (d'après Fernie et Bauwe, 2020<sup>20</sup>). En bleu, la voie métabolique principale de la photorespiration impliquant la participation des chloroplastes, des peroxysomes et des mitochondries. En noir, des voies métaboliques cytosoliques pouvant être impliquées dans la photorespiration. PGLP : 2-phospho glycolate phosphatase, GOX : glycolate oxydase, CAT : catalase, GGAT : glutamate:glyoxylate aminotransferase, SGAT : serine:glyoxylate aminotransferase, GCS : glycine cleavage system, SHMT : serine hydroxymethyltransferase, HPR1(2) : hydroxypyruvate reductase 1 (2), GLYK : glycerate kinase, cGLYK : GLYK cytosolique, MDH : malate déhydrogenase, ASAT : aspartate aminotransferase, THF : tetrahydrofolate, TPT : triose phosphate translocator, DIT : Dicarboxylate transporter, OAA : acide oxaloacétique, PLGG1 : glycolate/glycerate antiporteur, BASS6 : glycolate exporter, GS : glutamine synthetase, FdGltS : Glutamate synthase, DTC : dicarboxylate/tricarboxylate carrier.

#### a. BOU, un transporteur mitochondrial de glutamate

Afin d'identifier des gènes codant des transporteurs impliqués dans la photorespiration, nous avons adopté une approche gène candidat. Les données transcriptomiques disponibles publiquement (CSB.DB, comprehensive system-biology database) ont été utilisées pour ce criblage. Comme preuve de concept, tous les gènes codant des enzymes de la photorespiration ont été examinés. Sur les 20 gènes connus à la date de cette analyse, tous

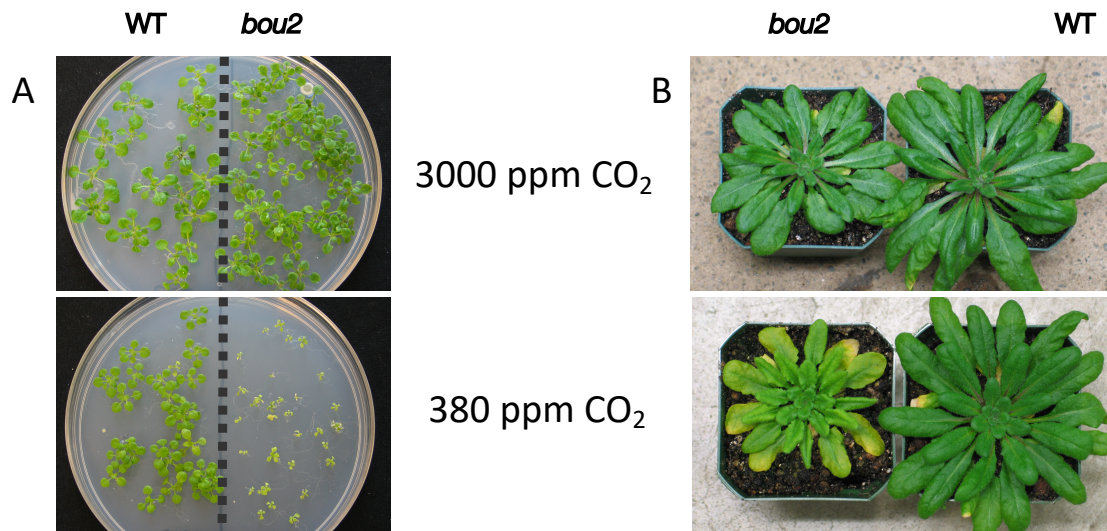
sont fortement co-régulés et forme un cluster robuste excepté les gènes codant une glycérate kinase (GLYK), une glutamate synthase (GLS1/GLU1), une catalase (CAT2) et les sous unités H2 et L1 de la glycine décarboxylase (notée GCS dans la Fig. 13) (Tableau II).

**Tableau II.** Gènes codant des enzymes impliquées dans la photorespiration. Les gènes indiqués en noir sont fortement co-régulés. Les gènes indiqués en bleu sont non co-régulés avec le reste des gènes et ne font pas parti du cluster fort de co-régulation identifié.

AGI	Gene Description	Localization	Reference
At5g36700	2-Phosphoglycolate Phosphatase (PGLP1)	Chloroplast	Schwarte & Bauwe (2007)
At1g80380	D-Glycerate 3-kinase (GLYK)	Chloroplast	Boldt et al. (2005)
At5g35630	Glutamine synthetase (GS2/GLN2)	Chloroplast	Somerville et al. (1980)
At5g04140	Fd-dependent glutamate synthase (GLS1/GLU1)	Chloroplast	Coschigano et al. (1998)
At5g12860	2-Oxoglutarate:malate antiporter (DiT1/OMT1)	Chloroplast	Tanigushi et al. (2002)
At5g64290	Glutamate:malate antiporter (DiT2.1/DCT1)	Chloroplast	Renne et al. (2003)
At3g14420	Glycolate oxidase (GOX2)	Peroxisomes	Fahnenstich et al. (2008)
At1g23310	Glu:glyoxylate aminotransferase (AOAT1/GGAT)	Peroxisomes	Igarashi et al. (2003)
At2g13360	Ala/Ser:glyoxylate aminotransferase (AGT1/SGT1)	Peroxisomes	Liepman et al. (2001)
At4g35090	Catalase (CAT2)	Peroxisomes	Queval et al. (2007)
At1g68010	Hydroxypyruvate reductase (HPR)	Peroxisomes	Murray et al. (1989)
At2g35370	Glycine decarboxylase H protein (GDH1)	Mitochondria	Bauwe et al. (2003)
At2g35120	Glycine decarboxylase H protein (GDH2)	Mitochondria	Bauwe et al. (2003)
At1g32470	Glycine decarboxylase H protein (GDH3)	Mitochondria	Bauwe et al. (2003)
At3g17240	Glycine decarboxylase L protein (LPD1)	Mitochondria	Bauwe et al. (2003)
At1g48030	Glycine decarboxylase L protein (LPD2)	Mitochondria	Bauwe et al. (2003)
At4g33010	Glycine decarboxylase P protein (GDP1)	Mitochondria	Engel et al. (2007)
At2g26080	Glycine decarboxylase P protein (GDP2)	Mitochondria	Engel et al. (2007)
At1g11860	Glycine decarboxylase T protein (GDT1)	Mitochondria	Bauwe et al. (2003)
At4g37930	Serine hydroxymethyltransferase (SHMT1)	Mitochondria	Moreno et al. (2005)

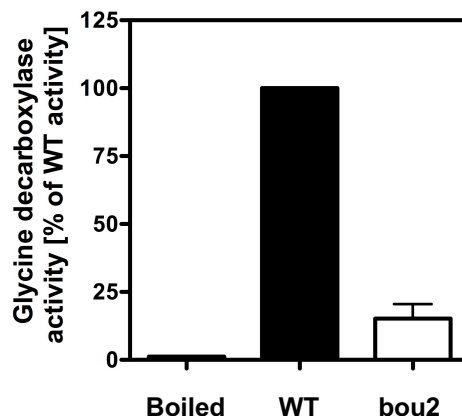
La base CSB.DB a été utilisé pour identifier des gènes supplémentaires co-régulés avec ce cluster fort. Les meilleurs candidats ont été triés de manière à ne conserver que les protéines possédant au moins 2 domaines hydrophobes et donc potentiellement intégrés dans les bicouches phospholipidiques (régions transmembranaires de type hélice  $\alpha$  prédites) en utilisant la base de données ARAMEMNON (plant membrane protein database<sup>21</sup>). Ces protéines représentaient des transporteurs candidats. Cette analyse a permis de sélectionner 32 candidats. Mon travail a consisté en la caractérisation du gène At5g46800 (AtBOU). Ce gène était annoté comme « putative mitochondrial carrier protein ». La protéine possède 3 à 6 domaines transmembranaires prédits et appartient à la famille MCF (mitochondrial carrier family). La protéine BOU avait été initialement décrite, par homologie à une protéine humaine, comme un transporteur à carnitine/acylcarnitine permettant la fourniture d'acétate à la mitochondrie<sup>22</sup>. En conditions de jours longs, les plantules, dont le gène est muté, ne sont pas capables de se développer de manière autotrophique et ne peuvent produire de vraies feuilles. L'ajout de sucres dans le milieu permet la complémentation de ce phénotype. Nous avons isolé un mutant (*bou2*, GABI Kat line 079D12) et analysé son phénotype en conditions

photorespiratoires (teneur en CO<sub>2</sub> de l'atmosphère, 380 ppm CO<sub>2</sub>) et non photorespiratoire (atmosphère enrichie en CO<sub>2</sub>, 3000 ppm CO<sub>2</sub>). Le mutant *bou2* présente un phénotype typique des mutants de gènes impliqués dans la photorespiration. En effet, en conditions photorespiratoires, les plantules KO âgées de 2 semaines sont pâles et significativement plus petites que les plantes sauvages (Fig. 14A). Lorsque les plantes KO sont cultivées avec de fortes concentrations en CO<sub>2</sub>, elles se développent normalement (Fig. 14A). Ce phénotype est également visible sur des plantes plus âgées. Lorsque les plantules sont cultivées en conditions non photorespiratoires pendant 6 semaines puis transférées pendant 1 semaine en conditions photorespiratoires, les rosettes du mutant *bou2* deviennent chlorotiques et les feuilles les plus vieilles sénescencent (Fig. 14B).



**Figure 14.** Caractérisation phénotypique du mutant *bou2*. A : Les plantules ont été mises à germer soit en conditions photorespiratoires (air ambiant, 380 ppm) soit en conditions non photorespiratoires (3000 ppm). B : Les plantules ont été cultivées pendant 6 semaines en conditions de fortes concentrations en CO<sub>2</sub> (3000 ppm) puis transférées pendant 1 semaine sous air ambiant (380 ppm) ou maintenues dans les conditions initiales.

L'analyse métabolique du mutant *bou2* a révélé une altération du métabolisme de la glycine, de la sérine, du glutamate, de l'oxoglutarate, du glyoxylate et du glycolate ce qui étayait l'hypothèse d'un rôle de BOU dans la photorespiration. L'implication de BOU dans le transport de la glycine et de la sérine a tout d'abord été suspecté. Des expériences ont été réalisées en ce sens mais aucune activité de transport impliquant ces molécules n'a été mise en évidence. Des expériences supplémentaires ont montré que l'activité de la glycine décarboxylase était fortement réduite dans le mutant *bou2* par rapport aux plantes sauvages. Celle-ci représentait 15.2% de celle mesurée dans les sauvages (Fig. 15).



**Figure 15.** Caractérisation de l'activité GDC dans des mitochondries isolées à partir de plantes sauvages ou du mutant *bou2*. Les mitochondries ont été isolées à partir de plantes âgées de 4 semaines et perméabilisées avec du triton X-100. Une fraction aliquote est ensuite incubée en présence de 0.2 mM de glycine marqué au C<sup>14</sup> pendant 15 min et la libération de <sup>14</sup>CO<sub>2</sub> est mesurée.

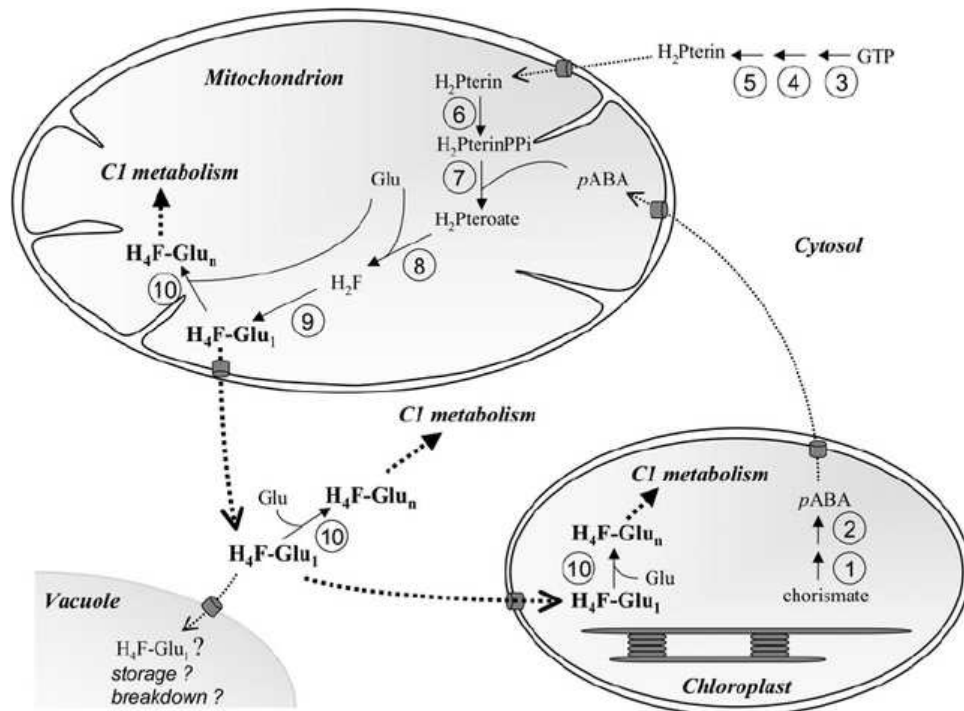
En complément de ces mesures d'activités, une analyse du complexe GDC par western blot a été réalisée et a révélé une altération de celui-ci. Ces résultats ont conduit à émettre l'hypothèse que l'un des cofacteurs de la GDC pouvait être limitant et entraîné un défaut de la complexation des sous-unités (Tableau III). Ces résultats ont été publiés en l'état sans connaître véritablement les substrats de ce transporteur.

**Tableau III.** Liste des cofacteurs nécessaires au fonctionnement de la GDC, lieux de leur biosynthèse et listes des précurseurs nécessaires lorsque leur synthèse a lieu dans la mitochondrie.

<b>GDC subunit</b>	<b>Cofactor</b>	<b>Final step cofactor biosynthesis</b>	<b>Predicted transport processes across mitochondrial membrane</b>
P protein	PLP	Cytosol	PLP Pyridoxine Pyridoxal
H protein	Lipoic acid	Mitochondria	Malonic acid Pyruvate "Sulfur" donor (cysteine, SAM) CoenzymeA (Valine, ketopantoate, CoA)
T protein	THF	Mitochondria	H <sub>2</sub> Pterin, pABA, glutamate
L protein	FAD	Cytosol	FAD Quinolate
L protein	NAD	Cytosol / Mitochondria	NaMN NAD <sup>+</sup>

Il avait déjà été montré que la diminution de la biosynthèse de l'acide lipoïque ainsi qu'une perturbation du recyclage du tetrahydrofolate (THF) entraînaient une diminution de l'activité GDC<sup>23,24</sup>.

Une publication récente (2018) a démontré que BOU était impliqué dans l'importation du glutamate dans la mitochondrie<sup>25</sup>. Il s'agit d'un antiport glutamate/H<sup>+</sup>. Le glutamate est nécessaire dans la mitochondrie pour la synthèse du THF de même que la dihydroptérine (H<sub>2</sub>Pterin) et l'acide p-aminobenzoïque (pABA)<sup>26</sup>. Le glutamate est utilisé par la dihydrofolate réductase (DHFR) en présence de dihydrofolate (H<sub>2</sub>F) pour former le THF-glutamate et pour la polyglutamylation du THF-glutamate en THF-glutamate<sub>n</sub>, une réaction catalysée par la folyl-polyglutamate synthétase (FPGS). Le THF-glutamate<sub>n</sub> est la molécule utilisée pour le métabolisme des composés à 1C dont le THF nécessaire à la GDC et à la sérine hydroxyméthyl transférase (SHMT, Fig. 16). Ceci explique le phénotype observé lors de l'analyse du mutant *bou2* et explique les résultats rapportés dans notre publication de 2013<sup>27</sup>.



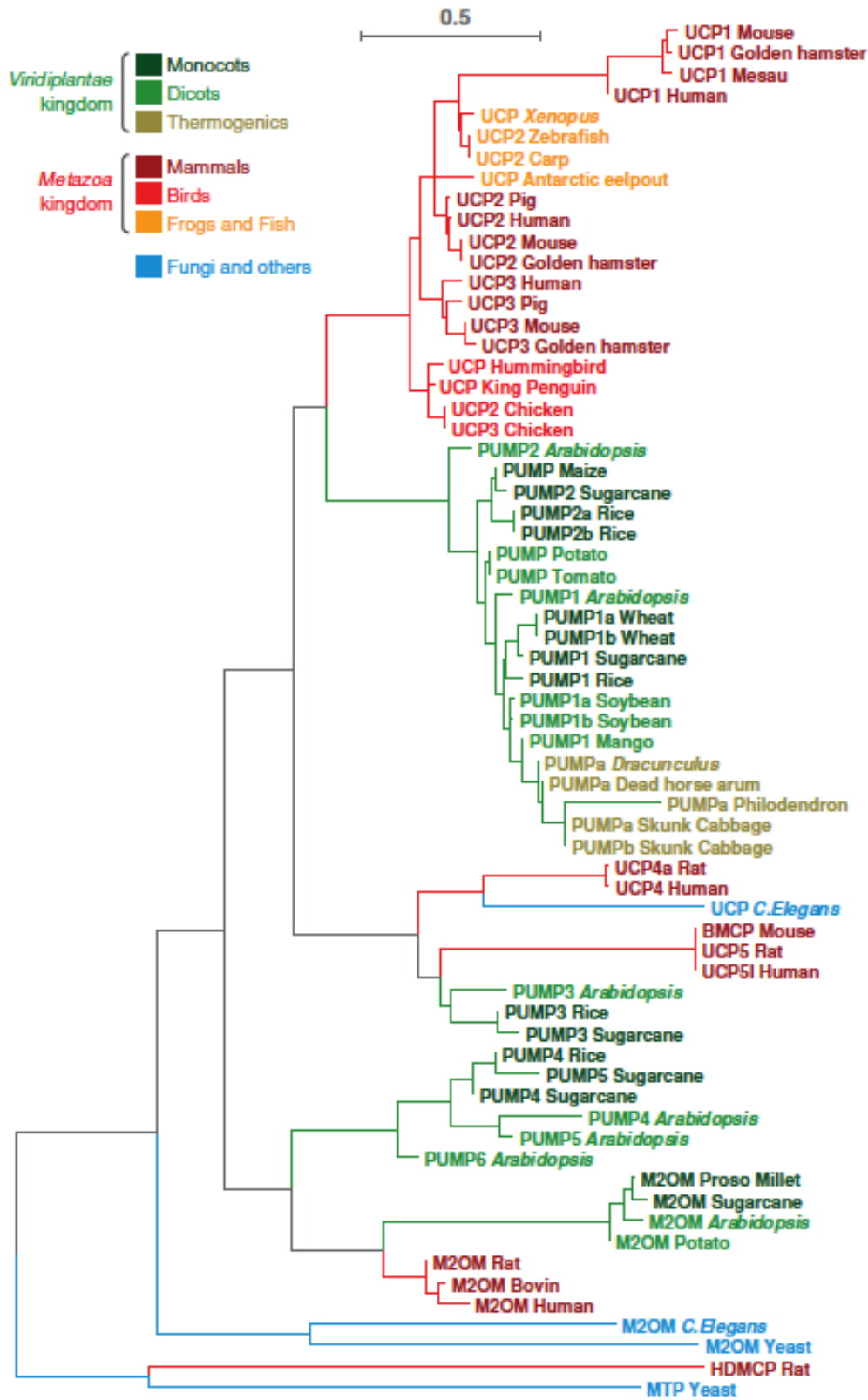
**Figure 16.** Métabolisme du THF chez les plantes et sa compartimentation (d'après Sahr et al. 2005<sup>26</sup>). Le THF utilisé par la GDC est synthétisé dans la mitochondrie en présence de dihydroptérine, d'acide p-aminobenzoïque et de glutamate. BOU assure l'importation de glutamate dans la mitochondrie. 1 : aminodeoxychorismate synthase (ADC synthase), 2 : aminodeoxychorismate lyase (ADC lyase), 3 : GTP-cyclohydrolase I (GTPCHI), 4 : nudix hydrolase, 5 : dihydroneopterin aldolase (DHNA), 6 : hydroxymethyldihydropterin pyrophosphokinase (HPPK), 7 : dihydropteroate synthase (DHPS), 8 : dihydrofolate synthetase (DHFS), 9 : dihydrofolate reductase (DHFR), 10 : folylpolyglutamate synthetase (FPGS).

Ce projet a été initié avant mon arrivée dans l'équipe du Pr. Andreas Weber (criblage des gènes candidats). J'ai effectué la caractérisation physiologique du mutant bou2 (profilage métabolique, mesure de l'activité GDC, analyse de la GDC par Western blot). Ce projet a fait l'objet de 2 publications. L'une portait sur l'analyse du mutant et l'autre sur la mise au point d'une méthode de dosage haut-débit du glyoxylate et de l'ammonium, des marqueurs de la dérégulation de la photorespiration, chez *A. thaliana* (publications n°4 et 12).

#### b. UCP1 et UCP2 des transporteurs mitochondriaux d'aspartate, de glutamate et d'acides dicarboxyliques

L'importation du pyruvate dans la mitochondrie est essentielle pour l'oxydation complète du glucose qui requiert un couplage entre la glycolyse cytosolique et le cycle de Krebs mitochondrial mais également pour la lipogénèse, la glycogénèse et la synthèse d'acides aminés tels que la valine, la leucine et l'isoleucine. Chez les plantes, le cycle de Krebs joue un rôle très important pour l'oxydation des substrats respiratoires et la fourniture d'énergie et de pouvoir réducteur mais également dans de nombreuses voies métaboliques en fournissant des substrats. En effet, chez les plantes, à la lumière, le cycle de Krebs aurait plus un rôle anabolique que catabolique et ne serait que très peu cyclique<sup>28</sup>. Le transport de pyruvate dans la mitochondrie apparaît donc primordial pour le métabolisme de la cellule végétale. Malgré ce rôle majeur, le transporteur impliqué n'a jamais été identifié sur le plan moléculaire et ceci chez tous les organismes. Pour identifier ce transporteur, en utilisant *A. thaliana* comme plante modèle, nous avons adopté la même démarche que pour l'identification des gènes impliqués dans la photorespiration. Notre hypothèse était que l'expression du gène codant ce transporteur devait être corrélée avec celles des gènes codant les enzymes du cycle de Krebs et les acteurs de la chaîne respiratoire. A l'issue de ce crible, les candidats ont été triés sur les mêmes critères que précédemment c'est à dire la présence de domaine transmembranaire

dans la protéine prédite. A l'issue de ce criblage, 2 gènes ont été retenus : At3g54110 et At5g58970.

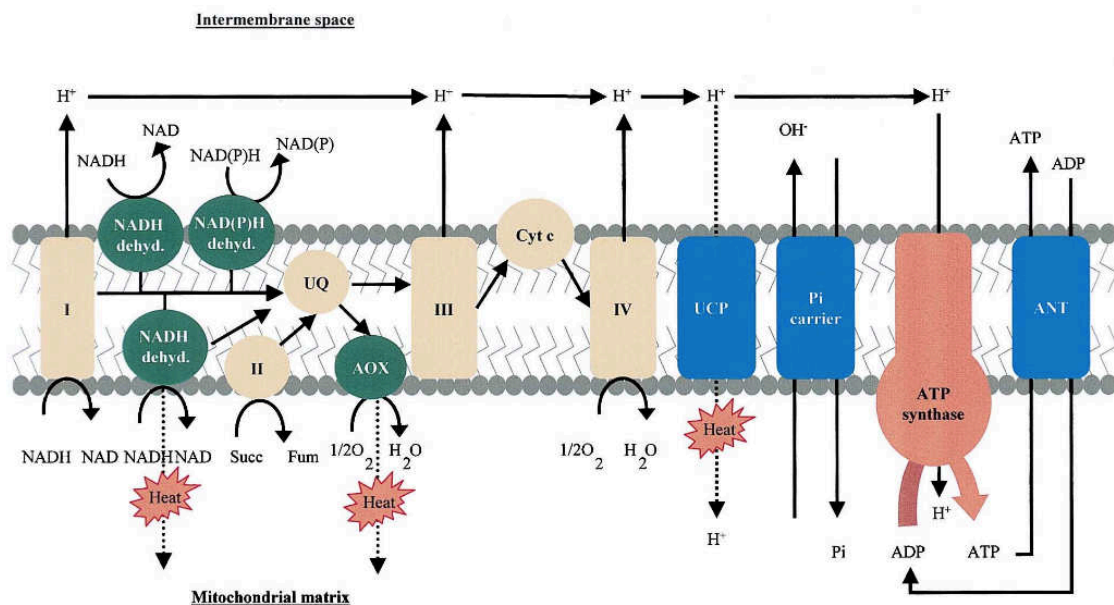


**Figure 17.** Arbre phylogénétique des protéines annotées comme protéines découplantes (d'après Vercesi et al., 2006<sup>29</sup>).



Ces 2 gènes codent 2 protéines appartenant à la famille MCF, la famille de transporteurs mitochondriaux à laquelle appartient BOU. Ces protéines sont appelées des UCPs (Uncoupling proteins ; protéines découplantes). Il s'agit de AtUCP1 et AtUCP2 connus également sous les noms de AtPUMP1 et AtPUMP2, respectivement. Il existe 6 UCPs chez Arabidopsis. Ces gènes ont été annotés par homologie avec les séquences protéiques de mammifères (Fig. 17).

Il existe, par exemple, 6 UCPs chez l'homme (hUCP1-6). Chez les mammifères, UCP1 a été décrite comme étant une pompe à protons de la membrane interne de mitochondrie qui permet le découplage de la phosphorylation oxydative<sup>30,31</sup>. Cette protéine catalyse la dissipation du gradient de protons à travers la membrane et court-circuite l'ATP synthétase. Son activité permet dans certaines conditions physiologiques d'accélérer la respiration qui, n'étant pas couplée à la production d'ATP, devient un processus thermogénique (Fig. 18).

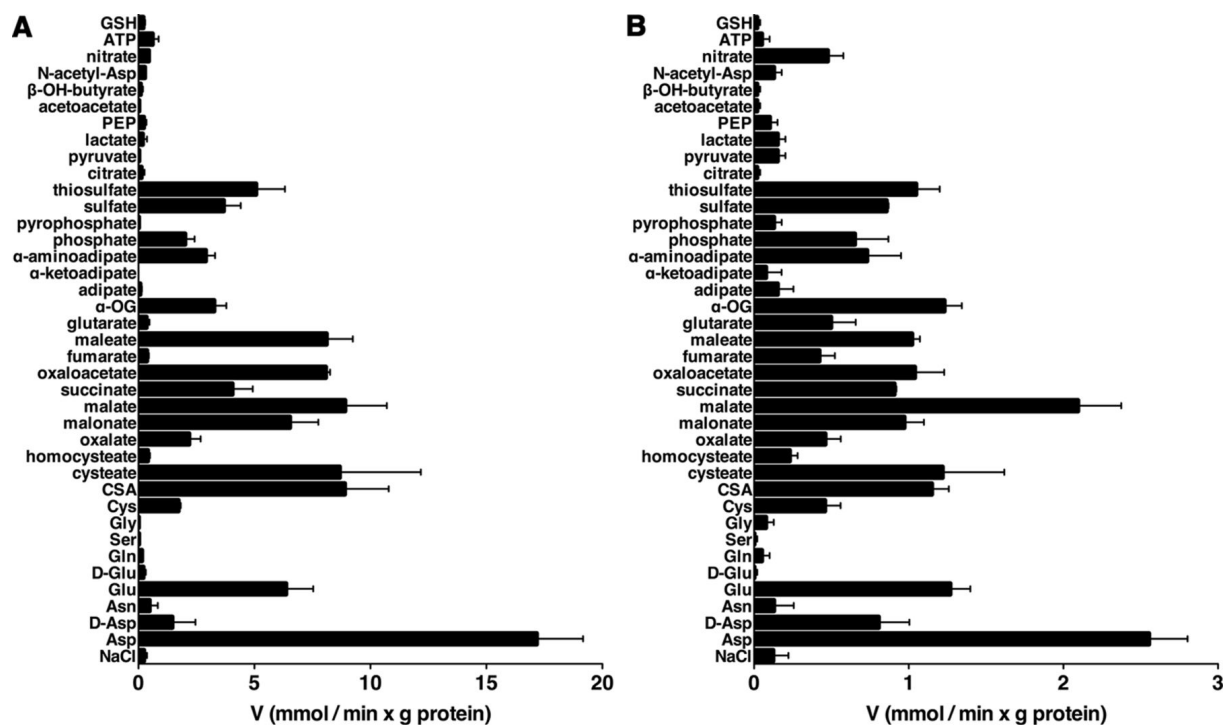


**Figure 18.** Représentation schématique de la membrane interne de la mitochondrie. La protéine UCP permet de dissiper le gradient de protons sous la forme de chaleur au détriment de la phosphorylation de l'ADP par l'ATP synthétase (d'après Laloi, 1999<sup>32</sup>).

Cependant, la fonction de AtUCP4-6 (PUMP4-6) a été identifiée plus récemment. Il s'agit de transporteurs d'acides dicarboxyliques (malate, oxaloacétate, succinate, Pi, sulfate, thiosulfate et sulfite) annotés par conséquent DIC1-3 (Dicarboxylate carrier)<sup>33</sup>. De même, hUCP2 est un transporteur de molécules à 4 carbones (aspartate, malate, malonate, oxaloacétate, Pi et sulfate)<sup>34</sup>.

Ces éléments constituaient des éléments en faveur de l'implication de AtUCP1 et AtUCP2 dans le transport du pyruvate au travers de la membrane interne de la mitochondrie. C'est cette hypothèse que nous avons testée. Le rôle biochimique et physiologique de ces 2 gènes a été étudié *in vitro* et *in planta*. L'activité de transport des protéines recombinantes a été analysé *in vitro* et la fonction physiologique de ces protéines a été élucidée en utilisant des simples mutants et un double mutant de ces gènes.

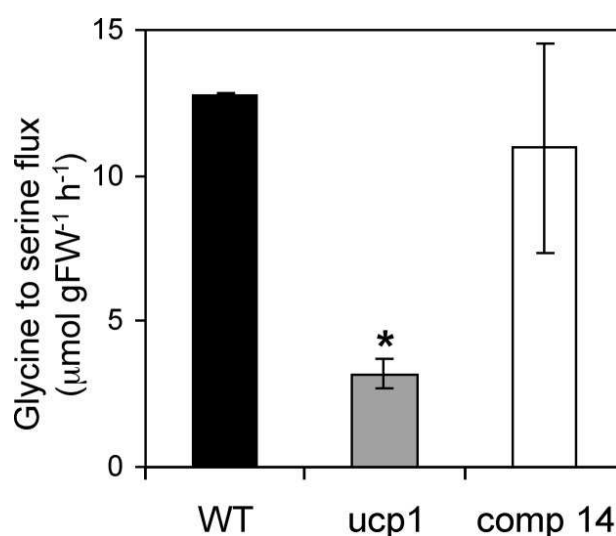
Les 2 protéines ont été produites chez *Escherichia coli* puis après purification reconstituées dans des liposomes constitués de phospholipides. Les 2 protéines sont capables de catalyser le transport de plusieurs substrats dont des acides aminés (l'aspartate et le glutamate), des acides dicarboxyliques (le malate et l'oxaloacétate et le 2-oxoglutarate), le sulfate, le phosphate et le thiosulfate mais pas le pyruvate ce qui invalidait notre hypothèse de départ (Fig. 19).



**Figure 19.** Spécificité de substrats de AtUCP1 et AtUCP2. Les protéoliposomes ont été chargés avec différents substrats (10 mM) au cours de leur préparation puis mis en présence d'une solution d'aspartate marqué au  $^{14}\text{C}$  (0.8 mM), celui-ci étant le meilleur substrat de AtUCP1 et AtUCP2. A l'issue de l'incubation, les liposomes sont rincés puis la radioactivité est mesurée. CSA : cysteinesulfonic acid, PEP : phosphoenolpyruvate,  $\alpha$ -OG :  $\alpha$ -ketoglutarate (d'après Monné et al. 2017<sup>35</sup>).

Le profilage métabolique des simples mutants et du double mutant a montré une perturbation des teneurs en acides aminés et en acides organiques. Ceci corrobore l'implication de AtUCP1 et AtUCP2 dans le métabolisme de ces molécules.

Une analyse du mutant *ucp1* publiée en 2006 a permis de confirmer et de préciser le rôle potentiel de ces gènes<sup>36</sup>. En effet, il a été montré que, chez ce mutant, l'intensité de la photorespiration était plus faible que chez les plantes sauvages ce qui entraînait, en conséquence, une diminution de l'activité photosynthétique (Fig. 20).



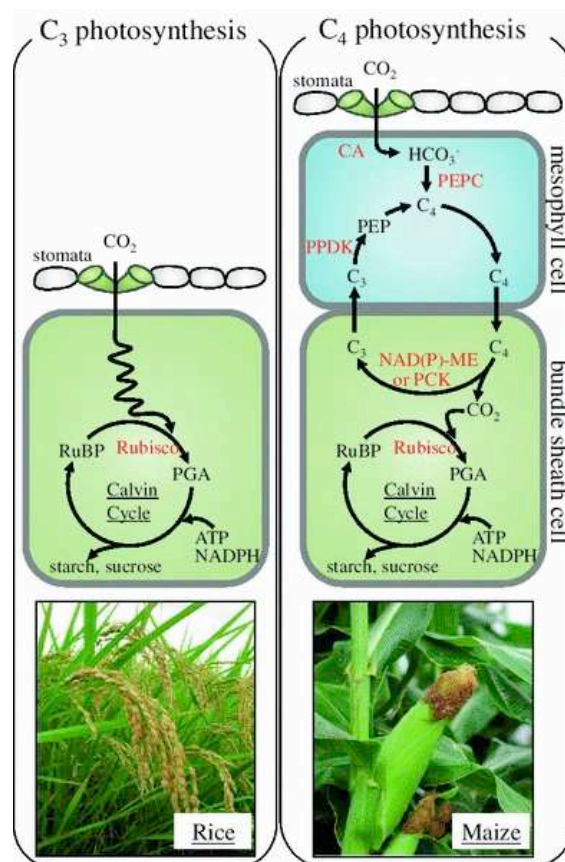
**Figure 20.** Intensité de la photorespiration mesurée *in vivo* en analysant le taux de conversion de la glycine en sérine. WT : arabidopsis sauvage ; *ucp1* : mutant *ucp1* ; comp 14 : mutant *ucp1* complémenté avec le gène AtUCP1.

Il est postulé que UCP1 et UCP2 serait des transporteurs catalysant le transport antiport d'aspartate et de malate afin d'exporter du pouvoir réducteur de la mitochondrie vers le cytosol. En effet, la GDC réduit dans la mitochondrie du  $\text{NAD}^+$  en NADH et l'hydroxypyruvate réductase utilise du NADH pour convertir l'hydroxypyruvate en glycérate dans le peroxysome. Ceci implique un transfert du pouvoir réducteur de la mitochondrie vers le peroxysome. UCP1 et UCP2 pourrait être responsable du transport de ce pouvoir réducteur vers le cytosol (Fig. 13). Ce rôle permettrait d'expliquer le phénotype observé chez le mutant *ucp1*.

La caractérisation fonctionnelle des UCPs d'*Arabidopsis* a fait l'objet d'une publication (publication n° 16).

## 2.2. Caractérisation des éléments moléculaires à l'origine de la transition d'une photosynthèse de type C3 vers un type C4

La photosynthèse de type C4 est une caractéristique biologique complexe qui permet aux plantes une acquisition plus rapide de biomasse mais également de vivre dans des conditions plus contraignantes que les plantes avec une photosynthèse de type C3. Le principe de la photosynthèse de type C4 expliquant ces avantages reposent sur la séparation spatiale de la fixation du  $\text{CO}_2$  et de son assimilation par la RubisCO. Ce système permet un enrichissement de la concentration en  $\text{CO}_2$  au niveau de la RubisCO et donc une diminution de son activité oxygénase au profit de son activité carboxylase. La génération du phosphoglycolate est ainsi réduite et, par conséquent, la perte d'énergie, de carbone et d'azote au cours de la photorespiration. La phosphoenolpyruvate carboxylase (PEPC) catalyse la fixation primaire du  $\text{CO}_2$  dans les cellules du mésophylle pour former une molécule à 4C (d'où le terme de photosynthèse de type C4) qui sera transportée vers les cellules de la gaine où le  $\text{CO}_2$  sera libéré au niveau de la RubisCO (Fig. 21).

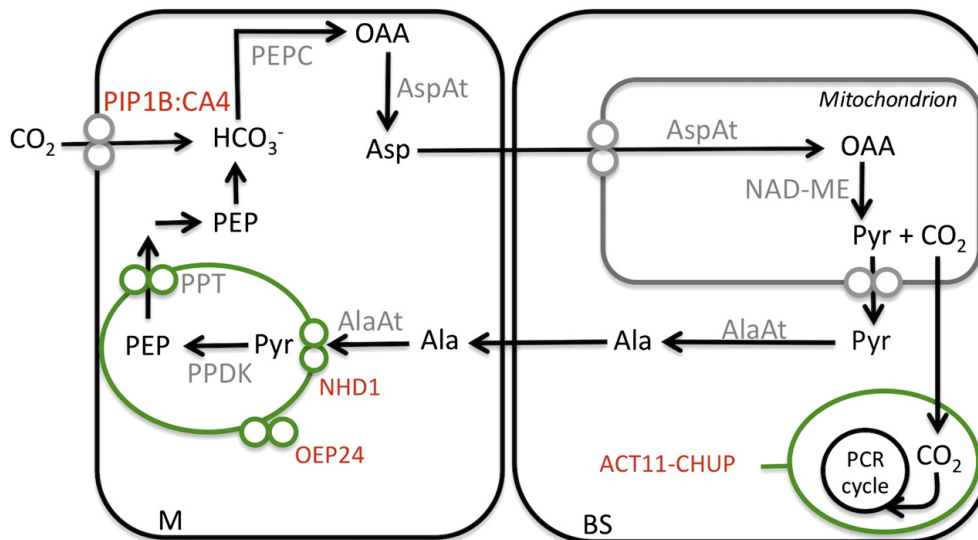


**Figure 21.** Comparaison du principe de la photosynthèse dite de type C3 et de type C4 (d'après Yamori et al., 2014<sup>37</sup>). PGA : phosphoglyceric acid, CA : carbonic anhydrase, PEPC : phosphoenolpyruvate carboxylase, NAD(P)-ME : NAD(P) malic enzyme, PCK : phosphoenolpyruvate carboxykinase, PPDK : pyruvate phosphate dikinase.

Cette séparation tissulaire est présente chez la plupart des plantes mais chez certaines espèces, la photosynthèse de type C4 a lieu au sein d'une même cellule et ces plantes ont été qualifiées de « single cell C4 plants »<sup>38</sup>.

Trois types de photosynthèse de type C4 sont décrits : à enzyme malique à NAD (NAD-ME), à enzyme malique à NADP (NADP-ME) et à phosphoenolpyruvate carboxykinase (PCK) (Fig. 21). Si les étapes enzymatiques de ce métabolisme sont bien décrites, il reste à ce jour beaucoup de zones d'ombre. En particulier, l'intérêt de la description approfondie de ce métabolisme repose sur le fait que ce métabolisme est apparu au moins 45 fois au cours de l'évolution et ceci représente un cas remarquable d'évolution convergente<sup>39</sup>. De plus, afin de répondre aux besoins croissants en produits végétaux (augmentation de la démographie, diminution des surfaces cultivables), la conversion des plantes cultivées, pour la plupart de type C3, en plantes de type C4 représente une alternative intéressante. Ceci repose sur une compréhension complète des déterminants moléculaires participant à ce métabolisme (biologie cellulaire, développement des feuilles, biochimie). Dans ce projet, nous avons choisi de comparer les transcriptomes de *Cleome spinosa* (photosynthèse de type C3) et *Cleome gynandra* (photosynthèse de type C4 à NAD-ME), 2 espèces du même genre et proches sur le plan phylogénétique d'*A. thaliana*. De plus, le génome de ces 2 espèces est séquencé et annoté ce qui facilite l'interprétation des données de séquençage haut-débit. En comparant les transcriptomes de ces 2 espèces, l'objectif était d'établir une liste exhaustive des facteurs moléculaires nécessaires à la mise en place de la photosynthèse de type C4.

Ce travail a permis de révéler un ensemble de facteurs déjà décrits dans la littérature mais également de découvrir de nouveaux déterminants moléculaires discriminants dont 4 exemples sont présentés ci-dessous (Fig. 22).



**Figure 22.** Représentation schématique des éléments participant à la photosynthèse de type C4 (NAD-ME type) basée sur l'analyse des données transcriptomiques (d'après Bräutigam et al., 2011<sup>40</sup>). Les protéines décrites précédemment sont indiquées en gris et les protéines découvertes dans cette étude en rouge. PIP1B:CA4 : PIP1B plasma membrane aquaporin:membrane-tethered carbonic anhydrase, OAA : oxaloacetic acid, PEPC : phosphoenolpyruvate carboxylase, AspAT : Aspartate aminotransferase, Asp : aspartate, AlaAT : Alanine aminotransferase, PEP : phosphoenolpyruvate, PPDK : pyruvate phosphate dikinase, PPT : PEP phosphate translocator, NAD-ME : NAD malic enzyme, Pyr : pyruvate, Ala : alanine, OEP24 : chloroplast outer membrane protein 24, ACT11-CHUP : ACTIN11-CHUP1 complex, NHD1 : putative plastidic proton:sodium symporter.

Les gènes codant les enzymes spécifiques au métabolisme de type C4 sont surexprimés comme attendu : PEPC, AspAt, NAD-ME, AlaAt, PPDK. PPT, le transporteur permettant la sortie du PEP du chloroplaste des cellules du mésophylle est également surexprimé chez *C. gynandra* (20 fois). L'abondance des transcrits de gènes impliqués dans d'autres métabolismes est également différente entre les 2 types de plantes. Par exemple, les gènes

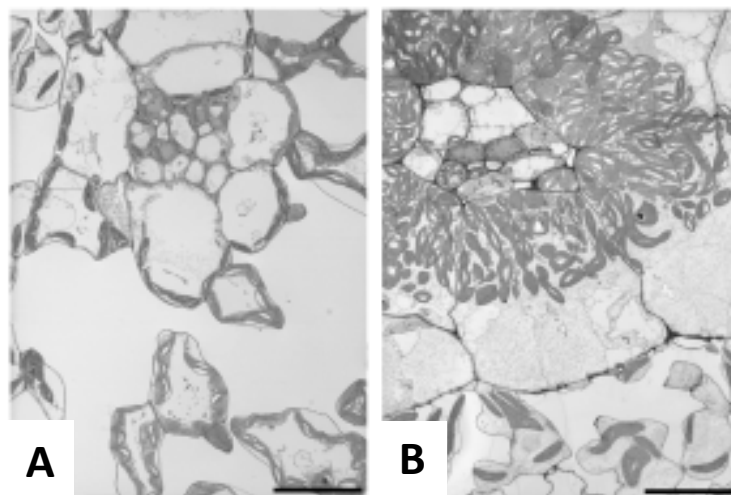
impliqués dans la production d'ATP par le flux cyclique d'électrons sont surexprimés chez *C. gynandra* notamment le PSI, le complexe cytochrome b6/f et les gènes impliqués dans le flux cyclique d'électrons eux-mêmes. La demande d'ATP et de pouvoir réducteur est augmenté pour alimenter le cycle en C4. En effet, le recyclage de l'accepteur primaire de carbone, le PEP, nécessite de l'ATP supplémentaire (PPDK). A l'inverse, les gènes impliqués dans la photorespiration et le cycle de Calvin et Benson sont moins exprimés chez *C. gynandra*. Il est bien connu que chez la plupart des plantes en C4, il y a moins de RubisCO par rapport aux plantes en C3<sup>41</sup>. De même, la photorespiration est beaucoup moins active chez les plantes en C4<sup>42,43</sup>. L'expression de la grosse sous-unité et de la petite sous-unité de la RubisCO est réduite de 1.5 et 1.4 fois, respectivement, chez *C. gynandra* par rapport à *C. spinosa*.

Outre ces gènes dont la fonction est déjà connue, d'autres gènes candidats ont été mis à jour.

NHD1 est un transporteur qui est exprimé 16 fois plus dans les feuilles de *C. gynandra* que dans celles de *C. spinosa*. Il s'agit d'un symport H<sup>+</sup>/Na<sup>+</sup> hypothétique. L'importation de pyruvate dans les cellules du mésophylle est dépendante du sodium (symport pyruvate/Na<sup>+</sup>)<sup>44,45</sup>. NHD1 pourrait intervenir dans le maintien de l'homéostasie du sodium dans le chloroplaste (sortie du Na<sup>+</sup>).

L'expression du système PIP1B :CA4 est également plus forte chez *C. gynandra*. PIP1B est une protéine insérée dans la membrane plasmique qui pourrait être impliquée dans le transport du CO<sub>2</sub> (expression multipliée par 20) et CA4 est une anhydrase carbonique (expression multipliée par 6) qui assurerait la conversion du CO<sub>2</sub> en bicarbonate. Ce système permettrait de faciliter l'acquisition du CO<sub>2</sub> et sa mise à disposition pour la PEPC.

Les chloroplastes des cellules de la gaine chez *C. gynandra* sont plus gros que ceux des cellules de la gaine de *C. spinosa* et comme chez beaucoup de C4, ils sont positionnés proches des nervures alors que chez *C. spinosa*, ils sont localisés sur la périphérie des cellules à l'opposé des nervures (Fig. 23)<sup>46,47</sup>.



**Figure 23.** Cellules de la gaine de *C. spinosa* (A) et *C. gynandra* (B). Chez *C. spinosa*, les chloroplastes sont disposés de manière centrifuge alors que chez *C. gynandra*, ils sont en position centripètes et sont beaucoup plus gros (d'après Marshall et al., 2007<sup>46</sup>).

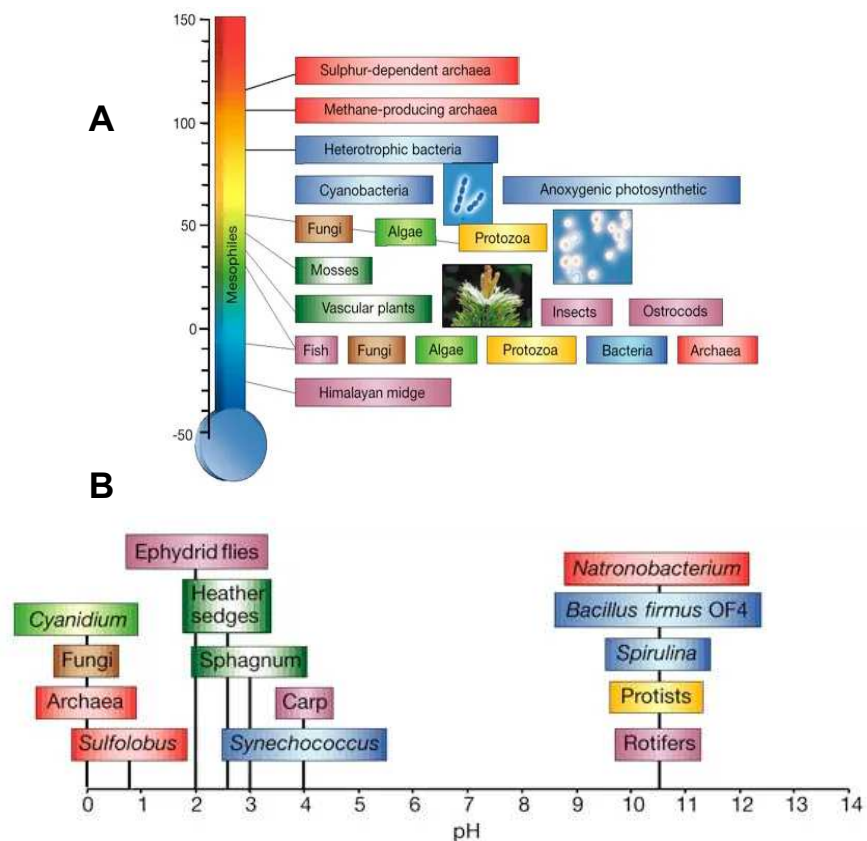
Une réduction des transcrits codant CHUP1 et ACTIN11 est observée chez *C. gynandra*. Chez *A. thaliana*, CHUP1 (chloroplast unusual positioning1) est localisée dans l'enveloppe externe du chloroplaste et cette protéine possède un motif de liaison à l'actine. CHUP1 permet d'éviter l'agrégation des chloroplastes<sup>48</sup>. Le positionnement différentiel des chloroplastes dans les cellules de la gaine et les cellules du mésophylle chez l'éleusine (*Eleusine coracana*) et le maïs nécessite le système actomyosin<sup>49</sup>. Comme AtCHUP1 est impliqué dans le positionnement homogène des chloroplastes lors de traitement lumineux grâce à une

interaction avec le cytosquelette<sup>48</sup>, il est possible que la faible expression de CHUP1 et ACTIN11 chez *C. gynandra* favorise un positionnement centripète des chloroplastes.

Ce travail a fait l'objet d'une publication en 2011 (publication n°10).

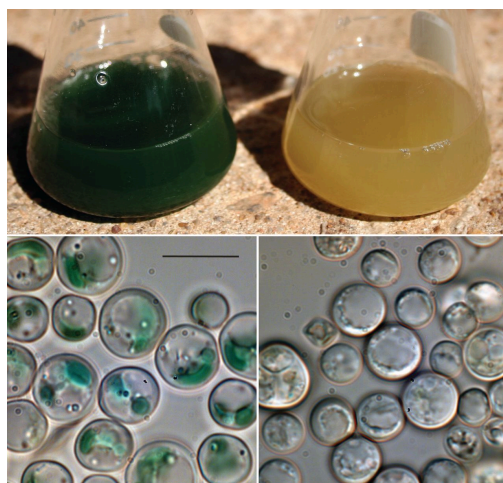
### 2.3. Séquençage du premier génome d'algue rouge unicellulaire

*Galdieria sulphuraria* est une algue rouge unicellulaire (eucaryote) appartenant à la famille des *Cyanidiophyceae*. Les algues rouges de cette famille sont les principaux organismes photosynthétiques qui colonisent les environnements extrêmes tels que les environnements très chauds ou très acides aux cotés de bactéries et d'archées mais où les bactéries photosynthétiques sont absentes<sup>50</sup>. Les *Cyanidiophyceae* peuvent se développer à des pHs très faibles (pH 0.05 à 4) et à des températures allant jusqu'à 56°C (Fig. 24)<sup>51,52</sup>.



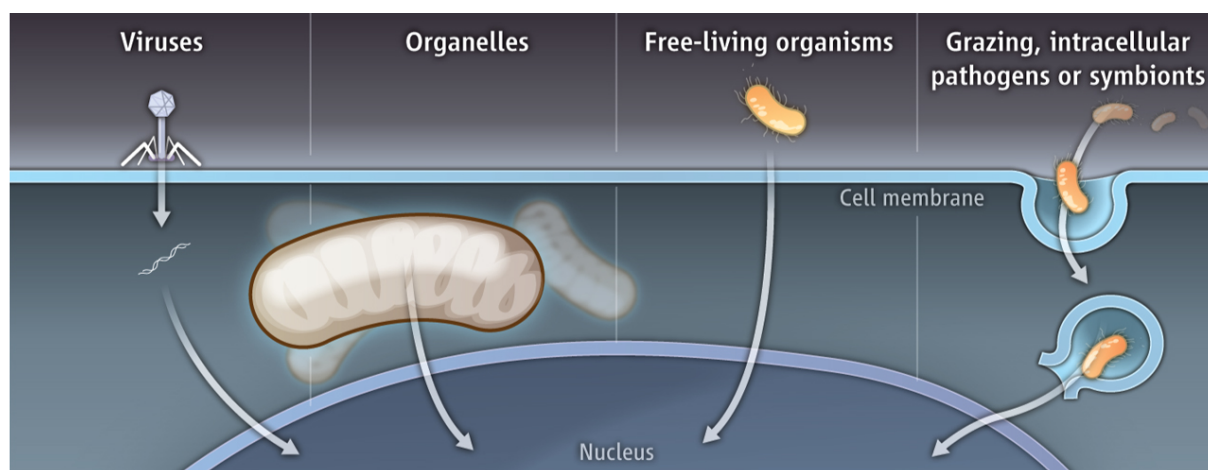
**Figure 24.** Zones de température (A) et de pH (B) pour laquelle la vie est possible (d'après Rothschild et Mancinelli, 2001<sup>51</sup>).

*G. sulphuraria* est un organisme extrémophile (colonisation d'environnements hostiles) qui tolère également les fortes concentrations en sels et en métaux et qui possède une flexibilité métabolique originale. Un organisme extrémophile est capable de survivre et de se développer dans des environnements extrêmes. Les conditions extrêmes regroupent les facteurs physiques extrêmes (température, radiations, pression...) et les facteurs géochimiques extrêmes (sécheresse, salinité, pH...). *G. sulphuraria* peut se développer en autotrophie mais également en hétérotrophie en présence de plus de 50 sources de carbone telles que des sucres, des polyols, des intermédiaires du cycle de Krebs et des acides aminés (Fig. 25). Cette flexibilité métabolique représente un cas presque unique dans le monde vivant<sup>52</sup>. Dans des niches contenant de fortes concentrations en arsenic, aluminium, cadmium, mercure et autres métaux toxiques, *G. sulphuraria* représente jusqu'à 90% de la biomasse totale et presque la totalité de la biomasse eucaryote<sup>50</sup>.



**Figure 25.** Culture autotrophique (à gauche) et hétérotrophique (à droite) de cellules de *Galdieria sulphuraria*. Cultures cellulaires (en haut) et images en microscopie photonique (en bas ; barre = 10 µm) de cellules cultivées en conditions de lumière continue en absence de glucose ou à l'obscurité en présence de 200 mM de glucose (d'après Schönknecht et al. 2013<sup>53</sup>).

Afin de tenter d'expliquer quels étaient les mécanismes moléculaires à l'origine de ces particularités physiologiques et métaboliques, nous avons entrepris le séquençage de son génome. La taille du génome est de 13.7 Mb. L'innovation chez les eucaryotes est le fruit de duplications de gènes suivies de néo-fonctionnalisations ce qui conduit à une expansion des familles de gènes au cours de l'évolution<sup>54</sup>. Chez les archées et les bactéries, l'adaptation et l'évolution s'opère au travers de transferts horizontaux de gènes à partir d'autres organismes<sup>55</sup>. Ce mécanisme d'évolution a pendant longtemps été considéré spécifique des procaryotes à l'exception de la transmission des éléments transposables chez des espèces proches et de transfert intracellulaire des mitochondries et des plastides vers le génome nucléaire<sup>56,57</sup>. Plus récemment, le transfert horizontal de gènes à partir d'endosymbiontes et d'organismes libres a également été démontré chez certains eucaryotes unicellulaires<sup>58-60</sup>. Ces études démontrent que les transferts horizontaux de gènes participent à la diversification génétique des eucaryotes au cours de l'évolution (Fig. 26).

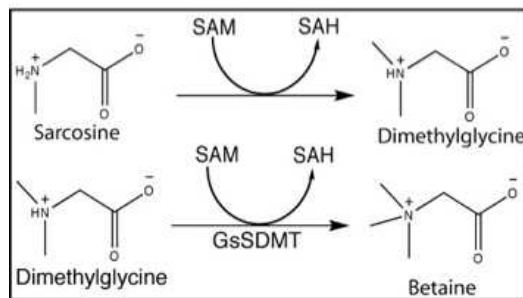


**Figure 26.** Différentes voies pour le transfert de gènes. Les génomes eucaryotes peuvent intégrer de l'ADN étranger de différentes manières (d'après Rocha, 2013<sup>61</sup>).

Nous avons démontré, à partir de l'analyse du génome, que chez *G. sulphuraria*, au moins 75 transferts de gènes à partir de bactéries ou d'archées se sont produits au cours de l'évolution. Après duplication et néo-fonctionnalisation, 5% des gènes du génome sont issus de ces événements. Plusieurs des événements mis à jour sont à mettre en relation avec

l'extrémophilie de *G. sulphuraria* et de son incroyable flexibilité métabolique. Je ne présenterai ici que 2 exemples.

L'halophilie extrême de cette algue rouge peut être liée à ces transferts de gènes. La résistance au stress salin (NaCl) implique l'exclusion du Na<sup>+</sup> du cytosol et l'augmentation de l'osmolarité grâce à l'accumulation de solutés compatibles dans le cytosol. En plus de plusieurs antiports Na<sup>+</sup>:H<sup>+</sup> d'origine eucaryote, le génome de *G. sulphuraria* code 2 transporteurs Na<sup>+</sup>:H<sup>+</sup> qui sont probablement d'origine bactérienne d'après les arbres phylogénétiques. De plus, des gènes codant la sarcosine diméthylglycine méthyltransferase (SDMT) sont probablement issus de transfert à partir de cyanobactéries. Cette enzyme catalyse la formation de glycine bêtaïne à partir de sarcosine chez *G. sulphuraria* (Fig. 27)<sup>62</sup>. Nous avons, en effet, montré que *G. sulphuraria* produisait et accumulait de la glycine bêtaïne et que cette accumulation était accrue en condition de stress salin.



**Figure 27.** Réactions catalysées par la sarcosine diméthylglycine méthyltransferase (GsSDMT) et permettant la synthèse de glycine bêtaïne chez *G. sulphuraria* (d'après McCoy et al. 2009<sup>62</sup>). Le donneur de groupement méthyle est la S-adenosyl méthionine (SAM).

De même, les analyses phylogénétiques indiquent que des transferts de gènes pourraient être à l'origine des particularités métaboliques de *G. sulphuraria*. Des perméases à acétate semblent provenir de génomes bactériens alors que des transporteurs à acides aminés, polyamines et cations organiques proviendraient d'archées thermo-acidophiles. La croissance hétérotrophe avec le glycérol comme seule source de carbone serait liée à la présence d'une famille de 5 transporteurs de glycérol et d'une famille de 3 glycérol déshydrogénases. Ces 2 familles seraient issues de transferts horizontaux de gènes.

Nous avons ainsi montré, pour la première fois chez un eucaryote, que des transferts horizontaux de gènes chez *G. sulphuraria* expliqueraient les formidables capacités d'adaptation à des environnements hostiles et à des conditions trophiques variées de cette algue rouge unicellulaire.

L'ensemble de ces résultats ont fait l'objet d'une publication (publication n°13).



## IV- Métabolisme et rôles des acides hydroxycinnamiques

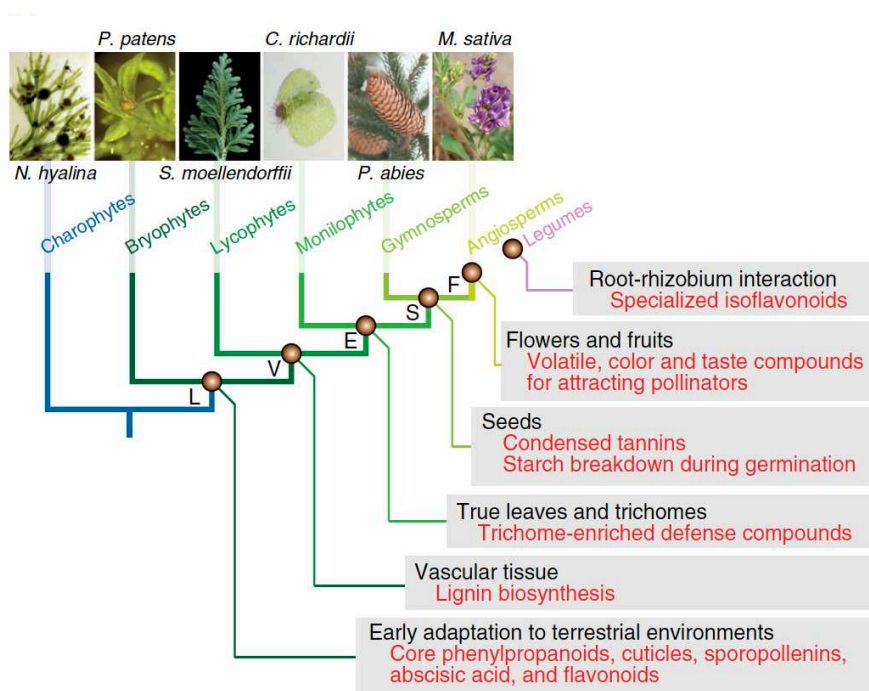
J'ai été recruté à l'Université de Lille<sup>1</sup> en septembre 2009. Notre équipe, Métabolites Secondaires et Stress Abiotiques (MSSA), appartenait à cette époque à l'UMR 1281 SADV (Stress et Adaptations des Végétaux Cultivés). En 2014, à la suite d'une restructuration, notre équipe a rejoint l'Institut Charles Viollette (Laboratoire Régional en Agroalimentaire et Biotechnologies) lors de sa création. Nous avons alors intégré l'équipe ASQV (Adaptation aux Stress et Qualité des Végétaux). En 2020, l'Institut Charles Viollette a intégré l'UMRT 1158 BioEcoAgro. Notre équipe fait maintenant partie de l'équipe 5 intitulée métabolites secondaires d'origine végétale. Cette équipe appartient au pôle 2 de l'UMRT : Molécules d'origine végétale et microbienne : de l'identification à la bioproduction.

J'ai initié des approches de génomique fonctionnelle dans le but de caractériser les voies de biosynthèse des composés phénoliques chez la chicorée, une espèce emblématique de la région Hauts-de-France. Dans ce contexte, notre équipe a généré des ressources moléculaires importantes facilitant ce type d'approche (séquençage du génome de la chicorée en collaboration avec nos partenaires industriels, librairies d'ESTs par séquençage haut débit, article 18).

### 1. Introduction

Le métabolome est l'ensemble des métabolites présents dans une cellule, un tissu ou un organisme. Ainsi, sont considérés comme des métabolites, les molécules organiques dont la masse est inférieure à 1500 Da. Ces molécules sont les substrats, les intermédiaires ou les produits du métabolisme, ce dernier étant défini comme l'ensemble des réactions biochimiques à l'origine de la transformation des molécules organiques au sein d'un organisme vivant.

L'évolution de la lignée verte est largement associée à l'apparition de nouvelles voies métaboliques et notamment celles engendrant la synthèse de métabolites spécialisés. Au cours de la colonisation de l'environnement terrestre, pour survivre aux conditions souvent hostiles et se défendre contre les animaux et les microorganismes évoluant en parallèle, les plantes ont opté pour une stratégie unique en étoffant leur arsenal métabolique ce qui a conduit à la production d'une large panoplie de molécules non présentes chez leurs ancêtres aquatiques. Ces composés qualifiés de métabolites spécialisés sont souvent accumulés de manière très restreinte au sein du règne végétal voir d'une même famille ou espèce. Des rôles primordiaux sont aujourd'hui attribués à ces métabolites pourtant longtemps qualifiés de secondaires en opposition aux métabolites primaires dont le rôle est essentiel pour la croissance et la reproduction de la plante. Leur nombre a constamment progressé au cours des processus de conquête du milieu terrestre par les végétaux il y a 500 millions d'années et leur rôle majeur dans les grandes étapes de l'évolution est établi (Fig. 28)<sup>63</sup>. Ainsi, le terme de métabolites spécialisés est, aujourd'hui, préféré pour refléter à la fois le rôle majeur de ces molécules dans la biologie de la plante (fonctions spécialisées) et le fait que leur répartition taxonomique est souvent très restreinte.

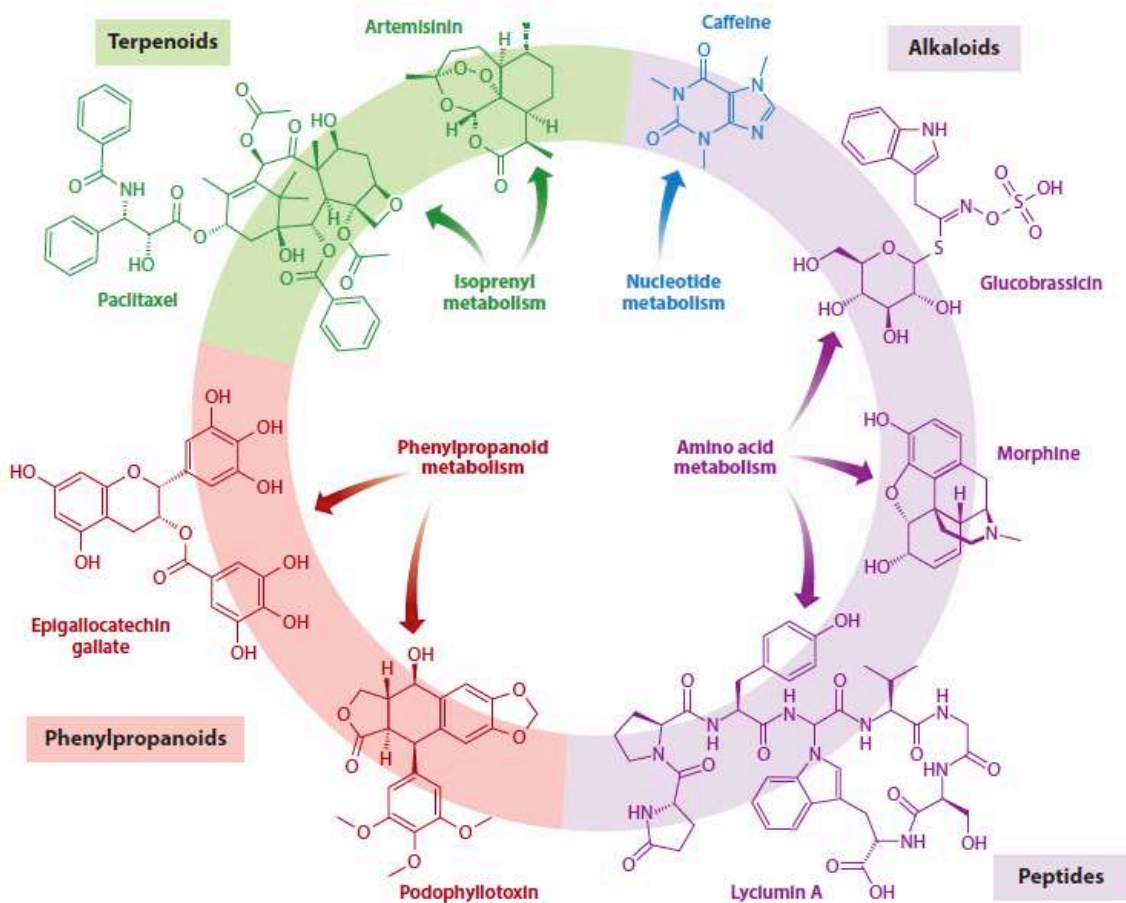


**Figure 28.** Implication du métabolisme spécialisé des plantes dans les grandes étapes de l'évolution (d'après Weng, 2014)<sup>63</sup>.

Ces molécules sont impliquées dans des processus aussi variés que l'attraction de pollinisateurs, la protection contre les stress abiotiques tels que les rayonnements UV ou la dessiccation, les interactions entre plantes, la défense contre les herbivores ou encore l'attraction des microorganismes symbiotiques fixateurs d'azote atmosphérique<sup>64</sup>. Le nombre de métabolites spécialisés accumulés dans le règne végétal est estimé aux environs de 200000 ce qui, au vu des résultats récents de profilage métabolique, semble être une sous-estimation importante<sup>65</sup>. De plus, seuls 15% des 350000 espèces de plantes référencées (des estimations annoncent l'existence de plus de 500000 espèces de plantes) ont été analysés pour cette caractéristique<sup>66</sup>.

D'un point de vue métabolique, ces métabolites prennent leur source au sein du métabolisme primaire<sup>67</sup>. Trois familles de métabolites spécialisés sont classiquement distinguées : les terpènes ou isoprénoïdes, les composés azotés (dont les alcaloïdes) et les composés phénoliques également nommés phénylpropanoïdes (Fig. 29).

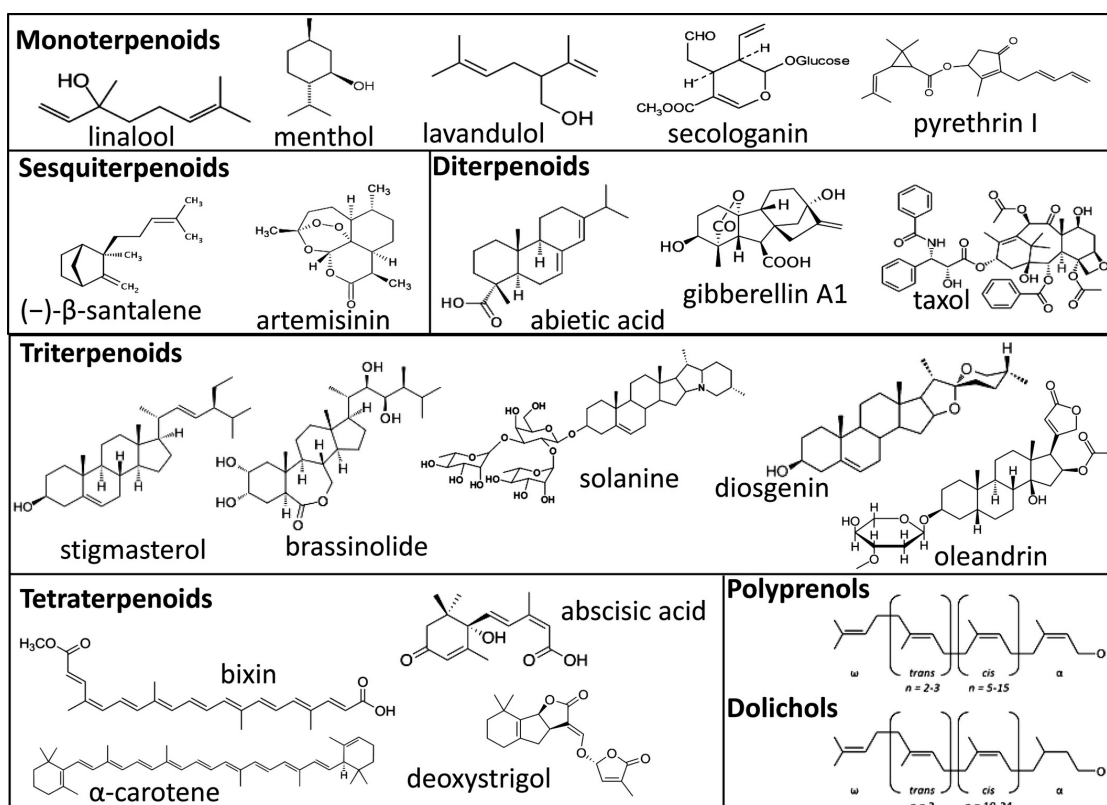
Une diversité structurale énorme existe au sein de chaque classe et un niveau de diversification supplémentaire est le résultat de la combinaison de structures issues de ces 3 classes. Par exemple, le paclitaxel (taxol), un anticancéreux, est issu de la combinaison de molécules issues du métabolisme des alcaloïdes, des terpènes et des phénylpropanoïdes.



**Figure 29.** Les principales classes de métabolites spécialisés des plantes et leur origine métabolique. Ils sont dérivés de précurseurs issus du métabolisme primaire dont le métabolisme des isoprénoides, des phénylpropanoïdes, des acides aminés et des nucléotides (d'après Jacobowitz et Weng, 2020<sup>68</sup>).

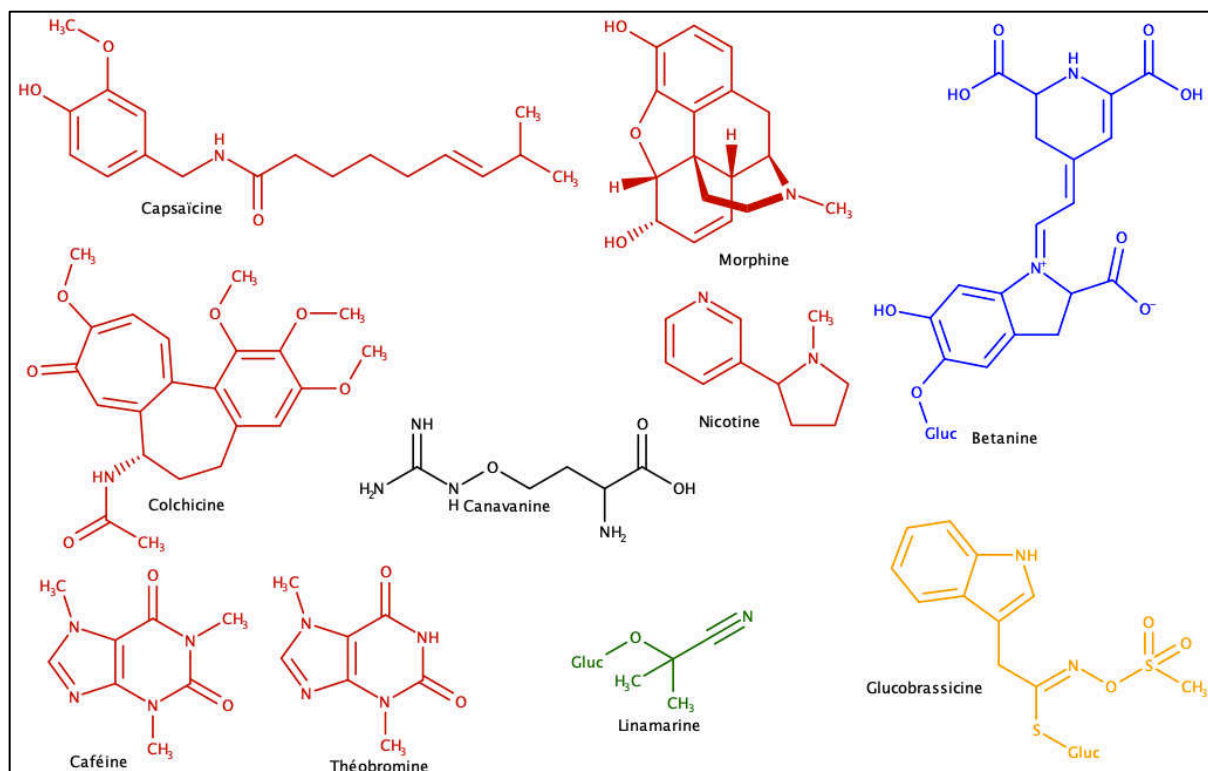
Les terpènes appartiennent à une classe hautement diversifiée d'hydrocarbures. Ils sont formés à partir d'unités isomériques à 5 carbones, le diméthylallylpyrophosphate (DMAPP) et l'isopenténylpyrophosphate (IPP), issus du métabolisme primaire. L'IPP et le DMAPP sont synthétisés par la voie du mévalonate (MVA) et la voie du méthylérythritol phosphate (MEP). La voie du MVA a été localisée dans le cytosol mais des éléments tendent à montrer qu'une partie au moins pourrait être située dans les peroxysomes<sup>69</sup>. La voie du MEP est localisée dans les chloroplastes<sup>70</sup>. Le DMAPP et l'IPP sont condensés pour former des isoprènes à chaînes longues et linéaires contenant un nombre de carbones multiple de 5, sous l'action de prényl transférase. Ces réactions d'addition font intervenir la géranyl diphosphate synthase (GPPS) qui produit le géranyl diphosphate (GPP, C10), la farnésyl diphosphate synthase produisant le farnésyl pyrophosphate (FPP, C15) et la géranylgéranyl diphosphate synthase (GGPPS) conduisant au géranylgéranyl diphosphate (GGPP, C20). La squalène synthase, quant à elle, catalyse la synthèse de squalène (C30) et la phytoène synthase celle de phytoène (C40). Ces 5 composés constituent les squelettes pour la synthèse des terpènes cycliques ou linéaires, sous l'action de terpènes synthases, dont la classification dépend de la longueur des chaînes des précurseurs. Ainsi, les chaînes à 5, 10, 15, 20, 30 et 40 carbones constituent, après métabolisation, les hémiterpènes, les monoterpènes, les sesquiterpènes, les diterpènes, les triterpènes et les tetraterpènes, respectivement. Au-delà, on parle de polyterpènes. Ces structures de base sont ensuite modifiées pour former une multitude de composés terpéniques aux structures très variées après hydroxylation, déshydratation, oxydation, cyclisation ou conjugaison (Fig. 30). De nombreux terpènes appartiennent aux métabolites primaires. C'est

le cas, par exemple, des chlorophylles, des ubiquinones, des plastoquinones et des phytohormones (gibbérellines, cytokinines, acide abscissique)<sup>70</sup>.



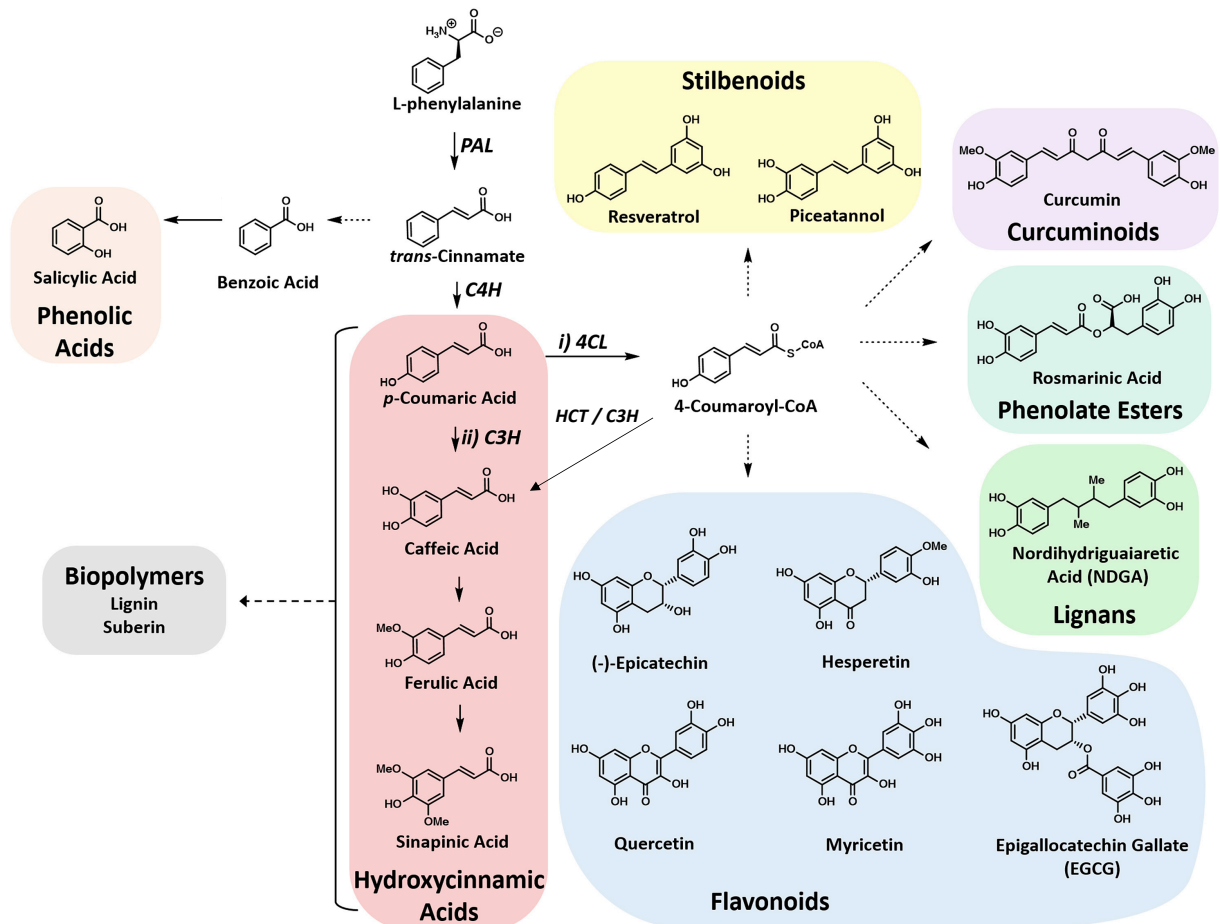
**Figure 30.** Exemples de terpènes produits par les plantes. Les monoterpènes et les sesquiterpènes sont généralement volatiles quand ils ne sont pas conjugués tandis que les terpènes de classe supérieures ne sont généralement pas volatiles (d'après Pichersky et Raguso, 2018<sup>71</sup>).

Les composés azotés regroupent les alcaloïdes aux propriétés pharmaceutiques et médicinales exceptionnelles, les acides aminés non protéinogènes, les bétalaïnes, les glycosides cyanogéniques et les glucosinolates (Fig. 31). Les alcaloïdes, aux structures chimiques souvent très complexes, comptent des molécules bien connues telles que la nicotine, la morphine, la caféine, la cocaïne, la colchicine ou la vinblastine. Plus de 12000 alcaloïdes ont été identifiés chez les plantes. Même s'ils sont distribués de manière ubiquitaire chez les plantes, ils sont particulièrement abondants dans certaines familles telles que les *Solanaceae*, les *Manispermaceae*, les *Papaveraceae*, les *Berberidaceae* et les *Fabaceae*. Ils sont classés en 3 catégories dépendamment de leur origine biosynthétique et de leur structure chimique. Les alcaloïdes vrais comportent un hétérocycle azoté et sont synthétisés à partir d'acides aminés tels que l'ornithine, l'aspartate, la lysine, la tyrosine ou le tryptophane. Plusieurs classes d'alcaloïdes vrais sont définies de par la nature des cycles azotés. Les proto-alcaloïdes dérivent également des acides aminés mais ne comportent pas d'hétérocycle azoté. La colchicine, qui est synthétisée à partir de la phénylalanine, appartient à cette classe d'alcaloïde. Les pseudo-alcaloïdes présentent dans leur structure un hétérocycle azoté mais ne dérivent pas des acides aminés. C'est le cas, par exemple, de la théobromine et de la caféine synthétisées à partir de la guanine. Comme leur nom l'indique, les acides aminés non protéinogènes sont des acides aminés ne participant pas à la synthèse des protéines et qui sont des homologues structuraux d'acides aminés protéinogènes, comme par exemple la canavanine, l'azetidine-2-carboxylic acid ou la mimosine. Ils sont synthétisés et accumulés principalement dans les graines des Fabacées. Les gluconisolates sont des dérivés soufrés d'acides aminés majoritairement présents dans la famille des Brassicacées. Les glycosides cyanogéniques sont issus d'acides aminés et sont responsables lors de leur hydrolyse de la libération de composés à groupement nitrile dont le cyanure.



**Figure 31.** Structures chimiques de composés azotés appartenant à la famille des alcaloïdes (en rouge), des glucosinolates (en orange), des glycosides cyanogéniques (en vert), des acides aminés non protéinogènes (en noir) et des bêtaïnes (en bleu). Gluc : Glucose.

Les composés phénoliques ont en commun la présence d'un (ou plusieurs) cycle(s) benzénique(s) hydroxylé(s). Ils sont, pour la plupart, synthétisés à partir de la phénylalanine, acide aromatique produit par la voie chloroplastique du shikimate. La désamination oxydative de la phénylalanine par la phénylalanine ammonia lyase (PAL) conduit à la formation de l'acide cinnamique, un composé avec un squelette carboné en C6-C3. La réaction catalysée par la PAL constitue la première étape de la voie dite des phénylpropanoïdes. Un des intermédiaires de cette voie, le coumaroyl-CoA, constitue un précurseur essentiel à la synthèse de la quasi-totalité des composés phénoliques dont les acides hydroxycinnamiques (acide caféique, acide férulique, acide sinapique etc), la lignine, les flavonoïdes, les coumarines, les tanins et les stilbènes (Fig. 32). Sa synthèse implique l'action successive de la C4H (cinnamate 4-hydroxylase) qui convertit l'acide cinnamique en acide coumarique et de la 4CL (4-coumarate : CoA ligase) qui permet la synthèse de coumaroyl-CoA. Chez les monocotylédones, la PAL possède une activité TAL (tyrosine ammonia lyase)<sup>72</sup>. Cette activité permet de convertir directement la tyrosine, un autre acide aminé aromatique issu de la voie du shikimate, en acide coumarique. Comme évoqué précédemment les composés phénoliques ont joué un rôle très important dans la conquête du milieu aérien.



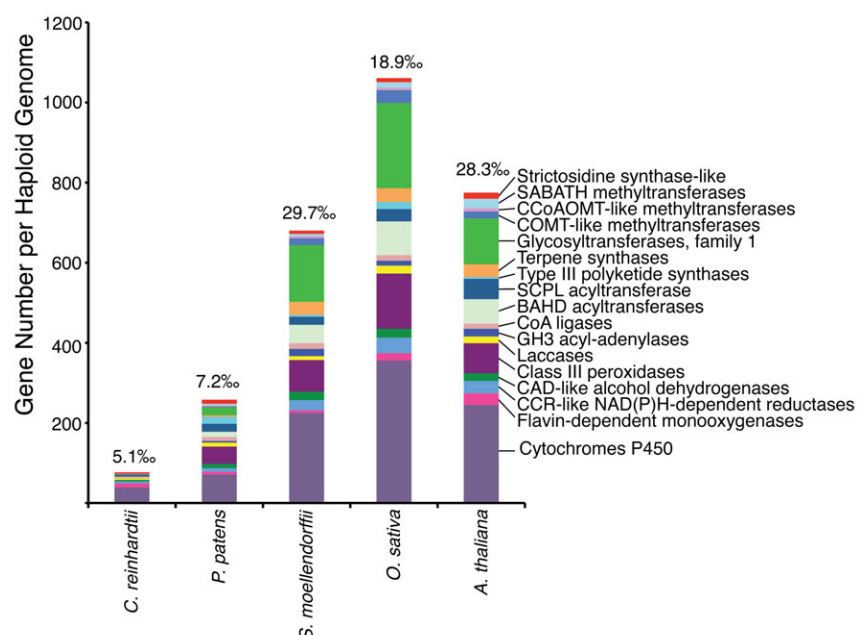
**Figure 32.** Voie de biosynthèse simplifiée des différentes classes de composés phénoliques à partir de la phénylalanine. Les flèches pleines indiquent une seule étape enzymatique tandis que les flèches en pointillés désignent plusieurs étapes enzymatiques. Le coumaroyl-CoA représente un carrefour important pour la synthèse de ces composés. PAL : phénylalanine ammonia lyase, C4H : cinnamate 4-hydroxylase, 4CL : 4-coumarate : CoA ligase, C3H : p-coumarate 3'-hydroxylase, HCT : hydroxycinnamoyl-CoA shikimate/quinatate hydroxycinnamoyl transférase (d'après Kolaj et al., 2018<sup>73</sup>).

L'évolution des voies métaboliques à l'origine de cette diversification chimique a fait l'objet d'études depuis plusieurs décennies<sup>74</sup>. Au cours des 25 dernières années, un nombre important d'enzymes du métabolisme spécialisé ont été démontrées comme étant issues d'enzymes intervenant dans les voies du métabolisme primaire (Tableau IV)<sup>63</sup>. Ce phénomène est qualifié de recrutement d'enzymes du métabolisme primaire. En effet, l'un des moteurs de l'innovation métabolique est la duplication de gènes préexistants. A la suite de ces duplications, la pression de sélection peut être relâchée sur une des deux copies, l'autre conservant sa fonction initiale, et des modifications génomiques successives peuvent aboutir soit à la modification des séquences codantes d'enzymes préexistantes ou à la modification des séquences régulatrices des gènes entraînant une variation spatiale et/ou temporelle de leur expression. Dans ce dernier cas, les enzymes sont exprimées dans de nouveaux contextes biochimiques et de nouvelles combinaisons substrats-produits peuvent émerger. Dans ces 2 cas, ces phénomènes de duplication peuvent aboutir à des phénomènes de néofonctionnalisation autrement dit à une diversification métabolique conduisant à l'apparition de nouvelles voies métaboliques. Beaucoup de gènes issus de duplication sont également perdus au cours du temps<sup>75</sup>. Ces mécanismes ont abouti à une expansion des familles des enzymes du métabolisme spécialisé. Parmi ces familles, on peut citer la famille des cytochromes p450 (245 représentants chez *A. thaliana*) qui est issue de l'évolution d'une stérol 14-déméthylase (enzyme du métabolisme primaire). Une autre famille dont il sera question dans la suite de ce manuscrit est bien représentée : la famille BAHD. Il s'agit d'une famille d'acyltransférases.

**Tableau IV.** Familles d'enzymes majeures du métabolisme spécialisé des plantes et leur correspondance au sein du métabolisme primaire. Le nombre de gènes dans les génomes de l'algue verte *Chlamydomonas reinhardtii* et de la plante à fleurs *Arabidopsis thaliana* codant des enzymes de chacune des familles est listé pour comparaison (d'après Weng, 2014<sup>63</sup>).

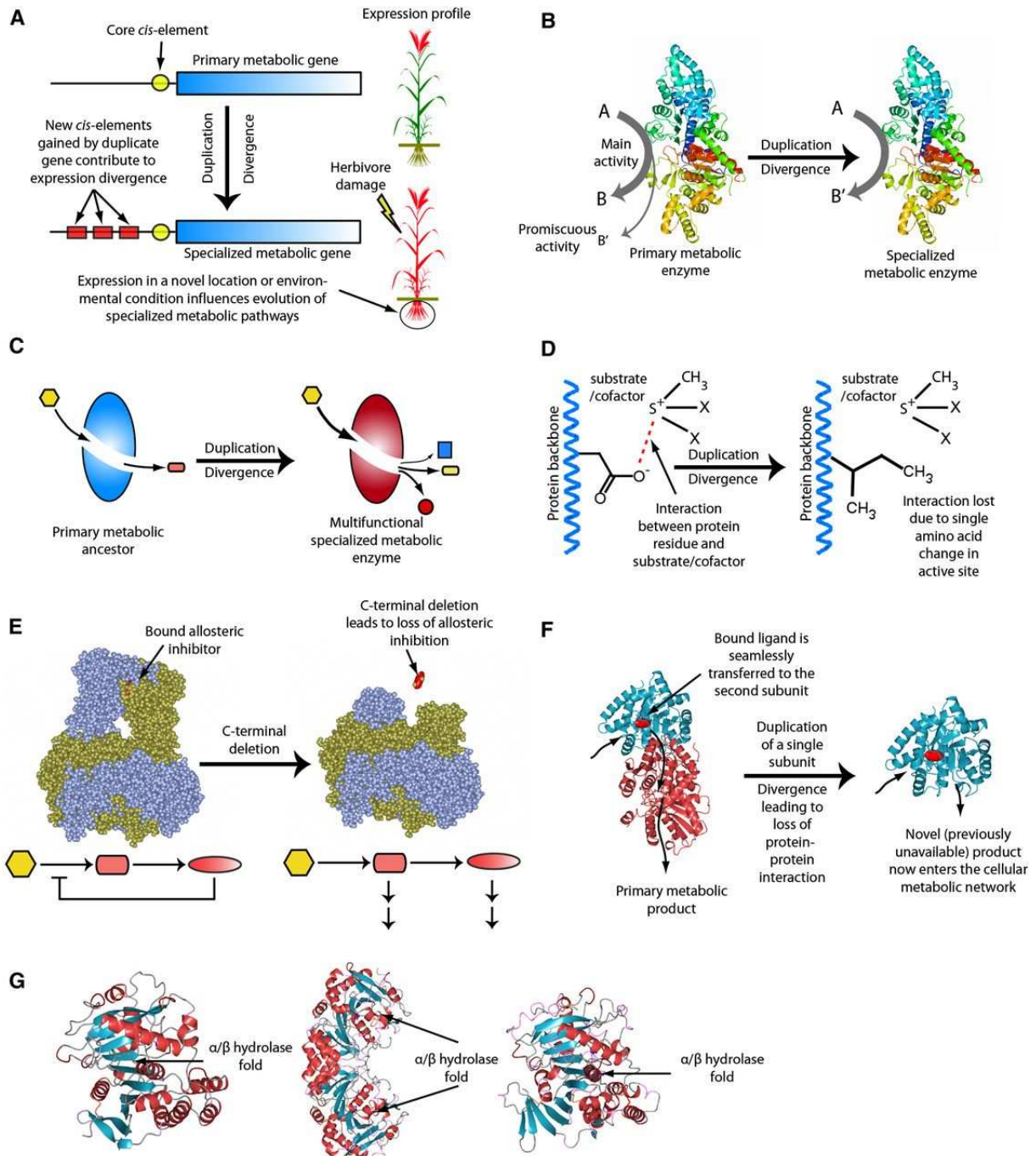
Major enzyme families involved in plant specialized metabolism	Gene number		Cousins in primary metabolism
	<i>C. reinhardtii</i>	<i>A. thaliana</i>	
Cytochromes P450	39	245	Sterol 14-demethylase
Glycosyltransferase family 1	1	115	Lipopolysaccharide glycosyltransferase
2-oxoglutarate-dependent dioxygenase	0	84	phytanoyl-CoA dioxygenase
Class III peroxidase	0	75	Cytochrome c peroxidase
BAHD acyltransferase	0	61	Carnitine acetyltransferase
SCPL acyltransferase	7	51	Serine carboxypeptidase
Terpene synthase	0	32	Farnesyl pyrophosphate synthase
CCR-like NAD(P)H-dependent reductase	5	31	3-β-Hydroxysteroid dehydrogenase
Flavin-dependent monooxygenase	10	29	trimethylamine oxidase
SABBATH methyltransferase	0	24	Catechol-O-methyltransferase
CAD-like alcohol dehydrogenase	5	19	Alcohol dehydrogenase
GH3 acyl-adenylase	0	19	Long-chain fatty acyl-CoA synthetase
COMT-like methyltransferase	1	17	Catechol-O-methyltransferase
Strictosidine synthase-like	1	15	Gluconolactonase
Acyl-CoA ligase	4	13	Long-chain fatty acyl-CoA synthetase
CCoAOMT-like methyltransferase	1	8	Catechol-O-methyltransferase
Type III polyketide synthase	0	4	β-ketoacyl-ACP synthase III

En conséquence, au cours de l'évolution des plantes terrestres, la proportion du génome codant des enzymes du métabolisme spécialisé n'a cessé de croître ce qui s'est traduit par une augmentation continue du nombre de métabolites spécialisés chez les plantes. (Fig. 33).



**Figure 33.** Proportion de gènes codant des enzymes du métabolisme spécialisé (% indiqués sur les bars) par rapport au nombre total de gènes au cours de l'évolution (d'après Weng et Joel, 2013<sup>76</sup>).

Le processus de recrutement des enzymes du métabolisme primaire et de divergence peut s'opérer en suivant plusieurs voies tels que le changement de la régulation transcriptionnelle ou de la régulation allostérique des enzymes, la modification de la spécificité catalytique (une activité auxiliaire catalysée par l'enzyme du métabolisme primaire devient l'activité prépondérante) ou des modifications structurales plus complexes (changement des interactions protéine-protéine entraînant une modification de la structure de la protéine) (Fig. 34).

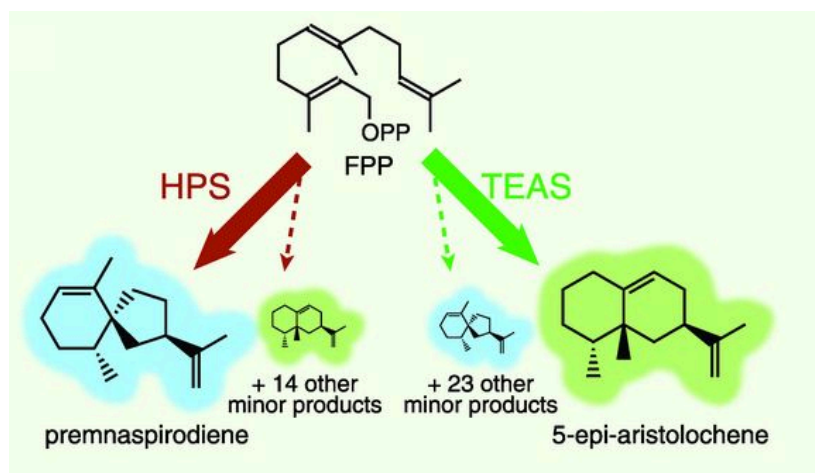


**Figure 34.** Mécanismes conduisant à la divergence de gènes dupliqués et à leur recrutement consécutif par le métabolisme spécialisé. A : après duplication, le gène dupliqué acquiert de nouveaux éléments cis-régulateurs ce qui conduit à des profils d'expression différents par rapport au gène initial. Des divergences aux niveaux spatial, temporel ou de la sensibilité à l'environnement peut conduire à l'apparition de nouvelles voies métaboliques. B : une enzyme du métabolisme primaire possède une activité principale et une activité auxiliaire. Après duplication, une copie du gène conserve et se spécialise dans cette activité auxiliaire. C : Une enzyme du métabolisme primaire utilise un substrat pour produire une seule molécule. Après duplication, une copie conserve le même substrat mais catalyse la formation de plusieurs produits. D : Le changement d'un seul acide aminé de la séquence d'une enzyme dupliquée supprime l'interaction avec le substrat original ou un cofacteur ce qui empêche l'activité initiale et ouvre la voie à de nouvelles activités. E : La suppression de la régulation allostérique par le produit final d'une voie dans l'enzyme dupliquée perturbe l'équilibre des réactions dans cette voie et modifie la disponibilité en substrats dans la cellule. F : Une interaction protéine-protéine canalise la circulation des molécules d'une voie entre 2 enzymes ce qui prévient la libération des intermédiaires. La modification de cette interaction dans l'enzyme dupliquée permet la libération des intermédiaires alors disponibles pour participer à d'autres réactions catalysées par d'autres enzymes. G : Les enzymes modélisées ont des séquences primaires très différentes. Cependant, elles conservent le même repliement  $\alpha/\beta$  ce qui leur permet de maintenir des activités similaires (d'après Moghe et Last, 2015<sup>64</sup>).



De par la nature aléatoire de ces processus de néo-fonctionnalisation, de nombreux métabolites spécialisés sont produits spécifiquement chez certaines familles ou espèces et la capacité à les produire peut être conservée ou non en fonction des avantages évolutifs potentiellement conférés par ces molécules. Cette caractéristique des métabolites spécialisés est souvent utilisée en classification au travers ce qu'on appelle la chimio-taxonomie. Par exemple, les lactones sesquiterpéniques sont une caractéristique chimique de la famille des Astéracées et la nicotine se retrouve exclusivement chez le genre *Nicotiana* à quelques exceptions près où celle-ci est présente à l'état de traces<sup>77</sup>. D'autres molécules sont plus largement distribuées comme les acides chlorogéniques néanmoins présents majoritairement, d'un point de vue quantitatif, chez les Solanacées, les Rubiacées et les Astéracées.

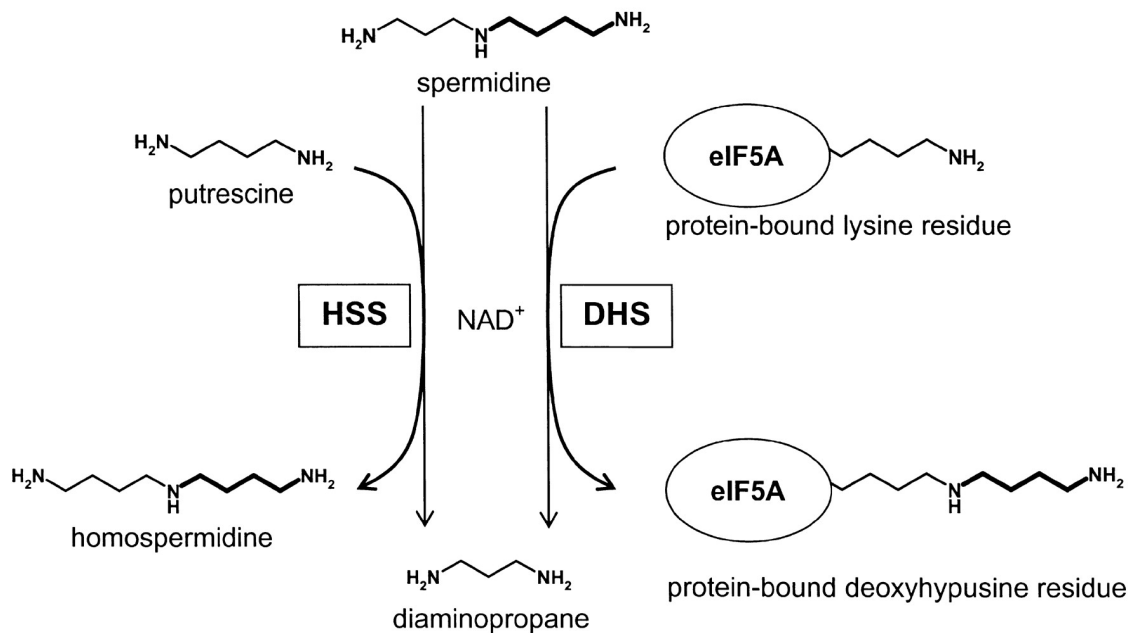
D'une manière générale, ce qui caractérise les enzymes du métabolisme spécialisé est leur relative promiscuité métabolique en comparaison à leurs équivalents du métabolisme primaire<sup>78</sup>. De même, alors que les enzymes du métabolisme primaire possèdent une très grande efficacité catalytique de manière à optimiser les flux métaboliques, celle du métabolisme spécialisé est en général bien moins efficace (30 fois en moyenne<sup>79</sup>). Bien souvent, les enzymes de ce métabolisme catalysent la formation de plusieurs produits minoritaires en parallèle de leurs activités principales. Par exemple, les terpènes synthases sont des enzymes dont on a démontré la capacité à produire une multitude de produits ou sous-produits (Fig. 35).



**Figure 35.** Exemples d'enzymes présentant une très faible spécificité de substrats. HPS : premnaspirodiene synthase de *Hyocymus muticus*, TEAS : 5-epiaristolochene synthase de *Nicotiana tabacum* (d'après Weng et al. 2012<sup>80</sup>).

Comme indiqué, outre une modification des propriétés catalytiques, les événements de néo-fonctionnalisation sont souvent accompagnés de modifications des propriétés structurales des enzymes (interactions entre sous-unités) et de l'expression des gènes. Sur ce dernier point, un exemple connu est celui de l'apparition du métabolisme des alcaloïdes pyrrolizidiniques impliqués dans la défense des plantes contre les herbivores. Ce métabolisme est lié à la néo-fonctionnalisation d'une déoxyhypusine synthase (DHS) en homospermidine synthase (HSS). La DHS catalyse la modification post-transcriptionnelle d'un facteur d'initiation (eIF5A) essentiel à la croissance de la plante et la HSS, la synthèse d'homospermidine, précurseur de tous les alcaloïdes de type pyrrolizidiniques (Fig. 36). Ces 2 enzymes utilisent la spermidine comme co-substrat. Alors que la DHS est exprimée de façon ubiquitaire dans les cellules en division, le gène codant la HSS est exprimé, de manière spécifique, dans les racines ce qui témoigne d'une modification de la régulation au cours du processus de néo-fonctionnalisation. De plus, il a été montré que chez 2 espèces proches appartenant aux Astéracées, la régulation transcriptionnelle pouvait être différentes. Ainsi, chez *Eupatorium cannabinum* (*Eupatorieae*), la HSS est exprimée de manière uniforme dans le parenchyme cortical de la racine mais pas dans l'endoderme et l'épiderme. Chez *Senecio vernalis* et *Senecio jacobaea* (*Senecioneae*),

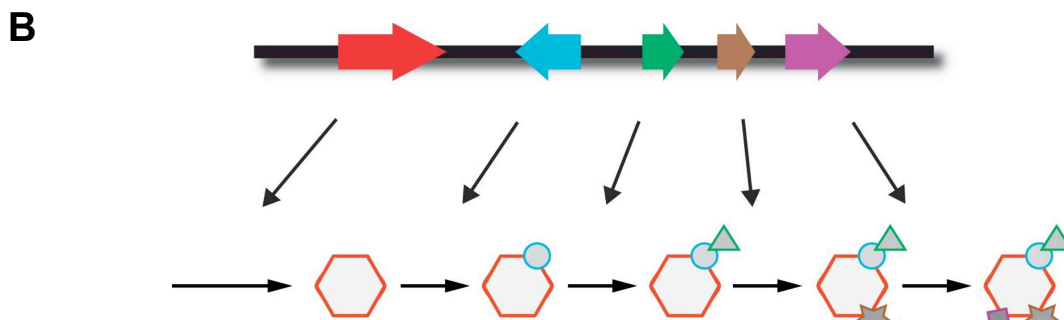
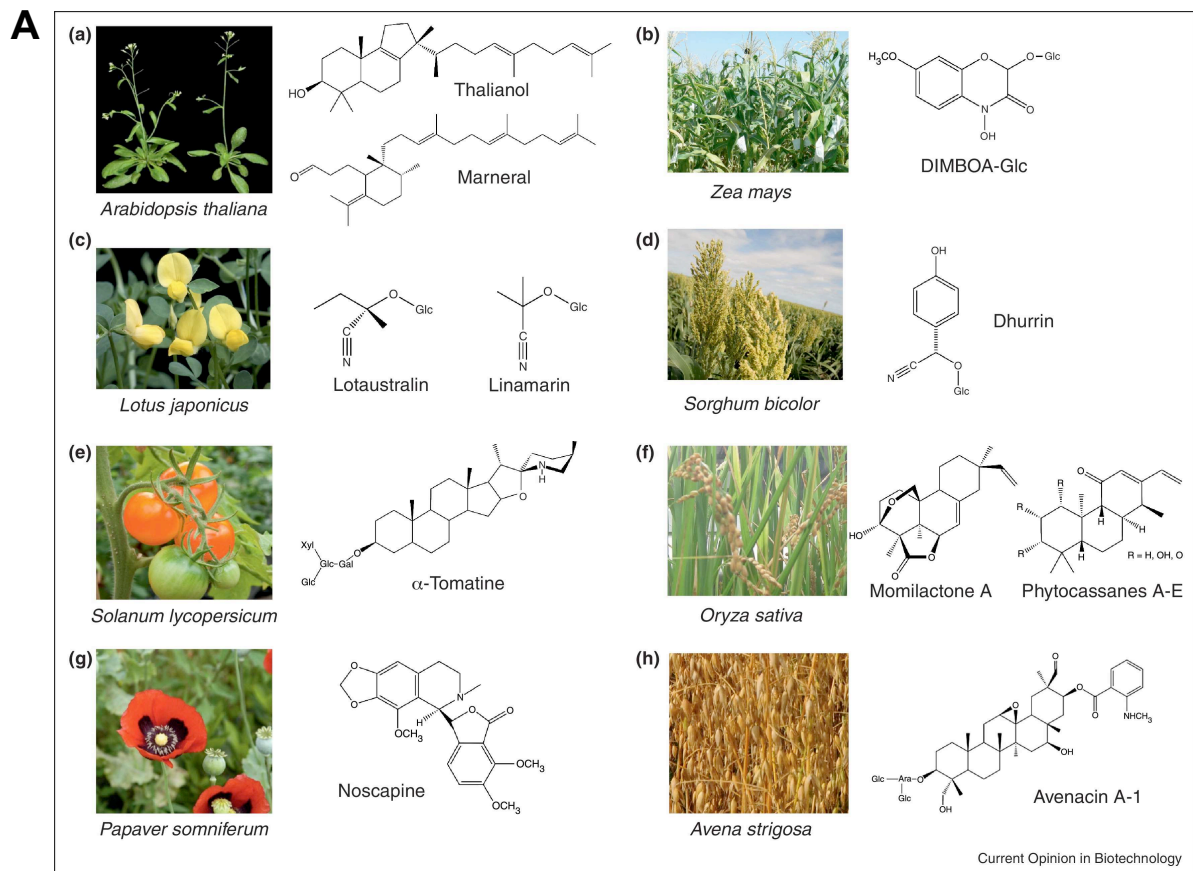
ce gène est exprimé dans un groupe de cellules spécialisées de l'endoderme et dans le parenchyme cortical des racines<sup>81</sup>.



**Figure 36.** Réactions catalysées par la déoxyhypusine synthase (DHS) et l'homospermidine synthase (HSS) chez les Astérocées (d'après Anke et al. 2004<sup>81</sup>).

Une autre originalité des enzymes du métabolisme spécialisé est que leurs gènes sont parfois organisés en cluster dans le génome. Les gènes codant tout ou partie d'une voie métabolique sont associés physiquement mais chaque gène possède son propre système de transcription ce qui différencie cette organisation des opérons des procaryotes. Cette caractéristique est connue depuis longtemps chez les champignons et les bactéries. Le premier exemple de ce type d'organisation des voies de biosynthèse des métabolites spécialisés a été décrite pour la première fois en 1997 chez le maïs<sup>82</sup>. Ce cluster permet la synthèse de 2,4-dihydroxy-7-méthoxy-1,4-benzoxazin-3-one (DIMBOA). Cette voie métabolique a été complétée plus tard. Depuis plusieurs clusters de gènes impliqués dans la biosynthèse de métabolites spécialisés ont été décrits. Ces clusters ont été découverts chez plusieurs espèces de plantes appartenant à la fois aux Dicotés et aux Monocotés et participent à la synthèse de différentes classes de molécules notamment les terpènes, les alcaloïdes et les glycosides cyanogéniques (Fig. 37). Une caractéristique commune de ces clusters est la co-localisation dans le génome d'au moins 3 gènes non homologues impliqués dans la même voie métabolique<sup>83</sup>. L'un des gènes code l'enzyme, dite enzyme signature, qui définit le squelette de base de la molécule et un nombre variable de gènes codent les enzymes responsables de la modification de ce squelette (Fig. 37)<sup>84</sup>. Les gènes signatures au sein de ces clusters de gènes ont émergé à partir de gènes du métabolisme primaire par duplication de gènes et néo-fonctionnalisation<sup>85</sup>. Le gène signature amorce alors la formation du groupe de gènes par le recrutement de gènes supplémentaires afin de constituer une voie métabolique complète. Les clusters décrits dans la littérature ont une taille de 35 à 270 kb et regroupent de 3 à 10 gènes. La majorité des gènes contenus dans ces clusters sont co-régulés ce qui permet la coordination spatio-temporelle de l'expression d'une voie métabolique complète. Pour expliquer l'avantage évolutif de cette organisation en clusters, 2 hypothèses sont formulées. Tout d'abord, comme évoqué, ceci permet une coordination fine de l'expression de ces gènes impliqués dans la même voie métabolique. Bien que non contrôlés par les mêmes promoteurs, la structure dynamique de la chromatine pourrait permettre une régulation concertée de tous ces gènes proches physiquement<sup>84</sup>. D'autre part, cet assemblage particulier permettrait une transmission à la descendance plus

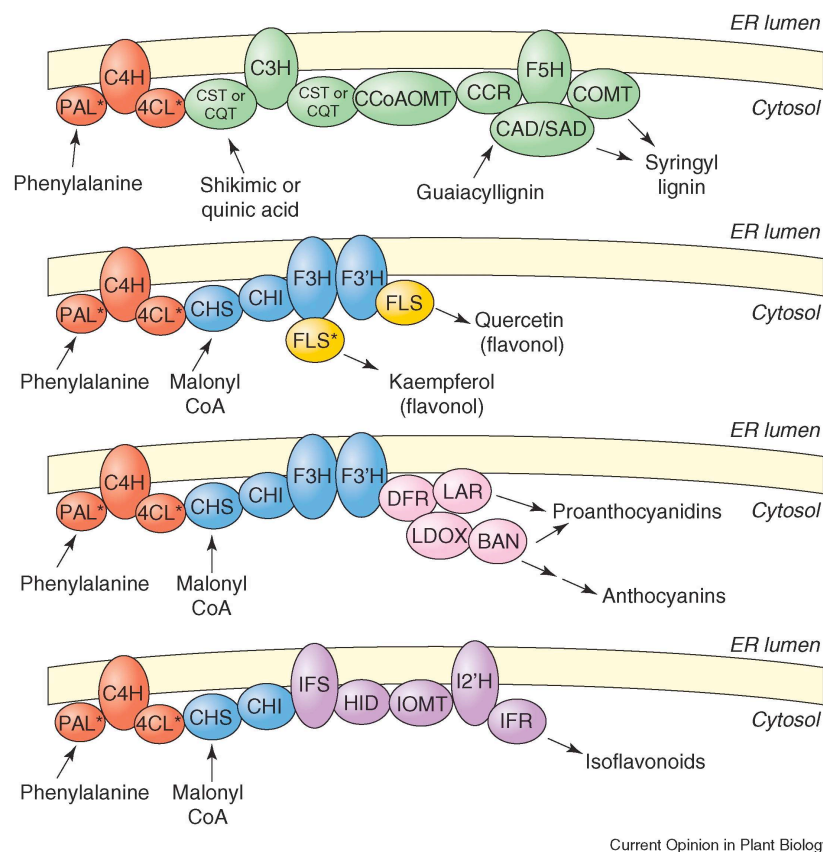
efficace de l'ensemble des gènes associés à une voie métabolique. Ceci est particulièrement important lorsque des intermédiaires de ces voies sont toxiques.



**Figure 37.** Cluster de gènes impliqués dans la synthèse de métabolites spécialisés chez les plantes. A : Exemples de métabolites spécialisés dont la synthèse est catalysée par des enzymes codées par des gènes organisés en clusters (d'après Nützmann et Osbourn, 2014<sup>83</sup>). DIMBOA-Glc : 2,4-dihydroxy-7-méthoxy-1,4-benzoxazin-3-one. B : Principe d'assemblage en cluster de gènes dont les produits participent à la biosynthèse d'un métabolite spécialisé. En rouge, le gène qui code l'enzyme qui catalyse la synthèse du squelette de base. Les autres gènes codent des enzymes permettant la modification de ce squelette (d'après Nützmann et Osbourn, 2013<sup>83</sup>).

Bien que non spécifique du métabolisme spécialisé, un certain nombre d'exemples indiquent une organisation en métabolons (complexes multi-enzymatiques) des enzymes de ce métabolisme. Pour optimiser la croissance et le développement des plantes, l'activité métabolique est hautement coordonnée à l'échelle de la plante entière, des organes, des tissus, des cellules, des organelles et au niveau moléculaire. Au niveau cellulaire, l'approvisionnement efficace en substrats des enzymes est favorisé par la compartimentation des cellules en différents organites et en sous structures de ces derniers (micro-compartimentation). Ceci permet un rapprochement physique des enzymes et de leurs substrats et ainsi une augmentation localisée de leurs concentrations. Au niveau moléculaire,

la formation de métabolons permet d'améliorer encore plus l'efficacité du métabolisme en facilitant la transmission des intermédiaires d'une enzyme à une autre, en augmentant localement la concentration en substrats et en évitant la fuite des intermédiaires vers d'autres voies métaboliques ou leur dégradation chimique lorsque ceux-ci diffusent dans la matrice cellulaire. La canalisation des flux d'intermédiaires permet également de protéger la cellule de la toxicité éventuelle de ceux-ci. De plus, les métabolons sont des structures dynamiques et leur assemblage/désassemblage permet de finement réguler le métabolisme en réponse aux contraintes environnementales ou à la demande physiologique de la plante, par exemple. Les métabolons s'organisent autour d'enzymes ancrées à la membrane, les cytochromes P450. Plusieurs exemples de voies métaboliques impliquées dans le métabolisme spécialisé et organisés en métabolons ont été décrits notamment pour la biosynthèse des composés phénoliques (Fig. 38).

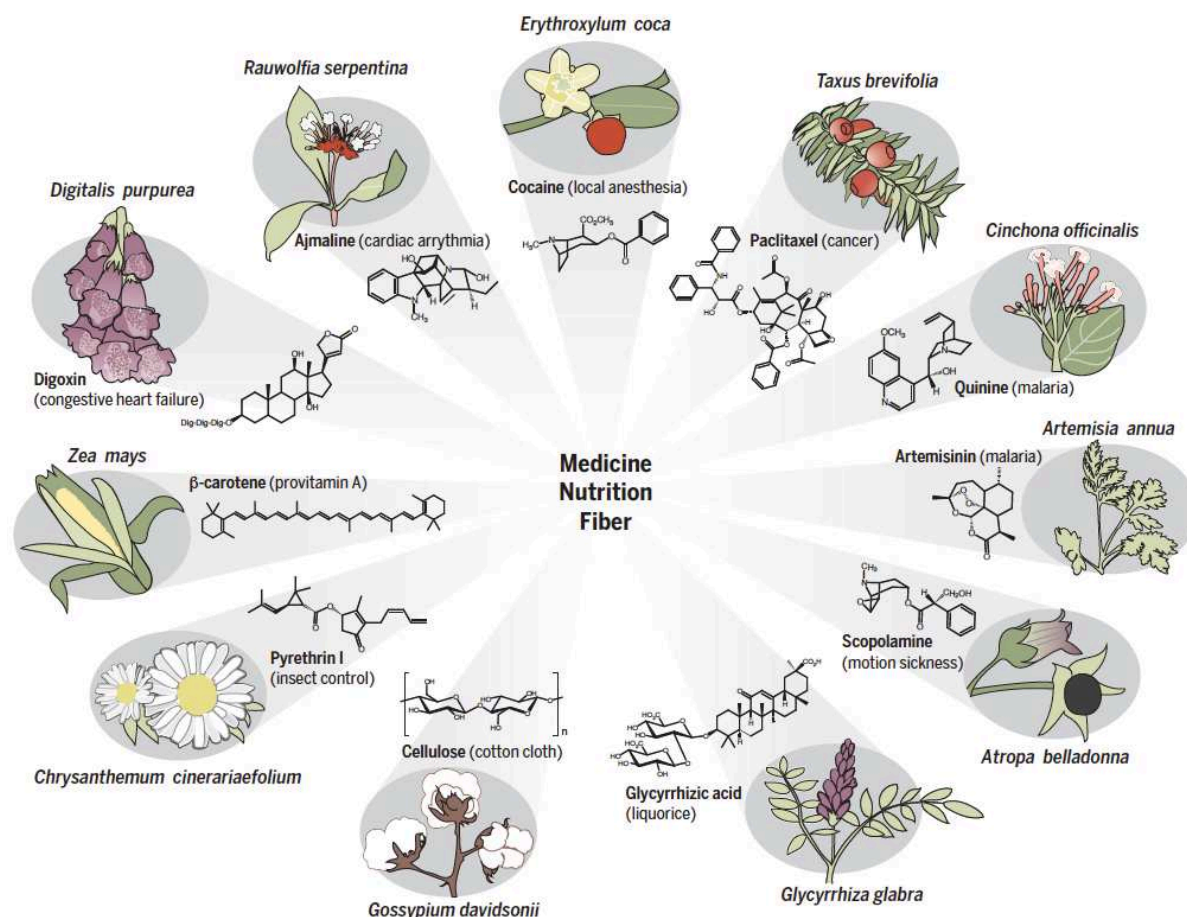


Current Opinion in Plant Biology

**Figure 38.** Quelques exemples d'enzymes impliquées dans la synthèse de composés phénoliques et organisées en métabolons (d'après Jorgensen et al., 2005)<sup>86</sup>.

Les différentes caractéristiques inhérentes au métabolisme spécialisé (promiscuité catalytique des enzymes, organisation des gènes en cluster, formation des métabolons) offrent des opportunités intéressantes pour les approches biotechnologiques.

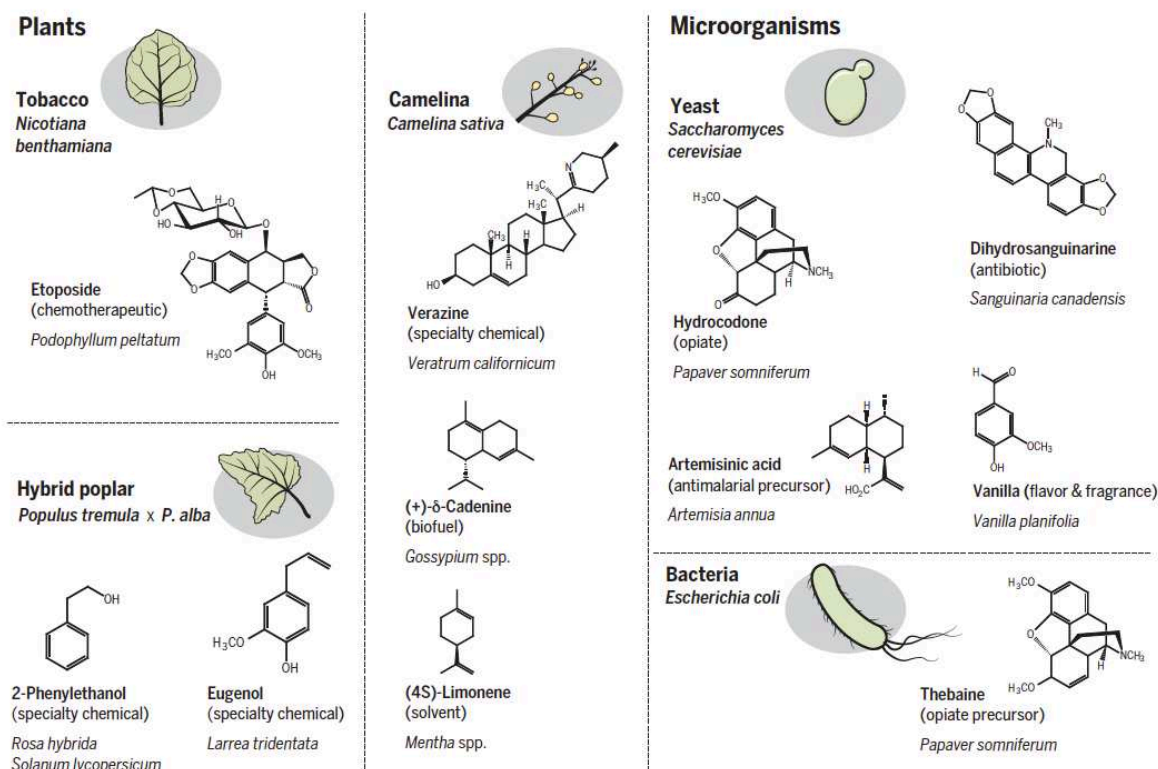
La diversité structurale des métabolites spécialisés du règne végétal en fait un réservoir intarissable de molécules bioactives largement utilisées notamment dans les domaines de la médecine, de la pharmacologie, de la cosmétique, de l'industrie ou encore de l'agronomie (Fig. 39). Parmi les molécules d'intérêt issues du métabolisme spécialisé, on peut citer les pyréthrinés (insecticide), le resvératrol qui entre dans la composition de nombreux produits cosmétiques (antioxydant), le taxol (anticancéreux), l'artémisinine (antimalarique), la codéine et la morphine (analgésiques). Ces quelques molécules ont une importance économique fondamentale. De nos jours, près de 25% des molécules utilisées en médecine sont issues des molécules d'origine naturelle.



**Figure 39.** Exemples de plantes accumulant des métabolites d'intérêt et leur utilisation (d'après Wurtzel et Kutchan, 2016<sup>87</sup>).

L'utilisation des métabolites spécialisés est souvent freinée par les difficultés d'approvisionnement. En effet, leur complexité structurale (nombreux centres chiraux) rend leur synthèse chimique très compliquée et peu rentable. De plus, ces composés sont souvent présents en très faible concentration dans la plante à l'état naturel et chez un nombre limité d'espèces. En effet, ces composés, souvent impliqués dans la défense de la plante, sont synthétisés au détriment d'autres molécules primaires participant à la croissance de la plante. Pour limiter le coût énergétique, leur biosynthèse dépend de mécanismes constitutifs ou inductibles. Lorsque leur synthèse est constitutive, leur accumulation est souvent très localisée soit dans des organes ou des tissus très spécifiques. Les tissus jeunes en développement sont souvent riches en ces molécules tandis que les organes matures en contiennent peu ce qui répond à la théorie de la défense optimale (optimal defense theory<sup>11</sup>). Lorsqu'elle est inductible, si les plantes ne sont pas cultivées dans des conditions particulières (exposition à des agresseurs par exemple), alors la concentration en molécules dans les tissus est très faibles. Ainsi, leur obtention par extraction est, le plus souvent, économiquement non viable et généralement couteuse sur le plan écologique (purification à partir d'une matrice complexe et utilisation de solvants organiques toxiques, mise en danger d'espèces végétales). De même, le potentiel de la majorité des molécules ne peut être évalué de par la faiblesse des quantités obtenues, insuffisantes pour mener à bien les tests de bioactivité<sup>88</sup>. Une valorisation optimale de ces molécules implique d'être capable de moduler les quantités disponibles. Que ce soit des démarches de sélection variétale (sélection assistée par marqueurs, édition de génome, transformation génétique) ou des stratégies d'ingénierie métabolique en système hétérologue procaryotes (bactéries) ou eucaryotes (levures, plantes) qui sont envisagées, la description de tous les déterminants moléculaires impliqués est un prérequis indispensable. Lorsque les voies métaboliques sont complètement établies, il est alors possible de faire produire les molécules d'intérêt ou leurs précurseurs en systèmes hétérologues. Plusieurs exemples ont déjà été décrits dans la littérature (Fig. 40). Dans le cas du métabolisme primaire, l'étude peut

être réduite à l'étude de plantes modèles tels que *Arabidopsis* pour les dicotylédones, le riz pour les monocotylédones ou le peuplier pour les ligneux. A l'inverse, le métabolisme spécialisé peut difficilement être appréhendé de manière exhaustive chez ces seules espèces. En effet, la distribution sporadique de tel ou tel composé au sein du règne végétal implique l'étude chez les espèces accumulatrices. De même, des phénomènes d'évolution convergente fréquents lorsque l'on parle de métabolites spécialisés renforcent la nécessité de s'intéresser à plusieurs espèces quand il s'agit d'élucider une voie de biosynthèse<sup>89</sup>.

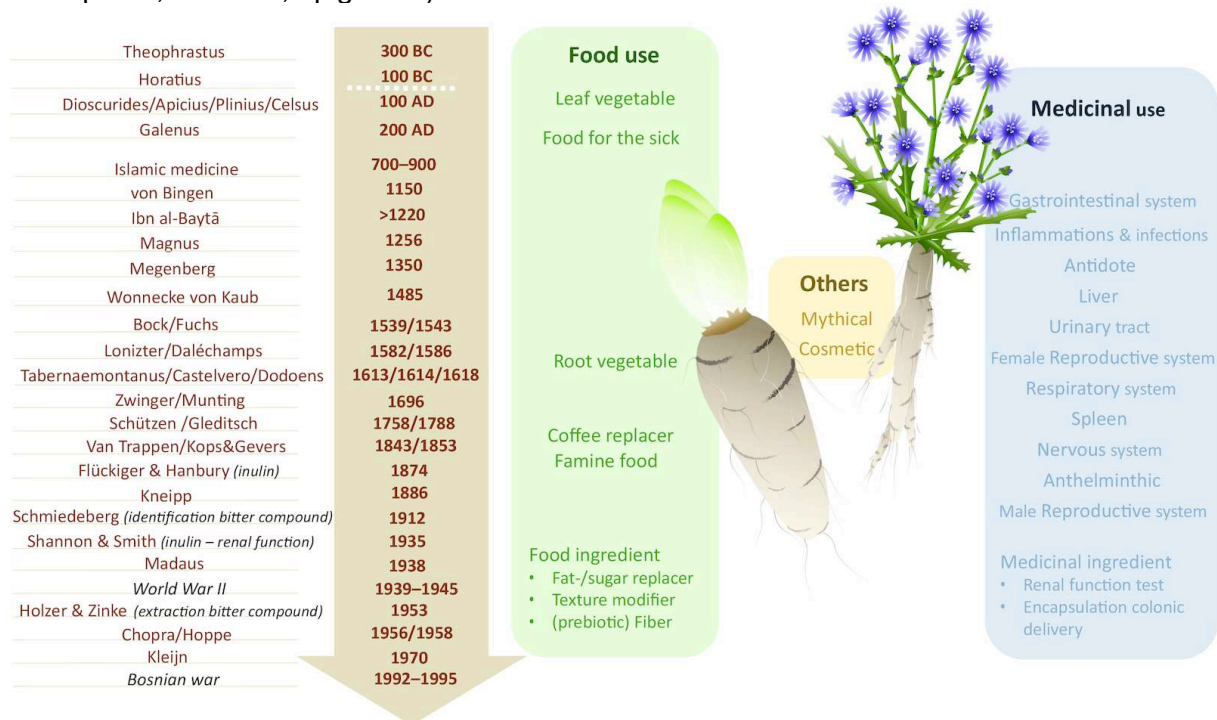


**Figure 40.** Exemples de métabolites spécialisés produits en système hétérologue (microorganismes ou plantes). Les organismes à partir desquels les gènes ont été isolés sont indiqués en italique sous les structures chimiques (d'après Wurtzel et Kutchan, 2013<sup>87</sup>).

Parmi ces métabolites spécialisés, les composés phénoliques possèdent des propriétés antimicrobiennes, antifongiques, antioxydantes et présentent donc potentiellement un intérêt pour l'industrie agroalimentaire, pharmaceutique ou dans le domaine de la protection des végétaux. Dans ce contexte, la chicorée représente un modèle régional pertinent.

La chicorée (*Cichorium intybus*), une Astéracée, est une espèce emblématique de la région Hauts-de-France qui comporte 4 variétés cultivées classées selon leur utilisation<sup>90</sup> : i) var. *foliosum*, utilisée pour la production d'endives, ii) var. *sylvestre* pour la production de salade (Trévise, Vérone, pain de sucre), iii) var. à fourrage et iv) var. *sativum* connue également sous le nom de chicorée industrielle ou chicorée à racines cultivée pour ses racines dont on extrait l'inuline et qui sont également utilisées, de manière plus minoritaire, pour la production de boissons après séchage ou de substitut de café après torréfaction. La France est le premier producteur européen de chicorée à café avec plus de 90% de la production localisée dans le nord du pays. *C. intybus* est principalement cultivée en Europe, en Asie, en Amérique du Sud, en Afrique du Sud et en Australie. Les variétés cultivées de chicorée sont bisannuelles alors que la chicorée sauvage est pérenne. Si la culture de chicorée est aujourd'hui majoritairement destinées à l'industrie agro-alimentaire, son utilisation en tant que plantes médicinales remonte quant à elle à l'Égypte ancienne, dans le bassin méditerranéen d'où elle est originaire.

La chicorée est utilisée depuis très longtemps en tant que plante médicinale, alimentaire ou pour d'autres applications (Fig. 41). Sur le plan pharmacologique, on lui attribue, entre autres, des activités antioxydantes, antibactériennes, vermifuges, antimalariques, gastroprotectrices, analgésiques, antitumorales, anti-inflammatoires, antidiabétiques ou encore hépatoprotectrices<sup>90</sup>. Parmi les composés bioactifs de la chicorée, des métabolites spécialisés ont été identifiés. Il s'agit notamment des lactones sesquiterpéniques (lactucine, lactucopricrine et 8-déoxylactucine et leurs dérivés dihydro-) et de composés phénoliques tels que la cichoriine (6,7-dihydrocoumarine 7-glucoside) et les flavonoïdes (quercétine, kaempférol, lutéoline, apigénine).

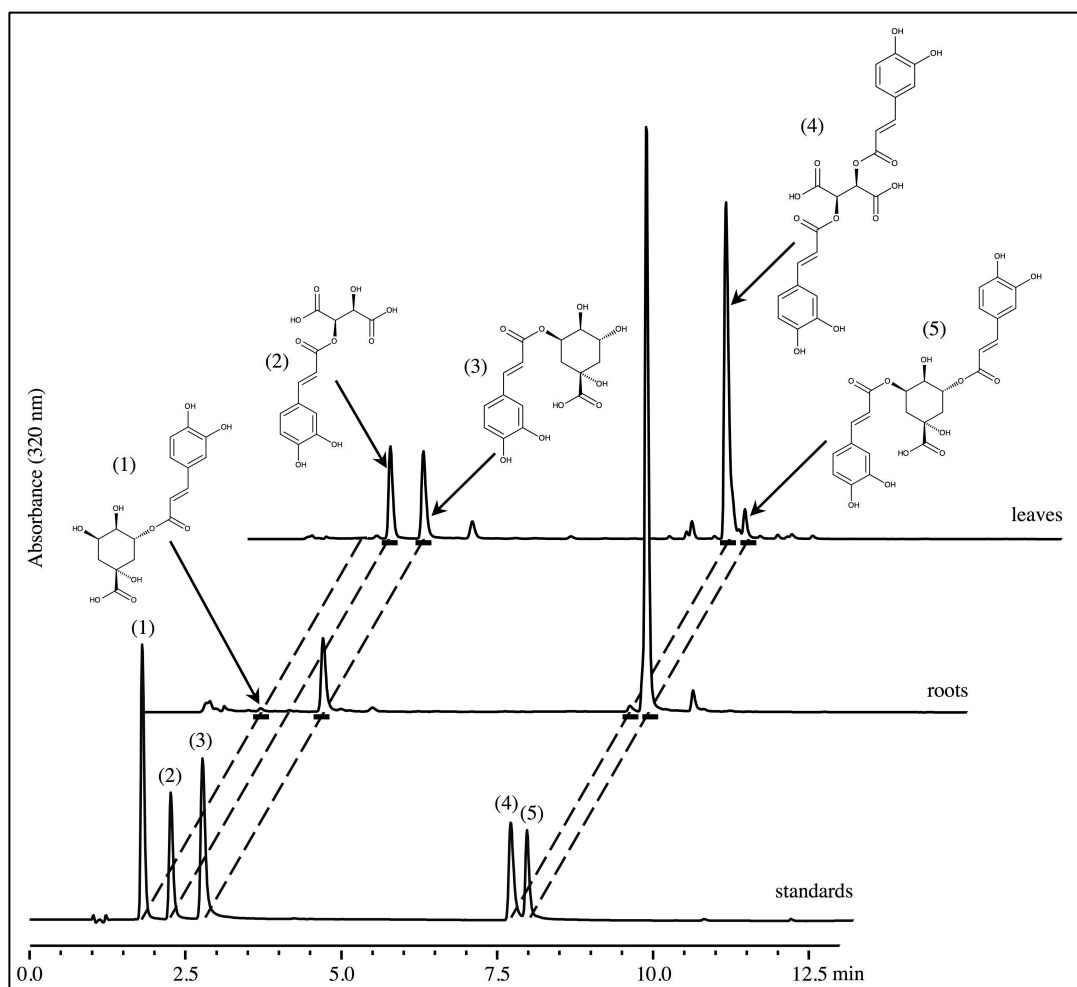


**Figure 41.** Utilisations de la chicorée au cours de l'histoire. Ces utilisations concernent à la fois les variétés sauvages et cultivées de *C. intybus* L. BC : before christ, AD : *anno domini* (d'après Puhlmann et de Vos, 2020<sup>91</sup>).

Le laboratoire s'intéresse depuis plusieurs décennies à cette plante pour répondre à différentes questions biologiques et a tissé des liens très étroits avec le monde de l'industrie notamment les Sociétés Florimond-Desprez (sélectionneur de chicorée industrielle, entre autres) et Leroux (torréfacteur de chicorée).

Nous avons au laboratoire mis en évidence la capacité de la chicorée à co-accumuler une panoplie originale d'esters d'acide caféique : les acides chlorogénique (acide 3-caféoylquinique, CQA), isochlorogénique (acide 3,5-dicaféoylquinique, diCQA), caftarique (acide caféoyltartrique, CTA) et chicorique (acide dicaféoyltartrique, diCTA)(Fig. 42)<sup>92</sup>. Ce dernier peut représenter jusqu'à 75% des composés phénoliques solubles de la feuille<sup>93</sup>. Le CQA, le diCQA et autres isomères de position (acide 5-caféoylquinique et acide 1,4-dicaféoylquinique par exemple) sont appelés communément les acides chlorogéniques. Si les rôles physiologiques de ces quatre dérivés d'acide caféique sont mal connus, ils sont cependant de plus en plus étudiés en raison de leur potentiel thérapeutique, très souvent lié à leur pouvoir antioxydant. Ces molécules sont notamment décrites pour leurs activités hépatoprotectrice, neuroprotectrice, antiproliférative, anti-inflammatoire, antidiabétique ou encore antivirale<sup>94-101</sup>. Ces activités offrent des opportunités intéressantes pour le traitement de maladies. Par exemple, du fait de leur activité neuroprotectrice largement décrite dans la littérature, le CQA et le diCQA sont envisagés pour le traitement de la maladie d'Alzheimer<sup>102-104</sup>. De plus, le diCQA et d'autres acides chlorogéniques ont également des propriétés antivirales (virus de l'immunodéficience humaine (HIV), virus de la grippe A, virus de l'herpès, les adénovirus, virus coxsackie B3 et le virus respiratoire syncytial)<sup>105-107</sup>. Le diCQA a également été identifié comme une molécule hypotensive en inhibant l'angiotensin I-converting enzyme (ACE). Cette activité

inhibitrice pourrait constituer une solution pour la protection des cellules pulmonaires infectées par le SARS-CoV-2 car le virus utilise une angiotensin II converting enzyme (ACE2), une protéine homologue de ACE, comme point d'entrée dans la cellule, ce qui modifie l'équilibre ACE/ACE2<sup>108</sup>.

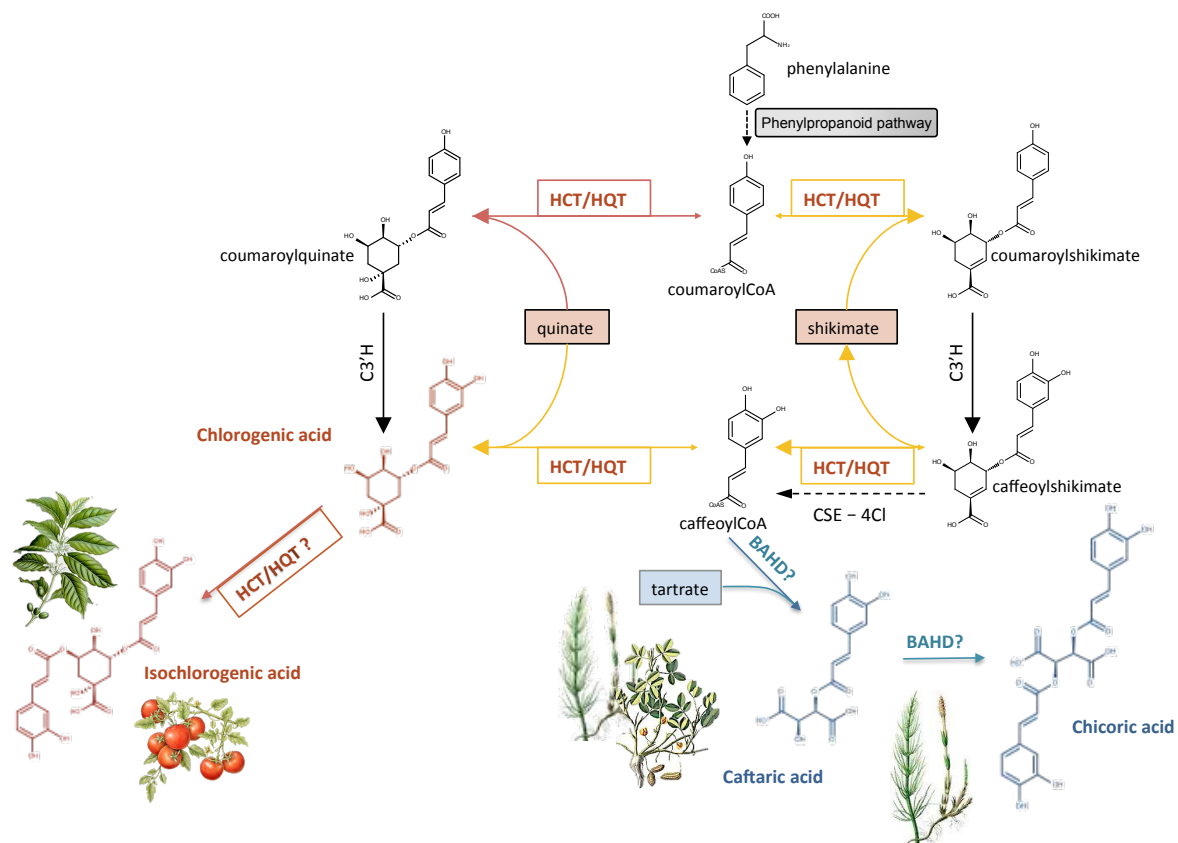


**Figure 42.** Chromatogrammes (HPLC-UV) d'extraits méthanoliques de feuilles et de racines de plantules de chicorée et d'un mélange de standards commerciaux. Les structures chimiques des composés phénoliques sont indiquées. (1) acide néochlorogénique (5-CQA), (2) acide catfarique, (3) acide chlorogénique (3-CQA), (4) acide chicorique, (5) acide isochlorogénique (3,5-diCQA)(d'après Legrand et al., 2016<sup>92</sup>).

Si la voie de biosynthèse du CQA est bien décrite chez plusieurs espèces, celles des autres composés le sont beaucoup moins (Fig. 43)<sup>109-113</sup>. L'importance socio-économique de la chicorée dans la région Hauts-de-France associée à ces particularités chimiques et aux propriétés biologiques des molécules produites font de la chicorée un modèle pertinent pour répondre à cette problématique. Le CQA est obtenu à partir de la phénylalanine issue de la voie du shikimate *via* la voie des phénylpropanoïdes (Fig. 32). Ces réactions impliquent, entre autres, 2 acyltransférases : HCT (hydroxycinnamoyl-CoA :shikimate/quinatate hydroxycinnamoyltransferase) et HQT (hydroxycinnamoyl-CoA :quinatate/shikimate hydroxycinnamoyltransferase). Ces enzymes diffèrent du point de vue de leur affinité de substrats. HCT est plus affine pour le shikimate tandis que HQT préfère le quinate. Cependant, les 2 acides sont utilisés par ces enzymes. Les mécanismes réactionnels responsables de la synthèse des autres composés ont été décrits plus ou moins partiellement chez d'autres espèces accumulatrices (Fig. 43). Le CTA serait synthétisé selon un mécanisme semblable à celui du CQA. Une étude menée chez la prêle a ainsi montré que la synthèse de CTA s'effectue à partir de l'acide tartrique auquel est ajouté un caféoyl-CoA par l'action de la CTT (hydroxycinnamoyl-CoA :meso-tartrate hydroxycinnamoyl transferase)<sup>114</sup>. Le diCQA est obtenue par transfert du groupement caféoyle d'un caféoyl-CoA sur un CTA par une CCT

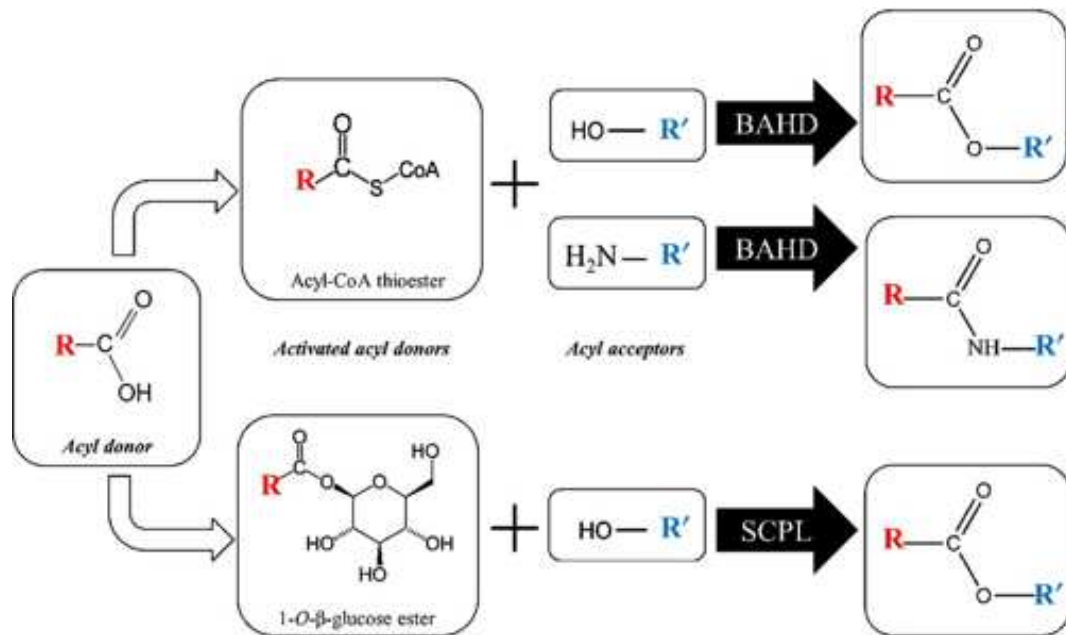


(hydroxycinnamoyl-CoA :cafféoyl-meso-tartrate hydroxycinnamoyltransferase)<sup>114</sup>. De même, une réaction identique catalysée par une HTT (hydroxycinnamoyl-CoA :tartric acid hydroxycinnamoyl transferase) permet la synthèse de CTA chez l'arachide<sup>115</sup>. Ces activités enzymatiques ont été décrites *in vitro* à partir d'extraits protéiques complexes ou de protéines purifiées mais les gènes impliqués n'ont pas été identifiés. La synthèse de diCQA serait, quant à elle, réalisée à partir de caféoyl-CoA et de CQA par une HCT chez le café et à partir de CQA seul grâce à une HQT chez la tomate<sup>116,117</sup>. Cependant, ces activités décrites *in vitro* à partir de protéines recombinantes restent très faibles et l'implication de ces enzymes dans la production de diCQA *in planta* reste à démontrer. Une activité chlorogénate :chlorogénate transférase a également été mise en évidence chez la patate douce (*Ipomoea batatas*) mais l'enzyme responsable, bien que purifiée, n'a pas été caractérisée au niveau moléculaire<sup>118</sup>. Cette enzyme transférerait le groupement caféoyle d'un CQA sur un autre CQA pour former du diCQA. Sur la base de ces résultats, la biosynthèse des 3 esters d'acide hydroxycinnamique (diCQA, CTA et diCTA) chez la chicorée impliquerait l'addition successive *via* une réaction d'acylation, d'un ou plusieurs acide(s) caféique(s) sur une molécule d'acide tartrique ou sur une molécule d'acide quinique.



**Figure 43.** Voies métaboliques impliquées dans la biosynthèse des acides hydroxycinnamiques chez les plantes. C3'H : p-coumaroyl ester 3' hydroxylase, CSE : caffeoylshikimate esterase, 4CL : 4-coumarate :CoA ligase, HCT : hydroxycinnamoyl-CoA :shikimate/quininate hydroxycinnamoyltransferase, HQT : hydroxycinnamoyl-CoA :quininate/shikimate hydroxycinnamoyltransferase, BAHD : membre de la famille BAHD.

Deux familles d'acyltransférases sont décrites comme intervenant dans le métabolisme spécialisé chez les plantes : les BAHDs (nommées d'après les quatre premières enzymes de cette famille caractérisées) et les SCPLs (Serine CarboxyPeptidase-Like)<sup>119</sup>. Ces réactions d'acylation impliquent l'utilisation d'un groupement acyle activé, riche en énergie, qui est transféré sur une molécule acceptrice. Les SCPLs et les BAHDs utilisent des esters de glucose ou des thioesters de CoA, respectivement, comme donneurs d'acyles pour réaliser des réactions d'estérification ou d'amidification (Fig. 44). Les réactions décrites précédemment suggèrent que les enzymes recherchées sont plus susceptibles d'appartenir à la famille des BAHDs, famille à laquelle appartiennent d'ailleurs les HCTs et les HQTs.



**Figure 44.** Réactions d'acylation catalysées par les enzymes de la famille des BAHDs ou des SCPLs (d'après Bontpart et al., 2015<sup>119</sup>).

La famille des BAHDs représente l'une des familles d'enzymes les plus vastes du métabolisme spécialisé (Tableau IV et Fig. 33). Selon les espèces, entre 50 et 120 gènes sont prédits pour coder des BAHDs et plus de 80 enzymes ont été caractérisées sur le plan biochimique<sup>120-124</sup>. Ces enzymes de masses comprises entre 45 à 55 kDa présentes 2 motifs peptidiques caractéristiques : le motif DFGWG situé près de l'extrémité C-terminale qui joue un rôle structural et le motif HXXXD en position médiane qui est directement impliqué dans l'activité catalytique de l'enzyme<sup>111,125-129</sup>. Sur la base des homologies de séquences peptidiques, les enzymes de cette famille sont classées en 5 clades<sup>120,122</sup>.

L'objectif des travaux initiés depuis mon recrutement était donc l'identification et la caractérisation des gènes impliqués dans la synthèse des esters d'acide caféique chez la chicorée industrielle. Considérant les caractéristiques biochimiques de ces BAHDs et les résultats rapportés chez d'autres espèces, c'est naturellement au sein de cette famille que ce sont portés nos efforts pour tenter d'identifier les acteurs responsables de la biosynthèse de ces esters d'acide caféique. Ces travaux ont permis d'identifier les gènes codant les enzymes impliquées dans la synthèse de CQA chez *C. intybus* (*CiHCT1*, *CiHCT2*, *CiHQT1*, *CiHQT2* et *CiHQT3*), d'identifier 2 gènes responsables de l'accumulation de composés originaux, des dérivés de spermine, dans la paroi du grain de pollen de *C. intybus* (*CiSHT1* et *CiSHT2*) et enfin d'identifier un gène qui code une enzyme catalysant la dernière étape de la synthèse du diCQA chez *I. batatas* (*IbICS*).

## 2. Bilan des activités de recherche

Depuis mon recrutement en 2009, mes travaux ont porté sur **l'identification et la caractérisation des déterminants moléculaires à l'origine de la synthèse des dérivés d'acide caféique chez la chicorée.**

Pour atteindre cet objectif, une démarche de génétique inverse a été entreprise. S'appuyant sur l'hypothèse selon laquelle les enzymes impliquées appartiendraient à la famille des BAHDs et plus exactement au clade Vb de cette famille (clade où l'on retrouve les HCTs et les HQTs), un criblage du génome a été effectué pour identifier ces séquences chez la chicorée. Cette recherche a permis d'identifier 3 HQTs et 2 HCTs impliquées dans la synthèse de CQA chez la chicorée. De même, 2 gènes annotés *CiSHT1* et *CiSHT2* ont été caractérisés ce qui a conduit à l'identification d'une activité originale en lien avec une diversification métabolique spécifique des Astéracées.

En parallèle, une approche biochimique a été entreprise pour identifier l'enzyme impliquée dans la synthèse de diCQA. Repartant des travaux antérieurs effectués sur la patate douce<sup>118</sup>, une protéine a été isolée et séquencée. A partir de criblage de banques d'ESTs, un gène, *IbICS*, a été cloné et une méthodologie de bioproduction de diCQA a pu être développée.

### 2.1. Développement de l'outil qRT-PCR chez la chicorée

Des outils génétiques ont été largement développés chez la chicorée. Des banques ESTs incomplètes étaient déjà disponibles à mon arrivée au laboratoire (banque publique notamment). Nous avons élargi ces banques et nous possédons aujourd'hui des banques ESTs assez exhaustives réalisées par les technologies de séquençage nouvelle génération (Illumina, 454) sur des échantillons très variés : racines, feuilles, cultures cellulaires de racines, appareil reproducteur. De plus, nous avons accès au génome de la chicorée *via* nos collaborations avec les partenaires industriels (Florimond-Desprez, Capelle en Pévèle, France).

Que ce soit pour valider l'implication d'un gène candidat dans un processus ou pour valider des résultats obtenus par les technologies d'analyse transcriptomique haut-débit, l'approche PCR quantitative en temps réel (qRT-PCR) s'est imposée comme une méthode de choix de par sa précision, sa rapidité et sa sensibilité. La fiabilité des résultats obtenus avec une telle approche dépend de plusieurs facteurs notamment la qualité des ARNs, le rendement de la synthèse des ADN complémentaires, ou encore l'efficacité de la réaction de PCR. Afin de standardiser les résultats obtenus et tenter de minimiser les biais introduits, le recours à des gènes de référence, choisis selon une procédure rigoureuse, est fondamental (MIQE Guidelines)<sup>130</sup>. Un gène de référence est un gène dont l'expression est stable dans les conditions expérimentales à l'étude. Le choix de ces gènes est très important et doit se faire dès lors qu'une nouvelle manipulation est envisagée (tissus différents, environnements de culture différents, stades développementaux différents...). Ces gènes de référence sont utilisés afin de normaliser les résultats obtenus ce qui permet de s'affranchir des variations dues au processus expérimental ne reflétant donc pas une réalité biologique.

La justesse des résultats de qRT-PCR repose donc en grande partie sur un travail préliminaire de sélection et de validation de gènes de références appropriés aux supports biologiques étudiés. Nous avons ainsi testé la stabilité des gènes les plus couramment utilisés chez les plantes dans différents tissus de chicorée<sup>131</sup>. Dix-huit gènes de référence potentiels ont été retenus en criblant les banques ESTs de chicorée par BLASTX avec des gènes orthologues d'*A. thaliana* classiquement utilisés comme gènes de référence (Tableau V)<sup>132</sup>. Il s'agit de gènes impliqués dans différents processus cellulaires tels que des facteurs de transcription, des enzymes ubiquitaires ou des éléments du cytosquelette. Sur la base des évaluations préliminaires (efficacité de l'amplification, spécificité des amorces, valeurs des Ct), 12 gènes ont été retenus. Leur expression a été mesurée sur différents types de tissus de chicorée (racines et feuilles d'âges différents de jeunes plantules et cultures cellulaires de racines

élicitées ou non au méthyljasmonate) et leur stabilité évaluée en utilisant 3 algorithmes (GeNorm, NormFinder, BestKeeper).

**Tableau V.** Description des gènes de référence candidats et comparaison avec les gènes orthologues d'*A. thaliana*. Les gènes retirés de l'analyse après filtrage préliminaire sont soulignés (d'après Delporte et al. 2015<sup>131</sup>).

Gene abbreviation	Accession number	Gene description	Arabidopsis ortholog	Arabidopsis blastX E-value
<i>UBQ10</i>	KP752078	Ubiquitin 10	At4g05320	6.00E-150
<i>EF1<math>\alpha</math></i>	KP752079	Elongation Factor 1- $\alpha$	At5g60390	0
<i>GAPDH</i>	DT210806.1	Glyceraldehyde 3-phosphate dehydrogenase	At1g13440	7.00E-172
<i>UBC9</i>	FL679405.1	Ubiquitin conjugating enzyme 9	At4g27960	7.00E-84
<i>ACT2</i>	KP752080	Actin 2	At3g18780	1.00E-31
<i>TIP41</i>	EH681285.1	TIP41 like protein	At4g34270	9.00E-105
<i>UBC</i>	EH709287.1	Ubiquitin-conjugating enzyme	At5g25760	8.00E-88
<i>PP2AA3</i>	EH675923.1	Protein phosphatase 2A subunit A3	At1g13320	1.00E-48
<i>bHLH</i>	KP752081	Transcription factor bHLH	At4g38070	5.00E-29
<i>CYP5</i>	EH688850.1	Cyclophilin 5	At2g29960	3.00E-78
<i>PROF</i>	DT2111339.1	Profilin	At2g19760	5.00E-57
<i>TUB <math>\beta</math>9</i>	KP752082	Tubulin $\beta$ -9	At4g20890	8.00E-87
<i>SAND</i>	KP752083	SAND family protein	At2g28390	0
<i>Clath</i>	EH684142.1	Clathrin adaptor complex subunit	At5g46630	0
<i>ACT7</i>	EH674624.1	Actin 7	At5g09810	0
<i>PP2AA2</i>	FL675704.1	Protein phosphatase 2A subunit A2	At3g25800	0
<i><math>\alpha</math>TUB</i>	EH691312.1	Tubulin $\alpha$	At5g19780	0
<i><math>\beta</math>TUB</i>	KP752084	Tubulin $\beta$ -3	At5g62700	0

Ce travail a permis d'établir une liste de gènes de référence potentiels pouvant être, par la suite, réévalués pour s'adapter aux conditions à l'étude (Tableau VI).

**Tableau VI.** Classement des gènes de référence établis après analyse par les 3 algorithmes pour les 2 types de modèles à l'étude (cultures cellulaires de racines et plantules).

Gene	Cell cultures					Seedlings				
	geNorm rank	Normfinder rank	BestKeeper rank	Total score	Overall rank	geNorm rank	Normfinder rank	BestKeeper rank	Total score	Overall rank
<i>TIP41</i>	3	2	3	8	1	4	6	6	16	4
<i>PP2AA2</i>	2	1	9	12	2	3	3	5	11	3
<i>UBC</i>	4	3	5	12	2	5	4	8	17	5
<i>SAND</i>	5	4	4	13	3	2	2	3	7	2
<i>PP2AA3</i>	1	9	6	16	4	7	7	7	21	6
<i>CYP5</i>	8	7	2	17	5	6	5	10	21	6
<i>UBQ10</i>	9	10	1	20	6	9	8	4	21	6
<i>ACT2</i>	6	5	10	21	7	11	11	/	33	/
<i>Clath</i>	7	8	8	23	8	1	1	1	3	1
<i>PROF</i>	10	6	7	23	8	8	9	9	26	8
<i><math>\beta</math>TUB</i>	11	11	/	33	/	10	10	2	22	7
<i>ACT7</i>	12	12	/	36	/	12	12	/	35	/

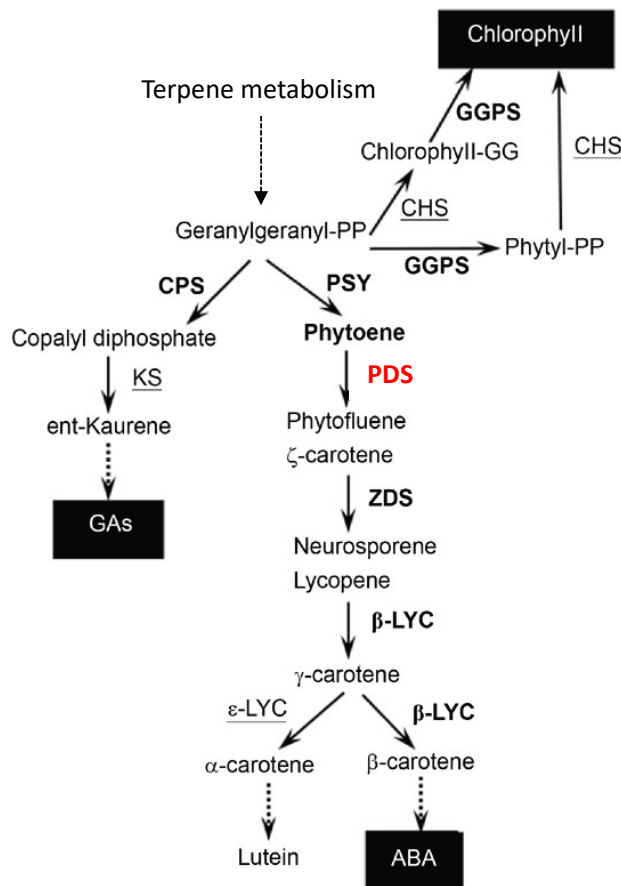
Ces gènes ont été utilisés à plusieurs reprises dans les publications du laboratoire visant à caractériser des gènes<sup>92,133</sup>.

Ce travail réalisé par une étudiante que j'ai encadrée (Marianne Delporte, stage de fin d'étude Ingénieur) en collaboration avec un étudiant en thèse que j'ai co-encadré (Guillaume Legrand) a fait l'objet d'une publication.

Delporte M., Legrand G., Hilbert J.L., Gagneul D., 2015. Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. *Front. Plant Sci.*, 6: 651.

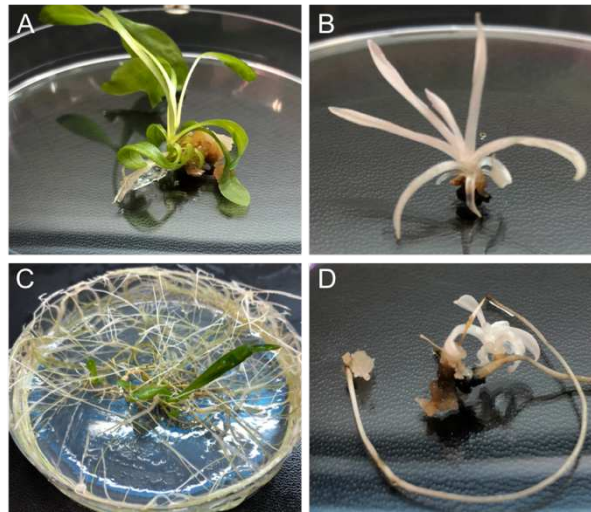
## 2.2. Edition du génome de chicorée par la technologie CRISPR/Cas9

L'édition du génome est un outil très important pour la validation de la fonction des gènes. Afin d'étayer la participation du produit d'un gène dans la synthèse d'un métabolite, il nous est apparu essentiel de disposer d'une telle approche chez la chicorée. Il existe de nombreuses techniques récentes qui permettent de générer des mutations ciblées dans le génome. Nous avons choisi d'utiliser la technologie CRISPR-Cas9 qui est la méthode la plus utilisée du fait de sa simplicité, du faible coût de mise en œuvre, de sa polyvalence et de sa haute efficacité. Pour la validation de cet outil, le gène *PDS* a été ciblé. Celui-ci code une phytoène désaturase. Cette enzyme intervient dans la voie de biosynthèse des caroténoïdes. La mutation du gène codant cette enzyme occasionne une augmentation de la quantité de phytoène, le substrat de cette enzyme, dans les feuilles ce qui a pour conséquence l'inhibition des enzymes impliquées dans la biosynthèse de la chlorophylle (Fig. 45)<sup>134</sup>. Il en résulte une plante dépourvue de chlorophylle et présentant un phénotype albinos ou de photoblanchiment, caractère facilement détectable.



**Figure 45.** Voie de biosynthèse de la chlorophylle, des caroténoïdes et des acides gibbérelliques (d'après Qin et al., 2009<sup>134</sup>).

Un orthologue de chicorée a été isolé (*CiPDS*). Deux modes de transformation ont été testés : transformation de protoplastes médiée par le PEG et agroinfiltration. Les 2 méthodes de transformation sont fonctionnelles chez la chicorée comme en témoigne les différentes lignées obtenues à partir de protoplastes ou de plantes agro-infiltrées (Fig. 46).



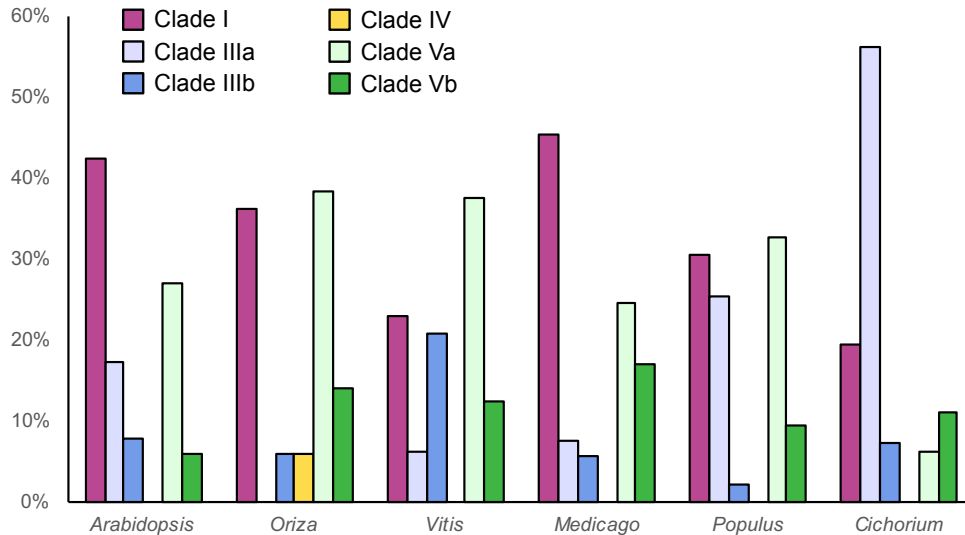
**Figure 46.** Phénotype de plantes dont le génome a été édité. A : Tige feuillée émergeant d'un cal sauvage (issu de protoplastes sauvages). B : Tige feuillée albinos émergeant d'un cal édité (issus de protoplastes transformés). C : Jeunes tiges émergeant de chevelus racinaires transformé avec la souche sauvage d'*A. rhizogenes* 15834. D : Tige albinos émergeant de chevelus racinaires transformés de manière à muter le gène *CiPDS* (d'après Bernard et al., 2019<sup>135</sup>).

Le protocole est établi et nous disposons donc d'un outil qui est utilisé pour invalider des gènes d'intérêt chez la chicorée. Ce travail d'optimisation a fait l'objet d'une publication.

Bernard G., Gagneul D., Alves Dos Santos H., Etienne A., Hilbert J.L., Rambaud C., 2019. Efficient genome editing using CRISPR/Cas9 technology in chicory. *Int. J. Mol. Sci.*, 20: 1155.

### 2.3. Identification et classement des membres de la famille BAHD chez la chicorée

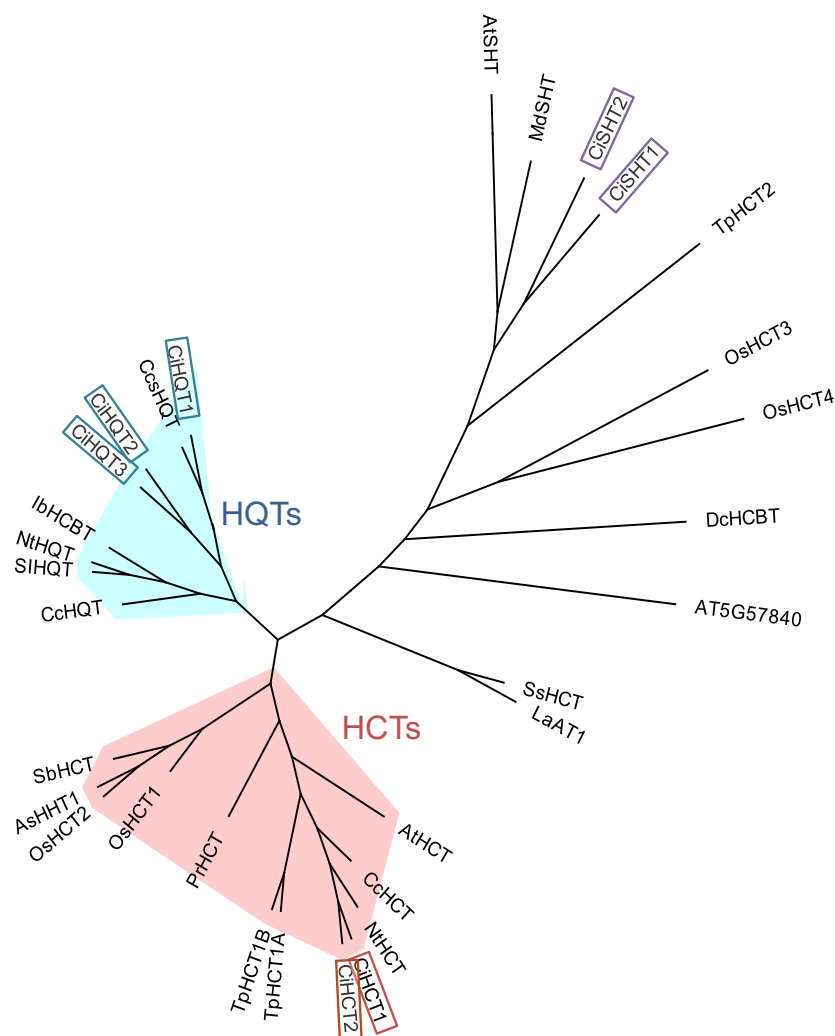
Afin d'identifier les BAHDs putatives de la chicorée, le criblage du génome a été effectué en se basant en premier lieu sur le motif DFGWG caractéristique des membres de cette famille (partie C-terminale de la protéine)<sup>120</sup>. Après un tblastn, les séquences nucléotidiques identifiées ont été individuellement parcourues à la recherche du motif HXXXD, impliqué dans l'activité catalytique de ces enzymes. La présence des 2 motifs ainsi que le poids moléculaire des protéines prédites (entre 40 et 60 kDa) ont permis de cibler les candidats potentiels. Cette recherche a permis d'identifier 81 séquences appartenant potentiellement à la famille des BAHDs. Ce nombre est légèrement supérieur à celui obtenu chez l'Artichaut (69 séquences identifiées) appartenant à la même famille botanique<sup>123</sup>. Chez *Arabidopsis*, *Oryza sativa*, *Vitis vinifera*, *Medicago truncatula* et *Populus trichocarpa*, 55, 84, 52, 50 et 74 séquences protéiques appartenant à cette famille ont été recensées<sup>122</sup>. L'analyse phylogénétique réalisée avec les séquences protéiques de chicorée et 80 séquences de BAHDs caractérisées biochimiquement a permis de répartir les 81 séquences protéiques identifiées dans 3 des 5 clades décrites précédemment chez d'autres espèces (clades I, III et V)<sup>120,122</sup> (Fig. 47). Le clade III et le clade V sont subdivisés en 2 sous-clades notés a et b. Le clade II tel que décrit dans d'autres travaux n'a pas été pris en compte dans notre étude, les enzymes qui ne présentent pas les caractéristiques structurales et catalytiques des BAHDs. Ce clade renferme des enzymes qualifiées de CER2-like impliquées dans l'élongation des acides gras à très longue chaîne et qui présentent des motifs caractéristiques des BAHDs profondément modifiés. Le clade IV ne contient pas de protéines de chicorée. Cependant, dans de nombreuses études, ce clade se retrouve inclus dans le sous-clade Vb.



**Figure 47.** Répartition par clade des séquences de chicorée et de 5 autres espèces analysées par Tuominen et al. (2011)<sup>122</sup>.

L'analyse de la répartition des protéines au sein de ces clades a permis de dégager une caractéristique propre à la chicorée (Fig. 47). En effet, plus de 50 % des protéines identifiées chez la chicorée appartiennent au clade IIIa (46 séquences). Ce sous-clade inclut majoritairement des BAHDs impliquées dans la synthèse d'alcaloïdes tels que la vindoline, la thébaïne ou encore la cocaïne mais contient également des protéines responsables de la synthèse d'autres classes de molécules telles que des esters volatils ou l'hydroxycinnamoyl spermidine. Au sein de ce clade IIIa, 29 des 46 séquences forment un sous-groupe au sein duquel aucune des BAHD précédemment caractérisées n'est incluse. Des analyses plus complètes ont démontré que ce sous-groupe pourrait constituer un nouveau clade spécifique des Astéracées.

L'objectif de notre travail était de rechercher des gènes candidats susceptibles de catalyser des réactions du même type que les HCTs et les HQTs. Ces enzymes appartiennent au sous-clade Vb qui contient des enzymes spécifiques du métabolisme des acides hydroxycinnamiques. Au sein de ce sous-clade, sept séquences ont été identifiées chez la chicorée et nous avons naturellement entrepris leur caractérisation fonctionnelle (Fig. 48).



**Figure 48.** Extrait de l'arbre phylogénétique des BAHDs correspondant au clade Vb. Les 7 séquences de chicorée sont encadrées. Leur nom a été donné à la suite des analyses présentées dans la suite de ce manuscrit.

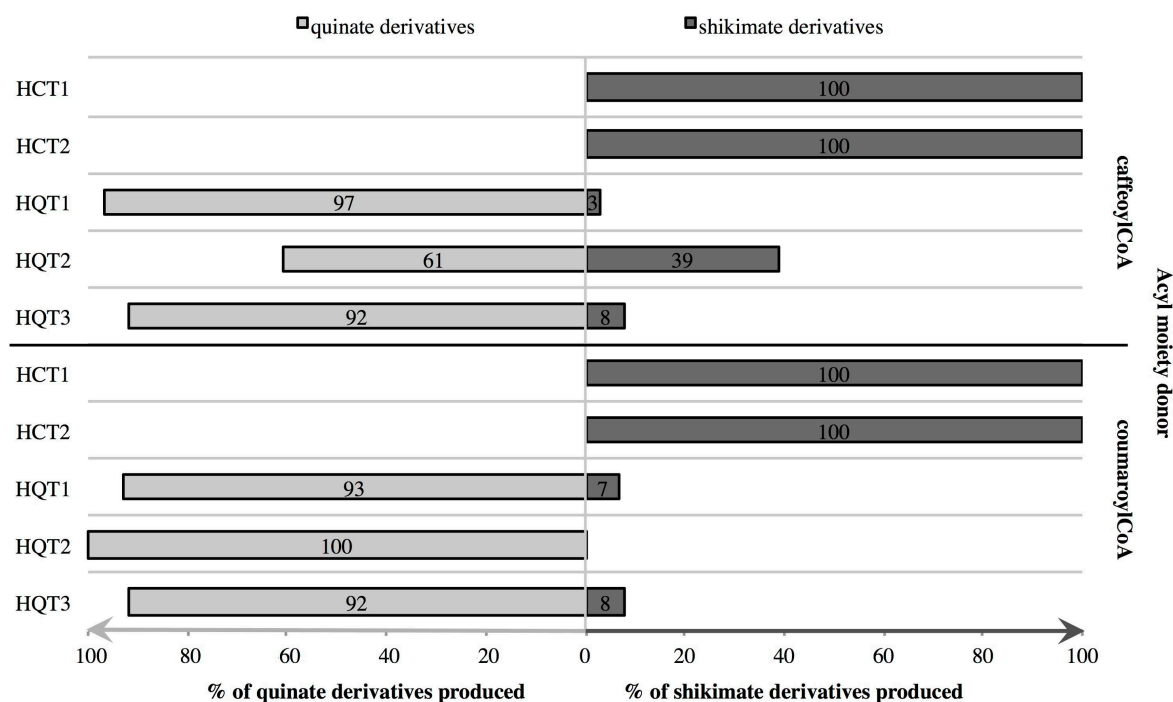
## 2.4. Caractérisation de la voie de biosynthèse du CQA chez la chicorée

Les analyses phylogénétiques conduites sur les BAHDs de chicorée ont permis d'identifier, dans le sous-clade Vb, 5 gènes prédits pour coder soit des HCTs (CiHCT1 et CiHCT2) soit des HQTs (CiHQT1, CiHQT2 et CiHQT3) (Fig. 48). Ces annotations ont été réalisées sur la base de l'homologie des protéines avec soit des HCTs ou des HQTs d'autres espèces chez lesquelles ces enzymes ont été caractérisées biochimiquement. Cependant, en considérant notamment le nombre de séquences homologues identifiées (ce qui sous-entend une redondance fonctionnelle) et la versatilité catalytique des enzymes de ce sous-clade, une caractérisation approfondie de ces gènes et de leur produit a été réalisée<sup>92</sup>. Leur implication dans la biosynthèse du CQA et de ses dérivés a été démontrée par des approches *in vitro* (production de protéines recombinantes et réalisation de tests d'activités enzymatiques) et *in vivo* (expression transitoire chez *Nicotiana benthamiana*).

Les 5 gènes ont été introduits chez *E. coli* (gènes placés en aval d'une étiquette histidine) et les protéines produites ont été purifiées par chromatographie d'affinité. L'analyse de leurs propriétés catalytiques *in vitro* a permis, tout d'abord, de confirmer que ces 5 enzymes étaient capables d'utiliser le quinate et le shikimate comme accepteurs d'acyle et les esters de CoA comme donneurs d'acyle (cinnamoyl-CoA, coumaroyl-CoA, caféoyl-CoA et feruloyl-CoA) pour former soit des acylquinates soit des acylshikimates. Les HCTs de chicorée se sont révélées plus affines pour le shikimate que pour le quinate et, à l'inverse, les HQTs utilisent



préférentiellement le quinate comme substrat (Fig. 49). Ces résultats biochimiques confirment l'arbre phylogénétique et sont en accord avec d'autres travaux menés chez d'autres espèces.



**Figure 49.** Tests de compétition pour déterminer les accepteurs de groupement acyle préférés des HCTs et HQTs. Les protéines recombinantes ont été incubées avec du caféoyl-CoA ou du coumaroyl-CoA et avec de l'acide quinique et de l'acide shikimique présent simultanément à la même concentration. Après 1 h, les produits de la réaction ont été quantifiés et leur concentration relative calculée (% de produits obtenus, d'après Legrand et al., 2016<sup>92</sup>).

La capacité de ces enzymes à utiliser d'autres substrats tels que le tartrate ou le malate a également été examinée. L'hypothèse était que l'une de ces enzymes pourrait catalyser la synthèse de CTA chez la chicorée. Aucune activité de ce type n'a pu être mise en évidence. L'implication des CiHCTs ou des CiHQTs dans la biosynthèse du diCQA n'a pas pu être démontrée dans nos conditions expérimentales. L'expression transitoire de HCT1 et de HQT1 chez *N. benthamiana* a permis de montrer que cette modification génétique induisait une augmentation de la teneur en 3-CQA. L'ensemble de ces résultats confirme l'existence de 2 HCTs et de 3 HQTs chez la chicorée. Des travaux similaires réalisées chez l'artichaut et des analyses génomiques menées au laboratoire suggèrent que cette redondance fonctionnelle est une caractéristique retrouvée chez toutes les Astéracées<sup>123</sup>. Cette question est abordée dans la partie projet du manuscrit.

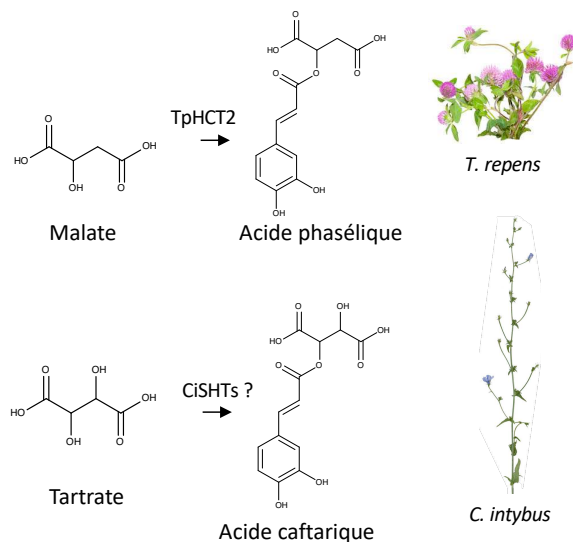
Ce travail a fait l'objet d'une publication dont je suis le dernier auteur et que 2 étudiants en thèse co-encadrés (Marianne Delporte et Guillaume Legrand) signent en premier et deuxième auteur. Le travail d'une étudiante encadrée dans le cadre de son master 2 (Chahinez Khelifi) est inclus dans ce manuscrit.

Legrand G., Delporte M., Khelifi C., Harant A., Vuylsteker C., Mörchen M., Hance P., Hilbert J.L., Gagneul D., 2016 Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.*, 7: 741.

## 2.5. Caractérisation de 2 enzymes impliquées dans la synthèse de phénolamides originales accumulées dans le manteau pollinique des Astéracées

L'analyse du clade Vb de famille BAHD de la chicorée a mis en évidence 2 séquences additionnelles annotées CiSHT1 et CiSHT2 qui partagent une similarité de séquence

importante avec les enzymes annotées SHT (MdSHT et AtSHT) et TpHCT2 (Fig. 48). TpHCT2 catalyse l'acylation du malate avec du caféoyl-CoA pour former l'acide phasélique (caféoylmalate) chez *Trifolium pratense*. Au vu de la similitude structurale entre l'acide phasélique et l'acide caftarique, les 2 protéines de chicorée identifiées constituaient des candidats intéressants pour la biosynthèse d'acide caftarique (Fig. 50). Cependant, les expériences menées en ce sens n'ont pas permis de démontrer une telle fonction.



**Figure 50.** Réaction catalysée par TpHCT2 et homologie structurale de l'acide phasélique et de l'acide caftarique ayant conduit à l'hypothèse que les CiSHTs proches sur le plan phylogénétique de TpHCT2 pourraient être impliquées dans la synthèse de l'acide caftarique chez la chicorée.

Les enzymes AtSHT et MdSHT, caractérisées chez *A. thaliana* et le pommier respectivement, correspondent toutes deux à des spermidine Hydroxycinnamoyl Transferases<sup>136,137</sup>. Ces enzymes catalysent la formation de trihydroxycinnamoyl spermidines, des phénolamides dont toutes les amines sont substituées, accumulées exclusivement au niveau du manteau du grain de pollen. Toutefois, des phénolamides partiellement substitués sont retrouvés dans tous les tissus des plantes. Les gènes *SHTs* sont exprimés spécifiquement dans le tapis des anthères. Il a été montré que l'accumulation de ces dérivés de spermidine était une caractéristique présente chez la quasi-totalité des Eudicotylédones en lien avec la présence d'enzymes de type SHT<sup>137</sup>.

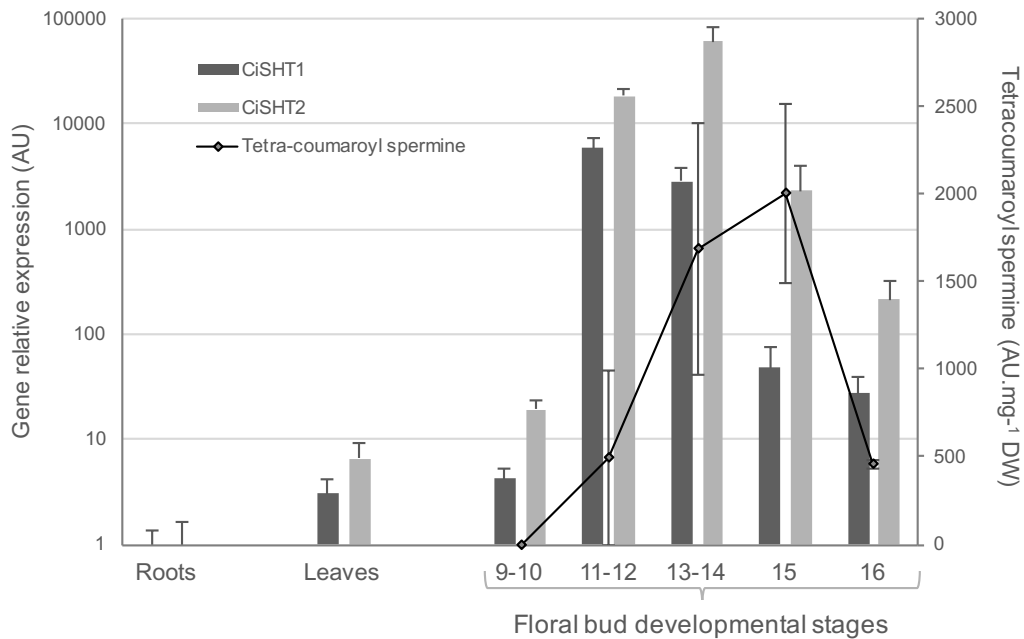
En collaboration avec Sébastien Besseau (BBV, Université de Tours) et Romain Larbat (LAE, Université de Lorraine), un article de revue reprenant ces éléments a été publié.

Roumani M., Besseau S., Gagneul D., Robin C., Larbat R., 2021. Phenolamides in plants: an update on their function, regulation, and origin of their biosynthetic enzymes. *J. Exp. Bot.*, in press.

L'hypothèse retenue était que les gènes, *CiSHT1* et *CiSHT2*, auraient une implication dans la synthèse de phénolamides accumulées dans la paroi du grain de pollen de la chicorée.

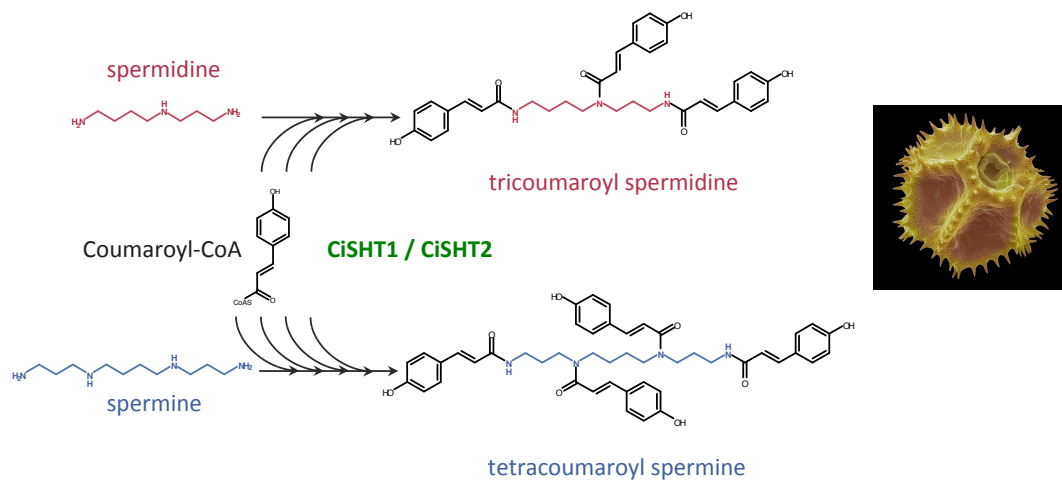
L'analyse des contenus en composés phénoliques dans les tissus reproducteurs de la chicorée a mis en évidence l'accumulation de tricoumaroyl spermidine mais celle-ci n'était pas la phénolamide majoritaire. Le composé majoritaire caractérisé et identifié par spectrométrie de masse était la N<sup>1</sup>,N<sup>5</sup>,N<sup>10</sup>,N<sup>14</sup>-tetracoumaroyl spermine, noté dans la suite du manuscrit tetracoumaroyl spermine (Fig. 51). Ce composé est spécifiquement accumulé dans les anthères et plus précisément dans la paroi du grain de pollen comme en témoigne l'analyse d'un « pollen wash » (Fig. 51). Cette extraction permet d'extraire spécifiquement les composés contenus dans la paroi du grain de pollen<sup>138,139</sup>. En plus, de la tetracoumaroyl spermine majoritaire et de la tricoumaroyl spermidine, l'analyse en spectrométrie de masse a révélé la présence, à l'état de traces, de N-coumaroyl(glucose)-N',N'',N'''-tricoumaroyl spermine.





**Figure 52.** Profil d'expression de CisHT1 et CisHT2 et d'accumulation de la tetracoumaroyl spermine dans des racines, des feuilles et des boutons de différents stades. L'expression des gènes a été déterminée par qRT-PCR.

Ces enzymes utilisent préférentiellement la spermine comme accepteur d'acyle et le coumaroyl-CoA comme donneur d'acyle (Fig. 53).



**Figure 53.** Voie de biosynthèse des dérivés de spermine et de spermidine accumulés dans la paroi du grain de pollen de la chicorée.

La diversification métabolique observée chez les Astéracées est donc liée à la néofonctionnalisation d'une enzyme de type SHT, elle-même vraisemblablement issue d'une HCT ancestrale, faisant potentiellement suite à une duplication de gène. L'ensemble de ces résultats concernant l'émergence de tetrahydroxycinnamoyl spermine chez les Astéracées ainsi que les processus biochimiques et évolutifs associés ont été publiés et l'originalité de ces résultats a été mise en avant dans un article paru dans le même journal<sup>144</sup>.

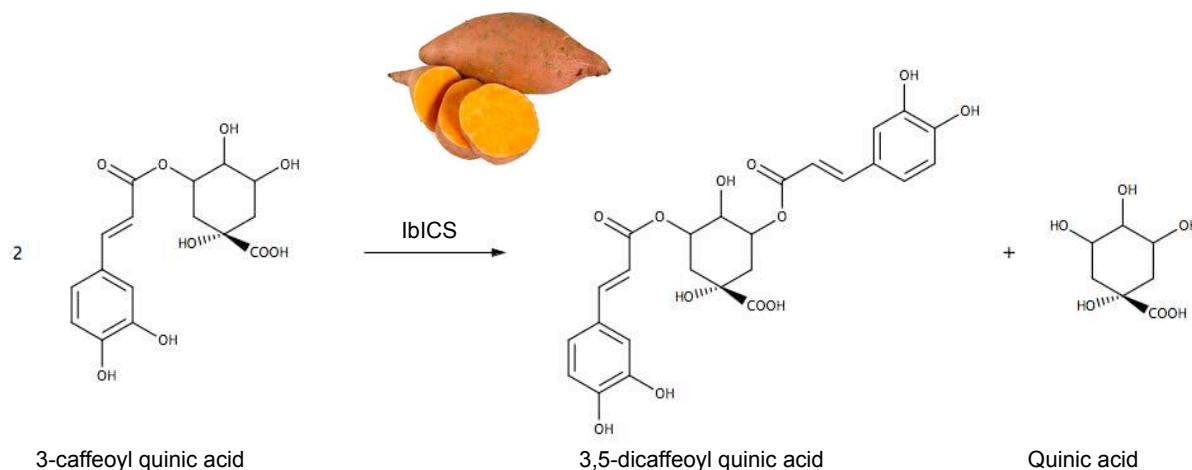
Delporte M., Bernard G., Legrand G., Hielscher B., Lanoue A., Molinié R., Rambaud C., Mathiron D., Besseau S., Linka N., Hilbert J.L., Gagneul D., 2018. A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in pollen coat of the Asteraceae family. *J. Exp. Bot.*, 69: 5365-5371.

A l'issue de ces travaux plusieurs questions restent en suspens. Le(s) rôle(s) des phénolamides du grain de pollen est (sont) à ce jour inconnu(s) de même que la signification évolutive du remplacement des dérivés de spermidine par des dérivés de spermine dans la paroi du grain de pollen des Astéracées. Ces molécules synthétisées dans le tapis des

anthères sont ensuite excrétées dans le locule des anthères puis déposées dans le manteau pollinique selon des mécanismes encore inconnus. Un transporteur de type ABC pourrait être impliqué pour le passage des phénolamides au travers de la membrane plasmique des cellules du tapis<sup>145</sup>. Il est également légitime de s'interroger sur la redondance fonctionnelle de ces 2 enzymes. Ce dernier point sera abordé dans la partie projet du manuscrit.

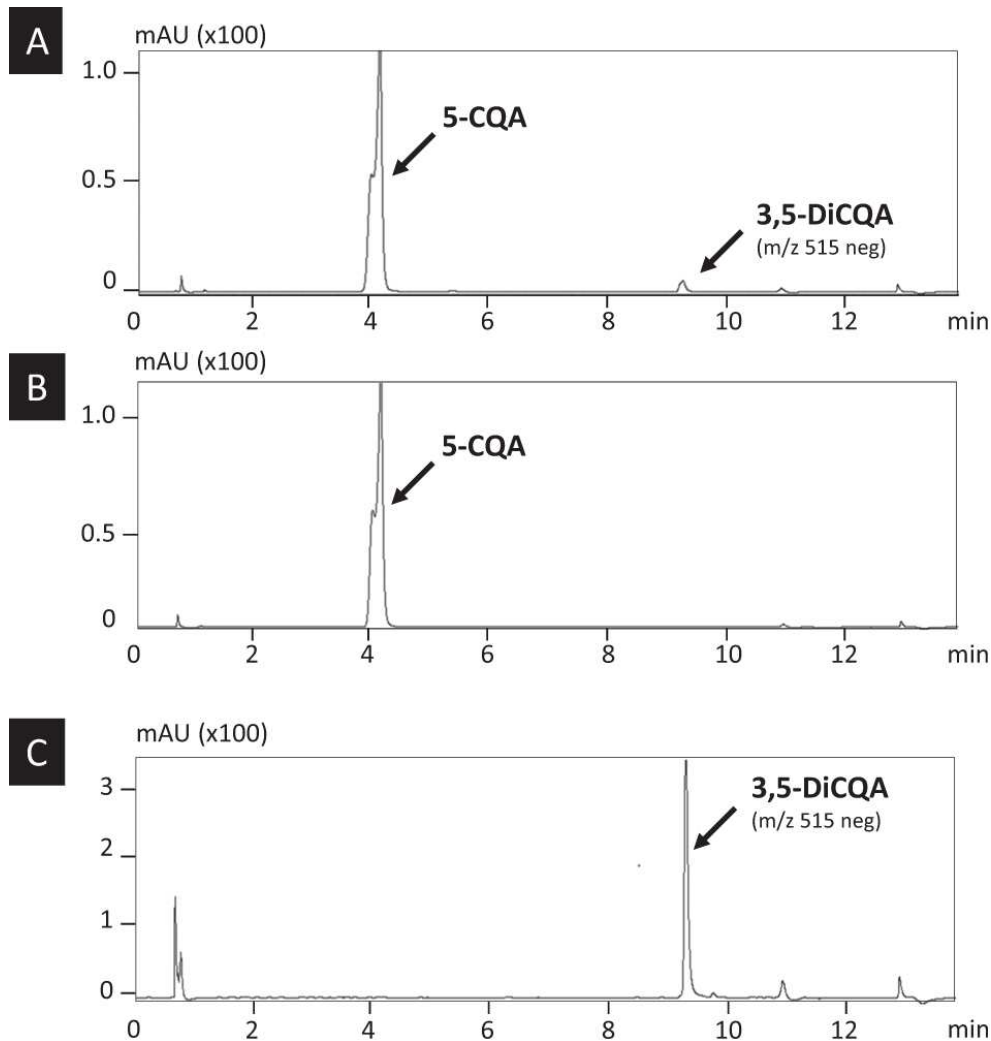
## 2.6. Caractérisation de la voie de biosynthèse du diCQA

Les travaux entrepris sur les membres du clade Vb de la famille des BAHDs chez la chicorée n'ont pas permis d'identifier un gène codant une enzyme impliquée dans la synthèse de diCQA. Des expériences réalisées *in vitro* à partir d'extraits protéiques brutes de chicorée attestent d'un mécanisme réactionnel identique à celui mis en évidence chez la patate douce (*I. batatas*, Fig. 54)<sup>118</sup>. Ainsi, cette enzyme utilise le CQA comme seul substrat. Le groupement caféoylé d'un CQA est transféré sur le quinate d'un deuxième CQA pour former le diCQA. Les approches de génétique inverse n'ayant pas permis d'identifier un gène codant une enzyme susceptible de catalyser ce type de réaction, nous avons entrepris une approche de recherche de gènes candidats par purification de protéines. Ce projet constituait l'activité principale de Guillaume Legrand que j'ai co-encadré (2011-2015).



**Figure 54.** Dernière étape de la synthèse du diCQA catalysée par IbICS chez la patate douce. Un mécanisme identique a pu être identifiée chez la chicorée à partir d'extraits protéiques brutes.

Afin de bénéficier d'un approvisionnement en matériel végétal non limitant pour ce type de stratégie, nous avons opté pour le tubercule de patate douce comme matériel de départ. Les tubercules ont été régulièrement achetés en supermarché. Afin d'induire la production de la protéine ou au moins en augmenter l'activité, les tubercules sont découpés en disque de 1 cm de diamètre et d'environ 5 mm d'épaisseur. Ces disques sont ensuite placés en incubation à 25°C à l'obscurité pendant 48 h sur papier filtre humidifié. Ce traitement permet l'augmentation de la teneur en diCQA et de l'activité enzymatique recherchée<sup>118</sup>. Les protocoles décrits par Villegas et al. (1987)<sup>118</sup> ont été utilisés et adaptés de manière à obtenir un extrait protéique enrichi. A l'issue d'un ensemble de séparations chromatographiques ayant permis l'obtention d'un extrait protéique enrichi mais complexe, l'activité recherchée a pu être détectée après séparation des protéines sur gel SDS-PAGE, découpage du gel en plusieurs fractions et extraction du gel. Les fractions présentant une activité significative ont été analysées par LC/MS et des peptides identifiés. En collaboration avec les équipes du Prof. Alain Hehn (LEA, Nancy, France) et du Prof. Frédéric Bourgaud (Plant Advanced Technology, Nancy, France), une séquence codante candidate notée *IbICS* pour IsoChlorogenate Synthase d'*I. batatas* a pu être identifiée. Son rôle dans la synthèse du diCQA à partir de CQA a pu être démontré *in vitro* et *in planta* (Fig. 55).



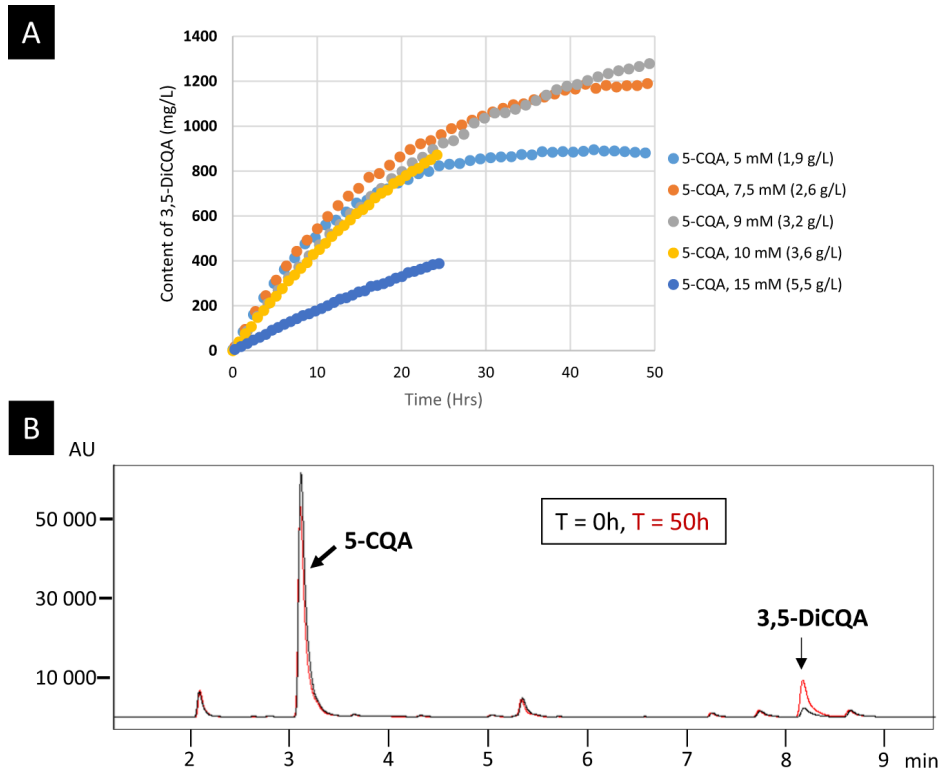
**Figure 55.** Analyses par HPLC des produits obtenus après incubation en présence de 5-CQA d'un extrait protéique de tabac (B) ou d'une enzyme recombinante purifiée, IbICS (produite en expression transitoire chez le tabac, A). C : standard commercial de 3,5-diCQA.

Cette enzyme appartient au sous-groupe SGNH de la famille des GDSL lipases/esterases. La famille des GDSL lipases/esterases contient des enzymes qui pour la plupart catalysent l'hydrolyse ou la synthèse d'esters. C'est le cas, par exemple, de la sinapine esterase (BnLIP2) de *Brassica napus* et de l'acetylajmalan esterase (RsAAE) de *Rauvolfia serpentina* qui catalysent l'hydrolyse de la sinapine et de l'acétylmorajmaline, respectivement<sup>146,147</sup>. D'autres membres catalysent des réactions d'acylation. C'est le cas de SICGT, une enzyme isolée chez *Solanum lycopersicum*, qui transfère le groupement caféoylé d'un CQA sur le glucarate pour former le caféoylglucarate<sup>148</sup>. IbICS constitue un exemple supplémentaire de GDSL lipase-like acyltransferases. IbICS comporte les particularités structurales de la famille des GDSL et du sous-groupe des SGNH (Fig. 56).

IbICS	1	MAAA-----APFSSLLGVLMLLASIGYAAA-----C-YS	AVFGF <b>GDSL</b> TDAGNLI	R
BnLIP2	1	MASSLKKLITSFLLFFFYTNIVASSEPSCR-----R-YK	SIISF <b>GDSI</b> ADTGNYL	H
RsAAE	1	MG-----FARLLHLVFSLLVFAGITNGLI-----CPFD	SIYQL <b>GDSF</b> SDTGNLI	R
SlCGT	1	MALGMRVVVLHLLISLLLLQLIKGDDDMVIKLEEPKLLKCGID	RIFQF <b>GDSL</b> SDTGNCLR	R
I ▼				
IbICS	46	LEPDGTVPHMYFPPYGETYFD	<b>KPTGR</b> CSDGRLIVDLIAQHYGLPLPPPSIPASLEGDKN-	
BnLIP2	51	LSDVNHPPQAAFLPYGETFFSVPTGRNSDGRLIIDFIAEFLGLPYVPPYF-----GSQNV		
RsAAE	46	LPPDGPTFTA AHFPYGETFFPTPTGRCS DGRLIIDFIATALNLP LLNPYL-----QQNV		
SlCGT	61	ESYCGAQTKTGKLPYGMNFYQ	<b>NATGR</b> CS DGFIILDYIAMECGLPLLNP SL-----EENA	
II				
IbICS	105	RLRAGVNFVAVGSRALDAQFYEQ	RIFDFTVTVNSMTDQLNWFKELLPSLC-SSPHDCKEL	
BnLIP2	106	SFEQGVNFVAVYGATALDRAFFIEKGI VSDFTNVSLSVQLNTFKQILPTLCASSRDCREM		
RsAAE	100	SFRHGVNFVAVAGATALDRSFLAARGVQVSDIHSHLSAQLNWFRTYLG SIC-STPKCSNK		
SlCGT	115	DFSHGVNFVAVSGATALSAEYLI SRDIAMSF TNSLSVQMRWMS S YFKSVC---SNDCAKY		
III				
IbICS	164	LDGSI	<b>FVLGE</b> FGGNDYTHSLLSGKGINVIRPFIPIVTRAIAQAVHELVELGAR	TVMIPSV
BnLIP2	166	LGDSII	LMGEIGGNDYNYPPFFEDKSINEIKELTPLIIKAI SDAIVDLIDLGGKTF LVPGS	
RsAAE	159	LKNALF	FILGNIGNNDVNYAF-PNRTIEEIRAYVPFITEAVANATREIIRLGGSRVIVPGI	
SlCGT	172	LENSIF	FLIGEIGGDDVTYVGFKQKPIEEVRRIVPDIVKNI IHSVRTVIGFGATRILVPGN	
IV				
IbICS	224	LPLGCAASYLTYYPSPNPDDYDEL-GCLIWVNEMASYHNQLLQQQLAHVRELHPYANIVF		
BnLIP2	226	FPGGCSAAYLTLFQTAKEEDYDPLTGCLPWLNDFGKHHDEQLKTEIKRLRKRYPHVNI IY		
RsAAE	218	FPIGCVARNLNF LNFFPDGDKDDL-GCLSSLNLSIYFNLSLFQRALASLSIEFPQAVI IY		
SlCGT	232	FPSGCFPIILTLYMNDSS TVYDEYH-CAEEWNNFTISYNLLQQS IHELNEEYPNIS IY		
V				
IbICS	283	ADIYNAAMEVYQGPDTYGFSGGAL-RACCGGGPYNFNTSAQCGNEGAT	-ACDNPSSYVN	
BnLIP2	286	ADYNSLYRLYQEPTKYGFKNRPL-AACCGVGGQYNF TIGE ECGYEGVG- YCQNPSEYIN		
RsAAE	277	ADYNAWRFLFRNGPALGSNSTSLLKCCCGIGGPYNDPDRECGSRGVE- VCPNP TQYIQ		
SlCGT	291	GDYYNAYWLLR NAVALGFNKKTLQISCCGIGGEYNYTESRRCGKPGAE	KACADPSSYLS	
VI				
IbICS	341	WDGYHLTE	EAAAYQCIT TGLLEG-----PYTYPRMKNLCRFDSRKT SRVAQI---	
BnLIP2	344	WDGYHLTE	EAAHQKMAHGILNG-----PYAAPAFNWSC-LDAASVDNESSFGS	
RsAAE	336	WDGTHFTQ	AAYRRVAEYVIPGI I KALKCSYSNIQPF LREGEGRQALRLNERE	
SlCGT	351	WDGSHLTO	KAYGWITKWLIDDILP-----QLNCRV	
VII				

**Figure 56.** Alignement de la séquence protéique déduite de IbICS avec des protéines de la famille GDSL lipase-like caractérisées biochimiquement. IbICS a été alignée avec BnLIP2 (AAX59709), RsAAE (AY762990) et SlCGT (FR667689). Le motif conservé GX<sub>3</sub>XXDXG est représenté en caractères gras. La séquence peptidique signal prédite de toutes les enzymes est indiquée en italique. Les blocs conservés de la famille SGNH (I, II, III et V) sont encadrés. Les acides aminés formant la triade catalytique sont indiqués par des flèches noires. Les séquences peptidiques identifiées par spectrométrie de masse après digestion à la trypsine et ayant permis d'identifier la séquence génique sont indiquées en rouge et en bleu.

Outre la caractérisation biochimique de cette protéine, un système efficace de bioproduction de diCQA par *Pichia pastoris* exprimant notre séquence candidate a été développé. La protéine modifiée de manière à être excrétée dans le milieu de culture permet la conversion de 60% du CQA commercial ajoutée. De même, l'ajout dans le milieu extérieur d'un extrait de café vert riche en CQA, en remplacement du CQA pur, a permis de multiplier par 7.5 la concentration en diCQA (Fig. 57).



**Figure 57.** Synthèse de 3,5-diCQA utilisant *P. pastoris* comme châssis. A. Cinétique d'accumulation du 3,5-diCQA en présence de différentes concentrations de 5-CQA commercial. B. Evaluation de l'efficacité de la conversion en présence d'un extrait de café vert contenant une forte concentration en 5-CQA.

Ce système est en cours d'optimisation pour être utilisé à l'échelle industrielle. L'ensemble de ces travaux a fait l'objet d'un dépôt de brevet (dépôt mars 2020) et d'une publication.

Miguel S., Legrand G., Duriot L., Delporte M., Menin B., Michel C., Olry A., Chataigné G., Salwinski A., Bygdell J., Vercaigne D., Wingsle G., Hilbert J.L., Bourgaud F., Hehn A., Gagneul D., 2020. A GDSL lipase-like from *Ipomoea batatas* catalyzes efficient production of 3,5-diCQA when expressed in *Pichia pastoris*. *Comm. Biol.*, 3: 673.

Miguel S., Bourgaud F., Salwinski A., Signe C., Hehn A., Hilbert J.L., Gagneul D., Legrand G., brevet déposé le 13 mars 2020. Procédé de production de l'acide 3,5-dicaféoylquinique (3,5-DCQ) par une levure (*Pichia pastoris*) exprimant une enzyme GDSL estérase/lipase recombinante d'*Ipomoea batatas*.



### 3. Projets de recherche

Les projets que je souhaiterais continuer à développer à court, moyen et long terme concernent :

- La **caractérisation des SHT-like**. Il s'agit, en particulier, d'approfondir les spécificités et caractéristiques catalytiques de ces enzymes en lien avec leur origine évolutive et de rechercher les rôles de leurs produits, les phénolamides.
- L'élucidation de la **voie de biosynthèse du diCQA** chez la chicorée. Cette étude permettra de déterminer si l'apparition de ce métabolisme est le résultat d'une évolution convergente.
- La caractérisation des **enzymes de la famille des acyltransférases BAHD appartenant au clade IIIa** et, en particulier, au sous clade mise en évidence au cours de nos travaux et spécifique des Astéracées.
- La caractérisation de la **voie de biosynthèse du CTA et du diCTA** chez la chicorée. Ces dérivés ont des propriétés biologiques intéressantes et la caractérisation de leur métabolisme est un prérequis indispensable à leur valorisation.
- La compréhension de l'apparente **redondance fonctionnelle des HCTs et HQTs** chez la chicorée.
- La caractérisation du **métabolisme des flavonoïdes C-glycosides** chez le lin. Cette étude, en lien avec d'autres approches conduites par des collaborateurs, permettra de déterminer si ces molécules contribuent à la tolérance au froid du lin.

#### 3.1. Caractérisation fonctionnelle et biochimique des SHT-like

Cet objectif s'inscrit dans le projet CHIC41H (Chicory for one health, Equipe mixte Laboratoire-entreprise, Entreprise Florimond-Desprez-Institut Charles Viollette). Je suis responsable de l'axe 3 qui concerne les Phénolamides.

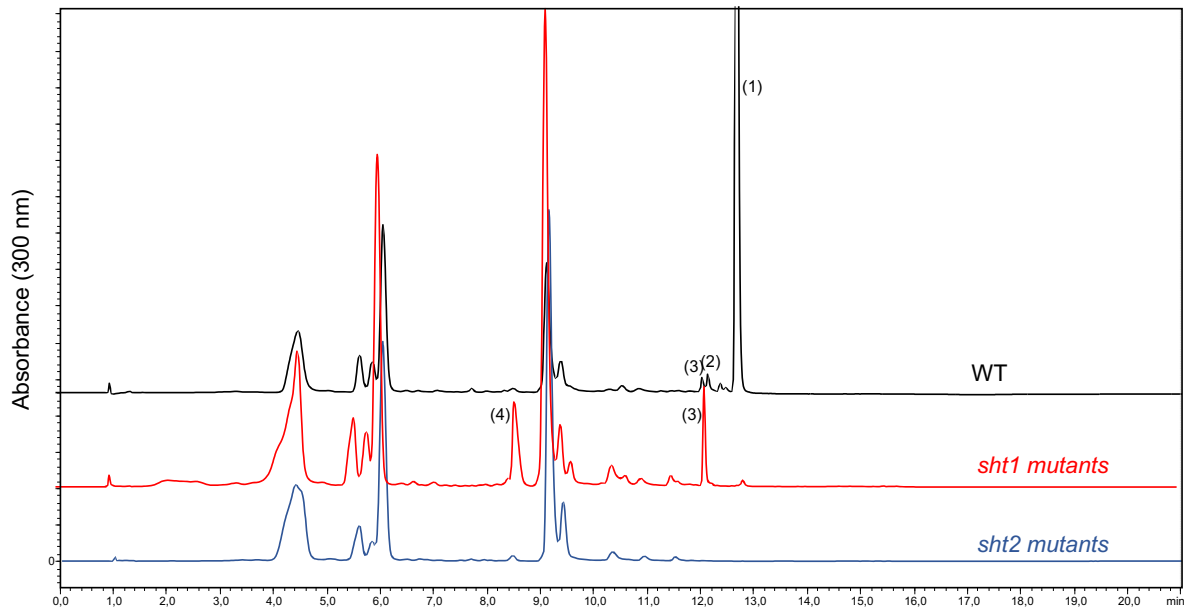
##### a. Spécificité catalytique de CiSHT1 et CiSHT2

###### Action catalytique séquentielle de CiSHT1 et CiSHT2

Les caractérisations biochimiques de CiSHT1 et CiSHT2 ont démontré leur capacité à catalyser la formation des phénolamides accumulées dans le manteau pollinique. Cependant, CiSHT1 est apparue comme une enzyme peu active *in vitro*. Il est donc possible que son activité principale n'ait pas été révélée au cours de notre étude précédente ou que les conditions expérimentales optimales n'aient pas été testées<sup>133</sup>. Par ailleurs, les tentatives de validation fonctionnelle *in planta* de CiSHT1 sont restées infructueuses que ce soit *via* la complémentation du mutant *sh1* d'Arabidopsis<sup>136</sup> ou *via* sa surexpression dans des chevelus racinaires de chicorée. La surexpression de CiSHT2 dans des chevelus racinaires de chicorée (tissus qui n'accumulent pas de phénolamides) induit une production et une accumulation de ces composés ce qui confirme sa fonction dans la biosynthèse de phénolamides. Cependant, dans ces conditions, la phénolamide majoritaire est la dicoumaroyl spermine et la tetracoumaroyl spermine n'est pas détectée. D'autre part, la complémentation du mutant *sh1* d'Arabidopsis avec CiSHT2 ne restaure que partiellement le phénotype métabolique sauvage. Les phénolamides complètement substituées sont peu représentées dans le cas des dérivés de spermidine voire absentes dans le cas de la spermine. Enfin, lorsque CiSHT2 est mise en contact avec de la spermine et du coumaroyl-CoA, *in vitro*, elle produit majoritairement la dicoumaroyl spermine, les composés tri- et tetrahydroxycinnamoylés étant très peu représentés. Ceci nous a conduit à émettre l'hypothèse que CiSHT1 et CiSHT2 interviendraient de manière successive pour catalyser les acylations successives de la spermine. C'est cette hypothèse que nous nous sommes proposés de tester dans un premier temps.

Le mutant *Cisht1*, *Cisht2* et le double mutant *Cisht1/Cisht2* ont été générés au laboratoire en utilisant la technologie d'édition de génome CRISPR/Cas9 décrite récemment chez la chicorée<sup>149</sup>. L'analyse des extraits méthanoliques de boutons floraux des mutants tend à

démontrer que CiSHT2 catalyserait les 3 premières étapes d'acylation tandis que CiSHT1 acylerait l'amine restante de la spermine (Fig. 58).



**Figure 58.** Chromatogrammes d'extraits méthanoliques de boutons floraux de plantes sauvages, de plantes *sht1* ou *sht2*. Chacune de ces lignées a été transformée avec la souche *A. rhizogenes* portant ou non (dans le cas du WT) un vecteur afin d'éditer le génome. 1 : tetracoumaroyl spermine, 2 : tricoumaroyl spermidine, 3 : tricoumaroyl spermidine, 4 : dicoumaroyl spermine.

Ces résultats ont été confirmés chez le tabac, *N. benthamiana*, qui n'accumule pas naturellement de dérivés de spermidine ou de spermine dans les feuilles. L'expression transitoire de *CiSHT1* est sans effet. Celle de *CiSHT2* conduit à une accumulation significative des produits intermédiaires, mono et dicoumaroyl spermidine ou spermine. La coexpression de *CiSHT1* et *CiSHT2* conduit à l'accumulation majoritaire de tricoumaroyl spermidine et de tetracoumaroyl spermine en quantité plus faible. Dans le cas de la coexpression, les dérivés non complètement acylés ont complètement disparu. Une situation équivalente est obtenue lorsque ces 2 gènes sont introduits chez la levure, *Saccharomyces cerevisiae*. Afin de valider cette hypothèse chez *A. thaliana*, le mutant *sht* déjà complété avec *CiSHT2* sera transformé avec *CiSHT1*. L'ensemble de ces résultats confirme l'hypothèse selon laquelle les 2 enzymes seraient nécessaires à l'obtention de spermine complètement substituée *in vivo*.

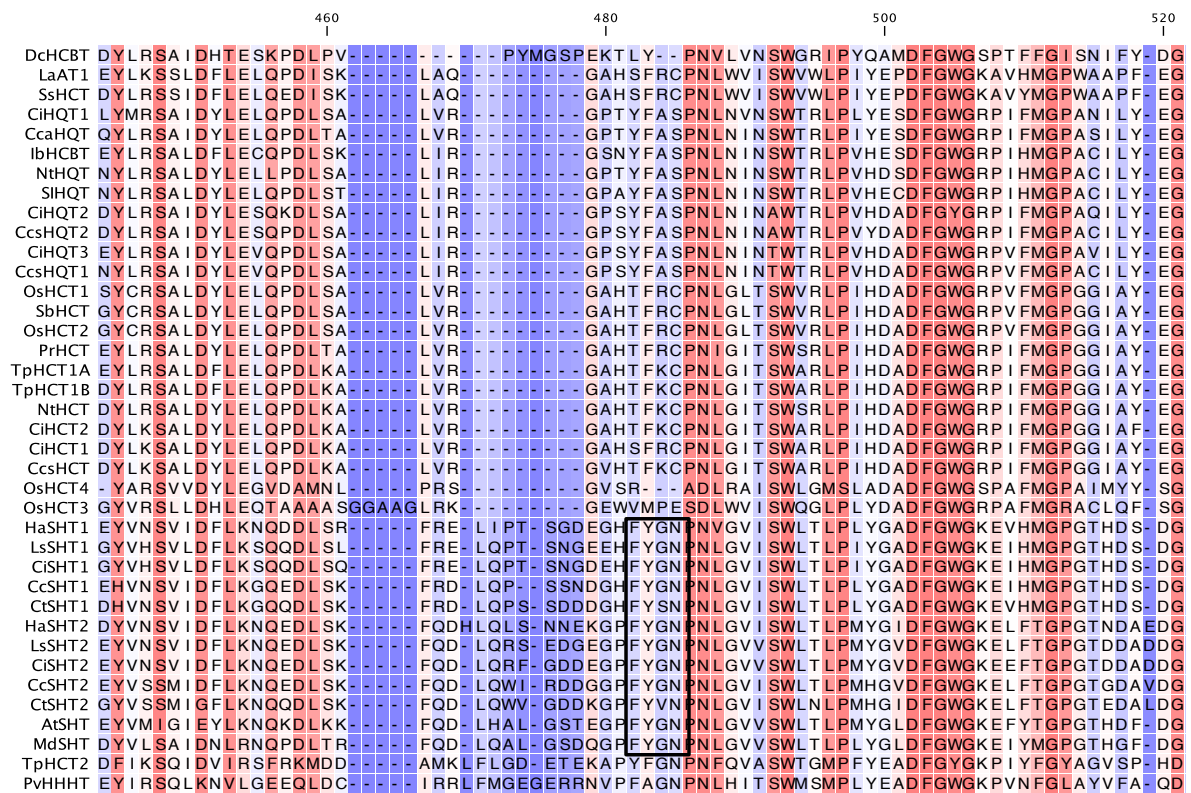
Les affinités des enzymes pour les différents intermédiaires semblent donc différentes. L'évaluation *in vitro* de ces affinités est une tâche qui est rendue difficile au vu de la disponibilité des substrats chez les fournisseurs de produits chimiques. Ceci sera possible lorsque nous aurons purifié les différents composés obtenus par voie enzymatique ou dans les organismes modifiés génétiquement (chevelus racinaires, *N. benthamiana*, levures). Le produit majoritaire obtenu avec des chevelus racinaires surexprimant *CiSHT2*, la dicoumaroyl spermine, a déjà été purifié après plusieurs étapes de chromatographie. Sa caractérisation par RMN et par spectrométrie de masse indique qu'il s'agirait de la spermine dont les 2 amines secondaires sont acylées. Ce substrat purifié sera utilisé *in vitro* avec *CiSHT1*. Il est prévu de faire de même avec la monocoumaroylspermine et la tricoumaroyl spermine.

Les essais préliminaires réalisés *in vitro* en mélangeant les 2 enzymes recombinantes en présence de spermine et de coumaroyl-CoA sont pour le moment restés infructueux. Il est également possible que ces enzymes soient organisées *in vivo* en métabolons avec d'autres enzymes de la voie des phénylpropanoïdes, ce qui sera compliqué à reproduire *in vitro*. Il a été démontré que les enzymes de la voie générale des phénylpropanoïdes que sont la PAL, la C4H, la 4Cl ainsi que la C3'H et HCT peuvent s'assembler en métabolons<sup>150,151</sup>. La présence de ce type de structures dynamiques dans le tapis des anthères pourra être recherchée *via*

l'utilisation d'approches telles que le FRET, la méthode yeast-2-hybrids ou la co-immunoprécipitation<sup>152</sup>.

### Analyse structurale des SHT-like

Il est admis que les SHT-like d'*A. thaliana* et de *Malus domestica* dérivent d'une HCT ancestrale qui aurait subi une duplication. L'une des copies aurait conservé sa fonction tandis que l'autre copie, à l'issue d'une série de modifications aléatoires, aurait acquis une nouvelle fonction à savoir la synthèse de phénolamides dans le tapis des anthères. Chez la chicorée (et plus généralement les Astéracées), une duplication plus récente d'un gène SHT-like ancestral aurait conduit à l'apparition de 2 gènes qui se seraient diversifiés en *SHT1* et *SHT2* dont les produits catalysent les acylations successives de la spermine. En effet, nous avons démontré que l'apparition des dérivés de spermine dans la paroi du grain de pollen à la place des dérivés de spermidine coïncide avec la présence des 2 enzymes SHT-like. Un examen approfondi des séquences des enzymes de la sous-famille SHT-like a été effectué afin de tenter d'expliquer cette diversité métabolique au niveau structural<sup>133</sup>. Aucun motif particulier n'a été détecté au sein de cette sous-famille. Cependant, l'alignement des séquences des SHT-like avec celles des autres membres du clade Vb a permis d'identifier une insertion peptidique placée juste avant le motif DFGWG caractéristique de la famille BAHD (Fig. 59).



**Figure 59.** Analyse des séquences des protéines de la sous-famille SHT-like appartenant au clade Vb des BAHDs. Seule une partie de l'alignement est représentée. Un des motifs caractéristiques des BAHD est souligné en gris (DFGWG). L'insertion peptidique commune à l'ensemble des membres de la sous-famille SHT-like est soulignée en noir tandis que le motif YFGN spécifique des enzymes impliquées dans la synthèse des phénolamides est encadré.

Cependant, la séquence de cette insertion est très variable et ce fragment peptidique est également retrouvé dans les séquences de TpHCT2 (hydroxycinnamoyl-CoA:malic acid transferase from *Trifolium pratense*) et PvHHHT (hydroxycinnamoyl-CoA:tetrahydroxyhexandioic acid transferase from *Phaseolus vulgaris*). Ces 2 enzymes ont des fonctions métaboliques distinctes des SHT-like. Elles pourraient dériver d'un même ancêtre commun que les SHTs mais auraient évolué différemment ce qui aurait conduit à la synthèse d'autres composés tel que l'acide phasélique chez *T. pratense* qui n'accumule pas

de phénolamides dans la paroi du grain de pollen. Une recherche de motif à l'aide de MEME a permis d'identifier un motif FYGN spécifique des enzymes acylant des polyamines (Fig. 59). Ce motif est très conservé chez les protéines acylant des polyamines. Celui-ci est légèrement modifié chez les enzymes T<sub>p</sub>HCT2 et P<sub>v</sub>H<sub>H</sub>H<sub>T</sub> (YFGN et FAGN pour T<sub>p</sub>HCT2 et P<sub>v</sub>H<sub>H</sub>H<sub>T</sub>, respectivement). L'implication de ce motif dans les spécificités métaboliques des SHT-like sera étudiée par mutagenèse dirigée. T<sub>p</sub>HCT2 et AtSHT seront manipulés génétiquement de manière à tester l'importance de ce motif (remplacement des motifs). Les expériences de mutagenèse dirigée par PCR ont été initiées et les protéines recombinantes modifiées sont en cours de production. Les propriétés catalytiques (spécificité de substrats en particulier) des enzymes recombinantes seront évaluées *in vitro* et *in vivo* chez *N. benthamiana*. Ces résultats devraient contribuer à attribuer ou non une participation de ce motif à la spécificité catalytique de ces enzymes.

Une analyse structurale par mutagenèse dirigée de CiSHT1 et CiSHT2 devrait permettre d'identifier les caractéristiques structurales à l'origine de la diversification catalytique de ces 2 enzymes. En effet, nos résultats indiquent que CiSHT2 aurait conservé les caractéristiques catalytiques des SHTs du type AtSHT et MdSHT tandis que CiSHT1 aurait acquis la capacité à acyler la di- et la tricoumaroylspermine. Des résultats obtenus *in vitro* montrent, en effet, que AtSHT présenterait les mêmes caractéristiques catalytiques que CiSHT2. CiSHT1 et CiSHT2 présentent 68% d'identité (72.9% au niveau nucléotidique). Afin de déterminer les motifs responsables de la diversification des activités catalytiques de ces enzymes, des protéines chimères ont été créées. Dans un premier temps, les chimères ont été définies de manière à séparer les différents domaines d'intérêt et à échanger ces régions entre CiSHT1 et CiSHT2 (isolation de 3 régions : régions contenant le motif HXXXD, les motifs DFGWG et FYGN et une région médiane). Les 6 gènes chimériques ont été clonés par PCR et seront introduits de manière transitoire chez le tabac. L'analyse des profils métaboliques du tabac devrait permettre de détecter des régions contribuant à la spécificité catalytique de ces enzymes. A l'issue de ces expérimentations, il sera possible de segmenter ces régions et *in fine* cibler des acides aminés spécifiques. Afin de faciliter cette approche, un travail récent d'analyse structurale de 2 protéines d'*Arabidopsis* (AtSDT et AtSHT) sera exploité<sup>153</sup>.

#### Lien entre pools de polyamines libres et produits finaux

Les résultats obtenus *in vitro* et *in vivo* montrent que les 2 enzymes de chicorée peuvent acyler à la fois la spermidine et la spermine. En outre, ces enzymes ne semblent pas plus affines pour la spermine que pour la spermidine. Or, la phénolamide majoritaire dans la paroi du pollen de chicorée est la tetracoumaroyl spermine. A l'inverse, dans les feuilles de tabac agroinfiltrées avec ces gènes, la phénolamide accumulée majoritairement est un dérivé de spermidine. L'examen des pools de polyamines libres présentes dans ces différents tissus végétaux devrait permettre de comprendre le rôle de la disponibilité en produits sur les profils d'accumulation des phénolamides (anthères de chicorée, anthères d'*Arabidopsis*, feuilles de *N. benthamiana* et levures). Chez *Arabidopsis*, le ratio spermidine/spermine dans les anthères est largement en faveur de la spermidine (5 fois plus de spermidine que de spermine) selon les données publiées<sup>154</sup>. Pour le dosage des polyamines libres, une collaboration a été mise en place avec le Prof. Alain Bouchereau (Université de Rennes1). Les résultats préliminaires montrent que dans les boutons floraux de chicorée, le ratio spermidine/spermine est largement en faveur de la spermine (ratio de 0.8). Sans surprise, dans les feuilles de tabac, le contenu en spermidine (1 nmol.g<sup>-1</sup> DW) est plus élevé que celui de la spermine (0.12 nmol.g<sup>-1</sup> DW), ce qui donne un ratio calculé de 8. Ainsi, la disponibilité en substrats ou l'abondance relative en substrats, et pas la spécificité de substrat, déterminerait la nature du produit final. L'examen des teneurs en phénolamides dans le mutant *sht* d'*Arabidopsis* complémenté avec CiSHT1 et CiSHT2 et le double mutant de chicorée *sht1/sht2* complémenté avec AtSHT devrait permettre de vérifier ou d'infirmer cette hypothèse.

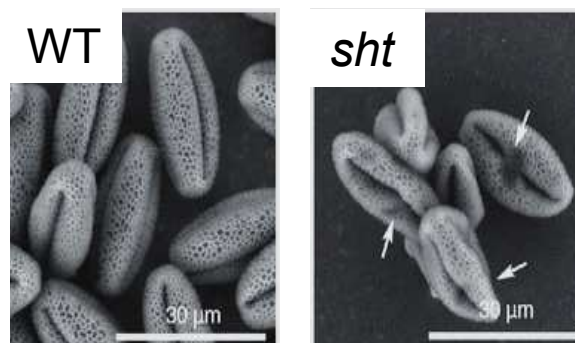
Si en plus d'une néo-fonctionnalisation des enzymes, le pool en polyamines est déterminant pour expliquer la diversification métabolique observée chez les Astéracées, il sera intéressant

de retracer la chronologie de ces événements au cours de l'évolution : le métabolisme des polyamines a-t-il été modifié avant, après ou conjointement à la néo-fonctionnalisation ? Pour ceci, nous proposons de réaliser des profils de polyamines libres dans les anthères d'espèces de la famille des Astéracées et familles proches en lien avec la duplication du gène *SHT* ancestral.

#### b. Rôles physiologiques des dérivés de spermine

La signification physiologique et évolutive de l'accumulation majoritaire chez les Astéracées des dérivés de spermine dans la paroi du grain de pollen reste énigmatique, tout comme l'est celle des dérivés de spermidine chez la quasi-totalité des Eudicotylédones<sup>133,144</sup>. Cette spécificité métabolique des Astéracées pourrait constituer une caractéristique supplémentaire expliquant le succès évolutif de cette famille. Cependant, les rôles des phénolamides dans la paroi du grain de pollen restent à ce jour inconnus.

La caractérisation phénotypique du mutant *sht* d'*Arabidopsis* a conduit à l'observation d'un grain de pollen déformé en lien avec une absence totale de phénolamides (Fig. 60)<sup>136</sup>. Ce phénotype n'affecte cependant pas la fertilité du grain de pollen ni sa résistance aux UVs. Concernant les UVs, il est important de noter que leurs effets n'ont été évalués que sur la capacité à germer des grains de pollen, sur la croissance du tube pollinique ainsi que sur leur fertilité. Les dommages que pourraient occasionner les UVs sur le support de l'information génétique n'ont pas été évalués. De même, il a été montré que cette modification métabolique n'avait aucun effet sur l'homéostasie des polyamines libres<sup>154</sup>. Les phénolamides pourraient jouer un rôle dans la protection contre les stress abiotiques et biotiques, d'antioxydant, dans les interactions plante-insectes pollinisateurs, dans l'interaction stigmate-pollen et dans les interactions plantes-microorganismes<sup>155</sup>. Cependant, les démonstrations sont manquantes. C'est ce que nous nous proposons d'examiner en utilisant les mutants d'*Arabidopsis* et de chicorée dont la synthèse en phénolamides est abolie.



**Figure 60.** Comparaison de la morphologie de grains de pollen d'*Arabidopsis* sauvage (WT) et de grains de pollen du mutant *sht* (d'après Grienemberger et al. 2012)<sup>136</sup>.

#### Expériences In planta

Comme indiqué précédemment, chez *Arabidopsis* l'absence de phénolamides dans la paroi du grain de pollen n'a pas révélé de phénotypes particuliers. Cependant, *Arabidopsis* est une plante autogame dont la pollinisation se fait souvent dans une fleur fermée et donc le grain de pollen est peu ou pas soumis aux pressions de l'environnement. Afin de déterminer si l'accumulation de spermine chez les Astéracées confère un avantage évolutif et tenter de comprendre le rôle de l'accumulation des phénolamides dans le grain de pollen, le double mutant *sht1/sht2* de chicorée sera utilisé. Des analyses biologiques du grain de pollen seront réalisées. Un protocole de germination des grains de pollen de chicorée *in vitro* est en cours d'optimisation au laboratoire. Cette méthodologie devrait permettre de tester l'impact de différents traitements sur la viabilité du grain de pollen.

Pour tester le rôle des phénolamides dans la protection de son information génétique, une collaboration a été initiée avec Jean Molinier (IBMP, Strasbourg). Les UVs représentent un

agent génotoxique qui conduit à des lésions pouvant se traduire par des perturbations de l'intégrité génomique et, par conséquent, mettant en danger la survie de l'organisme. Le grain de pollen et, en particulier, la cellule reproductrice est le vecteur de l'information génétique. Sa protection contre la lumière est donc primordiale. Dans ce contexte, une contribution des phénolamides à la protection contre les UVs expliquerait leur maintien chez la quasi-totalité des Eudicotylédones et, ainsi, l'avantage évolutif conféré par le remplacement des dérivés de spermidine par ceux de spermine chez les Astéracées. En effet, l'effet écran procuré par des phénolamides tetrasubstitués est plus important que celui de composés ne comportant que trois acides hydroxycinnamiques. C'est cette hypothèse que nous voulons tester.

Des grains de pollen issus de lignées sauvages ou modifiées génétiquement pour abolir la synthèse de ces molécules seront exposés aux UVs et l'intégrité de l'ADN analysée.

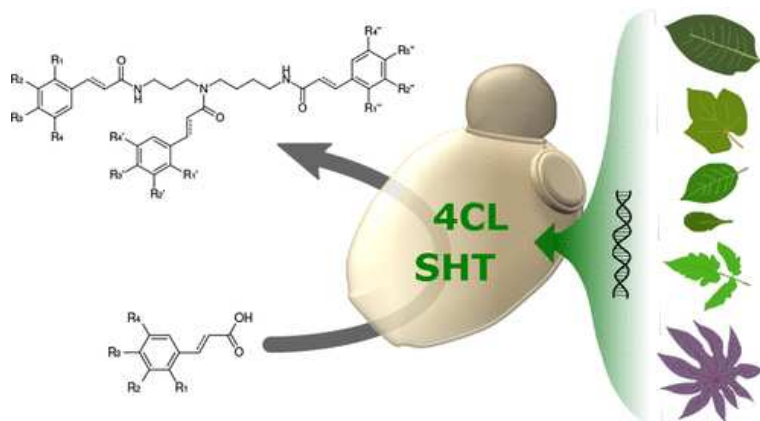
### Bioactivité des phénolamides

L'examen des propriétés biologiques des phénolamides apparaît pertinent pour répondre à 2 problématiques : déterminer le rôle physiologique des phénolamides en complément des manipulations exposées ci-dessus et examiner les possibilités de valorisation de ces composés (intérêt cosmétique, pharmaceutique, médical...).

Afin d'évaluer les propriétés biologiques de ces molécules, des démarches de purification à partir de capitule de marguerite (*Lecanthenum vulgare*) ont été initiées. Après une extraction suivie de plusieurs étapes de partition liquide/liquide basées sur un protocole décrit, par ailleurs, pour la purification de la trihydroxycinnamoyl spermidine<sup>156</sup>, ainsi que la mise au point d'une méthode de purification par HPLC semi-préparative, quelques milligrammes de tetracoumaroyl spermine purifiée ont été obtenus. La méthode de purification est désormais au point et devrait permettre de réaliser des purifications à plus grandes échelles afin de réaliser des tests d'activités *in vitro*. A ce stade, cette méthode a permis d'obtenir une quantité suffisante de molécules pour fournir au laboratoire un standard pour la chromatographie analytique. Si cette méthode permettra d'obtenir des quantités suffisantes pour la réalisation de tests préliminaires de bioactivité, la valorisation de ces composés selon cette approche semble inadéquate. En effet, pour un approvisionnement en quantité et en continu, la méthode retenue n'est pas adaptée. C'est pourquoi, d'autres solutions d'approvisionnements sont envisagées. Ainsi, la production en système hétérologue a été envisagée.

Pour cette approche, la levure de boulanger (*S. cerevisiae*) a été retenue comme châssis. La voie métabolique permettant la synthèse de phénolamides à partir d'acide coumarique fourni de manière exogène a été introduite chez *S. cerevisiae* et les premiers résultats sont encourageants. Comme évoqué précédemment, nous sommes en mesure de produire de la tetracoumaroyl spermine par ingénierie métabolique chez *S. cerevisiae*. Ceci nécessite l'introduction de 3 gènes : une 4CL (At4CL5 d'arabidopsis) et les 2 SHTs de chicorée et la fourniture exogène d'acide coumarique. D'autre part, la souche utilisée, ATCC 4005833, est une souche dont le gène PAD1 est invalidé. Ce gène code une coumarate décarboxylase. Ainsi, une production non négligeable de tetracoumaroyl spermine est obtenue. Cependant, ce n'est pas le produit majoritaire et les expériences ont montré que les dérivés de spermidine étaient les plus abondants et que les intermédiaires incomplètement substitués étaient bien représentés. D'autre part, les produits formés étaient accumulés dans les cellules ce qui ne facilite pas les procédures de purification. Il s'agira donc d'optimiser le système pour se soustraire d'un approvisionnement en acide coumarique, de favoriser l'excrétion dans le milieu extracellulaire des composés complètement substitués et de diminuer le ratio spermidine/spermine afin de favoriser la synthèse des composés dérivés de la spermine. L'introduction d'une TAL (tyrosine ammonia lyase) permettrait de lever le premier verrou. La TAL catalyse la transformation de la tyrosine, un acide aminé synthétisé par la levure, en acide coumarique. Ainsi, cette modification génétique permettrait de s'affranchir de l'approvisionnement en acide coumarique. Afin, de favoriser l'excrétion des composés, un changement de souche est envisagé. En effet, des travaux menés en collaboration avec Sébastien Besseau (BBV, Université de Tours) ont montré que certaines souches excrétaient

les phénolamides. Dans cette étude, une souche de *S. saccharomyces*, WT303, a été manipulée de manière à produire une panoplie de dérivés de la spermidine en utilisant des 4CL et des SHTs de différentes origines végétales (Fig. 61).



**Figure 61.** Illustration de la stratégie utilisée pour produire des dérivés de spermidine chez *S. cerevisiae*. L'approvisionnement d'acide coumarique à la souche WT303 dont le gène *PAD1* a été muté et chez laquelle, les gènes *4CL* et *SHT* de différentes origines (*At4CL1*, *At4CL2* ou *At4CL5* et *AtSHT*, *NtSHT*, *SISHT*, *RcoSHT*, *VvSHT* ou *RcaSHT*) ont été introduits. Cette souche excrète les produits synthétisés (d'après Perrin et al., 2021<sup>157</sup>).

Ces travaux ont fait l'objet d'une publication.

Perrin J., Kulagina N., Unlubayir M., Munsch T., Carqueijeiro I., Dugé de Bernonville T., De Craene J.O., Clastre M., St-Pierre B., Giglioli-Guivarc'h N., Gagneul D., Lanoue A., Courdavault V., Besseau S., 2021. Exploiting spermidine N-hydroxycinnamoyl transferase diversity and substrate promiscuity to produce various tri-hydroxycinnamoyl spermidines and analogs in engineered yeast. *ACS Synth. Biol.*, in press.

L'introduction d'une SPMS (spermine synthase) surexprimée devrait permettre de modifier le ratio spermidine/spermine et ainsi favoriser la synthèse des composés dérivés de la spermine. Des essais d'incubation avec ajout exogène de spermine de la souche transformée avec *At4CL5*, *CiSHT1* et *CiSHT2* ont montré que la biosynthèse et l'accumulation de dérivés de spermine étaient accrues (non publié).

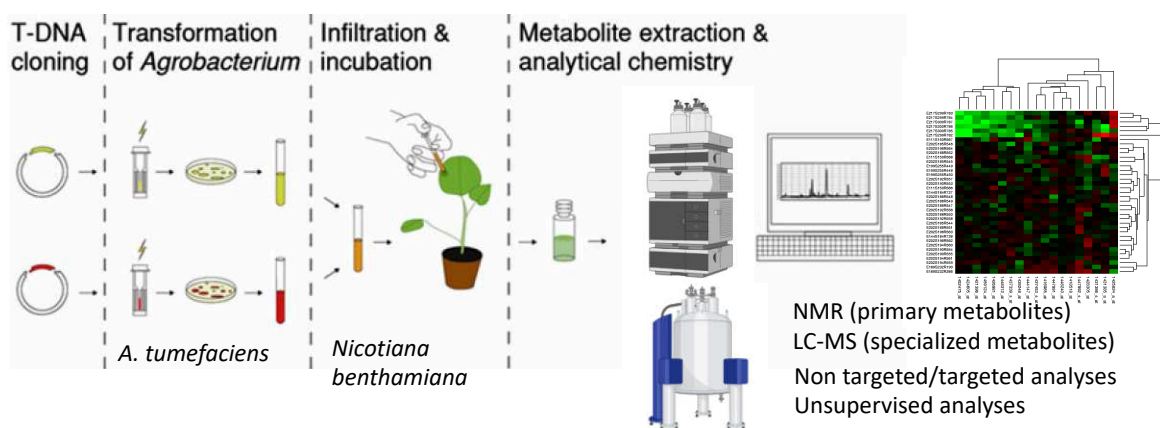
L'optimisation d'une souche de levure devrait donc permettre de produire le composé d'intérêt en quantité et son excrétion dans le milieu extracellulaire, ce qui représente un avantage certain pour sa purification. La mise en place d'un tel système devra être optimisée et une montée en échelle devrait permettre de collecter une quantité de molécules compatible avec la réalisation de tests d'activités biologiques.

### 3.2. Métabolisme et rôle du diCQA chez la chicorée

Le gène *lbICS* codant l'enzyme responsable de la dernière étape de la synthèse du diCQA a été identifié chez la patate douce (Fig. 54). Nos expériences menées avec des extraits protéiques bruts de chicorée ont montré que le même mécanisme réactionnel existait chez cette espèce. Afin de déterminer s'il s'agit d'une convergence évolutive ou s'il s'agit d'un système homologue, une démarche de génétique inverse a été entreprise chez la chicorée. Les bibliothèques d'ESTs ainsi que le génome de chicorée ont été criblés par tblastn en utilisant la séquence protéique lbICS. Trois gènes ont été retenus à l'issue de cette recherche : *CiGDSL1*, *CiGDSL2* et *CiGDSL3*. Les gènes candidats sont en cours de caractérisation *in vitro* et *in planta*. Si une séquence est identifiée comme codant une ICS, le rôle physiologique du diCQA chez la chicorée pourra être examiné par des approches génétiques : surexpression du gène, délétion du gène par la technologie CRISPR-Cas9.

### 3.3. Caractérisation des enzymes du sous-groupe spécifique du clade IIIa des BAHDs

Comme évoqué en 2.3, le criblage des bibliothèques de chicorée pour les protéines de la famille des BAHDs, a permis de mettre en évidence une proportion inattendue de protéines du clade IIIa (46 protéines, 50% des BAHDs). D'autre part, un sous-groupe spécifique des Astéracées au sein de ce clade a été déterminé. Il représente 29 protéines sur les 46 protéines classées au sein du clade IIIa. Ce sous-groupe fera l'objet d'investigations. Pour tenter d'attribuer une fonction à ces 29 gènes, une approche haut-débit est nécessaire. C'est pour cela que le projet **DECRIP** a été développé. Il a eu pour objet le **DE**veloppement d'un pipeline pour le **CR**iblage de gènes im**PL**iqués dans la synthèse de molécules à haute valeur ajoutée. Ce projet financé par la SFR Condorcet pour une durée de 1 an (2019) impliquait des partenaires de l'équipe BIOPI à l'UPJV (Roland Molinié et Jean-Xavier Fontaine) et de PFA à l'UPJV (David Mathiron). Ce partenariat associe des compétences en biologie moléculaire, en chimie analytique et en chimométrie. Il reposait sur l'analyse non ciblée de modification métabolique induite par l'expression transitoire de gènes de fonction inconnue chez *N. benthamiana* (Fig. 62).



**Figure 62.** Démarche développée dans le programme DECRIP (modifié d'après Bach et al. 2014)<sup>158</sup>.

Les avantages du tabac par rapport à d'autres modèles notamment microbiens et végétaux sont multiples : il a une croissance rapide, sa transformation pour une expression transitoire est une méthode rapide et qui ne nécessite pas une technicité importante, il est possible de co-exprimer plusieurs gènes en même temps, le tabac partage avec les autres plantes les mêmes caractéristiques de compartimentation cellulaire et le tabac synthétise naturellement une panoplie de substrats susceptibles d'être utilisés par les gènes candidats. Le pipeline a été optimisé en utilisant des gènes de fonctions connues. Par exemple, le gène *TpHCT2* (qui catalyse la dernière étape de la voie de biosynthèse de l'acide phasélique, acide caféoyl malique) a été introduit chez le tabac. A l'issue de cette modification, les modifications métaboliques induites ont été évaluées. L'accumulation de coumaroylmalate et de caféoylmalate a ainsi pu être mise en évidence. Ceci atteste de la solidité de la méthode. Ce projet a permis de mettre en place les outils moléculaires, analytiques et statistiques nécessaires à cette démarche.

Ce même pipeline sera utilisé pour la recherche sans *a priori* de la fonction de gènes non caractérisés.

### 3.4. Caractérisation du métabolisme du CTA et du diCTA

#### Implication de membres de la famille BAHD

L'élucidation des voies de biosynthèse de CTA et de diCTA n'a pas abouti jusqu'à maintenant. Cependant, plusieurs arguments sont en faveur de l'implication d'enzymes de la famille des BAHDs. C'est pourquoi, dans un premier temps, le sous-clade Vb a été ciblé sur la base du mécanisme réactionnel et de la proximité structurale des substrats. Mais il n'est pas exclu que



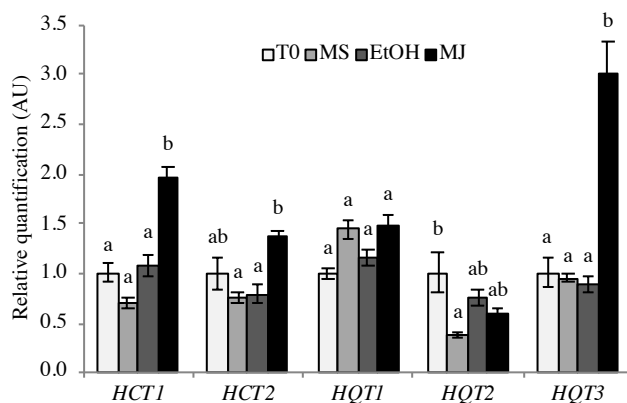
ces enzymes appartiennent à un autre clade de cette famille. Des enzymes relativement éloignées au sein de cette même famille peuvent, en effet, partager des activités similaires. Par exemple, AtSCT et AtSDT appartiennent au clade Va tandis que AtSHT appartient au clade Vb. Ces 3 enzymes catalysent l'acylation de polyamines. Comme indiqué précédemment, le criblage de l'ensemble des BAHDs de chicorée a permis d'identifier 81 protéines dont seulement 7 ont été caractérisées (2 CiHCTs, 3 CiHQTs et 2 CiSHTs). L'analyse de l'expression de ces gènes par qRT-PCR dans des conditions d'accumulation contrastées de ces molécules pourrait permettre d'isoler de nouveaux gènes candidats. Des expériences ont déjà été initiées dans ce sens.

### Mise en évidence du mécanisme réactionnel

Les mécanismes de synthèse des esters d'acide tartrique chez la chicorée peuvent être différents de ceux précédemment mis en évidence chez la prêle et l'arachide. Des phénomènes d'évolution convergente sont peut-être à l'origine de l'émergence ou de la diversification d'une autre famille d'enzymes impliquée dans la synthèse de ces dérivés d'acide caféique. Cette hypothèse est d'autant plus à considérer que les tests d'activité entrepris en alimentant des extraits protéiques bruts de chicorée avec de l'acide tartrique et de l'hydroxycinnamoyl-CoA n'ont à ce jour pas permis de confirmer le mécanisme réactionnel déjà proposé chez d'autres espèces. Des expériences complémentaires seront entreprises en fournissant notamment des acyles activés autres que les thioesters de CoA. Le CQA et les esters de glucose pourront notamment être testés. D'un point de vue génomique, d'autres familles d'acyltransférases pourront être criblées notamment la famille SCPL ou la famille GDSL.

### 3.5. Analyse de la redondance fonctionnelle des HCTs et des HQTs

Trois HQTs et 2 HCTs ont été caractérisées chez la chicorée<sup>92</sup>. Si leurs rôles dans la biosynthèse des dérivés d'acide shikimique et d'acide quinique ont été démontrés, la redondance fonctionnelle de ces enzymes soulève des interrogations. Une telle redondance a également été décrite chez l'Artichaut<sup>123</sup>. Chez la chicorée, une induction de culture cellulaire de racine au méthyljasmonate entraîne une augmentation significative de l'expression de CiHQT3 (Fig. 63)<sup>92</sup>. Cette apparente redondance pourrait donc être un moyen de contrôle des flux métaboliques dans des situations particulières.



**Figure 63.** Expression de *HCT1*, *HCT2*, *HQT1*, *HQT2* et *HQT3* dans des cultures cellulaires de racines de chicorée cultivées dans différentes conditions. T0 : cultures cellulaires avant élicitation, MS : cultures maintenues pendant 24 h sur le milieu témoin, EtOH : cultures traitées avec de l'éthanol pendant 24 h, MJ : cultures traitées avec du MeJA 50µM (dilué dans l'éthanol) pendant 24 h.

Des travaux supplémentaires sont nécessaires pour comprendre les rôles physiologiques précis de chacun de ces gènes : environnement particulier, répartition tissulaire particulière, stade de développement. Pour tenter d'apporter des éléments de réponses, la délétion de chacun de ces gènes en utilisant la technologie d'édition de gène CRISPR-Cas9 est envisagée. Des simples ou multiples mutants pourront également être générés et testés dans différentes

conditions. Une analyse de promoteurs à l'aide du gène rapporteur GUS sera entreprise afin d'appréhender de manière plus précise l'expression de ces différents gènes au niveau spatial.

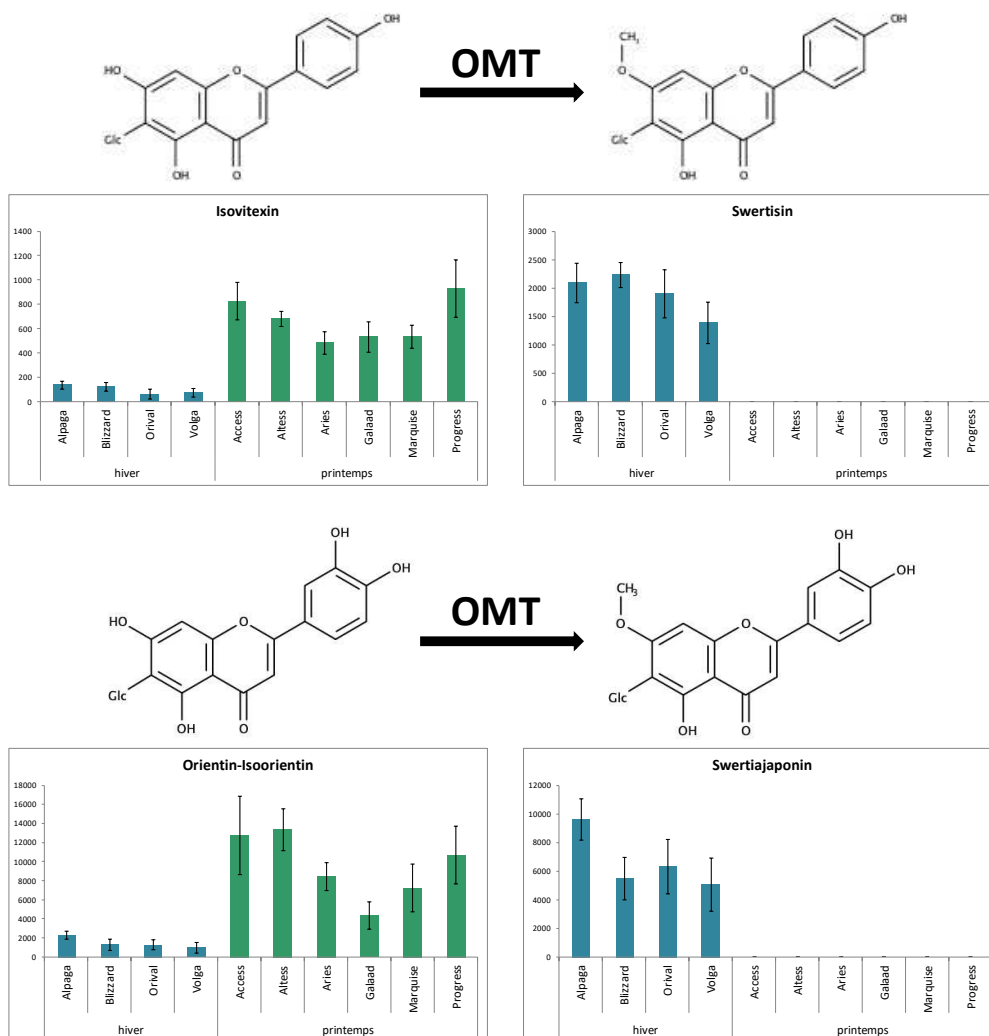
### **3.6. Caractérisation du métabolisme des flavones C-glycosides chez le lin**

Ce volet s'inscrit dans le projet CRISTALIN (Dispositif STIMuLE région Hauts-de-France, coordinateur : Anthony Quéro, UPJV, Amiens) : Contribution des flavones C-glucoside dans la tolérance au stress froid du lin. Ce projet est le fruit d'une collaboration établie avec l'équipe BIOPI (UPJV, Amiens) qui a rejoint notre UMRT en 2020, la plateforme analytique de l'UPJV (PFA, Amiens), le Centre de Ressources Régionales en Biologie Moléculaire de l'UPJV (CRRBM, Amiens) et le semencier Linea (Grandvilliers).

La filière agronomique souffre actuellement d'une sous-production chronique à l'échelle mondiale. Les prédictions annoncent une accentuation de ce problème dans les années à venir à cause de l'accroissement de la population d'une part et à des besoins nouveaux (agrocarburants, agromatériaux) d'autre part. Une augmentation de 70 % de la production alimentaire mondiale serait nécessaire d'ici 2050 pour répondre aux besoins. Le défi proposé à la filière agronomique est donc de taille, d'autant que les études sur les changements climatiques globaux, effectuées par le groupe d'experts intergouvernemental sur l'évolution du climat (IPCC), prédisent des températures extrêmes qui risquent de favoriser l'extension des zones climatiquement peu favorables au détriment des terres arables<sup>159,160</sup>. Pour les plantes de grande culture, l'augmentation de la fréquence de ces accidents climatiques va nécessiter l'obtention de variétés plus tolérantes aux conditions défavorables pour assurer un rendement suffisant. L'apparition de nouvelles variétés est donc nécessaire pour s'adapter à l'environnement de demain.

Chez le lin, il existe actuellement des variétés qualifiées d'hiver et de printemps. Le semis du lin d'hiver intervient entre le 15 septembre et la fin d'octobre et la récolte s'effectue de juillet à mi-août. Pour le lin de printemps, le semis est réalisé entre la fin février et la fin mars et la récolte intervient généralement un mois plus tard que celle du lin d'hiver. Ce décalage dans la conduite culturale permet au lin d'hiver d'éviter les périodes de sécheresse au moment de la floraison et de stabiliser les rendements. En 2018, le lin d'hiver couvrait 70 à 80% des surfaces cultivées. Pourtant peu de variétés d'hiver sont disponibles actuellement. Les nombreux travaux de sélection génétique n'ont pas permis d'obtenir un grand nombre de variétés tolérantes au froid. L'absence d'informations sur la différence de comportement dans la tolérance au froid entre les variétés d'hiver et de printemps a empêché pendant de nombreuses années l'identification de biomarqueurs qui pourraient faciliter ce travail de sélection.

Le profilage des composés phénoliques présents dans les feuilles de lin d'hiver et de lin de printemps a été réalisé en collaboration avec l'équipe BIOPI et la plateforme PFA (UPJV). En combinant la LC-MS et la RMN, 18 composés ont été identifiés et leur structure établie<sup>161</sup>. Parmi les molécules identifiées, 5 flavonoïdes C-glycoside l'ont été chez le lin pour la première fois. En particulier, 2 d'entre eux, la swertisine et la swertiajaponine étaient les composés phénoliques majoritaires dans les feuilles de 3 variétés d'hiver alors qu'ils étaient absents dans les feuilles de 3 variétés de printemps (Fig. 64). Ce résultat suggère que ces 2 composés pourraient être impliqués dans la tolérance au stress froid chez le lin même si cette hypothèse peut paraître à ce stade très naïve. Parallèlement, les teneurs en leurs 2 précurseurs hypothétiques, l'isovitexine et l'isorientine, étaient plus faibles dans les variétés d'hiver (Fig. 64). Ce résultat suggère que les gènes codant l'enzyme (ou les enzymes) responsable(s) de la O-méthylation de l'isovitexine et de l'isorientine en swertisine et swertiajaponine, respectivement, est (ou sont) invalidée(s) dans les variétés de printemps ou que les activités des enzymes produites sont modifiées.



**Figure 61.** Structures chimiques et teneurs des flavonoïdes C-glycoside dont les contenus sont différents entre les variétés de lin d'hiver (en bleu) et celles de printemps (en vert). Les relations substrats / produits indiquées par des flèches sont hypothétiques. OMT : O-méthyl transférase.

Ces travaux ont fait l'objet d'une publication.

Tchoumtchoua J., Mathiron D., Pontarin N., Gagneul D., van Bohemen A.I., Otogo N'ngang E., Mesnard F., Petit E., Fontaine J.X., Molinié R., Quérou A., 2019. Phenolic profiling of flax highlights contrasting patterns in winter and spring varieties. *Molecules*, 24: 1-14.

C'est dans ce contexte que s'inscrit le projet CRISTALIN. L'objectif de ce projet est d'étudier le lien entre la présence de flavones C-glucoside et la tolérance au stress froid. Pour cela, le nombre de variétés criblées sera élargi, d'autres biomarqueurs potentiels seront étudiés et la teneur des flavones C-glucoside sera suivie en condition de stress froid. Ce lien établi au niveau métabolique, un lien au niveau moléculaire sera également établi en remontant aux gènes impliqués dans la biosynthèse de ces composés. C'est sur ce volet du projet que j'interviens.

Nous supposons que le gène codant l'enzyme serait non exprimé chez la variété sensible et qu'une même enzyme catalyserait les méthylations pour produire ces 2 composés. Une population RILs (Recombinant Inbred Lines, F8) a été établie par notre partenaire Linea en croisant une variété tolérante au froid (Violin) et accumulant les 2 composés méthylés et une variété sensible au froid (Evea) et n'accumulant pas ces dérivés. A partir de ces RILs, une approche QTL a été engagée afin de corréler ce trait phénotypique avec un ensemble de marqueurs génétiques. Cette analyse QTL a permis de dégager une seule région fortement associée au trait phénotypique recherché. Cette zone couvre 700 kb. Sous cette zone, se

cachent 161 gènes. Un criblage des gènes en fonction des annotations prédites a permis d'identifier 5 gènes codant des enzymes catalysant des O-méthylations. Une cartographie plus fine devrait permettre de réduire encore le nombre de gènes candidats. A l'issue de ce travail, les gènes identifiés devront être validés. Ceci sera basé, en premier lieu, sur l'analyse de l'expression de ces gènes dans le parent sensible et dans le parent résistant par qRT-PCR. Si cette approche permettra peut-être de détecter un gène candidat fort, l'éventualité d'un gène exprimé mais dont le produit n'est pas fonctionnel n'est pas à exclure. En parallèle, des mesures d'activité méthyl transférase à partir d'extrait protéique brut sera entreprise pour valider l'absence d'une activité de ce type chez les variétés sensibles. La validation définitive des gènes candidats reposera sur la production et la purification de protéines recombinantes suivies de la mesure d'activité *in vitro* en utilisant l'isovitexine et l'isoorientine comme substrats. La validation par transgénèse (mutation par CRISPR-Cas9 du gène chez les variétés résistantes ou introduction du gène chez les variétés sensibles) est difficilement envisageable, le lin étant une espèce peu encline à la transformation génétique.

#### 4. Conclusion

Ce document dresse un panorama des projets auxquels j'ai participé au cours de la première partie de ma carrière de chercheur (doctorat, post-doctorats et activités d'enseignant-chercheur depuis mon recrutement à Lille en 2009) et ceux que je souhaite poursuivre ou développer dans l'avenir.

Le document répertorie un ensemble de projets dont les échéances s'étalent sur le court, le moyen et le long terme. Bien entendu, il ne s'agit pas de projets que je mènerai seul. Ils s'articulent autour d'un réseau de collaborations locales, nationales ou internationales. Ces projets s'appuient sur des compétences présentes au laboratoire ou auprès de partenaires extérieurs à l'Université de Lille. Les aspects biologie moléculaire et biochimie sont largement maîtrisés au laboratoire. Je bénéficie, dans ce cadre, d'une technicienne (Amandine Watrin) financée par le CPER Alibiotech et d'une doctorante (Julie Buges). De même, chaque année, je bénéficie du support d'étudiants de différents niveaux (BTS, L3, M1 et M2). Caroline Rambaud (Maître de Conférences, HDR) est experte dans le domaine de la biotechnologie en particulier sur la chicorée et contribuera largement aux expérimentations relatives à ces approches. Les analyses métaboliques ciblées ou non ciblées par spectrométrie de masse et RMN sont réalisées en partenariat avec l'équipe BIOPI de l'UPJV (Roland Molinié, Jean-Xavier Fontaine) et la plateforme analytique de l'UPJV (David Mathiron). Dans le cadre de la création de l'UMRT BioEcoAgro, l'équipe BIOPI nous a rejoint ce qui facilite les échanges technologiques et scientifiques. Au sein de cette UMRT, je dispose d'un environnement technologique et scientifique qui favorisera grandement la réalisation des projets proposés.

Tous les projets proposés concernent la caractérisation de voies métaboliques des composés phénoliques afin de disposer des outils moléculaires nécessaires à la compréhension de leurs rôles physiologiques, à l'amélioration variétale ou à la valorisation de métabolites d'intérêt. Bien que mes activités actuelles soient éloignées de ma formation initiale de physiologiste, les résultats générés ou à générer devraient me permettre d'aborder des problématiques relatives à la physiologie. Je dispose grâce à mon parcours professionnel de compétences en physiologie (doctorat), en biologie moléculaire (post-doctorats) et en biochimie qui seront mises à profit pour mener à bien les objectifs que je me suis fixés.



## V. Bibliographie

1. Flowers, T. J. & Yeo, A. R. Breeding for salinity tolerance in crop plants: where next? *Aust. J. Plant Physiol.* **22**, 875–884 (1998).
2. Tester, M. & Davenport, R. Na<sup>+</sup> tolerance and Na<sup>+</sup> transport in higher plants. *Ann. Bot.* **91**, 503–527 (2003).
3. Munns, R. & Tester, M. Mechanisms of Salinity Tolerance. *Annu. Rev. Plant Biol.* **59**, 651–681 (2008).
4. Chapman, V. J. The new perspectives in the halophytes. *Q. Rev. Biol.* **17**, 291–311 (1942).
5. Flowers, T. J. & Colmer, T. D. Salinity tolerance in halophytes. *New Phytol.* **179**, 945–963 (2008).
6. Horie, T., Karahara, I. & Katsuhara, M. Salinity tolerance mechanisms in glycophytes: An overview with the central focus on rice plants. *Rice* **5**, 1–18 (2012).
7. Abdelly, C., Bouchereau, A., Flowers, T. & Savoure, A. Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. *Ann. Bot.* **115**, 433–447 (2015).
8. Szekely, G. *et al.* Duplicated P5CS genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. *Plant J.* **53**, 11–28 (2008).
9. Sharma, S. & Verslues, P. E. Mechanisms independent of abscisic acid (ABA) or proline feedback have a predominant role in transcriptional regulation of proline metabolism during low water potential and stress recovery. *Plant, Cell Environ.* **33**, 1838–1851 (2010).
10. Szabados, L. & Savouré, A. Proline: a multifunctional amino acid. *Trends Plant Sci.* **15**, 89–97 (2010).
11. McKey, D. Adaptive patterns in alkaloid physiology. *Am. Nat.* **108**, 305–320 (1974).
12. Rhoades, D. Evolution of plant chemical defense against herbivores. in *Herbivores: their interactions with secondary plant metabolites*. (eds. Rosenthal, G. A. & Janzen, D. H.) 4–55 (Academic Press, Orlando, FL, 1979).
13. Rathinasabapathi, B., Fouad, W. M. & Sigua, C. A.  $\beta$ -Alanine betaine synthesis in the Plumbaginaceae. Purification and characterization of a trifunctional, S-adenosyl-L-methionine-dependent N-methyltransferase from *Limonium latifolium* leaves. *Plant Physiol.* **126**, 1241–1249 (2001).
14. Raman, S. B. & Rathinasabapathi, B.  $\beta$ -alanine N-methyltransferase of *Limonium latifolium*. cDNA cloning and functional expression of a novel N-methyltransferase implicated in the synthesis of the osmoprotectant  $\beta$ -alanine betaine. *Plant Physiol.* **132**, 1642–1651 (2003).
15. Rathinasabapathi, B., Sigua, C., Ho, J. & Gage, D. A. Osmoprotectant  $\beta$ -alanine betaine synthesis in the Plumbaginaceae: S-adenosyl-L-methionine dependent N-methylation of  $\beta$ -alanine to its betaine is via N-methyl and N,N-dimethyl  $\beta$ -alanines. *Physiol. Plant.* **109**, 225–231 (2000).
16. Duhaze, C., Gagneul, D., Leport, L., Larher, F. R. & Bouchereau, A. Uracil as one of the multiple sources of beta-alanine in *Limonium latifolium*, a halotolerant beta-alanine betaine accumulating Plumbaginaceae. *Plant Physiol. Biochem.* **41**, 993–998 (2003).
17. Duhazé, C., Gouzerh, G., Gagneul, D., Larher, F. & Bouchereau, A. The conversion of spermidine to putrescine and 1,3-diaminopropane in the roots of *Limonium tataricum*. *Plant Sci.* **163**, (2002).
18. Somerville, S. C. & Ogren, W. L. An Arabidopsis thaliana mutant defective in chloroplast dicarboxylate transport. *Proc. Natl. Acad. Sci.* **80**, 1290–1294 (1983).
19. Renné, P. *et al.* The Arabidopsis mutant *dct* is deficient in the plastidic glutamate/malate translocator DIT2. *Plant J.* **35**, 316–331 (2003).
20. Fernie, A. R. & Bauwe, H. Wasteful, essential, evolutionary stepping stone? The multiple personalities of the photorespiratory pathway. *Plant J.* 1–12 (2020).
21. Schwacke, R. *et al.* ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiol.* **131**, 16–26 (2003).
22. Lawand, S. *et al.* Arabidopsis *abou de soufflé*, which is homologous with mammalian carnitine acyl carrier, is required for postembryonic growth in the light. *Plant Cell* **14**, 2161–2173 (2002).
23. Ewald, R., Kolukisaoglu, Ü., Bauwe, U., Mikkat, S. & Bauwe, H. Mitochondrial protein lipoylation does not exclusively depend on the mtKAS pathway of de novo fatty acid synthesis in Arabidopsis. *Plant Physiol.* **145**, 41–48 (2007).
24. Collakova, E. *et al.* Arabidopsis 10-formyl tetrahydrofolate deformylases are essential for photorespiration. *Plant Cell* **20**, 1818–1832 (2008).
25. Porcelli, V. *et al.* Molecular identification and functional characterization of a novel glutamate transporter in yeast and plant mitochondria. *Biochim. Biophys. Acta - Bioenerg.* **1859**, 1249–

- 1258 (2018).
26. Sahr, T., Ravanel, S. & Rébeillé, F. Tetrahydrofolate biosynthesis and distribution in higher plants. *Biochem. Soc. Trans.* **33**, 758–762 (2005).
  27. Eisenhut, M. *et al.* Arabidopsis A BOUT de SOUFFLE is a putative mitochondrial transporter involved in photorespiratory metabolism and is required for meristem growth at ambient CO<sub>2</sub> levels. *Plant J.* **73**, (2013).
  28. Sweetlove, L. J., Beard, K. F. M., Nunes-Nesi, A., Fernie, A. R. & Ratcliffe, R. G. Not just a circle: Flux modes in the plant TCA cycle. *Trends Plant Sci.* **15**, 462–470 (2010).
  29. Vercesi, A. E. *et al.* Plant Uncoupling Mitochondrial Proteins. *Annu. Rev. Plant Biol.* **57**, 383–404 (2006).
  30. Klingenberg, M. & Winkler, E. The reconstituted isolated uncoupling protein is a membrane potential driven H<sup>+</sup> translocator. *EMBO J.* **4**, 3087–3092 (1985).
  31. Nicholls, D. G. The physiological regulation of uncoupling proteins. *Biochim. Biophys. Acta - Bioenerg.* **1757**, 459–466 (2006).
  32. Laloi, M. Plant mitochondrial carriers: An overview. *Cell. Mol. Life Sci.* **56**, 918–944 (1999).
  33. Palmieri, L. *et al.* Molecular identification of three Arabidopsis thaliana mitochondrial dicarboxylate carrier isoforms: Organ distribution, bacterial expression, reconstitution into liposomes and functional characterization. *Biochem. J.* **410**, 621–629 (2008).
  34. Vozza, A. *et al.* UCP2 transports C<sub>4</sub> metabolites out of mitochondria, regulating glucose and glutamine oxidation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 960–965 (2014).
  35. Monné, M. *et al.* Uncoupling proteins 1 and 2 (UCP1 and UCP2) from Arabidopsis thaliana are mitochondrial transporters of aspartate, glutamate and dicarboxylates. *J. Biol. Chem.* **2**, jbc.RA117.000771 (2018).
  36. Sweetlove, L. J. *et al.* Mitochondrial uncoupling protein is required for efficient photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 19587–19592 (2006).
  37. Yamori, W., Hikosaka, K. & Way, D. A. Temperature response of photosynthesis in C<sub>3</sub>, C<sub>4</sub>, and CAM plants: Temperature acclimation and temperature adaptation. *Photosynth. Res.* **119**, 101–117 (2014).
  38. Voznesenskaya, E. V., Franceschi, V. R., Kiirats, O., Freitag, H. & Edwards, G. E. Kranz anatomy is not essential for terrestrial C<sub>4</sub> plant photosynthesis. *Nature* **414**, 543–546 (2001).
  39. Sage, R. F. The evolution of C<sub>4</sub> photosynthesis. *New Phytol.* **161**, 341–370 (2004).
  40. Bräutigam, A. *et al.* An mRNA blueprint for C<sub>4</sub> photosynthesis derived from comparative transcriptomics of closely related C<sub>3</sub> and C<sub>4</sub> species. *Plant Physiol.* **155**, (2011).
  41. Ku, M. S. B., Schmitt, M. R. & Edwards, G. E. Quantitative determination of RuBP carboxylase-oxygenase protein in leaves of several C<sub>3</sub> and C<sub>4</sub> plants. *J. Exp. Bot.* **30**, 89–98 (1979).
  42. Leegood, R. C. C<sub>4</sub> photosynthesis: Principles of CO<sub>2</sub> concentration and prospects for its introduction into C<sub>3</sub> plants. *J. Exp. Bot.* **53**, 581–590 (2002).
  43. Chollet, R. & Ogren, W. L. Regulation of photorespiration in C<sub>3</sub> and C<sub>4</sub> species. *Bot. Rev.* **41**, 137–179 (1975).
  44. Aoki, N., Ohnishi, J. I. & Kanai, R. Proton/pyruvate cotransport into mesophyll chloroplasts of C<sub>4</sub> plants. *Plant Cell Physiol.* **35**, 801–806 (1994).
  45. Aoki, N. & Kanai, R. Reappraisal of the Role of Sodium in the Light- Dependent Active Transport of Pyruvate into Mesophyll Chloroplasts of C<sub>4</sub> Plants. *Plant Cell Physiol.* **38**, 1217–1225 (1997).
  46. Marshall, D. M. *et al.* Cleome, a genus closely related to Arabidopsis, contains species spanning a developmental progression from C<sub>3</sub> to C<sub>4</sub> photosynthesis. 886–896 (2007). doi:10.1111/j.1365-313X.2007.03188.x
  47. Voznesenskaya, E. V. *et al.* Physiological, anatomical and biochemical characterisation of photosynthetic types in genus Cleome (Cleomaceae). *Funct. Plant Biol.* **34**, 247–267 (2007).
  48. Oikawa, K. *et al.* Chloroplast Unusual Positioning1 Is Essential for Proper Chloroplast Positioning. *Plant Cell* **15**, 2805–2815 (2003).
  49. Kobayashi, H. *et al.* Differential positioning of C<sub>4</sub> mesophyll and bundle sheath chloroplasts: Recovery of chloroplast positioning requires the actomyosin system. *Plant Cell Physiol.* **50**, 129–140 (2009).
  50. Doemel, B. W. N. & Br, D. The Physiological Ecology. 17–32 (1971).
  51. Rothschild, L. J. & Mancinelli, R. C. Life in extreme environments. *Life Extrem. Environ.* **409**, 1–450 (2007).
  52. Barbier, G. & Oesterhelt, C. Comparative genomics of two closely related unicellular thermoacidophilic red algae, Galdieria sulphuraria and Cyanidioschyzon merolae, reveals the molecular. *Plant Physiol.* **137**, 460–474 (2005).



53. Schonknecht, G. *et al.* Gene Transfer from Bacteria and Archaea Facilitated Evolution of an Extremophilic Eukaryote. *Science (80- )*. **339**, 1207–1210 (2013).
54. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, 97–108 (2010).
55. Treangen, T. J. & Rocha, E. P. C. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* **7**, (2011).
56. Sanchez-Gracia, A., Maside, X. & Charlesworth, B. High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet.* **21**, 200–203 (2005).
57. Shih, P. M. *et al.* Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1053–1058 (2013).
58. Bowler, C. *et al.* The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* **456**, 239–244 (2008).
59. Nowack, E. C. M. & Grossman, A. R. Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5340–5345 (2012).
60. Dunning Hotopp, J. C. *et al.* Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science (80- )*. **317**, 1753–1756 (2007).
61. Rocha, E. P. C. With a little help from prokaryotes. *Science (80- )*. **339**, 1154–1155 (2013).
62. McCoy, J. G. *et al.* Discovery of sarcosine dimethylglycine methyltransferase from *Galdieria sulphuraria*. *Proteins Struct. Funct. Bioinforma.* **74**, 368–377 (2009).
63. Weng, J. The evolutionary paths towards complexity: a metabolic perspective. *New Phytol.* **201**, 1141–1149 (2014).
64. Moghe, G. D. & Last, R. L. Something Old , Something New: Conserved Enzymes and the Evolution of Novelty in Plant Specialized Metabolism. *Plant Physiol.* **169**, 1512–1523 (2015).
65. Dixon, R. A. & Strack, D. Phytochemistry meets genome analysis, and beyond. *Phytochemistry* **62**, 815–816 (2003).
66. Cragg, G. M. & Newman, D. J. Natural products: A continuing source of novel drug leads. *Biochim. Biophys. Acta - Gen. Subj.* **1830**, 3670–3695 (2013).
67. Vogt, T. Phenylpropanoid Biosynthesis. *Mol Plant* **3**, 2–20 (2010).
68. Jacobowitz, J. R. & Weng, J.-K. Exploring Uncharted Territories of Plant Specialized Metabolism in the Postgenomic Era. *Annu. Rev. Plant Biol.* **71**, annurev-arplant-081519-035634 (2020).
69. Sapir-Mir, M. *et al.* Peroxisomal localization of arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. *Plant Physiol.* **148**, 1219–1228 (2008).
70. Vranova, E., Coman, D. & Gruissem, W. Network Analysis of the MVA and MEP Pathways for Isoprenoid Synthesis. *Annu. Rev. Plant Biol.* **64**, 665–700 (2013).
71. Pichersky, E. & Raguso, R. A. Why do plants produce so many terpenoid compounds? *New Phytol.* **220**, 692–702 (2018).
72. Barros, J. *et al.* Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* **2**, 1–9 (2016).
73. Kolaj, I., Imindu Liyanage, S. & Weaver, D. F. Phenylpropanoids and Alzheimer’s disease: A potential therapeutic platform. *Neurochem. Int.* **120**, 99–111 (2018).
74. Jensen, R. A. Enzyme Recruitment in Evolution of New Function. *Annu. Rev. Microbiol.* **30**, 409–425 (1976).
75. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science (80- )*. **290**, 1151–1155 (2000).
76. Weng, J. & Joseph, P. Chemodiversity in *Selaginella*: a reference system for parallel and convergent metabolic evolution in terrestrial plants. **4**, 1–17 (2013).
77. Wink, M. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* **64**, 3–19 (2003).
78. Leong, B. J. & Last, R. L. Promiscuity, impersonation and accommodation: evolution of plant specialized metabolism. *Curr. Opin. Struct. Biol.* **47**, 105–112 (2017).
79. Bar-Even, A. *et al.* The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).
80. Weng, J. K. *et al.* The rise of chemodiversity in plants. *Science (80- )*. **336**, 1667–1670 (2012).
81. Anke, S., Niemuller, D., Moll, S., Hansch, R. & Ober, D. Polyphyletic Origin of Pyrrolizidine Alkaloids within the Asteraceae. Evidence from Differential Tissue Expression of Homospermidine Synthase. *Plant Physiol.* **136**, 4037–4047 (2004).
82. Frey, M. *et al.* Analysis of a chemical plant defense mechanism in grasses. *Science (80- )*. **277**, 696–699 (1997).

83. Nützmann, H.-W. & Osbourn, A. Gene clustering in plant specialized metabolism. *Curr. Opin. Biotechnol.* **26**, 91–99 (2014).
84. Osbourn, A. Secondary metabolic gene clusters: Evolutionary toolkits for chemical innovation. *Trends Genet.* **26**, 449–457 (2010).
85. Chu, H. Y., Wegel, E. & Osbourn, A. From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. *Plant J.* **66**, 66–79 (2011).
86. Jorgensen, K. *et al.* Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant Biol.* **8**, 280–291 (2005).
87. Wurtzel, E. T. & Kutchan, T. M. Plant metabolism, the diverse chemistry set of the future. *Science (80-. )*. **353**, 1232–1236 (2016).
88. De Luca, V., Salim, V., Atsumi, S. M. & Yu, F. Mining the biodiversity of plants: A revolution in the making. *Science* **336**, 1658–1661 (2012).
89. Pichersky, E. & Lewinsohn, E. Convergent Evolution in Plant Specialized Metabolism. *Annu. Rev. Plant Biol.* **62**, 549–566 (2011).
90. Street, R., #xe9, A., e, Sidana, J. & Prinsloo, G. Cichorium intybus: Traditional Uses, Phytochemistry, Pharmacology, and Toxicology. *Evidence-Based Complement. Altern. Med.* **2013**, 13 (2013).
91. Puhlmann, M.-L. & de Vos, W. M. Back to the Roots: Revisiting the Use of the Fiber-Rich Cichorium intybusL. Taproots. *Adv. Nutr.* 1–12 (2020). doi:10.1093/advances/nmaa025
92. Legrand, G. *et al.* Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.* **7**, 741 (2016).
93. Innocenti, M. *et al.* Evaluation of the phenolic content in the aerial parts of different varieties of Cichorium intybus L. *J. Agric. Food Chem.* **53**, 6497–6502 (2005).
94. Tusch, D. *et al.* Chicoric acid, a new compound able to enhance insulin release and glucose uptake. *Biochem. Biophys. Res. Commun.* **377**, 131–5 (2008).
95. Kuban-Jankowska, A., Sahu, K. K., Gorska, M., Tuszyński, J. A. & Wozniak, M. Chicoric acid binds to two sites and decreases the activity of the YopH bacterial virulence factor. *Oncotarget* **7**, 2229–2238 (2016).
96. Lee, N. Y. *et al.* Effect of Chicoric Acid on Mast Cell-Mediated Allergic Inflammation in Vitro and in Vivo. *J. Nat. Prod.* **78**, 2956–2962 (2015).
97. Bel-Rhliid, R. *et al.* Hydrolysis of chicoric and caftaric acids with esterases and lactobacillus johnsonii in vitro and in a gastrointestinal model. *J. Agric. Food Chem.* **60**, 9236–9241 (2012).
98. Koriem, K. M. M. & Soliman, R. E. Chlorogenic and Caftaric Acids in Liver Toxicity and Oxidative Stress Induced by Methamphetamine. *J. Toxicol.* **2014**, 583494 (2014).
99. Zhao, Y., Wang, J., Balleve, O., Luo, H. & Zhang, W. Antihypertensive effects and mechanisms of chlorogenic acids. *Hypertens. Res.* **35**, 370–374 (2012).
100. Liang, N. & Kitts, D. D. Role of chlorogenic acids in controlling oxidative and inflammatory stress conditions. *Nutrients* **8**, 1–20 (2015).
101. Zanello, P. R. *et al.* Quinic acid derivatives inhibit dengue virus replication in vitro. *Virology* **12**, 1–13 (2015).
102. Han, J., Miyamae, Y., Shigemori, H. & Isoda, H. Neuroprotective effect of 3,5-di-O-caffeoylquinic acid on SH-SY5Y cells and senescence-accelerated-prone mice 8 through the up-regulation of phosphoglycerate kinase-1. *Neuroscience* **169**, 1039–1045 (2010).
103. Gray, N. E. *et al.* Caffeoylquinic acids in Centella asiatica protect against  $\beta$ -amyloid toxicity. *J. Alzheimers Dis.* **40**, 359–373 (2014).
104. Kang, J. Y. *et al.* Reversal of Trimethyltin-Induced Learning and Memory Deficits by 3,5-Dicaffeoylquinic Acid. *Oxid. Med. Cell. Longev.* **2016**, (2016).
105. Sinisi, V. *et al.* Chlorogenic Compounds from Coffee Beans Exert Activity against Respiratory Viruses. *Planta Med.* **83**, 615–623 (2017).
106. Li, Y., But, P. P. H. & Ooi, V. E. C. Antiviral activity and mode of action of caffeoylquinic acids from Schefflera heptaphylla (L.) Frodin. *Antiviral Res.* **68**, 1–9 (2005).
107. Shi, S., Huang, K., Zhang, Y., Zhao, Y. & Du, Q. Purification and identification of antiviral components from Laggera pterodonta by high-speed counter-current chromatography. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **859**, 119–124 (2007).
108. Wu, Y. Compensation of ACE2 Function for Possible Clinical Management of 2019-nCoV-Induced Acute Lung Injury. *Virology* **35**, 256–258 (2020).
109. Sonnante, G. *et al.* Novel Hydroxycinnamoyl-Coenzyme A Quinate Transferase Genes from Artichoke Are Involved in the Synthesis of Chlorogenic Acid. *Plant Physiol.* **153**, 1224–1238 (2010).
110. Comino, C. *et al.* Isolation and functional characterization of a cDNA coding a

- hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. *BMC Plant Biol.* **7**, 14 (2007).
111. Lallemand, L. a. *et al.* A Structural Basis for the Biosynthesis of the Major Chlorogenic Acids Found in Coffee. *Plant Physiol.* **160**, 249–260 (2012).
  112. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* **278**, 95–103 (2003).
  113. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat. Biotechnol.* **22**, 746–754 (2004).
  114. Hohlfeld, M., Veit, M. & Strack, D. Hydroxycinnamoyltransferases Involved in the Accumulation of Caffeic Acid Esters in Gametophytes and Sporophytes of *Equisetum arvense*. *Plant Physiol.* **111**, 1153–1159 (1996).
  115. Sullivan, M. L. Perennial peanut (*Arachis glabrata* Benth.) leaves contain hydroxycinnamoyl-CoA:tartaric acid hydroxycinnamoyl transferase activity and accumulate hydroxycinnamoyl-tartaric acid esters. *Planta* **239**, 1091–1100 (2014).
  116. Lallemand, L. A. *et al.* A Structural Basis for the Biosynthesis of the Major Chlorogenic Acids Found in Coffee. *Plant Physiol.* **160**, 249–260 (2012).
  117. Moglia, A. *et al.* Dual catalytic activity of hydroxycinnamoyl-CoA quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicaffeoylquinic acids. *Plant Physiol.* **166**, 1777–87 (2014).
  118. Villegas, R. J. A., Shimokawa, T., Okuyama, H. & Kojima, M. Purification and characterization of chlorogenic acid: chlorogenate caffeoyl transferase in sweet potato roots. *Phytochemistry* **26**, 1577–1581 (1987).
  119. Bontpart, T., Cheynier, V., Ageorges, A. & Terrier, N. BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. *New Phytol.* **208**, 695–707 (2015).
  120. D'Auria, J. C. Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* **9**, 331–340 (2006).
  121. Yu, X.-H. H., Gou, J.-Y. Y. & Liu, C.-J. J. BAHD superfamily of acyl-CoA dependent acyltransferases in *Populus* and *Arabidopsis*: bioinformatics and gene expression. *Plant Mol. Biol.* **70**, 421–442 (2009).
  122. Tuominen, L. K., Johnson, V. E. & Tsai, C.-J. Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among *Populus* paralogues. *BMC Genomics* **12**, 236 (2011).
  123. Moglia, A. *et al.* Genome-Wide Identification of BAHD Acyltransferases and In vivo Characterization of HQT-like Enzymes Involved in Caffeoylquinic Acid Synthesis in Globe Artichoke. *Front. Plant Sci.* **7**, 1424 (2016).
  124. Ma, C. *et al.* Genome-wide analysis and characterization of molecular evolution of the HCT gene family in pear (*Pyrus bretschneideri*). *Plant Syst. Evol.* **303**, 71–90 (2017).
  125. Suzuki, H., Nakayama, T. & Nishino, T. Proposed mechanism and functional amino acid residues of malonyl-CoA:anthocyanin 5-o-glucoside-6"-O-malonyltransferase from flowers of *Salvia splendens*, a member of the versatile plant acyltransferase family. *Biochemistry* **42**, 1764–1771 (2003).
  126. Bayer, A., Ma, X. & Stöckigt, J. Acetyltransfer in natural product biosynthesis - Functional cloning and molecular analysis of vinorine synthase. *Bioorganic Med. Chem.* **12**, 2787–2795 (2004).
  127. Walker, A. M. *et al.* Elucidation of the Structure and Reaction Mechanism of Sorghum Hydroxycinnamoyltransferase and Its Structural Relationship to Other Coenzyme A-Dependent Transferases and Synthases. *Plant Physiol.* **162**, 640–651 (2013).
  128. Unno, H. *et al.* Structural and Mutational Studies of Anthocyanin Malonyltransferases Establish the Features of BAHD Enzyme Catalysis. *J. Biol. Chem.* **282**, 15812–15822 (2007).
  129. Ma, X., Koepke, J., Panjikar, S., Fritsch, G. & Stöckigt, J. Crystal Structure of Vinorine Synthase, the First Representative of the BAHD Superfamily. *J. Biol. Chem.* **280**, 13576–13583 (2005).
  130. Bustin, S. A. *et al.* The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).
  131. Delporte, M., Legrand, G., Hilbert, J.-L. & Gagneul, D. Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. *Front. Plant Sci.* **6**, 651 (2015).
  132. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W.-R. Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in

- Arabidopsis. *Plant Physiol.* **139**, 5–17 (2005).
133. Delporte, M. *et al.* A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in pollen coat of the Asteraceae family. *J. Exp. Bot.* **69**, 5365–5371 (2018).
  134. Qin, G. *et al.* Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in Arabidopsis by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res.* **17**, 471–482 (2007).
  135. Bernard, G. *et al.* Efficient genome editing using CRISPR/Cas9 technology in chicory. *Int. J. Mol. Sci.* **20**, (2019).
  136. Grienenberger, E. *et al.* A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant J.* **58**, 246–259 (2009).
  137. Elejalde-Palmett, C. *et al.* Characterization of a spermidine hydroxycinnamoyltransferase in *Malus domestica* highlights the evolutionary conservation of trihydroxycinnamoyl spermidines in pollen coat of core Eudicotyledons. *J. Exp. Bot.* **66**, 7271–7285 (2015).
  138. Murphy, D. J. The extracellular pollen coat in members of the Brassicaceae: Composition, biosynthesis, and functions in pollination. *Protoplasma* **228**, 31–39 (2006).
  139. Piffanelli, P., Ross, J. H. E. & Murphy, D. J. Biogenesis and function of the lipidic structures of pollen grains. *Sex. Plant Reprod.* **11**, 65–80 (1998).
  140. Eliasova, A., Poracka, V., Palove-Balang, P., Imrich, J. & Repcak, M. Accumulation of tetracoumaroyl spermine in *Matricaria chamomilla* during floral development and nitrogen deficiency. *Zeitschrift Fur Naturforsch. C-a J. Biosci.* **67**, 58–64 (2012).
  141. Yamamoto, A. *et al.* A new nonpeptide tachykinin NK1 receptor antagonist isolated from the plants of compositae. *Chem. Pharm. Bull.* **50**, 47–52 (2002).
  142. Granica, S. *et al.* Novel stilbenoids, including cannabispiradienone glycosides, from *Tragopogon tommasinii* (Asteraceae, Cichorieae) and their potential anti-inflammatory activity. *Phytochemistry* **117**, 254–266 (2015).
  143. Lam, S. C. *et al.* Chemical characteristics of different parts of *Coreopsis tinctoria* in China using microwave-assisted extraction and high-performance liquid chromatography followed by chemometric analysis. *J. Sep. Sci.* **39**, 2919–2927 (2016).
  144. Vogt, T. Unusual spermine-conjugated hydroxycinnamic acids on polle: function and evolutionary advantage. *J. Exp. Bot.* **69**, 5311–5318 (2018).
  145. Quilichini, T. D., Samuels, a L. & Douglas, C. J. ABCG26-Mediated Polyketide Trafficking and Hydroxycinnamoyl Spermidines Contribute to Pollen Wall Exine Formation in Arabidopsis. *Plant Cell* **26**, 4483–4498 (2014).
  146. Clauß, K., Baumert, A., Nimtz, M., Milkowski, C. & Strack, D. Role of a GDSL lipase-like protein as sinapine esterase in Brassicaceae. *Plant J.* **53**, 802–813 (2008).
  147. Ruppert, M. *et al.* Functional expression of an ajmaline pathway-specific esterase from *Rauvolfia* in a novel plant-virus expression system. *Planta* **222**, 888–898 (2005).
  148. Teutschbein, J. *et al.* Identification and Localization of a Lipase-like Acyltransferase in Phenylpropanoid Metabolism of Tomato (*Solanum lycopersicum*). *J. Biol. Chem.* **285**, 38374–38381 (2010).
  149. Bernard, G. *et al.* Efficient genome editing using CRISPR/Cas9 technology in chicory. *Int. J. Mol. Sci.* **20**, (2019).
  150. Bassard, J.-E. *et al.* Protein-protein and protein-membrane associations in the lignin pathway. *Plant Cell Online* (2012). doi:10.1105/tpc.112.102566
  151. Achnine, L., Blancaflor, E. B., Rasmussen, S. & Dixon, R. A. Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* **16**, 3098–109 (2004).
  152. Bassard, J. E. & Halkier, B. A. How to prove the existence of metabolons? *Phytochem. Rev.* **17**, 211–227 (2018).
  153. Wang, C. *et al.* Structural and Biochemical Insights Into Two BAHD Acyltransferases (AtSHT and AtSDT) Involved in Phenolamide Biosynthesis. *Front. Plant Sci.* **11**, (2021).
  154. Fellenberg, C., Ziegler, J., Handrick, V. & Vogt, T. Polyamine Homeostasis in Wild Type and Phenolamide Deficient Arabidopsis thaliana Stamens. *Front. Plant Sci.* **3**, 1–11 (2012).
  155. Fellenberg, C. & Vogt, T. Evolutionarily conserved phenylpropanoid pattern on angiosperm pollen. *Trends Plant Sci.* **20**, 212–218 (2015).
  156. Kite, G. C. *et al.* Acyl spermidines in inflorescence extracts of elder (*Sambucus nigra* L., Adoxaceae) and elderflower drinks. *J. Agric. Food Chem.* **61**, 3501–3508 (2013).
  157. Perrin, J. *et al.* Exploiting Spermidine N-HydroxycinnamoylTransferase diversity and substrate promiscuity to produce various tri-hydroxycinnamoyl spermidines and analogs in engineered

- yeast. *ACS Synth. Biol.* **in press**, (2021).
158. Bach, S. S. *et al.* Plant Isoprenoids. **1153**, 245–255 (2014).
  159. Schwilk, D. W. & Keeley, J. E. A plant distribution shift: Temperature, drought or past disturbance? *PLoS One* **7**, (2012).
  160. Joshi, R., Singla-Pareek, S. L. & Pareek, A. Engineering abiotic stress response in plants for biomass production. *J. Biol. Chem.* **293**, 5035–5043 (2018).
  161. Tchoumtchoua, J. *et al.* Phenolic Profiling of Flax Highlights Contrasting Patterns in Winter and Spring Varieties. *Molecules* **24**, 1–14 (2019).



## VI- Annexes

### Publications les plus significatives



Miguel S., Legrand G., Duriot L., Delporte M., Menin B., Michel C., Olry A., Chataigné G., Salwinski A., Bygdell J., Vercaigne D., Wingsle G., Hilbert J.L., Bourgaud F., Hehn A., Gagneul D., 2020. A GDSSL lipase-like from *Ipomoea batatas* catalyzes efficient production of 3,5-diCQA when expressed in *Pichia pastoris*. *Comm. Biol.*, 3: 673.

Delporte M., Bernard G., Legrand G., Hielscher B., Lanoue A., Molinié R., Rambaud C., Mathiron D., Besseau S., Linka N., Hilbert J.L., Gagneul D., 2018. A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in pollen coat of the Asteraceae family. *J. Exp. Bot.*, 69: 5365-5371.

Legrand G., Delporte M., Khelifi C., Harant A., Vuylsteker C., Mörchen M., Hance P., Hilbert J.L., Gagneul D., 2016 Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.*, 7: 741.

Gagneul D., Ainouche A., Duhazé C., Lugan R., Larher F.R. and Bouchereau A., 2007. A reassessment of the function of the so-called compatible solutes in the halophytic *Plumbaginaceae Limonium latifolium*, *Plant Physiol.*, 144, 1598-1611.

## A GDSL lipase-like from *Ipomoea batatas* catalyzes efficient production of 3,5-diCQA when expressed in *Pichia pastoris*

Sissi Miguel<sup>1,6</sup>, Guillaume Legrand<sup>2,6</sup>, Léonor Duriot<sup>1</sup>, Marianne Delporte<sup>2</sup>, Barbara Menin<sup>3</sup>, Cindy Michel<sup>1</sup>, Alexandre Olry<sup>3</sup>, Gabrielle Chataigné<sup>2</sup>, Aleksander Salwinski<sup>1</sup>, Joakim Bygdell<sup>4</sup>, Dominique Vercaigne<sup>2</sup>, Gunnar Wingsle<sup>5</sup>, Jean Louis Hilbert <sup>2,7</sup>, Frédéric Bourgaud<sup>1,7</sup>, Alain Hehn <sup>3,7</sup>✉ & David Gagneul<sup>2,7</sup>✉

The synthesis of 3,5-dicaffeoylquinic acid (3,5-DiCQA) has attracted the interest of many researchers for more than 30 years. Recently, enzymes belonging to the BAHD acyl-transferase family were shown to mediate its synthesis, albeit with notably low efficiency. In this study, a new enzyme belonging to the GDSL lipase-like family was identified and proven to be able to transform chlorogenic acid (5-O-caffeoylquinic acid, 5-CQA, CGA) in 3,5-DiCQA with a conversion rate of more than 60%. The enzyme has been produced in different expression systems but has only been shown to be active when transiently synthesized in *Nicotiana benthamiana* or stably expressed in *Pichia pastoris*. The synthesis of the molecule could be performed in vitro but also by a bioconversion approach beginning from pure 5-CQA or from green coffee bean extract, thereby paving the road for producing it on an industrial scale.

<sup>1</sup>Plant Advanced Technologies, Vandœuvre-lès-Nancy, France. <sup>2</sup>UMR Transfrontalière BioEcoAgro N° 1158, Univ. Lille, INRAE, Univ. Liège, UPJV, ISA, Univ. Artois, Univ. Littoral Côte d'Opale, ICV - Institut Charles Viollette, 59000 Lille, France. <sup>3</sup>Université de Lorraine-INRAE, LAE, 54000 Nancy, France.

<sup>4</sup>Chemistry Department, Umeå University, 90183 Umeå, Sweden. <sup>5</sup>Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, 90183 Umeå, Sweden. <sup>6</sup>These authors contributed equally: Sissi Miguel, Guillaume Legrand. <sup>7</sup>These authors jointly supervised this work: Jean Louis Hilbert, Frédéric Bourgaud, Alain Hehn, David Gagneul. ✉email: [Alain.Hehn@univ-lorraine.fr](mailto:Alain.Hehn@univ-lorraine.fr); [David.Gagneul@univ-lille.fr](mailto:David.Gagneul@univ-lille.fr)



Plants can be seen as factories that produce a large variety of specialized molecules displaying interesting biological properties. Among them, phenolic compounds are reported for their human health benefits and especially for their antioxidant and anti-inflammation properties<sup>1</sup>. A well-known example is chlorogenic acid (5-*O*-caffeoylquinic acid, 5-CQA, CGA)<sup>2–5</sup> formed by the esterification of caffeic acid with quinic acid. This molecule can be further combined with a second caffeic acid to produce dicaffeoyl quinic acids like 3,4-dicaffeoylquinic acid or its isomer 3,5-dicaffeoylquinic acid (3,5-DiCQA)<sup>2–5</sup>. 5-CQA and 3,5-DiCQA were detected in many higher plants, and among them, the main sources are coffee, potato and vegetables<sup>6–9</sup>. According to epidemiological studies, their daily intake is between 100 and 224 mg for 5-CQA and between 0.2 and 7.7 mg for 3,5-DiCQA<sup>6–8</sup>. The increasing complexity of these caffeate derivatives provides additional properties to the molecules, such as a higher solubility or stability, but also new biological activities. Many reports highlight the health benefits of 5-CQA and 3,5-DiCQA<sup>10,11</sup>. For instance, 5-CQA prevents low-density lipid oxidation, whereas 3,5-DiCQA has antihepatotoxic and anti-inflammatory activities. Due to their neuroprotective activity largely described in the literature, 5-CQA and 3,5-DiCQA are considered for the treatment of Alzheimer's disease<sup>12–16</sup>. Additionally, 3,5-DiCQA and other chlorogenic acids have been associated to anti-virus-cell fusion properties against numerous human viruses like human immunodeficiency virus, influenza A, herpes simplex 1 and 2, Coxsackie B3, adenovirus and respiratory syncytial viruses<sup>17–19</sup>. In parallel to these anti-viral activities, 3,5-DiCQA has also been assigned as a potent anti-inflammatory compound against lipopolysaccharide-induced acute lung injury through the inhibition of human neutrophil elastase activity, making it a potential new therapeutic candidate against acute viral pneumonia diseases<sup>20,21</sup>. Finally, 3,5-DiCQA has also been identified as a hypotensive molecule, inhibiting angiotensin I-converting enzyme (ACE)<sup>22,23</sup>. This ACE-inhibiting activity could help protect the lung cells infected by SARS-CoV-2 since the virus is using angiotensin II-converting enzyme (ACE2), a protein homolog to ACE, as a cell entry point, which, in return, generates a disruption in the ACE/ACE2 balance<sup>24</sup>.

The synthesis of 5-CQA is well documented, and four pathways have been described in the literature<sup>2,4</sup> (Fig. 1). In the first pathway, *p*-coumaroyl-CoA is esterified with quinic acid by a hydroxycinnamoyl-CoA: shikimate/quinic transferase (HCT), leading to *p*-coumaroylquinic acid<sup>2–4</sup>. This molecule is subsequently hydroxylated by a phenolic ester 3' hydroxylase (C3'H) to generate 5-CQA<sup>25</sup>. A second pathway involves a hydroxycinnamoyl-CoA: quinate hydroxycinnamoyl transferase (HQT) using caffeoyl-CoA and quinic acid as substrates to produce 5-CQA<sup>26</sup>. This second route seems to be the main pathway in tomato<sup>27</sup>. A third route was reported in the late eighties involving hydroxycinnamoyl D-glucose as an alternative substrate for the acylation of quinic acid<sup>28</sup>. The enzyme, a hydroxycinnamoyl D-glucose: quinate hydroxycinnamoyl transferase, was purified from the root of sweet potato, but no further molecular clues are available at present. Recently, Barros and colleagues showed that an enzyme, a coumarate 3-hydroxylase (C3H), catalyzes the direct 3-hydroxylation of 4-coumarate to produce caffeate that is further converted by a 4-hydroxycinnamate: CoA ligase (4CL) in caffeoyl-CoA and fused to a quinic acid to produce 5-CQA<sup>29</sup>. Finally, caffeate can be produced from caffeoyl shikimate due to a caffeoyl shikimate esterase (CSE)<sup>30</sup>.

The mechanism involved in the production of 3,5-DiCQA, although being widely investigated over the last 30 years, has not been fully elucidated. In 1987, Villegas et al. performed the first protein purification attempts from sweet potato tissues (*Ipomoea batatas*). The study performed by these researchers led to the characterization of the enzymatic properties of a chlorogenic acid:

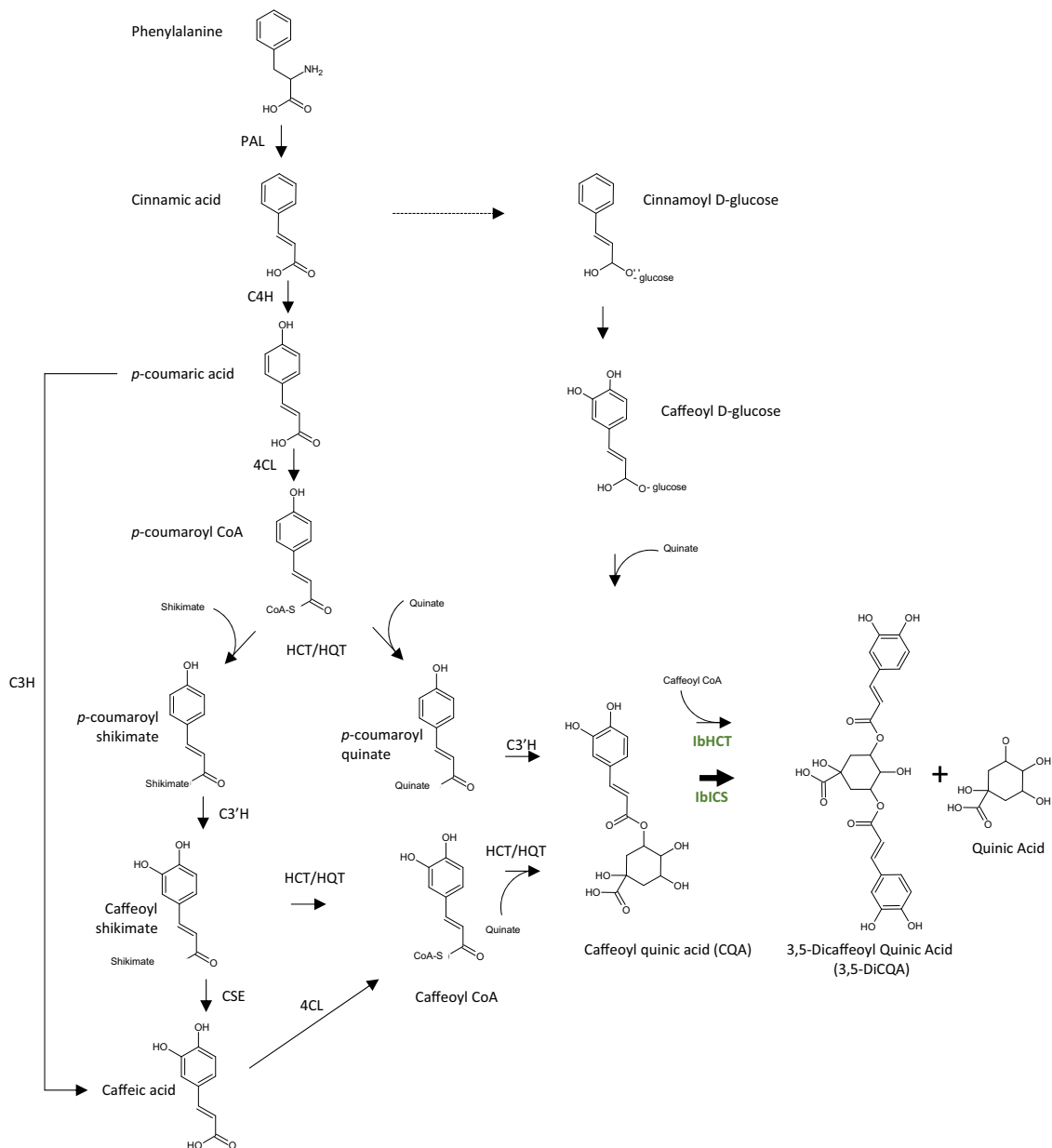
chlorogenate caffeoyl transferase directly involved in the synthesis of this molecule using 5-CQA as the sole substrate<sup>31</sup>. The isolated enzyme consisted of a single polypeptide with a molecular mass of 25 kDa and a pI of 4.6. The gene encoding this enzyme has not been further investigated. More recently, several genes were characterized, and the corresponding enzymes were reported to catalyze the production of 3,5-DiCQA. Hence, an HCT isolated from coffee and an HQT from tomato were shown to be able to catalyze the esterification of 5-CQA in the presence of caffeoyl CoA or 5-CQA alone, respectively, when expressed in *Escherichia coli*<sup>27,32</sup>. HCT and HQT are members of the BAHD family of plant-specific acyl-CoA-dependent acyltransferases<sup>33</sup>. However, the catalytic efficiency monitored *in vitro* is insufficient to explain the high levels of 3,5-DiCQA present in several plant species. Additionally, the abovementioned BAHD transferases do not share the same biochemical properties as the enzyme previously isolated by Villegas<sup>31</sup>. It might therefore be possible that the biosynthesis of 3,5-DiCQA in *I. batatas* relies on a different mechanism that remains to be identified at the molecular level.

In this report, we carried out a combined proteomic and transcriptomic approach to identify and functionally characterize two enzymes involved in the final step of the synthesis of 3,5-DiCQA from 5-CQA in *I. batatas*: IbHCT (hydroxycinnamoyl-CoA: shikimate/quinic hydroxycinnamoyl transferase) belonging to the BAHD acyltransferase and IbICS (isochlorogenate synthase) to the GDSL lipase-like family. IbICS exhibits the highest efficient chlorogenic acid:chlorogenate caffeoyl transferase activity. A bioproduction process using *Pichia pastoris* was successfully established, and a valuable tool to produce 3,5-DiCQA at high levels in industrial applications was initiated.

## Results

**Identification and functional characterization of IbHCT.** BAHD acyltransferases have been previously shown to be involved in the synthesis of 3,5-DiCQA in *Coffea canephora* and *Solanum lycopersicum*. Therefore, our investigation first began by focusing on the identification of orthologous genes in *I. batatas*. Using the amino acid sequences of HCT from robusta coffee (ID: ABO47805) and tomato SIHQT (ID: Q70G32) as queries, a tBLASTn search was conducted on the GenBank Database. We focused on a putative IbHCT (protein ID: BAJ14794). This protein shares 85% amino acid identity with the HCT from *C. canephora* that was described to be involved in the synthesis of 3,5-DiCQA. This enzyme also shares 84% with an HCT from *Nicotiana tabacum* (ID: CAD47830) and 82% with that of *Cynara cardunculus* (ID: AAZ80046). The protein contains the HXXXD and DFGWG motifs characteristic of the acyltransferases of the BAHD family (Supplementary Fig. 1). Based on the sequence available, the coding sequence was amplified from RNA extracted from *I. batatas* roots. The resulting sequence shows six different amino acids compared to the registered sequence. Since these amino acid mismatches were reproducible through several cloning experiments and were localized in nonconserved regions, this sequence was used for the following experiments (Supplementary Fig. 1). Since an HQT has been described for the same activity in tomato, we also cloned a gene encoding a putative IbHQT (Protein ID: BAA87043) for which the sequence was already available in public databases.

To perform functional characterization of this enzyme, it was expressed in *E. coli* BL21 (DE3) strains and, after purification, incubated in the presence of caffeoyl-CoA and quinic acid at 30 °C and pH 7.0 for 1 h, as described in Comino et al.<sup>3</sup> (Fig. 2a). In the presence of IbHCT, a new product (P1), absent in control reactions, was detected (Fig. 2a.i). P1 was identified as 5-CQA on the basis of UV spectra, mass spectra (MS) and retention time in

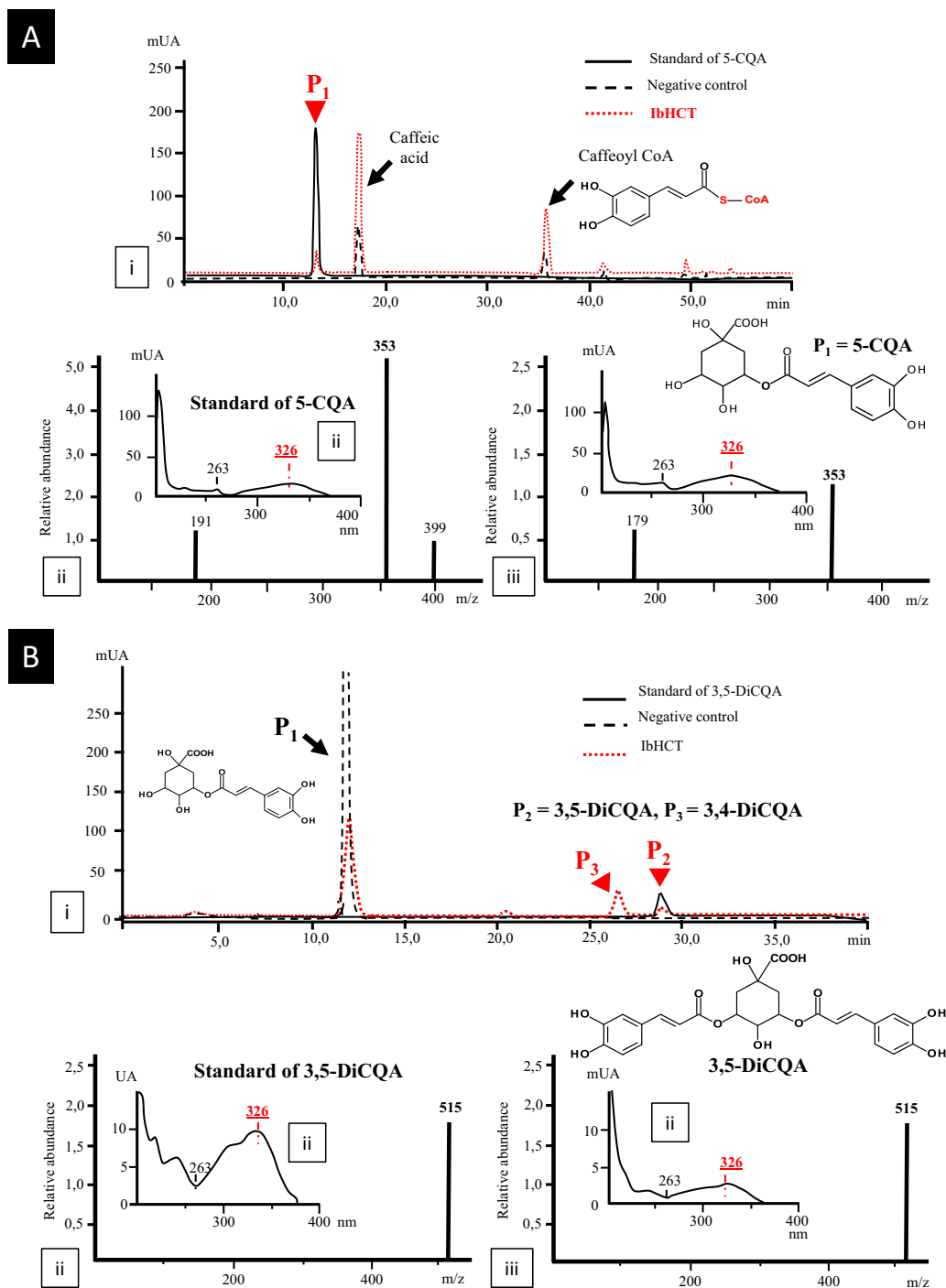


**Fig. 1 Simplified biosynthetic pathway of 3,5-DiCQA. IbHCT and IbCS (in green) are two enzymes of *Ipomoea batatas* characterized in this study.** PAL L-phenylalanine ammonia-lyase, C4H cinnamate 4-hydroxylase, C3H 4-coumarate 3-hydroxylase, 4CL 4-coumaroyl-CoA ligase, HCT/HQT 4-hydroxycinnamoyl CoA: shikimate/quinate hydroxycinnamoyl transferase, C3'H 4-coumaroyl shikimate/quinate 3'-hydroxylase, CSE caffeoyl shikimate esterase, IbCS isochlorogenate synthase from *I. batatas*, IbHCT HCT from *I. batatas*.

comparison to an authentic standard (Fig. 2a.ii). P1 gave a  $[M-H]^-$  ion at a mass-to-charge ratio ( $m/z$ ) of 353 and exhibited a maximum absorption peak at 326 nm (Fig. 2a.iii). These results showed that IbHCT has the ability to synthesize 5-CQA as HCTs from other plant species<sup>2-5</sup>. An additional peak (not present in the initial reaction mixtures) corresponding to caffeic acid also appeared in the positive and control reactions. This molecule may arise from caffeoyl-CoA or 5-CQA degradation, as reported by Kojima and Kondo<sup>34</sup> and Lallemand et al.<sup>32</sup>. In addition to these results, incubation of recombinant IbHCT with 5-CQA and caffeoyl-CoA at pH 6.0 led to the production of two new compounds (P2 and P3; Fig. 2b.i) absent in the negative control after 1 h of incubation at 30 °C. P2 was identified as 3,5-DiCQA based on UV spectra (maximum absorption at 326 nm) and MS spectra ( $[M-H]^-$  ion at a  $m/z$  of 515) in comparison with an authentic standard (Fig. 2b.ii and iii). P3 eluted before 3,5-DiCQA

and was identified as 3,4-DiCQA by mass spectrometry (not shown). P3 may result from the isomerization of 3,5-DiCQA (Supplementary Fig. 11), as previously described in other studies<sup>32</sup>. These results demonstrated that IbHCT is able to catalyze both 5-CQA and 3,5-DiCQA synthesis in vitro. Regarding IbHQT, only a very slight activity towards the synthesis of p-coumaroylquinic acid was monitored. No production of DiCQA was detected under our experimental conditions.

HCTs are often described to be involved in the synthesis of various hydroxycinnamoyl esters. The specificity of the enzyme in in vitro conditions with cinnamoyl-, p-coumaroyl-, caffeoyl- or feruloyl-CoA as acyl donors in combination with quinate or shikimate as acceptors was assessed. As expected, in the presence of the enzyme cinnamoyl-quinic acid (P4 in Supplementary Fig. 2 A), p-coumaroyl-quinic acid (P5 in Supplementary Fig. 2B), cinnamoyl-shikimate (P6 in Supplementary Fig. 3A), p-coumaroyl-shikimate



**Fig. 2** Enzymatic conversion of caffeoyl CoA and quinic acid in 5-CQA. Synthesis of 5-CQA from caffeoyl CoA and quinic acid (a) and of 3,5-DiCQA from caffeoyl-CoA and 5-CQA (b) catalyzed by IbHCT after a 1 h incubation at 30 °C at pH 7.0 or 6.0, respectively. (i) HPLC profile of the reaction mixtures after incubation with the recombinant protein, of the standards or of reaction mixtures incubated in the absence of the recombinant protein; (ii) UV spectrum and MS profile of standards; (iii) UV spectrum and MS profile of metabolization products.

(P7, P8 and P9 in Supplementary Fig. 3B) and caffeoyl-shikimate (P10 in Supplementary Fig. 3C) were detected in the reaction mixtures containing the different combinations of substrates. However, derivatives with feruloyl entities were not detected in the presence of quinate or shikimate.

This study led us to three main conclusions. First, IbHCT has a broad substrate activity as orthologous enzymes from coffee and tobacco<sup>2,32</sup>. With the exception of the production of 3,5-DiCQA and its isomer, 3,4-DiCQA, when the enzyme was incubated with

caffeoyl-CoA and quinate, we could not identify any additional peak that could correspond to dicaffeoyl shikimic acid or any other dihydroxycinnamoyl acids. Second, in a preliminary experiment, the content of 3,5-DiCQA was higher in roots grown in hydroponic conditions without nitrogen (2 g of 3,5-DiCQA / 100 g dry weight) than in medium supplemented with nitrogen (1 g of 3,5-DiCQA / 100 g dry weight) (Supplementary Fig. 4). The results obtained regarding the low efficiency of IbHCT are insufficient to explain the high concentration of 3,5-DiCQA in

**Table 1 Analyses of the peptides obtained by trypsin digestion of the proteins collected by SDS-PAGE.**

Acrylamide slice ID displaying an activity	Number of different peptides	Number of related proteins
3	261	25
4	230	27
5	437	64
6	779	52

Slices of SDS-PAGE exhibiting 3,5-DiCQA synthesis activity were analyzed by LC MS/MS. Numbers of corresponding proteins identified in public databases are indicated.

*I. batatas* roots grown under nitrogen depletion conditions. Third, the molecular mass of the enzyme (50 kDa) does not match the value of the chlorogenic acid:chlorogenate caffeoyl transferase described previously by Villegas et al.<sup>31</sup>. Therefore, another strategy was initiated in an attempt to identify additional enzymes involved in the synthesis of 3,5-DiCQA.

### Identification and functional characterization of IbICS

cDNA library construction. Little genetic information concerning *I. batatas* was available in public databases, which led us to generate a RNA-seq library. Total RNA was extracted from the roots of plants deprived of nitrogen. As mentioned above, in roots of plants grown under nitrogen-deprived conditions, the amounts of 3,5-DiCQA doubled compared to the control conditions. RNAs were extracted from roots of these plants to construct a 454-based de novo RNA-seq library (ID: PRJNA647243). This library was normalized to ensure representation of most of the genes expressed in these tissues. Clustering and assembly of these singletons led to the identification of 65264 contigs. These contigs were annotated and functionally assigned using Blast2GO. A Gene Ontology analysis realized on the library showed 8454 different annotations (Supplementary Fig. 5A), among which 37% were assigned to catalytic activities (Supplementary Fig. 5B). To assess the coverage of our library, we searched and confirmed the presence of IbHCT along with 13 additional sequences of putative acyltransferases belonging to the BAHD family. These enzymes were distributed in almost all BAHD acyltransferase clades (Supplementary Fig. 6).

*Preparation of IbICS enriched protein fractions and sequence analysis.* Based on the work performed by Villegas and collaborators, slices of *I. batatas* tubers were incubated in the dark for 48 h at 25 °C to induce 3,5-DiCQA accumulation and promote an increase in target enzyme activity<sup>31</sup>. Proteins from an acetone powder were extracted and subjected to different separation steps to identify the enzyme involved in 3,5-DiCQA biosynthesis. The enrichment procedure was only followed with the most active fractions after each step. In the final step, the proteins were separated by SDS-PAGE, and the proteins were collected in 8 different acrylamide gel fragments (~1 cm). The proteins extracted from each acrylamide piece were tested, and 3,5-DiCQA synthase activity could be highlighted in fragments #4 and #5, which correspond to proteins with an apparent molecular weight of approximately 30 kDa. The proteins present in these two bands, as well as the proteins included in the surrounding fragments (#3 and #6), were digested with trypsin and further analyzed using a UHPLC MS/MS approach. This analysis provided several hundred peptide sequences (Table 1, and Supplementary Table 1) that have been further identified and assigned to 25 (fragment #3) up to 64 (fragment #5) homologous proteins present in public databases.

*Identification of a full-length cDNA encoding IbICS.* The peptides identified by UHPLC MS/MS were used in a tBLASTn search of the de novo normalized RNA-seq library obtained from *I. batatas* roots to subsequently identify the corresponding coding sequences. A total of 24 proteins were identified in the four selected gel samples (Supplementary Table 1), and among these proteins, a sequence matching with a predicted *Ipomoea nil* GDSL esterase/lipase At1g28590-like (XM\_019313071.1) was detected in fragments #3, #4 and #5. This result led us to identify IbICS, a 1158 bp coding sequence (CDS) (Genbank ID: MT291823) (Supplementary Fig. 7) encoding a putative 385-amino-acid-long protein (Supplementary Fig. 7 and Fig. 3) with a calculated molecular mass of 41.93 kDa and pI of 4.92. In silico analysis performed with the SignalP and TargetP programs revealed a predicted 25-amino acid long N-terminal signal that might target IbICS into the secretory pathway. Removing this sequence leads to a protein of 39.54 kDa with an unchanged pI. A GDSL motif was detected between amino acids 34 and 37 of the deduced IbICS amino acid sequence (Fig. 3). A sequence alignment with three additional characterized GDSL lipase-like enzymes demonstrated a notably high sequence identity. Within GDSL, IbICS belongs to the subclass of SGNH hydrolases that are characterized by the presence of four conserved amino acids in four conserved blocks (I, II, III and V). A phylogenetic analysis showed that these enzymes clustered with GDSL identified in *Rauvolfia serpentina* (AAW88230), *S. lycopersicum* (CBV37053.1) and *Alopecurus myosuroides* (CAG27610). All these enzymes were shown to be involved in specialized metabolism (Fig. 4)<sup>35</sup>.

*Functional characterization of IbICS.* To determine the function of IbICS, the CDS was cloned into two different inducible expression systems based on *E. coli* and *Saccharomyces cerevisiae*. To analyze the effective expression of the protein in both systems, we added a polyhistidine tag at the C-terminus of the recombinant protein. The production of the protein was efficient in both systems (Fig. 5a, b). To further characterize the function, some enzymatic assays were performed in vitro in the presence of 5-CQA under the same conditions as previously described with the protein purified from the *I. batatas* tissues. Unfortunately, none of the proteins displayed any 3,5-DiCQA synthase activity under these conditions. We extended our investigation to a *Nicotiana benthamiana* plant expression system. The infiltrated leaves were harvested 96 h post infiltration, and the crude extract was used to perform incubations in the presence of 5-CQA at 28 °C for 22 h. UHPLC-MS analysis showed the appearance of a new product (P11) that has been identified as 3,5-DiCQA (molecular weight at 516) by mass spectrometry (mass in negative mode at 515) (Supplementary Fig. 8). This product was not detected when the crude extract was prepared from control leaves, which led us to assume that the enzyme is active in this system.

*Kinetic parameters and substrate specificity of the recombinant protein.* The protein expressed in *N. benthamiana* leaves was further purified using a His-Trap Ni column. The presence and identity of the recombinant protein were assessed by immunoblot analysis using an anti-His antibody (Fig. 5c). The activity of the purified recombinant protein was measured in vitro using 5-CQA as a substrate. The products of the reaction analyzed by UHPLC-MS confirmed the production of 3,5-DiCQA (Fig. 6a) with an m/z of 515 comparable to a commercial standard (Fig. 6c). No product was detected when 5-CQA was incubated in the presence of a protein extract of *N. benthamiana* infiltrated with an empty plasmid (Fig. 6b). The optimal pH and temperature were set at pH 6.3 and 39.9 °C, respectively (Supplementary Fig. 9). The  $K_m$  was determined at 3.5 mM (Supplementary Fig. 9). The activity of IbICS with various substrates including acyl donors (various hydroxycinnamate

<i>I. batatas</i>	MAAAP--SS	SLLGVLMLLA	SIGYA-----	-----AAC	--YSAVFGFG
<i>B. napus</i>	--MKKL--LS	FFLSTLFLLT	VVNSE-----	-----TTC	RNFKSIISFG
<i>R. serpentina</i>	MGFAR---LL	HLVFSLLVFA	GITNG-----	-----LIC	P-FDSIYQLG
<i>S. lycopersicum</i>	MALGMRVVVL	LLIISLILLQ	LIKGDMMVI	KLEEPKPKCC	G-IDRIFQFG
		▼		II	
<i>I. batatas</i>	<b>DSLTDAGNLI</b>	RLEPDGTVPH	MYFPPYGETY	<b>FKPTGRCS</b>	GRLIVDLIAQ
<i>B. napus</i>	<b>DSIADTGNLL</b>	GLSDPNLKP	VAFPPYGETF	FHHPTGRFSN	GRLIIDFLAE
<i>R. serpentina</i>	<b>DSFSDTGNLI</b>	RLPPDGPTFT	AAHFPYGETF	PGTPTGRCS	GRLIIDFLAT
<i>S. lycopersicum</i>	<b>DSLSDTGNCL</b>	RESYCGAQT	TGKLPYGMNF	YQNTATGRCS	GFILLDYIAM
<i>I. batatas</i>	<b>HYGLPLPPPS</b>	<b>IPASLEGDKN</b>	R-LRAGVNFA	<b>VVGSRALDAQ</b>	<b>FYEQR</b> DIFDFT
<i>B. napus</i>	FLGFPLVPP-	----FYGSQN	ANFEKGVNFA	VGGATALEPS	VLEERGIHFA
<i>R. serpentina</i>	ALNLPLLNP-	-----YLQQN	VSFRHGVNFA	VAGATALDRS	FLAARGVQVS
<i>S. lycopersicum</i>	ECGLPLLNP-	-----SLEEN	ADFSHGVNFA	VSGATALSAE	YLISRDIAMS
				III	
<i>I. batatas</i>	VTNVSMTDQL	NWFKELLPSL	CSSPHDCKEL	<b>LDGSI</b> <b>FVLGE</b>	<b>FGGNDYTHSL</b>
<i>B. napus</i>	YTNVSLGVQL	QSFKDSLPLN	CGSPTDCRHM	IENALILMGE	IGGNDYNYPL
<i>R. serpentina</i>	DIHSHLSAQL	NWFRTYLGSI	CSTPKECSNK	LKNALFILGN	IGNNDVNYAF
<i>S. lycopersicum</i>	FTNSSLSVQM	RWMSSYFKSV	CSN--DCAKY	LENSIFLIGE	IGGDDVTYGF
<i>I. batatas</i>	<b>LSGKGINVIR</b>	<b>PFPIVTRAI</b>	<b>AQAVHELVEL</b>	<b>GARTVMIPSV</b>	LPLGCAASYL
<i>B. napus</i>	FLGKPIEEIR	ELVPLVITTI	PSAITELIGM	GGRTFLVPGE	FPIGCAVIYL
<i>R. serpentina</i>	PN-RTIEEIR	AYVPFITEAV	ANATREIIRL	GGSRVIVPGI	FPIGCVARNL
<i>S. lycopersicum</i>	KQGKPIEEVR	RIVPDIVKNI	IHSVRTVIGF	GATRILVPGN	FPSGCFPIIL
<i>I. batatas</i>	TYYPSPNPDD	YDELGCLIWV	NEMASYHNQL	LQQQLAHVRE	LHPYANIVFA
<i>B. napus</i>	TLYKTPNKEA	YDSSGCLKWL	NEFAVYHDDQ	LQAELNKLRR	LYPHVNIYYA
<i>R. serpentina</i>	NFLNFFPDGD	KDDLGLCLSL	NNLSIYFNSL	FQALASLSI	EFFQAVIYYA
<i>S. lycopersicum</i>	TLYMNDSSTV	YDEYHCAEEW	NNFTISYNNL	LQQSIHELNE	EYPNISIIYG
				V	
<i>I. batatas</i>	DIYNAAMEVY	QGPDTYGFSG	GALR-ACCGG	GGPYNFNTSA	<b>QCGNE</b> <b>GAT-A</b>
<i>B. napus</i>	DYYNALLRLS	QEPTKFGFID	RALP-ACCG-	-----	-FGEK <b>GME-C</b>
<i>R. serpentina</i>	DYYNAWRFLF	RNGPALGSNS	TSLLKCCCGI	GGPYNYPDR	ECGSR <b>GVP-V</b>
<i>S. lycopersicum</i>	DYYNAYWLL	RNAVALGFNK	KTLQISCCGI	GGEYNYTESR	RCGK <b>GAEKA</b>
		▼			
<i>I. batatas</i>	<b>CDNPSSYVNW</b>	<b>DGYHLTEAAY</b>	QCITTGLLEG	PYTYPRMKNL	CRFDSRKTSR
<i>B. napus</i>	<b>CSGPSKYVSW</b>	<b>DSVHMTAAY</b>	RFMAEGVLKG	PYAIPPFDWS	CLNPEIKNSG
<i>R. serpentina</i>	<b>CPNPTQYIQW</b>	<b>DGTHFTQAA</b>	RRVAEYVIPG	I IKALKCSYS	NIQPFLEGE
<i>S. lycopersicum</i>	<b>CADPSSYLSW</b>	<b>DGSHLTQKAY</b>	GWITKWLIDD	ILPQLNCRV-	-----
<i>I. batatas</i>	VAQI-----	-			
<i>B. napus</i>	SSDTKSII--	-			
<i>R. serpentina</i>	GRQALRLNER	E			
<i>S. lycopersicum</i>	-----	-			

**Fig. 3 Alignment of the deduced amino acid sequence of IbICS with biochemically characterized GDSL lipase-like proteins.** The IbICS peptide sequence was deduced from cDNA and aligned with sinapine esterase from *Brassica napus* (AAX59709), acetylajmaline esterase from *Rauvolfia serpentina* (AY762990) and chlorogenate: glutarate caffeoyltransferase from *Solanum lycopersicum* (FR667689). The conserved GX<sub>2</sub>XXDXG motif is illustrated in bold. The predicted N-terminal leader sequences of all enzymes are shown in italics. Black box: conserved blocks in the SGNH-hydrolyase family (I, II, III and V). Amino acid residues forming the catalytic triad in the consensus sequences of blocks I and V are marked by black triangles. Sub-sequences identified by mass spectrometric sequencing after trypsin digestion are marked in red or in blue.

derivatives) and other acyl acceptors (5-CQA and quinic acid) was tested, but no activity was observed under our experimental conditions (Supplementary Fig. 10).

**Development of a system amenable to produce 3,5-DiCQA at industrial level.** To establish a suitable process for mass production of 3,5-DiCQA, we switched from the plant system to *P. pastoris* as an alternative expression factory. The IbICS CDS was cloned into the pICZaA integrative plasmid that contains the *S. cerevisiae* alpha secretor factor sequence favouring the excretion of the protein into the culture medium. The recombinant

*P. pastoris* strain was cultured under standard conditions, and protein expression was induced with methanol after 96 h. These conditions led to the production of the protein detectable in the culture medium (Fig. 5d). In a second step, different amounts of pure 5-CQA were added to the medium and the production of 3,5-DiCQA was monitored. The highest bioconversion rate of 5-CQA (61.3%) was obtained with 7.5 mM 5-CQA (Fig. 7a), leading to the synthesis of 2.3 mM of DiCQA (instead of 3.75 mM expected for 100% efficiency). Using a higher concentration of 5-CQA (over 10 mM) led to an inhibitory effect, and a lower concentration decreased the yields of 3,5-DiCQA produced



**Fig. 4 Phylogenetic tree of GDSL lipase/esterase protein sequences available in public databases.** The tree was constructed using the neighbour joining method and adapted from Chepyshko et al.<sup>35</sup>.

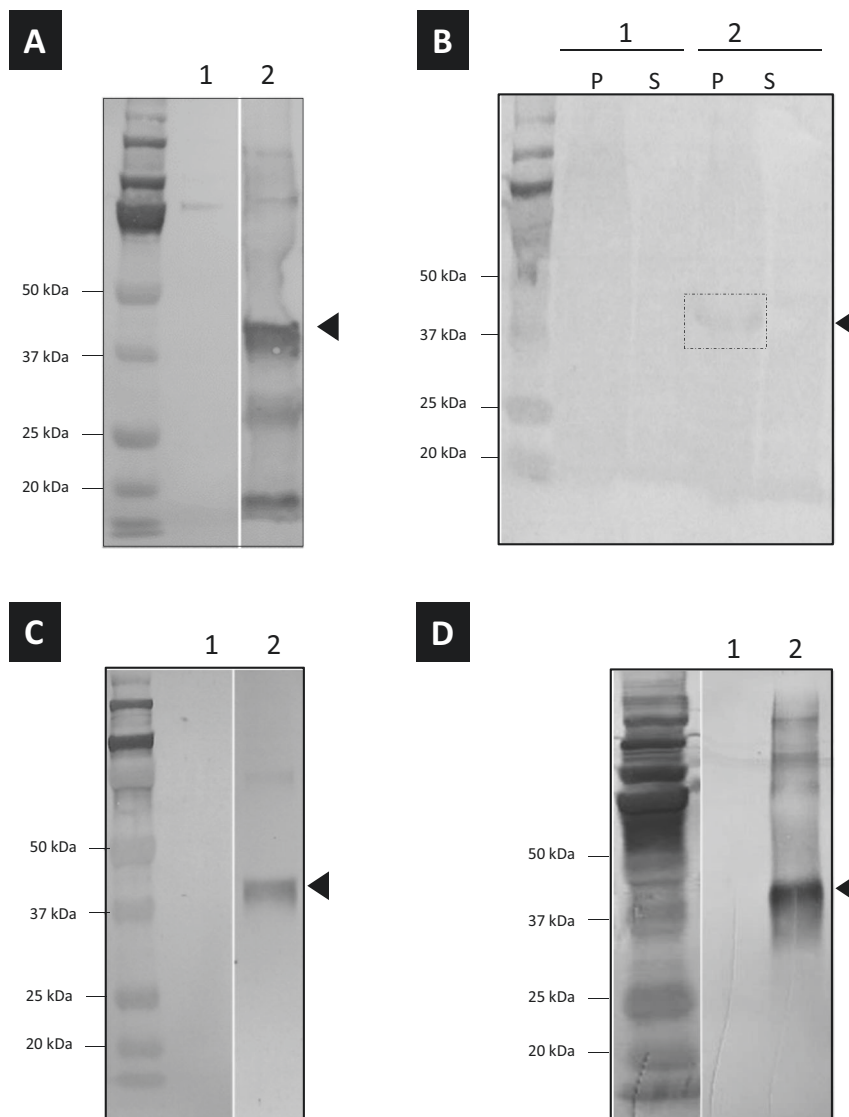
(Fig. 7a). A 3,5-DiCQA-specific synthesis could also be highlighted when adding a green coffee extract containing 5-CQA (Fig. 7b). The bioconversion allowed a 7.5-fold increase of 3,5-DiCQA in comparison to native extract with a transformation efficiency estimated at 36%. The difference in efficiency in comparison with pure 5-CQA described above might be related to the complex composition of the green coffee bean extract where compounds structurally close to 5-CQA or other unrelated molecules might exert an inhibitory activity against IbICS.

**Discussion**

Given the interest of 3,5-DiCQA, many scientists have been searching for more than 30 years for the gene(s) responsible for its biosynthesis. In this report, our phytochemical analyses show that a decrease in nitrogen availability increased the

concentrations of 5-CQA acid and 3,5-DiCQA in *I. batatas* plant root tissues. The amount of 3,5-DiCQA can reach almost 2 g for 100 g of dried roots. Such results have already been described by Galieni and collaborators, who have evidenced in lettuce that the accumulation of phenolic acids was significantly enhanced, especially 5-CQA and 3,5-DiCQA<sup>36</sup>. Consequently, we carried out a combined proteomic and transcriptomic approach in this tissue to identify and functionally characterize two enzymes involved in the final step of the synthesis of 3,5-DiCQA from 5-CQA: IbHCT and IbICS.

In this study, an HCT isolated from *I. batatas* was shown to be able to synthesize 3,5-DiCQA using chlorogenic acid and caffeoyl CoA as substrates in addition to its main catalytic activity towards 5-CQA production. Such enzymatic activity has already been demonstrated for an HCT isolated from coffee. The synthesis of



**Fig. 5 Expression analysis of IbICS in different heterologous expression systems.** Western blot analysis of proteins produced and purified from **a** *E. coli*, **b** *S. cerevisiae*, **c** *N. benthamiana* and **d** *P. pastoris*. For each sample, 1 corresponds to the purification performed on cells/tissues transformed with an empty vector. 2 corresponds to the purification performed on cells/tissues transformed with a plasmid containing the *IbICS* CDS. P: pellet of *S. cerevisiae* culture after protein extraction; S: supernatant. Parts of gels separated by a space were grouped together and lined up to facilitate the visualization of results.

3,5-DiCQA was found to be optimal at pH 6, while the synthesis of chlorogenic acid was optimal at pH 7 (not shown). This difference could be explained by the fact that the activity of the enzyme may change depending on the cell compartment. It is therefore likely that the activity of 3,5-DiCQA synthesis takes place in an acidic compartment of the cell, such as the vacuole<sup>27</sup>. We also observed that the reaction leading to the synthesis of 3,5-DiCQA was notably slow. Indeed, the product became detectable after a 1 h incubation time. These data are consistent with the results described by Lallemand et al. for coffee HCT<sup>32</sup>. Finally, small amounts of 3,5-DiCQA were produced in vitro, which did not allow quantification. According to several studies<sup>2–4</sup>, HCTs perform reversible reactions. It is therefore possible that the low reaction yields are related to an equilibrium between the precursor and the product. However, the low catalytic performance of HCTs for the synthesis of 3,5-DiCQA makes it difficult to explain the accumulation of high concentrations of this molecule in plants like sweet potato. Some authors suggest that 5-CQA synthesized in the cytoplasm is further transferred into vacuoles, where it is stored as 3,5-DiCQA. This side activity

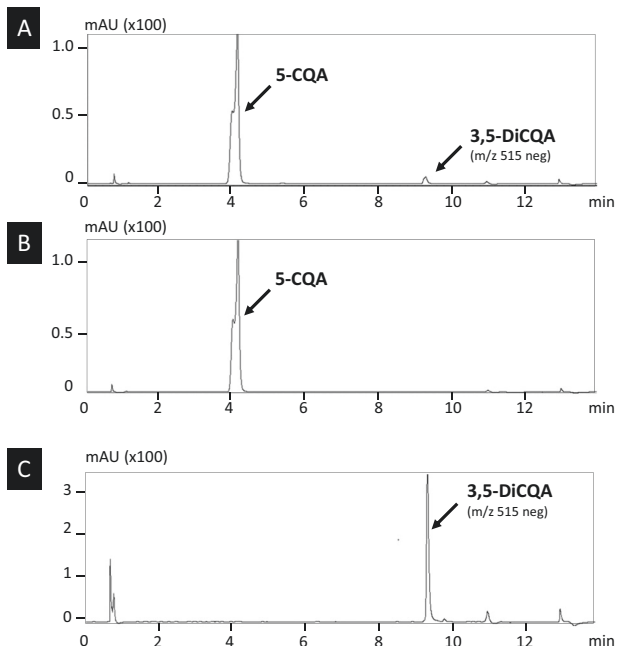
has been referred to as “moonlight activity” by Moglia and collaborators<sup>37</sup>.

Using a proteomic/transcriptomic approach, a completely new player in the synthesis of these DiCQAs was isolated. *IbICS* is a gene encoding an enzyme belonging to a subgroup of the GDSL family called SGNH hydrolase. The name of this enzyme derives from the four strictly conserved Ser-Gly-Asn-His residues in four conserved blocks: I, II, III and V<sup>38</sup>. It has been reported that this enzyme is generally involved in the regulation of development, morphogenesis and the synthesis of specialized metabolites and compounds involved in defence mechanisms<sup>35</sup>. Our experiments showed that this enzyme could be the main actor in the synthesis of 3,5-DiCQA in sweet potato. Whereas the experiments performed with IbHCT led to the production of trace amounts of 3,5-DiCQA, we were able to convert nearly 80% of the 5-CQA that was added to the reaction mix containing IbICS in in vitro experiments. An enzyme using chlorogenate as an acyl donor (a chlorogenate-dependent caffeoyl transferase) was previously described in tomato roots, but no molecular evidence was provided<sup>39</sup>. In this case, the enzyme transfers the caffeoyl group of

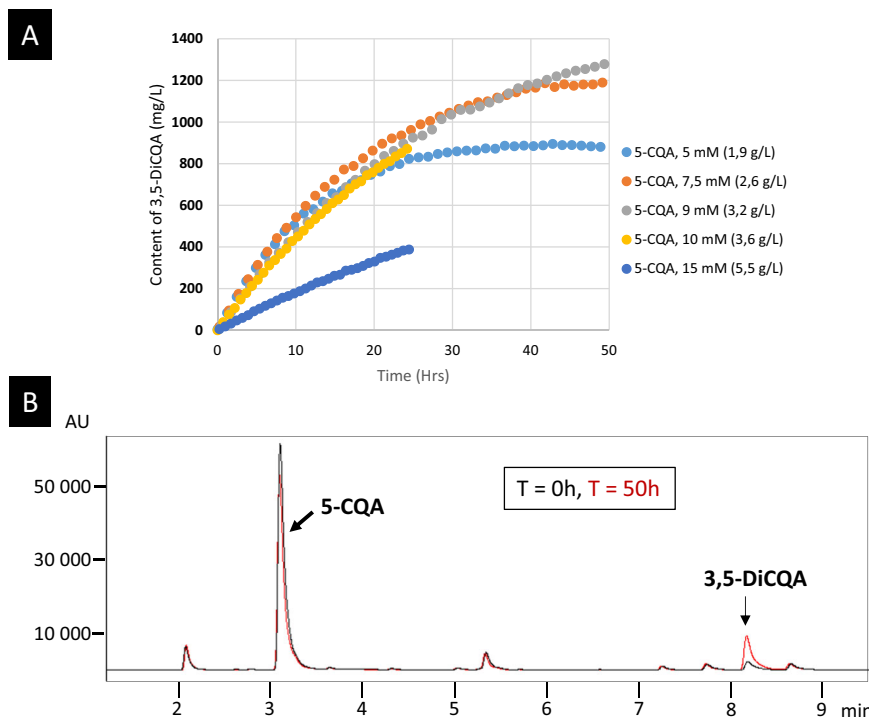
5-CQA to ethanol to produce ethyl caffeate. This enzyme is inhibited by PMSE, which is a feature common to characterized members of the GDSL family and suggests that this enzyme could

also belong to the GDSL family<sup>31,40,41</sup>. This finding indicates an evolutionarily driven metabolic diversification that was already highlighted for SICGT<sup>42</sup>. BAHD and SCPL acyltransferases are currently considered the two main families of acyltransferases involved in phenylpropanoid metabolism<sup>40</sup>. In the future, it might be relevant to add the GDSL family if additional examples of such activity are evidenced.

The expression of IbICS has been successfully realized in four heterologous expression systems. Under standard conditions, the expression in *S. cerevisiae* and *E. coli* was effective for both systems but no activity in 3,5-DiCQA synthase was detected. This finding could be related to a poor solubilization of the proteins or to post-translational modification. Regarding the second possibility, an in silico analysis using the online GlycoEP tool (<http://crdd.osdd.net/raghava/glycoep/submit.html>) revealed 3 putative N-glycosylation sites (N136, N319 and N334) and 7 putative O-glycosylation sites (T51, T69, T234, T320, T347, T355 and T356). The third expression system based on agroinfiltration of *N. benthamiana* leaves led to the production of a functional enzyme able to convert 5-CQA into 3,5-DiCQA. This result is consistent with other studies. For example, Tan et al. could not express a functional *Brassica napus* GDSL esterase (BnGLIP) in bacteria or yeast but succeeded in *N. benthamiana* leaves. In this system, these researchers obtained up to 50 µg of recombinant protein from 1 g of leaves<sup>40</sup>. Similarly, expression of SICGT, a GDSL lipase from tomato, failed to be active in *E. coli* and yeast but was successfully produced and active in tobacco plants<sup>42</sup>. The plant expression system has also been used for the expression of CpLIP2, a GDSL lipase isolated from *Carica papaya*<sup>41</sup>. 3,5-DiCQA has been identified as a potential anti-viral and anti-inflammatory candidate molecule but it is scarcely available at industrial scale considering the low concentrations found in plants<sup>43</sup>. Hence, producing this molecule in high amounts is an interesting challenge. In our hands, the best method was the *P. pastoris* expression platform. The main advantages of this



**Fig. 6 HPLC analyses of the in vitro metabolization product of 5-CQA.** **a** 3,5-DiCQA Commercial standard. **b** Incubation of 5-CQA with a protein mix prepared from leaves of *N. benthamiana* infiltrated with recombinant *A. tumefaciens* transformed with an empty plasmid. **c** Incubation of 5-CQA with purified IbICS-HIS produced in *N. benthamiana* leaves. Analyses were performed at 330 nm. The identity of 3,5-DiCQA was confirmed by MS.



**Fig. 7 3,5-DiCQA synthesis using a bioconversion strategy with *P. pastoris* as a host platform.** **a** Time course of incubation of 3,5-DiCQA in the presence of different concentrations of pure 5-CQA. **b** Evaluation of the conversion efficiency in the presence of green coffee extract mix containing a high intrinsic amount of 5-CQA. Incubation was performed for 50 h. The analysis was performed at 330 nm.



system are that the protein is excreted in the culture medium and the IbICS produced is active. Other GDSL lipases were described to be active and are secreted out of the plant tissue. For example, JNP1, a GDSL lipase/esterase isolated from *Jacaranda mimosifolia*, a tropical tree, could be highlighted in the floral nectar<sup>44</sup>. GLIP1, another enzyme belonging to this same family, has been detected in the secretome of *Arabidopsis thaliana*. In this case, the enzyme is produced in response to salicylic acid treatments<sup>45</sup>.

The proposed *Pichia*-based bioconversion system is rather efficient since 61.3% of the 5-CQA could be converted, leading to a final concentration of 1.2 g/L 3,5-DiCQA when starting from 2.6 g/L 5-CQA. This system could be further improved by using a fed batch approach consisting in adding substrate during the process without reaching substrate toxic concentration, which seems to be a limiting factor. Another possible process improvement might be to replace the pure substrate molecule and use a plant extract harbouring high concentration of 5-CQA. This last strategy was tested in our experiments but the conversion rate of the 5-CQA present in a green coffee extract could only reach 36%. This result might be related to the presence of numerous specialized metabolites that can display an inhibitory effect on the GDSL enzyme. Finally, IbICS is soluble and active in aqueous solution. However, in our complex culture medium, with an improper pH and temperature, there might be a decline in the enzymatic activity, despite preservative buffers. The production yield might, for example, be increased through enzyme immobilization approaches that would preserve it from degradation. Such approaches were investigated in several cases and were described to provide interesting benefits, such as convenient handling, easy and efficient recovery of the product and reuse of the same enzyme batch in several reactions of conversion<sup>46,47</sup>. This finding was reported for a lipase isolated from *Candida rugosa* and immobilized on magnetic hollow mesoporous silica microspheres. In this case, the enzyme was used with fatty acids and triglycerides in a solvent-free system. The authors obtained a conversion rate of 90%, and the enzyme could be reused in 50 successive cycles.

### Other potential players involved in 3,5-DiCQA biosynthesis.

Our results also raise new questions. The investigation of the biochemical properties of IbICS revealed some discrepancies with the pioneering work of Villegas et al.<sup>31</sup>. First, during the purification process, ICS activity was detected in protein fractions precipitated with 40% ammonium sulphate, while in a previous report, the enzyme precipitated between 40 and 70% saturation. Second, IbICS has a calculated molecular weight of 39.54 kDa and a pI of 4.92, whereas the chlorogenic acid:chlorogenate caffeoyl transferase isolated by Villegas et al. had a deduced molecular weight of 25 kDa and a pI of 4.6. Third, the kinetic parameters are highly different. We found a  $K_m$  of 3.5 mM and an optimal pH of 6.3, while Villegas determined a  $K_m$  of 0.87 mM and an optimal pH of 5.0, with the enzyme losing 50% of its activity at pH 6.8. Taken together, these differences between the two studies suggest that additional enzymes might be involved in the outstanding efficient synthesis of 3,5-DiCQA in sweet potato roots. In our study, we overcome a new step thanks to the identification of IbICS. However, given the complexity of specialized metabolism, it is possible that we have not reached the end of the story. Further experiments are needed to test new hypotheses, such as the deletion of *IbICS* in sweet potato, through genome editing.

### Methods

**Plant material.** *I. batatas* were grown in hydroponic conditions (Supplementary Fig. 4). Young cuttings of sweet potato roots were soaked in water for 2 weeks to induce root generation. During the following 3 weeks, the plants were cultivated either in nitrogen-enriched medium (15:10:30, N:P:K, PlantProd) or

in nitrogen-free medium (0:15:40, PlantProd) with electroconductivity adjusted each day to 1.0–1.2 and a pH of 5.6–5.8.

For enzyme purification, sweet potato tubers were purchased at a local supermarket. Discs (1 cm diameter) were prepared from tuber slices (5 mm thick) using a cork borer, rinsed with distilled water and immediately placed on moist filter paper in closed plastic boxes. Discs were incubated at 25 °C in the dark for 48 h and subsequently used as a source for identifying the enzyme.

*N. benthamiana* seeds were sown on soil and cultivated in culture rooms under a 16 h/8 h day/night photoperiod with artificial light (70  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) at 26 °C and 70% humidity. Three-week-old plantlets were transplanted into individual pots and were used after 3–4 weeks for agroinfiltration experiments, as described in Munakata et al.<sup>48</sup>.

**Cloning, expression in *E. coli* and purification of IbHCT.** To amplify the *IbHCT* coding sequence (ID: BAJ14794), we used *IbHCT\_Fw*: 5'-CGGTTTGGATC-CAATGAAGATCAGCGTGAAGGAC-3' and *IbHCT\_Rev*: 5'-CGCTCGAGAA-TATCATACAGGAAGACTCTTTGAAG-3' primers. First-strand cDNAs were synthesized from RNA extracted from *I. batatas* roots cultivated in nitrogen-free medium and using the SuperScript™ III One-Step RT-PCR System with Platinum™ Taq High Fidelity DNA Polymerase (Invitrogen™, Thermo Fisher Scientific). The PCR products were ligated into the pCR8 plasmid (Invitrogen™, Thermo Fisher Scientific) and further transferred between the *NdeI* and *XhoI* restriction sites in the pET28b expression vector (Addgene, <https://www.addgene.org/>). Expression in *E. coli* and purification of the proteins were performed as described by Comino et al.<sup>3</sup>.

**Enzymatic in vitro assays of IbHCT.** Synthesis of 5-CQA, cinnamoyl-shikimate, *p*-coumaroyl-shikimate and caffeoyl-shikimate from cinnamoyl-, *p*-coumaroyl- and caffeoyl-CoA, respectively, with quinic or shikimic acid was performed in a volume of 400  $\mu\text{L}$  with 0.6 mM of ester-CoA (*p*-coumaroyl-, cinnamoyl-, caffeoyl-CoA), 0.6 mM of quinic acid (or shikimic acid) and 15 mg/mL of purified enzyme in 0.1 M NaPi at pH 7. The reactions were incubated at 30 °C during 1 h. The reactions were stopped by adding of 10  $\mu\text{L}$  of mix acetonitrile/HCl 2 M and the solution was mixed and centrifuged at 10,000  $\times g$  for 10 min at room temperature. Products were then purified by ethyl acetate extraction for HPLC analyses.

Synthesis of 3,5-DiCQA from caffeoyl-CoA and 5-CQA was performed in a volume of 200  $\mu\text{L}$  with 0.1 mM of caffeoyl-CoA, 0.1 mM of 5-CQA and 15 mg/ml of purified enzyme in 0.1 M NaPi at pH6. The reaction was incubated at room temperature during 15 h, stopped by adding of 5  $\mu\text{L}$  of mix acetonitrile/HCl 2 M, mixed and centrifuged at 10,000  $\times g$  for 10 min at room temperature. Product was directly analyzed in HPLC and MS.

**Construction of an RNA-seq library.** Total messenger RNA (mRNA) was extracted from *I. batatas* roots cultivated in nitrogen-free medium using the Plant RNeasy™ kit (Qiagen). A de novo normalized RNA-seq library was constructed using GS FLX + chemistry by Eurofins MWG Operon (Ebersberg, Germany). Deep sequencing of 600,000 ESTs was performed. Clustering and assembly of the reads were performed by the supplier.

### Preparation of the IbICS-enriched protein fraction and monitoring of ICS

**activity.** Sweet potato tuber discs (200 g) were chopped with a Waring Blender in the presence of precooled acetone (–20 °C) (4 times 30 s runs with 30 s cooling-off periods). The homogenate was filtered through filter paper (Whatman), and the residue was washed with cold acetone (–20 °C) until the flow-through became transparent. After drying at room temperature, the acetone powder was homogenized in extraction buffer (50 mM Hepes-KOH, 10% glycerol (v/v), 2 mM EDTA, 10 mM  $\text{MgCl}_2$ , 1 mM DTT, 1 mM PMSF, 10% PVPP (m/v), pH 7.5). The suspension was further filtered on Miracloth and centrifuged at 12,000  $\times g$  at 4 °C for 20 min. The proteins present in the supernatant were precipitated using ammonium sulphate (40% saturation) before being suspended in buffer A (50 mM Tris-HCl, pH 7.0). After centrifugation to remove insoluble residues (4000  $\times g$ , 5 min, 4 °C), the extract was loaded on a Sephadex G-50 column equilibrated with buffer A, and the proteins eluted with buffer A (0.4 mL  $\text{min}^{-1}$ ). One-millilitre fractions were collected. The fractions containing enzyme activity were pooled and concentrated using a Vivaspinn 6 centrifugal concentrator (100,000 MWCO, Sartorius). The protein solution was applied to a DEAE Sepharose anion exchange column equilibrated with buffer A and eluted with a step gradient of NaCl (from 0.15 to 1 M) prepared in 50 mM Tris-HCl, pH 7.0, at a flow rate of 1 mL  $\text{min}^{-1}$ . Active fractions were pooled and concentrated through a Vivaspinn 6 centrifugal concentrator. SDS-PAGE analysis was conducted as described elsewhere<sup>49</sup>. After electrophoresis, one lane of the gel was stained with Coomassie blue, while another lane was washed twice with 2% Triton (v/v) to remove SDS, which might interfere with the enzyme activity. To localize enzyme activity, the two lanes were identically and separately cut into 8 slices (between 20 and 75 kDa, 1 cm height). The slices of the lane prewashed with Triton were ground, buffer A was added to solubilize the proteins, and thereafter, enzyme activity was measured in each preparation. Slices containing activities (4 slices) were identified, and the corresponding slices of the lane stained with Coomassie blue were subjected to LC/MS/MS analysis.

For the standard assay of IbICS activity, enzyme solutions were incubated at 30 °C for 15 min in 35 mM acetate buffer, pH 5.0, containing 1 mM 5-CQA and 10 mM MgCl<sub>2</sub> in a total volume of 100 µL. Reactions were stopped by placing the tubes in boiling water for 10 min. After centrifugation at 12,000 ×g for 3 min at 4 °C and filtration (0.45 µm filter), the reaction mixtures were analyzed by HPLC to quantify 3,5-DiCQA.

**Protein MS analyses.** The proteins collected by SDS-PAGE were digested in the gel in the presence of 50 mM ammonium bicarbonate, 0.01% Protease Max (Promega) and 12 ng µL<sup>-1</sup> of sequencing grade trypsin (Promega). After 2 h at 37 °C, the liquid phase of the digestion was acidified with trifluoroacetic acid (final concentration: 0.5% (v/v)). Samples were cleaned with an off-line C18 column before analysis by mass spectrometry.

The peptides were analyzed by ESI-LC/MS/MS using a nano LC system (Waters) coupled to a SynaptG2 mass spectrometer. Separation was performed using an Acquity BEH C18 nano reverse phase (150 cm) column (Waters) at a flow rate of 300 nL min<sup>-1</sup> with a mobile phase that consisted of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile). Chromatographic separation was achieved using a 180 min linear gradient from 0 to 50% solvent B. Processing, deconvolution and peptide detection were performed using Data Analysis software (Protein Lynx global server, Waters).

**Bioinformatic analysis.** The FASTA file provided by MWG Eurofins Operon (Ebersberg, Germany) was imported as a library into BioEdit software (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). The peptide sequences obtained by MS analysis were compared with the RNA data available in the cDNA library using the Basic Local Alignment Search Tool (tBLASTn) to identify the corresponding coding sequences. A global analysis of the RNA-seq database was performed with the Blast2GO platform.

The prediction of putative N-terminal signal peptide (SP) was performed using SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) and TargetP 2.0 (<http://www.cbs.dtu.dk/services/TargetP/>). The molecular weight and isoelectric point (pI) of the protein and the number of amino acids were calculated using the ExPASy compute pI/MW tool ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). To predict putative amino acid sites for glycosylation, we used the NetNGlyc 1.0 and NetOGlyc 4.0 servers (<http://www.cbs.dtu.dk/services/NetNGlyc/> and <http://www.cbs.dtu.dk/services/NetOGlyc/>) at the Center for Biological Sequence Analysis website (<http://www.cbs.dtu.dk/services/>), as well as the GlycoEP tool (<http://crdd.osdd.net/raghava/glycoep/submit.html>).

The phylogenetic analyses were performed using Clustal W<sup>50</sup>. Pairwise distance matrices were generated using MEGA X<sup>51</sup>. Phylogenetic trees were reconstructed using the neighbour-joining algorithm from MEGA X.

#### Cloning of IbICS for heterologous expression in *E. coli*, *S. cerevisiae*, *N. benthamiana* and *P. pastoris*.

To amplify the IbICS coding sequence (CDS), we used IbICS\_Fw: 5'-ATGGCCGCGCAGCTCCTTCTCCTCGTTGCTTG-3'; IbICS\_Rev: 5'-TTAGTGGTGGTGGTGGTGGTGAATTTGGGCAACTCGAGATGTTTTCGCA-3' primers. The reverse primer included the addition of a 6xHis tag at the C-terminal end of the protein (underlined in the sequence). The CDS was amplified from total RNA using the SuperScript™ III One-Step RT-PCR System with Platinum™ Taq High Fidelity DNA Polymerase (Invitrogen™, Thermo Fisher Scientific). The PCR product was ligated into the TA cloning vector pCR™8/GW/TOPO™ system (Invitrogen™, Thermo Fisher Scientific) and subsequently sequenced. The gene was further cloned by LR recombination using LR Clonase II™ (Invitrogen™, Thermo Fisher Scientific) into pEAQ-HT-DEST1<sup>52</sup> and pYedP60GW<sup>53</sup> for expression in *N. benthamiana* and *S. cerevisiae*, respectively. The gene without signal peptide was amplified thanks to IbICS\_Rev associated with IbICS\_Ecoli\_Fw: 5'-ATGTGCTATTCGCGCGTTTTCGGC-3' from pEAQ-HT-IbICS, ligated into pCR™8/GW/TOPO™ system and cloned by LR recombination into p0GWA<sup>54</sup> for expression in *E. coli*.

To express IbICS in *P. pastoris* and direct the protein in the extracellular space, the putative N-terminal peptide signal of IbICS was replaced by the α-factor secretion signal of *S. cerevisiae*. The CDS was amplified with IbICS\_Pichia\_Fw: 5'-GAATTCGCTATTCGCGCGTTTTCGGC-3' and IbICS\_Pichia\_Rev: 5'-TCTAGAAAAATTTGGGCAACTCGAGATGTTTTCGAG-3' primers. Additional EcoRI and XbaI restriction sites were integrated (underlined nucleotides) to allow insertion into the pPICZaA vector (Invitrogen™, Thermo Fisher Scientific).

#### Expression, purification and in vitro assays of recombinant IbICS produced in *N. benthamiana*.

The recombinant pEAQ-HT-IbICS plasmid (and the negative control corresponding to pEAQ-HT without the gene of interest) was introduced into *Agrobacterium tumefaciens* EHA105 using the freeze-thaw method<sup>55</sup>. The recombinant bacteria were selected on YEB medium (10 g L<sup>-1</sup> beef extract, 5 g L<sup>-1</sup> yeast extract, 10 g L<sup>-1</sup> peptone, 15 g L<sup>-1</sup> sucrose, 0.5 g L<sup>-1</sup> MgSO<sub>4</sub>, 20 g L<sup>-1</sup> at pH 7.2) supplemented with 100 mg L<sup>-1</sup> rifampicin and 30 mg L<sup>-1</sup> kanamycin at 28 °C for 3 days.

For transient expression in *N. benthamiana*, recombinant agrobacteria were cultured in YEB medium containing 100 mg L<sup>-1</sup> rifampicin and 30 mg L<sup>-1</sup>

kanamycin at 28 °C for 2 days at 180 rpm. Four hours before infiltration, acetosyringone was added to the cultures at a final concentration of 100 µM. The agrobacteria were centrifuged for 15 min at 5000 ×g, and the pellet was washed twice with YEB medium to remove antibiotics. The pellet was finally suspended in infiltration buffer (10 mM MES, pH 5.6) at a cell density corresponding to OD<sub>600</sub> 0.8 ± 0.1. The whole aerial parts of *N. benthamiana* plants were immersed in the agrobacteria suspension and subjected to vacuum infiltration at 15 mbar for 20 min. Finally, the infiltrated plants were cultivated under a 16 h/8 h day/night photoperiod at 26 °C and 70% humidity. Leaves were harvested 6 days post infiltration and stored at -80 °C until use.

Total protein extract was obtained by grinding the leaves in the presence of extraction buffer (20 mM sodium phosphate, 0.5 M NaCl, pH 7.4) using a Polytron™ PT2500E (Kinematica). The crude plant extract was clarified by centrifugation at 14,000 ×g for 10 min at 4 °C. The supernatant was further filtered through a 0.2 µm filter and loaded onto 1 mL HisTrap HP columns (GE Healthcare). Due to elution with 500 mM imidazole, one-millilitre fractions were collected and individually analyzed by SDS-PAGE. The fractions harbouring the target protein were pooled, diafiltered with 5 volumes of phosphate buffer and concentrated using centrifugal 10 kDa molecular mass cutoff filters (Amicon™ Ultra 0.5 mL Centrifugal Filters — PMNL 10 kDa, MILLIPORE). Purified proteins were stored at -20 °C.

The standard enzymatic in vitro assay was performed in a volume of 50 µL with 0.2–0.4 µg of purified enzyme. For evaluation of the optimal pH, the reaction mixture was incubated for 30 min at 25 °C in a buffer (0.1 M acetic acid, 0.1 M Tris, 20 mM MES) at pH values ranging from 4.0 to 9.0 with a single 5-CQA concentration of 10 mM. To evaluate the best substrate concentration, the 5-CQA concentration was modulated from 0 to 30 mM. The reactions were stopped by adding 150 µL absolute ethanol, and the solution was mixed and centrifuged at 14,000 ×g for 10 min at 4 °C. The supernatant was analyzed by UHPLC-MS. Reactions were run in triplicate. K<sub>m</sub> value was determined in triplicate by fitting Michaelis–Menten curves using SigmaPlot Software.

**Expression of IbICS in *E. coli* and *S. cerevisiae*.** Expression in *E. coli* was realized in the BL21 Rosetta strain, and the culture was performed as for CchCT described in Comino et al.<sup>3</sup>. Expression in *S. cerevisiae* was performed in the WAT11 yeast strain<sup>56</sup>, as described for cytochrome P450s<sup>57</sup>.

#### Expression and in vivo bioconversion of 5-CQA to 3,5-DiCQA by recombinant IbICS produced by *P. pastoris*.

The *P. pastoris* X-33 strain (Invitrogen™, Thermo Fisher Scientific) was transformed as specified by the supplier with pPICZaA-IbICS or empty vector pPICZaA used as a negative control. The culture of recombinant *P. pastoris* was performed as described by the supplier. Briefly, a colony was picked into 25 mL of culture medium including glycerol at 28 °C at 250 rpm for 16–18 h. The yeast was centrifuged at 5000 ×g, and the pellet was suspended in 250 mL of culture medium buffered at pH 6.0 and containing 0.5% methanol as an inducer. The induction was maintained for 3 days and supplied with methanol every 24 h. Three days after the onset of induction, 10 mM 5-CQA (final concentration) was added directly to the culture medium, and 3 days later, the supernatant of *Pichia* culture was sampled and analyzed by UHPLC-MS.

#### Determination of protein concentrations and immunoblot analysis.

Protein concentrations of the extracts were determined using the Qubit™ Fluorometer (Invitrogen™, Thermo Fisher Scientific) with the Qubit™ protein assays kit (Invitrogen™, Thermo Fisher Scientific) according to the supplier's instructions. For immunodetection, 25 µL of protein mix was analyzed by SDS-PAGE. The separated proteins were blotted on a polyvinylidene fluoride membrane (Membrane PVDF 0.45 µm Amersham™ Hybond™ P, GE Healthcare). His-tagged IbICS was detected using a rabbit primary polyclonal anti-6xHistidine Epitope Tag (Cat#R1181, Acris) at a 1:5000 dilution in PBS (137 mM NaCl, 2.7 mM KCl, 8.1 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) and a secondary anti-rabbit antibody conjugated to alkaline phosphatase activity (Cat#A0418–1ML, Anti-Rabbit IgG (whole molecule)–Alkaline Phosphatase antibody produced in goat, Sigma-Aldrich) diluted to 1:6000 in PBS. Revelation was performed with NBT/BCIP (Promega, USA) as substrates.

**Chromatography analyses.** Quantification of 5-CQA and 3,5-DiCQA was performed by:

- UHPLC Shimadzu Prominence XR system (Shimadzu, Kyoto, Japan) with a PDA detector (applied range: 190–800 nm, detection at 300 nm) using a LiChrospher™ 100 RP-18 reverse phase column (250 mm×4.6 mm, 5 µm, Merck, Darmstadt, Germany), maintained at 30 °C during all analyses. The mobile phase was composed of water, pH 2.6 adjusted using orthophosphoric acid (A) and pure methanol (B), delivered at 0.7 mL/min with the gradient of B phase as follows: 23–55% (0–43 min); 55–70% (43–50 min); 70% (50–55 min), 70–23% (55–60 min), 23% (60–65 min);
- UHPLC Shimadzu Nexera X2 system (Shimadzu, Kyoto, Japan) with a PDA detector (applied range: 220–370 nm, detection at 300 nm) coupled to a mass spectrometer LCMS2020 (electrospray ionization in a negative ion mode, *m/z*

100–1000) using a Kinetex Biphenyl reverse phase column (150 mm × 2.1 mm, 2.6 µm, Phenomenex, Torrance, CA, USA), maintained at 40 °C during all analyses. The mobile phase was composed of water containing 0.1% formic acid (A) and pure acetonitrile (B), delivered at 0.5 ml/min with a gradient of B phase as follows: 5–25% (0–10 min); 25–90% (10–10.5 min) 90% (10.5–12 min), 90–5% (12–12.1 min), and 5% (12.1–14.1 min).

The contents of 3,5-DiCQA and 5-CQA were determined using the standards of 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The concentrations of other hydroxycinnamate derivatives were expressed as 3,5-DiCQA equivalents.

### Data availability

The data that support the findings of this study have been deposited in Genbank (<https://www.ncbi.nlm.nih.gov>). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The IBCS CDS has the ID: MT291823.

Received: 26 April 2020; Accepted: 9 October 2020;

Published online: 13 November 2020

### References

- Biesalski, H. K. Polyphenols and inflammation: basic interactions. *Curr. Opin. Clin. Nutr.* **10**, 724–728 (2007).
- Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* **278**, 95–103 (2003).
- Comino, C. et al. Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. *BMC Plant Biol.* **7**, 1–14 (2007).
- Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat. Biotechnol.* **22**, 746–754 (2004).
- Legrand, G. et al. Identification and characterization of five baoh acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.* **7**, 741 (2016).
- Perez-Jimenez, J. et al. Dietary intake of 337 polyphenols in French adults. *Am. J. Clin. Nutr.* **93**, 1220–1228 (2011).
- Grosso, G. et al. Dietary polyphenols are inversely associated with metabolic syndrome in Polish adults of the HAPIEE study. *Eur. J. Nutr.* **56**, 1409–1420 (2017).
- Julia, C. et al. Cluster analysis of polyphenol intake in a French middle-aged population (aged 35–64 years). *J. Nutr. Sci.* **5**, e28 (2016).
- Nascimento-Souza, M. A., de Paiva, P. G., Pérez-Jiménez, J., do Carmo Castro Franceschini, S. & Ribeiro, A. Q. Estimated dietary intake and major food sources of polyphenols in elderly of Viçosa, Brazil: a population-based study. *Eur. J. Nutr.* **57**, 617–627 (2018).
- Kurata, R., Adachi, M., Yamakawa, O. & Yoshimoto, M. Growth suppression of human cancer cells by polyphenolics from sweetpotato (*Ipomoea batatas* L.) leaves. *J. Agric. Food Chem.* **55**, 185–190 (2007).
- Zhu, F., Cai, Y. Z., Yang, X. S., Ke, J. X. & Corke, H. Anthocyanins, hydroxycinnamic acid derivatives, and antioxidant activity in roots of different chinese purple-fleshed sweetpotato genotypes. *J. Agric. Food Chem.* **58**, 7588–7596 (2010).
- Kim, J. Y. et al. Neuroprotection of *Ilex latifolia* and caffeoylquinic acid derivatives against excitotoxic and hypoxic damage of cultured rat cortical neurons. *Arch. Pharm. Res.* **35**, 1115–1122 (2012).
- Han, J., Miyamae, Y., Shigemori, H. & Isoda, H. Neuroprotective effect of 3,5-di-O-caffeoylquinic acid on SH-SY5Y cells and senescence-accelerated-prone mice 8 through the up-regulation of phosphoglycerate kinase-1. *Neuroscience* **169**, 1039–1045 (2010).
- Gray, N. E. et al. Caffeoylquinic acids in *Centella asiatica* protect against  $\beta$ -amyloid toxicity. *J. Alzheimers Dis.* **40**, 359–373 (2014).
- Kang, J. Y. et al. Reversal of trimethyltin-induced learning and memory deficits by 3,5-dicaffeoylquinic acid. *Oxid. Med. Cell Longev.* **2016**, <https://doi.org/10.1155/2016/6981595> (2016).
- Anggreani, E. & Lee, C. Y. Neuroprotective effect of chlorogenic acids against Alzheimer's disease. *Int J. Food Sci. Nutr. Diet.* **6**, 330–337 (2017).
- Li, Y., But, P. P. H. & Ooi, V. E. C. Antiviral activity and mode of action of caffeoylquinic acids from *Schefflera heptaphylla* (L.) Frodin. *Antivir. Res.* **68**, 1–9 (2005).
- Shi, S., Huang, K., Zhang, Y., Zhao, Y. & Du, Q. Purification and identification of antiviral components from *Laggera pterodonta* by high-speed counter-current chromatography. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **859**, 119–124 (2007).
- Sinisi, V. et al. Chlorogenic compounds from coffee beans exert activity against respiratory viruses. *Plant. Med.* **83**, 615–623 (2017).
- Chen, Y.-L. et al. *Ilex kaushue* and its bioactive component 3,5-dicaffeoylquinic acid protected mice from lipopolysaccharide-induced acute lung injury. *Sci. Rep.* **6**, 1–12 (2016).
- Gao, X. et al. Anti-inflammatory effects of neochlorogenic acid extract from mulberry leaf (*Morus alba* L.) against LPS-stimulated inflammatory response through mediating the AMPK/Nrf2 signaling pathway in A549 cells. *Molecules* **25**, 1385 (2020).
- Oh, H., Kang, D.-G., Lee, S. & Lee, H.-S. Angiotensin converting enzyme inhibitors from *Cuscuta japonica* Choisy. *J. Ethnopharmacol.* **83**, 105–108 (2002).
- Ishiguro, K., Yoshimoto, M., Tsubata, M. & Takagaki, K. Hypotensive effect of sweetpotato tops. *J. Jpn. Soc. Food Sci. Technol.* **54**, 45–49 (2007).
- Wu, Y. Compensation of ACE2 function for possible clinical management of 2019-nCoV-induced acute lung injury. *Virology* **35**, 256–258 (2020).
- Schoch, G. et al. CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J. Biol. Chem.* **276**, 36566–36574 (2001).
- Comino, C. et al. The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway. *BMC Plant Biol.* **9**, 1–13 (2009).
- Moglia, A. et al. Dual catalytic activity of hydroxycinnamoyl-CoA quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicaffeoylquinic acids. *Plant Physiol.* **166**, 1777–1787 (2014).
- Villegas, R. J. A. & Kojima, M. Purification and characterization of hydroxycinnamoyl D-glucose - quinate hydroxycinnamoyl transferase in the root of sweet-potato, *Ipomoea batatas* Lam. *J. Biol. Chem.* **261**, 8729–8733 (1986).
- Barros, J. et al. 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nat. Commun.* **10**, 1–11 (2019).
- Vanholme, R. et al. Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* **341**, 1103–1106 (2013).
- Villegas, R. J. A., Shimokawa, T., Okuyama, H. & Kojima, M. Purification and characterization of chlorogenic acid: chlorogenate caffeoyl transferase in sweet potato roots. *Phytochem.* **26**, 1577–1581 (1987).
- Lallemant, L. A. et al. A Structural basis for the biosynthesis of the major chlorogenic acids found in coffee. *Plant Physiol.* **160**, 249–260 (2012).
- St-Pierre, B. & De Luca, V. in *Recent Advances in Phytochemistry* Vol. 34 (eds. Romeo, J. T., Ibrahim, R., Varin, L. & De Luca, V.) 285–315 (Elsevier, 2000).
- Kojima, M. & Kondo, T. An enzyme in sweet-potato root which catalyzes the conversion of chlorogenic acid, 3-caffeoylquinic acid, to isochlorogenic acid, 3,5-dicaffeoylquinic acid. *Agric. Biol. Chem.* **49**, 2467–2469 (1985).
- Chepyshko, H., Lai, C.-P., Huang, M., Liu, J.-H. & Shaw, J.-F. Multifunctionality and diversity of GDSL esterase/lipase gene family in rice (*Oryza sativa* L. *japonica*) genome: new insights from bioinformatics analysis. *BMC Genomics* **13**, 309 (2012).
- Galièni, A. et al. Effects of nutrient deficiency and abiotic environmental stresses on yield, phenolic compounds and antiradical activity in lettuce (*Lactuca sativa* L.). *Sci. Hortic.* **187**, 93–101 (2015).
- Moglia, A. et al. Isolation and mapping of a C3H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress. *Plant Cell Rep.* **28**, 963–974 (2009).
- Akoh, C. C., Lee, G. C., Liaw, Y. C., Huang, T. H. & Shaw, J. F. GDSL family of serine esterases/lipases. *Prog. Lipid Res.* **43**, 534–552 (2004).
- Negrel, J., Javelle, F. & Morandi, D. Detection of a plant enzyme exhibiting chlorogenate-dependant caffeoyltransferase activity in methanolic extracts of arbuscular mycorrhizal tomato roots. *Plant Physiol. Biochem.* **66**, 77–83 (2013).
- Tan, X. et al. Characterization and expression of a GDSL-like lipase gene from *Brassica napus* in *Nicotiana benthamiana*. *Protein J.* **33**, 18–23 (2014).
- Gutierrez-Ortega, A. et al. Identification, cloning and expression of a new GDSL lipase from *Carica papaya*. *Process Biochem.* **55**, 70–76 (2017).
- Teutschbein, J. et al. Identification and localization of a lipase-like acyltransferase in phenylpropanoid metabolism of tomato (*Solanum lycopersicum*). *J. Biol. Chem.* **285**, 38374–38381 (2010).
- Moglia, A. et al. Stress-induced biosynthesis of dicaffeoylquinic acids in globe artichoke. *J. Agric. Food Chem.* **56**, 8641–8649 (2008).
- Kram, B. W., Bainbridge, E. A., Perera, M. A. D. N. & Carter, C. Identification, cloning and characterization of a GDSL lipase secreted into the nectar of *Jacaranda mimosifolia*. *Plant Mol. Biol.* **68**, 173–183 (2008).
- Oh, I. S. et al. Secretome analysis reveals an *Arabidopsis* lipase involved in defense against *Alternaria brassicicola*. *Plant Cell* **17**, 2832–2847 (2005).
- Sheldon, R. A. Enzyme immobilization: the quest for optimum performance. *Adv. Synt. Cat.* **349**, 1289–1307 (2007).
- Zheng, M. et al. A mixed-function-grafted magnetic mesoporous hollow silica microsphere immobilized lipase strategy for ultrafast transesterification in a solvent-free system. *RSC Adv.* **5**, 43074–43080 (2015).

48. Munakata, R. et al. Molecular cloning and characterization of a geranyl diphosphate-specific aromatic prenyltransferase from lemon. *Plant Physiol.* **166**, 80–90 (2014).
49. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 3rd edn (Cold Spring Harbor Laboratory Press, 1989).
50. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
51. Kumar, S., Nei, M., Dudley, J. & Tamura, K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* **9**, 299–306 (2008).
52. Peyret, H. & Lomonosoff, G. P. The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants. *Plant Mol. Biol.* **83**, 51–58 (2013).
53. Krieger, C. et al. The CYP71AZ P450 Subfamily: A Driving Factor for the Diversification of Coumarin Biosynthesis in Apiaceous Plants. *Front. Plant Sci.* **9**, 820 (2018).
54. Busso, D., Delagoutte-Busso, B. & Moras, D. Construction of a set gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Anal. Biochem.* **343**, 313–321 (2005).
55. Chen, H., Nelson, R. S. & Sherwood, J. L. Enhanced recovery of transformants of *Agrobacterium tumefaciens* after freeze-thaw transformation and drug selection. *BioTechniques* **16**, 664–668 (1994).
56. Pompon, D., Louerat, B., Bronnie, A. & Urban, P. Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* **272**, 51–64 (1996).
57. Larbat, R. et al. Molecular cloning and functional characterization of psoralen synthase, the first committed monooxygenase of furanocoumarin biosynthesis. *J. Biol. Chem.* **282**, 542–554 (2007).

### Acknowledgements

The authors would like to thank Clément Charles for his technical support. This work was financially supported by the “Bioprolor2” project (Région Grand-Est) (to A.H.) and by the “Impact Biomolécules” project of the “Lorraine Université d’Excellence” (Investissements d’avenir–ANR) (to A.H.). Experiments performed at Charles Viollette Research Institute were supported by an ALIBIOTECH program (2015–2020) obtained from the CPER/FEDER Région Hauts-de-France. Marianne Delporte and Guillaume Legrand were supported by doctoral fellowships from the doctoral school 104 SMRE.

### Author contributions

M.S. performed the genetic molecular cloning, the bioinformatic analysis and the enzymatic assays in the 4 expression systems for IbICS. L.G., D.M., C.G., V.D., J.L.H. and G.D. purified the protein from the roots of *I. batatas*. B.J. and W.G. performed the MS protein analyses. L.D. cloned, expressed the IbHCT and realized the functional characterization. M.B. constructed the *I. batatas* cDNA library. O.A., S.A. and M.S. performed the enzymatic characterization of IbICS. M.S. realized the experimentations with green coffee extracts and S.A. performed the UHPLC-MS investigations of IbICS. M.S., F.B., H.A., and G.D. wrote the article. H.J.L., B.F., H.A. and G.D. supervised the project and are the principal investigators.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s42003-020-01387-1>.

**Correspondence** and requests for materials should be addressed to A.H. or D.G.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020



RESEARCH PAPER

# A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in the pollen coat of the Asteraceae family

Marianne Delporte<sup>1,✉</sup>, Guillaume Bernard<sup>1</sup>, Guillaume Legrand<sup>1</sup>, Björn Hielscher<sup>2</sup>, Arnaud Lanoue<sup>3</sup>, Roland Molinié<sup>4</sup>, Caroline Rambaud<sup>1</sup>, David Mathiron<sup>5</sup>, Sébastien Besseau<sup>3</sup>, Nicole Linka<sup>2</sup>, Jean-Louis Hilbert<sup>1</sup> and David Gagneul<sup>1,\*,✉</sup>

<sup>1</sup> EA 7394, USC INRA 1411, Institut Charles Viollette (ICV), Agro-food and Biotechnology Research Institute, Université de Lille, INRA, ISA, Univ. Artois, Univ. Littoral Côte d'Opale, Cité Scientifique, 59655 Villeneuve d'Ascq, France

<sup>2</sup> Institute for Plant Biochemistry and Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University, Universitätsstrasse 1, Düsseldorf, 40225, Germany

<sup>3</sup> Biomolécules et Biotechnologies Végétales, EA 2106, Université de Tours, 37200 Tours, France

<sup>4</sup> Biologie des Plantes & Innovation (EA 3900 BIOPI), Université de Picardie Jules Verne, 80037 Amiens Cedex 1, France

<sup>5</sup> Plateforme Analytique (PFA), Université de Picardie Jules Verne, 80039 Amiens Cedex 1, France

\* Correspondence: [david.gagneul@univ-lille.fr](mailto:david.gagneul@univ-lille.fr)

Received 13 April 2018; Editorial decision 6 August 2018; Accepted 6 August 2018

Editor: Dabing Zhang, Shanghai Jiao Tong University, China

## Abstract

In eudicotyledons, accumulation of trihydroxycinnamoyl spermidine that is restricted to the pollen wall constitutes an evolutionary conserved trait. However, the role of this compound, which is synthesized by the BAHD enzyme spermidine hydroxycinnamoyl transferase (SHT), is still a matter of debate. Here, we show that this particular phenolamide is replaced by tetrahydroxycinnamoyl spermine in the pollen coat of the Asteraceae. Phylogenetic analyses combined with quantitative RT-PCR experiments allowed the identification of two homologous genes from *Cichorium intybus* (chicory) putatively involved in its metabolism. *In vitro* biochemical characterization of the two enzymes, named CiSHT1 and CiSHT2, confirmed the capability of recombinant proteins to synthesize spermine as well as spermidine derivatives. The wild-type metabolic phenotype was partially restored in an *Arabidopsis* *sht* mutant expressing CiSHT2. Strikingly, the transgenic plants also accumulated spermine derivatives that were absent in the wild-type. Overexpression of CiSHT2 in chicory hairy roots led to the accumulation of spermine derivatives, confirming its *in vivo* function. Complementary sequence analyses revealed the presence of an amino acid motif typical of the SHTs among the BAHD enzyme family. Our results highlight a recent neofunctionalization among the SHTs that has promoted the emergence of new phenolamides in the Asteraceae, which could potentially have contributed to the evolutionary success of this family.

**Keywords:** Acylated spermine, BAHD acyltransferases, *Cichorium intybus*, metabolic diversification, pollen coat, phenolamides, SHTs.

## Introduction

Evolution within the plant kingdom is highly correlated with the appearance of new metabolic pathways, and notably those leading to the synthesis of specialized (secondary) metabolites. These are responsible for the tremendous adaptability of plants to their challenging environment, and their number has increased constantly since terrestrialization. The

enzymes required for their biosynthesis have consequently expanded in number during the evolution of the green plant lineage (Weng and Noel, 2013). This has mostly been due to gene duplication events, which may have resulted either from whole-genome duplications or, more probably, from local duplications (Chae et al., 2014). Substrate flexibility is a common trait for enzymes involved in these pathways, which leads to rapid neofunctionalization (Weng and Noel, 2012; Weng et al., 2012). Consequently, a number of specialized metabolites occur in a lineage-specific manner and they participate in plant speciation by conferring to the host species the ability to adjust to a particular environment. Nevertheless, some of these metabolites, such as flavonoids, are widely conserved across species, which reveals a positive Darwinian selection and hence highlights their essential functions (Weng and Noel, 2013).

Evolution of the reproductive structures was a key element during colonization of the land habitat and emergence of pollen was a major event, as well as the appearance of the pollen coat among angiosperms. The pollen coat is produced by the anther tapetum (Pacini and Hesse, 2005). This lipidic matrix is then deposited on the surface of the pollen grain where it fills the exine cavities. The pollen coat provides a wide range of functions, including protection against biotic and abiotic stresses, stigma-pollen recognition mechanisms, and pollinator attraction. Its composition can be highly variable among species (Quilichini et al., 2015). However, trihydroxycinnamoyl spermidines are phenolamides that are found in the pollen coat of nearly all Eudicots (Elejalde-Palmett et al., 2015). Phenolamides, also referred to as hydroxycinnamic acid amides or phenylamides, are specialized metabolites resulting from the conjugation of aliphatic or aromatic amines with phenolic acids. They are mostly known for their involvement in floral induction and development, or as defense compounds (Bassard et al., 2010), but fully substituted spermidines ( $N^1, N^5, N^{10}$ -trihydroxycinnamoyl spermidines) are often considered separately due to their specific location in the pollen coat. However, to date, no clear function has been assigned to these compounds, although they probably have a role in determining the ultrastructure of sporopollenin (Elejalde-Palmett et al., 2015; Fellenberg and Vogt, 2015; Aloisi et al., 2016). A lack of pollen-coat phenolamides in an Arabidopsis mutant leads to abnormal formation of the pollen grain but has no effects on its viability or fertility (Grienenberger et al., 2009). Nevertheless, the fact that these particular phenolamides have been maintained over the evolution of the green plant lineage suggests that they have a major significance.

The appearance of trihydroxycinnamoyl spermidine in pollen coat was linked with the emergence of a new class of enzymes, spermidine hydroxycinnamoyl transferases (SHTs), which are specifically expressed in the tapetum of the anther (Grienenberger et al., 2009). These enzymes belong to the clade Vb of the superfamily of BAHD acyltransferases (St-Pierre and De Luca, 2000; Tuominen et al., 2011). This family, named after the four enzymes that were first characterized, is subdivided into eight clades and encompasses acyltransferases that catalyze either O-acylation or N-acylation in an acyl-CoA-dependent

manner (D'Auria, 2006). To date, all the enzymes described that promote phenolamide synthesis from aliphatic amines fall within this BAHD family, as follows: agmatine coumaroyltransferase (ACT) in barley and Arabidopsis (Burhenne et al., 2003; Muroi et al., 2009); spermidine disinnapoyltransferase (SDT) and spermidine dicoumaroyltransferase (SCT) in Arabidopsis (Luo et al., 2009); hydroxycinnamoyl-CoA:spermidine acyltransferase (DH29) and hydroxycinnamoyl-CoA:putrescine acyltransferase (AT1) in *Nicotiana attenuata* (Onkokesung et al., 2012); putrescine hydroxycinnamoyl transferase (PHT) in potato and rice (Tanabe et al., 2016; Yogendra et al., 2017); and SHT from Arabidopsis and *Malus domestica* (Grienenberger et al., 2009; Elejalde-Palmett et al., 2015). The last two of these enzymes are the only ones so far that have been identified as catalysing acylation of secondary amino groups to synthesize fully substituted polyamines (i.e. trihydroxycinnamoyl spermidines).

A few studies have reported another type of fully substituted polyamine, namely tetrahydroxycinnamoyl spermine. This spermine-derived phenolamide was first described in the pollen and/or flowers of *Matricaria chamomilla* and four other Asteraceae species (Yamamoto et al., 2002). It has more recently been found in flowers of *Coreopsis tinctoria* (also Asteraceae) (Lam et al., 2016) as well as in a few other species belonging to the same family (Božičević et al., 2017). Tetracoumaroyl spermine has been identified as the first natural tachykinin NK1 receptor antagonist. As such, this amine could be a good candidate as a drug for the treatment of illnesses such as depression, anxiety, insomnia, and certain cancers (Yamamoto et al., 2002; Park et al., 2017). Another study has reported that synthetic tetracoumaroyl spermine has potent inhibitory effects on HIV-1 protease (Ma et al., 2001). This type of spermine-derivative has also been reported in the composition of bee pollen, where they are considered as an important antioxidant (Ohta et al., 2007). Thus, tetrahydroxycinnamoyl spermines appear to be promising natural medicinal compounds. Accumulation of tetracoumaroyl spermine in *M. chamomilla* has been evaluated during flower ontogenesis and found to be especially concentrated in pollen (Eliašová et al., 2012). This has led us to the hypothesis that tetrahydroxycinnamoyl spermines may replace trihydroxycinnamoyl spermidines in the pollen coat of certain species.

Since the only reports we could find about tetrahydroxycinnamoyl spermine were related to species belonging to the Asteraceae, we first decided to conduct a screen of members of this family for spermine-based phenolamides. Because of the availability of public and in-house genomic resources for chicory (*Cichorium intybus*), we began by examining this species before broadening our analyses to cover the entire family. We used chicory to identify the molecular determinants involved in the metabolic diversification that we observed among the Asteraceae. Our starting assumption was that the last step of the biosynthesis of tetrahydroxycinnamoyl spermine would be related to that of trihydroxycinnamoyl spermidine. Accordingly, we identified and cloned two genes encoding BAHD acyltransferases that are phylogenetically close to previously characterized SHTs. Further investigations demonstrated their

involvement in tetrahydroxycinnamoyl spermine biosynthesis and suggested that the appearance of these new metabolites in the Asteraceae family resulted from the neofunctionalization of a duplicated gene.

## Material and methods

### Plant material

*Cichorium intybus* (L.) line 'CC', an auto-fertile and rapid-cycling line (Habarugira *et al.*, 2015), and the 'clone 17' line (Florimond Desprez, Cappelle en Pévèle, France) were used in this study. The Arabidopsis ecotype Columbia 0 (Col-0) and the *sh1* mutant line SALK\_055511, homozygous for T-DNA insertion, were used for complementation and overexpression experiments (Grienenberger *et al.*, 2009). Plants were grown in a greenhouse under a 16/8 h light/dark cycle.

Floral buds from other species were collected from the countryside around Lille (northern France). To broaden the range of sampling, species from the herbarium collection were kindly provided by the French National Museum of Natural History (Paris, France).

### Metabolite extraction

Lyophilized plant material was powdered and resuspended in a methanol/water/acetic acid mixture (75/23/2, v/v/v). The mixtures were then incubated in the dark under agitation for 1 h at 4 °C. Homogenates were clarified by centrifugation (14000 g, 4 °C, 10 min) and passed through a 0.45- $\mu$ m filter. Methanol washes of pollen were used to extract phenolics from the pollen coat. Pollen grains were collected in water, the samples were then centrifuged (3000 g, 5 min), and the pollen pellets were resuspended in 100% methanol and vigorously shaken for 30 s. Methanolic wash was collected by centrifugation (3000 g, 5 min) and filtrated (0.45- $\mu$ m filter).

### Metabolite analyses by HPLC/DAD and compound identification by UPLC/MS/MS

Metabolite analysis was carried out using a Prominence HPLC system (Shimadzu) consisting of a quaternary pump (LC-20AD) and a UV-visible diode-array detector (DAD; SPD-20A). Analyses were performed on a 100  $\times$  4.6 mm Kinetex 2.6- $\mu$ m PFP 100 Å column (Phenomenex, <http://www.phenomenex.com/>). The chromatographic separation was performed using water (solvent A) and acetonitrile (solvent B), both acidified with 0.1% ortho-phosphoric acid. The solvents were delivered at a flow rate of 1.1 ml min<sup>-1</sup> and the oven temperature was set at 45 °C. The HPLC conditions for analysis of chicory floral buds were as follow: start at 10% solvent B, 7 min gradient to 30% solvent B, followed by 3 min gradient to 70%, and 1.5 min of isocratic 70% solvent B, then a 1-min gradient to return to 10% solvent B, and 8.5 min of isocratic re-equilibration at 10% solvent B. For analysis of hairy roots, complemented Arabidopsis lines, and enzymatic reactions, the HPLC gradient was slightly modified so that the following conditions were applied: start at 5% solvent B, 6 min

gradient to 20% solvent B followed by 1 min isocratic at 20% solvent B, and 1 min gradient to 25%, then a 4-min gradient to 70%, and 1 min isocratic 70% solvent B were applied before returning to 5% solvent B and 7 min re-equilibration.

The UPLC-MS/MS method for phenolamide analyses in the pollen and floral buds of the various species was adapted from Elejalde-Palmett *et al.* (2015). The chromatography system used was an ACQUITY UPLC (Waters, [www.waters.com](http://www.waters.com)). Separation was performed using a Waters ACQUITY HSS T3 C18 column (150  $\times$  2.1 mm, 1.8- $\mu$ m) with a flow rate of 0.4 ml min<sup>-1</sup> at 55 °C. The injection volume was 5  $\mu$ l. The mobile phase consisted of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile). Chromatographic separation was achieved using an 18-min linear gradient from 10% to 60% solvent B. MS detection was performed by using a XEVO TQD mass spectrometer equipped with an electrospray ionization (ESI) source controlled by Masslynx 4.1 software (Waters) in both positive and negative full-scan mode ( $m/z$  50–2000). The capillary and sample cone voltages were 3000 V and 30 V, respectively. The cone and desolvation gas flow rates were 60 l h<sup>-1</sup> and 800 l h<sup>-1</sup>, respectively. Targeted data collection was carried in selected ion monitoring (SIM) mode for  $N^1, N^5, N^{10}$ -tricoumaroyl spermidine ([M+H]<sup>+</sup>;  $m/z$  584; RT=10.8 min) and  $N^1, N^5, N^{10}, N^{14}$ -tetracoumaroyl spermine ([M+H]<sup>+</sup>;  $m/z$  787; RT=11.8 min). Additional trihydroxycinnamoyl spermidine and tetrahydroxycinnamoyl spermine derivatives were detected in the total ionic current based on predicted mass-to-charge ratios of theoretical phenolamides. Candidates were further confirmed by analysis of mass fragmentation.

The UPLC-high-resolution MS (HRMS) and MS/MS methods for the profiling and the characterization of chicory floral buds, hairy roots, and enzymatic reaction mixtures were developed on an ACQUITY UPLC I-Class system (Waters) coupled to a Vion IMS quadrupole time-of-flight (QToF) mass spectrometer (Waters), equipped with an ESI source and an additional sprayer for the reference compound (Lockspray). Separation was performed using a KINETEX Biphenyl (100  $\times$  2.1 mm, 1.7- $\mu$ m) column (Phenomenex) heated at 40 °C with a mobile phase that consisted of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in methanol). For each sample, the injection volume was 1  $\mu$ l and the chromatographic separation was achieved using a 5-min linear gradient from 10% to 90% solvent B with a flow rate of 0.5 ml min<sup>-1</sup> fully directed toward the ESI source of the QToF instrument.

ESI-HRMS and MS/MS data were recorded in the positive ionization mode. The source and desolvation temperatures were 120 °C and 450 °C, respectively. Nitrogen was used as a drying and nebulizing gas at flow rates of 50 l h<sup>-1</sup> and 800 l h<sup>-1</sup>, respectively. Typically, the capillary voltage was 3 kV, the sampling cone voltage was 20V, and the source offset was 20V. Lock mass corrections using [M+H]<sup>+</sup> at  $m/z$  556.2771 of a leucine-enkephalin solution (100 pg  $\mu$ l<sup>-1</sup> in 50/50 acetonitrile/water, v/v, added with 0.1% formic acid) were applied for accurate mass measurements and determination of elemental composition. The mass range in full-scan mode was 50–2000 Da and spectra were recorded at 0.2 s per scan in the profile mode at

a resolution of 40 000 (full width half maximum). Three compounds, termed C1, C2, and C3, were identified in the positive ionization mode according to their respective  $[M+H]^+$  at  $m/z$  787.3710 (RT=4.72 min), 949.4229 (RT=4.42 min), and 584.2746 (RT=4.39 min). To obtain deeper structural information, MS/MS experiments were performed on  $[M+H]^+$  ions by optimizing the collision energies to be 28 eV, 34 eV, and 22 eV for C1, C2, and C3, respectively. Data acquisition and processing were performed with UNIFI 1.8 software.

### RNA extraction

Total RNA was isolated from *C. intybus* or Arabidopsis using a Nucleospin RNA Kit (Macherey–Nagel), including a DNase treatment. The yield and purity of the total RNA samples were determined using an Experion Automated Electrophoresis System (Bio–Rad). First-strand cDNAs were synthesized from 1  $\mu$ g of DNase-treated total RNA using an iScript cDNA synthesis kit (BioRad).

### Quantitative RT-PCR

For determination of expression patterns, *TIP41* (TIP41 like protein) and *Clath* (clathrin adaptor complex subunit) were used as reference genes to normalize the data. Evaluation of the reference genes and quantitative RT-PCR (qPCR) were conducted as previously described (Delporte et al., 2015). The gene-specific primers are listed in Supplementary Table S1 at JXB online.

### Identification and cloning of SHT-like sequences from chicory

Putative SHT-like sequences from chicory were identified in the EST database (<https://www.ncbi.nlm.nih.gov/nucest>). The full-length cDNAs of *CiSHT1* or *CiSHT2* with or without the stop codon were amplified from chicory floral bud cDNA by PCR using a proofreading polymerase (PrimeSTAR HS DNA polymerase, Takara). The PCR primers designed for these amplifications are listed in Supplementary Table S1. The entry clones were obtained through recombination of the PCR products with either pDONR207 or pDONR221 (Invitrogen), as follows: pDONR221\_*SHT1*, pDONR221\_*SHT2*, pDONR207\_*SHT1*, pDONR207\_*SHT1*-stop, pDONR207\_*SHT2*, pDONR207\_*SHT2*-stop. The consensus sequences were established after sequencing of independent clones.

### Expression, purification, and in vitro assays of recombinant proteins

The entry clones pDONR221-*SHT1* and pDONR221-*SHT2* were used to introduce the ORFs of *CiSHT1* and *CiSHT2* into the pDEST17 expression vector (Invitrogen) by LR recombination to produce pDEST17-*SHT1* or pDEST17-*SHT2*. Recombinant proteins with an N-terminal 6xHis-tag were expressed in *Escherichia coli* BL21-CodonPlus-(DE3)-RIL cells and purified as previously described (Legrand et al., 2016).

The standard *in vitro* assay was performed in a total volume of 50  $\mu$ l containing 1  $\mu$ g of purified protein, 50 mM phosphate

buffer (pH 9.0), 400  $\mu$ M acyl donor, and 2 mM acyl acceptor. After incubation at 28 °C for 1 h, the reaction was stopped by adding 15  $\mu$ l acetonitrile and 1  $\mu$ l 12 N HCl. As no production of free acid was observed, diminution of the CoA ester quantified after 1 h (in the linear phase of the enzymatic reaction) was used to compare acyltransferase activities. pH optima were determined by performing the assay in 50 mM phosphate buffer at pH 5.0 to 9.0, or in 50 mM Tris-HCl at pH 7.0 to 10.0.

Reverse reaction assays were conducted with phenolamides extracted from chicory pollen as substrates. To set up this reaction, 1 mM of CoASH and 1  $\mu$ g of protein were added in 50 mM phosphate buffer, pH 9.0, in a final volume of 50  $\mu$ l. The production of coumaroyl-CoA was evaluated after incubation at 28 °C for 3 h. The reaction mixtures were then centrifuged at 14 000 g for 5 min and filtered through a 0.45- $\mu$ m filter before being used for HPLC or HPLC-MS/MS analysis.

Cinnamoyl-CoA, coumaroyl-CoA, caffeoyl-CoA, and feruloyl-CoA were prepared enzymatically as previously described (Legrand et al., 2016).

### Chicory hairy root generation and growing conditions

The pB2GW7 binary expression vector was used. pDONR207\_*SHT1* and pDONR207\_*SHT2* were used to generate pB2GW7\_*SHT1* and pB2GW7\_*SHT2* by LR recombination. *Agrobacterium rhizogenes* strain 15834 (kindly provided by Marc Buée, INRA, Nancy, France) was transformed with the expression vectors by electroporation, as described in Nagel et al. (1990). Hairy root induction was performed on leaves of ‘clone 17’ using *A. rhizogenes* (wild-type or transformed with either pB2GW7\_*SHT1* or pB2GW7\_*SHT2*). Three independent lines were isolated. Hairy roots were maintained in liquid half-strength media (Murashige and Skoog, 1962) with added sucrose (30 g l<sup>-1</sup>) in flasks on an orbital shaker at an agitation speed of 110 rpm with a light source providing about 5 W m<sup>-2</sup> on a 16/8 h light/dark cycle at 25/20 °C. Supplementation experiments were carried out in triplicate on 20-ml subcultures seeded with ~300 mg of root fragments. Spermine was added to the medium to reach a final concentration of 1.5 mM 12 d later. Roots were sampled 6 d after the addition of spermine and immediately frozen in liquid nitrogen until metabolite extraction. Production of spermine derivatives was further quantified using free coumaric acid as a standard.

### Complementation of the Arabidopsis sht mutant with CiSHT1 or CiSHT2

For complementation experiments, the expression vector pB2GW7 was modified to replace the 35S promoter with the TA29 promoter for specific expression in the anther (Koltunow et al., 1990; Karimi et al., 2002). The TA29 promoter was amplified from *Nicotiana tabacum* gDNA using the primers listed in Supplementary Table S1. pB2GW7 was digested with SacI and SpeI restriction enzymes to remove the 35S promoter and the previously amplified TA29 promoter was inserted using an In-Fusion Cloning kit (Takara). The resulting expression vector, named pB2GW7(TA29), was sequenced to verify the correct insertion of the TA29 promoter.



pDONR207\_SHT1 and pDONR207\_SHT2 were then used to generate pB2GW7(TA29)\_SHT1 and pB2GW7(TA29)\_SHT2 by LR recombination.

*Agrobacterium tumefaciens* strain GV220 was transformed by electroporation with either pB2GW7(TA29)\_SHT1 or pB2GW7(TA29)\_SHT2. The floral dip method (Clough and Bent, 1998) was then used to transform the Arabidopsis *sh1* mutant line SALK\_055511. Phosphinotricin-resistant plants were selected and four independent homozygous lines per construction were further isolated by progeny segregation on selective medium.

### Subcellular localization experiments

For subcellular localization assays, the Gateway-compatible binary vectors pUBC-Dest and pUBN-Dest were used (Grefen *et al.*, 2010), which allow fusion of yellow fluorescent protein (YFP) in the C- and N-terminals of the protein of interest, respectively. The vectors pUBC\_SHT1, pUBN\_SHT1, pUBC\_SHT2, and pUBN\_SHT2 were obtained by LR recombination with the entry clones pDONR207\_SHT1-stop, pDONR207\_SHT1, pDONR207\_SHT2-stop, and pDONR207\_SHT2, respectively.

The *A. tumefaciens* strain GV3101 (pMP90) was transformed via heat-shock with the localization vectors pUBC\_SHT1, pUBC\_SHT2, pUBN\_SHT1, or pUBN\_SHT2 (Koncz and Schell, 1986; Höfgen and Willmitzer, 1988) and then 6–8-week-old *Nicotiana benthamiana* plants were transiently transformed by agroinfiltration of the youngest leaves (Waadt and Kudla, 2008). Protoplasts were isolated 2 d after infiltration and directly used for microscopy (<http://genetics.mgh.harvard.edu/sheenweb/>).

The *A. tumefaciens* strain GV220 was transformed by electroporation with either pUBN\_SHT1 or pUBC\_SHT1 and used to stably transform Arabidopsis Col-0 using the floral dip method (Clough and Bent, 1998). Transformants were then selected based on their resistance to phosphinotricin. Seeds collected from the transformants were grown *in vitro* and fluorescence was observed in roots from 7-d-old seedlings. For plasmolysis, roots were treated in a 500 mM NaCl solution for 5 min before observation (Hou *et al.*, 2005).

Protoplasts from *N. benthamiana* or Arabidopsis root cells were observed using a Zeiss LSM 780 Confocal Microscope and Zeiss ZEN software. The following excitation/emission wavelength settings were used: YFP (514/493–550 nm) and chlorophyll *a* (488 /640–710 nm). Images were processed using Fiji software (<https://fiji.sc/>) and Adobe Photoshop CS6.

### Sequence analysis

Sequence data from this article can be found in the GenBank data library under the accession numbers listed in Supplementary Table S2. Alignments and the phylogenetic tree were generated using CLC Sequence Viewer 7 software (Qiagen). Sequences were aligned using MUSCLE (with default parameters) and the phylogenetic tree was constructed using the neighbor-joining method (Jukes–Cantor model) and bootstrap analysis (1000 replications). Conserved motif determination was conducted using the MEME suite (Bailey *et al.*, 2009).

## Results

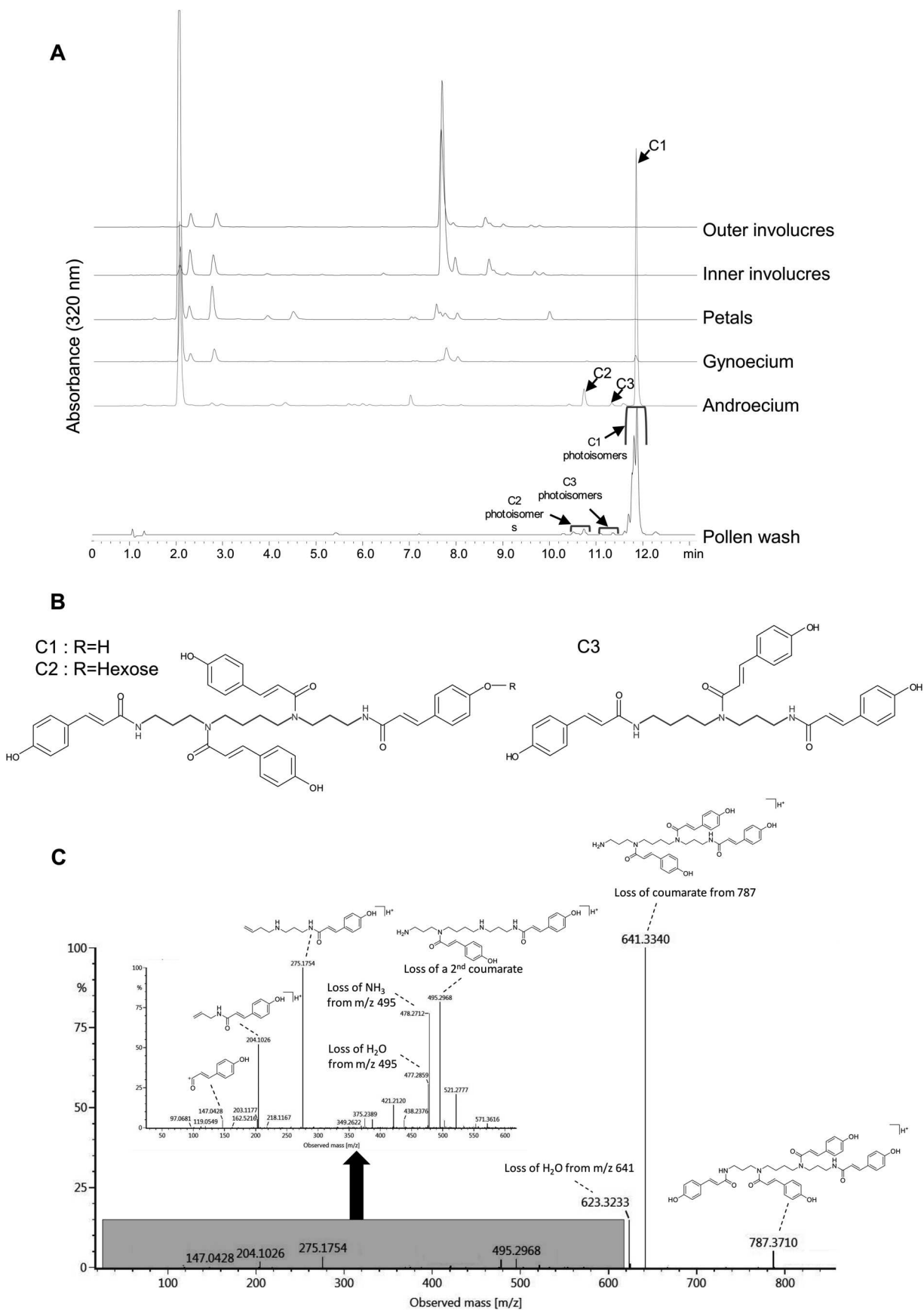
### Identification and characterization of pollen coat phenolamides in *C. intybus*

Chicory floral buds were dissected and methanolic extracts were analysed by HPLC-DAD in order to identify and localize accumulated phenolamides. The results indicated the presence of three compounds (termed C1, C2, and C3) restricted to the androecium that exhibited UV spectra typical of phenolamides (Fig. 1A, Supplementary Figs S1–S3). These unknown compounds were further investigated by mass spectrometry. In their positive-ion electrospray ionization mass spectra, C1 gave  $[M+H]^+$  a  $[M+H]^+$  ion at a mass-to-charge ratio ( $m/z$ ) of 787, C2 gave a protonated ion at  $m/z$  of 949, and C3 gave a protonated ion at  $m/z$  of 584 (Fig. 1B). The molecular formula of the major compound C1 was determined to be  $C_{46}H_{51}N_4O_8$ , and it was tentatively identified as  $N^1, N^5, N^{10}, N^{14}$ -tetracoumaroyl spermine. The fragmentation pattern of compound C1 (Fig. 1C) showed the presence of fragment ions at  $m/z=147$ , indicating the presence of coumaroyl residues. The ion at  $m/z=641$  could be assigned to the loss of a coumaroyl residue, whereas the one at  $m/z=495$  could have arisen from the loss of a second coumaroyl residue. The major ions at  $m/z$  values of 275 and 204 probably resulted from the cleavage of the  $N^5-C^6$  and  $N^{10}-C^{11}$  bonds in the spermine moiety. Similarly, C2 and C3 were identified as *N*-coumaroyl(glucose),  $N', N'', N'''$ -tricoumaroyl spermine and  $N^1, N^5, N^{10}$ -tricoumaroyl spermidine, respectively (Supplementary Fig. S2, S3). The position of the glycoside residue on C2 was not investigated further.

Analyses of the methanolic washes of the pollen grains were then conducted to confirm that these androecium-specific compounds were constituents of the pollen coat: the pollen coat is extractable by organic solvents without any alteration of the intracellular content of the pollen grain (Piffanelli *et al.*, 1998; Murphy, 2006). The resulting chromatogram (Fig. 1A) revealed the presence of the three phenolamides previously identified in the androecium, as well as their photoisomerization products. Indeed, previous studies on trihydroxycinnamoyl spermidine have shown the high photosensitivity of this type of molecule, with photoisomers readily appearing after the flower opens (Sobolev *et al.*, 2008).

### Identification of tetrahydroxycinnamoyl spermine as a conserved constituent of pollen coat in the Asteraceae

Our results from chicory showed that trihydroxycinnamoyl spermidine could be replaced by tetrahydroxycinnamoyl spermine as the major phenolamide of the pollen coat. Moreover, previously published studies of other species belonging to the Asteraceae have suggested that this may be specific to this family (Yamamoto *et al.*, 2002; Lam *et al.*, 2016). Therefore we undertook a targeted metabolic profiling of either floral buds or pollen coats among a wide range of Asteraceae and species belonging to related families. In order to achieve the greatest possible diversity, freshly collected samples and samples from herbarium specimens were analysed. We first established that time-dependent



**Fig. 1.** Identification of phenolamides present in chicory floral buds. (A) Stacked HPLC chromatograms of methanolic extracts obtained from different parts of the floral buds, and of a methanol wash of pollen. (B) Structures of the compounds identified as C1 ( $N^1,N^5,N^{10},N^{14}$ -tetracoumaroyl spermine), C2 ( $N$ -coumaroyl(glucose)- $N',N'',N'''$ -tricotmaroyl spermine), and C3 ( $N^1,N^5,N^{10}$ -tricotmaroyl spermidine) that were found in the pollen coat of chicory. The hexose position in C2 was not confirmed. (C) Fragmentation of C1 according to LC/MS/MS.

decomposition of the phenolamides was negligible by comparing the contents of phenolamides of freshly cut flowers immediately frozen in liquid nitrogen and lyophilized (see Methods) to those of flowers conserved in herbaria for decades (M. Delporte *et al.*, unpublished results). The results showed that tetrahydroxycinnamoyl spermine was found in almost all the Asteraceae species tested, with spermine amino groups substituted with either coumaric or caffeic acid (Fig. 2). The identity of the four spermine derivatives was confirmed by LC/MS (Supplementary Fig. S4 and Table S3). These compounds were mostly accumulated together with spermidine derivatives as well. In the nearest families to the Asteraceae that we investigated (i.e. Calyceraceae, Argophyllaceae, Stylidiaceae), no spermine derivatives were

identified, whereas spermidine derivatives were found with the exception of the Stylidiaceae.

#### Identification of candidate genes for tetrahydroxycinnamoyl spermine biosynthesis

Given the structural similarity between trihydroxycinnamoyl spermidine and tetrahydroxycinnamoyl spermine, candidate genes for the biosynthesis of spermine derivatives could be phylogenetically close to those involved in the biosynthesis of spermidine derivatives in the tapetum of the anther, i.e. SHTs. BLAST searches for SHT-like sequences in the Asteraceae family resulted in the identification of two sequences in each of five species: *Helianthus annuus*, *Lactuca sativa*, *C. intybus*,

Order	Family	Subfamily	Genus	Tetra-HCSpm	
Asterales	Stylidiaceae		Stylidium*	nd	
	Argophyllaceae		Argophyllum*	nd	
	Calyceraceae		Calycera*	nd	
	Asteraceae	Barnadesioideae	Barnadesia*	✓	
			Mutisioideae	Onoseris*	✓
				Mutisia*	nd
		Stiffieae	Stiffia*	✓	
		Wunderlichioideae	Wunderlichia*	nd	
			Nouelia*	✓	
		Gochnatieae	Gochnatia*	✓	
		Asteraceae	Carduoideae	Centaurea	✓
				Cirsium	✓
				Carduus	✓
				Cynara	✓
				Picris*	✓
			Cichorioideae	Leontodon	✓
	Lapsana			✓	
	Taraxacum			✓	
	Cichorium			✓	
	Lactuca			✓	
Asteroideae	Leontopodium		✓		
	Senecio		✓		
	Bellis		✓		
	Leucanthemum		✓		
	Matricaria*	✓			
	Achillea	✓			
	Chrysanthemum	✓			
Ambrosia	✓				
Sanvitalia	✓				
Helianthus	✓				

**Fig. 2.** Occurrence of tetrahydroxycinnamoyl spermine in floral buds among Asteraceae and close families. Methanolic extracts of flowers from various species (at least one for each represented genus) were screened for tetrahydroxycinnamoyl spermine using HPLC/DAD or UPLC/MS/MS. Tetra-HCSpm, tetrahydroxycinnamoyl spermine; nd, not detected; \*, extracts from herbarium samples.

*Carthamus tinctorius*, and *Cynara cardunculus*. Phylogenetic analyses of these 10 amino acid sequences among other characterized BAHDs belonging to the clade Vb resulted in the identification of two clusters for the Asteraceae SHT-like sequences that we had identified (Fig. 3). For each species, we named one sequence as SHT1 and the other as SHT2. The two sequences from *C. intybus*, CiSHT1 and CiSHT2, were selected for further characterization. The associated genes, *CiSHT1* and *CiSHT2*, were 1356-bp and 1359-bp long, respectively, and shared 72.9% nucleotide sequence identity. No intron was identified for either of these genes. The deduced protein sequences of 451 and 452 amino acids for CiSHT1 and CiSHT2, respectively, had calculated molecular weights of 50.2 and 50.4 kDa, and harboured the two peptide motifs HXXXD and DFGWG conserved among the BAHD family (St-Pierre and De Luca, 2000). In addition, the two proteins shared 63% and 66% similarity, respectively, with AtSHT.

### CiSHT1 and CiSHT2 expression patterns are correlated with phenolamide accumulation in chicory flowers

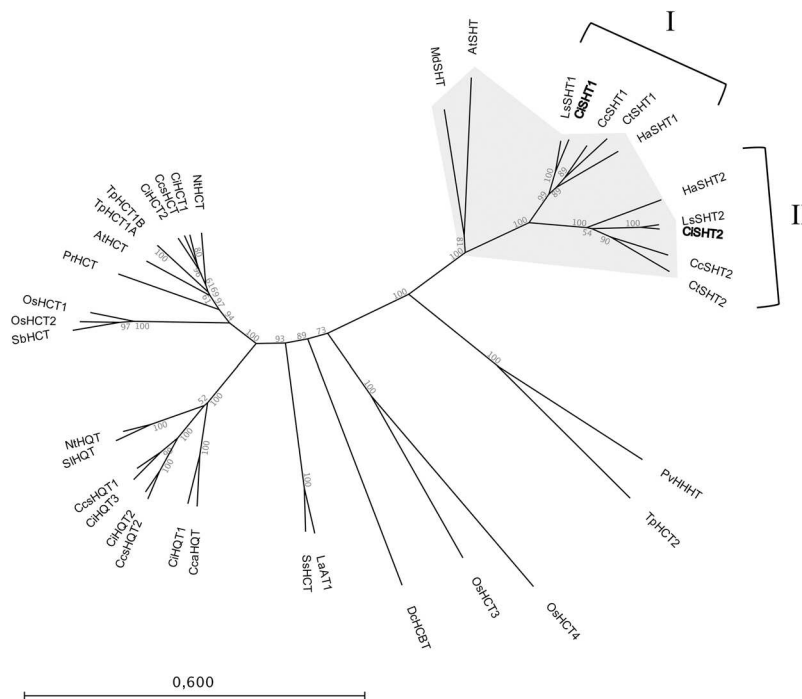
The contents of the phenolamides that we designated as C1, C2, and C3 were measured in roots, leaves, and flower buds. Flower buds were sampled at different developmental stages, beginning from the formation of the tetrads of microspores (corresponding to stages 9–10) to anthesis (stage 16; Habarugira et al., 2015).

Tetracoumaroyl spermine was the major phenolamide identified at every stage of the floral bud development and its accumulation was restricted to the floral parts of the plant (Fig. 4A).

Its content increased dramatically until the flowers opened and then decreased. This drop was probably due to petal formation and pollen release, as recently observed in *M. chamomilla* (Park et al., 2017). In an attempt to assign a function to *CiSHT1* and *CiSHT2*, their expression patterns were examined in the same tissues by qPCR (Fig. 4B). For both genes, the results showed either nearly undetectable or weak expression in roots, leaves, and in the youngest floral buds (stage 9–10) when the tapetum was not fully differentiated. Transcript accumulation was then increased by up to 10 000-fold. Transcripts started to decrease between stages 14 and 15, which corresponded to the degeneration of the tapetum. The expression profiles were thus in accordance with the accumulation of tetracoumaroyl spermine during floral bud development, supporting the hypothesis of the involvement of CiSHT1 and/or CiSHT2 in phenolamide biosynthesis in chicory. Moreover, *CiSHT1* and *CiSHT2* transcript accumulation throughout flower development was similar to that previously described for *AtSHT* and *MdSHT* (Grienenberger et al., 2009; Elejalde-Palmett et al., 2015).

### CiSHT2 shows a net in vitro activity towards the production of tetracoumaroyl spermine

In order to examine the catalytic activities of CiSHT1 and CiSHT2, recombinant proteins with an N-terminal 6xHis-tag were produced in *E. coli* and purified by affinity chromatography. Their purity and identity were verified by SDS-PAGE and western blot analysis using an anti-His antibody. Both purified proteins produced a major band at the expected molecular weight, i.e. ~50 kDa (Fig. 5). The activities of the recombinant proteins were first assayed against spermidine and spermine as the acyl acceptor and

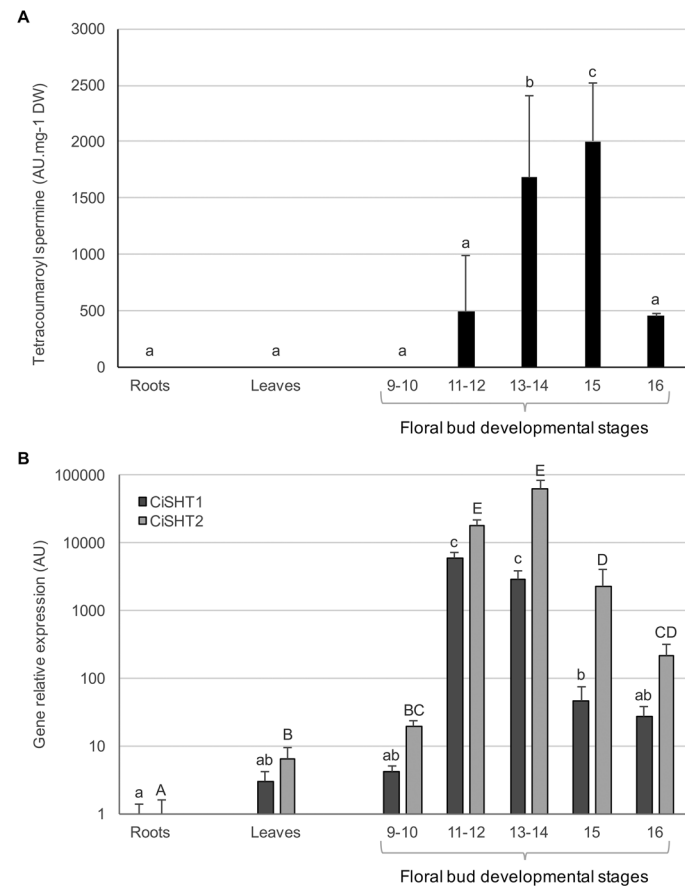


**Fig. 3.** Phylogenetic analysis of the ten SHT-like amino acid sequences identified in five species of the Asteraceae family and other biochemically characterized members of the BAHD family clade Vb. Two clusters of SHT-like sequences are indicated for the Asteraceae sequences. See [Supplementary Table S2](#) for details of protein names and accession numbers.

against coumaroyl-CoA as the acyl donor. Production of tricoumaroyl spermidine and tetracoumaroyl spermine was achieved with both CiSHT1 and CiSHT2. During activity

monitoring, the partially substituted polyamines appeared first in the reaction media but then diminished over time in favor of the fully substituted products, indicating the sequential nature of the reaction (Fig. 6). Spermine appeared to be the preferential polyamine substrate of CiSHT2: its activity measured with spermidine was 56% lower than that measured with spermine. The relative activity of CiSHT1 was much weaker than that of CiSHT2, making it difficult to determine the substrate preference. Other acyl acceptors were examined following the same standard *in vitro* assay procedure. Although the activities were very weak for both the enzymes, positive results were obtained after overnight incubation with diamine putrescine and, more surprisingly, with quinic and shikimic acids (all supplied at 10 mM), and with glycerol, although only at an extremely high concentration (100 mM). With regards to acyl donors, as already shown for AtSHT (Grienenberger *et al.*, 2009), all the common CoA-esters tested (cinnamoyl-CoA, caffeoyl-CoA, and feruloyl-CoA) were used by both enzymes to produce substituted spermidine or spermine.

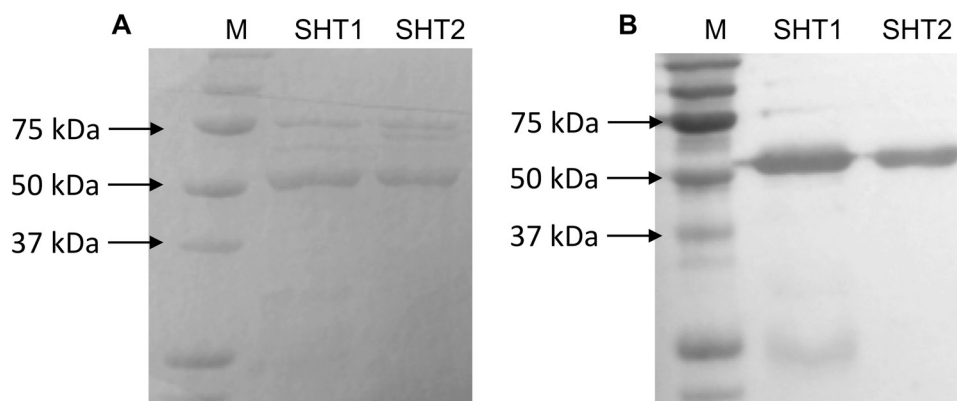
The pH optimum for the recombinant enzyme activity was determined over a range of 5.0–10.0 using spermine or spermidine and coumaroyl-CoA as substrates, and was found to be 9.0, which matched values for previously identified BAHD N-acyltransferases (Luo *et al.*, 2009; Schmidt *et al.*, 2015). The ability to catalyse the reverse reaction was also examined; several BAHD enzymes, including those from chicory that belong to the same clade Vb, have been demonstrated to be freely reversible (Legrand *et al.*, 2016). Both CiSHT1 and CiSHT2 were tested in the presence of CoASH and phenolamides extracted from chicory pollen but no production of coumaroyl-CoA was detected under our experimental conditions.



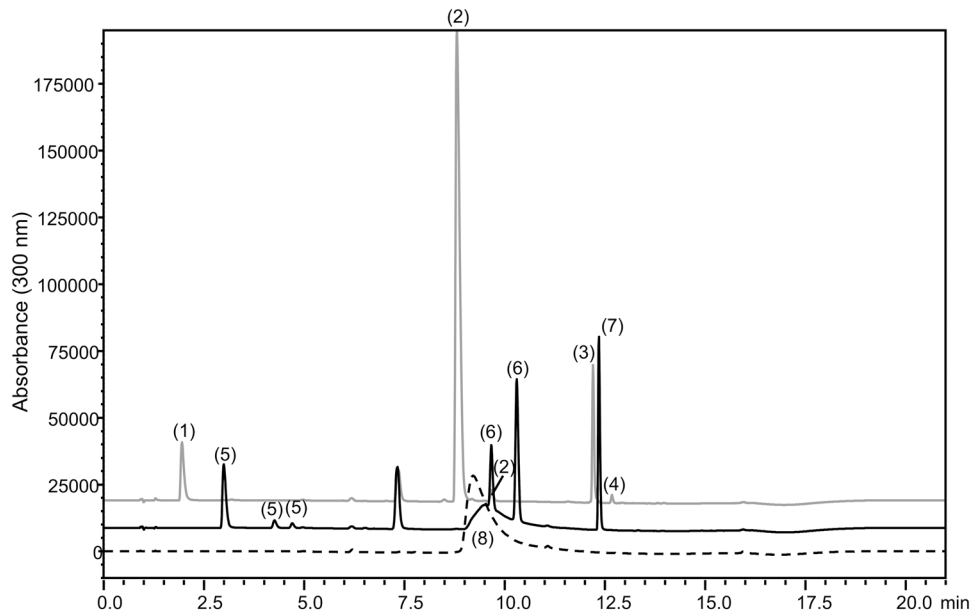
**Fig. 4.** (A) Tetracoumaroyl spermine contents and (B) expression profiles of CiSHT1 and CiSHT2 in roots, leaves, and in floral buds at different developmental stages. Tetracoumaroyl spermine was not detected in roots and leaves. Gene expression was determined by qPCR: expression levels are normalized to the expression of the reference genes *TIP41* and *Clath* and are presented relative to that of roots. Note the logarithmic scale for the expression data. Data are means ( $\pm$ SD) from three independent biological replicates. Different letters indicate significant differences as determined using ANOVA ( $P < 0.05$ ). The numbers for floral bud developmental stages are as in Habarugira *et al.* (2015).

#### Overexpression of CiSHT2 in chicory hairy roots leads to production of spermine derivatives

*In vitro* characterization of the two enzymes highlighted a substrate flexibility often observed in specialized metabolism. In order to verify the *in vivo* biochemical functions of CiSHT1 and CiSHT2, the corresponding genes were



**Fig. 5.** Expression and purification of recombinant CiSHT1 and CiSHT2. (A) SDS-PAGE separation of recombinant proteins. The gel was stained with Coomassie blue. (B) Immunoblot analysis confirming the expression of the recombinant proteins. The proteins were detected with an anti-His antibody. M, protein markers.



**Fig. 6.** HPLC chromatograms of *in vitro* assays of recombinant CiSHT2 in the presence of coumaroyl-CoA and spermine (grey line) or spermidine (black line), or without an acyl acceptor (dotted line). The identities of the numbered peaks were confirmed by mass spectrometry: (1) monocoumaroyl spermine; (2) dicoumaroyl spermine; (3) tricoumaroyl spermine; (4) tetracoumaroyl spermine; (5): monocoumaroyl spermidine; (6) dicoumaroyl spermidine; (7) tricoumaroyl spermidine; and (8) coumaroylCoA.

stably overexpressed in chicory hairy roots and production of phenolamides was then evaluated when the roots were supplemented with spermine. In hairy root lines overexpressing *CiSHT1* no differences with the controls were observed in terms of production of phenolic compounds. However, production of acylated spermines was confirmed by UPLC-MS/MS analysis in lines overexpressing *CiSHT2*. These spermine derivatives were quantified in three independent lines (Table 1). Dicoumaroyl spermine was the most abundant and accounted for 78% of total phenolamides. These results confirmed the ability of CiSHT2 to promote spermine-based phenolamide synthesis *in vivo*, but the fully substituted compound was not detected. The negative results observed with the lines transformed with *CiSHT1* were in accordance with the very weak enzymatic activity observed *in vitro*.

#### *CiSHT2* partly complements the phenolamide deficiency in the pollen coat of the *Arabidopsis* *sht* mutant and also leads to the production of spermine derivatives

The *Arabidopsis sht* T-DNA insertion mutant previously described as lacking trihydroxycinnamoyl spermidine conjugates in the pollen coat (Grienenberger et al., 2009; Elejalde-Palmett et al., 2015) was used to overexpress either *CiSHT1* or *CiSHT2* under control of the TA29 promoter, which allows overexpression in tapetum cells (Koltunow et al., 1990). As observed in the previous studies (Grienenberger et al., 2009; Elejalde-Palmett et al., 2015), dihydroxyferuloyl sinapoyl spermidine (and its associated biosynthetic intermediates) was the main phenolamide in wild-type flowers (Table 2). The overexpression of *CiSHT1* did not lead to a different chemotype compared to the mutant, but a panel of different

phenolamides was observed in the lines overexpressing *CiSHT2* (*sht/CiSHT2*). The production of dihydroxyferuloyl sinapoyl spermidine was weakly restored to 3.8% of the wild-type level. Moreover, the most abundant intermediate was monoferuloyl spermidine. This demonstrated the inability of CiSHT2 to fully acylate spermidine and highlighted the difference between this enzyme and the two previously characterized SHTs from *Arabidopsis* and *M. domestica* (Grienenberger et al., 2009; Elejalde-Palmett et al., 2015). In addition to spermidine derivatives, spermine derivatives were also found in the *sht/CiSHT2* lines. Thus, diferuloyl and triferuloyl spermines were detected but if tetrahydroxycinnamoyl spermine was present it was below the level of detection (Table 2). However, production of spermine conjugates in *Arabidopsis*, a species that does not produce these particular phenolamides, again demonstrated the ability of CiSHT2 to acylate spermine *in vivo*.

#### *CiSHT1* and *CiSHT2* are targeted to the nucleocytoplasmic compartment

The expression of *CiSHT1* and *CiSHT2* early in the development of the stamen suggested that expression was restricted to tapetum cells, as has been demonstrated in *Arabidopsis* (Grienenberger et al., 2009). The accumulation of phenolamides in the pollen coat thus implies the export of either the final molecules or the enzymes implicated in their synthesis. It has been assumed that, like most of the BAHD acyltransferases that have been characterized, SHT-like enzymes should be cytoplasmic, but to date this has not actually been assessed. Consequently, we investigated the subcellular localization of CiSHT1 and CiSHT2 using two constructions. pUB vectors were used to fuse YFP at the C- or N-terminals of CiSHT1 and CiSHT2, and then *YFP::SHT1*, *SHT1::YFP*, *YFP::SHT2*, and *SHT2::YFP* were transiently expressed in *N. benthamiana* leaves. Fluorescence was

**Table 1.** Contents of spermine derivatives in chicory hairy roots overexpressing CiSHT2 grown in media supplemented with spermine 1.5 mM for 6 d

Hairy root line	Phenolamide content (nmol mg <sup>-1</sup> DW)			
	Monocoumaroyl spermine	Dicoumaroyl spermine	Tricoumaroyl spermine	Tetracoumaroyl spermine
Line 1	1.15 ± 0.08	6.12 ± 0.80	0.31 ± 0.02	nd
Line 2	1.08 ± 0.19	6.40 ± 0.26	0.33 ± 0.03	nd
Line 3	1.54 ± 0.14	6.51 ± 0.44	0.27 ± 0.01	nd
Mean	1.26 ± 0.25	6.35 ± 0.21	0.30 ± 0.03	nd

No spermine conjugates were detected in control lines. Values are means of three independent biological replicates (±SD). nd, not detected.

**Table 2.** Identification and relative quantification of phenolamides in *Arabidopsis* floral buds by LC/MS in the wild-type (WT) and *sht* and *sht/CiSHT2* lines

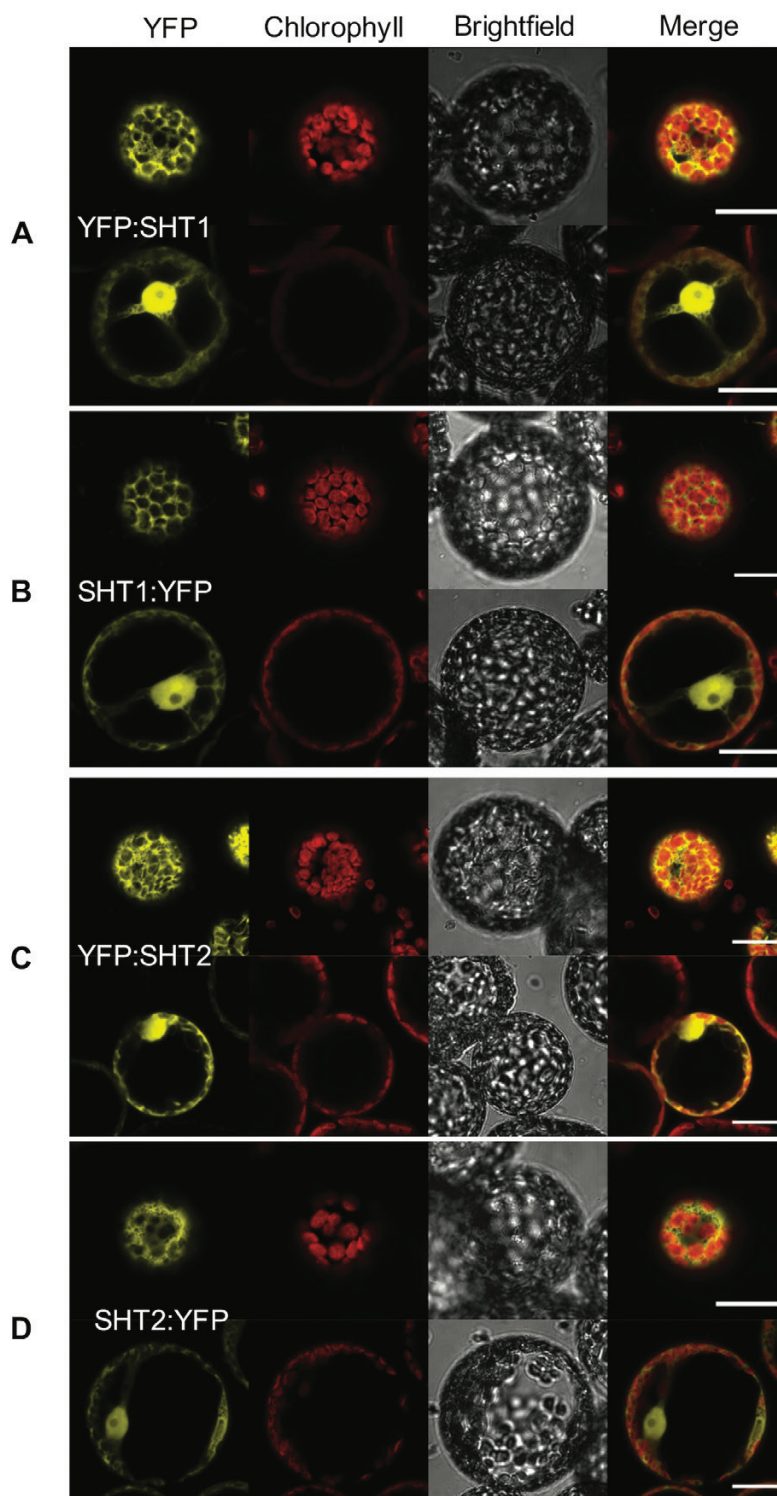
<i>R<sub>t</sub></i> (min)	<i>m/z</i> ES(-)	<i>m/z</i> ES(+)	Compound	Relative quantity (%)		
				WT	<i>sht</i>	<i>sht/CiSHT2</i>
2.31	320 [M-H] <sup>-</sup>	322 [M+H] <sup>+</sup>	Feruloyl spermidine	4.2	nd	92.5
		177 [M+H-Spd] <sup>+</sup>				
		146 [M+H-Fer] <sup>+</sup>				
5.89	553 [M-H] <sup>-</sup>	555 [M+H] <sup>+</sup>	Diferuloyl spermine	nd	nd	9
		379 [M+H-Fer] <sup>+</sup>				
		305 [M+H-Fer-C3 chain] <sup>+</sup>				
		234 [M+H-Fer-C3NC4 chain] <sup>+</sup>				
		177 [M+H-Fer-Spm] <sup>+</sup>				
8.84	496 [M-H] <sup>-</sup>	498 [M+H] <sup>+</sup>	Diferuloyl spermidine	14.2	nd	2.4
		322 [M+H-Fer] <sup>+</sup>				
		177 [M+H-Fer-Spd] <sup>+</sup>				
10.73	729 [M-H] <sup>-</sup>	731 [M+H] <sup>+</sup>	Triferuloyl spermine	nd	nd	11
		555 [M+H-Fer] <sup>+</sup>				
		234 [M+H-Fer-C3NC4 chain] <sup>+</sup>				
		177 [M+H-Fer-Fer-Spm] <sup>+</sup>				
10.86	720 [M-H] <sup>-</sup>	722 [M+H] <sup>+</sup>	Trihydroxyferuloyl spermidine	13.9	nd	nd
		530 [M+H-OHfer] <sup>+</sup>				
		338 [M+H-OHfer-OHfer] <sup>+</sup>				
		193 [M+H-OHfer-OHfer-Spd] <sup>+</sup>				
		734 [M-H] <sup>-</sup>				
11.78	734 [M-H] <sup>-</sup>	736 [M+H] <sup>+</sup>	Dihydroxyferuloyl sinapoyl spermidine	100	nd	3.8
		544 [M+H-OHfer] <sup>+</sup>				
		338 [M+H-OHfer-Sin] <sup>+</sup>				
		278 [M+H-OHfer-OHfer-C3 chain] <sup>+</sup>				
		264 [M+H-OHfer-OHfer-C4 chain] <sup>+</sup>				
		250 [M+H-OHfer-Sin-C4 chain] <sup>+</sup>				
		207 [M+H-OHfer-OHfer-Spd] <sup>+</sup>				
		193 [M+H-OHfer-Sin-Spd] <sup>+</sup>				

*R<sub>t</sub>*, retention time. For relative quantification, phenolamide signals were normalized to the dihydroxyferuloyl sinapoyl spermidine signal. Values are presented for one line per construct and are representative of four independent lines. nd, not detected.

observed in protoplasts obtained from the transformed leaves. For the four constructs, the predominant fluorescence was cytoplasmic, as evidenced by a strong signal detected at the periphery of the cells and in the cytoplasmic strands (Fig. 7). A fluorescent signal was also observed in the nucleus. Moreover, there was no signal in the chloroplasts, as shown by the merged images of YFP fluorescence and chlorophyll auto-fluorescence. This clearly demonstrated that the two proteins shared a nucleo-cytoplasmic localization.

This localization was further demonstrated in additional experiments conducted with *YFP::SHT1* and *SHT1::YFP*

stably expressed in *Arabidopsis*. Confocal microscopic examination was conducted on root cells, because this tissue was free of interference by the chlorophyll auto-fluorescence. Observations made on independent positive lines confirmed the nucleo-cytoplasmic localization of CiSHT1 (Supplementary Fig. S5): the fluorescent signal was clearly visible at the periphery of the cells and in nucleus-like structures. Signal localization after retraction of the vacuole by plasmolysis was associated with a strong signal inside the cell, corroborating the cytosolic localization (Supplementary Fig. S5).



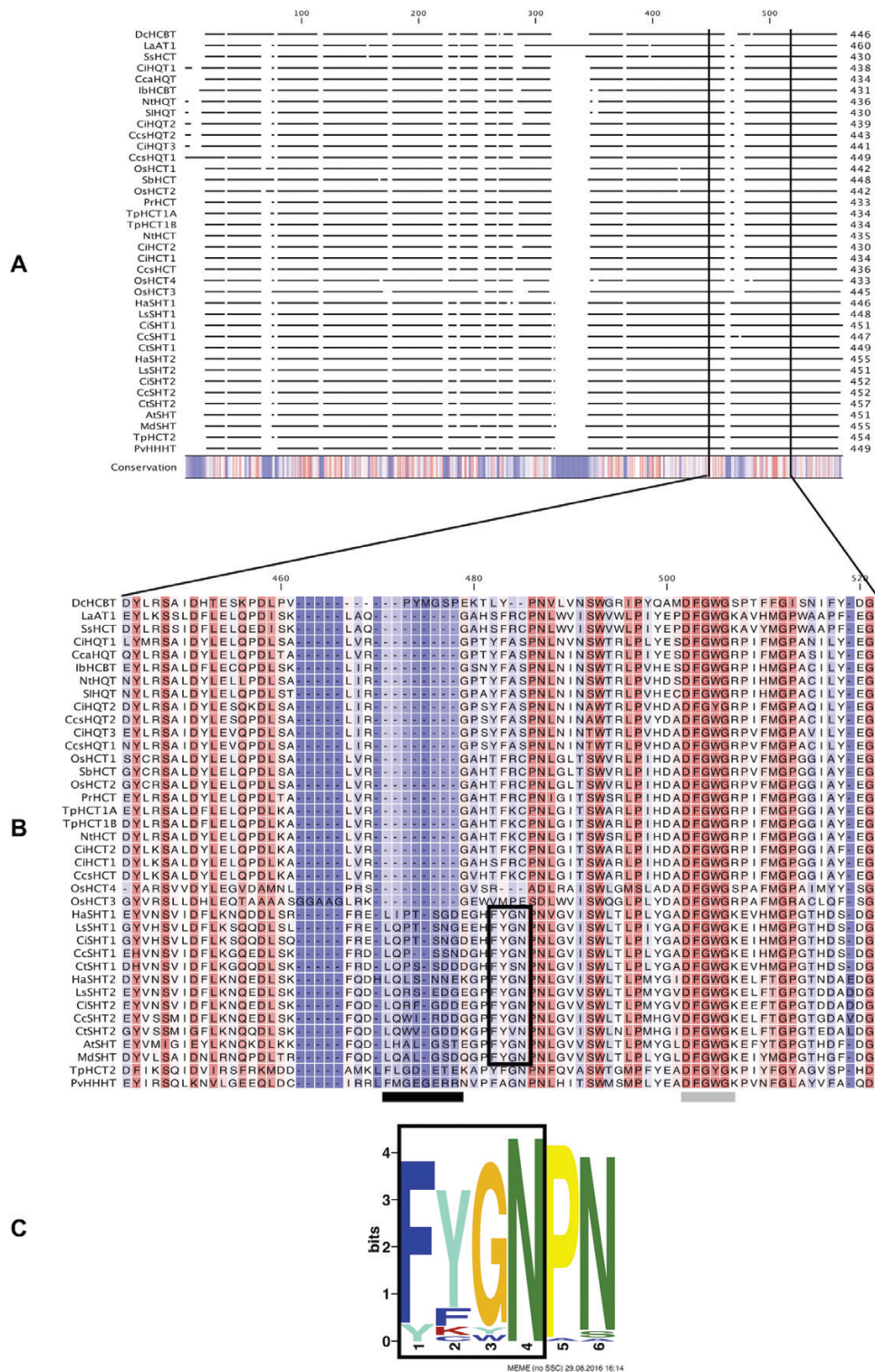
**Fig. 7.** Subcellular imaging of yellow fluorescent protein (YFP)-tagged constructs transiently expressed in *N. benthamiana* leaves. The images of the leaf cell protoplasts using confocal microscopy show the fluorescence pattern for (A) *YFP::SHT1*, (B) *SHT1::YFP*, (C) *YFP::SHT2*, and (D) *SHT2::YFP*. Scale bars are 20  $\mu$ m.

Sequence analyses indicate a specific motif shared by members of the *SHT* subfamily

Our results demonstrated that the appearance of spermine-based phenolamides in the pollen coat of species belonging to the Asteraceae family coincided with the presence of two

*SHT*-like enzymes (Fig. 3). The high sequence identity shared by these two enzymes whatever the species (68% on average) suggested that they could have arisen from the duplication of an ancestral *SHT* gene, after which the neofunctionalization of the one termed *SHT2* would have led to the emergence of





**Fig. 8.** Sequence analysis of the SHT-like enzymes. (A) Full alignment of the ten SHT-like sequences identified in Asteraceae species with other characterized members of the BAHD family clade Vb. (B) Detail of a particular portion of the alignment showing the BAHD motif DFGWG (grey bar), an amino acid insertion (black bar), and the FYGN motif conserved in the SHT-like sequences (black box). (C) Result of a motif discovery conducted with the MEME Suite on all the publicly available predicted SHT-like sequences.

tetrahydroxycinnamoyl spermine in the pollen coat. Sequences of the characterized enzymes from *Arabidopsis*, *M. domestica*, and *C. intybus* together with orthologues identified in *L. sativa*, *Ca. tinctorius*, *Cy. cardunculus*, and *H. annuus* were aligned in an attempt to link the neofunctionalization observed in this

SHT-like family with some particular structural features (Grienenberger *et al.*, 2009; Elejalde-Palmett *et al.*, 2015; this current study). In terms of sequences, no clear differences were observed that could explain the functional divergence. However, alignment of the SHT-like sequences with other

characterized members of the clade Vb highlighted the insertion of a few amino acid residues slightly upstream of the DFGWG motif conserved in the BAHD family (Fig. 8A, B). This fragment did not seem to be very conserved and was also present in the Fabaceae sequences TpHCT2 (hydroxycinnamoyl-CoA:malic acid transferase from *Trifolium pratense*) and PvHHHT (hydroxycinnamoyl-CoA:tetrahydroxyhexandioic acid transferase from *Phaseolus vulgaris*). However, a 4-amino acid motif, FYGN, appeared to be characteristic of the SHT-like sequences. A conserved motif search was then conducted using MEME with all the predicted SHT-like sequences publicly available, and resulted in the identification of the same FYGN motif (Fig. 8C).

## Discussion

The accumulation of trihydroxycinnamoyl spermidine in the pollen coat is an evolutionary conserved trait among eudicotyledons. This suggests it has an essential function that is independent from specific environmental constraints; however, our data highlight a metabolic diversification of these particular compounds in the Asteraceae family. Thus, tetrahydroxycinnamoyl spermine was accumulated in addition to trihydroxycinnamoyl spermidine in the pollen coat of most members of the Asteraceae family that we analysed (Fig. 2). This diversification could potentially give an evolutionary advantage to these species. With more than 23500 species, the Asteraceae family represents nearly 10% of the angiosperms and is one of the largest families of flowering plants, colonizing most parts of the earth except for Antarctica (Panero and Crozier, 2016). The worldwide success of this family is often linked to its specialized metabolism in addition to its flower architecture, complex reproductive biology, and habit plasticity (Calabria et al., 2007). The metabolic diversification of the pollen coat could constitute a supplemental feature contributing to its evolutionary success. In *Arabidopsis* and *M. domestica*, trihydroxycinnamoyl spermidine is synthesized by the enzyme SHT, which belongs to the clade Vb of the BAHD acyltransferases (as defined by Tuominen et al., 2011). Notably, hydroxycinnamoyl-CoA:shikimate/quinic acid hydroxycinnamoyl transferases (HCTs), involved in the synthesis of lignin precursors and therefore distributed ubiquitously in higher plants, also belong to this clade. The other members of the clade probably originated from duplication and neofunctionalization of ancestral HCTs, as it has been suggested for rosmarinic acid synthase (RAS) and also for SHTs (Petersen et al., 2009; Elejalde-Palmett et al., 2015). Mining of the genomic data for five members of the Asteraceae family revealed the presence of two AtSHT homologues in each species (Fig. 3). The presence of two SHT-like would undoubtedly have favoured the emergence of new pollen coat phenolamides, releasing the selective pressure apparently applied on these compounds. The proteins, which we term SHT1 and SHT2, grouped into two distinct classes in the phylogenetic tree (Fig. 3). Both CiSHT1 and CiSHT2 were shown to promote biosynthesis of phenolamides *in vitro*. The ability of CiSHT1 and CiSHT2 to accept acyl acceptors other than polyamines was tested, and the weak but positive results obtained with quinate and shikimate provide

an argument in favour of their assumed evolutionary origin, i.e. the neofunctionalization of an ancestral duplicated HCT.

Expression profiles of *CiSHT1* and *CiSHT2* were also examined and demonstrated a clear correlation between gene expression and phenolamide accumulation in the androecium (Fig. 4). Moreover, whilst the overexpression of *CiSHT1* in *planta* gave results that were in accordance with the weak activity that had already been observed *in vitro*, the overexpression of *CiSHT2* in chicory hairy roots and in *Arabidopsis* *sht* mutant demonstrated the ability of this enzyme to produce hydroxycinnamoyl spermines *in vivo* (Tables 1, 2); however, the fully substituted spermine was not the major phenolamide, in contrast to the observations made in chicory floral buds and pollen. This may suggest that there are specific factors within the biosynthetic pathway in the tapetum of the Asteraceae that act in favor of the accumulation of tetrahydroxycinnamoyl spermine. Polyamine availability could be an element of such a specific metabolic environment. Indeed, it is known that spermidine is by far the most abundant polyamine in *Arabidopsis* flowers (Tassoni et al., 2000, 2008), and this could explain why spermine conjugates were not efficiently produced in flowers of the *sht/CiSHT* transgenic lines. This hypothesis was supported by assays conducted on hairy roots, which suggested that spermine concentration was of paramount importance. Indeed, in chicory hairy roots overexpressing *CiSHT2*, the production of polyacylated spermines was significantly enhanced by exogenous supplementation with spermine (data not shown). From this perspective, it would be interesting to quantify the different polyamines in flowers of chicory.

Spatial organization of the metabolic pathway could be another factor in explaining the accumulation only of fully substituted spermine in the pollen coat of chicory. Hence we examined the subcellular localization of the two enzymes by transient expression in *N. benthamiana* leaves, which clearly demonstrated a nucleo-cytoplasmic localization of the SHTs (Fig. 7, Supplementary Fig. S5). The members of the BAHD family are predicted to be localized in the cytosol due to the absence of a peptide signal (D'Auria, 2006). Similar subcellular localization was previously confirmed in *Arabidopsis* for HCT and in *Cy. cardunculus* for a hydroxycinnamoyl-coenzyme A:quinic acid hydroxycinnamoyl transferase (HQT) responsible for chlorogenic acid biosynthesis (Bassard et al., 2012; Moglia et al., 2014). This cytosolic localization excludes the possibility of the involvement of a macro-compartmentation, which could have explained the metabolic fluxes leading to the accumulation of only fully substituted spermine in the pollen coat. Several proteins involved in the phenylpropanoid pathway are also assembled in so-called metabolons (Ralston and Yu, 2006; Laursen et al., 2015), a type of supramolecular complex that is the main form of micro-compartment and that is supposed to improve channeling of compounds between consecutive enzymes. Assembly of metabolons relies on several parameters (e.g. availability and concentrations of the constituents) and it could be assumed that these conditions were not met in the hairy roots and in the *Arabidopsis* floral buds. This could explain the differences between phenolamides accumulated in these overexpression systems and the chicory floral buds,

but this hypothesis would need to be assessed. The subcellular localization results that we obtained in this study also suggested that, once synthesized, phenolamides have to be exported outside of the tapetal cells to be deposited on the surface of the pollen grains. The mechanisms involved in this deposition are not known. Most of the pollen coat components are released after programmed cell death of the tapetum; however, a study by Quilichini *et al.* (2014) has suggested the involvement of an ABC transporter in the export of the pollen coat phenolamides. Transport of either spermidine or spermine derivatives could involve different systems, and the lack of efficient transport of the particular spermine-based phenolamides produced in our *in vivo* assays may explain the weak concentration that we observed. Indeed, feedback inhibition may occur, thus explaining why tetracoumaroyl spermine was not produced more efficiently *in vitro* (Fig. 6).

To gain insights into the structural characteristics of these enzymes and try to elucidate the evolution within the clade Vb, sequence analysis was performed. This revealed the presence of a specific motif, FYGN, in the amino acid sequence of the enzymes belonging to the SHT family (Fig. 8). This motif, associated with the insertion of few amino acids upstream, was located slightly before the DFGWG motif of the BAHD family and may be involved in the ability of the SHTs to synthesize fully substituted polyamines. However, the lack of structural data makes it difficult to confirm of this hypothesis. Enzymes that clustered close to SHTs, namely TpHCT2 from *T. pratense* and PvHHHT from *P. vulgaris*, have also been found to exhibit the insertion observed in the SHT family but not the SHT conserved motif (Sullivan, 2009, 2017). These enzymes could have arisen from the same ancestor as SHTs but then evolved differently, leading to the synthesis of new compounds such as phasic acid in *T. pratense* while phenolamides are absent from the pollen coat in this species (Elejalde-Palmett *et al.*, 2015). To gain further insights into the functional role of this motif, site-directed mutagenesis experiments would have to be carried out.

To the best of our knowledge, this study reports for the first time the identification and characterization of an enzyme catalysing the *N*-acylation of spermine. Acylation of spermine with hydroxycinnamoyl-CoA has previously been described in crude protein extracts of *Aphelandra tetragona*; however, the enzymes were only able to produce monocoumaroyl spermine, a precursor of the alkaloid aphelandrine, and no molecular characterization was achieved (Hedberg *et al.*, 1996). Complementary experiments are still needed to identify all the determinants involved in the production of the fully substituted polyamines that are observed in floral buds. Indeed, the results that we obtained *in vitro* and *in vivo* suggested that the affinity of CiSHT2 for the different intermediates could not by itself explain this specific metabolic pattern. Such information should help towards being able to fully exploit the potential of these enzymes for the biological synthesis of valuable molecules that have many diverse applications.

The role of trihydroxycinnamoyl spermidine in the pollen coat has been questioned in several studies and some assumptions have been made but not yet confirmed. The abnormal

formation of pollen grains of the *sht* mutant described in Arabidopsis points to an apparent involvement in sporopollenin structure. However, the lack of phenolamides has not been shown to lead to an alteration of pollen fertility, at least under the conditions tested (Grienenberger *et al.*, 2009). These compounds are also thought to play a role in the protection of pollen against biotic and abiotic stresses, and in interactions with pollinators and with the stigma (Fellenberg and Vogt, 2015). The presence of spermine derivatives instead of spermidine derivatives in the pollen coat of the Asteraceae raises questions. We hypothesize that this metabolic diversification could confer an evolutionary advantage to the Asteraceae in challenging environments. More studies are required to gain further insights into the role of these molecules. The deletion of these genes in chicory through genome editing, which is part of our ongoing research, should help to answer this question.

## Supplementary data

Supplementary data are available at *JXB* online.

Fig. S1. Proposed fragmentation scheme of the compound designated as C1 and its UV spectra.

Fig. S2. LC/MS/MS fragmentation of the compound designated as C2 and its UV spectra.

Fig. S3. LC/MS/MS fragmentation of the compound designated as C3 and its UV spectra.

Fig. S4. LC/MS/MS fragmentation of spermine derivatives found in Asteraceae flower buds.

Fig. S5. Subcellular imaging of YFP-tagged constructs stably expressed in Arabidopsis.

Table S1. Sequences of primers used in this study.

Table S2. Description of proteins used in the phylogenetic analysis.

Table S3. Identification of spermine derivatives in Asteraceae flower buds.

## Acknowledgments

The authors thank Florian Jabbour (Herbarium of the Muséum National d'Histoire Naturelle, Paris, France) for sampling and for providing herbarium samples. MD and GL were supported by doctoral fellowships from the doctoral school 104 SMRE, while GB was supported by a doctoral fellowship from University of Lille1 and the Région Hauts de France. BH was funded by DFG grant GRK 1525. Experiments performed at Charles Viollette Institute were supported by an Alibiotech grant (2016–2020) obtained from the CPER/FEDER program.

## References

- Aloisi I, Cai G, Serafini-Fracassini D, Del Duca S. 2016. Polyamines in pollen: from microsporogenesis to fertilization. *Frontiers in Plant Science* **7**, 155.
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research* **37**, W202–W208.
- Bassard JE, Mutterer J, Duval F, Werck-Reichhart D. 2012. A novel method for monitoring the localization of cytochromes P450 and other endoplasmic reticulum membrane associated proteins: a tool

- for investigating the formation of metabolons. The FEBS Journal **279**, 1576–1583.
- Bassard JE, Ullmann P, Bernier F, Werck-Reichhart D.** 2010. Phenolamides: bridging polyamines to the phenolic metabolism. *Phytochemistry* **71**, 1808–1824.
- Božičević A, De Mieri M, Nassenstein C, Wiegand S, Hamburger M.** 2017. Secondary metabolites in allergic plant pollen samples modulate afferent neurons and murine tracheal rings. *Journal of Natural Products* **80**, 2953–2961.
- Burhenne K, Kristensen BK, Rasmussen SK.** 2003. A new class of N-hydroxycinnamoyltransferases. Purification, cloning, and expression of a barley agmatine coumaroyltransferase (EC 2.3.1.64). *The Journal of Biological Chemistry* **278**, 13919–13927.
- Calabria LM, Emerenciano VP, Ferreira MJP, Scotti MT, Mabry TJ.** 2007. A phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. *Natural Product Communications* **2**, 277–285.
- Chae L, Kim T, Nilo-Poyanco R, Rhee SY.** 2014. Genomic signatures of specialized metabolism in plants. *Science* **344**, 510–513.
- Clough SJ, Bent AF.** 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**, 735–743.
- D'Auria JC.** 2006. Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology* **9**, 331–340.
- Delporte M, Legrand G, Hilbert JL, Gagneul D.** 2015. Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. *Frontiers in Plant Science* **6**, 651.
- Elejalde-Palmett C, de Bernonville TD, Gievarec G, Pichon O, Papon N, Courdavault V, St-Pierre B, Giglioli-Guivarc'h N, Lanoue A, Besseau S.** 2015. Characterization of a spermidine hydroxycinnamoyltransferase in *Malus domestica* highlights the evolutionary conservation of trihydroxycinnamoyl spermidines in pollen coat of core Eudicotyledons. *Journal of Experimental Botany* **66**, 7271–7285.
- Elišová A, Poracká V, Paľove-Balang P, Imrich J, Repečák M.** 2012. Accumulation of tetracoumaroyl spermine in *Matricaria chamomilla* during floral development and nitrogen deficiency. *Zeitschrift für Naturforschung C* **67**, 58–64.
- Fellenberg C, Vogt T.** 2015. Evolutionarily conserved phenylpropanoid pattern on angiosperm pollen. *Trends in Plant Science* **20**, 212–218.
- Grefen C, Donald N, Hashimoto K, Kudla J, Schumacher K, Blatt MR.** 2010. A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *The Plant Journal* **64**, 355–365.
- Grienenberger E, Besseau S, Geoffroy P, Debayle D, Heintz D, Lapiere C, Pollet B, Heitz T, Legrand M.** 2009. A BAHD acyltransferase is expressed in the tapetum of *Arabidopsis* anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *The Plant Journal* **58**, 246–259.
- Habarugira I, Hendriks T, Quillet MC, Hilbert JL, Rambaud C.** 2015. Effects of nuclear genomes on anther development in cytoplasmic male sterile chicories (*Cichorium intybus* L.): morphological analysis. *The Scientific World Journal* **2015**, 529521.
- Hedberg C, Hesse M, Werner C.** 1996. Spermine and spermidine hydroxycinnamoyl transferases in *Aphelandra tetragona*. *Plant Science* **113**, 149–156.
- Höfgen R, Willmitzer L.** 1988. Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acids Research* **16**, 9877.
- Hou X, Tong H, Selby J, Dewitt J, Peng X, He ZH.** 2005. Involvement of a cell wall-associated kinase, WAKL4, in *Arabidopsis* mineral responses. *Plant Physiology* **139**, 1704–1716.
- Karimi M, Inzé D, Depicker A.** 2002. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends in Plant Science* **7**, 193–195.
- Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RB.** 1990. Different temporal and spatial gene expression patterns occur during anther development. *The Plant Cell* **2**, 1201–1224.
- Koncz C, Schell J.** 1986. The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Molecular & General Genetics* **204**, 383–396.
- Lam SC, Liu X, Chen XQ, Hu DJ, Zhao J, Long ZR, Fan B, Li SP.** 2016. Chemical characteristics of different parts of *Coreopsis tinctoria* in China using microwave-assisted extraction and high-performance liquid chromatography followed by chemometric analysis. *Journal of Separation Science* **39**, 2919–2927.
- Laursen T, Møller BL, Bassard JE.** 2015. Plasticity of specialized metabolism as mediated by dynamic metabolons. *Trends in Plant Science* **20**, 20–32.
- Legrand G, Delporte M, Khelifi C, Harant A, Vuylsteker C, Mörchen M, Hance P, Hilbert JL, Gagneul D.** 2016. Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Frontiers in Plant Science* **7**, 741.
- Luo J, Fuell C, Parr A, Hill L, Bailey P, Elliott K, Fairhurst SA, Martin C, Michael AJ.** 2009. A novel polyamine acyltransferase responsible for the accumulation of spermidine conjugates in *Arabidopsis* seed. *The Plant Cell* **21**, 318–333.
- Ma CM, Nakamura N, Hattori M.** 2001. Inhibitory effects on HIV-1 protease of tri-*p*-coumaroylspermidine from *Artemisia caruifolia* and related amides. *Chemical & Pharmaceutical Bulletin* **49**, 915–917.
- Moglia A, Lanteri S, Comino C, Hill L, Kneivt D, Cagliero C, Rubiolo P, Bornemann S, Martin C.** 2014. Dual catalytic activity of hydroxycinnamoyl-coenzyme A quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicafeoylquinic acids. *Plant Physiology* **166**, 1777–1787.
- Murashige T, Skoog F.** 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum* **15**, 473–497.
- Muroi A, Ishihara A, Tanaka C, Ishizuka A, Takabayashi J, Miyoshi H, Nishioka T.** 2009. Accumulation of hydroxycinnamic acid amides induced by pathogen infection and identification of agmatine coumaroyltransferase in *Arabidopsis thaliana*. *Planta* **230**, 517–527.
- Murphy DJ.** 2006. The extracellular pollen coat in members of the Brassicaceae: composition, biosynthesis, and functions in pollination. *Protoplasma* **228**, 31–39.
- Nagel R, Elliott A, Masel A, Birch RG, Manners JM.** 1990. Electroporation of binary Ti plasmid vector into *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. *FEMS Microbiology Letters* **67**, 325–328.
- Ohta S, Fujimaki T, Uy MM, Yanai M, Yukiyoishi A, Hirata T.** 2007. Antioxidant hydroxycinnamic acid derivatives isolated from Brazilian bee pollen. *Natural Product Research* **21**, 726–732.
- Onkokesung N, Gaquerel E, Kotkar H, Kaur H, Baldwin IT, Galis I.** 2012. MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-coenzyme A:polyamine transferases in *Nicotiana attenuata*. *Plant Physiology* **158**, 389–407.
- Pacini E, Hesse M.** 2005. Pollenkitt – its composition, forms and functions. *Flora - Morphology, Distribution, Functional Ecology of Plants* **200**, 399–415.
- Panero JL, Crozier BS.** 2016. Macroevolutionary dynamics in the early diversification of Asteraceae. *Molecular Phylogenetics and Evolution* **99**, 116–132.
- Park SB, Song K, Kim YS.** 2017. Tetra-*cis/trans*-coumaroyl polyamines as NK<sub>1</sub> receptor antagonists from *Matricaria chamomilla*. *Planta Medica International* **4**, e43–e51.
- Petersen M, Abdullah Y, Benner J, et al.** 2009. Evolution of rosmarinic acid biosynthesis. *Phytochemistry* **70**, 1663–1679.
- Piffanelli P, Ross JHE, Murphy DJ.** 1998. Biogenesis and function of the lipidic structures of pollen grains. *Sexual Plant Reproduction* **11**, 65–80.
- Quilichini TD, Grienenberger E, Douglas CJ.** 2015. The biosynthesis, composition and assembly of the outer pollen wall: a tough case to crack. *Phytochemistry* **113**, 170–182.
- Quilichini TD, Samuels AL, Douglas CJ.** 2014. ABCG26-mediated polyketide trafficking and hydroxycinnamoyl spermidines contribute to pollen wall exine formation in *Arabidopsis*. *The Plant Cell* **26**, 4483–4498.
- Ralston L, Yu O.** 2006. Metabolons involving plant cytochrome P450s. *Phytochemistry Reviews* **5**, 459–472.
- Schmidt GW, Jirschtzka J, Porta T, et al.** 2015. The last step in cocaine biosynthesis is catalyzed by a BAHD acyltransferase. *Plant Physiology* **167**, 89–101.

- Sobolev VS, Sy AA, Gloer JB.** 2008. Spermidine and flavonoid conjugates from peanut (*Arachis hypogaea*) flowers. *Journal of Agricultural and Food Chemistry* **56**, 2960–2969.
- St-Pierre B, De Luca V.** 2000. Evolution of acyltransferase genes: origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. *Recent Advances in Phytochemistry* **34**, 285–315.
- Sullivan M.** 2009. A novel red clover hydroxycinnamoyl transferase has enzymatic activities consistent with a role in phaselic acid biosynthesis. *Plant Physiology* **150**, 1866–1879.
- Sullivan ML.** 2017. Identification of bean hydroxycinnamoyl-CoA:tetrahydroxyhexanedioate hydroxycinnamoyl transferase (HHHT): use of transgenic alfalfa to determine acceptor substrate specificity. *Planta* **245**, 397–408.
- Tanabe K, Hojo Y, Shinya T, Galis I.** 2016. Molecular evidence for biochemical diversification of phenolamide biosynthesis in rice plants. *Journal of Integrative Plant Biology* **58**, 903–913.
- Tassoni A, Franceschetti M, Bagni N.** 2008. Polyamines and salt stress response and tolerance in *Arabidopsis thaliana* flowers. *Plant Physiology and Biochemistry* **46**, 607–613.
- Tassoni A, Van Buuren M, Franceschetti M, Fornalè S, Bagni N.** 2000. Polyamine content and metabolism in *Arabidopsis thaliana* and effect of spermidine on plant development. *Plant Physiology and Biochemistry* **38**, 383–393.
- Tuominen LK, Johnson VE, Tsai CJ.** 2011. Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among *Populus paralogues*. *BMC Genomics* **12**, 236.
- Waadt R, Kudla J.** 2008. In planta visualization of protein interactions using bimolecular fluorescence complementation (BiFC). *Cold Spring Harbor Protocols* **2008**, doi:10.1101/pdb.prot4995.
- Weng JK, Noel JP.** 2012. The remarkable pliability and promiscuity of specialized metabolism. *Cold Spring Harbor Symposia on Quantitative Biology* **77**, 309–320.
- Weng JK, Noel JP.** 2013. Chemodiversity in *Selaginella*: a reference system for parallel and convergent metabolic evolution in terrestrial plants. *Frontiers in Plant Science* **4**, 119.
- Weng JK, Philippe RN, Noel JP.** 2012. The rise of chemodiversity in plants. *Science* **336**, 1667–1670.
- Yamamoto A, Nakamura K, Furukawa K, Konishi Y, Ogino T, Higashiura K, Yago H, Okamoto K, Otsuka M.** 2002. A new nonpeptide tachykinin NK<sub>1</sub> receptor antagonist isolated from the plants of Compositae. *Chemical & Pharmaceutical Bulletin* **50**, 47–52.
- Yogendra KN, Sarkar K, Kage U, Kushalappa AC.** 2017. Potato NAC43 and MYB8 mediated transcriptional regulation of secondary cell wall biosynthesis to contain *Phytophthora infestans* infection. *Plant Molecular Biology Reporter* **35**, 519–533.



# Identification and Characterization of Five BAHD Acyltransferases Involved in Hydroxycinnamoyl Ester Metabolism in Chicory

Guillaume Legrand, Marianne Delporte, Chahinez Khelifi, Adeline Harant, Christophe Vuylsteker, Monika Mörchen, Philippe Hance, Jean-Louis Hilbert and David Gagneul\*

Agro-Food and Biotechnology Research Institute, EA 7394, Charles Viollette Research Institute, Université de Lille 1, Villeneuve d'Ascq, France

## OPEN ACCESS

### Edited by:

David Caparros-Ruiz,  
Centre for Research in Agricultural  
Genomics, Spain

### Reviewed by:

Fumiya Kurosaki,  
University of Toyama, Japan  
Jules Beekwilder,  
Plant Research International –  
Wageningen UR, Netherlands

### \*Correspondence:

David Gagneul  
david.gagneul@univ-lille1.fr

### Specialty section:

This article was submitted to  
Plant Metabolism  
and Chemodiversity,  
a section of the journal  
Frontiers in Plant Science

**Received:** 30 March 2016

**Accepted:** 16 May 2016

**Published:** 06 June 2016

### Citation:

Legrand G, Delporte M, Khelifi C,  
Harant A, Vuylsteker C, Mörchen M,  
Hance P, Hilbert J-L and Gagneul D  
(2016) Identification  
and Characterization of Five BAHD  
Acyltransferases Involved  
in Hydroxycinnamoyl Ester  
Metabolism in Chicory.  
Front. Plant Sci. 7:741.  
doi: 10.3389/fpls.2016.00741

Chicory (*Cichorium intybus*) accumulates caffeic acid esters with important significance for human health. In this study, we aim at a better understanding of the biochemical pathway of these bioactive compounds. Detailed metabolic analysis reveals that *C. intybus* predominantly accumulates caffeic and chicoric acids in leaves, whereas isochlorogenic acid (3,5-diCQA) was almost exclusively accumulated in roots. Chlorogenic acid (3-CQA) was equally distributed in all organs. Interestingly, distribution of the four compounds was related to leaf age. Induction with methyljasmonate (MeJA) of root cell suspension cultures results in an increase of 3-CQA and 3,5-diCQA contents. Expressed sequence tag libraries were screened using members of the BAHD family identified in *Arabidopsis* and tobacco as baits. The full-length cDNAs of five genes were isolated. Predicted amino acid sequence analyses revealed typical features of BAHD family members. Biochemical characterization of the recombinant proteins expressed in *Escherichia coli* showed that two genes encode HCTs (hydroxycinnamoyl-CoA:shikimate/quinic hydroxycinnamoyltransferases, HCT1 and HCT2) whereas, three genes encode HQTs (hydroxycinnamoyl-CoA:quinic hydroxycinnamoyltransferases, HQT1, HQT2, and HQT3). These results totally agreed with the phylogenetic analysis done with the predicted amino acid sequences. Quantitative real-time polymerase chain reaction analysis of gene expression indicated that HQT3, HCT1, and HCT2 might be more directly associated with CQA accumulation in cell culture in response to MeJA elicitation. Transient expression of HCT1 and HQT1 in tobacco resulted in a higher production of 3-CQA. All together these data confirm the involvement of functionally redundant genes in 3-CQA and related compound synthesis in the *Asteraceae* family.

**Keywords:** caffeic acid esters, BAHD family, acyltransferases, chlorogenic acid, chicory (*Cichorium intybus*), functionally redundant genes

## INTRODUCTION

Plants accumulate a wide range of specialized metabolites with a large diversity of chemical types. Phenolic compounds are recognized for their health benefit effects and are the most important dietary antioxidants. They have drawn increasing attention due to their marked effect in the prevention of various oxidative stress associated diseases. Among them, hydroxycinnamic acid

derivatives, particularly hydroxycinnamoyl esters, are widely distributed in the plant kingdom (Petersen et al., 2009). For example, a caffeoyl moiety can be combined to quinic acid to form either moncaffeoylquinic acid (chlorogenic acid, CQA) as found in tobacco (*Nicotiana tabacum* L.) or in tomato (*Solanum lycopersicum* L.; Niggeweg et al., 2004) or dicaffeoylquinic acid (isochlorogenic acid, diCQA) as in coffee (*Coffea* spp.), tomato or sweet potato (*Ipomea batatas* L.; Kojima and Kondo, 1985; Lallemand et al., 2012; Moglia et al., 2014). In lettuce (*Lactuca sativa* L.) or red clover (*Trifolium pratense* L.), the caffeoyl group is attached to malic acid to form phaselic acid (Sullivan, 2009; Mai and Glomb, 2013). Esters of tartaric acid such as moncaffeoyltartaric acid [caftaric acid (CTA)] are also found in grape vine (*Vitis vinifera* L.), perennial peanut (*Arachis glabrata* L.) and several members of the *Asteraceae* like purple coneflower (*Echinacea purpurea* L.) which also contains dicaffeoyltartaric acid (chicoric acid, diCTA; Singleton et al., 1986; Perry et al., 2001; Sullivan and Foster, 2013). Chicory (*Cichorium intybus* L.) is a member of the *Asteraceae* family used for a long time in traditional medicine. This is notably due to the accumulation of high-value health promoting compounds such as CQA, diCQA, CTA, and diCTA (Kandeler and Ullrich, 2009; Bahri et al., 2012). Indeed, many health benefit effects are attributed to these molecules when isolated (Tousch et al., 2008; Koriem and Soliman, 2014; Yasir et al., 2016). In addition to their notable dietary role, these molecules are important compounds with multiple roles in plants. They are involved in plant protection against abiotic (UV, oxidative stress) and biotic (insects, pathogens) stresses. Indeed, in *V. vinifera*, CTA accumulation was shown to be related to the presence of UV radiation supporting a putative role of this molecule in UV protection (Del-Castillo-Alonso et al., 2014). In globe artichoke (*Cynara cardunculus* L.), it was also hypothesized that diCQA could play a role in UV protection (Moglia et al., 2008). Increased CQA accumulation in *S. lycopersicum* was shown to improve antioxidant activity and resistance to infection by *Pseudomonas syringae* (Niggeweg et al., 2004). Furthermore, CQA as well as feruloylquinic acid were also identified as protective agents against thrips in chrysanthemum (*Dendranthema grandiflora*; Leiss et al., 2009).

A better understanding of CQA, diCQA, CTA, and diCTA metabolic pathways is of paramount importance to develop agronomic, genetic, or biotechnological tools for higher production of these high-value bioactive compounds. In this context, *C. intybus* could be a model species considering that the metabolism of these molecules is highly interconnected (Figure 1). The aromatic parts of the esters are synthesized via the phenylpropanoid pathway. Most enzymes involved in the first part of this pathway are known for years (Vanholme et al., 2010). The entry point is the aromatic amino acid phenylalanine (Phe) arising from the shikimate pathway. Deamination of Phe by Phe ammonia lyase (PAL) leads to cinnamic acid. Cinnamate-4-hydroxylase and 4-coumarate coenzyme A (CoA) ligase (4CL) generate *p*-coumaroyl-CoA from cinnamic acid. Thereafter, hydroxycinnamoyltransferases (HCTs) convert the CoA-thioester to coumaroylquinic acid or coumaroylshikimate which is subsequently hydroxylated

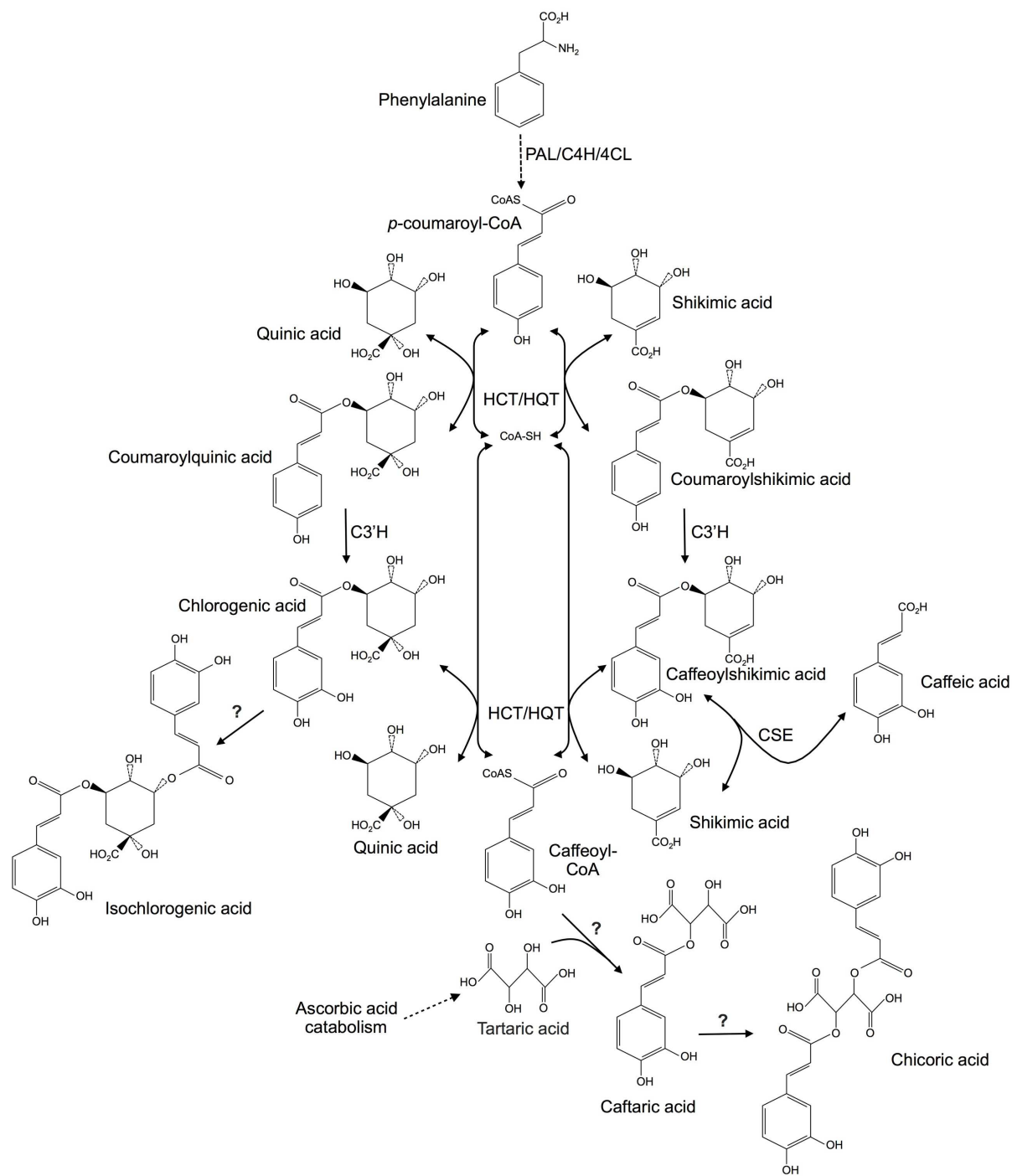
by *p*-coumarate-3'-hydroxylase (C3'H) to form the caffeoyl derivatives. Two types of transferases have been identified. Hydroxycinnamoyl-CoA: shikimate/quinic acid HCTs use shikimate as a preferred acyl acceptor whereas hydroxycinnamoyl-CoA: quinic acid hydroxycinnamoyltransferases (HQTs) exhibit preference for quinate (Lallemand et al., 2012). Experimental evidence has shown that HCTs may be involved in the synthesis of precursors for lignin synthesis whereas HQTs may be directly involved in CQA synthesis toward its accumulation (Hoffmann et al., 2003; Niggeweg et al., 2004). HQT and HCT activities are fully reversible: quinate or shikimate derivatives can be converted in presence of CoA to the free acid plus *p*-coumaroyl-CoA or caffeoyl-CoA. Recently, a caffeoylshikimate esterase (CSE) that releases caffeic acid and shikimate has been identified in *Arabidopsis* (Figure 1; Vanholme et al., 2013). HCTs and HQTs belong to the BAHD superfamily of plant-specific acyl-CoA dependent acyltransferases (St-Pierre and De Luca, 2000; D'Auria, 2006; Yu et al., 2009; Tuominen et al., 2011). In *I. batatas*, diCQA arises from the transfer of a caffeoyl group from a CQA to a second CQA (Villegas et al., 1987). An enzyme called chlorogenic acid: chlorogenate caffeoyl transferase catalyzes this reaction. The protein was partially characterized but the corresponding gene has not been identified (Villegas et al., 1987). Recent study identified a HCT from *Coffea canephora* able to convert CQA to diCQA (Lallemand et al., 2012). A *S. lycopersicum* HQT was also shown to be involved in the formation of diCQA from CQA (Moglia et al., 2014). In *A. glabrata* and *Equisetum arvense*, it was shown that CTA is synthesized through condensation of a caffeoyl-CoA with tartaric acid which is a typical mechanism of the BAHD family members (Hohlfeld et al., 1996; Sullivan, 2014). Furthermore chicoric acid was demonstrated *in vitro* to be synthesized by transfer of a caffeoyl moiety of caffeoyl-CoA to CTA in *E. arvense* (Hohlfeld et al., 1996). Pathways involved in CTA and diCTA are only supported by biochemical experiments and to date, no molecular data are available. To summarize, all these molecules could arise through the action of acyltransferases of the BAHD family.

In chicory, knowledge on the genetic control of these metabolic pathways is still limited. We report in this paper the isolation, the cloning, and the biochemical and functional characterization of five full-length cDNA sequences encoding HCTs or HQTs in *C. intybus*. The properties of these enzymes are consistent with a role in CQA production and detailed analysis confirms that both types of enzymes are subsequently needed to ensure CQA synthesis and accumulation.

## MATERIALS AND METHODS

### Plant Material and Growth Conditions

Seeds of chicory, *C. intybus* L. var Orchies, were provided by Florimond-Desprez (Cappelle en Pève, France). They were germinated for 7 days in Petri dishes on paper humidified with half strength MS medium at pH 6.0 in a growth chamber (Photoperiod, 16 h: 8 h, light: dark; thermoperiod, 24°C: 18°C, day: night; Murashige and Skoog, 1962). The 7 days-old seedlings were then transferred hydroponically to culture pots containing



**FIGURE 1 | Putative metabolic pathways involved in hydroxycinnamic acid biosyntheses in plants.** Question marks indicate enzymes that have been characterized biochemically but which are not supported by molecular data. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaroyl CoA ligase; HCT, hydroxycinnamoyl-CoA:shikimate/quinic acid hydroxycinnamoyltransferase; HQT, hydroxycinnamoyl-CoA:quinic acid hydroxycinnamoyltransferase; C3'H, *p*-coumaroyl ester 3'-hydroxylase; CSE, caffeoyl shikimate esterase.

half strength MS solution. The volume of the solution was daily adjusted with fresh medium and the medium completely renewed weekly. For analysis of HCT and HQT mRNA levels, for hydroxycinnamoyl ester contents and for full length cDNA

cloning, either 6 or 4 weeks-old seedlings were individually harvested. Three biological replicates were sampled. Roots and leaves were separately collected and for the 4 weeks-old seedlings, leaves were divided in three samples: old leaves, medium age



leaves (intermediate leaves) and newly expanding leaves (young leaves). Samples were immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until needed. Chicory cell cultures were prepared as described in Delporte et al. (2015).

### Isolation of Full-Length cDNA Sequences

Putative BAHD members of clade Vb were identified by searching publicly available databases with AtHCT and NtHQT. Full-length cDNAs of CiHCT1, CiHCT2, CiHQT1, CiHQT2, and CiHQT3 were first subcloned into a TA cloning vector (pGEM-T easy, Promega). DNA sequences were amplified by PCR using proofreading polymerase (PrimeSTAR HS DNA polymerase, Takara). The PCR primers designed for these amplifications are listed in Supplementary Table S1. The sequences of independent clones were determined in their entirety on both strands and the consensus sequences established. One clone per gene was kept and was used for following cloning (pGEMT-easy-HCT1, pGEMT-easy-HCT2, pGEMT-easy-HQT1, pGEMT-easy-HQT2, and pGEMT-easy-HQT3).

### Phylogenetic Tree of Acyltransferases and Multiple Sequence Alignment

Hydroxycinnamoyltransferases cDNAs from chicory were translated to the corresponding amino acid sequences. Phylogenetic tree and multiple alignments were generated by CLC Sequence Viewer 7 software.

### Expression, Purification, and *In Vitro* Assays of Recombinant Proteins

Full-length cDNAs of the five genes were amplified using a proofreading polymerase (PrimeSTAR HS DNA polymerase, Takara) using the previously mentioned sequenced clones (pGEMT-easy-HCT1, 2 and pGEMT-easy-HQT1, 2, or 3) as the templates and the primers listed in Supplementary Table S1. The entry clones (pDONR221-HCT1 or 2 and pDONR221-HQT1, 2, or 3) were obtained through recombination of the PCR products with pDONR221 (Invitrogen). For further expression analysis, the open reading frames were introduced into pDEST17 expression vector (Invitrogen) by LR recombination to produce pDEST17-HCT1 or 2 and pDEST17-HQT1, 2, or 3. Recombinant proteins with an N-terminal 6xHis-tag were expressed in *Escherichia coli* BL21-CodonPlus-(DE3)-RIL cells following induction with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside. After induction, cells were harvested by centrifugation and pellets stored at  $-80^{\circ}\text{C}$ . Cell pellets were resuspended in the extraction buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , 500 mM NaCl, 20 mM imidazole pH 7.4) added with 1 mg  $\text{ml}^{-1}$  lysozyme and the suspensions were incubated on ice for 30 min. After sonication on ice, samples were centrifuged (10000 g,  $4^{\circ}\text{C}$ , 30 min) and the supernatants were loaded onto 1 ml HisTrap HP columns (GE Healthcare) and processed according to the manufacturer's procedures. The positive fractions were identified by SDS-PAGE, pooled, desalted with a PD-10 desalting column (Pharmacia), and concentrated using vivaspin sample concentrators with 30 kDa molecular mass cutoff (GE

Healthcare). Purified proteins were stored at  $-20^{\circ}\text{C}$  until needed.

The standard enzymatic *in vitro* assay was performed in a volume of 50  $\mu\text{L}$  containing acyl donor (hydroxycinnamoyl-CoA) and acyl acceptor (shikimate or quinate) at various concentrations, 0.1–1  $\mu\text{g}$  of purified enzyme and 50 mM phosphate buffer pH 7.1. The reactions were incubated at room temperature and stopped by adding 10  $\mu\text{l}$  formic acid 10%. The reaction mixture were then filtered through a 0.45  $\mu\text{m}$  filter (Pall GHP, VWR) and analyzed by HPLC-UV.  $K_m$  and  $V_{max}$  values were determined in triplicates by fitting Michaelis-Menten curves directly using GraphPad-Prism Software. Saturating concentration of acyl-CoAs was set at 800  $\mu\text{M}$ . For evaluation of the optimal pH, the reaction mixture was incubated at room temperature in 50 mM phosphate buffer at pH ranging from 5.0 to 9.0 or in 50 mM acetate buffer at pH ranging from 4.0 to 5.0. Reactions were run in triplicates.

### Enzymatic Production and Purification of Hydroxycinnamoyl-CoA Compounds

*p*-coumaroyl-CoA, caffeoyl-CoA, feruloyl-CoA, and cinnamoyl-CoA were prepared enzymatically from each acid and purified (Beuerle and Pichersky, 2002; Rautengarten et al., 2010). Purified *Nicotiana benthamiana* 4-coumarate: CoA ligase 1 (Nb4CL1) recombinant protein from *E. coli* was used to synthesize the CoA-thioesters following the procedure described by Rautengarten et al. (2010). The plasmid for protein expression was kindly provided by Pr. Sheller (Emeryville, CA, USA). The CoA-thioesters were then purified using a SPE cartridge (Beuerle and Pichersky, 2002). Yields were 97, 90, 95, and 87% for caffeoyl-CoA, *p*-coumaroyl-CoA, feruloyl-CoA, and cinnamoyl-CoA, respectively.

### gDNA Extraction, Amplification, and Sequencing

gDNA was extracted from fully expanded leaves using the Nucleospin Plant II kit (Macherey-Nagel). DNA sequences were amplified by PCR using proofreading polymerase (PrimeSTAR HS DNA polymerase, Takara). The PCR primers designed for these amplifications are listed in Supplementary Table S1. PCR products were sequenced to determine intron sequences and sizes.

### RNA Extraction and cDNA Synthesis

Total RNA was isolated from the different parts of young seedlings and from culture cells as described in Delporte et al. (2015). The yield and purity of total RNA were determined using the Experion Automated Electrophoresis System (Bio-Rad). One  $\mu\text{g}$  of DNase treated total RNA was used for cDNA synthesis with the Reverse Transcriptase Superscript III RNase H kit and oligo(dT)<sub>20</sub> primer (Invitrogen).

### qRT-PCR Analysis

Gene specific primers (listed in Supplementary Table S1) were designed using Primer 3 software using recommended parameters (Udvardi et al., 2008; melting temperature =  $60 \pm 1^{\circ}\text{C}$ ;

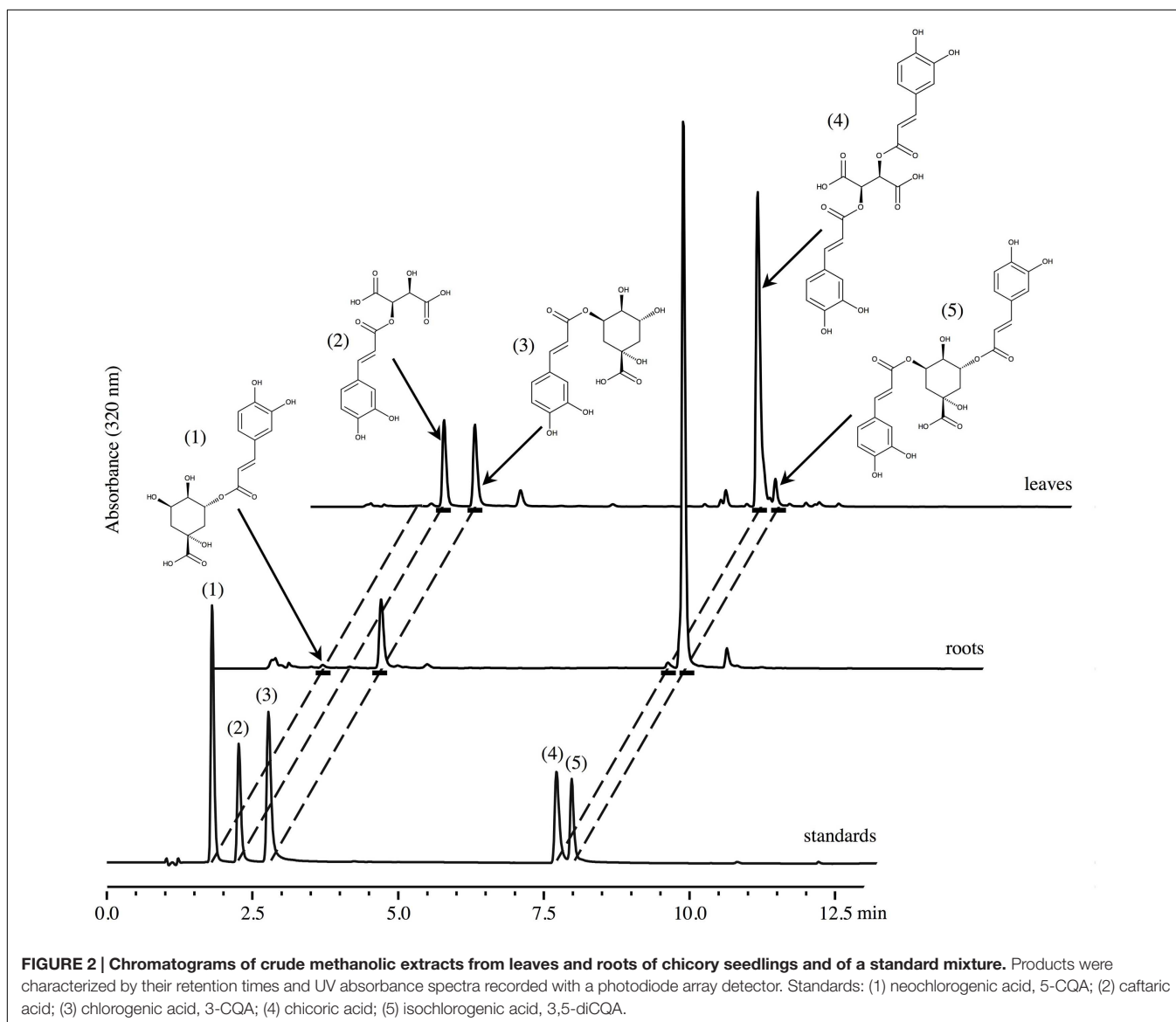
length of 18–25 nucleotides; 40–60% GC; length of amplicon from 60 to 150 bp). qRT-PCR was carried out using iQ SYBR Green Supermix (Bio-Rad). Reactions were set up in a 20  $\mu$ l total volume containing cDNA equivalent to 100 ng of total RNA and 5  $\mu$ M of each primer. Signals were normalized using geometric mean of *CLATH* and *SAND* mRNA levels for seedling experiments and of *TIP41* and *PP2AA2* mRNA levels for cell culture experiments as described by Delporte et al. (2015). Initial denaturation was 95°C for 3 min followed by 40 cycles including 95°C for 10 s and 60°C for 30 s. A standard dissociation protocol was run at the end of each run to ensure that each amplicon was a single product. Each reaction was run in duplicates. Control PCRs were run with non retro-transcribed RNA to check for gDNA contamination. Calculation of relative expression was done according to the Pfaffl equation (Pfaffl, 2001).

## Determination of Protein Concentrations, SDS-PAGE Analysis, and Immunoblot Analysis

Protein concentrations of the extracts were determined using Bio-Rad Protein Assay using bovine serum albumin as a standard. SDS-PAGE and immunoblot analysis was conducted as described elsewhere (Sambrook et al., 1989). The following antibody combination were used for immunodetection: penta-His antibody (Qiagen)/alkaline phosphatase (AP)-conjugate anti-mouse IgG (Promega). To estimate molecular masses, the Precision Plus Protein All blue Standards ladder (Bio-Rad) was used.

## Transient Expression in *N. benthamiana*

Previously obtained clones pGEMT-easy-HCT1 and pGEMT-easy-HQT1 were used to amplify full-length cDNA of *HCT1*



and *HQT1* with the primers described in Supplementary Table S1. Each PCR product was introduced into pDONR221 by recombination. Open reading frame were then introduced into the expression vector pB2GW7 by LR recombination (Karimi et al., 2002). Empty pB2GW7 was also generated by first cutting the vector with EcoRV (NEB) to remove the RfA containing the *ccdB* gene. Then, the empty vector was re-circularized using T4 DNA Ligase (NEB). The resulting vector pB2GW7-HCT1, pB2GW7-HQT1, empty pB2GW7 as well as the pEAQ-HT vector, harboring the gene encoding the silencing inhibitor protein p19 (Sainsbury et al., 2009), were individually introduced into the *Agrobacterium tumefaciens* strain GV2260 by electroporation. These four recombinant strains were grown overnight in YEB medium at 28°C, with shaking at 200 rpm. Thereafter, the cells were pelleted by centrifugation at 5000 g for 15 min and resuspended in activation buffer containing 10 mM MgCl<sub>2</sub> and 150 µg ml<sup>-1</sup> acetosyringone before incubation at room temperature for 3 h. For agroinfiltration, the strain containing the pEAQ-HT was used alone or combined with strain transformed with either pB2GW7-HQT1 or pB2GW7-HCT1 or empty pB2GW7. Whatever the case, final OD<sub>600</sub> was 0.8 (0.3 pEAQ-HT + 0.5 of the other construct). Each construct or construct combination was used to infiltrate the abaxial air space of two individual leaves per plant. Four 6 weeks-old plants per construct were used. After 4 days, the infiltrated leaves were collected and immediately frozen in liquid nitrogen before freeze-drying.

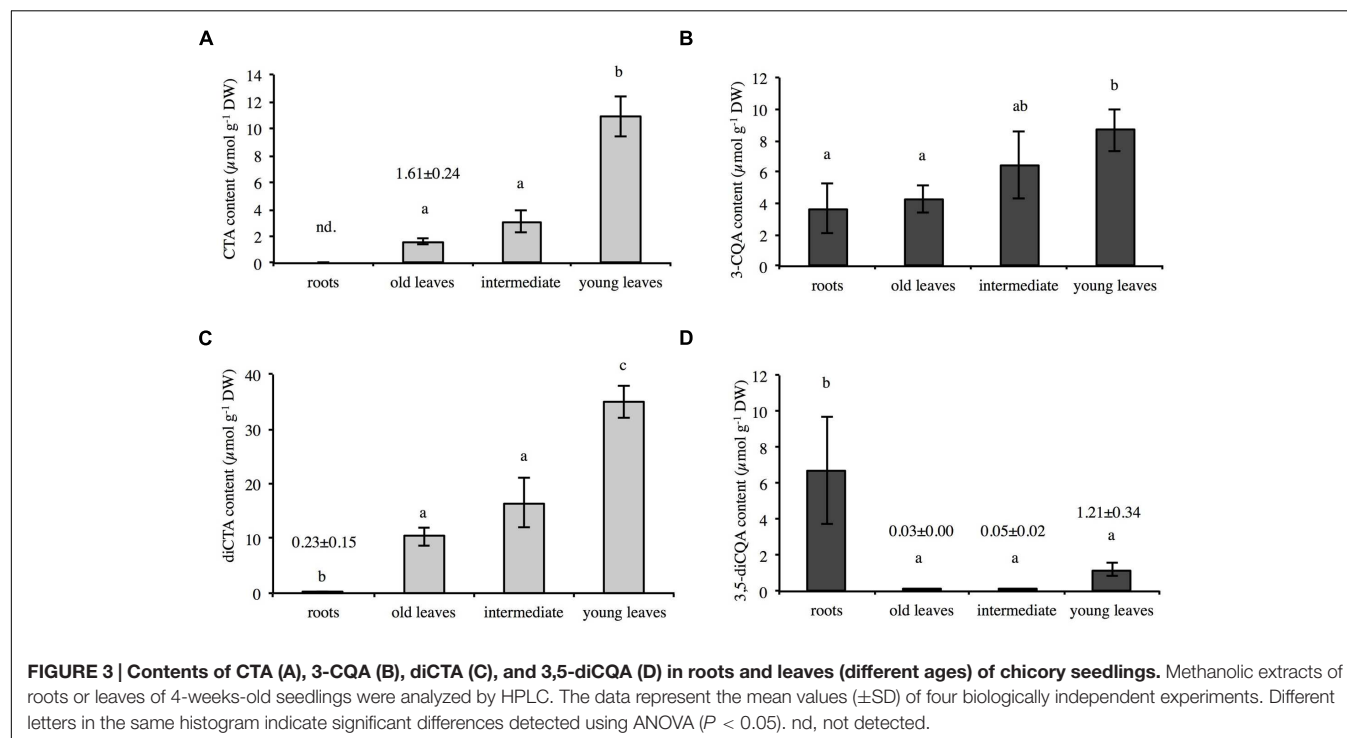
## HPLC Analysis of Polyphenols

The lyophilized plant material was powdered and 50 mg were resuspended in 1 ml of a methanol/water/acetic acid

mixture (75/23/2, v/v/v). The mixtures were then incubated under agitation for 12 h at 4°C. Homogenates were clarified by centrifugation (14000 g, 4°C, 10 min) and 300 µl of supernatant were transferred in a new tube. After addition of 150 µl chloroform and 150 µl water, the suspensions were thoroughly shaken and centrifuged (14000 g, 5 min, 20°C). The upper phase was collected and passed through a 0.45 µm filter and 5 µl aliquots were analyzed on a 100 mm × 4.6 mm Kinetex 2.6 µm PFP 100 Å column (Phenomenex). For determination of phenolics contents in chicory and tobacco and for kinetics parameter of the enzymes, the chromatographic separation was performed using water (solvent A) and acetonitrile (solvent B) both acidified with 0.1% ortho-phosphoric acid. The solvents were delivered at a flow rate of 1.1 ml min<sup>-1</sup> and the oven temperature was set at 45°C with start condition at 10% solvent B. The HPLC conditions were as follow: 7 min gradient to 30% solvent B followed by 3 min gradient to 70% and 1.5 min of isocratic 70% solvent B. Then, 1 min gradient to return to 10% solvent B and 8.5 min of isocratic re-equilibration at 10% solvent B. Phenolics were characterized by cochromatography of pure synthetic compounds and quantified making reference to individual external calibration curves at 320 nm.

## Statistics

Statistical analysis were conducted using R 3.2.2 for Mac and used to determined between extract variation for both chicory and tobacco extract (R Core Team, 2015). In accordance with need, ANOVA or Student's test were used.



## RESULTS

### Profiling of the Main Caffeic Acid Esters in Chicory Tissues

The amounts of CQA, diCQA, CTA, and diCTA were first determined in 6 weeks-old chicory seedlings grown hydroponically in half strength MS medium (Murashige and Skoog medium). Methanolic extracts of roots or leaves were analyzed by HPLC-UV. Typical chromatograms are presented in **Figure 2** and quantifications indicated in Supplementary Table S2. Notably, only trace amounts of diCTA and no CTA could be detected in roots whereas diCQA was poorly abundant in leaves. CQA, a putative precursor of all three other molecules, did not allow to discriminate between organs. In roots, diCQA was the most abundant compound ( $6.7 \pm 3.0 \mu\text{mol g}^{-1}$  DW, i.e., 62.6% of total caffeate derivatives) while in leaf, diCTA was the most abundant ( $18.9 \pm 2.2 \mu\text{mol g}^{-1}$  DW, i.e., 63.4% of total caffeate derivatives). Detailed analyses show that chicory predominantly accumulates one of the CQA isomers, 3-CQA (i.e., chlorogenic acid according to the actual CAS nomenclature), and 3,5-diCQA one of the diCQA isomers. However, 5-CQA (i.e., neochlorogenic acid, actual CAS nomenclature) could also be detected in trace amount in chicory roots but not in leaves. Overall leaf phenolic contents were about three times higher than those in roots ( $29.8$  and  $10.7 \mu\text{mol g}^{-1}$  DW, respectively). To extend our analysis, leaves of 4 weeks-old seedlings were separately collected and analyzed. Interestingly, levels of CTA, diCTA, 3-CQA, and 3,5-diCQA were dependent on leaf ages (**Figure 3**). The highest contents were found in the youngest leaves. For instance, CTA content was seven times higher in the youngest leaves than in the older ones (**Figure 3A**).

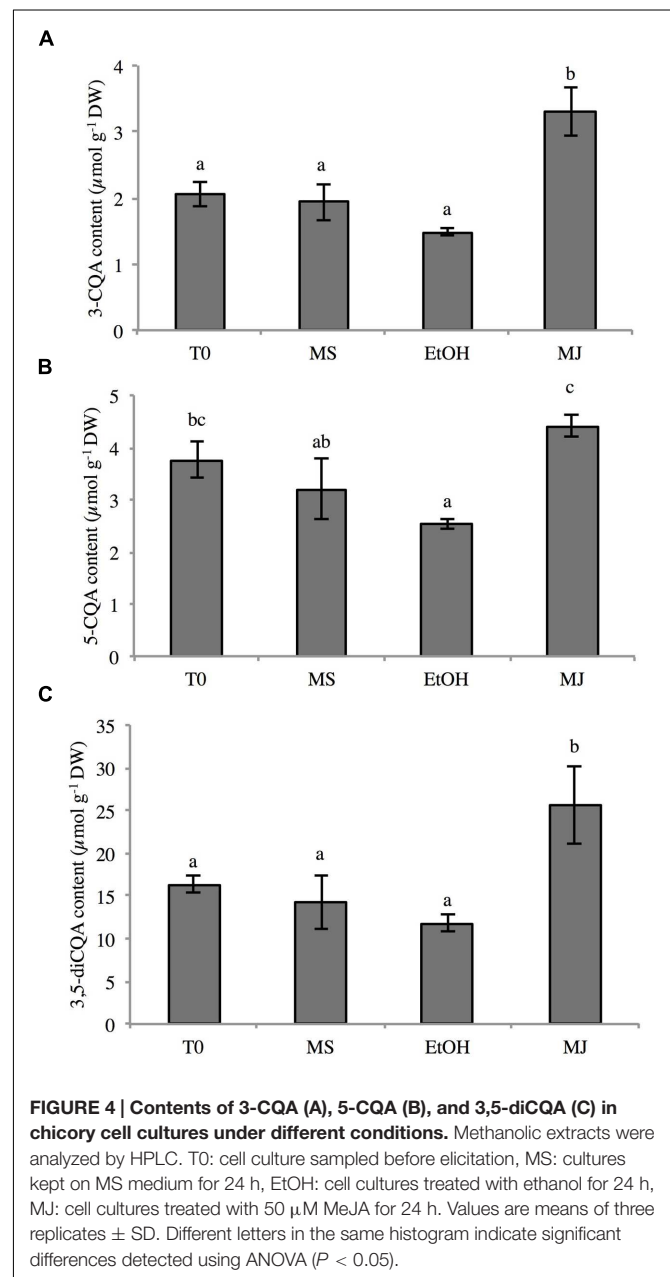
### Modulation of Caffeic Acid Ester Contents by MeJA Elicitation in Root Culture Cells

As an inducer of jasmonates, methyljasmonate (MeJA) regulates a diverse set of physiological and developmental processes. It has been observed that addition of MeJA can significantly alter the production of specialized metabolites (De Geyter et al., 2012; Wasternack and Hause, 2013). Phenylpropanoid compounds were shown to accumulate in *N. tabacum* cell cultures upon MeJA elicitation (Gális et al., 2006). In our lab, we have shown that elicitation of chicory cell cultures with MeJA readily enhances 3-CQA and 3,5-diCQA production and accumulation (unpublished results). Thus, chicory root cell cultures were treated for 24 h with  $50 \mu\text{M}$  MeJA diluted in ethanol (EtOH) and the contents of 3-CQA, 5-CQA, and 3,5-diCQA were analyzed (**Figure 4**). As expected, upon addition of MeJA, the concentrations of 3-CQA and 3,5-diCQA were 2.2 fold higher than in the control cells grown in the presence of the sole ethanol. The levels of 3-CQA and 3,5-diCQA were  $1.5$  and  $11.8 \mu\text{mol g}^{-1}$  DW, respectively in the cells treated with EtOH and  $3.3$  and  $25.6 \mu\text{mol g}^{-1}$  DW in the cells elicited with MeJA. As previously described in chicory roots, 5-CQA was also detected in this material but, surprisingly, in quantity quite similar to that of 3-CQA (**Figure 4**). Despite a variation in 5-CQA quantity less pronounced than these of

the other quantified phenolics, significant difference between elicited cells and cells treated with EtOH is notable ( $4.4$  and  $2.5 \mu\text{mol g}^{-1}$  DW respectively). No CTA or diCTA could be detected in this suspension culture. This is in agreement with the seedling data.

### Identification, Cloning, and Sequence Analysis of Five Hydroxycinnamoyl Transferase Genes in Chicory

To identify sequences of putative members of the BAHD family from chicory, the amino acid sequences of HCT from *Arabidopsis thaliana*, i.e., AtHCT (GenBank accession



AED95744) and of HQT from *N. tabacum*, i.e., NtHQT (GenBank accession CAE46932) were used as the queries in a tBLASTn search of the EST sequences publicly available. The ESTs showing the highest similarity to either AtHCT or NtHQT were selected for further analysis. This search identified 15 ESTs (GenBank accessions EH696409, FL673648, DT213617, EH686422, EH674585, EH692212, FL682645, EH691266, EH677208, EH692526, EH705216, EH706054, EH682505, EH692394, and EH701989) that were further assembled in four contigs. According to the homology of the predicted amino acid sequences to either the AtHCT or NtHQT, they were named *HCT1*, *HCT2* (GenBank accession KT222892), *HQT1* (GenBank accession KT222893), and *HQT2* (GenBank accession KT222894). Examination of predicted amino acid sequences showed that *HCT2*, *HQT1*, and *HQT2* were the full-length cDNA. No additional overlapping ESTs were identified to extend the *HCT1* sequence. For this reason, a chicory bacterial artificial chromosome library was screened and one positive clone was identified (Gonthier et al., 2010). The 5' sequence was extended by sequencing to get the full-length sequence (GenBank accession KT222891). To get more putative sequences, the recently released 454 reference assemblies of *C. intybus* was interrogated and an additional HQT, *HQT3* (comp5746\_C0-seq1, GenBank accession KT222895), was identified (Hodgins et al., 2014).

The *HCT1* coding region shares 81% nucleotide sequence identity with that of *HCT2*. *HQT2* coding region shares 74 and 66% nucleotide sequence identity with that of *HQT3* and *HQT1*, respectively. *HQT1* and *HQT3* share 67% nucleotide sequence identity. Comparison of *HCT* coding region with that of *HQT* gave a maximum score of 61%. Gene sizes were 1305, 1293, 1317, 1320, and 1326 bp for *HCT1*, *HCT2*, *HQT1*, *HQT2*, and *HQT3*, respectively. *HCT1* encodes a protein of 434 amino acids with a calculated molecular mass of 48.1 kDa. The peptide product of *HCT2* consists of 430 amino acid residues and has a predicted molecular mass of 47.7 kDa. The 438-amino acid protein encoded by *HQT1* has a predicted molecular mass of 48.6 kDa. *HQT2* and *HQT3* encode proteins of 439 and 441 amino acids, respectively, with calculated molecular masses of 48.6 and 48.8 kDa.

Phylogenetic analysis of predicted amino acid sequences confirm our first analysis, i.e., *HCT1* and *HCT2* group together with biochemically characterized HCTs whereas *HQT1*, *HQT2*, and *HQT3* group with characterized HQTs (Figure 5A). All five proteins contain the two conserved motifs, HXXXDG and DFGWG that are observed among BAHD acyltransferases (Figure 5B). The first one, implicated in the active site, is strongly conserved whereas the second one could be slightly altered. *HCT1* and *HCT2* contain the same motif HHAADG in the middle part of the protein whereas *HQT1* and *HQT2* share the motif HTLSDG. For *HQT3*, a unique motif HTLADG was found. In the C-terminal part of the proteins, all proteins harbor the same DFGWG motif except *HQT2* which amino sequence is DFGYG. Sequencing and analyses of genomic sequences show that *HCT1*, *HQT1*, *HQT2*, and *HQT3* exhibit a single intron whereas *HCT2* has none. The introns are 242, 1203, 1293, and 873 bp in length for *HCT1*, *HQT1*, *HQT2*, and *HQT3*,

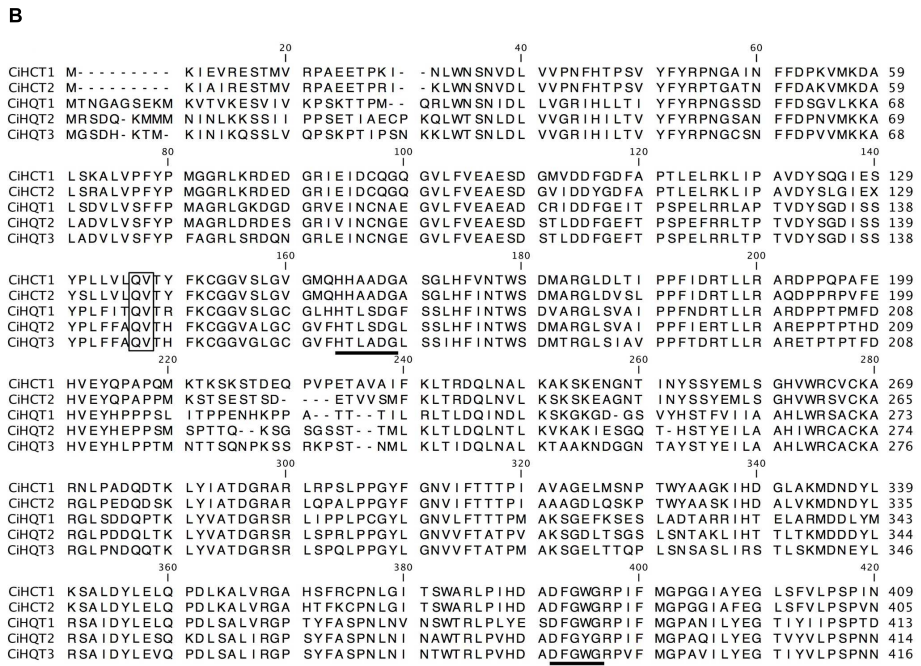
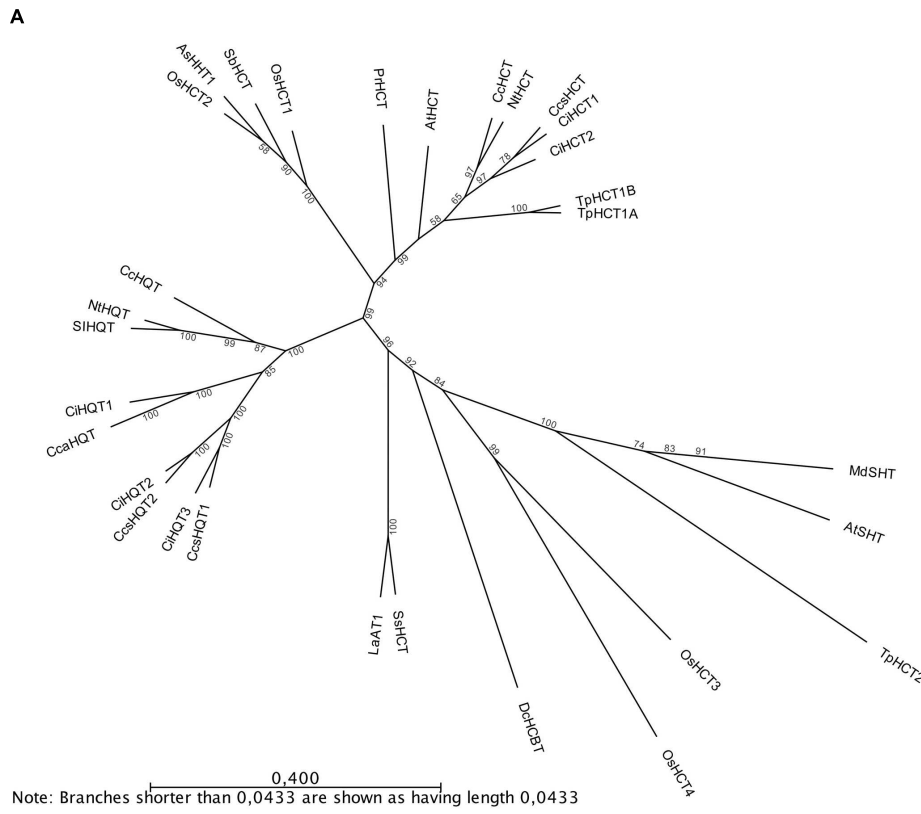
respectively. They harbor the features of the conserved intron "Q" found in many members of the BAHD family (St-Pierre and De Luca, 2000). The intron position in the coding region corresponds to an insertion between a Gln residue and a Val residue of the predicted amino acid sequences (Figure 5B).

## Expression of HCTs and HQTs in *E. coli* and Evaluation of Their Activity *In Vitro*

To study the catalytic activities of the chicory HCT and HQT proteins and to confirm the phylogenetic analysis, their entire coding regions were cloned into pDEST-17 vectors for expression in *E. coli*. Proteins were recovered in the soluble fractions and purified using His-Trap Ni columns. Presence, identity and purity of the recombinant proteins were then validated by SDS-PAGE and by immunoblot analysis using anti-His antibody (Figure 6). All proteins give a single band at the expected molecular mass, i.e., about 50 kDa (His<sub>6</sub>-HCTs or His<sub>6</sub>-HQTs). Activities of recombinant proteins were measured *in vitro* using either *p*-coumaroyl-CoA or caffeoyl-CoA as acyl donor and quinate or shikimate as the acceptor of the transferase reaction. The products of the reaction were analyzed by HPLC. In the presence of the enzymes, coumaroylquinate, coumaroylshikimate, caffeoylquinate, and caffeoylshikimate were detected in the reaction mixtures containing different combinations of substrates (not shown). No product could be detected in the absence of proteins or in the absence of the free acid. To verify their substrate preference, reaction mixtures containing 5 mM quinic acid plus 5 mM shikimic acid were run in the presence of 0.4 mM caffeoyl-CoA or 0.4 mM *p*-coumaroyl-CoA. When both acids were provided at the same concentration, HCTs only produce caffeoylshikimate or coumaroylshikimate whereas HQTs produce predominantly caffeoylquinate or coumaroylquinate (Figure 7). These results clearly demonstrate that HCTs from chicory strongly prefer shikimic over quinic acid whereas HQTs prefer quinic acid. These biochemical data support the phylogenetic tree and are in accordance with previously published reports dealing with HCTs or HQTs from other species. In additional experiments, we have shown that HCTs and HQTs can also use feruloyl-CoA and cinnamoyl-CoA as acyl donors but at much smaller rates (not shown). The activity of HCTs and HQTs toward other potential substrates (spermidine, tartaric acid) was also tested but no activity was observed in our experimental conditions.

These reactions were shown to be fully reversible in other systems (Hoffmann et al., 2003; Comino et al., 2007, 2009). In our experiments, when recombinant HCTs or HQTs were incubated in the presence of 3-CQA and CoA, caffeoyl-CoA was detected except for *HCT2* (not shown).

The optimal pHs of *HCT1* and *HQT1* were evaluated in the presence of caffeoyl-CoA and shikimate or quinate, respectively (Supplementary Figure S1). Optimum pHs are 7.5 for *HCT1* and 5.0 for *HQT1*. Activity of *HCT1* readily decreases at acidic or alkaline pHs. At pH 6.5 and 8.0, reaction rates are 61%. For *HQT1*, the activity slowly declines at alkaline pHs. At pH 7.0, activity is still 60% whereas at acidic pH, reaction rates rapidly decline (21% at pH 4.0).



**FIGURE 5 | Phylogenetic analysis and sequence comparison of the five members of the BAHD family identified in this study. (A)** Phylogenetic tree based on the amino acid sequences of HCTs and HQTs from chicory and biochemically characterized members of BAHD family members. Only the clade Vb is represented. The phenogram was generated by the neighbor-joining method following alignment by the MUSCLE algorithm (with default parameters) using CLC Sequence Viewer 7 software (bootstrap method, 500 replications). The length of the line indicates the relative distance between nodes thanks to the scale bar provided. **(B)** Amino acid sequence comparison of chicory HCTs and HQTs. Black bars indicate the conserved BAHD acyltransferase domains (HXXXDG and DFGWG). The intron (Continued)

**FIGURE 5 | Continued**

"Q" when present is located between the Q and V residue of the predicted amino acid sequence (black frame). CcsHCT: HCT from *Cynara cardunculus* var. *scolymus* (AFL93686), CiHCT1: this work (KT222891), CiHCT2: this work (KT222892), NiHCT: HCT from *Nicotiana tabacum* (CAD47830), CcHCT: HCT from *Coffea canephora* (CAJ40778), TpHCT1A: HCT from *T. pratense* (ACI16630), TpHCT1B: HCT from *T. pratense* (ACI28534), AtHCT: HCT from *Arabidopsis thaliana* (AED95744), PrHCT: HCT from *Pinus radiata* (ABO52899), SbHCT: HCT from *Sorghum bicolor* (4KE4\_A), OsHCT1: HCT from *Oryza sativa* (NP\_001053225.1), OsHCT2: HCT from *O. sativa* (NP\_001047408.1), AsHHT1: hydroxyanthranilate/hydroxycinnamoyl transferase from *Avena sativa* (BAC78633), CcsHQT1: HQT from *C. cardunculus* var. *scolymus* (CAM84302), CcsHQT2: HQT from *C. cardunculus* var. *scolymus* (CAR92145), CcaHQT: HQT from *C. canephora* (ABO77957), NiHQT: HQT from *N. tabacum* (CAE46932), SiHQT: HQT from *Solanum lycopersicum* (CAE46933), SsHCT: HCT from *Solenostemon scutellarioides* (CAK55166), LaAT1: alcohol acyltransferase 1 from *Lavandula angustifolia* (ABI48360), DcHCBT: anthranilate *N*-hydroxycinnamoyl/benzoyl transferase from *Dianthus caryophyllus* (CAB06430), TpHCT2: HCT from *Trifolium pratense* (ACI16631), OsHCT3: Hydroxycinnamoyl shikimate/glycerol transferase from *O. sativa* (NP\_001056998), OsHCT4: Hydroxycinnamoyl shikimate/glycerol transferase from *O. sativa* (NP\_001057003), AtSHT: spermidine *N*-hydroxycinnamoyl transferase from *A. thaliana* (AEC06845), MdSHT: spermidine *N*-hydroxycinnamoyl transferase from *Malus domestica* (ALF00095).

## Kinetic Parameters of Recombinant Proteins

Kinetic parameters of the enzyme were determined using caffeoyl-CoA or *p*-coumaroyl-CoA as the acyl-CoA donors and either quinate for HQTs or shikimate for HCTs as acyl acceptors (Table 1). The reactions showed typical Michaelis-Menten kinetics with increasing concentrations of acyl acceptor. The  $K_m$  values and  $V_{max}$  values were calculated from triplicates by the Lineweaver-Burk method. As shown in Table 1A, HCTs have better affinity for shikimate in the presence of saturating concentration of *p*-coumaroyl-CoA than in the presence of saturating concentration of caffeoyl-CoA (for HCT1,  $K_m = 320 \pm 40$  and  $8000 \pm 600 \mu\text{M}$ , respectively). On the contrary, HQT1 has better affinity for quinate in the presence of caffeoyl-CoA (Table 1B,  $K_m = 160 \pm 34 \mu\text{M}$  with caffeoyl-CoA and  $3800 \pm 413 \mu\text{M}$  with *p*-coumaroyl-CoA). HQT2 and HQT3 behave similarly and have about the same affinity for quinate in the presence of either caffeoyl-CoA or *p*-coumaroyl-CoA.

## Chicory HCTs and HQTs Are Differentially Expressed

The expression pattern of chicory HCTs and HQTs were examined using quantitative real-time PCR (qRT-PCR) in different parts of chicory seedlings shown to accumulate contrasted levels of phenolic compounds and in culture cells

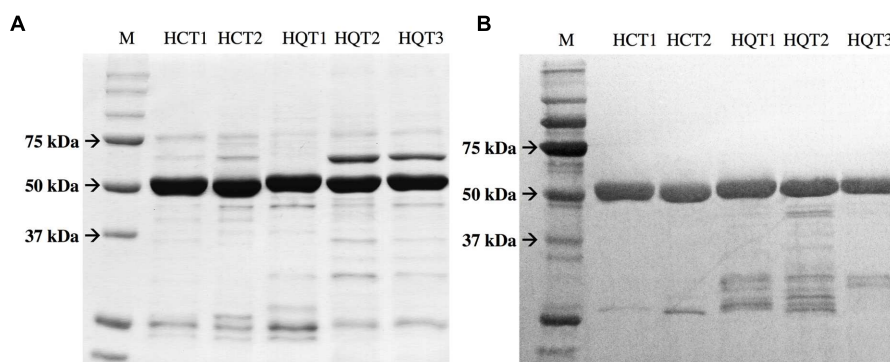
treated or not with MeJA. Primers were design to detect specifically HCT1, HCT2, HQT1, HQT2, or HQT3. The expression data of the target genes were normalized to CLATH and SAND mRNA levels for seedling experiments and to TIP41 and PP2AA2 mRNA levels for cell culture experiments in agreement with previous work done on the same plant material (Delpote et al., 2015).

The five genes were expressed in all of the investigated tissues (roots, old leaves, intermediate leaves, and young leaves; Figure 8A). Each gene shows a tissue-specific pattern of accumulation. Nevertheless no significant difference could be detected. HCT1 is the most expressed in roots whereas HCT2 and HQT1 expressions are the highest in the old leaves. HQT2 and HQT3 expressions are the highest in the intermediate leaves and in the youngest leaves, respectively.

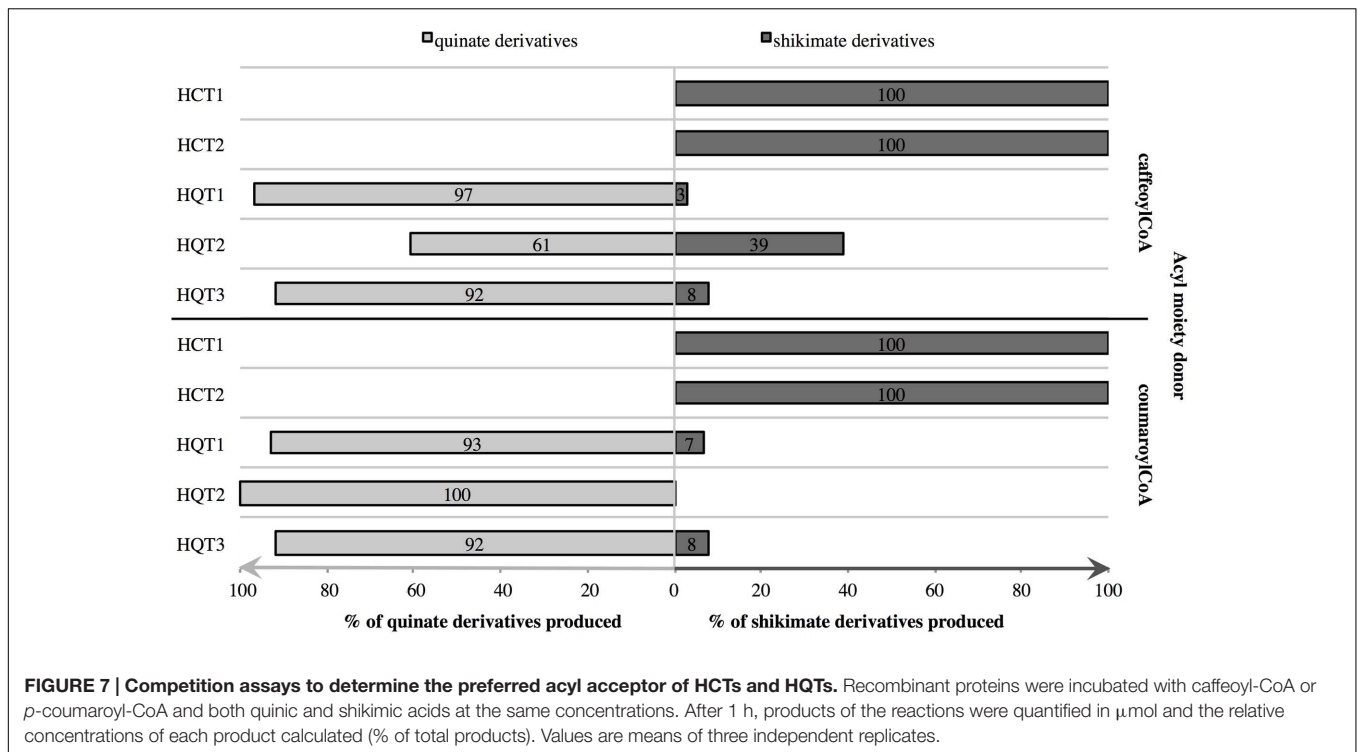
In cell culture, MeJA treatment up-regulates the mRNA abundance of HCT1, HCT2, and HQT3 (Figure 8B). For these three genes, mRNA abundance in MeJA treated cells was about 2-, 1.5-, and 3-fold higher respectively than in control cells. The expression of the two other acyltransferases (HQT1 and HQT2) was rather stable in all investigated conditions.

## Functional Analysis of HCT1 and HQT1

In order to assess the function of HCTs and HQTs *in planta*, HCT1 and HQT1 were transiently expressed in *N. benthamiana*, a species known to accumulate large amounts of CQA



**FIGURE 6 | Expression and purification of recombinant HCT and HQT proteins. (A)** SDS-PAGE separation of recombinant HCT and HQT. The gel was stained with Coomassie blue. **(B)** Immunoblot analysis confirming the expression of the recombinant proteins. The proteins were detected with anti-His antibody. M: Markers.



(Niggeweg et al., 2004). *N. benthamiana* plants were co-infiltrated with *A. tumefaciens* containing constructs of HCT1 (or HQT1) or the p19 gene. As control, leaves were infiltrated with the p19 vector only or the empty pB2GW7 vector (see Materials and Methods for the details). After 4 days, leaves were collected and the amounts of CQA measured.

3-CQA was the most abundant CQA isomer in leaves of tobacco. In leaves transiently expressing HQT1 or HCT1, levels of 3-CQA were higher than in leaves infiltrated with the empty vector or the p19-harboring plasmid (Figure 9). The increase of 3-CQA levels was more pronounced in HCT1 inoculated plants. Leaves infiltrated with the empty

**TABLE 1 | Kinetic parameters of HCTs and HQTs from chicory.**

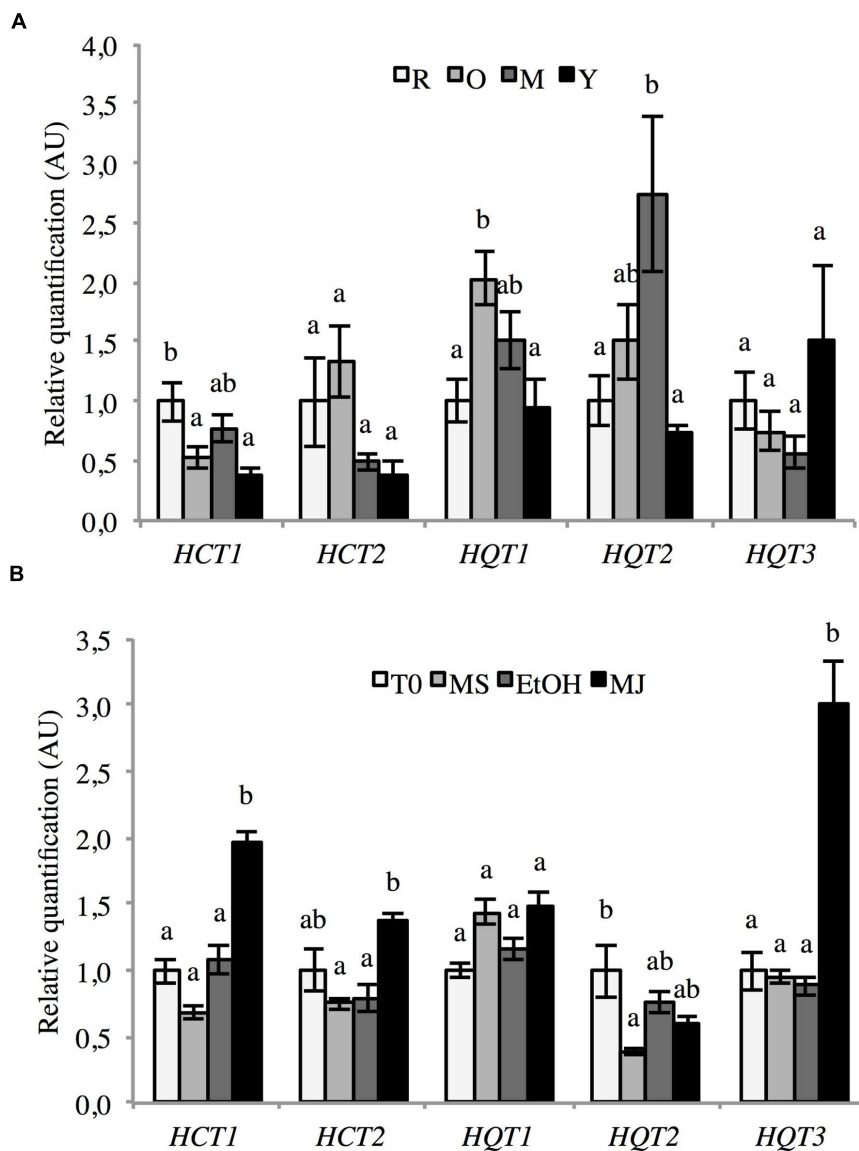
(A) HCT1 and HCT2.							
Varying substrate	Saturating substrate	HCT1			HCT2		
		$K_m \pm SD$ $\mu\text{M}$	$V_{max} \pm SD$ $\text{nkat mg}^{-1}$	$V_{max}/K_m$ $\text{nkat mg}^{-1}\mu\text{M}^{-1}$	$K_m \pm SD$ $\mu\text{M}$	$V_{max} \pm SD$ $\text{nkat mg}^{-1}$	$V_{max}/K_m$ $\text{nkat mg}^{-1}\mu\text{M}^{-1}$
Shikimate	<i>p</i> -coumaroyl-CoA	$320 \pm 40$	$237 \pm 9$	0.74	$802 \pm 63$	$303 \pm 9$	0.38
Shikimate	Caffeoyl-CoA	$8000 \pm 600$	$618 \pm 34$	0.08	$5080 \pm 360$	$695 \pm 31$	0.14

(B) HQT1, HQT2, and HQT3.										
Varying substrate	Saturating substrate	HQT1			HQT2			HQT3		
		$K_m \pm SD$ $\mu\text{M}$	$V_{max} \pm SD$ $\text{nkat mg}^{-1}$	$V_{max}/K_m$ $\text{nkat mg}^{-1}\mu\text{M}^{-1}$	$K_m \pm SD$ $\mu\text{M}$	$V_{max} \pm SD$ $\text{nkat mg}^{-1}$	$V_{max}/K_m$ $\text{nkat mg}^{-1}\mu\text{M}^{-1}$	$K_m \pm SD$ $\mu\text{M}$	$V_{max} \pm SD$ $\text{nkat mg}^{-1}$	$V_{max}/K_m$ $\text{nkat mg}^{-1}\mu\text{M}^{-1}$
Quinate	<i>p</i> -coumaroyl-CoA	$3800 \pm 413$	$903 \pm 49$	0.24	$637 \pm 65$	$149 \pm 6$	0.23	$416 \pm 240$	$74 \pm 11$	0.18
Quinate	Caffeoyl-CoA	$162 \pm 34$	$113 \pm 6$	0.70	$613 \pm 53$	$134 \pm 4$	0.22	$318 \pm 71$	$165 \pm 10$	0.52

The Michaelis constants ( $K_m$ ) were determined using various acid concentrations and saturating concentrations of acyl-CoA (800  $\mu\text{M}$ ). Data are summarized as the arithmetic mean  $\pm$  SD of three independent experiments.





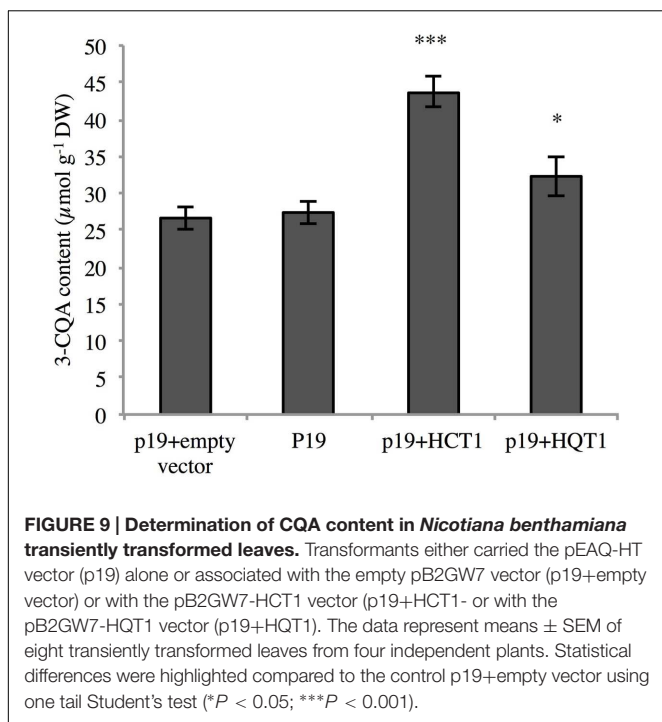
**FIGURE 8 | Expression of HCT1, HCT2, HQT1, HQT2, and HQT3 in various chicory tissues.** mRNA abundance in different parts of 6-weeks-old seedlings (A) and culture cells elicited or not with MeJA (B) was determined by qRT-PCR. (A) Expression levels were normalized to the expression of CLATH and SAND and presented relative to that of Roots. Y: young leaves, M: intermediate leaves, O: Old leaves, R: roots. (B) Expression levels were normalized to the expression of TIP41 and PP2AA2 and presented relative to that of T0. T0: cells sampled before elicitation, MS: cultures kept on MS medium for 24 h, EtOH: cell cultures treated with ethanol for 24 h, MJ: cell cultures treated with 50  $\mu$ M MeJA for 24 h. Error bars indicate SEM. Different letters above histograms indicate statistical difference highlighted using ANOVA ( $P$ -value < 0.05).

vector or the p19-harboring vector accumulate similar levels of 3-CQA.

## DISCUSSION

Hydroxycinnamoyl esters play essential roles in plant physiology especially in plant-environment interactions (Sullivan, 2014). Furthermore, these compounds have significance in the area of human health. They are present in fruits and vegetables and

they represent important antioxidant molecules with multiple applications for cosmetic, pharmaceutical and food industries. Industrial chicory accumulates four main caffeic esters, i.e., 3-CQA, 3,5-diCQA, CTA, and diCTA (this work; Kandeler and Ullrich, 2009; Bahri et al., 2012). An additional CQA isomer, i.e., 5-CQA known as neochlorogenic acid was also detected in root tissues (this work). Its content was quite low in seedling roots (0.0035% DW, i.e., 37 times lower than 3-CQA content in the same tissue) but was significantly higher in root cell cultures (from 0.088 to 0.156% DW in our set



of experiments). These molecules exhibit an original tissue distribution pattern. Tartaric acid derivatives are predominantly accumulated in aerial parts and diCQA is mainly located in roots. The contents of 3-CQA is equally distributed between organs whereas 5-CQA, although at low level, is exclusively located in roots. We presume that these specific localizations must rely on specific tissue functions that largely remain to be elucidated. In addition, detailed analysis of aerial parts revealed that contents of all phenolic acids were related to leaf age. Young leaves accumulate the highest levels of CTA, diCTA, 3-CQA, and diCQA whereas oldest leaves accumulate the lowest levels of these compounds. Total level of molecules reaches  $56 \mu\text{mol g}^{-1}$  DW in the youngest expanding leaves whereas levels were 26 and  $16 \mu\text{mol g}^{-1}$  DW in intermediate leaves and in oldest leaves, respectively. This could suggest a role of these compounds in biotic or abiotic stress protection, these tissues being the most valuable parts of the plants. This is in favor of the so-called optimal defense theory. This theory suggests that plants accumulate more constitutive defense metabolites in tissues that are crucial in the fitness of the plant (McKey, 1974; Rhoades, 1979). Notably CQA was shown to have a role in plant protection against thrips in *Chrysanthemum* (Leiss et al., 2009). This theory was already suggested to explain the original tissue pattern accumulation of phenolamides (phenylpropanoid-polyamine conjugates) in *Nicotiana attenuata* (Kaur et al., 2010). These phenylpropanoid derivatives were preferentially allocated to the young leaves presumably to prevent pathogen or herbivore attacks. Whether, this original pattern of accumulation in chicory seedlings relies on more active synthesis of these molecules in young expanding leaves, more active catabolism of these molecules in oldest leaves or more efficient

transport of phenolics to the growing tissues still remain to be determined.

To take advantage of this original chemical composition (traditional breeding, synthetic biology, genetic engineering), a full understanding of the relevant biosynthetic pathways is required. Evidences in other species prompted us to focus our analysis on the clade Vb of the BAHD superfamily of plant-specific acyl-CoA dependent acyltransferases (St-Pierre and De Luca, 2000). This clade was shown to contain proteins involved in the synthesis of CQA and related compounds as well as more diverse products like phaselic acid, triferuloyl spermidine and feruloyl glycerol (Grienberger et al., 2009; Sullivan and Zarnowski, 2011; Kim et al., 2012; Elejalde-Palmett et al., 2015). The full-length open reading frames of five candidate genes were cloned. On the basis of their high similarity with the already functionally characterized hydroxycinnamoyl-transferase sequences (HCT or HQT) and of the phylogenetic analysis, they were named *HCT1*, *HCT2*, *HQT1*, *HQT2*, and *HQT3*. They share typical features of the members of the BAHD family (D'Auria, 2006). Phylogenetic clustering can give a clue of the function. Nevertheless function need to be established on a biochemical demonstration of activity. Clade Vb of the BAHD contain members with very versatile catalytic specificities (Grienberger et al., 2009; Sullivan and Zarnowski, 2011; Kim et al., 2012; Elejalde-Palmett et al., 2015). All five proteins were shown to be able to use either *p*-coumaroyl-CoA or caffeoyl-CoA as an acyl donor and quinic acid or shikimic acid as an acyl acceptor. Substrate specificities and associated kinetic parameters are consistent with those of HCTs or HQTs isolated and cloned in other species (Hoffmann et al., 2003; Niggeweg et al., 2004; Sonnante et al., 2010). Abilities to catalyze the reverse reactions were also examined and confirmed for four of the five enzymes as found in other species.

Phenylpropanoid contents usually rise when the plant is submitted to MeJA elicitation (Gális et al., 2006). We found out that treatment of chicory cell cultures with MeJA leads to an increase of 3-CQA and 3,5-diCQA amounts. Furthermore, in contrast to root or leaf extract where only one major isomer of CQA (i.e., 3-CQA) could be detected, in suspension culture, 5-CQA was detected in equivalent quantity to 3-CQA. We observed that treatment of cell cultures with MeJA induces the expression of the genes encoding HQT3 and the two HCTs with a stronger induction for *HCT2* and even more for *HQT3*. This suggests the major involvement of HQT3 in the higher production and accumulation of 3-CQA in response to MeJA treatment. No relationship between 3-CQA or 3,5-diCQA accumulation in young leaves (compare to the oldest leaves) and HCT or HQT expression was shown.

The involvement of HQTs and HCTs in 3-CQA production was further confirmed by *in vivo* functional analysis. *HCT1* and *HQT1* were transiently overexpressed in tobacco. The results clearly demonstrate that *HCT1* and *HQT1* are involved in the synthesis of CQA. This is in accordance with previous studies dealing with HQT from artichoke, tobacco, or tomato overexpressed transiently or stably in tobacco (Niggeweg et al., 2004; Sonnante et al., 2010). Surprisingly, *HCT1* transient expression promoted the accumulation of 3-CQA at higher

level than that of HQT1. Considering substrate affinity of these two enzymes *in vitro*, we are prone to suggest that *in planta* pathway for 3-CQA production toward its accumulation occurs in two steps. First caffeoyl-CoA is synthesized through the successive action of HCT, C3'H and either HCT or HQT (reverse reaction). Subsequently, HQT uses the CoA ester to produce 3-CQA. This assumption is also supported by the transcriptional analysis. Indeed 3-CQA accumulation promoted by MeJA induction is concomitant to an increase of the expression of HCTs and HQT3. If HQT was the only required enzymatic step, one could expect the induction of the sole HQT expression. This route for the synthesis of CQA was already favored in artichoke (Sonnante et al., 2010). Docking and modeling experiments showed that HQTs preferred quinate as a substrate whereas HCTs preferred shikimate. Better efficiency of artichoke C3'H in the conversion of coumaroylshikimate into caffeoylshikimate also favored this hypothesis (Moglia et al., 2009). Preference of HCTs for shikimate and coumaroyl-CoA was also demonstrated by structural analysis in *Sorghum bicolor* (Walker et al., 2013). SbhCT activity was proven to be quite restricted to the production of coumaroylshikimate. All together these data infer that esterification by HCT should be the limiting catalytic step in CQA synthesis. Flux analysis should confirm this assumption. Alternatively, CQA synthesis promoted by HQT could be limited by substrate availability due to HQT specific subcellular localization. Optimal pH of this enzyme was shown to be around 5. This could fit with a vacuolar compartmentalization of this enzyme. Such localization of HQT was already shown in *S. lycopersicum* (Moglia et al., 2014). In this context, HQT activity rate could be limited by CoA-ester or quinic acid import into vacuole.

The presence of several homologous genes might relate various independent separations of them during plant evolution. Therefore, in chicory, HQT enzymes and HCT enzymes seem to be encoded by a gene family of at least three members and two members, respectively. This seems to be a general trend in the *Asteraceae* family since three HQTs and one HCT were found in artichoke and several members of each family are present in sunflower and in lettuce (Sonnante et al., 2010). The presence of multiple isoforms of each family might ensure an optimization of fluxes toward the accumulation of different metabolites. Especially, in chicory, in addition to CQA accumulation, CQA might serve as a caffeoyl donor for the production of diCQA, CTA, and diCTA. Further experiments are needed to decipher the physiological roles of these enzymes. Tissue and subcellular localization

of the products of these genes could provide clues of their function.

This work is the first contribution in the understanding of the genetic basis of the hydroxycinnamate biosynthesis in *C. intybus*. Future research will be devoted to extend the investigation on CQA accumulation but also to the characterization of the CTA and diCTA biochemical pathway as well as that of the diCQA, which remains incomplete at least in the *Asteraceae*.

## AUTHOR CONTRIBUTIONS

GL, MD, CK, MM, and DG carried out the molecular genetic studies. AH, CV, and PH set up the cell culture system and performed preliminary experiments. GL, MD, CK, and DG carried out the biochemical experiments. DG and J-LH planned and supervised the work. DG wrote the article with contributions of all the authors. All authors read and approved the final manuscript.

## FUNDING

The present work was financed by a public-private partnership “Groupement d’Intérêt Scientifique” GENOCHIC (EA 7394 Agro-food and Biotechnology Research Institute, Charles Viollette Research Institute, University Lille 1 – Florimond-Desprez Veuve et fils SAS – Leroux SAS). Fundings were also provided by BPI France and the Région Nord Pas-de-Calais to QUALICHIC and GLYCACHIC projects. GL and MD were supported by a doctoral fellowship from the doctoral school 104 SMRE.

## ACKNOWLEDGMENTS

The authors thank Sylvain Legrand for his helpful advice on bioinformatics and gratefully acknowledge Najia Voedts for cell culture maintenance. Authors would like to thank Professor George Lomonosoff (John Innes Center) for providing the pEAQ-HT expression vector.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.00741>

## REFERENCES

- Bahri, M., Hance, P., Grec, S., Quillet, M.-C., Trotin, F., Hilbert, J.-L., et al. (2012). A “Novel” protocol for the analysis of hydroxycinnamic acids in leaf tissue of chicory (*Cichorium intybus* L., Asteraceae). *Sci. World J.* 2012:142983. doi: 10.1100/2012/142983
- Beuerle, T., and Pichersky, E. (2002). Enzymatic synthesis and purification of aromatic coenzyme a esters. *Anal. Biochem.* 302, 305–312. doi: 10.1006/abio.2001.5574
- Comino, C., Hehn, A., Moglia, A., Menin, B., Bourgaud, F., Lanteri, S., et al. (2009). The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway. *BMC Plant Biol.* 9:30. doi: 10.1186/1471-2229-9-30
- Comino, C., Lanteri, S., Portis, E., Acquadro, A., Romani, A., Hehn, A., et al. (2007). Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. *BMC Plant Biol.* 7:14. doi: 10.1186/1471-2229-7-14
- D’Auria, J. C. (2006). Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* 9, 331–340. doi: 10.1016/j.pbi.2006.03.016
- De Geyter, N., Gholami, A., Goormachtig, S., and Goossens, A. (2012). Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends Plant Sci.* 17, 349–359. doi: 10.1016/j.tplants.2012.03.001

- Del-Castillo-Alonso, M. -Á., Diago, M. P., Monforte, L., Tardaguila, J., Martínez-Abaigar, J., and Núñez-Olivera, E. (2014). Effects of UV exclusion on the physiology and phenolic composition of leaves and berries of *Vitis vinifera* cv. *Graciano*. *J. Sci. Food Agric.* 95, 409–416. doi: 10.1002/jsfa.6738
- Delporte, M., Legrand, G., Hilbert, J.-L., and Gagneul, D. (2015). Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. *Front. Plant Sci.* 6:651. doi: 10.3389/fpls.2015.00651
- Elejalde-Palmett, C., Dugé De Bernonville, T., Glevarec, G., Pichon, O., Papon, N., Courdavault, V., et al. (2015). Characterization of a spermidine hydroxycinnamoyltransferase in *Malus domestica* highlights the evolutionary conservation of trihydroxycinnamoyl spermidines in pollen coat of core Eudicotyledons. *J. Exp. Bot.* 66, 7271–7285. doi: 10.1093/jxb/erv423
- Gális, I., Simek, P., Narisawa, T., Sasaki, M., Horiguchi, T., Fukuda, H., et al. (2006). A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. *Plant J.* 46, 573–592. doi: 10.1111/j.1365-313X.2006.02719.x
- Gonthier, L., Bellec, A., Blassiau, C., Prat, E., Helmstetter, N., Rambaud, C., et al. (2010). Construction and characterization of two BAC libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L., Asteraceae). *BMC Res. Notes* 3:225. doi: 10.1186/1756-0500-3-225
- Grienenberger, E., Besseau, S., Geoffroy, P., Debayle, D., Heintz, D., Lapierre, C., et al. (2009). A BAHd acyltransferase is expressed in the tapetum of *Arabidopsis* anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant J.* 58, 246–259. doi: 10.1111/j.1365-313X.2008.03773.x
- Hodgins, K. A., Lai, Z., Oliveira, L. O., Still, D. W., Scascitelli, M., Barker, M. S., et al. (2014). Genomics of compositae crops: reference transcriptome assemblies and evidence of hybridization with wild relatives. *Mol. Ecol. Resour.* 14, 166–177. doi: 10.1111/1755-0998.12163
- Hoffmann, L., Maury, S., Martz, F., Geoffroy, P., and Legrand, M. (2003). Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* 278, 95–103. doi: 10.1074/jbc.M209362200
- Hohlfeld, M., Veit, M., and Strack, D. (1996). Hydroxycinnamoyltransferases involved in the accumulation of caffeic acid esters in gametophytes and sporophytes of *Equisetum arvense*. *Plant Physiol.* 111, 1153–1159. doi: 10.1104/pp.111.4.1153
- Kandeler, R., and Ullrich, W. R. (2009). Symbolism of plants: examples from European-Mediterranean culture presented with biology and history of art. *J. Exp. Bot.* 60, 3973–3974. doi: 10.1093/jxb/erp248
- Karimi, M., Inzé, D., and Depicker, A. (2002). Gateway vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* 7, 193–195. doi: 10.1016/S1360-1385(02)02251-3
- Kaur, H., Heinzel, N., Schöttner, M., Baldwin, I. T., and Gális, I. (2010). R2R3-NaMYB8 regulates the accumulation of phenylpropanoid-polyamine conjugates, which are essential for local and systemic defense against insect herbivores in *Nicotiana attenuata*. *Plant Physiol.* 152, 1731–1747. doi: 10.1104/pp.109.151738
- Kim, I. A., Kim, B.-G., Kim, M., and Ahn, J.-H. (2012). Characterization of hydroxycinnamoyltransferase from rice and its application for biological synthesis of hydroxycinnamoyl glycerols. *Phytochemistry* 76, 25–31. doi: 10.1016/j.phytochem.2011.12.015
- Kojima, M., and Kondo, T. (1985). An enzyme in sweet-potato root which catalyzes the conversion of chlorogenic acid, 3-caffeoylquinic acid, to isochlorogenic acid, 3,5-dicafeoylquinic acid. *Agric. Biol. Chem.* 49, 2467–2469. doi: 10.1271/abb1961.49.2467
- Korim, K. M. M., and Soliman, R. E. (2014). Chlorogenic and caftaric acids in liver toxicity and oxidative stress induced by methamphetamine. *J. Toxicol.* 2014:583494. doi: 10.1155/2014/583494
- Lallemand, L. A., Zubietta, C., Lee, S. G., Wang, Y., Acajao, S., Timmins, J., et al. (2012). A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. *Plant Physiol.* 160, 249–260. doi: 10.1104/pp.112.202051
- Leiss, K. A., Maltese, F., Choi, Y. H., Verpoorte, R., and Klinkhamer, P. G. L. (2009). Identification of chlorogenic acid as a resistance factor for thrips in chrysanthemum. *Plant Physiol.* 150, 1567–1575. doi: 10.1104/pp.109.138131
- Mai, F., and Glomb, M. A. (2013). Isolation of phenolic compounds from iceberg lettuce and impact on enzymatic browning. *J. Agric. Food Chem.* 61, 2868–2874. doi: 10.1021/jf305182u
- McKey, D. (1974). Adaptive patterns in alkaloid physiology. *Am. Nat.* 108, 305–320. doi: 10.1086/282909
- Moglia, A., Comino, C., Portis, E., Acquadro, A., De Vos, R. C. H., Beekwilder, J., et al. (2009). Isolation and mapping of a C3H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress. *Plant Cell Rep.* 28, 963–974. doi: 10.1007/s00299-009-0695-1
- Moglia, A., Lanteri, S., Comino, C., Acquadro, A., de Vos, R., and Beekwilder, J. (2008). Stress-induced biosynthesis of dicafeoylquinic acids in globe artichoke. *J. Agric. Food Chem.* 56, 8641–8649. doi: 10.1021/jf801653w
- Moglia, A., Lanteri, S., Comino, C., Hill, L., Knevič, D., Cagliero, C., et al. (2014). Dual catalytic activity of hydroxycinnamoyl-CoA quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicafeoylquinic acids. *Plant Physiol.* 166, 1777–1787. doi: 10.1104/pp.114.251371
- Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x
- Niggeweg, R., Michael, A. J., and Martin, C. (2004). Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat. Biotechnol.* 22, 746–754. doi: 10.1038/nbt966
- Perry, N. B., Burgess, E. J., and Glennie, V. L. (2001). Echinacea standardization: analytical methods for phenolic compounds and typical levels in medicinal species. *J. Agric. Food Chem.* 49, 1702–1706. doi: 10.1021/jf001331y
- Petersen, M., Abdullah, Y., Benner, J., Eberle, D., Gehlen, K., Hüchering, S., et al. (2009). Evolution of rosmarinic acid biosynthesis. *Phytochemistry* 70, 1663–1679. doi: 10.1016/j.phytochem.2009.05.010
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, 2002–2007. doi: 10.1093/nar/29.9.e45
- R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. R Found. Vienna: Statistics Computing.
- Rautengarten, C., Baidoo, E., Keasling, J. D., and Scheller, H. V. (2010). A simple method for enzymatic synthesis of unlabeled and radiolabeled hydroxycinnamate-CoA. *Bioenergy Res.* 3, 115–122. doi: 10.1007/s12155-010-9085-3
- Rhoades, D. (1979). “Evolution of plant chemical defense against herbivores,” in *Herbivores: Their Interactions with Secondary Plant Metabolites*, eds G. A. Rosenthal and D. H. Janzen (Orlando, FL: Academic Press), 4–55.
- Sainsbury, F., Thuenemann, E. C., and Lomonosoff, G. P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* 7, 682–693. doi: 10.1111/j.1467-7652.2009.00434.x
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 3rd Edn. New York, NY: Cold Spring Harbor Laboratory Press.
- Singleton, V. L., Zaya, J., and Trousdale, E. K. (1986). Caftaric and coutaric acids in fruit of *Vitis*. *Phytochemistry* 25, 2127–2133. doi: 10.1016/0031-9422(86)80078-4
- Sonnante, G., D’Amore, R., Blanco, E., Pierri, C. L., De Palma, M., Luo, J., et al. (2010). Novel hydroxycinnamoyl-coenzyme a quinate transferase genes from artichoke are involved in the synthesis of chlorogenic acid. *Plant Physiol.* 153, 1224–1238. doi: 10.1104/pp.109.150144
- St-Pierre, B., and De Luca, V. (2000). “Evolution of acyltransferase genes: Origin and diversification of the BAHd superfamily of acyltransferases involved in secondary metabolism,” in *Recent Advances in Phytochemistry*, eds J. Romeo, R. Ibrahi, L. Varin, and V. De Luca (Amsterdam: Elsevier Science Ltd), 285–315. doi: 10.1016/S0079-9920(00)80010-6
- Sullivan, M. (2009). A novel red clover hydroxycinnamoyl transferase has enzymatic activities consistent with a role in phaselic acid biosynthesis. *Plant Physiol.* 150, 1866–1879. doi: 10.1104/pp.109.136689
- Sullivan, M. L. (2014). Perennial peanut (*Arachis glabrata* Benth.) leaves contain hydroxycinnamoyl-CoA:tartaric acid hydroxycinnamoyl transferase activity and accumulate hydroxycinnamoyl-tartaric acid esters. *Planta* 239, 1091–1100. doi: 10.1007/s00425-014-2038-x
- Sullivan, M. L., and Foster, J. L. (2013). Perennial peanut (*Arachis glabrata* Benth.) contains polyphenol oxidase (PPO) and PPO substrates that can reduce post-harvest proteolysis. *J. Sci. Food Agric.* 93, 2421–2428. doi: 10.1002/jsfa.6050
- Sullivan, M. L., and Zarnowski, R. (2011). Red Clover HCT2, a hydroxycinnamoyl-coenzyme a: malate hydroxycinnamoyl transferase, plays a crucial role in

- biosynthesis of phasic acid and other hydroxycinnamoyl-malate esters in vivo. *Plant Physiol.* 155, 1060–1067. doi: 10.1104/pp.110.166793
- Tousch, D., Lajoix, A.-D., Hosi, E., Azay-Milhau, J., Ferrare, K., Jahannault, C., et al. (2008). Chicoric acid, a new compound able to enhance insulin release and glucose uptake. *Biochem. Biophys. Res. Commun.* 377, 131–135. doi: 10.1016/j.bbrc.2008.09.088
- Tuominen, L. K., Johnson, V. E., and Tsai, C.-J. (2011). Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among *Populus* paralogues. *BMC Genomics* 12:236. doi: 10.1186/1471-2164-12-236
- Udvardi, M. K., Czechowski, T., and Scheible, W.-R. (2008). Eleven golden rules of quantitative RT-PCR. *Plant Cell* 20, 1736–1737. doi: 10.1105/tpc.108.061143
- Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., et al. (2013). Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* 341, 1103–1106. doi: 10.1017/CBO9781107415324.004
- Vanholme, R., Ralph, J., Akiyama, T., Lu, F., Pazo, J. R., Kim, H., et al. (2010). Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in *Arabidopsis*. *Plant J.* 64, 885–897. doi: 10.1111/j.1365-3113X.2010.04353.x
- Villegas, R. J. A., Shimokawa, T., Okuyama, H., and Kojima, M. (1987). Purification and characterization of chlorogenic acid: chlorogenate caffeoyl transferase in sweet potato roots. *Phytochemistry* 26, 1577–1581. doi: 10.1016/S0031-9422(00)82248-7
- Walker, A. M., Hayes, R. P., Youn, B., Vermerris, W., Sattler, S. E., and Kang, C. (2013). Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme a-dependent transferases and synthases. *Plant Physiol.* 162, 640–651. doi: 10.1104/pp.113.217836
- Wasternack, C., and Hause, B. (2013). Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in *Annals of Botany*. *Ann. Bot.* 111, 1021–1058. doi: 10.1093/aob/mct067
- Yasir, M., Sultana, B., Nigam, P. S., and Owusu-Apenten, R. (2016). Antioxidant and genoprotective activity of selected cucurbitaceae seed extracts and LC-ESIMS/MS identification of phenolic components. *Food Chem.* 199, 307–313. doi: 10.1016/j.foodchem.2015.11.138
- Yu, X.-H. H., Gou, J.-Y. Y., and Liu, C.-J. J. (2009). BAHD superfamily of acyl-CoA dependent acyltransferases in *Populus* and *Arabidopsis*: bioinformatics and gene expression. *Plant Mol. Biol.* 70, 421–442. doi: 10.1007/s11103-009-9482-1

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Legrand, Delporte, Khelifi, Harant, Vuylsteker, Mörchen, Hance, Hilbert and Gagneul. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# A Reassessment of the Function of the So-Called Compatible Solutes in the Halophytic Plumbaginaceae *Limonium latifolium*<sup>1[W][OA]</sup>

David Gagneul<sup>2</sup>, Abdelkader Aïnouche, Claire Duhazé, Raphaël Lugan, François Robert Larher, and Alain Bouchereau\*

Unité Mixte de Recherche 6026 Interactions Cellulaires et Moléculaires (D.G., C.D., R.L., F.R.L., A.B.) and Unité Mixte de Recherche 6553 Ecosystèmes, Biodiversité, Evolution (A.A.), Université de Rennes 1, Centre National de la Recherche Scientifique, Campus de Beaulieu, 35042 Rennes cedex, France

The compatible solute hypothesis posits that maintaining osmotic equilibrium under conditions of high salinity requires synthesis of organic compounds, uptake of potassium ions, and partial exclusion of NaCl. To assess whether osmotic adaptation in *Limonium latifolium* proceeds according to this hypothesis, a comprehensive analysis of solute accumulation during NaCl treatments was conducted. Determination of prevailing inorganic ions and establishment of the metabolic profiles for low  $M_r$  organic substances revealed that contrary to the mentioned hypothesis the major contributors to osmolarity were constituted by inorganic solutes. Independent of salinity, only 25% of this osmolarity resulted from organic solutes such as Suc and hexoses. Proline (Pro),  $\beta$ -alanine betaine, and choline-*O*-sulfate were minor contributors to osmolarity. Compatible inositols also occurred, especially *chiro*-inositol, characterized for the first time in this species, to our knowledge. Principal component analysis showed that only a limited number of metabolic reconfigurations occurred in response to dynamic changes in salinity. Under such conditions only sugars, *chiro*-inositol, and Pro behave as active osmobalancers. Analysis of metabolic profiles during acclimatization to either mild salinity or nonsaline conditions showed that organic solute accumulation is predominantly controlled by constitutive developmental programs, some of which might be slightly modulated by salinity. Osmolarity provided under such conditions can be sufficient to maintain turgor in salinized seedlings. Compartmental analysis of Pro and  $\beta$ -alanine betaine in leaf tissues demonstrated that these solutes, mainly located in vacuoles under nonsaline conditions, could be partly directed to the cytosol in response to salinization. Thus they did not conform with the predictions of the compatible solute hypothesis.

Due to either water loss or induced processes responsible for enhancement of the total number of osmotically active particles, higher plants are able to increase their osmolarity in response to hyperosmotic conditions. Accumulation of such particles, collectively termed osmolytes, contributes to osmotic adjustment (OA), which is needed for survival and resuming of growth at after-stress recovery. Active OA has been recognized as one of the key determinants of tolerance to salinity and other osmotic stresses encountered by higher plants (Jamaux et al., 1997). Inorganic and/or

organic solutes could be involved in such function and depending on their origin, their production is associated with different energetic requirements partly responsible for inhibitory effects on growth rate. It is widely believed that osmolytes do not exert, per se, damaging effects on cell compartments where they are located.

Primarily OA could be achieved with ions such as  $K^+$ ,  $Na^+$ ,  $NO_3^-$ , or  $Cl^-$  when available in the root environment (Shabala and Lew, 2002) through stress-induced activation of uptake and translocation of these substances. At the cellular level further accumulation of ions in the vacuole is required to prevent the deleterious effects that could result from enhancement of ionic strength in the cytoplasm. Despite the energy requirements for ion uptake and further compartmentation in the vacuole (Yamaguchi and Blumwald, 2005), such active OA is considered as relatively inexpensive because carbon skeletons are not needed for formulation of the osmotic complement. Alternatively, OA could result from accumulation of low  $M_r$  organic compounds belonging, according to their chemical structure, to a restricted number of classes like saccharides, polyhydroxyalcohols, organic acids, amino acids, betaines, and tertiary sulfonium substances. These compounds are diverted from primary metabolic pathways either

<sup>1</sup> This work was supported by the French Ministry of National Education and Technological Research (to D.G.).

<sup>2</sup> Present address: Department of Plant Biology, Michigan State University, East Lansing, MI 48824.

\* Corresponding author; e-mail alain.bouchereau@univ-rennes1.fr; fax 33-02-23-23-69-15.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors ([www.plantphysiol.org](http://www.plantphysiol.org)) is: Alain Bouchereau (alain.bouchereau@univ-rennes1.fr).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

[www.plantphysiol.org/cgi/doi/10.1104/pp.107.099820](http://www.plantphysiol.org/cgi/doi/10.1104/pp.107.099820)

directly or through short specific pathways (Rhodes et al., 2002) that lead to the conversion of precursor(s) to metabolically inactive substances. This type of OA is considered as expensive.

Some of the organic osmolytes that do not disrupt proper functioning of organelles and are assumed to be preferentially located in nonvacuolar compartments of plant cells have also been termed compatible solutes (Brown and Simpson, 1972; Yancey et al., 1982; Bohnert and Jensen, 1996). Organic acids and charged amino acids that could induce damages to cellular components are not regarded as compatible. Hexoses that exert direct and indirect effects on carbon metabolism should also be excluded (Smeeckens and Rook, 1997; Lalonde et al., 1999; Rolland et al., 2002). Based on the results of *in vitro* experiments, compatible solutes have also been assumed to be involved in osmoprotection of functional components of plant cells subjected to increased osmolarity through their stabilizing or chaperoning properties (Yancey et al., 1982; Diamant et al., 2001). However, direct evidences for such functions are still lacking despite attempts to improve growth under saline conditions by genetic engineering of metabolic pathways involved in metabolism of accumulated substances. Such elegant solutions to improve crop performances have actually been marginally beneficial and this might be due to metabolic constraints that limit overproduction of compatible solutes in transgenic plants (Nuccio et al., 1999). It might also result from their noncompatibility in transgenic types. Compatibility of these substances seems indeed to be restricted to the wild genotypes that naturally produce them (Gibon et al., 1997; Romero et al., 1997; Sulpice et al., 1998, 2002; Bohnert and Shen, 1999; Hare et al., 2002).

This model for intracellular compartmentation of inorganic ions and other osmolytes in salinized tissues of higher plants has emerged from the pioneering works of Flowers (1972), Stewart and Lee (1974), Storey and Wyn Jones (1975), and Wyn Jones et al. (1977). These authors anticipated that in plants coping with salinity the distribution of osmolytes between cell compartments was expected to provide osmotic balance between the cytosol and the vacuole on the one hand and between the cytosol and the apoplast on the other hand. Some of the premises of the compatible solute theory that had been coined by Brown and Simpson (1972) to explain the function of glycerol accumulated by yeasts (*Saccharomyces cerevisiae*) in response to osmotic stress, have been put to the test with higher plants subjected to saline conditions. These approaches have shown that some typical compatible solutes and ions such as  $\text{Na}^+$  and  $\text{Cl}^-$  were indeed preferentially compartmentalized, respectively, in the extracellular compartments and the vacuoles. However, according to Wyn Jones and Gorham (2002), these statements should now be regarded with caution since great plasticity occurs at both the tissue and the subcellular levels. These authors suggested that vacuoles are not inert balloons, some of them being filled with inorganic ions

while others might contain large amounts of organic osmolytes. The prevailing model for allocation of osmolytes at the subcellular level might also be modulated by the degree of vacuolation of plant cells that depends on cell expansion and changes induced by salinity to the volume fraction of the vacuoles (Chang et al., 1996).

The biochemical diversity of organic osmolytes accumulated by the salt-excreting halophyte *Plumbaginaceae* (Larher and Hamelin, 1975; Rhodes and Hanson, 1993) is rarely met in other halophytes. It has attracted our attention since it raises questions about the biological significance of these substances in terms of evolutionary adaptation to salinity in comparison to apparently more simple status occurring in other halophytic families (Tipirdamaz et al., 2006). The first purpose of this metabolic study, based on efficient analytical methods, was to extend the characterization of sugars, polyols, and organic acids that could also be involved in OA of *Limonium latifolium*. Their contribution to osmolarity has been compared to that of nitrogenous solutes previously characterized (Hanson et al., 1994; Bouchereau et al., 1999). Second, we have tried to discriminate between metabolites that are constitutively accumulated from those that behave as stress metabolites, their amount being modulated by changes in salinity. Finally, we have assessed the validity of some of the premises of the compatible solute hypothesis in terms of compartmentation of these substances as well as its apparent plasticity in response to salinity, using the Pro and the  $\beta$ -Ala betaine (AB) cases as examples.

## RESULTS

### Suc and Inositol Isomers as Major Organic Osmotica in Shoots and Roots of *L. latifolium*

The amounts of soluble low  $M_r$  organic compounds and those of  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cl}^-$ , and  $\text{NO}_3^-$  were first determined in seedlings grown for 10 d in the presence of either 300 mM NaCl or the reference medium. As anticipated from relevant signals in the  $^1\text{H-NMR}$  spectrum performed on a crude extract obtained from shoots of control seedlings (Supplemental Fig. S1), nitrogenous substances known to accumulate in the *Plumbaginaceae* were also detected in this species. However, a number of other prominent signals were contained in this spectrum, suggesting the presence of other abundant organic solutes.

The amounts of major components and that of some of their precursors are shown in Table I. Gln was the most abundant nitrogenous solute in shoots and roots of control seedlings, whereas Pro became quantitatively prominent in those treated with NaCl. The betaines choline-O-sulfate (COS) and  $\beta$ -AB were abundant in shoots and roots and this was independent of growth conditions. Gly betaine (GB) was not detectable. Citric and malic acids were abundant in nonsalinized seedlings. The most abundant organic solutes

**Table 1.** The main organic and inorganic solutes occurring in *L. latifolium*

Three-month-old seedlings were treated or not for 10 d with 300 mM NaCl. Free amino acids, quaternary ammonium compounds (QACs), nonstructural carbohydrates (NSCs), cyclitols, organic acids, Na<sup>+</sup>, K<sup>+</sup>, NO<sub>3</sub><sup>-</sup>, and Cl<sup>-</sup> were determined separately for shoots and roots. Values are means of three replicates ± SE. % Total, Percentage of the amount of individual compounds with respect to the total amount of solutes determined; OS, organic solutes; DW, dry weight. Bold values represent the total amounts of solutes and corresponding percentages per biochemical families.

Compound	Shoots				Roots			
	Control		Salinized		Control		Salinized	
	$\mu\text{mol g}^{-1} \text{ DW}$	% Total	$\mu\text{mol g}^{-1} \text{ DW}$	% Total	$\mu\text{mol g}^{-1} \text{ DW}$	% Total	$\mu\text{mol g}^{-1} \text{ DW}$	% Total
Gln	84.0 ± 6.9	2.6	32.8 ± 3.5	0.8	80.6 ± 10.0	2.8	26.4 ± 4.0	0.7
Glu	9.9 ± 2.0	0.3	6.6 ± 1.3	0.2	8.8 ± 1.3	0.3	5.9 ± 1.3	0.2
$\beta$ -Ala	0.6 ± 0.5	0.0	0.5 ± 0.1	0.0	0.5 ± 0.2	0.0	0.6 ± 0.5	0.0
Pro	38.3 ± 6.8	1.2	76.0 ± 5.3	1.8	18.7 ± 3.5	0.6	65.7 ± 3.9	1.8
Other amino acids	24.1 ± 2.8	0.7	24.5 ± 2.1	0.6	32.5 ± 1.1	1.1	26.3 ± 3.7	0.7
<b><math>\Sigma</math> amino acids</b>	<b>156.8 ± 7.4</b>	<b>4.9</b>	<b>140.4 ± 4.7</b>	<b>3.3</b>	<b>141.2 ± 5.8</b>	<b>4.9</b>	<b>125.0 ± 10.2</b>	<b>3.5</b>
$\beta$ -AB	50.0 ± 3.7	1.6	44.0 ± 8.3	1.0	16.9 ± 2.3	0.6	19.0 ± 2.6	0.5
Cho	9.8 ± 1.4	0.3	8.0 ± 1.5	0.2	6.0 ± 3.1	0.2	4.0 ± 1.7	0.1
COS	52.1 ± 5.8	1.6	49.0 ± 7.2	1.2	37.0 ± 12.4	1.3	41.1 ± 3.1	1.1
<b><math>\Sigma</math> QACs</b>	<b>111.9 ± 7.8</b>	<b>3.5</b>	<b>101.0 ± 16.9</b>	<b>2.4</b>	<b>59.9 ± 17.4</b>	<b>2.1</b>	<b>64.2 ± 6.8</b>	<b>1.8</b>
Suc	221.2 ± 40.4	6.9	259.4 ± 48.2	6.1	258.1 ± 63.7	8.9	462.8 ± 131.4	12.9
Fru	19.3 ± 3.5	0.6	45.0 ± 11.0	1.1	60.1 ± 5.9	2.1	99.5 ± 22.3	2.8
Glc	29.8 ± 2.8	0.9	67.4 ± 16.5	1.6	92.3 ± 15.4	3.2	130.5 ± 36.0	3.6
<i>c</i> -Inos	56.0 ± 4.6	1.7	102.1 ± 15.5	2.4	7.2 ± 0.8	0.2	21.3 ± 8.2	0.6
<i>m</i> -Inos	13.6 ± 1.6	0.4	15.0 ± 3.3	0.4	2.8 ± 0.2	0.1	6.7 ± 2.0	0.2
<b><math>\Sigma</math> NSCs and cyclitols</b>	<b>339.9 ± 41.0</b>	<b>10.5</b>	<b>488.8 ± 41.2</b>	<b>11.6</b>	<b>420.5 ± 64.1</b>	<b>14.4</b>	<b>721.0 ± 185.0</b>	<b>20.1</b>
Malate	68.0 ± 11.0	2.1	34.0 ± 4.0	0.8	35.0 ± 8.0	1.2	24.0 ± 6.0	0.7
Citrate	77.0 ± 4.0	2.4	53.0 ± 6.0	1.3	55.0 ± 7.0	1.9	17.0 ± 2.0	0.5
Other organic acids	12.0 ± 0.7	0.4	9.0 ± 0.7	0.2	11.0 ± 0.9	0.4	6.0 ± 0.3	0.2
<b><math>\Sigma</math> organic acids</b>	<b>157.0 ± 9.6</b>	<b>4.9</b>	<b>96.0 ± 5.3</b>	<b>2.3</b>	<b>101.0 ± 6.3</b>	<b>3.5</b>	<b>57.0 ± 4.2</b>	<b>1.6</b>
<b><math>\Sigma</math> OS</b>	<b>765.6 ± 42.5</b>	<b>23.8</b>	<b>826.3 ± 33.1</b>	<b>19.6</b>	<b>722.6 ± 69.8</b>	<b>24.8</b>	<b>957.1 ± 182.1</b>	<b>26.7</b>
Na <sup>+</sup>	23.3 ± 1.6	0.7	744.0 ± 233.8	17.6	19.3 ± 1.7	0.7	1,014.3 ± 63.5	28.3
K <sup>+</sup>	1,398.2 ± 142.5	43.4	1,063.8 ± 111.0	25.2	1,383.9 ± 169.4	47.6	516.5 ± 62.3	14.4
NO <sub>3</sub> <sup>-</sup>	138.5 ± 14.6	4.3	77.5 ± 9.2	1.8	419.8 ± 27.4	14.4	217.7 ± 12.1	6.1
Cl <sup>-</sup>	896.5 ± 60.1	27.8	1,512.7 ± 130.2	35.8	364.5 ± 40.8	12.5	877.6 ± 80.1	24.5
<b><math>\Sigma</math> inorganic solutes</b>	<b>2,456.6 ± 210.0</b>	<b>76.2</b>	<b>3,398.0 ± 250.8</b>	<b>80.4</b>	<b>2,187.4 ± 227.1</b>	<b>75.2</b>	<b>2,626.1 ± 181.8</b>	<b>73.3</b>

were Suc, Fru, Glc, and cyclitols. They represented more than 50% of the total amount of organic solutes determined in this study. The uncommon cyclitol *chiro*-inositol (*c*-Inos) was characterized in Plumbaginaceae on the basis of the chromatographic retention time and the specific fragmentation signature (mass-to-charge ratio) of its trimethylsilylated adduct as compared to the commercial trimethylsilylated standard. The level of this novel cyclitol was higher in shoots where it represented more than 16% of carbohydrates and cyclitols.

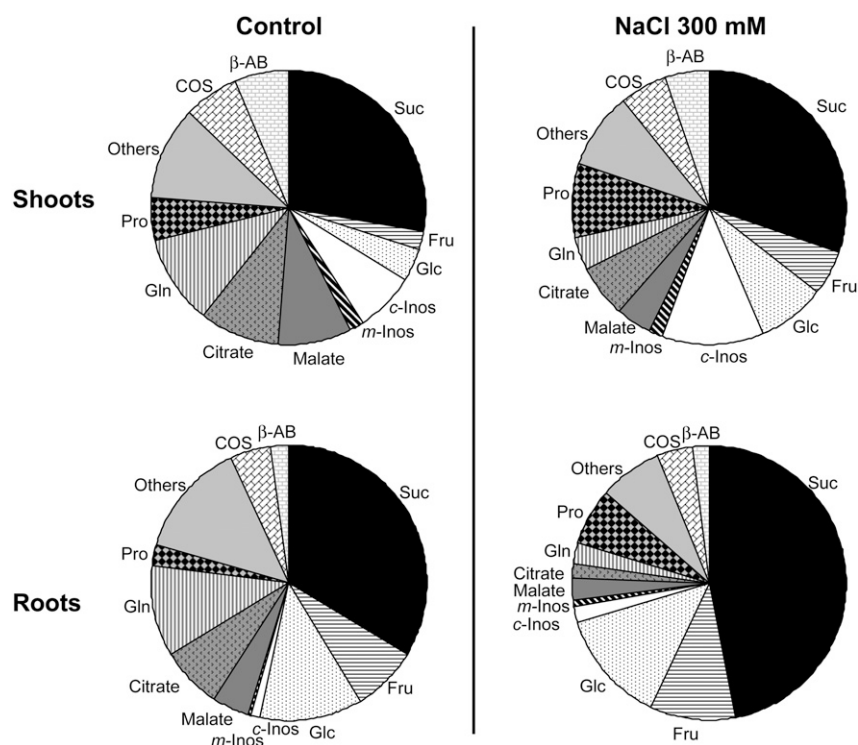
Assuming that all these organic substances behave as ideal osmotic particles, we calculated their relative efficiency in lowering the osmotic potential according to Munns and Weir (1981; Fig. 1). The osmotic potential collectively developed by them (i.e. one-third of the total solutes determined in this study) was lower in roots and shoots of salinized plants as compared to controls (from -0.28 MPa to -0.36 MPa and from -0.65 MPa to -0.70 MPa in roots and shoots, respectively). The major contributor to this calculated potential was Suc, which accounted for more than 30% of the total in roots and shoots under both conditions. Sugars plus cyclitols were actually responsible for more than 50% of the predicted osmotic potential due to organic solutes. Unexpectedly, the betaines COS and  $\beta$ -AB ac-

counted for less than 12% of this potential and Pro for less than 10%, even under saline conditions. Organic acids that contributed significantly to this calculated potential in control seedlings were found to decrease in those treated with NaCl. Finally, in spite of the remarkable contribution of Suc, more than 75% of the total osmotic potential that could arise from both types of osmolytes resulted from inorganic particles.

#### Variation of the Metabolic Phenotype along the Salt-Free Recovery Process in Shoots and Roots of *L. latifolium* following a Saline Treatment

To further investigate whether or not reconfiguration of metabolite profiles could take place in response to salinization and, if so, whether it was reversible, 3-month-old seedlings were treated for 10 d with 300 mM NaCl and then transferred to a reference NaCl-free medium for 10 d of recovery. In parallel, control seedlings were kept on the nonsaline reference medium. As described in "Materials and Methods," shoots and roots from bulks of seedlings were collected at different stages of the experiment (three representative bulks per stage), and were separately analyzed for 30 organic solute contents suspected to be involved in adaptive or





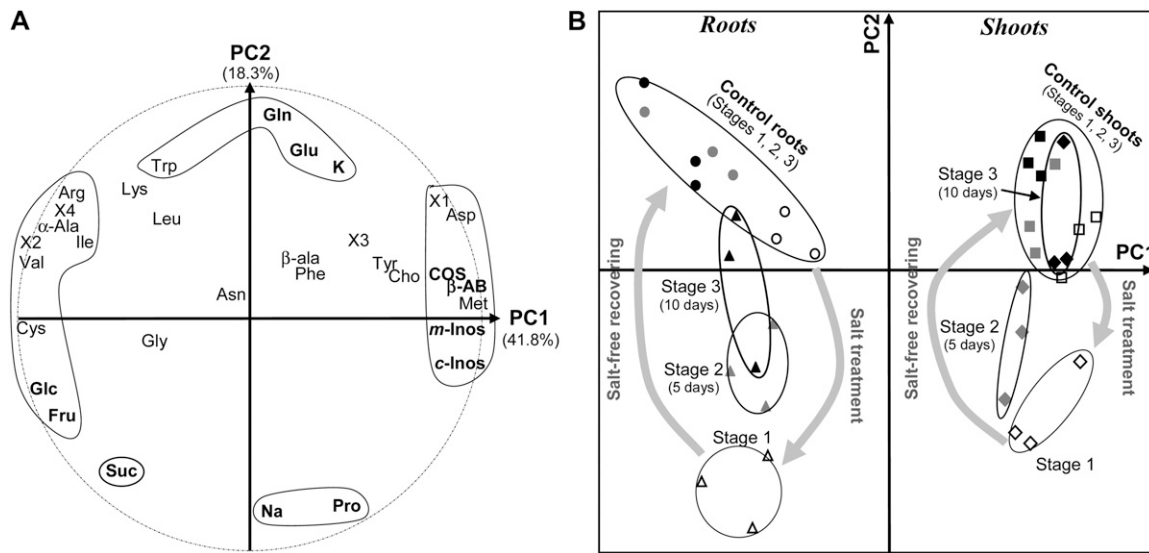
**Figure 1.** The relative contribution of organic solutes to osmolarity in *L. latifolium*. Three-month-old seedlings were treated or not for 10 d with 300 mM NaCl. Analyses were performed separately for shoots and roots.

stress responses, and two inorganic ones ( $\text{Na}^+$  and  $\text{K}^+$ ): stage 1, after 10 d NaCl treatment; stages 2 and 3, after 5 and 10 d of recovery in a salt-free reference medium, respectively. Also, three parallel bulks of shoots and roots of non-salt-treated seedlings kept in the reference medium were collected and analyzed at each of the three corresponding stages to serve as control. Principal component (PC) analysis (PCA) was then employed to examine the differences and similarities among 38 controls, salt-treated, and salt-free recovering seedling samples, respectively, based upon the variation of their metabolic phenotypes represented by 32 variable solutes.

The first two PC axes cover 60.1% of the total variation (PC1: 41.8%; PC2: 18.3%). Projection of the initial variable solutes in the plot defined by the first axes is presented in Figure 2A. This figure shows that the solutes do not vary in the same way and variously contribute to the significance of the PCs. Therefore, the positive part of the first component is strongly correlated with the highest values of a group of organic solutes, such as COS,  $\beta$ -AB, *c*-Inos, *myo*-inositol (*m*-Inos), Met, Asp, and choline (Cho). At the opposite, the negative part of PC1 is well correlated with the highest concentration of the amino acids Cys, Val,  $\alpha$ -Ala, Ile, and Arg, and most particularly of the sugars Glc, Fru, and partly Suc. Concerning PC2, its positive part is well correlated with the highest levels of Gln, Glu, and  $\text{K}^+$  (and secondarily Trp and Lys), whereas the negative part is defined by the highest values of Pro,  $\text{Na}^+$ , and also Suc. The other solutes contribute much less to the axes.

Relative to the same informative PC1 and 2, samples are clearly separated on both sides of PC1 into two main assemblages (Fig. 2B) corresponding to either the shoot or the root samples, regardless of their treatment and stage of development. According to the significance of the axes, the shoot samples share a metabolic profile characterized by the highest concentrations of COS and  $\beta$ -AB, *m*-Inos, and *c*-Inos, and at a lower level of some amino acids such as Met, Asp, X1, and Cho; at the opposite, the profiles of the root samples are characterized by the highest values of the amino acids Cys, Val,  $\alpha$ -Ala, Ile, and Arg, and most particularly of those for Glc, Fru, and partly Suc, and vice versa. While PC1 unambiguously distinguishes organs (shoots versus roots), it is obvious that the second PC (PC2) primarily discriminates groups of shoot and root samples according to their treatment, and secondarily to their developmental stage (1, 2, or 3). As can be observed from Figure 2B, both shoot and root samples are scattered along PC2, in decreasing order, from the control seedlings, with highest levels of Gln, Glu, and  $\text{K}^+$  (and secondarily Trp and Lys) in the positive direction, to the salt-treated samples, accumulating the highest contents of Pro,  $\text{Na}^+$ , and Suc in the opposite direction. In between are ranked the salt-free recovering samples, those recovering for 5 d closer to the salt-treated samples, and the ones recovering for 10 d closer to the control seedlings. Reversal to pretreatment (nonsaline) conditions was faster in shoots than in roots.

Three classes of metabolites arose, according to the changes observed in their concentrations, in response to successive up- and downshift osmotic stresses. Typical



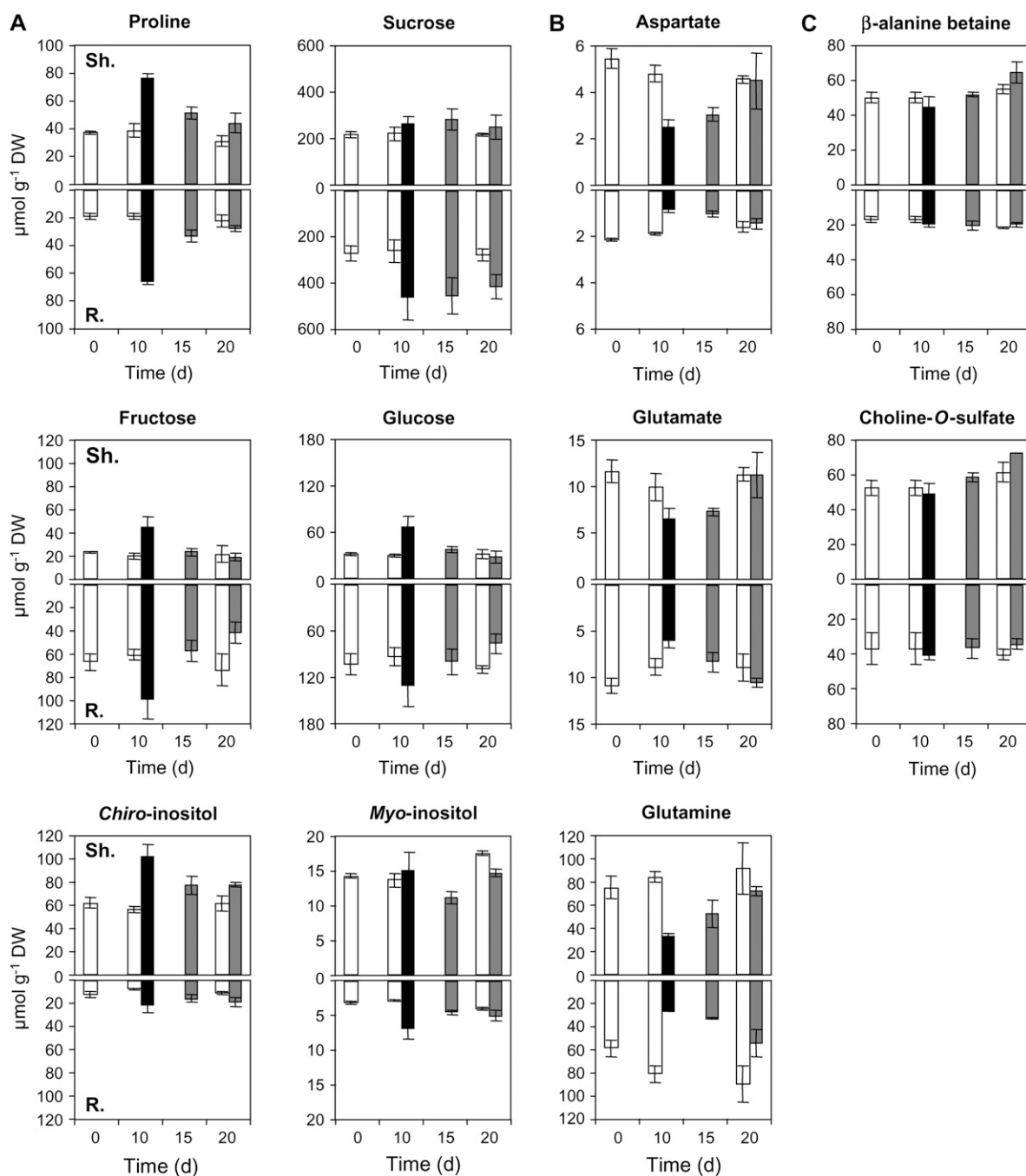
**Figure 2.** PCA showing the pattern of the metabolic phenotype variation along the salt-free recovery process in shoots and roots of *L. latifolium* seedlings after a saline treatment. Three-month-old seedlings, first treated for 10 d with 300 mM NaCl, were transferred for a 10 d period of recovery on a nonsaline reference medium. Shoots (indicated in the diagram by diamonds) and roots (indicated by triangles) from bulks of seedlings were collected at different stages of the experiment (three representative bulks per stage) and were separately analyzed for their organic (30) and inorganic (two) solute contents: stage 1 (indicated by white symbols), after 10 d treatment with 300 mM NaCl; stages 2 (gray symbols) and 3 (black symbols), after 5 and 10 d of recovery in reference medium, respectively. Also, three parallel bulks of shoots (indicated with squares) and roots (indicated by circles) of non-salt-treated seedlings grown in the reference medium were analyzed at each stage and included in the analysis to serve as control. Presented is the PCA diagram of the 36 analyzed samples based on 32 of their variable solute contents. The most remarkable solutes are indicated in bold. The first two PC axes cover 60.1% of the total variation (PC1: 41.8%; PC2: 18.3%). B shows the plot of the 36 samples in the two first PCs, whereas the variable solutes are loaded on the same components in A. X1 to 4 refer to unknown compounds revealed by the amino acid chromatographic profile.

changes for a restricted number of them are plotted in Figure 3. The first class (Fig. 3A) consisted of substances like Pro, Suc, Fru, Glc, *c*-Inos, and *m*-Inos, which were found to increase in response to salinization and to decrease after transfer to nonsaline conditions. They behave as true osmoregulators. The second one (Fig. 3B) is composed of substances whose amounts were inversely regulated. Thus Gln, Glu, and Asp decreased under saline conditions and increased in response to nonsaline ones. The last one (Fig. 3C) was constituted by the betaines, which exhibited amounts that were not significantly adjusted in response to changing experimental conditions.

#### Osmolyte Deposition as Related to Either Constitutive Processes or Salt-Induced Adjustments

To assess changes of the metabolic phenotype related to either vegetative growth or to stress responses shoots and roots of *L. latifolium* seedlings have been examined during acclimatization to either mild salinity or nonsaline conditions. Accordingly, 6-d-old seedlings, sowed in the same time and conditions, were divided in two sets, and each of them was grown hydroponically for 54 d in presence of either 100 mM NaCl or the nonsaline reference medium. Three individuals were collected from each of the salt-treated set

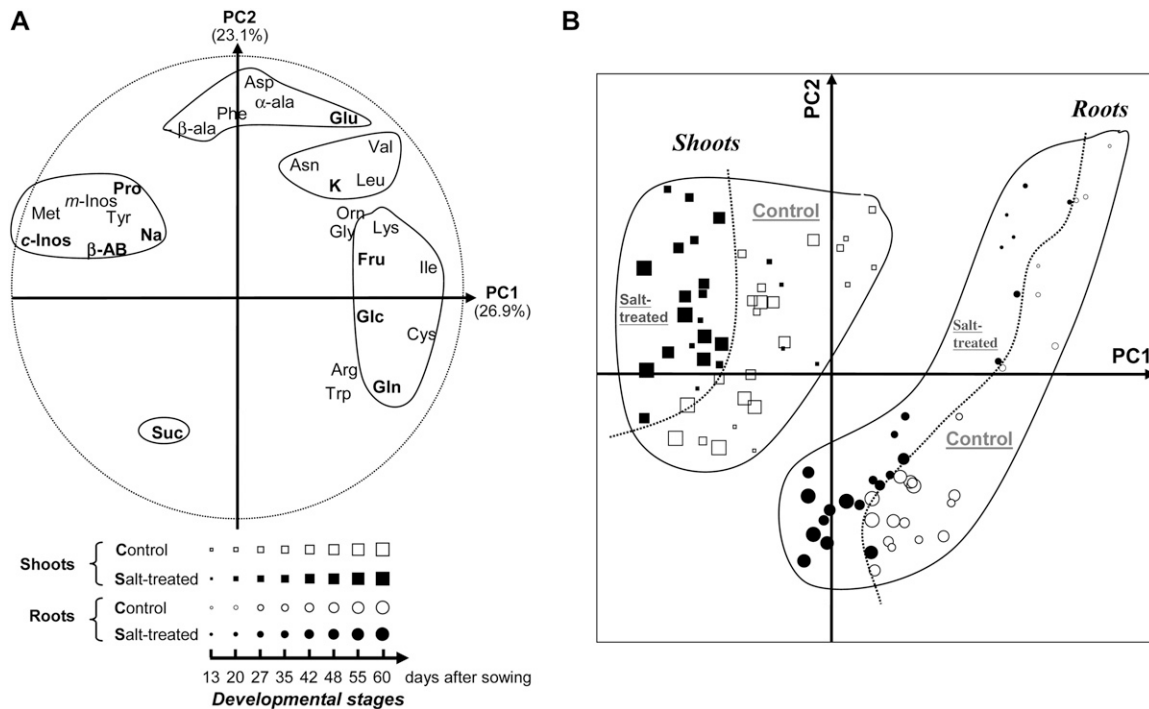
and the reference set at eight different developmental stages (13, 20, 27, 35, 42, 48, 55, and 60 d after sowing), resulting in a sampling of 48 individuals. Shoots and roots from each of these individuals were separately analyzed for 27 solute contents (including Na<sup>+</sup> and K<sup>+</sup>). The whole data set obtained, including 96 shoot and root samples (from 48 individuals) and 27 variables, was analyzed by PCA, regardless of their organ, their age, or their treatment status. The results based on the two most informative first PC (containing 50% of the total variation) are presented in Figure 4. The contribution of the 27 initial variables to the significance of the two first PC axes is shown in Figure 4A. The positive part of the first axis is mainly defined by the highest values of Ile, Cys, Gln, Glc, Fru, and at a lower level by other solutes, such as Lys, Arg, Trp, Val, Orn, and Leu. The negative part is clearly defined with the highest amount of *c*-Inos, Met, *m*-Inos,  $\beta$ -AB, Tyr, Pro, and Na<sup>+</sup>. Conversely, the second axis is strongly correlated with the highest amount of Asp,  $\alpha$ -Ala, Phe, Glu, and  $\beta$ -Ala, on one side, and with a high level of Suc, on the other side. According to this pattern, the two first PCs clearly discriminate the roots from the shoots in two assemblages among the 96 samples (see legend of Fig. 4), regardless of their developmental stage (Fig. 4B). The shoot samples exhibit higher levels of *c*-Inos and *m*-Inos, and of some amino acids or



**Figure 3.** Changes in the amounts of major organic contributors to osmolarity in *L. latifolium*. Three-month-old seedlings first treated for 10 d with 300 mM NaCl were transferred for a 10 d period of recovery on the reference medium. Shoots and roots collected at various moments of the experiment were analyzed separately (white bars, control plants; black bars, salt-treated plants; gray bars, salt-treated plants transferred to the reference medium devoid of NaCl). Sh., Shoots; R., roots. Values are means of three replicates  $\pm$  SE.

derivative such as Met,  $\beta$ -AB, Tyr, and Pro, whereas root samples mostly accumulate Glc and amino acids, such as Gln, Arg, Cys, and vice versa. As can be seen from Figure 4B, within each of these assemblages, salt-treated samples, on one side, are relatively well separated from the control samples, on the other side,

along the PC1. Partial analyses of shoot or root samples alone (data not shown) also support such separation between salt-treated samples, which accumulate remarkable amounts of Pro and *c*-Inos with the increasing Na<sup>+</sup> content and non-salt-treated samples, particularly characterized by a high amount of Gln. Thus,



**Figure 4.** PCA showing the pattern of the metabolic phenotype changes in shoots and roots of *L. latifolium* seedlings during acclimatization to saline or nonsaline condition. Six-day-old seedlings were divided in two sets and each grown hydroponically for 54 d in presence of either 100 mM NaCl or the nonsaline reference medium. Three individuals were collected from each of the salt-treated set (indicated by black symbols) and the reference set (indicated by white symbols) at eight different developmental stages (indicated by increasing sizes from 13–60 d after sowing). Shoots (indicated in the diagram by squares) and roots (indicated by circles) of each sample were separately analyzed for 27 solute contents (including  $\text{Na}^+$  and  $\text{K}^+$ ). Presented are the results generated by the PCA analysis of 96 shoot and root samples based on 27 of their variable solute contents. The first two PCs cover 50% of the total variation (PC1: 26.9%; PC2: 23.1%). The symbols used to represent the different saline versus nonsaline treatments, shoots versus roots parts of the seedlings, and the successive developmental stages are summarized in the legends. The loading of the variable solutes in the two first PCs is presented in A, whereas B shows the plot of the 96 samples in the same two first axes. The most remarkable solutes are indicated in bold in A.

the PC1 not only allows separating shoots from roots, but also to distinguish the treated from the nontreated organs based upon their metabolic profiles. Along the second axis, both the salt-treated and the control root samples evolve in the same way and show significant changes of their metabolite phenotype according to their developmental stage. Thus, the seedlings ranged from the youngest ones, mainly accumulating in their roots Glu,  $\text{K}^+$ , and some amino acids, such as Asn, Val, Leu, Orn, while the roots of the oldest seedlings are essentially characterized by the highest levels of Suc, and conversely, with intermediates in between these two extremes. Such remarkable correlation of the metabolite phenotype and vegetative growth is much less obvious in the case of shoots where the range of variation is narrowest and shows an overall OA independent from the vegetative growth.

More accurate survey of changes in concentration of some remarkable solutes involved actively or not in OA provides additional insights into this general pattern. It was first observed (Fig. 5, A and B) that total carbohydrates (Glc, Fru, Suc) plus cyclitols increased with age in shoots and roots whereas free amino acids

decreased (Fig. 5, C and D), this being independent of growth conditions. It was also apparent (Fig. 6) that during the whole period of acclimatization investigated, Suc behaved as a major solute in shoots of both types of seedlings and it accumulated progressively in roots. On the contrary, the amount of *c*-Inos remained rather low and stable in roots when it increased in shoots especially under saline conditions. In parallel, *m*-Inos decreased in shoots under both conditions of growth while it remained quite low and stable in roots. The Pro level, which remained high in shoots of treated seedlings during the whole experiment, was twice as low in shoots of control seedlings. Higher amounts of Pro also occurred in roots of treated seedlings especially at the first stages of vegetative growth. Changes in Gln concentration of shoots mimicked those previously mentioned for total amino acids, the values being higher in control seedlings.  $\beta$ -AB was already present at high levels in young seedlings and it tended to increase in older ones but these changes did not seem to be related to growth conditions. The same trends were observed for COS since its amount in shoots of both types of seedlings was found to be close to 60 and 80

$\mu\text{mol g}^{-1}$  dry weight in 35- and 60 d-old seedlings, respectively.

As shown in Figure 7, only the net accumulation of Pro increased to some extent in salinized seedlings whereas those of  $\beta$ -AB and *c*-Inos evolved in relationship with growth. As expected for salt-treated seedlings,  $\text{Na}^+$  accumulated during vegetative growth and this occurs at the expense of  $\text{K}^+$ . As a consequence, the absolute amount of both cations behaved similarly in control and salt-treated seedlings.

#### Changes in Subcellular Distribution of Pro and $\beta$ -AB as Related to Sudden Salinization

To specify the role(s) of  $\beta$ -AB and Pro in OA, the concentrations of these typical stress metabolites present in the cytosolic, vacuolar, and plastidic compartments in mesophyll cells of fully developed leaves were evaluated for 3-month-old seedlings that were either submitted or not for 10 d to a 300 mM NaCl salt treatment. The results shown in Table II represent the mean values of three independent fractionation trials. The changes in Pro and  $\beta$ -AB concentrations that were induced by NaCl in these compartments have been calculated with the assumption that cell volumes as well as compartment volumes were not modified in response to the salt treatment and according to the means of typical volumes determined for leaves of other species (Winter et al., 1993; Leidreiter et al., 1995). Under control conditions a large proportion of the investigated solutes was found in the vacuoles. Thus 62% of total Pro content and 94% of total  $\beta$ -AB content were located in this compartment, with the remaining fractions being compartmentalized in the chloroplasts. In tissues of salinized seedlings, 11% of total free Pro content was found in the chloroplasts, 11% was cytosolic, and the remaining was vacuolar. This strongly sug-

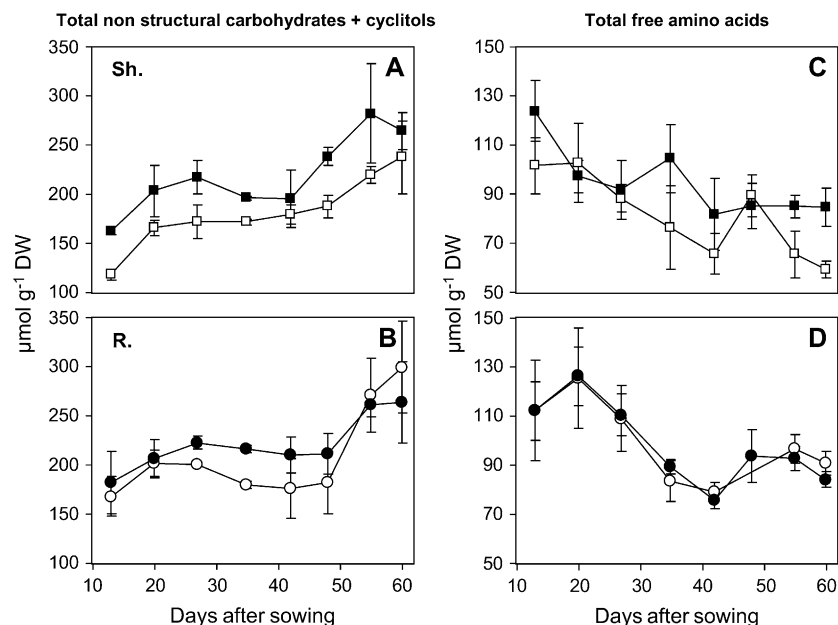
gested that a significant part of Pro newly synthesized in response to salt stress could be loaded in the non-vacuolar compartment.  $\beta$ -AB levels remained constant under salt stress, but its content increased in the cytosol reaching 18% of the total, whereas chloroplasts became free of it. In parallel, the amount of  $\beta$ -AB stored in vacuoles was found to decrease and this change might be related to a salt-dependent transfer of a significant part of the betaine accumulated in the vacuoles to the cytoplasm.

#### DISCUSSION

This study carried out with seedlings of the halophyte *L. latifolium* aimed first to investigate the profile of the major organic solutes occurring in its tissues and to assess their contribution to adjustment of osmolarity under saline conditions. Second, we tried to discriminate substances that actually behave as true osmoregulatory solutes from those that participate to the same function through passive processes. We have also analyzed the conditions of deposition of these solutes during acclimatization to either mild salinity or nonsaline conditions. Finally, we performed compartmental analysis on shoot tissues from seedlings treated or not with 300 mM NaCl to investigate cellular localization of Pro and  $\beta$ -AB. The metabolic data generated were expected to provide relevant arguments to assess the suitability of the compatible solute theory in this plant species.

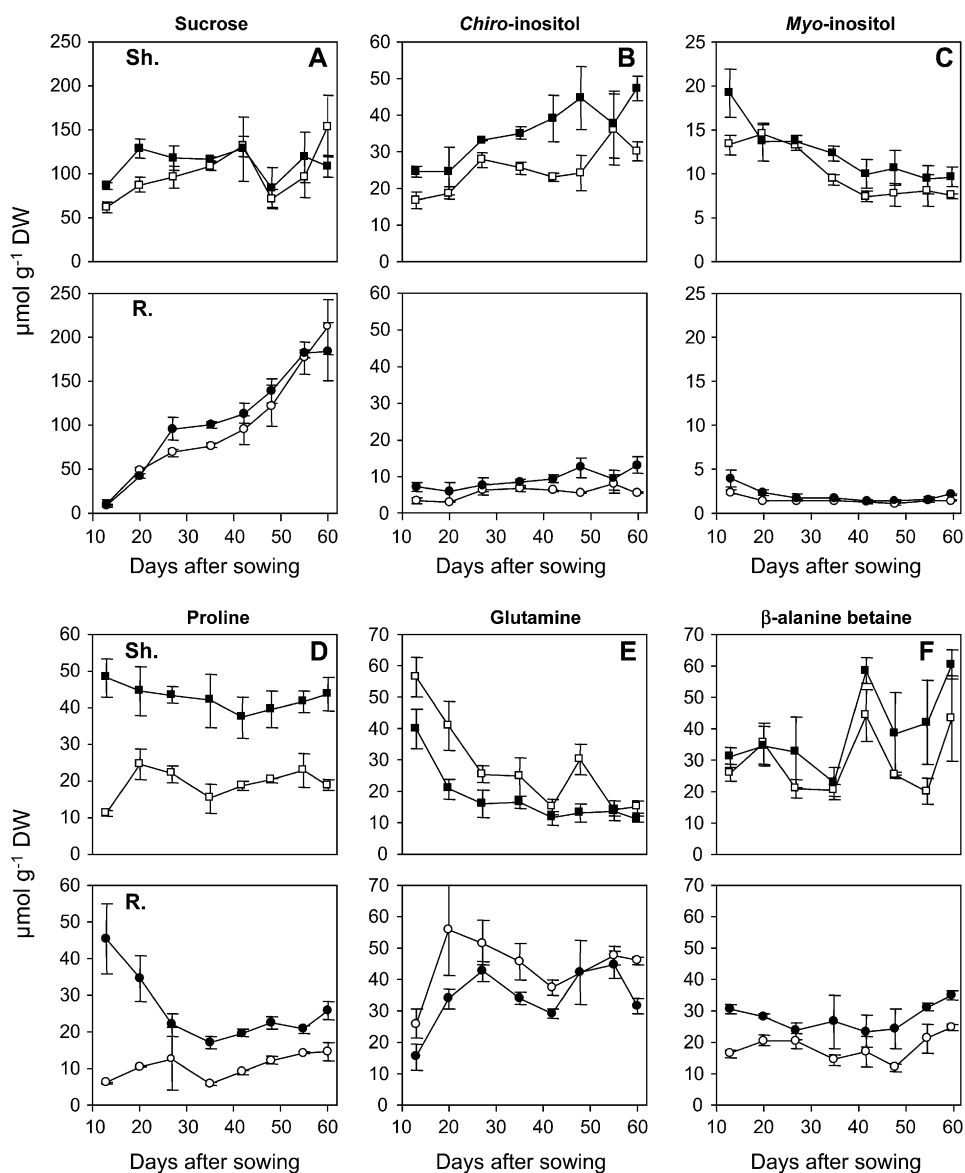
#### Organic Solutes Accumulated by *L. latifolium* Exhibit a More Complicated Status than Expected

Collectively the solutes determined in this study gave rise to a calculated osmotic potential close to  $-1.5$  and  $-3$  Mpa, respectively, in roots and shoots, which



**Figure 5.** Changes in the amounts of total non-structural carbohydrates + cyclitols contents and that of total free amino acids contents during acclimation of *L. latifolium* to saline or nonsaline conditions. Six-day-old seedlings were grown hydroponically for 54 d in presence of either 100 mM NaCl or the reference medium. Shoots and roots collected at various stages of the vegetative growth were analyzed separately. Symbols are as follows: circles, roots; squares, shoots. White symbols represent tissues issued from control seedlings and black ones from salinized seedlings (values are means of three replicates  $\pm$  SE).

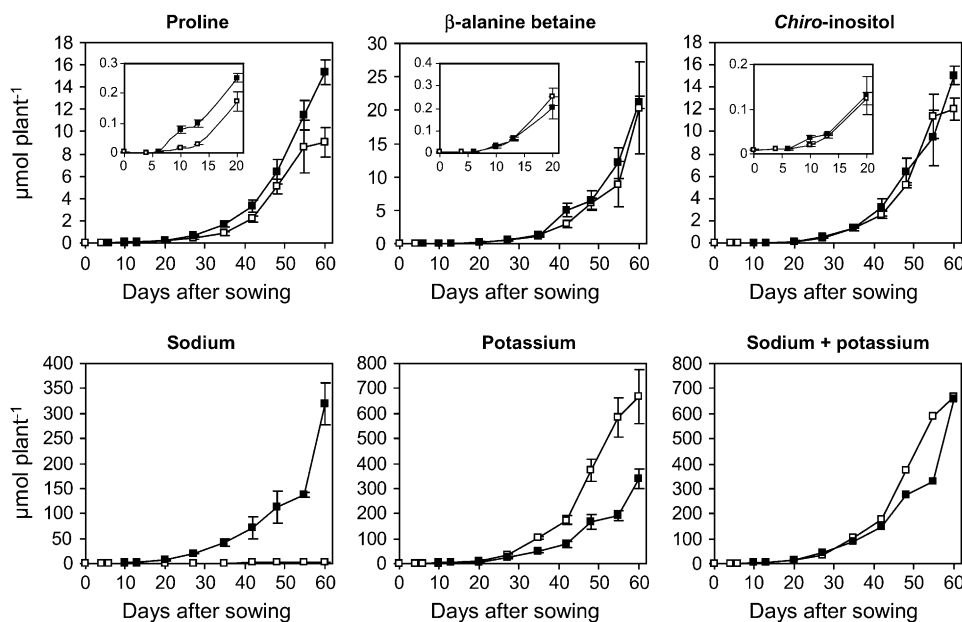
**Figure 6.** Changes in the relative amounts of a restricted number of organic solutes during acclimation of *L. latifolium* to saline or non-saline conditions. Six-day-old seedlings were grown hydroponically for 54 d in presence of either 100 mM NaCl or the reference medium. Shoots and roots collected at various stages of the vegetative growth were analyzed separately. Symbols are as follows: circles, roots; squares, shoots. White symbols represent tissues issued from control seedlings and black ones from salinized seedlings (values are means of three replicates  $\pm$  SE).



was in keeping with results obtained for *L. latifolium* and its interspecific hybrid with *Limonium caspia* by Alarcon et al. (1999). This was assumed to provide the seedlings with a water potential gradient steep enough to maintain turgor when subjected to sudden changes in external salinity. The range of organic osmolytes involved was broader than that previously shown to occur. Previous studies had emphasized the osmotic function of nitrogenous compounds such as Pro,  $\beta$ -AB, and COS (and GB in a restricted number of other *Limonium* species) but here we show that their contribution to the regulation of osmotic pressure remains relatively weak (Fig. 1). The major osmoticum constituted by free sugars, cyclitols, and organic acids has previously been overlooked. *L. latifolium* actually behaves as a glycohalophyte with a strong ability to allocate Suc and hexoses to combat salinity. This fits with the low relative growth rate determined for seedlings growing in the

presence of 100 mM NaCl ( $0.05 \text{ g g}^{-1}$  dry weight  $\text{d}^{-1}$ ). The cooccurrence of *c*-Inos and *m*-Inos is of special interest for salt tolerance because cyclitols and their methylated derivatives have already been described as compatible solutes in halophytic and glycophytic species (Smirnov and Cumbes, 1989; Larher et al., 1990; Bohnert et al., 1995; Popp and Smirnov, 1995). Pinitol, which results from methylation of *m*-Inos, was not detectable in *L. latifolium* while it was found to be associated with both isomers of inositol in *Limonium gmelini* (A. Bouchereau, unpublished data).

Stability of the profiles of major organic solutes all along the whole period of vegetative growth investigated indicates that long-term saline treatment did not induce (or suppress) the production of special solutes. However, regardless of the presence of NaCl, the free sugars + cyclitols fraction increased regularly while that of free amino acids decreased. Such imbalance



**Figure 7.** Changes in the absolute amounts of a restricted number of compatible solutes and that of  $\text{Na}^+$  and  $\text{K}^+$  during acclimation of *L. latifolium* to saline or nonsaline conditions. Six-day-old seedlings were grown hydroponically for 54 d in presence of either 100 mM NaCl or the reference medium. Shoots and roots collected at various stages of the vegetative growth were analyzed separately. Insets show results monitored during the first 20 d. Symbols are as follows: white squares, control plants; black squares, treated plants. The results are the means of three replicates ( $\pm$ SE).

between carbon and N metabolites may restrict relative growth rate and therefore play a part as a determinant of salt tolerance (Schulze and Chapin, 1987). Unfortunately this is not necessarily relevant for crop breeding because of its detrimental effect on agronomic yield.

#### Organic Osmolytes Occurring in *L. latifolium* Could Be Ranked into Two Major Classes According to Their Pattern of Deposition

Owing to robust metabolic analyses performed on seedlings subjected to saline treatments, we were able to discriminate between substances that are constitutively accumulated and those that behaved as stress metabolites. Exploratory data analysis by PCA revealed that the amount of certain metabolites that increased in response to salinization showed an opposite trend in response to transfer to nonsaline conditions (Figs. 2 and 3). Such biochemical flexibility is assumed to provide vectorial homeostasis in solute concentrations that

allow proper functioning of primary metabolism as well as fine control of the amount of solutes needed to maintain osmolarity. Some of these changes could mimic those of perfect osmolytes but we have to concede that they could also reflect fluctuating rates of solute consumption related to inhibition and resumption of growth. Independent of the cause underlying the observed effect, it could be inferred that free sugars and *c*-Inos apparently behave as ideal osmolytes. In performing such function Pro seems to have a minor importance because it accumulated only in relatively low amounts in response to saline upshift. The reversible changes observed in the Pro level might merely be related to successive damage and repair at the mitochondrial step of Pro oxidation by the Pro dehydrogenases that control the Pro/pyrroline-5-carboxylate cycle operating between the cytosol and the mitochondria (Larher et al., 2007). Surprisingly, for the uncommon betaines of *L. latifolium* their amount was not found to be significantly osmoregulated; thereby they behaved as passive osmolytes that did not meet all the requirements

**Table II.** Changes in compartmentation of Pro and  $\beta$ -AB in *L. latifolium* subjected to a salt shock

Hydroponically grown 3-month-old seedlings were treated for 10 days with 300 mM NaCl. Control seedlings were maintained for same time on the reference medium. Fully expanded leaves were collected 7 h after the onset of light period and further submitted to compartmental analysis. Total leaf tissue contents of these solutes were also shown. %, Percentage of Pro and  $\beta$ -AB with respect to their total amount; mM, concentrations as related to presumed volume of each cellular compartment; FW, fresh weight.

Organic Solutes	Total Amounts	Chloroplast		Cytosol		Vacuole	
	$\mu\text{mol g}^{-1}$ FW	%	mM	%	mM	%	mM
Pro							
Control	1.5	38	6.5	0	0	62	1.2
Salinized	17	11	23.4	11	47	78	18.7
$\beta$ -AB							
Control	13.5	6	10	0	0	94	17.9
Salinized	14.8	0	0	18	66	82	17

for typical compatible solutes. This does not preclude other beneficial functions like those attributed to GB (Murata et al., 1992; Hoekstra et al., 2001).

#### Compartmental Analysis Based on Nonaqueous Fractionation Revealed Unexpected Locations for Pro and $\beta$ -AB

In control plants both solutes were already abundant and preferentially associated with the vacuolar fraction. This apportionment was found to change in response to salinization. Thus salt-induced enhancement of the Pro amount benefits the cytosol and this might result from activation of the Glu pathway of Pro synthesis operating at this level (Hare et al., 1998). On the contrary, as the  $\beta$ -AB amount was not osmoregulated, the significant amount of this betaine found in the cytosol from salt-treated tissues must result from its transfer from the vacuole. However, enhancement of Pro and  $\beta$ -AB in the cytosol remained in the low range and thus it is suggested that they might behave as counteracting cytoprotectants because their actual concentration could be sufficient to strive against some of the damaging effects of salinization. In other words they could act as osmoprotectants in the cytosol and as osmolytes in the vacuole.

However, this scheme results from a speculative oversimplification deduced from analysis of complex crude extracts coming from a great variety of plant cells exhibiting specific anatomical and physiological attributes hidden by the destructive procedure used. In addition, the concentrations shown in Table II remain rough estimations because the changes that could be induced by salinity at the fraction of the total cell volume represented by the vacuolar and the non-vacuolar compartments have not been assessed. Salt-induced plasticity at this level has been reported for cultured plant cells derived from both halophytic and glycophytic species. Thus salinity has been shown to induce a rapid increase in vacuolar volume associated with activation of  $H^+$ -ATPase and vacuolar acid phosphatase without any change in the total cell volume (Mimura et al., 2003). Those changes could not solely increase osmotic concentration in the cytosol without the need of additional osmotic particles but also provide an enlarged vacuolar compartment for the storage of  $Na^+$  and  $Cl^-$ . These authors documented similar responses in barley (*Hordeum vulgare*) root meristematic cells but not in those of salt-sensitive species like pea (*Pisum sativum*) and tomato (*Solanum lycopersicum*). Chang et al. (1996) also demonstrated extensive vacuolation in cultured cells of tobacco (*Nicotiana tabacum*) treated with NaCl. On the contrary Binzel et al. (1988) demonstrated that tobacco cells adapted to 428 mM NaCl exhibited a very large decrease in both whole cell and vacuolar volumes associated with an increase in cytoplasmic one. At the whole plant level, salt tolerance could also partly rely on developing succulence in shoot tissues constituted by inflated cells with very large vacuoles allowing storage of NaCl in a diluted

environment but succulence does not seem to be induced by salinity in *L. latifolium*. In addition, due to the presence of the secretory cells of salt glands a large heterogeneity could also be predicted at the level of vacuolar volumes. Morphometric determinations performed by Faraday and Thomson (1986) on these special cells have indeed shown that the mean percent volume represented by the vacuoles reached 31.4%, which refers to the coarse evaluation of a very complicated status that has been done in this study. In addition, the abilities of various cell types to accumulate one or another organic solute belonging to those found in crude extracts of the whole shoots of *L. latifolium* also remain to be investigated.

#### Some of the Organic Osmolytes Occurring in *L. latifolium* That Do Not Behave as Active Osmobalancers May Be Involved in Other Counteracting Functions That Remain to Be Investigated

As a whole, organic solutes under investigation in this study are obviously involved in the colligative properties of cellular solutions where they are located. However some of them, exhibiting amounts hardly osmoregulated, could be also regarded as secondary plant products associated (or not) with metabolic dysregulations. We are rather prone to believe they might result from active responses intended to exert counteracting effects against damages directly or indirectly induced by NaCl. Such apparent discrepancy between the relevance of metabolic processes involved might have implications in deciphering the metabolic determinants of salt tolerance of halophytes.

First it appears that the ability to allocate a range of primary metabolites to strive against damages caused by salt stress and finally contribute to plant survival under salt stress could reflect traits of paramount importance in salt tolerance. Because these substances were produced by *L. latifolium* without any induction by salinity they actually behave as antistress metabolites preaccumulated to caution the whole plant against the osmotic stresses that could be encountered under salt marsh conditions. Similar metabolic traits have already been reported for *Thellungiella halophila* (Gong et al., 2005). This expensive strategy for anticipated OA is obviously used at the expense of biomass production. Nevertheless, under recovery conditions some of these compounds could be consumed to sustain renewal of growth.

Second, the physiological relevance of the special betaines occurring in *L. latifolium* deserves further investigations because some of their properties documented in this study do not conform to the premises of the compatible solute hypothesis. Our results are indeed at odds with those related for example to the putative function of GB in halophytic Chenopodiaceae (Tipirdamaz et al., 2006). Using methods different from that selected for this study for plant material fractionation, authors have reported that this betaine was mainly located in the extravacuolar compartment where it



could be acting as a major compatible cytosolute (Wyn Jones et al., 1977; Hall et al., 1978; Leigh et al., 1981; Matoh et al., 1987; Hanson, 1992). Clearly close structural relationship between GB and its higher homolog is not an adequate evidence to suspect similar functions in species belonging to different family of halophytes. In addition, reciprocal exclusion between GB and  $\beta$ -AB in the Plumbaginaceae is not necessarily founded on the need for a betaine to prevent deleterious stress effects at the cytosolic level.

Third, it became evident that detailed studies of the metabolic phenotypes expressed in halophytic plants subjected or not to saline conditions, using reliable analytical procedures and suitable statistical tools, deserve to be used more thoroughly to specify the actual interest of the compatible solute hypothesis to predict the role played by solutes accumulated. More realistic pictures should emerge from metabolomics approaches that can apprehend on the long range changes the amount of a wider range of metabolites occurring in both salinized and nonsalinized plants that could reflect genuine adaptive processes or stress responses as well as secondary metabolic pathways of unknown functions.

Finally, if the various organic osmolytes accumulated in tissues of *L. latifolium* accounted for the decline in water potential regardless of their compartmentation, this does not inevitably result from osmoregulatory responses that mitigate the damages provoked at the cellular level by salinity. With respect to the so-called popular compatible solutes consisting of Pro,  $\beta$ -AB, COS, and the cyclitols, they behave actually as minor regulators of intracellular water activity. The expected preferential localization of some of them in cytosol and chloroplasts does not prove correct. Thus it remains speculative whether their relative high amounts could be involved in salt tolerance or if they are just temporally associated with expression of more important traits for coping with salinity. Our findings highlight the question of the real value of their compatibility and suggest that it might be less misleading at this time to call them either compensatory solutes (Gilles, 1997) or counteracting solutes (Yancey, 2005) rather than compatible solutes.

## MATERIALS AND METHODS

### Plant Material and Growth Conditions

Seeds of *Limonium latifolium* were provided by Ball Ducretet. They were germinated (sowing day) for 4 d in 90 mm petri dishes on paper humidified with Hoagland solution (Hoagland and Arnon, 1938) in a growth chamber (14-h light, 10-h dark, 24°C/18°C day/night thermoperiod, light intensity 250  $\mu\text{mol m}^{-2} \text{s}^{-1}$ , relative humidity 75% day and 90% night). The 4-d-old seedlings were then transferred hydroponically to individual pots containing full strength Hoagland solution containing 5 mM  $\text{Ca}^{2+}$ . All treatments and samplings started at midday at the indicated dates. In all experiments, the volume of nutrient solution was daily adjusted with fresh medium and the medium completely renewed weekly.

For recovery experiments after saline treatments, 3-month-old plants grown individually as described above were divided into two sets. One set was transferred to the reference medium and the other to this medium added with

300 mM NaCl. After 10 d of salt treatment, treated plants were separated in two batches: one was immediately harvested and the second transferred to fresh growth medium free of NaCl for 10 more days and then collected. Root and leaf samples from control or salt-treated plants were taken at the 0th, 10th, 15th, and 20th days from onset of the experiment.

For the long-term treatment at mild salinity, seedlings (6-d-old) were divided into two sets. The first one, referred as control, was kept on the same medium and the second, referred as salinized, was transferred to the Hoagland solution added with 100 mM NaCl. Seedlings were harvested at sowing and 4, 6, 10, 13, 20, 27, 35, 42, 48, 55, and 60 d later.

### Plant Sampling and Sample Extraction

Experiments were set up in a completely randomized design. Three replicates were done for each treatment. For each replicate, five to 20 plants were harvested and pooled. Plants were harvested at midday at the indicated dates. Roots and shoots were separately collected, thoroughly rinsed with distilled water, dried, and immediately plunged in liquid nitrogen. Frozen tissues were then lyophilized for 72 h until dry, noting that the tissues were maintained in their frozen state through evaporative cooling during the lyophilization process. The dried material was powdered and stored at  $-80^\circ\text{C}$  until extraction. Dried crushed materials (up to 30 mg) were suspended in 96% ethanol containing 50  $\mu\text{M}$  norleucine and 50  $\mu\text{M}$   $\beta$ -D-phenylglucopyranoside (internal standards for amino acid and sugar profile normalization, respectively) and thoroughly shaken. Suspensions were heated at  $85^\circ\text{C}$  until complete evaporation of ethanol. The residues, resuspended with deionized water, were shaken at  $4^\circ\text{C}$  for 1 h. Homogenates were clarified by centrifugation (15,000g,  $4^\circ\text{C}$ , 20 min) and supernatants stored at  $-20^\circ\text{C}$  until analysis. Crude supernatants were used to quantify low  $M_r$  organic solutes without further purification and to determine the amount of  $\text{Na}^+$  and  $\text{K}^+$  solubilized in these extracts.

### Low Molecular Weight Organic Solutes Analysis

#### Sugars, Sugar Alcohols, and Organic Acids Derivatization and Analysis by Gas Capillary Chromatography

Derivatization and chromatography were achieved according to Adams et al. (1999). Calibration plots were constructed with external standards and peaks were attributed on the basis of their retention time. The percent recovery of metabolites through extraction, derivatization, storage, and quantification procedures was assessed using  $\beta$ -D-phenylglucopyranoside as an internal standard. To ensure identification of unknown compounds, analyses were performed by gas chromatography coupled with mass spectrometry (Agilent Technologies), the chromatographic procedure remaining unchanged. The ion source was adjusted to  $230^\circ\text{C}$ . Mass spectra were recorded at 2 scan  $\text{s}^{-1}$  with a scanning range of 25 to 500 mass-to-charge ratio.

#### Amino Acids Derivatization and Analysis by HPLC

Amino acids were characterized and quantified with HPLC after pre-column derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate (using the Waters AccQ-Fluor™ reagent kit) and reversed-phase liquid chromatographic separation as described by Cohen and Michaud (1993). Ten microliters of the crude aqueous extracts were reacted with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate using the procedure optimized by Bouchereau et al. (1999). Amino acids were characterized by cochromatography of pure synthetic compounds (Sigma) and quantified making reference to individual external calibration curves, recovery of successive procedures being assessed using internal standard norleucine. Our analytical device did not discriminate between Asn and Ser on the one hand and Arg and Thr on another. As a consequence, the amounts of both pairs of amino acids were expressed in terms of Asn and Arg, respectively.

#### Quaternary Ammonium Compounds Determination by $^1\text{H-NMR}$

Quaternary ammonium compounds were determined as described in Bouchereau et al. (1999). Aliquots of crude extracts were freeze dried. Just before analysis residues were redissolved in  $\text{D}_2\text{O}$  (99.9% deuterium) containing *tert*-butanol 0.5 mM as an internal standard. The butanol was used as a

reference both for chemical shift (1.2 ppm) and quantification of the signals.  $^1\text{H-NMR}$  spectra were recorded on a Bruker NMR spectrometer operating at  $^1\text{H}$  frequency of 300 MHz. The processing of the spectra was carried out using Mestrec 2.3 software.

### Inorganic Solutes Determination

$\text{Na}^+$  and  $\text{K}^+$  concentrations in appropriately diluted extracts were determined directly using a flame photometer (Jenway). Chloride has been determined through the colorimetric titration method of Schoenfeld and Lewellen (1964) using mercuric thiocyanate and ferric nitrate as reactants. Nitrate has been assayed through the colorimetric method of Robarge et al. (1983) via nitration of salicylic acid.

### Determination of Subcellular Metabolite Concentrations

Nonaqueous density gradient fractionation of leaves was performed according to Gerhardt and Heldt (1984) and Farré et al. (2001). Leaves were collected at midday. For the fractionation an exponential gradient (25 mL between 1.28 and 1.59 followed by a 2 mL cushion of  $\text{CCl}_4$ ) made using a gradient maker connected to a peristaltic pump was used. For marker enzymes assays, the dried sediments were homogenized for 5 min at  $4^\circ\text{C}$  in 1 mL of 100 mM potassium phosphate buffer pH 7.3 containing 20 mM sodium tetraborate, 1 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol, 2 mM EDTA, 1 mM  $\text{MgCl}_2$ , 10% (w/v) polyvinylpyrrolidone, and 20% (w/v) polyvinylpyrrolidone. Crude homogenates were centrifuged at 10,000g for 20 min and the precipitates were discarded. These extracts were used for enzymatic assays as described by Gerhardt and Heldt (1984). Metabolites were extracted and measured as described above.

The deconvolution approach described by Riens et al. (1991) was used for the evaluation of the subcellular distribution of metabolites between the vacuolar, cytosolic, and plastidic compartments. To calculate metabolite concentrations in millimolar from the analytical results, the volumes of the vacuole, chloroplasts, and cytosol were estimated according to the published values of volumes from spinach (*Spinacia oleracea*), barley (*Hordeum vulgare*), and potato (*Solanum tuberosum*) leaves (Winter et al., 1993; Leidreiter et al., 1995).

### Statistical Analysis

Two main experiments were conducted in this work to study the pattern of the metabolic phenotype changes: (1) along a salt-free recovery process after a saline treatment, and (2) during acclimatization to saline or nonsaline conditions, in shoots and roots of *L. latifolium* seedlings. These experiments resulted in large data sets, one containing 38 samples and 32 variable solutes, and the other including 96 samples and 27 variables. Usual statistical parameters and diagrams (mean values,  $\text{SES}$ , relative percentages, histograms, two-way scattered diagrams) have been used to characterize and estimate the variation of each metabolite, employing Minitab software (Windows version 13.31, Minitab Inc.).

Additionally, a multivariate approach, employing the PCA method, has been performed to give a synthetic view of the data and to identify the pattern and trends of the physiological behavior shown by *L. latifolium* in response to salinization or acclimatization, as revealed by changes of the metabolite profile following the treatments. PCA is a powerful statistical method, which allows estimation of overall similarity and difference levels among analyzed samples, based on a multidimensional data set (Sneath and Sokal, 1973). With the development of possibilities to access an increasing number of variable characters, multivariate analyses (including PCA) become tools of choice to detect which variables play a significant role in differentiating or clustering the samples, particularly in the metabolomic era (Fritz et al., 2006; Manetti et al., 2006). The main advantage of PCA is to compress the information carried by the original variables and their interrelationships (if any) into a smaller number of new synthetic variables called PCs, using a covariance matrix calculated from the initial data set (Sneath and Sokal, 1973). Thus, the first PC covers as much of the variation in the data as possible, the second is orthogonal to the first and covers as much of the remaining variability as possible, and so on. The plot defined by the first PCs (or axes) often contains most of the total variation carried by the initial data. Projection of the individual samples into this plot provides a general picture of their distribution along the PCs. Whereas loading of the variables (organic and inorganic solutes) in the same plot allows identifying variables that contribute most (or not) to the significance of the PCs, and further makes easier interpretation of

sample groupings, similarities, or differences. PCAs were performed using the R 1.9.1 statistical package (<http://www.r-project.org/>).

### Supplemental Data

The following materials are available in the online version of this article.

**Supplemental Figure S1.**  $^1\text{H-NMR}$  spectrum of a crude extract from shoots of 3-month-old seedlings of *L. latifolium*.

### ACKNOWLEDGMENTS

The authors thank Dr. M. Vaultier for allowing them to use the gas chromatography-mass spectrometry system of the department Chimie Synthèse Electrosynthèse Organiques of the University of Rennes 1 (France) as well as B. Plunian for technical assistance. They are grateful to Dr. P. Guenot (Centre Régional de Mesures Physiques de l'Ouest, University of Rennes 1, France) for helpful discussions and suggestions regarding metabolite analysis. They gratefully thank Dr. A. Weber (Michigan State University) and Dr. R.J. Neil Emery (Trent University, Ontario, Canada) for their useful comments on the manuscript and help with the english version. Thanks are also due to C. Monnier and N. Raimbeaud (University of Rennes 1, France) for technical assistance.

Received March 20, 2007; accepted April 6, 2007; published April 27, 2007.

### LITERATURE CITED

- Adams MA, Chen ZL, Landman P, Colmer TD (1999) Simultaneous determination by capillary gas chromatography of organic acids, sugars, and sugar alcohols in plant tissue extracts as their trimethylsilyl derivatives. *Anal Biochem* **266**: 77–84
- Alarcon JJ, Morales MA, Torrecillas A, Sanchez-Blanco MJ (1999) Growth, water relations and accumulation of organic and inorganic solutes in the halophyte *Limonium latifolium* cv. Avignon and its interspecific hybrid *Limonium caspia* × *Limonium latifolium* cv. Betlaard during salt stress. *J Plant Physiol* **154**: 795–801
- Binzel ML, Dana Hess F, Bressan RA, Hasegawa PM (1988) Intracellular compartmentation of ions in salt adapted tobacco cells. *Plant Physiol* **86**: 607–614
- Bohnert HJ, Jensen RG (1996) Strategies for engineering water stress tolerance in plants. *Trends Biotechnol* **14**: 89–97
- Bohnert HJ, Nelson DE, Jensen RG (1995) Adaptations to environmental stresses. *Plant Cell* **7**: 1099–1111
- Bohnert HJ, Shen B (1999) Transformation and compatible solutes. *Sci Hortic (Amsterdam)* **78**: 237–260
- Bouchereau A, Duhazé C, Martin-Tanguy J, Guégan JP, Larher F (1999) Improved analytical methods for determination of nitrogenous stress metabolites occurring in *Limonium* species. *J Chromatogr A* **836**: 209–221
- Brown AD, Simpson JR (1972) Water relations of sugar-tolerant yeasts: the role of intracellular polyols. *J Gen Microbiol* **72**: 589–591
- Chang PFL, Damsz B, Kononowicz AK, Reuveni M, Chen ZT, Xu Y, Hedges K, Tseng CC, Singh NK, Binzel ML, et al (1996) Alteration in some membrane structure and expression of a membrane-associated protein after adaptation to osmotic stress. *Physiol Plant* **98**: 505–516
- Cohen SA, Michaud DP (1993) Synthesis of a fluorescent derivatizing reagent, 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate, and its application for the analysis of hydrolysate amino acids *via* high-performance liquid chromatography. *Anal Biochem* **211**: 1–9
- Diamant S, Eliahu N, Rosenthal D, Goloubinoff P (2001) Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses. *J Biol Chem* **276**: 39586–39591
- Faraday CD, Thomson WW (1986) Structural aspects of the salt glands of the *Plumbaginaceae*. *J Exp Bot* **37**: 461–470
- Farré EM, Tiessen A, Roessner U, Geigenberger P, Trethewey RN, Willmitzer L (2001) Analysis of the compartmentation of glycolytic intermediates, nucleotides, sugars, organic acids, amino acids, and sugar alcohols in potato tubers using a nonaqueous fractionation method. *Plant Physiol* **127**: 685–700
- Flowers TJ (1972) The effect of sodium chloride on enzyme activities from four halophyte species of *Chenopodiaceae*. *Phytochemistry* **11**: 1881–1886

- Fritz C, Mueller C, Matt P, Feil R, Stitt M** (2006) Impact of the C-N status on the amino acid profile in tobacco source leaves. *Plant Cell Environ* **29**: 2055–2076
- Gerhardt R, Heldt HW** (1984) Measurement of subcellular metabolite levels in leaves by fractionation of freeze-stopped material in non aqueous media. *Plant Physiol* **75**: 542–547
- Gibon Y, Bessieres MA, Larher F** (1997) Is glycine betaine a non-compatible solute in higher plants that do not accumulate it? *Plant Cell Environ* **20**: 329–340
- Gilles R** (1997) “Compensatory” organic osmolytes in high osmolarity and dehydration stresses: history and perspectives. *Comp Biochem Physiol A* **117**: 279–290
- Gong Q, Li P, Ma S, Rupassara SI, Bohnert HJ** (2005) Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. *Plant J* **44**: 826–839
- Hall JL, Harvey DMR, Flowers TJ** (1978) Evidence for the cytoplasmic localization of betaine in leaf cells of *Suaeda maritima*. *Planta* **140**: 59–62
- Hanson AD** (1992) Compatible solute synthesis and compartmentation in higher plants. In CN Romero, CB Osmond, CL Bolis, eds, *Water and Life*. Springer-Verlag, Berlin, pp 52–60
- Hanson AD, Rathinasabapathi B, Rivoal J, Burnet M, Dillon MO, Gage DA** (1994) Osmoprotective compounds in the *Plumbaginaceae*: a natural experiment in metabolic engineering of stress tolerance. *Proc Natl Acad Sci USA* **91**: 306–310
- Hare PD, Cress WA, Van Staden J** (1998) Dissecting the roles of osmolyte accumulation during stress. *Plant Cell Environ* **21**: 535–553
- Hare PD, Cress WA, van Staden J** (2002) Disruptive effects of exogenous proline on chloroplast and mitochondrial ultrastructure in *Arabidopsis* leaves. *S Afr J Bot* **68**: 393–396
- Hoagland DR, Arnon DI** (1938) The water culture method for growing plants without soil. *Calif Agric Exptl Stat Circ* **98**: 1105–1114
- Hoekstra FA, Golovina EA, Buitnik J** (2001) Mechanism of plant desiccation tolerance. *Trends Plant Sci* **6**: 431–438
- Jamaux I, Steinmetz A, Belhassen E** (1997) Looking for molecular and physiological markers of osmotic adjustment in sunflower. *New Phytol* **137**: 117–127
- Lalonde S, Boles E, Hellman H, Patrick JW, Frommer WB** (1999) The dual function of sugar carriers: transport and sugar sensing. *Plant Cell* **11**: 707–726
- Larher F, Hamelin J** (1975) L'acide  $\beta$ -triméthylaminopropionique des rameaux de *Limonium vulgare* Mill. *Phytochemistry* **14**: 205–207
- Larher F, Quéméner B, Hervochon P** (1990) L'ajustement osmotique pendant la vie végétative de *Cicer arietinum*. *C R Acad Sci III* **312**: 55–61
- Larher FR, Gagneul D, Deleu C, Bouchereau A** (2007) The physiological functions of nitrogenous solutes accumulated by higher plants subjected to environmental stress. *Acta Hort* **729**: 33–41
- Leidreiter K, Kruse A, Heineke D, Robinson DG, Heldt HW** (1995) Subcellular volumes and metabolite concentrations in potato (*Solanum tuberosum* cv. Désirée) leaves. *Bot Acta* **108**: 439–444
- Leigh RA, Ahmad N, Wyn Jones RG** (1981) Assessment of glycinebetaine and proline compartmentation by analysis of isolated beet vacuoles. *Planta* **153**: 34–41
- Manetti C, Bianchetti C, Casciani L, Castro C, Di Cocco ME, Miccheli A, Motto M, Conti F** (2006) A metabolomic study of transgenic maize (*Zea mays*) seeds revealed variations in osmolytes and branched amino acids. *J Exp Bot* **57**: 2613–2625
- Matoh T, Watanabe J, Takahashi E** (1987) Sodium, potassium, chloride, and betaine concentrations in isolated vacuoles from salt-grown *Atriplex gmelini* leaves. *Plant Physiol* **84**: 173–177
- Mimura T, Hura-Hotta M, Tsujimura T, Ohnishi M, Miura M, Okazaki Y, Mimura M, Maeshima M, Washitani-Nemoto S** (2003) Rapid increase of vacuolar volume in response to salt stress. *Planta* **216**: 397–402
- Munns R, Weir R** (1981) Contribution of sugars to osmotic adjustments in elongating and expanded zones of wheat leaves during moderate water deficit at two light levels. *Aust J Plant Physiol* **8**: 93–105
- Murata N, Mohanty PS, Hayashi H, Papageorgiou GC** (1992) Glycine betaine stabilizes the association of extrinsic proteins with the photosynthetic oxygen-evolving complex. *FEBS Lett* **296**: 187–189
- Nuccio ML, Russel BL, Nolte KD, Rathinasabapathi B, Gage DA, Hanson AD** (1999) The endogenous choline supply limits glycine betaine synthesis in transgenic tobacco expressing choline monoxygenase. *Plant J* **16**: 487–496
- Popp M, Smirnoff N** (1995) Polyol accumulation and metabolism during water deficit. In N Smirnoff, ed, *Environment and Plant Metabolism*, BIOS Scientific Publishers LTD, Oxford, pp 199–215
- Rhodes D, Hanson AD** (1993) Quaternary ammonium and tertiary sulfonium compounds in higher plants. *Annu Rev Plant Physiol Plant Mol Biol* **44**: 357–384
- Rhodes D, Nadolska-Orczyk A, Rich PJ** (2002) Salinity, osmolytes and compatible solutes. In A Läuchli, U Lüttge, eds, *Salinity: Environment-Plants-Molecules*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 181–204
- Riens B, Lohaus G, Heineke D, Heldt HW** (1991) Amino acid and sucrose content determined in the cytosolic, chloroplastic, and vacuolar compartments and in the phloem sap of spinach leaves. *Plant Physiol* **97**: 227–233
- Robarge WP, Edwards A, Johnson B** (1983) Water and waste water analysis for nitrate via nitration of salicylic acid. *Soil Sci Plant Anal* **14**: 1207–1215
- Rolland F, Moore B, Sheen J** (2002) Sugar sensing and signalling in plants. *Plant Cell* **11**: 185–205
- Romero C, Bellés JM, Vaya JL, Serrano R, Culianez-Macia FA** (1997) Expression of the yeast *trehalose-6-phosphate synthase* gene in transgenic tobacco plants: pleiotropic phenotypes include drought tolerance. *Planta* **201**: 293–297
- Schoenfeld RG, Lewellen CJ** (1964) A colorimetric method for determination of serum chloride. *Clin Chem* **10**: 533–539
- Schulze ED, Chapin FS** (1987) Plant specialization to environments of different resource availability. In ED Schulze, H Zwolfer, eds, *Potentials and Limitations in Ecosystem Analysis*. Springer Verlag, Berlin, pp 120–148
- Shabala SN, Lew RR** (2002) Turgor regulation in osmotically stressed *Arabidopsis* epidermal root cells: direct support for the role of inorganic ion uptake as revealed by concurrent flux and cell turgor measurements. *Plant Physiol* **129**: 290–299
- Smeekens S, Rook F** (1997) Sugar sensing and sugar-mediated signal transduction in plants. *Plant Physiol* **115**: 7–13
- Smirnoff N, Cumbes QJ** (1989) Hydroxyl radical scavenging activity of compatible solutes. *Phytochemistry* **28**: 1057–1060
- Sneath PHA, Sokal RR** (1973) *Numerical Taxonomy. The Principles and Practice of Numerical Classification*. W. E. Freeman and Co., San Francisco
- Stewart GR, Lee JA** (1974) The role of proline accumulation in halophytes. *Planta* **120**: 279–289
- Storey R, Wyn Jones RG** (1975) Betaine and choline levels in plants and their relationship to NaCl stress. *Plant Sci Lett* **4**: 161–168
- Sulpice R, Gibon Y, Bouchereau A, Larher F** (1998) Exogenously supplied glycine betaine in spinach and rapeseed leaf discs: compatibility or non-compatibility? *Plant Cell Environ* **21**: 1285–1292
- Sulpice R, Gibon Y, Cornic G, Larher FR** (2002) Interaction between exogenous glycine betaine and the photorespiration pathway in canola leaf discs. *Physiol Plant* **116**: 460–467
- Tipirdamaz R, Gagneul D, Duhazé C, Aïnouche A, Monnier C, Özkum D, Larher F** (2006) Clustering of halophytes from an inland salt marsh in Turkey according to their ability to accumulate sodium and nitrogenous osmolytes. *Environ Exp Bot* **57**: 139–153
- Winter H, Robinson DG, Heldt HW** (1993) Subcellular volumes and metabolite concentrations in barley leaves. *Planta* **191**: 180–190
- Wyn Jones RG, Gorham J** (2002) Intra- and inter-cellular compartmentation of ions—a study in specificity and plasticity. In A Läuchli, U Lüttge, eds, *Salinity: Environment-Plants-Molecules*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 159–180
- Wyn Jones RG, Storey R, Leigh RA, Ahmad N, Pollard A** (1977) A hypothesis on cytoplasmic osmoregulation. In E Marrè, O Ciferri, eds, *Regulation of Cell Membrane Activities in Plants*. Elsevier, Amsterdam, pp 121–136
- Yamaguchi T, Blumwald E** (2005) Developing salt-tolerant crop plants: challenges and opportunities. *Trends Plant Sci* **10**: 615–620
- Yancey PH** (2005) Organic osmolytes as compatible, metabolic and counteracting cryoprotectants in high osmolarity and other stresses. *J Exp Biol* **208**: 2819–2830
- Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN** (1982) Living with water stress; evolution of osmolyte system. *Science* **217**: 1214–1222