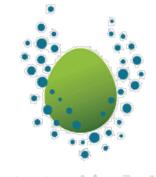


UNIVERSITE DE LILLE

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



THESE



Pour l'obtention du grade de

Docteur en Sciences de l'Université de Lille
Mention : Ingénierie des Fonctions Biologiques

*Contribution à la caractérisation du métabolisme
des phénolamides chez la chicorée : approches
biotechnologiques et édition ciblée.*

par BERNARD Guillaume

soutenue le 18 décembre 2019 devant la commission
d'examen :

Nathalie GIGLIOLI-GUIVARC'H, Professeur, Université de Tours	Rapporteur
Alain HEHN, professeur, Université de Lorraine	Rapporteur
Michèle BOITEL-CONTI, Professeur, Université de Picardie Jules Vernes	Examinateur
Nicolas SZYDLOWSKI, Chargé de recherche, CNRS, UGSF	Examinateur
Jean-Louis HILBERT, Professeur, Université de Lille	Président du jury
Caroline RAMBAUD, Maître de conférences, Université de Lille	Directeur de thèse

« La vraie sagesse, la vraie supériorité ne se gagne pas en luttant mais en laissant les choses se faire d'elles-mêmes. Les plantes qui résistent au vent se cassent, alors que les plantes souples survivent aux ouragans. » Épicure.

Résumé

Contribution à la caractérisation du métabolisme des phénolamides chez la chicorée : approches biotechnologiques et édition ciblée du génome

Cichorium intybus, communément nommée chicorée est une plante qui fut, dans un premier temps, consommée pour ses vertus thérapeutiques, puis sa culture s'est peu à peu industrialisée pour la production de substitut du café et la production d'inuline. Chez la chicorée, par une approche de génétique inverse, 2 gènes nommés *CiSHT1* et *CiSHT2*, proches des *Spermidine hydroxycinnamoyl transférases* (*SHTs*), ont été identifiés et semblent impliqués dans la voie de biosynthèse de la tetracoumaroyl spermine au niveau du tapetum des anthères de chicorée. L'expression de ces deux gènes et l'accumulation de tetrahydroxycinnamoyl spermine semblent spécifiques des plantes appartenant à la famille des Astéracées, mais le rôle de ces molécules est encore inexpliqué. Dans le but de valider la fonction de ces deux gènes, la technologie d'édition du génome : CRISPR/Cas9 a été mise en place chez la chicorée. La validation de cette technologie a été réalisée grâce à la mutation de la *phytoène désaturase* (*CiPDS*) via deux techniques de transformation, la transformation stable par *A. rhizogenes* et la transformation transitoire de protoplastes médiée par le polyéthylène glycol. La transformation stable médiée par *A. rhizogenes* étant plus efficace et la régénération de plantes plus rapide, elle a été utilisée pour générer des plantes mutantes *sht1* et *sht2*. L'étude de ces mutants a permis de valider l'implication de ces deux gènes dans la production de tetracoumaroyl spermine chez la chicorée et de montrer l'action combinée et séquentielle des deux enzymes. La production de cette molécule en systèmes homologue et hétérologues ainsi qu'une étude plus approfondie des mutants permettra peut-être de faciliter la compréhension de l'accumulation originale de cette molécule au niveau du grain de pollen, ainsi que de la valoriser à des fins thérapeutiques. De plus, la mise en place et la maîtrise de la technologie CRISPR/Cas9 chez la chicorée permettra d'étudier et d'améliorer rapidement de nombreux caractères chez cette plante cultivée.

Mots clés : BAHD, acyltransférases, phénolamides, CRISPR/Cas9, *A. rhizogenes*, *Cichorium intybus L.*

Abstract

Contribution to the characterisation of phenolamide metabolism in chicory: Biotechnological approaches and targeted genome editing

Cichorium intybus, commonly known as chicory, is a plant that was first consumed for its therapeutic properties, then its cultivation was gradually industrialized for the production of coffee substitute and inulin production. In chicory, by an inverted genetic approach, 2 genes named *CiSHT1* and *CiSHT2*, closed to the *Spermidine hydroxycinnamoyl transferases (SHTs)*, have been highlighted and seem to be involved in the biosynthesis pathway of tetracoumaroyl spermine in the chicory anther tapetum. The expression of these two genes and the accumulation of tetrahydroxycinnamoyl spermine appear to be specific to plants belonging to the Asteraceae family, but the role of these molecules is still unexplained. In order to validate the function of these two genes, genome editing technology: CRISPR/Cas9 was implemented in chicory. Validation of this technology was achieved through the edition of the phytoene desaturase gene (*CiPDS*), using *Agrobacterium rhizogenes*-mediated transformation and protoplast transfection methods. The *Agrobacterium rhizogenes*-mediated transformation being more effective and plant regeneration faster, it has been used to generate *sht1* and *sht2* mutant plants. The involvement of these two genes in the production of tetracoumaroyl spermine in chicory has been validated by the study of the mutant plants and has shown the combinatory and sequentially action of these two enzymes. The production of this molecule in homologous and heterologous systems and further study of the mutants may facilitate the understanding of the original accumulation of these molecules in pollen grains, as well as its value for industrial applications. In addition, the implementation and the mastery of CRISPR/Cas9 technology in chicory will allow to study and to quickly improve many traits in this cultivated plant.

Key words: BAHD, acyltransferases, phenolamides, CRISPR/Cas9, *A. rhizogenes*, *Cichorium intybus L.*

Remerciements

Je tiens à remercier la région des Hauts de France ainsi que l'Université de Lille pour avoir financé mon travail durant ces trois années de thèse.

Je remercie ensuite les membres du jury d'avoir accepté d'évaluer mon travail de thèse : Nathalie GIGLIOLI-GUIVARC'H et Alain HEHN en tant que rapporteurs ainsi que Michèle BOITEL-CONTI et Nicolas SZYDLOWSKI en tant qu'examinateurs.

Je remercie les membres de mon comité de thèse, Michèle BOITEL-CONTI et Fabien NOGUE, qui m'ont donné de précieux conseils et qui ont permis de mieux orienter mon projet durant nos rencontres.

Je remercie ensuite Jean-Louis HILBERT de m'avoir accueilli au sein de son équipe durant ces 3 ans et demi.

Je tiens à remercier particulièrement ma directrice de thèse Caroline RAMBAUD. Elle m'a permis de réaliser cette thèse, m'a fait confiance et m'a donné une grande liberté d'action pour la réalisation de ce projet. Merci pour votre disponibilité, vos conseils et vos précieuses corrections.

Je tiens également à remercier David GAGNEUL, pour ses conseils avisés, sa disponibilité, son implication et sa collaboration dans de nombreuses parties de ce projet.

Merci également à Marianne DELPORTE pour m'avoir initié à la recherche sur les SHTs et de m'avoir conseillé et aidé lors de nombreuses manipulations durant mon stage et ma première année de thèse.

Je remercie aussi tous les membres du 3ème étage du SN2, plus particulièrement Nasser Bahrman et Thierry Cadalen pour leur conseil et les bons moments passés lors des déjeuners quotidiens. Merci à Harmony, Amandine et Audrey pour leur contribution à ce travail.

Merci à mon co-loc Vincent de m'avoir supporté pendant ces 3 ans ainsi que pour sa bonne humeur et son amitié.

Enfin je tiens à remercier mes parents, ma sœur, ma famille et mes amis, pour leur soutien et leur présence dans les moments difficiles mais surtout pour les bons moments partagés lors de mes retours à la maison.

Sommaire

<i>Liste des abréviations</i>	1
<i>Liste des figures</i>	3
<i>Liste des tableaux</i>	8
INTRODUCTION GENERALE	10
CHAPITRE I : CARACTERISATION DES ENZYMES CISHT1 ET CISHT2 IMPLIQUEES DANS LA SYNTHESE DES PHENOLAMIDES ACCUMULEES DANS LE MANTEAU POLLINIQUE DE LA CHICOREE ET DES ASTERACEES.....	40
<i>Avant-propos</i>	42
A BAHD NEOFUNCTIONALIZATION PROMOTES TETRAHYDROXYCINNAMOYL SPERMINE ACCUMULATION IN THE POLLEN COATS OF THE ASTERACEAE FAMILY	46
1. INTRODUCTION.....	48
2. RESULTS.....	52
2.1. Identification and characterization of pollen coat phenolamides in <i>C. intybus</i>	52
2.2. Identification of tetrahydroxycinnamoyl spermine as a conserved constituent of pollen coat in the Asteraceae.....	54
2.3. Identification of candidate genes for tetracoumaroyl spermine biosynthesis.....	56
2.4. CiSHT1 and CiSHT2 expression patterns are correlated with phenolamide accumulation in chicory flowers.....	56
2.5. CiSHT2 shows a net <i>in vitro</i> activity towards the production of tetracoumaroyl spermine.....	58
2.6. Overexpression of CiSHT2 in chicory hairy roots leads to production of spermine derivates	60
2.7. CiSHT2 partly complements the phenolamide deficiency in the pollen coat of the <i>Arabidopsis sht</i> mutant and also leads to the production of spermine derivates.....	62
2.8. CiSHT1 and CiSHT2 are targeted to the nucleo-cytoplasmic compartment	62
2.9. Sequence analyses indicate a specific motif shared by members of the SHT subfamily.....	64
3. DISCUSSION	67
4. MATERIEL AND METHODS	71
4.1. Plant material.....	71
4.2. Metabolite extraction.....	71
4.3. Metabolite analyses HPLC/DAD and Compound identification by UPLC/MS/MS.....	71
4.4. RNA extraction	73
4.5. Quantitative RT-PCR.....	74
4.6. Identification and cloning of SHT-like sequences from chicory.....	74
4.7. Expression, purification, and <i>in vitro</i> assays of recombinant proteins	74
4.8. Chicory hairy root generation and growing conditions	75
4.9. Complementation of the <i>Arabidopsis sht</i> mutant with CiSHT1 or CiSHT2.....	76

4.10. <i>Subcellular localization experiments</i>	76
4.11. <i>Sequence analysis</i>	77
<i>Éléments complémentaires</i>	94
CHAPITRE II : MISE EN PLACE DE LA TECHNOLOGIE CRISPR/CAS 9 CHEZ LA CHICOREE	99
<i>Avant-propos</i>	100
EFFICIENT GENOME EDITING USING CRISPR/CAS9 TECHNOLOGY IN CHICORY	105
1. INTRODUCTION.....	107
2. RESULTS.....	110
2.1. <i>Prediction of U6 Promoter Sequences in Chicory</i>	110
2.2. <i>Target Selection and Plasmid Construction for the CRISPR/Cas9 System</i>	111
2.3. <i>Regeneration of CRISPR-Edited plants from protoplasts and from Hairy Roots lines</i>	113
2.4. <i>Identification of mutation events in genome edited plants</i>	114
2.5. <i>Transgene detection in albino plants</i>	117
3. DISCUSSION	118
4. MATERIALS AND METHODS	121
4.1. <i>Plant material</i>	121
4.2. <i>Sequence identification and gRNA design</i>	121
4.3. <i>Vector construction</i>	121
4.4. <i>rhizogenes transformation</i>	122
4.5. <i>Hairy root induction and maintenance</i>	123
4.6. <i>Protoplast isolation, transformation, and culture</i>	123
4.7. <i>DNA extraction, detection of mutation, and CAS9 amplification</i>	124
CHAPITRE III : MISE EN PLACE DE MUTANTS SHT1 ET SHT2 GRACE A L'UTILISATION DE LA TECHNOLOGIE CRISPR/CAS 9.....	137
<i>Avant-propos</i>	138
FULLY SUBSTITUTED SPERMINE-DERIVED PHENOLAMIDE SYNTHESIS REQUIRES SEQUENTIAL ACTION OF TWO BAHD ACYLTRANSFERASES IN CHICORY FLOWER BUDS	143
1. INTRODUCTION.....	144
2. RESULTS.....	147
2.1. <i>Generation of Cisht1 and Cisht2 chicory mutants</i>	147
2.2. <i>Metabolic analysis of genome edited plants</i>	149
2.3. <i>Transient expression of CiSHT1 and CiSHT2 in tobacco leaves</i>	150
2.4. <i>Reconstitution of the tetracoumaroyl spermine biochemical pathway in <i>S. cerevisiae</i></i>	151
2.5. <i>Polyamine availability controls the nature of the final product</i>	153
2.6. <i>The main CiSHT2 products is N⁵,N¹⁰-dicoumaroyl spermine</i>	153

3.	DISCUSSION	156
4.	MATERIALS AND METHODS	160
4.1.	<i>Plant material</i>	160
4.2.	<i>Design of sgRNA, construction of CRISPR/Cas9 vectors and A. rhizogenes transformation</i>	160
4.3.	<i>Hairy root culture and selection</i>	161
4.4.	<i>Genomic DNA extraction and High-Resolution Melting Temperature analysis</i>	162
4.5.	<i>Phenolic acid extraction</i>	162
4.6.	<i>Metabolite analyses by HPLC/DAD</i>	162
4.7.	<i>Extraction and analysis of free polyamines</i>	163
4.8.	<i>Transient expression in Nicotiana benthamiana</i>	163
4.9.	<i>Stable expression in Saccharomyces cerevisiae</i>	163
4.10.	<i>Dicoumaroyl spermine purification</i>	164
4.11.	<i>ESI-HRMS and NMR Analysis</i>	165
	<i>Éléments complémentaires</i>	175
1.	MATERIEL ET METHODES	175
1.1.	<i>Complementation du mutants sht d'Arabidopsis avec CiSHT1 et CiSHT2</i>	175
1.2.	<i>Génération de chevelus racinaires de chicorée surexprimant CiSHT1 et CiSHT2</i>	176
1.3.	<i>Addition de spermine et de phénylalanine dans le milieu de culture</i>	176
2.	RESULTATS ATTENDUS.....	176
	CONCLUSIONS ET PERSPECTIVES	179
	Références bibliographiques.....	190

Liste des abréviations

2,4-D	acide 2,4-dichlorophénoxyacétique
4CL	4 coumarate :CoA ligase
AA	Acide Aminé
C3'H	<i>p</i> -coumarate 3' -hydroxylase
C4H	cinnamate 4 -hydroxylase
Cas9	CRISPR associated protein 9
CoA	Coenzyme A
CQA	Caffeoyl quinic acid = acide chlorogénique
CTA	caffeoil tartric acid = acide caftarique
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
DH29	spermidine acyltransférase
diCQA	dicaffeoyl quinic acid = acide isochlorogénique
diCTA	dicaffeoyl tartric acid = acide chicorique
DNA	DesoxyriboNucleic Acid
FD	Florimond Desprez
HCT	Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase
HDR	Homologous Directed Repair
HPLC	High Performance Liquid Chromatography
HQT	Hydroxycinnamoyl-CoA quinate/shikimate hydroxycinnamoyltransferase
IPP	isopentenyl diphosphate
kDa	Kilo Dalton
KO	Knock Out
LB	Left Border
MeJa	Methyl Jasmonate
NHEJ	Non Homologous End Joining
OMT	O-methyl transferase
ORF	Open Reading Frame

PAL	Phénylalanine Ammonia Lyase
PAM	protospacer-adjacent motif
PDS	Phytoene desaturase
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
qRT-PCR	quantitative real-time PCR
QTL	Quantitative Trait Loci
RB	Right Border
RNA	RiboNucleic Acid
RNP	ribonucléoprotéique
SDT	Spermidine DisinapoylTransférase
SCPL	Serine CarboxyPeptidase-Like
SCT	Spermidine DicoumaroylTransférase
sgRNA	single guide RNA
SHT	Spermidine Hydroxycinnamoyl Transferase
TAL	Tyrosine Ammonia Lyase
TALENs	Transcription Activator-Like Effector Nuclease
T-DNA	Transfer DNA
UV	Ultraviolet
ZFNs	Zinc Finger Nucleases

Liste des figures

Introduction

Figure 1 : Plante botanique de *Cichorium intybus*.L.

Figure 2 : Espèces appartenant au genre *Cichorium* cultivées en Europe et leurs utilisations industrielles majoritaires.

Figure 3 : Expérience de Skoog et Miller (1953) et son interprétation sur le contrôle du développement des végétaux en culture *in vitro*.

Figure 4 : Structure schématique des plasmides Ri des différentes souches d'*Agrobacterium rhizogenes* d'après (Veena et Taylor 2007)

Figure 5 : Représentation schématique du T-DNA présent sur les plasmides Ri de type Agropine chez *Agrobacterium rhizogenes* (Ozyigit, Dogan and Tarhan, 2013)

Figure 6 : Organigramme pour la mise en place de chevelus racinaires ainsi que quelques applications (d'après Georgiev et al. 2007)

Figure 7 : Représentation générale des voies de biosynthèses des métabolites spécialisés.

Figure 8 : Stratégie pour la production de métabolites spécialisés grâce à la culture cellulaire (d'après Bourgaud et al 2001)

Figure 9 : Voie de biosynthèse de l'acide chlorogénique

Figure 10 : Réactions d'acylation catalysées par les enzymes de la famille des BAHD et des SCPL. (D'après Bonpart et al. 2015)

Figure 11 : Représentation du sous-clade Vb de l'arbre phylogénétique des BAHD.

Chapitre I

Figure 1.1 : Accumulation de trihydroxycinnamoyl spermidine dans le manteau pollinique des spermaphytes.

Figure 1.2 : Représentation schématique de la connexion entre la voie de biosynthèse des polyamines et le métabolisme des acides hydroxycinnamiques au niveau du manteau pollinique chez *Arabidopsis*.

Figure 1.3 : Activité *in vitro* d'AtSHT (A) et MdSHT (B) en présence de différents esters de CoA (D'après Grienberger et al. 2009 et Elejalde-Palmett et al. 2015).

Article I: A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in the pollen coats of the Asteraceae family

Figure 1: Identification of phenolamides present in chicory floral buds.

Figure 2: Occurrence of tetrahydroxycinnamoyl spermine in floral buds among Asteraceae and close families.

Figure 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.

Figure 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.

Figure 5: Expression and purification of recombinant *CiSHT1* and *CiSHT2*

Figure 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).

Figure 7: Subcellular imaging of yellow fluorescent protein (YFP)-tagged constructs transiently expressed in *N. benthamiana* leaves.

Figure 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

Supplementary Figure S1: Proposed fragmentation scheme of the compound designated as C1 (A) and its UV spectra (B).

Supplementary Figure S2: LC/MS/MS fragmentation of the compound designated as C2 (A) and its UV spectra (B).

Supplementary Figure S3: LC/MS/MS fragmentation of the compound designated as C3 (A) and its UV spectra (B).

Supplementary Figure S4: LC/MS/MS fragmentation of spermine derivatives found in Asteraceae flower buds.

Supplementary Figure S5: Subcellular imaging of YFP-tagged constructs stably expressed in *A. thaliana*.

Figure 1.4: représentation des différentes types monocoumaroyl spermidines (A ; B ; C) et monocoumaroyl spermine (1 ; 2).

Figure 1.5: (A) Analyses HPLC des chevelus racinaires sur-exprimant SHT2 après différentes conditions de cultures. (B) Evaluation des quantités de dérivés de spermine et de spermidine accumulés dans les chevelus racinaires de chicorée dans différentes conditions de culture.

Chapitre II

Figure 2.1: Représentation schématique de l'édition d'un gène grâce à la technologie CRISPR/Cas 9. (D'après Zhang et al 2016)

Figure 2.2: Représentation des voies de biosynthèses des caroténoïdes, de la chlorophylle et des gibberellines (GAs). La phytoène désaturase est encadrée en rouge. (D'après Qin et al. 2007)

Figure 2.3: Représentation schématique de la stratégie d'édition d'un gène en utilisant 2 sgRNAs et quelques événements de mutation possibles.

Article II: Efficient genome editing in chicory

Figure 1: Chicory U6 and *Arabidopsis* U6-26 gene sequence partial alignment.

Figure 2: Overview of the experimental design for *CiPDS* disruption.

Figure 3: Phenotype of genome edited plants.

Figure 4: Sequence analysis of the two alleles of the 10 albino hairy root lines.

Figure 5: Sequence analysis of the two alleles of 9 albino calli obtained after protoplast transformation.

Supplementary Figure S1: *CiPDS* DNA sequence, sgRNA target sites and the primer sites for PCR amplification of the fifth exon.

Supplementary Figure S2: DNA amplification of albino plant regenerated from hairy root lines by PCR with specific primer pair (E5-F and E5-R) surrounding the two target sites in the fifth exon of the *CiPDS*

Supplementary Figure S3: Amplification by PCR of DNA extracted from albino plants regenerated from calli with specific primer pair (E5-F and E5-R) surrounding the two target sites in the fifth exon of the *CiPDS*.

Supplementary Figure S4: Predicted amino acid sequences of the *CiPDS* from the 10 albino hairy root lines and their shoots.

Supplementary Figure S5: Predicted amino acid sequence of the CiPDS from the 9 albino calli and their shoots

Supplementary Figure S6: Partial sequence of the DNA fragments inserted at the target 2 site of the C8 mutant.

Supplementary Figure S7: DNA amplification of albino plants regenerated from calli by PCR with specific primer pair (C9-F and C9-R) hybridizing in CAS9 gene.

Supplementary Figure S8: Sequence of the cassette “CiU6-PlacZ ccdb-sgRNAscaffold” harboring by the plasmid pKanCiU6-1p-sgRNAscaffold.

Chapitre III

Article III: Fully substituted spermine-derived phenolamide synthesis requires sequential action of two BAHD acyltransferases in chicory flower buds

Figure 1: Schematic diagram of *CiSHT1* and *CiSHT2* gene structures and of vectors used to generate mutants.

Figure 2: Representative genotypes of *sht1* and *sht2* mutants.

Figure 3: Stacked HPLC chromatograms of extracts obtained from WT, *sht1*, *sht2* flower buds at stage 13-15.

Figure 4: Stacked HPLC chromatograms of extracts obtained from agroinfiltrated tobacco leaves.

Figure 5: Stacked HPLC chromatograms of extracts from transgenic yeast expressing different combination of genes.

Figure 6: LC/MS/MS spectra of purified N⁵, N¹⁰-dicoumaroyl spermine.

Figure 7: Chemical structure of N⁵, N¹⁰-dicoumaroyl spermine as deduced from NMR data.

Supplementary Figure S1: *CiSHT1* DNA sequence, sgRNA target sites and the primer sites for sequencing and HRM analysis.

Supplementary Figure S2: *CiSHT2* DNA sequence, sgRNA target sites and the primer sites for sequencing and HRM analysis.

Supplementary Figure S3: Mutation events occurred in *SHT1* gene at target 2 localisation in plants transformed with pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2.

Supplementary Figure S4: Representative genotype and predicted amino acid sequences of the double mutant *sht1/sht2* obtained with pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2.

Supplementary Figure S5: Predicted amino acid sequences of the *SHT1* and *SHT2* genes

Supplementary Figure S6: ^1H NMR spectra of $\text{N}^5, \text{N}^{10}$ -dicoumaroyl spermine.

Conclusion et perspectives

Figure 4.1: Représentation schématique de la connexion entre la voie de biosynthèse des polyamines et le métabolisme des acides hydroxycinnamiques au niveau du manteau pollinique chez la chicorée.

Figure 4.2: Analyse des séquences de certaines enzymes capable d'acyler les polyamines.

Liste des tableaux

Introduction

Tableau 1: Utilisation traditionnelle de l'espèce *Cichorium intybus* (D'après Street et al. 2013)

Tableau 2: Gènes des opines présents sur les différents plasmides Ri d'après (Vladimirov, Matveeva, et Lutova 2015)

Tableau 3: Les oncogènes *d'Agrobacterium rhizogenes*, les protéines qu'ils encodent, leurs fonctions et leurs effets sur le phénotype de la plante hôte (D'après Ozyigit, Dogan and Tarhan, 2013)

Tableau 4: Exemple de biomolécules produites par des systèmes *in vitro* de cultures d'organes (d'après Steingroewer et al. 2013)

Tableau 5: Exemple d'augmentation de production de métabolites spécialisés par élicitation (d'après Georgiev et al 2007)

Chapitre I

Article I: A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in the pollen coats of the Asteraceae family

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

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

Supplementary Table S1: Sequence of primers used in this study.

Chapitre II

Article II: Efficient genome editing in chicory

Table 1: Mutation analysis in the *CiPDS* gene. Analysis were performed on plants regenerated from protoplasts or hairy roots.

Tableau A1: Sequences of primers used in this study.

Chapitre III

Tableau 1: Exemples d'espèces végétales dont la résistance aux pathogènes a été améliorée grâce à l'utilisation de la technologie CRISPR/Cas9 (D'après Langner, Kamoun, et Belhaj 2018).

Article III: Fully substituted spermine-derived phenolamide synthesis requires sequential action of two BAHD acyltransferases in chicory flower buds

Table 3.1: ^1H and ^{13}C -NMR data of N⁵, N¹¹-dicoumaroyl spermine in methanol-d₄.

Supplementary table T1: Primer listing used in this study

Introduction Générale

La famille des Astéracées dont une des caractéristiques est l'inflorescence en capitule, représente environ 10% des plantes à fleurs terrestres, ce qui fait d'elle une des plus grandes familles de plantes appartenant aux angiospermes. Cette famille apparue il y a environ 70 millions d'années en Amérique du sud, a réussi à coloniser l'ensemble du globe. De nos jours il existe environ 23 000 espèces au sein des Astéracées, qui ont été classées dans de nombreuses sous-familles, dont les trois majoritaires sont : les *Astroideae*, dans laquelle on trouve, par exemple, le tournesol « *Helianthus annuus* », les *Carduoideae*, dans laquelle on trouve notamment l'artichaut « *Cynara cardunculus* » et les *Cichorioideae* dans laquelle sont présentes les laitues « *Lactuca* » et les chicorées « *Cichorium* ». (Barreda et al. 2015)

Le genre *Cichorium* regroupe 6 espèces : *intybus L.*, *endivia L.*, *pumilum Jacq.*, *calvum Sch.Bip. ex Asch.*, *spinosum L.*, *bottae* Deflers, qui se trouvent en majorité en Europe et en Asie. A l'état sauvage, *C. intybus*, *C. spinosum*, et *C. bottae* sont des espèces pérennes, les 3 autres étant des espèces annuelles.

Cichorium intybus L., plus communément nommée « chicorée » est à l'état sauvage, une plante qui peut atteindre 1 mètre de haut. Sa racine pivotante et charnue, peut mesurer 20 cm de longueur et a un diamètre d'environ 3 cm. Les feuilles inférieures sont divisées en dents aigües, les supérieures sont petites et allongées. De longues tiges portent de fins rameaux rigides formant un angle obtus avec la tige. Les fleurs sont bleues, ligulées en capitules. La chicorée possède un mode de pollinisation entomophile et favorise l'allofécundation, grâce à sa morphologie florale défavorable à l'autofécundation, la mise en place d'un système d'auto-incompatibilité sporophytique (Gonthier et al. 2013) ainsi que la mise en place d'un phénomène de compétition pollinique

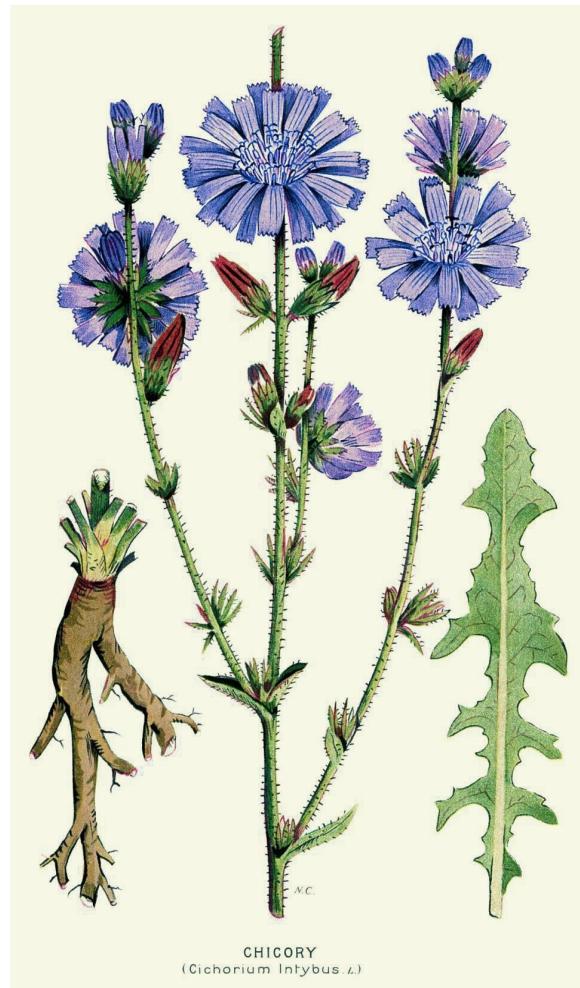


Figure 1 : Plante botanique de *Cichorium intybus*.L.

(Eenink 1982). Le fruit produit est un akène sans aigrette, couronné de minuscules écailles. La chicorée sauvage pousse couramment au bord des champs, dans les prairies et les fossés.

Les premières traces écrites mentionnant l'utilisation de la chicorée par l'homme ont été retrouvées sur le papyrus Ebers datant d'environ 1600 avant Jésus-Christ. En effet, les bienfaits culinaires de la chicorée y sont déjà retranscrits et les Égyptiens semblaient la consommer avant le repas pour faciliter la digestion (Short et al. 2015). Le poète latin Horace qui vécut au premier siècle avant Jésus-Christ cite la chicorée dans les Odes, « *Me pascunt olivae, Me cichorea, levesque malvae* », pouvant être traduit par « Pour moi, je me suffis d'olives, de chicorée et mauves légères » de nos jours s'apparentant à « vivre d'amour et d'eau fraîche ». Plus tard les Grecs et les Romains la considèrent comme une alliée pour guérir les maladies digestives et hépatiques. Ces vertus sont vantées dans *De materia medica*, ouvrage botanique écrit par le médecin grec Dioscoride (1er siècle avant J-C). Pline l'Ancien la mentionne également dans le 20ème livre, chapitre XXIX de son œuvre *Naturalis historia*. Par l'intermédiaire de leurs conquêtes, les romains vont la propager dans tout leur empire et c'est comme cela que cette plante méditerranéenne va, petit à petit, coloniser le nord de l'Europe.

Plus tard, en 812, Charlemagne promulgue le capitulaire *De Villis*, dans lequel il ordonne qu'un certain nombre de plantes agronomiques et/ou thérapeutiques soient cultivées dans les jardins des domaines impériaux et la chicorée en fait partie. Dans un même temps, en Hollande, des moines commencent à l'améliorer par sélection massale. Petit à petit, la chicorée devient de plus en plus populaire et durant la renaissance, il semble que toutes les parties de la plante soient utilisées pour leurs vertus thérapeutiques, que ce soit en Europe ou en Asie (Tableau 1).

Mais dans les siècles qui suivirent, les qualités médicinales de la chicorée sont peu à peu délaissées. En France au début des années 1600, C. Mollet la mentionne comme une alternative à la salade verte dans son livre « Théâtre des plans et jardinages » dans lequel il dit « Étiolées, les feuilles de chicorée sauvage sont excellentes en salade ». Ce n'est qu'en 1690 que les hollandais recommencent à l'utiliser, non pas pour ses vertus thérapeutiques mais comme substitut du café. La torréfaction industrielle commence et le breuvage se popularise dans le nord de l'Europe. C'est au début des années 1800 que la culture de la chicorée va connaître un essor considérable. En effet, après la bataille de Trafalgar en 1805, la flotte française de

Napoléon Bonaparte est décimée et les Anglais règnent en maîtres sur les océans et les mers du globe. Pour pallier cela, et essayer de fragiliser l'économie anglaise, Napoléon va instaurer le blocus de Berlin, en 1806, qui interdit tout commerce avec le royaume d'Angleterre.

Country	Traditional use(s)	Plant part(s)	Preparation(s)
Afghanistan	Malaria	Root	Aqueous extract
Bosnia and Herzegovina	Diarrhea, strengthening the prostate and other reproductive organs, pulmonary cancer, hangover, and purification of biliary tract	Aerial part, flowers, roots	Not stated
Bulgaria	Liver disorders, spasmolytic, cholesterol, antiseptic	Aerial	Decoction
India	Cholagogue stimulant for gastric secretion, hypoglycemic	Roots, aerial parts	Decoction
	Liver disorders	Seeds	
	Diabetes	Whole plant	Not stated
	Jaundice, liver enlargement, gout, and rheumatism	Root	Decoction
	Cough relief	Not stated	
Iran	Euppeptic, stomachic, depurative, choleretic, laxative, hypotension, tonic, and antipyretic	Whole plant	Not stated
	Blood cleansing	Leaves	Not stated
	High blood pressure	Leaves	Decoction
Italy	Blood purification, arteriosclerosis, antiarthritis, antispasmodic, digestive	Leaves/roots	Decoction
	Depurative	Whorls	Decoction
	Choleretic, hepatoprotective against jaundice, mild laxative, hypoglycemic	Leaves	Decoction, squashed fresh leaves
Jordan	Internal hemorrhage, sedative in typhoid	Whole plant	Cooking
Morocco	Renal disease	Aerial/roots	Not stated
	Kidney disorders, diabetes	Whole plant	Decoction
Pakistan	Diabetes	Roots	Decoction
Poland	Digestive complaints and lack of appetite	Roots	Tea
	Diarrhea	Flower	Infusion
Serbia	Diuretic, digestive, laxative, anti-inflammatory, liver complaints, reducing blood sugar	Roots	Decoction/tea
	Cholagogue, digestive, hypoglycemic	Aerial part/root	Not stated
South Africa	Jaundice, tonic	Leaves, stems, roots	
Turkey	Cancer, kidney stones	Roots	Decoction
	Wound healing	Leaf	Ointment
	Hemorrhoids, urinary disorders	Aerial	Tea

Tableau 1 : Utilisation traditionnelle de l'espèce *Cichorium intybus* (D'après Street et al. 2013)

La France va donc manquer de certaines matières premières provenant des colonies, comme le sucre, le coton et le café. La culture de la betterave va être développée pour la production de sucre et, dans une moindre mesure, la culture de la chicorée pour la production du substitut de café (Kandeler et Ullrich 2009). La culture de la chicorée va donc s'industrialiser ce qui va favoriser la recherche agronomique liée à cette plante. La chicorée industrielle *Cichorium intybus* L. var. *sativum* va petit à petit émerger de la recherche agronomique ainsi que de nouveaux cultivars sélectionnés pour leurs potentiels, dont le plus anciennement connu : « la chicorée de Magdebourg ». En 1853, par un pur hasard, un agriculteur de Bruxelles découvrit que si on laissait des racines de chicorée dans le noir, elles développaient des feuilles blanches imbriquées les unes dans les autres, la chicorée « Witloof » (terme flamand signifiant « feuilles blanches »), *Cichorium intybus* L. var. *foliosum* venait d'être inventée.

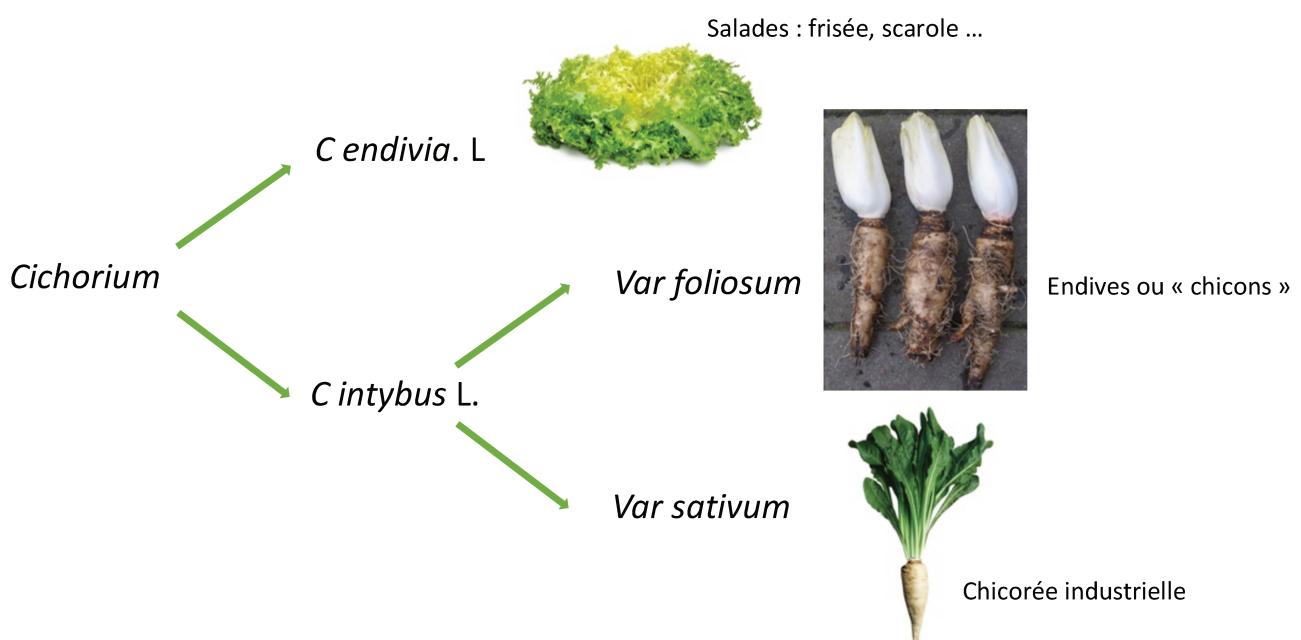


Figure 2 : Espèces appartenant au genre *Cichorium* cultivées en Europe et leurs utilisations industrielles majoritaires.

La culture de la chicorée va surtout se développer en France, en Belgique, en Allemagne et en Autriche. Entre 1872 et 1940, les surfaces cultivées de chicorées vont être multipliées par 15. Cette forte croissance étant due en partie à la capacité de la chicorée à croître dans des sols infestés de nématodes impropre à la culture de la betterave.

La chicorée industrielle est une plante bisannuelle, son cycle de développement est donc constitué de deux phases : une phase végétative, accomplie durant la première année, pour laquelle, après germination, la plante forme une rosette de feuilles mettant ensuite en place une racine tubérisée où des réserves azotées et carbonées seront stockées pour le passage de la mauvaise saison. Au printemps, après avoir subi cette étape de vernalisation, la phase reproductrice intervient. Les réserves vont être remobilisées pour la mise en place d'une hampe florale qui portera les futurs boutons inflorescentiels, cette étape est nommée montaison. Puis les boutons vont fleurir, être fécondés et donner la nouvelle semence. Pour augmenter les rendements, les agronomes vont essayer d'allonger la phase végétative. Pour cela la résistance à la montée à graine va être un caractère primordial à améliorer. En effet ce caractère est essentiel pour les zones du nord de l'Europe où le risque de vernalisation par le froid après semis est redouté. Les sélectionneurs vont surtout améliorer ce caractère par sélection massale, puis avec l'évolution des biotechnologies, notamment la culture *in vitro*, les scientifiques vont étudier et essayer de comprendre les mécanismes liés à ce phénomène.

Au début du 20ème siècle, le biologiste Allemand Haberlandt, énonce pour la première fois la propriété appelée totipotence cellulaire végétale. Cette propriété démontre que toutes les cellules végétales, quels que soient le tissu et l'organe dont elles proviennent, possèdent la capacité de régénérer un individu complet et identique à la plante mère. Cette capacité repose sur l'aptitude des cellules à la dédifférenciation. Des cellules différenciées peuvent redevenir des cellules non spécialisées, se multiplier et ensuite se spécialiser à nouveau en tous les différents types de cellules nécessaires à la création d'un nouvel individu. Cette propriété va devenir la base de la culture *in vitro* des cellules végétales. En 1926, le chercheur Went découvre l'auxine, qu'il qualifie comme une substance de croissance étant responsable de la spécialisation des cellules, la recherche sur les phytohormones est lancée. En 1939, Gautheret et White obtiennent à partir de tissu de carotte, un amas de cellules dédifférenciées : un cal, qui peut être cultivé indéfiniment dans le temps. Avec cette nouvelle découverte, démarre vraiment la culture *in vitro* ("dans du verre"). C'est après la seconde guerre mondiale que la culture *in vitro* va le plus évoluer. En 1956, Miller met au jour l'existence et les propriétés d'autres phytohormones de croissance, les cytokinines. Les phytohormones provoquent la division cellulaire et contrôlent la rhizogénèse ainsi que la caulogénèse. Grâce à cette découverte, Skoog et Miller,

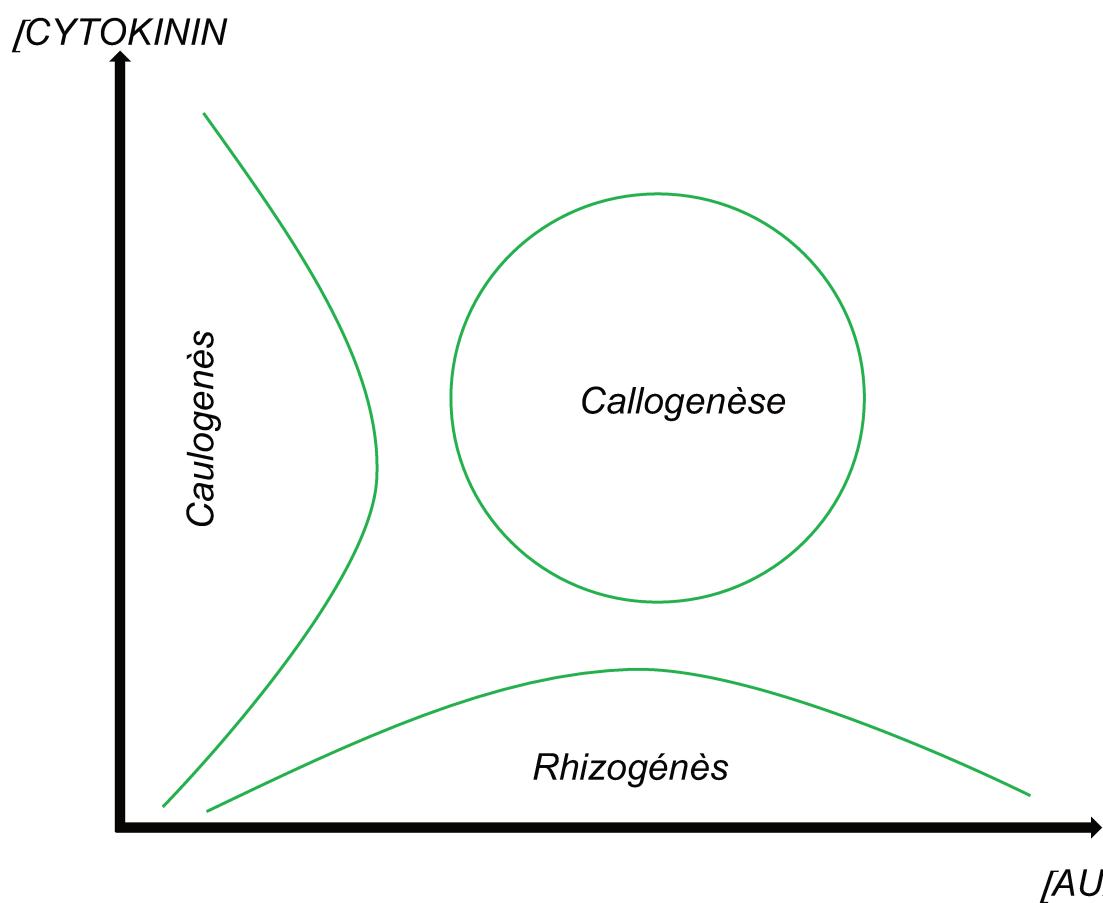
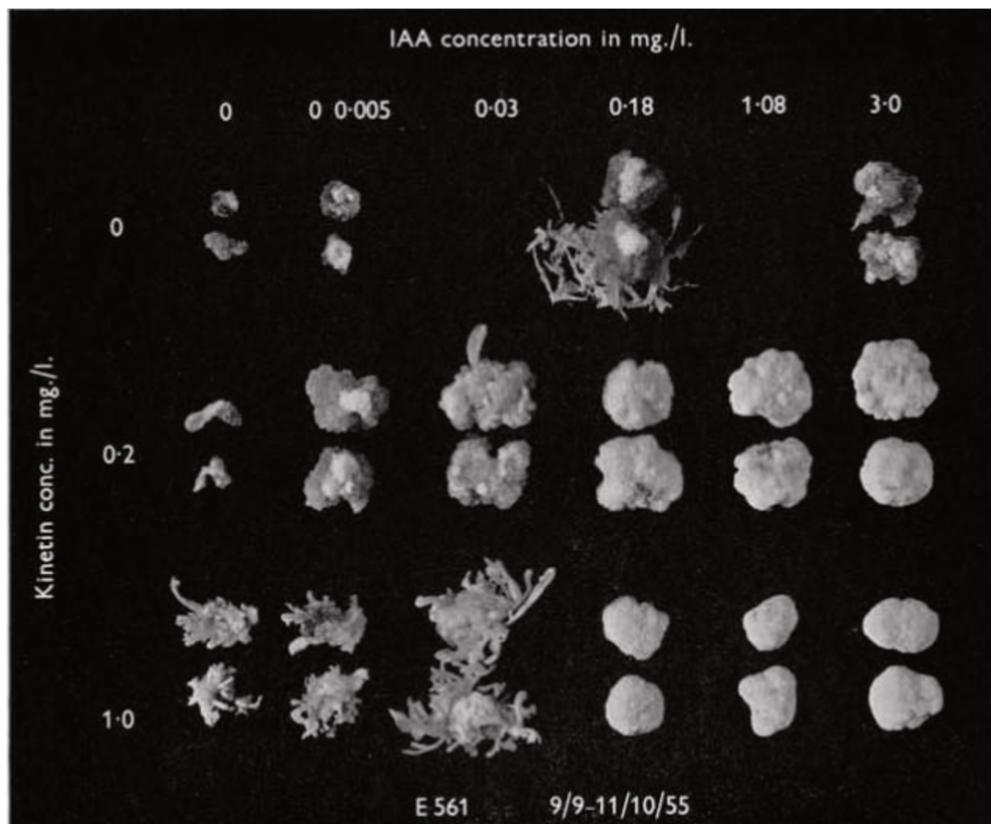


Figure 3 : Expérience de Skoog et Miller (1953) et son interprétation sur le contrôle du développement des végétaux en culture *in vitro*.

en 1957, vont mettre en évidence l'influence des auxines et des cytokinines sur les cellules cultivées *in vitro* (Figure 3) et montrer que les balances hormonales sont indispensables au développement d'une plante et que les hormones permettent les capacités de régénération des cultures. C'est en 1962 que le premier milieu de culture *in vitro* de base (auxines, cytokinines, sucres, minéraux) est créé par Murashige et Skoog (MS), lors de l'étude de la croissance de cultures de tissus de tabac (Murashige et Skoog 1962). Ce milieu va être à la base de pratiquement tous les milieux de culture utilisés en culture *in vitro*. Chez *C. intybus*, les premiers travaux de culture *in vitro* ont porté sur la régénération de plantes à partir de fragment de méristèmes racinaires et foliaires préalablement traités ou non par le froid (Nitsch et Nitsch 1964). Dans cette étude les époux Nitsch obtiennent pour la première fois des fleurs mises en place par des plantes régénérées à partir de tissus racinaires et foliaires cultivés sur du milieu MS modifié. Ces travaux permettront aussi l'étude des mécanismes liés à la vernalisation, notamment l'intervention des cytokinines dans ce mécanisme (Nitsch et al. 1966). Dans un même temps, Vasil et Hildebrandt, deux pionniers de la culture cellulaire *in vitro* étudient la production de biomasse à partir de culture de cals de nombreuses espèces. Ils pensaient que ces cultures pourraient devenir de potentielles sources nutritionnelles pour les astronautes (Vasil et Hildebrandt 1965). En 1966, ils mettent au point des milieux de culture à base de lait de coco, d'acide 2,4-dichlorophénoxyacétique (2,4-D) et d'acide naphtalène acétique (ANA) pour induire la callogénèse sur des embryons matures de chicorée (Vasil et Hildebrandt 1966). Gwozdz va par la suite analyser les effets de l'auxine IAA sur la croissance et l'organogénèse de cals provenant de racines de chicorée et montrera que l'IAA induit une forte croissance des cals ainsi que l'augmentation de la production d'ARN (Gwozdz 1973). La régénération de plantes à partir de cals de racine sera réalisée par Caffaro et al (1982) en utilisant du milieu de Heller (Heller 1953) additionné avec du 2,4-D (Profumo et al. 1985). La chicorée possède donc la capacité d'être multipliée par embryogenèse somatique, ce qui permet de la multiplier indéfiniment *in vitro*. Les plantes produites possèdent la même information génétique que la plante mère, ce sont des clones. Cette technique appelée multiplication végétative *in vitro* sera très importante dans les programmes de sélection pour la conservation de génotypes particuliers. Mais les inconvénients du clonage végétal sont, paradoxalement, les mêmes que ses avantages : la reproduction du génome à l'identique. Or, le brassage de gènes par l'intermédiaire de la reproduction sexuée est la base de la production de

nouveaux génotypes, phénomène très important pour l'évolution des espèces et leur adaptation à un environnement versatile.

Chez les cellules animales, depuis les années 1960 et les expériences de Georges Barski, il était possible de fusionner deux souches de cellules somatiques différentes. La fusion de cellules humaines avec des cellules de souris produisait des cellules hybrides possédant les unités cellulaires et le matériel génétique des deux cellules mères. Ces hybrides normalement impossibles à produire par voie sexuelle à cause des barrières génétiques, étaient viables et avaient la capacité de se multiplier. Ne faisant pas intervenir les gamètes, ce type de cellule fut nommé « hybride somatique » (Barski et al. 1960). Les frontières de l'hybridation sexuée transgressées, permettaient d'ouvrir de nouvelles perspectives en biologie expérimentale. Mais comparées aux cellules animales, les cellules végétales possèdent une paroi de nature pectocellulosique qui empêche tout phénomène de fusion. S'affranchir de cette barrière, associé au caractère totipotent des cellules végétales, pourrait permettre la production d'organismes hybrides entre des espèces végétales éloignées et donc la création de nouvelles espèces ou variétés. L'anglais Cocking, en 1960, est le premier à obtenir des cellules de racines de tomate (*Lycopersicum esculentum*) sans paroi en utilisant une cellulase fongique de *Myrothecium verrucaria* (Cocking 1960). Des protoplastes ont par la suite été isolés de nombreuses espèces de plantes. Le type d'explants, feuilles, cotylédons, hypocotyles, cals ou suspensions cellulaires, leur âge, le type de milieu et la solution enzymatique utilisés, se sont avérés comme des paramètres très importants pour la maîtrise de l'isolement de protoplastes chez les différentes espèces de plantes. Une fois leurs parois digérées, les protoplastes ne sont plus protégés des différences d'osmolarité entre les milieux extracellulaire et intracellulaire, le milieu doit donc aussi agir comme stabilisateur osmotique, en plus d'intervenir pour pallier les exigences nutritives impliquées dans les processus de prolifération. Les protoplastes de nombreuses espèces de plantes ont été isolés dans des milieux assez pauvres, osmotiquement ajustés avec des sucres et supplémentés avec des sources de carbone, de vitamines et d'hormones (Binding et al. 1980) puis ces milieux de culture ont été précisément adaptés en fonction du type de plante étudiée. La régénération de plantes à partir de protoplastes est aussi une étape difficile à maîtriser chez de nombreuses espèces, excepté pour les plantes de la famille des Solanaceae. Chez la chicorée, les premiers protoplastes ont été isolés en 1980 (Binding et al. 1980). En 1982, les époux Chupeau (Crépy, Chupeau, et Chupeau 1982) arrivent pour la

première fois à régénérer des plantes à partir de protoplastes de mésophylle de chicorée. La maîtrise de l'isolement et de la régénération de plantes à partir de protoplastes a permis à Rambaud et al. de réaliser la première fusion de protoplastes intraspécifique chez la chicorée pour produire des plantes tétraploïdes (Rambaud, Dubois, et Vasseur 1992). Par la suite, ce succès a ouvert les portes à une fusion intergénérique *Cichorium intybus* x *Helianthus annuus* via la technique du polyéthylène glycol (PEG). Cette hybridation somatique a permis de produire des cybrides dans lesquels la recombinaison des mitochondries a induit la stérilité male cytoplasmique. Ce matériel a été très utile aux sélectionneurs pour la production de graines de chicorées hybrides (Rambaud, Dubois, et Vasseur 1993). Une fusion interspécifique *Cichorium intybus* x *Cichorium endivia* a aussi été réalisée en 2007 (Cappelle et al. 2007). Petit à petit les expériences d'hybridations somatiques ont montré les limitations de la technique. L'importance des éliminations chromosomiques engendrées par des fusions trop audacieuses ainsi que des problèmes de régénération ont mis un frein aux projets ambitieux (Demarly et Sibi 1996).

Parallèlement à la recherche sur les biotechnologies végétales pour l'étude de la physiologie végétale et l'amélioration des plantes, les travaux sur l'ADN vont permettre d'entrevoir une possibilité pour le franchissement des barrières inter-espèces. En effet en 1953, James Watson et Francis Crick découvrent la structure en double hélice de l'ADN. Cette découverte va permettre de comprendre le fonctionnement de cette molécule fondamentale et dans les années 1960 l'universalité du code génétique va être démontrée (Nirenberg et Matthaei 1961; Nirenberg et al. 1966). La majorité des êtres vivants possède son code génétique sur la molécule d'ADN. Les scientifiques vont donc envisager que n'importe quel gène de n'importe quelle espèce va pouvoir être ajouté et fonctionner chez n'importe quelle autre espèce. Cette théorie va ouvrir la porte à de nouvelles possibilités d'amélioration génétique, notamment chez les plantes cultivées et chez les bactéries. Les premiers travaux d'ingénierie génétique *in vitro*, sont réalisés par Paul Berg, en 1972, en obtenant le premier ADN recombinant grâce à l'utilisation d'enzymes de restriction, le premier plasmide synthétique est créé (Jackson, Symons, et Berg 1972). Suite à cela les premières réussites de transformation bactérienne sont réalisées (Cohen et al. 1973).

Chez les végétaux, c'est grâce à l'étude d'une bactérie tellurique, *Agrobacterium tumefaciens*, que la première transformation va être réalisée. *Agrobacterium tumefaciens* est une bactérie responsable d'une maladie appelée la « galle du collet ».

Grâce à de nombreuses études, les chercheurs sont arrivés à identifier les mécanismes induisant cette maladie. Cette bactérie est attirée par les composés phénoliques libérés par les plantes lorsqu'elles ont été victimes d'une blessure. *Agrobacterium* a la capacité de se fixer sur les cellules accessibles au niveau de la lésion, et, à la suite de ce contact, ces cellules vont se multiplier de manière anarchique, formant une tumeur. Les cellules tumorales vont produire et libérer dans le sol des molécules que l'on nomme « opines ». Les différentes souches d'*Agrobacterium* peuvent être classées en fonction du type d'opines produites par les cellules tumorales. On trouve donc des souches à octopine, nopaline, leucinopine et succinamopine... Ces opines serviront à *Agrobacterium* de sources d'azote et de carbone. La bactérie a donc la capacité de détourner le métabolisme des cellules végétales pour leur faire produire une source d'énergie. En 1974, une découverte majoritaire permettra la compréhension moléculaire de ce phénomène. En effet cette bactérie possède un plasmide dans lequel se trouve les gènes responsables de l'induction de la galle du collet, c'est pour cela qu'il a été nommé « plasmide Ti » (Tumour-inducing) (Zaenen et al. 1974). *Agrobacterium* possède donc la capacité d'injecter un ADN dans une cellule végétale et de l'intégrer dans son génome, il réalise donc une transgénèse naturelle par l'intermédiaire de son plasmide Ti. Le plasmide Ti d'environ 200kb est composé de deux parties essentielles, nécessaires à la transformation génétique, la région de virulence et l'ADN de transfert ou T-DNA. Le T-DNA est le fragment d'ADN qui sera transféré à la cellule végétale. On trouve aussi une région possédant les gènes nécessaires à la synthèse des opines et une région possédant les oncogènes nécessaires à la synthèse d'auxines et de cytokinines. Ces deux phytohormones étant responsables de la callogénèse, les oncogènes sont donc responsables de la formation des tumeurs. Ce T-DNA possède aussi deux séquences de 24 pb à ses extrémités (LB = Left Border, RB = Right Border) très importantes pour l'excision, le transfert et l'intégration du T-DNA durant la transgénèse. En effet, Les composés phénoliques libérés par la plante blessée vont activer l'expression des gènes présents dans la région de virulence. Certaines des protéines produites vont permettre la fixation des *Agrobacterium* sur les cellules végétales, d'autres comme par exemple les protéines VIR vont permettre l'excision, le transfert, la protection et l'intégration du T-DNA dans le génome de la plante (Hooykaas et Schilperoort 1992). Une fois ce mécanisme connu, il a été possible d'envisager de le détourner pour transférer des gènes d'intérêt chez les plantes. Deux types de systèmes ont vu le jour :

- Le système *cis*, ou un gène est introduit par recombinaison homologue au sein d'un T-DNA déjà présent dans le plasmide Ti.
- Le système binaire, ou un gène est cloné entre les séquences LB et RB au sein d'un vecteur synthétique. Ce vecteur sera ensuite introduit dans une souche d'agrobactérie possédant un plasmide Ti avec une région *vir* mais sans T-DNA.

Le Tabac fut la première plante transformée avec un gène étranger de levure en 1983. Puis de nombreuses plantes de la famille des dicotylédones ont suivi, comme la tomate, la pomme de terre, etc. La transformation via *Agrobacterium tumefaciens* s'est ensuite normalisée pour de nombreuses espèces afin d'améliorer des caractères agronomiques, ou pour la recherche fondamentale. D'autres techniques dites directes, utilisant des protoplastes ont aussi vu le jour, notamment la transformation médiée par le PEG, l'électroporation et la micro-injection (Hernández et Guillén 2012). Mais le principal problème reste la difficulté de régénérer les plantes à partir de protoplastes transformés *in vitro* chez certaines espèces.

Chez la chicorée, la première transgénèse n'a pas été réalisée avec *Agrobacterium tumefaciens* mais avec une autre espèce de bactérie : *Agrobacterium rhizogenes*. *Agrobacterium rhizogenes* a aussi la capacité de transférer un T-DNA à l'intérieur du génome d'une plante *via* un mécanisme très proche de celui d'*Agrobacterium tumefaciens*, les symptômes de ce transfert sont caractérisés par une production anarchique de racines qui donnera les initiales Ri au plasmide pour « Roots inducing ». *A. rhizogenes* a récemment (2001) été rebaptisée *Rhizobium rhizogenes* en raison d'analyses moléculaires qui ont montré qu'elle était plus apparentée à *Rhizobium* qu'à *Agrobacterium* (Young et al. 2001). Les différentes souches de cette agrobactérie sont aussi classées en fonction du type d'opines qu'elles produisent, plus particulièrement en fonction des gènes d'opines présents sur le plasmide Ri (Tableau 2).

Plasmid type	Opine biosynthesis genes	Opines, synthesized
Agropine	T _L : <i>acs</i> (agrocinopine synthase) T _R : <i>mas1, mas2</i> (mannopine synthase), <i>ags</i> (agropine synthase)	Mannopine, mannopinoc acid; agropine, agropinic acid; agrocinopines C and D
Mannopine	<i>mas1, mas2</i> (mannopine synthases)	Mannopine, mannopinoc acid
Mikimopine	<i>mis</i> (mikimopine synthase)	Mikimopine, mikimopine-lactam
Cucumopine	<i>cus</i> (cucumopine synthase)	Cucumopine, cucumopine-lactam

Tableau 2 : Gènes des opines présents sur les différents plasmides Ri d'après (Vladimirov, Matveeva, et Lutova 2015)

Le plasmide Ri des souches de type mannopine : pRi8196, de type mikimopine : pRi1724 et de type cucumopine : pRi2659, possèdent un simple T-DNA en revanche le plasmide Ri des souches de type agropine : pRiA4, pRi1855, pRiHRI, pRi15834 possèdent deux T-DNA, un T_L -DNA (left T-DNA) et un T_R -DNA (right T-DNA) (Figures 4 et 5). (Veena et Taylor 2007 ; Ozyigit, Dogan and Tarhan, 2013). Ces deux T-DNA sont transférés et intégrés indépendamment dans le génome de la plante lors du processus d'infection. En plus des gènes impliqués dans le catabolisme des opines et la synthèse des hormones, on trouve aussi des gènes *rol* sur les T-DNA des plasmides Ri. Les gènes *rolA*, *rolB*, *rolC* et *rolD* sont des oncogènes, leur expression modifie le niveau des hormones endogènes au sein de la plante plus particulièrement la biosynthèse des auxines et des cytokinines. Leur expression est aussi directement liée au phénotype hairy root (chevelu racinaire). Cependant leurs rôles exacts sont encore mal connus mais leurs effets sur la plante hôte ont été étudiés (Tableau 3). Les racines issues de la transformation par *A. rhizogenes* peuvent être cultivées *in vitro* sans nécessiter de régulateurs de croissance et chez certaines espèces il est possible de régénérer des plantes à partir de lignées de chevelus racinaires (Tepfer 1990). Les chevelus racinaires peuvent donc être utilisés pour la production de biomasse racinaire ou pour la production de plantes transformées (Figure 6). Chez la chicorée, c'est en 1987 que les premiers chevelus racinaires ont été mis en place par Mugnier (Mugnier 1988). En 1991, Sun (L. Y. Sun et al. 1991) observa qu'il était possible de régénérer des plantes à partir de chevelus racinaires de chicorée. Ces plantes possèdent un phénotype altéré et ont la particularité de fleurir très rapidement. Comme chez la

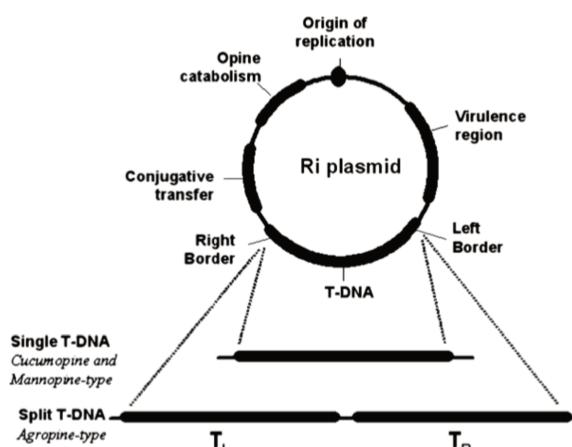


Figure 4: Structure schématique des plasmides Ri des différentes souches d'*Agrobacterium rhizogenes* d'après (Veena et Taylor 2007)

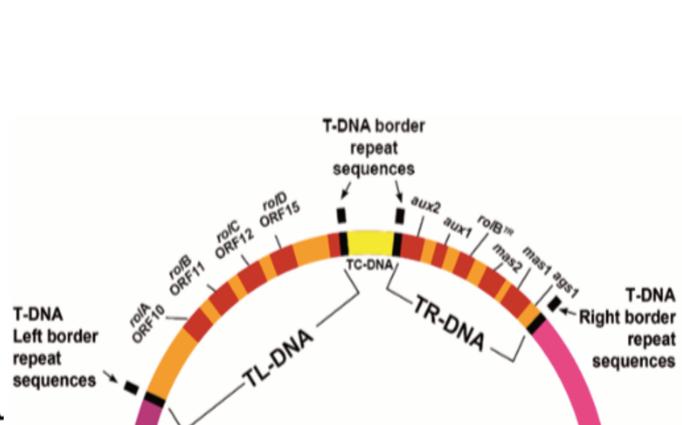


Figure 5 : Représentation schématique du T-DNA présent sur les plasmides Ri de type agropine chez *Agrobacterium rhizogenes* (Ozyigit, Dogan et Tarhan, 2013)

carotte, elles passent de plantes bisannuelles à plantes annuelles, effet direct de l'insertion du T-DNA du plasmide Ri de la bactérie. En ce qui concerne l'altération du phénotype, Limami montra qu'elle était directement liée à la diminution du contenu en arginine dans la plante (Limami 1998). Le transfert de gènes étrangers est aussi possible avec l'utilisation d'*A. rhizogenes*. Dans ce cas l'utilisation du système binaire ainsi qu'une souche sauvage d'agrobactéries est nécessaire. Le transfert du T-DNA du plasmide Ri possédant les gènes *rol* donnera le phénotype « hairy root », le T-DNA du vecteur binaire transférera le gène d'intérêt.

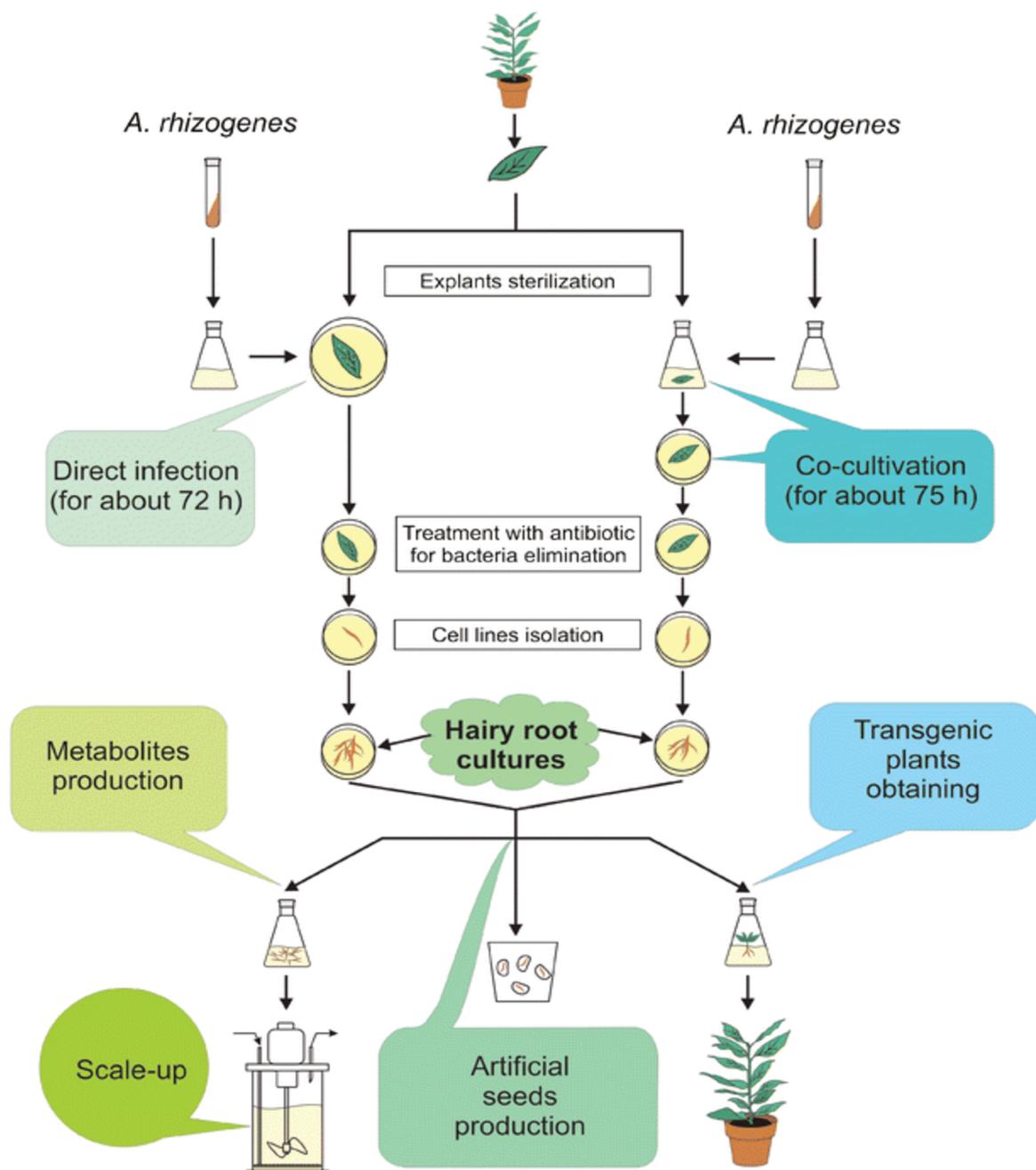


Figure 6 : Organigramme pour la mise en place de chevelus racinaires ainsi que quelques applications (d'après Georgiev et al. 2007)

Gene	Protein	Function	Phenotype
<i>rolA</i>	Sequence motif common in DNA-binding proteins Regulatory transcription factor	Inhibits cell elongation via diffusible factor Decreases hormone concentrations Increase sensitivity to auxin Modulating hormone physiology of GA Interfere polyamine metabolism Correlate with plasma membrane H ⁺ ATPase activity	Stunted growth, dark green wrinkled leaves with an altered length to width ratio, condensed inflorescences, retarded onset of flowering, compact reduced number of flowers
<i>rolB</i>	Localizes to plasma membrane	Alterations in the reception/transduction of the auxin signal Stimulates new meristem formation Induce secondary metabolism	Fast growth, root meristem neoformation, high branching and plagiotropism
<i>rolC</i>	rolC Phloem-specific expression in the root, low expression in the leaf, and no expression in the shoot tip	Reduces cell size Reduces abscisic acid (ABA), polyamine, and ethylene levels Formation of shoot meristems Regulate sugar metabolism and transport Stimulate the production of high levels of secondary metabolites Incapable	Increased branching, dwarfed plants with short internodes, reduced epidermal cell size in internodes, lanceolate leaves, early flowering, reduced flower size and reduced pollen production
<i>rolD</i>	Only expresses in Agropine type strains Cytosolic protein Exhibits poor tissue- or organ-specific expression	Incapable of inducing root formation on its own Provide defense response as a result of environmental stress	Increased flowering, reduced rooting, elongating and expanding tissues of each organ but not on apical meristem, callus growth giving rise to initiation of tumor resemble formation
<i>rolB^{TR}</i>		<i>rolB</i> homolog on TR-DNA in the agropine type Ri plasmid	Wrinkled leaves bent strongly downward formed shoots at the base of the stem and retarded growth

Tableau 3: Les oncogènes d'*Agrobacterium rhizogenes*, les protéines qu'ils encodent, leurs fonctions et leurs effets sur le phénotype de la plante hôte (D'après Ozyigit, Dogan et Tarhan, 2013)

Jusqu'au milieu de la deuxième partie du XXème siècle, les travaux sur les biotechnologies végétales (culture *in vitro*, cultures cellulaires, travaux sur les protoplastes) ont permis d'étudier et de comprendre le rôle des métabolites primaires dans les fonctions de base du développement des plantes, c'est à dire : les mécanismes de division cellulaire, de croissance, de respiration, de stockage et de reproduction. La transgénèse végétale a permis le franchissement des barrières inter-espèces. Tout cela a généralement permis d'améliorer le potentiel agronomique des plantes cultivées. A la fin du XXème siècle, avec l'arrivée des technologies analytiques comme par exemple la chromatographie et la spectroscopie de masse, le métabolisme dit « secondaire », en opposition au métabolisme primaire, va être de plus en plus étudié. Comparés aux principales molécules trouvées chez les plantes, les métabolites secondaires ont été dans un premier temps considérés comme des déchets issus de la dégradation des composés du métabolisme primaire, n'ayant pas un rôle essentiel à la survie de la plante. C'est grâce à l'accessibilité et à l'amélioration des techniques analytiques ainsi qu'à l'évolution de la biologie moléculaire, qu'il a clairement été démontré que les métabolites secondaires jouaient un rôle essentiel dans l'adaptation des plantes à leur environnement. Ces molécules ont été décrites comme étant produites en très petite quantité (<mg/g) par les plantes et ayant des activités antibiotique, antifongique et antivirale ; elles sont donc capables de protéger les plantes contre les pathogènes (phytoalexines). Elles peuvent aussi stopper la germination ou être toxiques pour les autres plantes (allélopathie). De plus elles jouent un rôle important dans la protection des plantes contre les UV (J. Li et al. 1993) ainsi que contre les animaux herbivores ou les insectes. Elles favoriseraient aussi les interactions avec les insectes pollinisateurs et les organismes symbiotiques (L. Yang et al. 2018). Au vu de leur très grande implication dans l'adaptation des plantes à leur environnement, le terme métabolisme secondaire a récemment été remplacé par métabolisme « spécialisé ».

Des milliers de molécules appartenant au métabolisme spécialisé ont été caractérisées et on estime qu'il en existerait des centaines de milliers. Malgré leur diversité structurale, ces molécules ont été classées en 3 familles : les composés azotés, les terpènes et les composés phénoliques. Concernant leur production, toutes ces molécules, quelle que soit leur classe, sont issues du métabolisme primaire. (Figure 7) (Borrelli et Trono 2016).

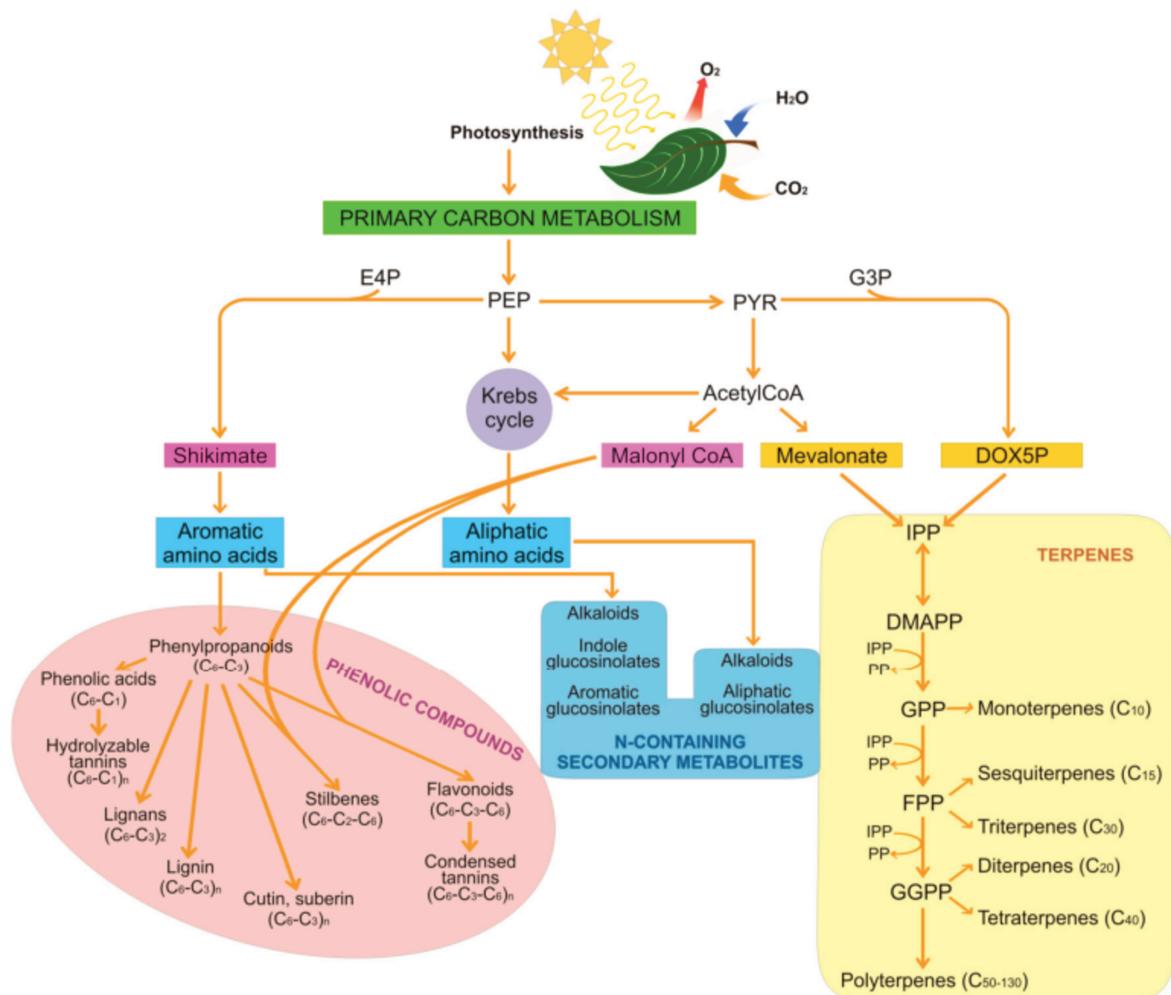


Figure 7 : Représentation générale des voies de biosynthèses des métabolites spécialisés. E4P, érythroose 4-phosphate; G3P, glycéradehyde 3-phosphate; PEP, phosphoénolpyruvate; PYR, pyruvate; DOX5P, désoxyxylulose 5-phosphate; IPP, isopentényl pyrophosphate; DMAPP, diméthylallyl pyrophosphate; PP, pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnésyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate (D'après Borrelli and Trono 2016)

Les métabolites spécialisés azotés regroupent les alcaloïdes et les glucosinolates :

- Les alcaloïdes représentent une grande classe de métabolites spécialisés avec plus de 12000 structures. L'azote qui provient d'acides aminés tels que la tyrosine, le tryptophane, de la lysine ou de l'acide aspartique, peut se trouver à l'intérieur (alcaloïdes vrais) ou à l'extérieur (pseudo-alcaloïdes) de l'hétérocycle. On peut aussi trouver des alcaloïdes possédant un hétérocycle azoté mais dont l'azote ne provient pas d'acides aminés (alcaloïdes stéroïdiens) (Ahmed et al. 2017). Ces molécules sont souvent toxiques pour les autres organismes et jouent donc un rôle important pour la défense de la plante contre les pathogènes et les herbivores.

-Les glucosinolates (ou hétérosides soufrés), représentent un petit groupe où environ 200 molécules sont répertoriées à l'heure actuelle. On les trouve particulièrement chez les plantes de la famille des Brassicacées. En fonction du type d'acide aminé dont ils dérivent, les glucosinolates peuvent être définis comme aliphatique, indole ou aromatique (Cartea et al. 2008).

Les terpènes représentent la plus grande classe de métabolites spécialisés avec plus de 50 000 structures différentes. Ils sont produits dans les tissus végétatifs et les fleurs, toutefois ils peuvent aussi être présents dans les racines de certaines plantes. Ils sont dans certains cas nommés « isoprénoïdes » car ils proviennent tous du même précurseur, l'isopentényl pyrophosphate (IPP), qui peut être synthétisé soit par la voie cytosolique du mélanovate, soit par la voie plastidiale du MEP/DOXP (2-C-méthylérythritol-4-phosphate / 1-désoxy-D-xylulose-5-phosphate) (Rohmer et al. 1999). Grâce à l'intervention de l'IPP isomérase, l'IPP peut être isomérisé en diméthylallyl pyrophosphate (DMAPP), et la condensation entre ce dernier et l'IPP donne le géranyl-pyrophosphate (GPP) qui est le précurseur de tous les monoterpènes. Par la suite, l'allongement de la chaîne carbonée par addition successive d'IPP permettra la synthèse des sesquiterpènes (C15), diterpènes (C20), triterpènes (C30), tétraterpènes (C40) et polyterpènes (C50-C130) (Gershenzon and Dudareva 2007). Ces molécules interviennent dans l'attraction des polliniseurs ainsi que la défense contre les pathogènes et herbivores. Certaines d'entre elles sont aussi impliquées dans le métabolisme primaire, notamment dans la synthèse des chlorophylles et des phytohormones (gibbérellines, cytokines).

Enfin les composés phénoliques qui regroupent environ 8000 structures différentes. Ils sont présents de façon ubiquitaire chez les plantes. Toutes les molécules appartenant à cette famille possèdent au moins un cycle benzénique hydroxylé. Ces composés découlent tous de la voie du shikimate car la plupart d'entre eux proviennent des deux acides aminés aromatiques que sont la phénylalanine et la tyrosine. La production de coumaroyl-CoA à partir de phénylalanine grâce à la voie des phénylpropanoïdes est une étape clé de leur voie de biosynthèse. En effet ce métabolite à une importance capitale, il est le précurseur de la quasi-totalité des composés phénoliques. Il interviendra dans la synthèse des flavonoïdes, des tannins, des lignines, etc (Figure 7). Les composés phénoliques sont impliqués dans de nombreux processus physiologiques chez les plantes, comme la pollinisation, la

résistance aux pathogènes et aux prédateurs ainsi que la résistance aux stress abiotiques (Cheynier et al. 2013).

Grâce à leur gigantesque diversité structurale, les métabolites spécialisés possèdent de nombreuses activités biologiques bénéfiques à la plante. Ces activités sont utilisées depuis des milliers d'années par l'homme au travers de la médecine traditionnelle. L'étude traditionnelle des effets d'extraits (infusion, décoction, etc) ou de la consommation directe des plantes par l'homme a permis de mettre au jour de nombreuses pharmacopées propres aux différents peuples du globe. Ces recueils retraçant l'expérience de l'homme avec les plantes sont de nos jours très utiles pour l'identification de plantes à haut potentiel. Les années 70 ont marqué le début des études phytochimiques destinées à identifier les métabolites spécialisés produits par les plantes.

La chicorée, étant historiquement une plante utilisée pour ses nombreuses vertus (Tableau 1), a donc fait l'objet de nombreuses investigations afin d'identifier les composés actifs pouvant expliquer ses propriétés antioxydante, antibactérienne, anti-inflammatoire, antidiabétique ou encore hépato-protectrice (Carazzone et al. 2013; El-Sayed et al. 2015; Sinkovic et al. 2015). Ainsi plus de 100 métabolites spécialisés ont été identifiés et isolés chez la chicorée, la majorité d'entre eux provenant de sa racine. La chicorée, comme de nombreuses espèces appartenant à la famille des Astéracées, contient une grande quantité de lactones sesquiterpéniques, la plupart d'entre elles étant produites dans les racines et les plus abondantes étant, la lactucine, la 8-déoxylactucine et la lactucopicrine (Rees and Harborne 1985; Willeman et al. 2014). Ces molécules mises en place par la plante pour se défendre contre les herbivores seraient responsables de l'amertume chez la chicorée (Van Beek et al. 1990; Kraker et al. 1998). On retrouve aussi chez la chicorée des composés phénoliques comme la 6,7-dihydroxycoumarine plus communément nommée cichoriine et des flavonoïdes comme des dérivés de quercétine, de kampférol et d'épigénine (Carazzone et al. 2013). Cependant les composés phénoliques majoritaires chez la chicorée, comme chez la plupart des astéracées sont des acides hydroxycinnamiques tels que les acides chlorogéniques (acide caféoylquinique ou CQA) et les acides isochlorogéniques (acide dicaféoylquinique ou diCQA). On y trouve aussi deux esters d'acide tartrique : l'acide caftarique (acide caféoyltartrique ou CTA) et l'acide chicorique (acide dicaféoyltartrique ou diCTA) (Legrand et al. 2016; Street, Sidana, et Prinsloo 2013). Au sein de la plante, le rôle de ces molécules est encore mal connu

mais de nombreuses études sur leurs activités ont montré qu'elles possédaient un fort pouvoir antioxydant directement lié à leurs activités anti-inflammatoire, hépatoprotectrice, neuroprotectrice ou encore antivirale (Heyman et al. 2015; Xu, Hu, et Liu 2012; Das et Vasudeva 2016; Zhao et al. 2011).

L'isolement et l'identification de composés nouveaux ont permis de découvrir et de produire de nouvelles molécules actives qui se sont avérées avoir un intérêt particulier dans le domaine de la pharmacologie. Ainsi de nombreuses molécules végétales ont été utilisées en médecine pour leurs propriétés antibactérienne, antifongique, anticancéreuse, etc. La découverte de l'artémisinine à partir de l'armoise (*Artemisia annua L.*), a permis de réduire de 40% les cas de malaria en Afrique. Un autre exemple bien connu est le paclitaxel produit par l'if (*Taxus brevifolia*) qui a permis de traiter de nombreux cancers depuis les années 1990. De nos jours les molécules bioactives sont très utilisées dans les secteurs de la pharmacologie, de la cosmétique, de l'agronomie et de la nutrition. Environ 25% des molécules médicinales utilisées sont issues de molécules d'origine naturelle. Au vu de la situation sanitaire actuelle (augmentation du nombre de cancers, résistances des agents pathogènes aux antibiotiques, etc.) la demande en molécules bioactives n'a cessé d'augmenter au cours de ces 30 dernières années.

Cependant, la plupart des molécules actives sont produites en très petites quantités par les plantes, ce qui rend leur obtention par extraction souvent non viable économiquement. De plus, du fait de leurs complexités structurales, la synthèse chimique de ces molécules est souvent impossible. La production de certaines molécules s'est donc pendant longtemps faite par la culture à grande échelle de plantes médicinales. Cependant, les plantes provenant de biotopes particuliers peuvent être difficiles à cultiver en dehors de leurs écosystèmes naturels. Les scientifiques et les biotechnologistes ont donc dû trouver des alternatives pour obtenir une biomasse élevée, riche en métabolites spécialisés. Ils se sont donc tournés vers les cultures cellulaires et la culture d'organes, notamment la culture de chevelus racinaires.

Les cultures cellulaires mises en place chez de nombreuses espèces, notamment pour l'étude du métabolisme primaire, ont pendant longtemps été considérées comme étant incapables de produire des métabolites spécialisés à cause de leur statut de cellules indifférenciées. Mais en 1975, Zenk et son équipe montrèrent que des cultures cellulaires de *Morinda citrifolia* pouvaient produire 2,5 grammes d'anthraquinone par

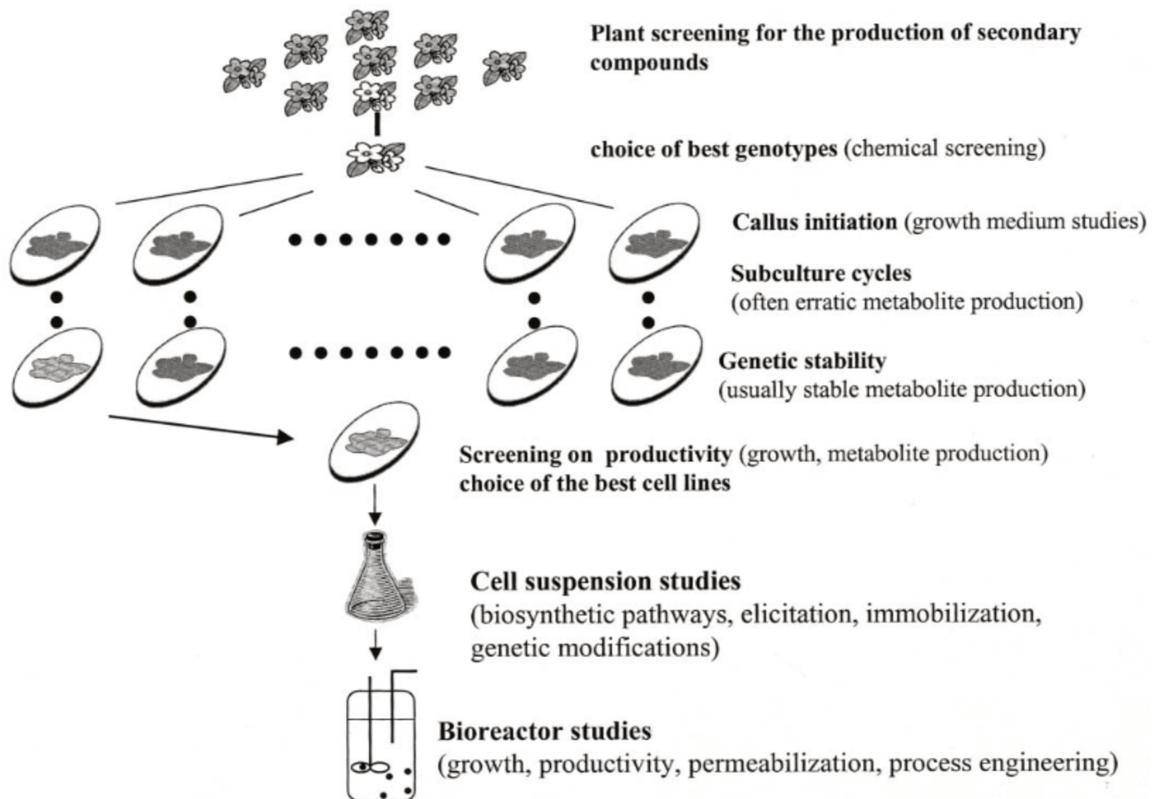


Figure 8 : Stratégie pour la production de métabolites spécialisés grâce à la culture cellulaire (d'après Bourgaud et al 2001)

litre de milieu de culture (Zenk, El-Askary, et Schulte 1975). Cette découverte marque le début de la culture cellulaire pour la production des métabolites spécialisés et son adaptation à la culture industrielle (Figure 8) (Bourgaud et al. 2001 ; Eibl 2018). Chez la chicorée, la mise en place de cultures cellulaires à partir de cals provenant de tissus racinaires, fait accroître la production d'acide isochlorogénique (3,5-diCQA) par 2. De plus, l'élicitation de ces cultures avec 50 µmol/litre (µM) de méthyljasmonate (MeJa) pendant 96h fait passer la production de 3,5-diCQA de 15µmol/gramme Matière Sèche (MS) à 65µmol/g MS (Legrand et al. 2016).

La culture d'organes s'est aussi avérée, dans certains cas, une bonne alternative pour la production de biomolécules. Les cultures de plantules et de chevelus racinaires sont les plus utilisées. Les chevelus racinaires ont la capacité de pouvoir être cultivés indéfiniment dans des milieux de culture ne possédant pas de phytohormones. Leur développement est très rapide et de nombreux bioréacteurs ont été mis en place pour la culture à grande échelle (Steingroewer et al. 2013). De plus, les gènes *rol* présents sur le T-DNA ont la capacité d'accroître la production de métabolites spécialisés (Bulgakov 2008). De nombreuses cultures d'organes ont été mises en place pour la

production de biomolécules (Tableau 4). Ces cultures peuvent aussi être élicitées pour augmenter la production (Tableau 5).

species	Phytochemicals	<i>In vitro</i> system
<i>Abrus precatorius</i> L.	Glycyrrhizin	Hairy roots
<i>Azadirachta indica</i> (L.) Adelb.	Azadirachtin	Hairy roots
<i>Beta vulgaris</i> L.	Betalains, phenolics	Hairy roots
<i>Catharanthus roseus</i> (L.) G.Don	Vindoline, vincristine, vinblastine	Shoots
<i>Datura stramonium</i> L.	Hyoscyamine	Hairy roots
<i>Hypericum hirsutum</i> L. <i>Hypericum maculatum</i> Crantz	Hypericin, pseudohypericin, hyperforin	Shoots
<i>Leucojum aestivum</i> L.	Galanthamine, lycorine	Shoots
<i>Ophiorrhiza rugosa</i> var. <i>decumbens</i>	Camptothecin	Shoots
<i>Pancratium maritimum</i> L.	Amaryllidaceae alkaloids	Shoots
<i>Picrorhiza kurroa</i> Royle ex Benth	Picrotin, picrotoxinin	Hairy roots
<i>Plumbago indica</i> L.	Plumbagin	Hairy roots
<i>Pueraria tuberosa</i> (Roxb. ex Willd.)	Puerarin	Shoots
<i>Rubia akane</i> Nakai	Alizarin, purpurin	Hairy roots
<i>Salvia tomentosa</i> Mill.	Oleanolic acid, ursolic acid	Hairy roots
<i>Stevia rebaudiana</i> Cav.	Steviol glycosides	Shoots
<i>Withania somnifera</i> (L.) Dunal.	Withanolide A	Hairy roots

Tableau 4 : Exemple de biomolécules produites par des systèmes *in vitro* de cultures d'organes (d'après Steingroewer et al. 2013)

Chez la chicorée, les travaux de Bais (Bais, George, et Ravishankar 1999 ; Bais 2000) ont permis de mettre en évidence la production de deux coumarines nommées « esculine » et « esculétine ». Ces deux molécules ayant un potentiel intérêt en cosmétique et en microbiologie. De plus, cette production peut être augmentée par l'ajout de polyamines (spermidine et spermine) dans le milieu ainsi qu'après élicitation avec des éliciteurs fongiques (Tableau 5). D'autres expériences ont montré que les chevelus racinaires de chicorée étaient capables de produire de grandes quantités d'acide hydroxycinnamique avec des concentrations d'acide isochlorogénique pouvant atteindre 5% de MS (Malarz, Stojakowska, et Kisiel 2013).

Elicitor	Hairy root culture	Metabolite	Elicitor concentration	Fold increase
Abiotic elicitors				
CdCl ₂	<i>Brugmansia candida</i>	Hyoscyamine	2 mM	1.35
Jasmonic acid	<i>Panax ginseng</i>	Ginsenoside	5.0 mg/l	4.0
Methyl jasmonate	<i>Panax ginseng</i>	Ginsenoside	22.4 mg/l	4.0
Vanadyl sulfate	<i>Ambrosia artemisiifolia</i>	Thiarubrine A	50 mg/l	8.0
Biotic elicitors				
<i>Enterobacter saka</i>	<i>Ammi majus</i>	Umbelliferone	15 ml susp/l	1.2
<i>Fusarium congl</i>	<i>Tagetes patula</i>	Total thiophenes	0.12 mg/ml	2.7
<i>Phytophtora para</i>	<i>Cichorium intybus</i>	Esculin esculentin	0.1% (v/v)	4.1/3.7
<i>S. aureus</i>	<i>Scopolia parviflora</i>	Scopolamine	133.3 ml/l	2.8
Yeast elicitor	<i>Salvia miltiorrhiza</i>	Rosmarinic acid	—	2.3
Yeast extract	<i>Brugmansia candida</i>	Scopolamine hyoscyamine	0.8 mg/ml	2.0

Tableau 5 : Exemple d'augmentation de production de métabolites spécialisés par élicitation
(d'après Georgiev et al 2007)

Malgré la forte demande en biomolécules possédant des activités recherchées en pharmacologie, seulement quelques-unes d'entre elles sont produites industriellement par culture végétale *in vitro* (Eibl 2018). Ceci peut être expliqué par les variations somaclonales dont peuvent être victimes les cultures ainsi que par une trop faible production de métabolites spécialisés comparée aux coûts de mise en place et de maintenance des procédés de culture *in vitro* à l'échelle industrielle. De plus les molécules du métabolisme spécialisé sont souvent produites par des cellules spécialisées comme par exemple des cellules laticifères qui produisent les latex (Castelblanque et al. 2016), des trichomes (Huchelmann, Boutry, et Hachez 2017), des cellules du tapis nourricier du grain de pollen (Grienemberger, Besseau, et Geoffroy 2009) etc. Cellules dont la culture *in vitro* est à l'heure d'aujourd'hui impossible. Dans le but d'optimiser la production de biomolécules, identifier les gènes clés des voies de biosynthèse pour les transférer dans des systèmes d'expression plus productifs (levures, bactéries) va donc devenir une bonne alternative. La naissance du génie métabolique va donc rendre l'étude et la compréhension des voies de biosynthèse fondamentales. De nombreux gènes impliqués dans le métabolisme spécialisé vont être identifiés, grâce à la purification et au séquençage de protéines. Puis avec l'évolution des technologies de séquençage et l'acquisition d'une grande

quantité de données génomiques chez les organismes modèles ou chez certaines plantes cultivées, cette stratégie va être délaissée au profit des stratégies de bio-informatique. En effet la recherche par homologie de séquences avec des enzymes aux activités similaires va permettre d'accélérer considérablement l'identification de gènes candidats.

Chez de nombreuses espèces, la voie de biosynthèse du CQA a bien été décrite (Lallemand et al. 2012 ; Kim et al. 2013 ; Comino et al. 2009). Comme pour la plupart des composés phénoliques, elle nécessite l'utilisation de coumaroyl-CoA provenant de la voie des phénylpropanoïdes, elle-même initiée par la phénylalanine. Cet acide aminé aromatique essentiel, est produit par la voie du shikimate, alimentée soit en érythrose-4-phosphate par la voie des pentoses phosphates, soit en phosphoénolpyruvate par la glycolyse. La phénylalanine, une fois synthétisée, emprunte la voie des phénylpropanoïdes proprement dite. Elle est catalysée en acide cinnamique grâce à la phénylalanine ammonialyase (PAL), qui est ensuite catalysé en acide *p*-coumarique par la cinnamate 4-hydroxylase (C4H), enzyme faisant partie de la famille des cytochromes P450. La dernière étape de cette voie est la catalyse de l'acide *p*-coumarique en coumaroyl-CoA grâce à la 4-hydroxycinnamoyl-CoA ligase (4CL), réaction faisant intervenir la coenzyme A (CoA). L'expression de ces trois enzymes (celle de la PAL plus particulièrement) peut être induite par une infection, une blessure, une exposition aux UVs, un stress hydrique, toutefois, le génotype est également un déterminant majeur de la variation de la teneur en polyphénols dans la plante (Pandino et al. 2014). Le coumaroyl-CoA va ensuite être pris en charge par des acyltransférases de type HCT (Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransférase) et HQT (Hydroxycinnamoyl-Coa quinate/shikimate hydroxycinnamoyltransférase) pour produire soit du coumaroylshikimate, soit du coumaroylquinate. Ces deux dérivés d'acide coumarique vont ensuite être hydroxylés grâce à l'intervention d'une autre enzyme de la famille des P450, la C'3H (*p*-coumarate 3'-hydroxylase), pour former deux dérivés d'acide caféique : du CQA ou du cafféoylshikimate. Ces deux esters d'acide caféique peuvent par la suite être dégradés par ces mêmes HCT/HQT pour donner du cafféoyl-CoA, qui servira de substrat dans la voie de biosynthèse des lignines (Figure 9). Le CQA peut donc être obtenu par l'intermédiaire de la voie shikimate/quinate-dépendante soit par la voie quinate-dépendante. La présence d'HCT de manière ubiquitaire chez les végétaux est directement liée à leur implication dans la synthèse de cafféoyl-CoA comme

précurseur de la synthèse des lignines. En revanche la présence d'HQT semble restreinte aux espèces accumulant du CQA.

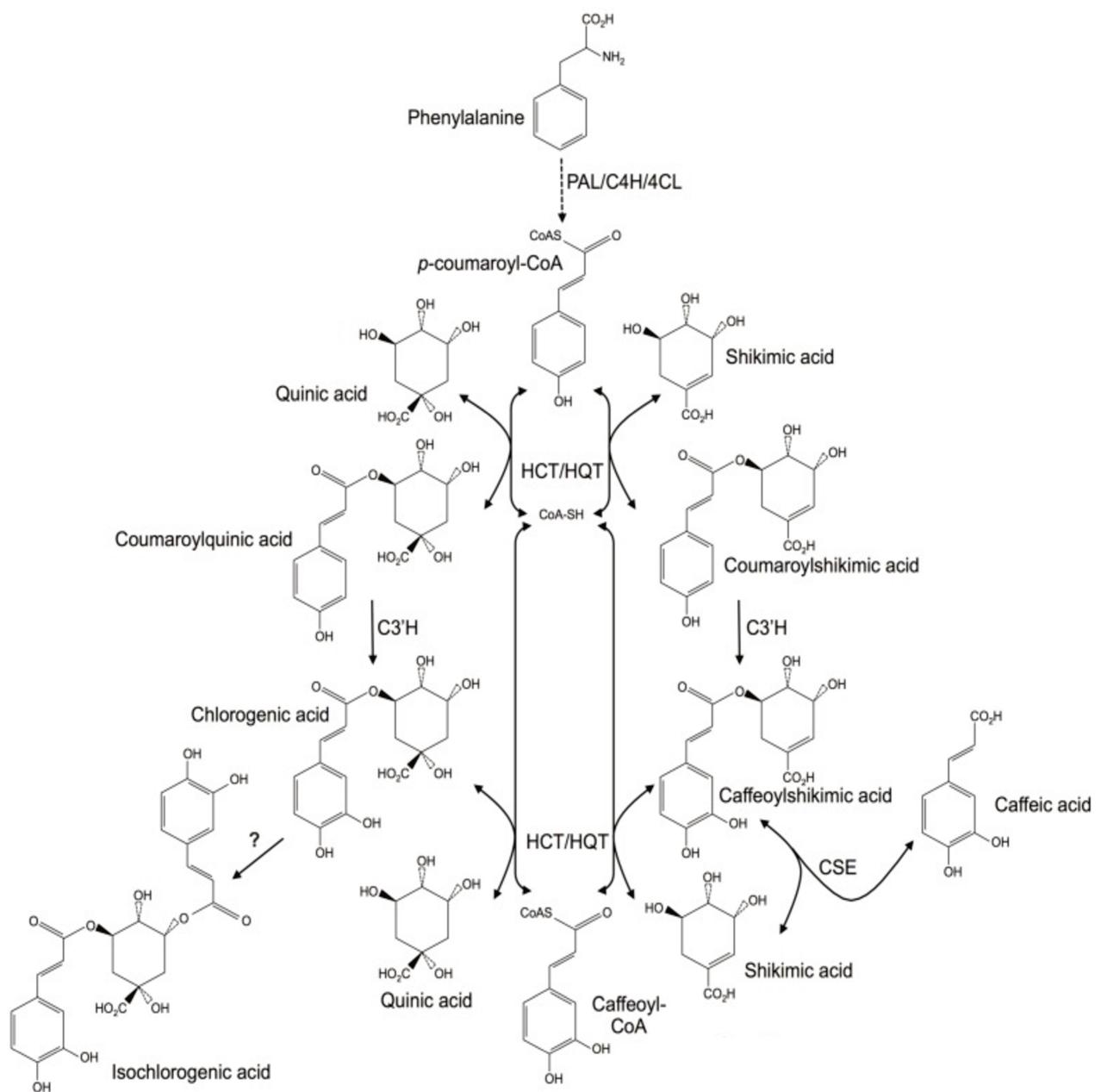


Figure 9 : Voie de biosynthèse de l'acide chlorogénique. Abréviations : PAL, Phénylalanine Ammonia Lyase ; C4H, Cinnamate 4-Hydroxylase ; 4CL, 4 Coumarate-CoA Lyase ; HCT, Hydroxycinamoyl-CoA shikimate/quinate hydroxycinnamoylTransférase ; HQT, Hydroxycinamoyl-CoA Quinate /shikimate hydroxycinnamoyltransférase ; C3'H, *p*-Coumarate 3'-Hydroxylase ; CSE, Cafféoyl Shikimate Estérase ; ? = Isochlorogénate Synthase, enzyme caractérisée biochimiquement mais dont la séquence génomique est encore inconnue (D'après Legrand et al 2016).

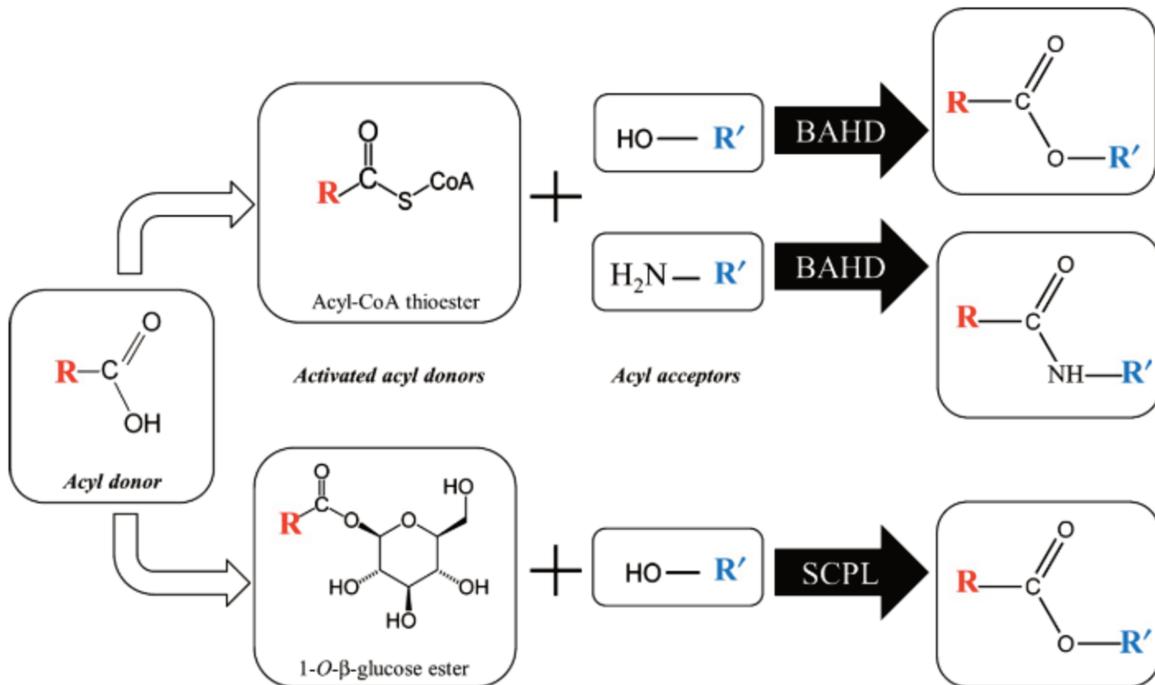


Figure 10 : Réactions d'acylation catalysées par les enzymes de la famille des BAHD et des SCPL. (D'après Bontpart et al. 2015)

En greffant une molécule de coumaroyl-CoA sur une molécule de shikimate ou de quinate les HCTs/HQTs réalisent donc des réactions d'acylation. Chez les végétaux, les réactions d'acylation sont très fréquentes au sein du métabolisme spécialisé. Elles confèrent aux molécules différentes propriétés comme la solubilité, la stabilité, ce qui influe directement sur leur transport et leur stockage ainsi que sur leurs activités (Gomez et al. 2009). D'un point de vue plus chimique, l'acylation est le transfert d'un groupement acyle préalablement activé, sur une molécule accepteuse (Figure 10). Deux principales familles d'enzymes réalisent des réactions d'acylation chez les plantes vasculaires : les Sérine CarboxyPeptidase-Like (SCPL) qui utilisent des esters de glucose comme donneurs d'acyle, et les enzymes de la famille des BAHD qui utilisent des thioesters de Coenzyme A, pour réaliser des réactions d'estéification ou d'amidification (Figure 10) (Bontpart et al. 2015). L'acronyme BAHD vient des initiales des 4 premières enzymes caractérisées (BEAT : Benzylalcohol O-acétyltransférase, AHCT : Anthocyanin O- hydroxycinnamoyltransférase, HCCT : N-Hydroxycinnamoyl/benzoyltransférase, DAT : Désacétylvindoline4-O-acétyltransférase) (St-Pierre et al. 1998). Ces enzymes sont des monomères compris entre 48 et 55 kDa, d'environ 445 acides aminés, qui partagent deux motifs conservés

HXXXD et DFGWG. Entre 50 et 120 gènes appartenant à cette famille ont été prédits dans les génomes d'*Arabidopsis* (Brassicaceae), de la luzerne tronquée (*Medicago truncatula*, Fabaceae) ou encore du peuplier (*Populus trichocarpa*, Salicaceae) et du poirier (*Pyrus bretschneideri*, Rosaceae) (Auria 2006 ; Yu et al. 2009 ; Tuominen, Johnson, et Tsai 2011 ; Ma et al. 2017). De tous ces gènes, environ 75 ont été caractérisés biochimiquement et génétiquement. Des expériences de cristallographie et/ou de mutagenèse dirigée menées sur les anthocyanin-malonyl transférases de sauge (*Salvia splendens*, Lamiaceae) et de chrysanthème (*Dendranthema morifolim*, Asteraceae), sur la vinorine synthase de *Rauvolfia serpentina* (Apocynaceae), et les HCTs de café (*Coffea canephora*, Rubiaceae) et de sorgho (*Sorghum bicolor*, Poaceae) (Suzuki, Nakayama, et Nishino 2003 ; Lallemand et al. 2012 ; Ma et al. 2005) ont permis d'obtenir la structure de ces enzymes et de situer les motifs conservés. Les BAHD présentent une structure globulaire composée de deux domaines de même taille reliés par une boucle. Ces deux domaines mettent en place un tunnel qui forme le site actif au milieu duquel on trouve le motif HXXXD. L'histidine de ce motif serait directement impliquée dans l'acte catalytique réalisé par ces enzymes (Haslam et al. 2017). Le motif DFGWG quant à lui serait plus impliqué dans la structure de l'enzyme. Des analyses phylogénétiques utilisant 46 enzymes biochimiquement et génétiquement caractérisées, ont permis de mettre en évidence 5 clades majeurs au sein de cette famille (D'Auria 2006). Cette classification a révélé des activités biochimiques pouvant être spécifiques au sein d'un clade mais pouvant aussi être différentes. La spécificité de substrat varie généralement d'un clade à l'autre mais peut aussi varier au sein d'un même clade. Par exemple les membres du clade V utilisent des substrats allant des terpénoïdes, aux alcools à moyenne chaîne en passant par les acides quiniques, alors que les membres du clade I agissent principalement sur les flavonoïdes. Une diversité similaire a également été notée pour les membres du clade III qui sont impliqués dans la formation d'alcaloïdes, d'esters et de flavonoïdes. D'un autre côté, l'utilisation de benzoyl-CoA semble limitée au clade V alors que les hydroxycinnamoyl-CoA ont été décrits comme des donneurs d'acyle au sein de nombreux clades. Tout ceci met en évidence la diversité des enzymes de la famille des BAHD, d'autant plus que certaines enzymes sont dotées d'une polyvalence catalytique qui peut varier considérablement. *In vitro*, certaines enzymes ont une large spécificité de substrat, de sorte que les produits qu'elles forment sont parfois déterminés par la disponibilité relative des substrats. De ce fait l'introduction au sein

d'une autre plante, d'un gène codant une enzyme possédant une large spécificité de substrat, pourrait former de nouveaux métabolites acylés, pouvant rendre complexe sa bonne caractérisation (D'Auria 2006).

L'étude et la caractérisation des gènes impliqués dans cette voie de biosynthèse a permis la production de CQA par des souches d'*Escherichia coli*. En effet la co-expression d'une 4CL d'*Oryza sativa* et d'une HCT de *Nicotiana tabacum* dans des souches d'*Escherichia coli* optimisées a permis une production de 450 mg/L de CQA (Kim et al. 2013).

Outre la production de biomolécules, la connaissance des gènes impliqués dans ces voies de biosynthèse, a un grand intérêt pour les nouvelles stratégies de sélection variétale et amélioration des plantes, liant la biologie moléculaire, la génomique et la métabolomique. En effet la tendance actuelle est à l'« aliment fonctionnel », riche en biomolécules ayant des bienfaits pour la santé. De ce fait la détection et la cartographie de locus de caractères quantitatifs (QTL) dans des populations en ségrégation ou des collections de matériel génétique fournissent des informations d'un grand intérêt pour la sélection assistée par marqueurs et l'amélioration génétique des plantes. Par conséquent, la cartographie des principaux QTL pour la teneur en composés phénoliques facilite l'incorporation de ce caractère dans des cultivars élites (Kaushik et al. 2015). Par exemple chez l'aubergine, les gènes codant pour les enzymes des PAL, C4H, 4CL, HCT, C3'H et HQT ont été cartographiés et il a été montré qu'excepté pour la 4CL et l'HCT, ces gènes n'étaient pas liés, ce qui d'après Gramazio et al, faciliterait la hiérarchisation des allèles favorables dans une même variété dans le but que celle-ci produise de plus grandes teneurs en acides chlorogéniques (Gramazio et al. 2014). De plus avec l'apparition de nouveaux outils d'édition des génomes comme la technologie CRISPR/Cas9, l'accumulation de données génomiques et l'identification de gènes candidats sont primordiales pour la création de nouvelles variétés.

Identifier des gènes et accumuler des données génomiques est donc une question de survie pour la sélection des petites espèces cultivées comme la chicorée. Ainsi chez la chicorée, 81 séquences appartenant potentiellement à la famille des BAHD ont été identifiées (thèse Mariane Delporte 2017). L'alignement de ces séquences avec des BAHD caractérisées chez d'autres espèces, a permis d'établir un arbre phylogénétique. Au sein de cet arbre on distingue 4 des 5 clades décrits par D'Auria (2006) et Tuominen et al. (2011), le clade II n'y étant pas représenté. Parmi ces 81

séquences, 7 séquences ont été identifiées comme appartenant au sous-clade Vb, sous-clade dans lequel on trouve les acyltransférases qui transfèrent toutes un acide hydroxycinnamique sur des accepteurs de nature variée, pouvant autant bien être des fonctions alcools que des fonctions amines. Sous-clade comprenant entre autres les HCTs et les HQTs caractérisées.

Sur ces 7 séquences, 2 HCTs et 3HQTs ont été caractérisées (Legrand et al. 2016). Les deux autres séquences semblent quant à elles un peu plus éloignées des HCTs et des HQTs, mais se rapprocheraient plus des enzymes annotées SHTs Spermidine hydroxycinnamoyl Transférases et annotées CiSHT1 et CiSHT2 (Figure 11).

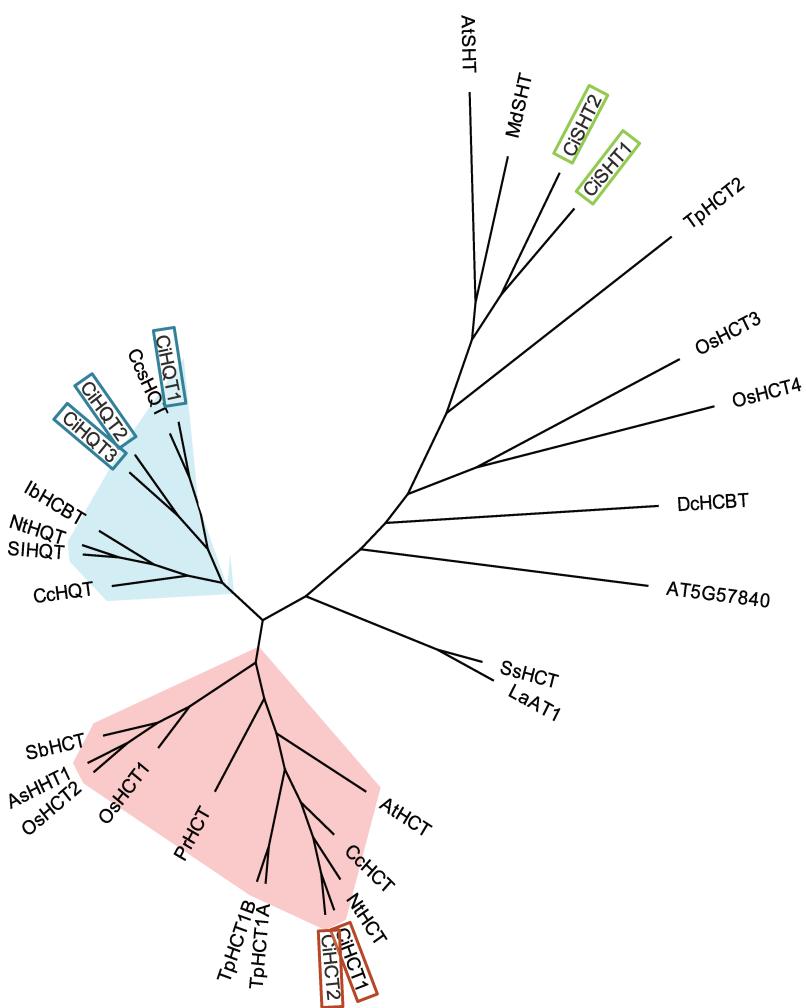


Figure 11 : Représentation du sous-clade V b de l'arbre phylogénétique des BAHD. Les séquences de chicorée sont encadrées en vert pour les SHTs, en rouge pour les HCTs et en bleu pour les HQT. (Thèse Mariane Delporte)

Ainsi durant la première partie de ma thèse j'ai contribué à la caractérisation des deux séquences annotées *CiSHT1* et *CiSHT2* chez la chicorée. Notamment en mettant en place les outils de validation fonctionnelle des gènes que sont les chevelus racinaires et en induisant la surexpression de ces gènes dans les chevelus racinaires de chicorée dans le but de valider leur fonction putative *in planta*. Cela constitue la première partie de ma thèse.

Les résultats obtenus grâce à la surexpression de ces gènes ne nous permettant pas de valider définitivement la fonction de ces gènes, nous avons alors décidé de créer des mutants Knock Out (KO) pour ces deux gènes, et donc de mettre au point le révolutionnaire outil d'édition du génome qu'est CRISPR/Cas9, chez la chicorée. Pour la mise au point de cet outil biotechnologique, nous avons choisi de cibler le gène de la *phytoène désaturase* de la chicorée (*CiPDS*). Les résultats seront exposés dans la seconde partie de ce mémoire.

Cet outil s'étant avéré fonctionnel chez la chicorée nous avons alors entrepris de créer des mutants KO pour les gènes *CiSHT1* et *CiSHT2* chez la chicorée. La caractérisation biochimique des mutants obtenus ainsi que des expériences complémentaires nous ont permis de finaliser la caractérisation de ces deux gènes. Les résultats seront exposés dans la troisième partie de ce mémoire.

Chapitre I : Caractérisation des enzymes CiSHT1 et CiSHT2 impliquées dans la synthèse des phénolamides accumulées dans le manteau pollinique de la chicorée et des Astéracées

Classification	Plant species	TriHCSpd
Spermaphyta		
Gymnospermae		
Coniferophyta		
Pinales		
<i>Picea abies</i>	○	
<i>Pinus pinaster</i>	○	
<i>Pseudotsuga menziesii</i>	○	
<i>Toreya nucifera</i>	○	
Taxales		
<i>Taxus baccata</i>	○	
Magnoliophyta		
Basal magnoliophyta		
Nymphaeales		
<i>Nymphaea alba</i>	●●●	
Mesangiospermae		
Magnoliids		
Magnoliales		
<i>Magnolia grandiflora</i>	○	
Monocotyledons		
Asparagales		
<i>Iris xiphium</i>	○	
Liliales		
<i>Tulipa gesneriana</i>	○	
<i>Lilium longiflorum</i>	○	
Alismatales		
<i>Spathiphyllum wallisii</i>	○	
Commelinids		
Poales		
<i>Zea mays</i>	○	
<i>Triticum aestivum</i>	○	
<i>Alopecurus pratensis</i>	○	
Zingiberales		
<i>Canna indica</i>	●	
Eudicotyledons		
Stem eudicotyledons		
Ranunculales		
<i>Papaver rhoeas</i>	○	
Core eudicotyledons		
Asterids		
Cornales		
<i>Hydrangea macrophylla</i>	●●●	
Ericales		
<i>Camellia japonica</i>	●●	
Lamiids		
Lamiales		
<i>Nepeta cataria</i>	●	
Solanales		
<i>Solanum lycopersicum</i>	●●●	
<i>Nicotiana tabacum</i>	●●●	
Gentianales		
<i>Catharanthus roseus</i>	●	
Campanulids		
Asterales		
<i>Helianthus annuus</i>	●●●	
Apiales		
<i>Daucus carota</i>	●	
Dipsacales		
<i>Lonicera caprifolium</i>	●	
Rosids		
Vitales		
<i>Vitis vinifera</i>	●●●	
Malvids		
Geraniales		
<i>Pelargonium grandiflorum</i>	○	
Brassicales		
<i>Arabidopsis thaliana</i>	●●	
Myrtales		
<i>Myrtus communis</i>	●●●	
Malvales		
<i>Tilia tomentosa</i>	○	
Fabids		
Rosales		
<i>Rosa Canina</i>	●●●	
<i>Malus domestica</i>	●●	
Malpighiales		
<i>Ricinus communis</i>	●●●	
Cucurbitales		
<i>Cucurbita pepo</i>	○	
Fagales		
<i>Castanea sativa</i>	●●●	
<i>Alnus glutinosa</i>	●●	
Fabales		
<i>Cytisus scoparius</i>	○	
<i>Trifolium pratense</i>	○	
<i>Mimosa pudica</i>	○	

Figure 1.1 : Accumulation de trihydroxycinnamoyl spermidine dans le manteau pollinique des spermaphytes.

••• Les phénolamides sont les composés phénoliques majoritaires

•• Les phénolamides sont en quantités équivalentes à celle d'autres composés phénoliques

• Les phénolamides sont des composés phénoliques minoritaires

Avant-propos

L'établissement d'un l'arbre phylogénétique à partir des 81 séquences de chicorée appartenant potentiellement à la famille des BAHD a permis l'identification de 7 séquences appartenant au sous-clade Vb. Sur ces 7 séquences, 2 ont été caractérisés par Legrand et al (2016) comme étant des HCTs et 3 comme étant des HQTs. Les deux autres séquences sont phylogénétiquement plus éloignées des HCTs/HQTs et se rapprochent plus des enzymes annotées SHTs, SpermidineHydroxycinnamoyl Transférases. Chez *Arabidopsis* et chez le pommier (*Malus domestica*) les enzymes annotées AtSHT et MdSHT ont été caractérisées, elles correspondent toutes deux à des Spermidine Hydroxycinnamoyl Transférases (Grienberger, Besseau, and Geoffroy 2009) (Elejalde-Palmett et al. 2015). Ces enzymes catalysent la formation de phénolamides, plus particulièrement des trihydroxycinnamoyl spermidines, qui sont synthétisées au niveau du tapetum des anthères, puis sont déposées et accumulées au niveau de la surface des grains de pollen durant l'étape de mort cellulaire programmée du tapetum. Tout cela se déroulant au cours de la maturation des grains de pollen. La présence de ces molécules serait retrouvée chez la quasi-totalité des Eudicotylédones en lien avec la présence d'enzymes SHT orthologues (Figure 1.1). Durant la réaction enzymatique d'acylation effectuée par ces SHTs, la spermidine qui est une polyamine possédant trois groupes amines, est N-acylée par trois acides hydroxycinnamiques sur ses trois azotes. Chez *Arabidopsis*, la synthèse de spermidine totalement substituée fait aussi intervenir des cytochromes P450 (CYP98A8/A9) ainsi qu'une O-méthyltransférase (AtSM1) réalisant des hydroxylations et la méthoxylation des groupements phénoliques pour finalement aboutir à la synthèse de dihydroxyféruloyl-sinapoyl spermidine, phénolamide majoritaire que l'on trouve à la surface des grains de pollen d'*Arabidopsis*. Cependant cette voie métabolique serait restreinte aux Brassicacées (Figure 1.2). Chez la plupart des autres espèces, la spermidine est plutôt conjuguée à des groupements moins complexes tels que des groupements coumaroyl- ou caféoyl- comme chez le pommier, chez lequel c'est la tricoumaroyl spermidine qui est retrouvée majoritairement (Elejalde-Palmett et al. 2015). La voie de biosynthèse des phénolamides est donc la connexion entre la voie de biosynthèse des polyamines et celles des acides hydroxycinnamiques (Figure 2). *In vitro* AtSHT et MdSHT sont capables de catalyser la formation de trihydroxycinnamoyl spermidine avec des acides hydroxycinnamiques de natures variées (Figure 1.3). Au vu de la faible spécificité des

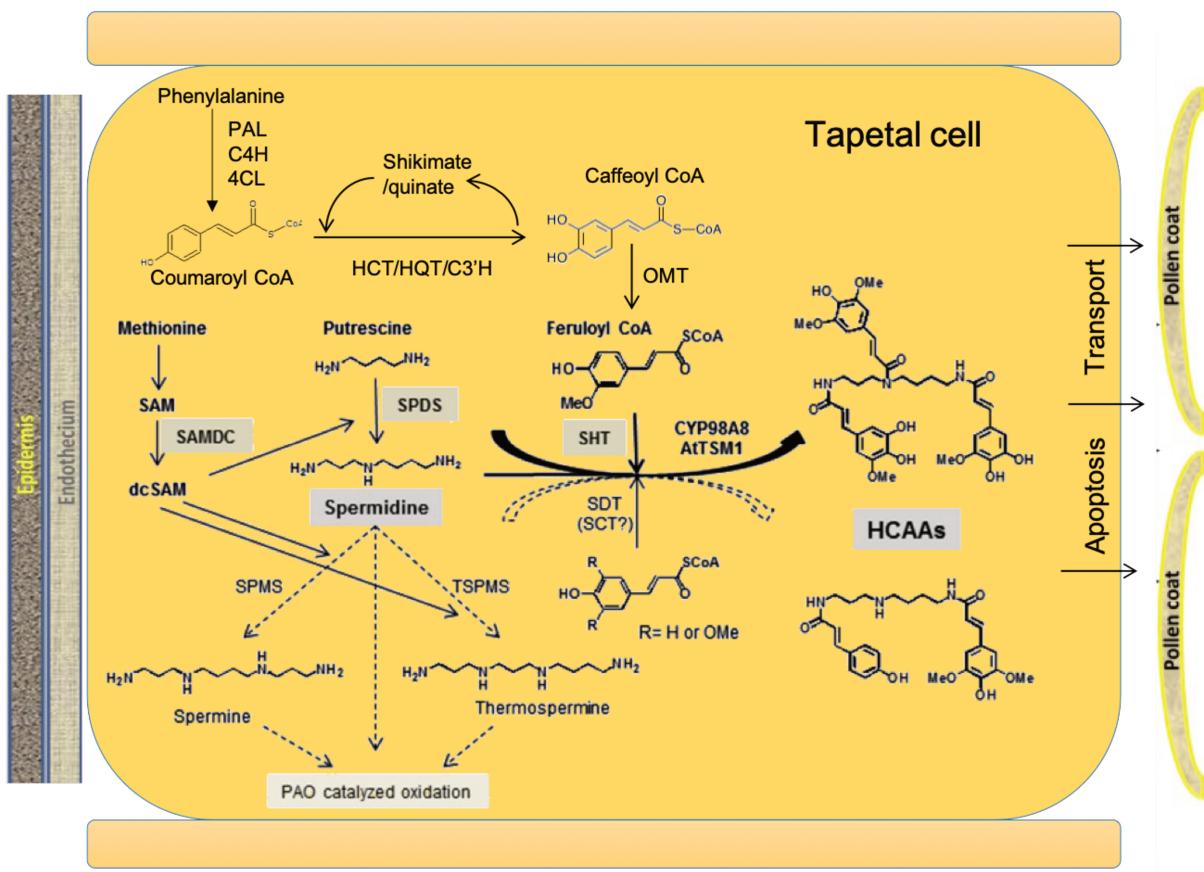


Figure 1.2 : Représentation schématique de la connexion entre la voie de biosynthèse des polyamines et le métabolisme des acides hydroxycinnamiques au niveau du manteau pollinique chez *Arabidopsis*. (D'après Fellenberg et Vogt 2012, 2015). Abréviations : SAM, S-AdénosylMéthionine ; dcSAM, décarboxy-S-AdénosylMéthionine ; SAMDC, S-adénosylMéthionine décarboxylase; SPDS, Spermidine Synthase; SPMS, Spermine Synthase; TSPMS, ThermoSpermineSynthase; SHT, Spermidine Hydroxycinnamoyl Transférase ; OMT, O-MethylTransférase ; HCAAs, HydroxyCinnamoylAcideAmides = Phénolamides ; PAO, Polyamine.

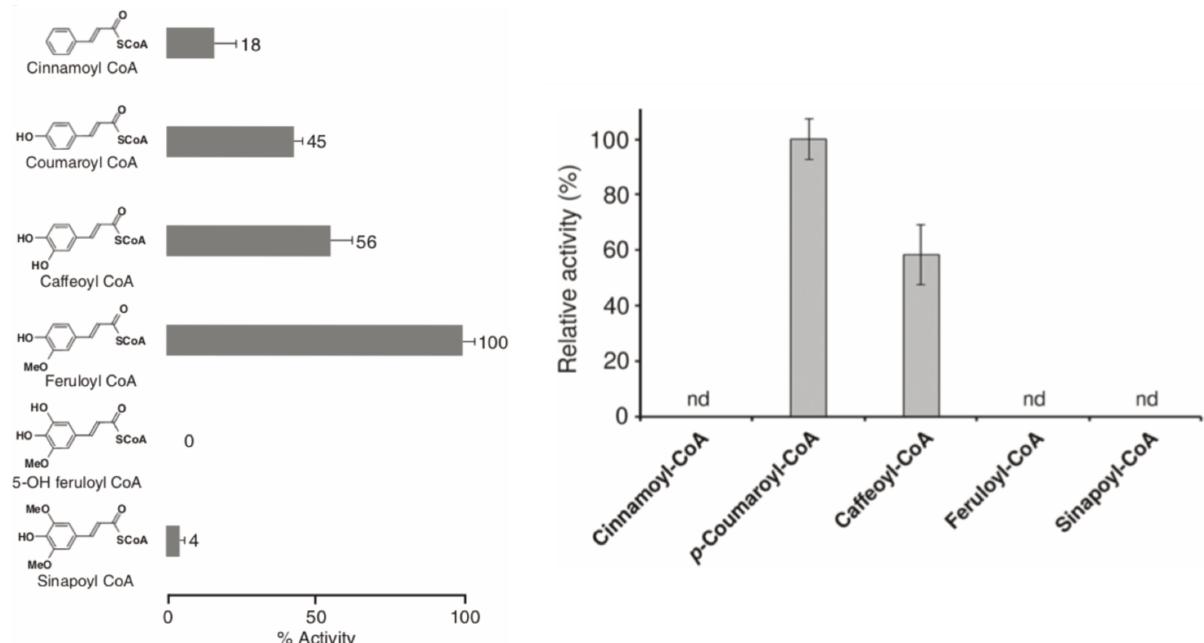


Figure 1.3 : Activité *in vitro* d'AtSHT (A) et MdSHT (B) en présence de différents esters de CoA (D'après Grieninger et al. 2009 et Elejalde-Palmett et al. 2015).

SHTs, la qualité des phénolamides accumulées sur le grain de pollen semble donc directement liée à la qualité et la quantité des polyamines et des acides hydroxycinnamiques présents au sein des cellules du tapis de l'anthère chez les différentes espèces de plantes.

D'autres enzymes comme les SCT (Spermidine DicoumaroylTransférase), SDT (Spermidine DisinapoylTransférase) et DH29 (spermidine acyltransférase) peuvent aussi acyler les fonctions amines de la spermidine mais seulement sur une ou deux des fonctions amines situées en bout de chaîne, c'est à dire les amines communément nommées « amines primaires ». Ces enzymes en plus d'être exprimées dans le tapetum, le seraient aussi dans les racines et dans les graines (Luo et al. 2009). Les SHTs sont donc les seules enzymes connues pouvant acyler les trois fonctions amines de la spermidine, elles ont donc la capacité d'acyler à la fois les amines primaires mais aussi les amines dites secondaires, situées à l'intérieur de la chaîne carbonée. La particularité des phénolamides retrouvées au niveau du grain de pollen n'est donc pas due à la nature des groupements qui substituent la spermidine mais à la capacité qu'ont les SHTs à produire des spermidines totalement substituées.

Concernant le rôle des phénolamides, Luo et al. (2009) ont montré que la SDT d'*Arabidopsis* était capable de réaliser la catalyse réverse pour former de la spermidine et du sinapoyl-CoA, à partir de disinapoyl spermidine et que cette activité était très élevée durant les premières étapes de la germination de la graine. Les spermidines di-substituées pourraient donc avoir un rôle de stockage des polyamines, qui seraient remobilisées durant la germination. En effet les polyamines jouent un rôle important dans de nombreux processus physiologiques, notamment dans l'organogénèse, la floraison ou encore la sénescence des feuilles, elles seraient aussi impliquées dans la réponse à des stress biotiques et abiotiques (Tiburcio et al. 2014) (Pál, Szalai, and Janda 2015). Les phénolamides comme des putrescine et spermidine mono et di-substituées joueraient aussi un rôle direct dans la croissance et le développement floral chez de nombreuses espèces, ainsi que chez la chicorée (Meurer, Wiermann, and Strack 1988) (Martin-Tanguy 1997). Ces molécules seraient aussi impliquées dans la défense des plantes contre les microorganismes (Muroi et al. 2009) ainsi que contre les insectes prédateurs (Kaur et al. 2010). En ce qui concerne les spermidines tri-substituées spécifiques du grain de pollen, leur rôle est encore mal connu. Le mutant *sht* mis en place chez *Arabidopsis*, ne produisant pas de tri-hydroxycinnamoyl spermine, produit des grains de pollen déformés, ce phénotype n'affectant pas la

fertilité du grains de pollen, ni sa résistance aux UVs (Grienenberger, Besseau, and Geoffroy 2009). En revanche ces molécules seraient dotées d'activités antioxydante et antifongique *in vitro* (Zamble et al. 2006)(Walters, Meurer-Grimes, and Rovira 2001) ce qui pourrait protéger le grain de pollen contre certains stress biotiques ou abiotiques, mais cela reste encore à démontrer. Ces molécules pourraient aussi jouer un rôle dans l'attraction des pollinisateurs, une étude sur la tricoumaroyl spermidine a montré que cette molécule était un important stimulant alimentaire pour insectes (Lin et Mullin 1999). Le rôle biologique de ces molécules n'étant pas encore exactement connu, et le fait qu'il existe deux séquences *SHTs* chez la chicorée, ont donc attiré toute notre attention au sein du laboratoire. Les investigations menées sur les phénolamides accumulées au niveau des grains de pollen de chicorée ont permis de mettre en évidence l'accumulation de tetracoumaroyl spermine en plus de la tricoumaroyl spermidine au niveau du manteau pollinique de la chicorée ainsi que chez de nombreuses Astéracées (notamment les travaux de thèse de Marianne Delporte, 2017), sans pour autant démontrer clairement que les deux gènes *SHT1* et *SHT2* mis en évidence chez la chicorée sont impliqués dans la synthèse de ces molécules. Afin de comprendre quelle pouvait être l'implication de ces deux gènes dans la synthèse des phénolamides accumulées dans les grains de pollen de chicorée, j'ai mis en place des expérimentations visant à obtenir la surexpression des deux gènes *CiSHT1* et *CiSHT2*, dans des chevelus racinaires de chicorée. Le but étant de valider leur fonction putative *in planta*, ainsi que de produire ces molécules par un procédé biotechnologique. Les résultats ont été publiés dans la revue *Journal of Experimental Botany* en 2018 et constituent ce premier chapitre.

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

Marianne Delporte¹, Guillaume Bernard¹, Guillaume Legrand¹, Björn Hielscher², Arnaud Lanoue³, Roland Molinié⁴, Caroline Rambaud¹, David Mathiron⁵, Sébastien Besseau³, Nicole Linka², Jean-Louis Hilbert¹ and David Gagneul^{1*}

¹ 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

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

³ Biomolécules et Biotechnologies Végétales, EA 2106, Université de Tours, 37200 Tours, France

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

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

Correspondence: david.gagneul@univ-lille.fr

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

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 synthetized 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 coats 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.

1. 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 catalyse 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 disinapoyltransferase (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.

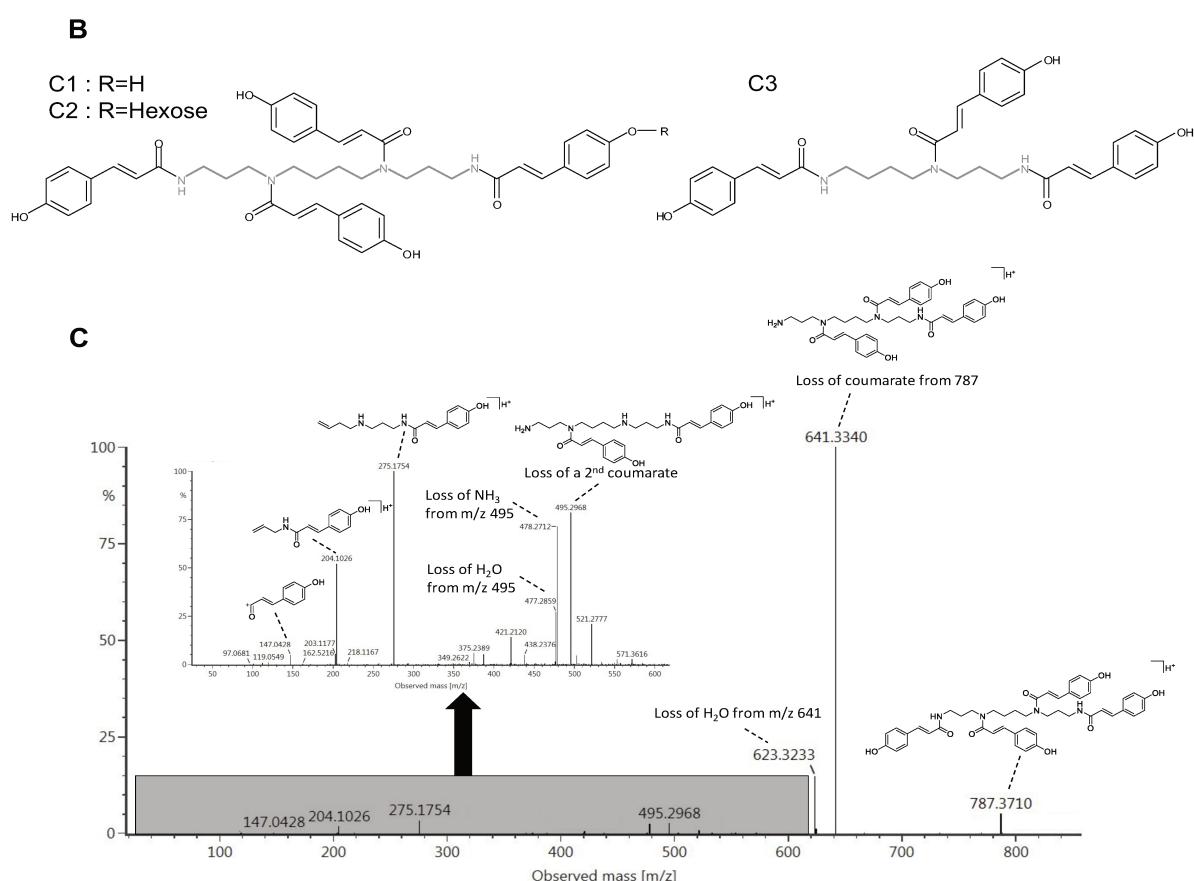
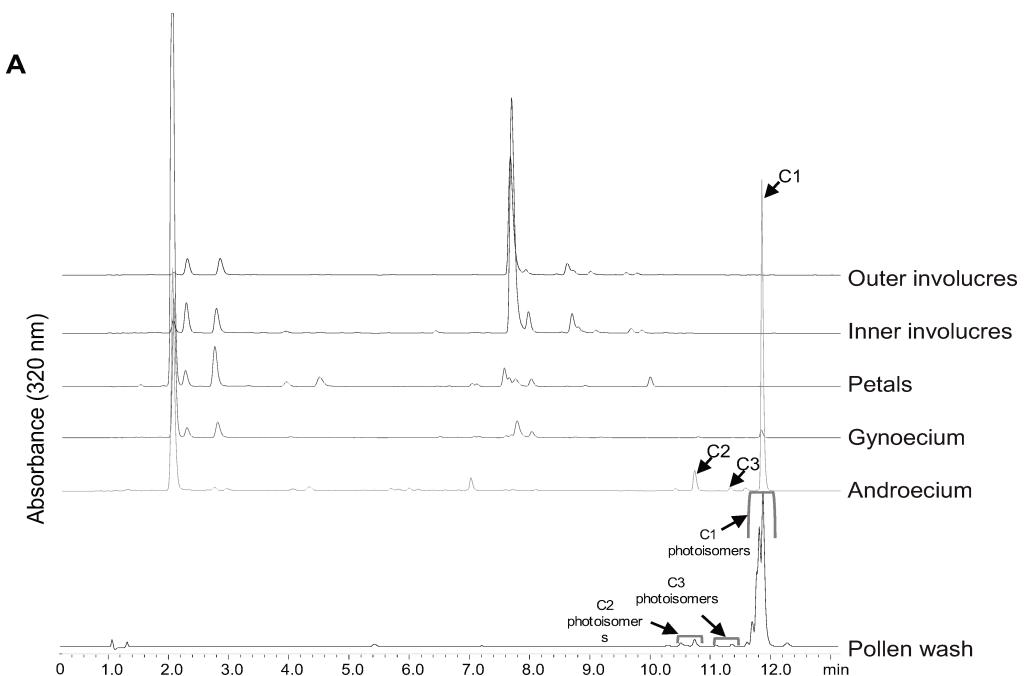


Figure 1: Identification of phenolamides present in chicory floral buds. (A) Stacked HPLC chromatograms of methanolic extracts obtained from different parts of the floral buds, and of a methanol wash of pollen. (B) Structures of the compounds identified as C1 (N^1,N^5,N^{10},N^{14} -tetracoumaroyl spermine), C2 (N-coumaroyl(glucose)- N',N'',N''' -tricoumaroyl spermine), and C3 (N^1,N^5,N^{10} -tricoumaroyl 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.

2. Results

2.1. 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 -C⁶ and N^{10} -C¹¹ bonds in the spermine moiety. Similarly, C2 and C3 were identified as *N*-coumaroyl(glucose), $N^{\prime},N^{\prime\prime},N^{\prime\prime\prime}$ -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).

Order	Family	Subfamily	Genus	Tetra-HCSpM
	Styliaceae		<i>Stylium</i> *	nd
	Argophyllaceae		<i>Argophyllum</i> *	nd
	Calyceraceae		<i>Calycera</i> *	nd
Asterales		Barnadesioideae	<i>Barnadesia</i> *	✓
		Mutisioideae	<i>Onoseris</i> *	✓
			<i>Mutisia</i> *	nd
		Stifftieae	<i>Stifftia</i> *	✓
		Wunderlichioideae	<i>Wunderlichia</i> *	nd
			<i>Nouelia</i> *	✓
		Gochnatieae	<i>Gochnatia</i> *	✓
	Asteraceae	Carduoideae	<i>Centaurea</i>	✓
			<i>Cirsium</i>	✓
			<i>Carduus</i>	✓
			<i>Cynara</i>	✓
			<i>Picris</i> *	✓
			<i>Leontodon</i>	✓
			<i>Lapsana</i>	✓
			<i>Taraxacum</i>	✓
			<i>Cichorium</i>	✓
			<i>Lactuca</i>	✓
		Cichorioideae	<i>Leontopodium</i>	✓
			<i>Senecio</i>	✓
			<i>Bellis</i>	✓
			<i>Leucanthemum</i>	✓
			<i>Matricaria</i> *	✓
			<i>Achillea</i>	✓
			<i>Chrysanthemum</i>	✓
			<i>Ambrosia</i>	✓
			<i>Sanvitalia</i>	✓
			<i>Helianthus</i>	✓
	Astroideae	Asteroideae		

↓

Dispersal out
of South
America

Figure 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.

2.2. 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 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.

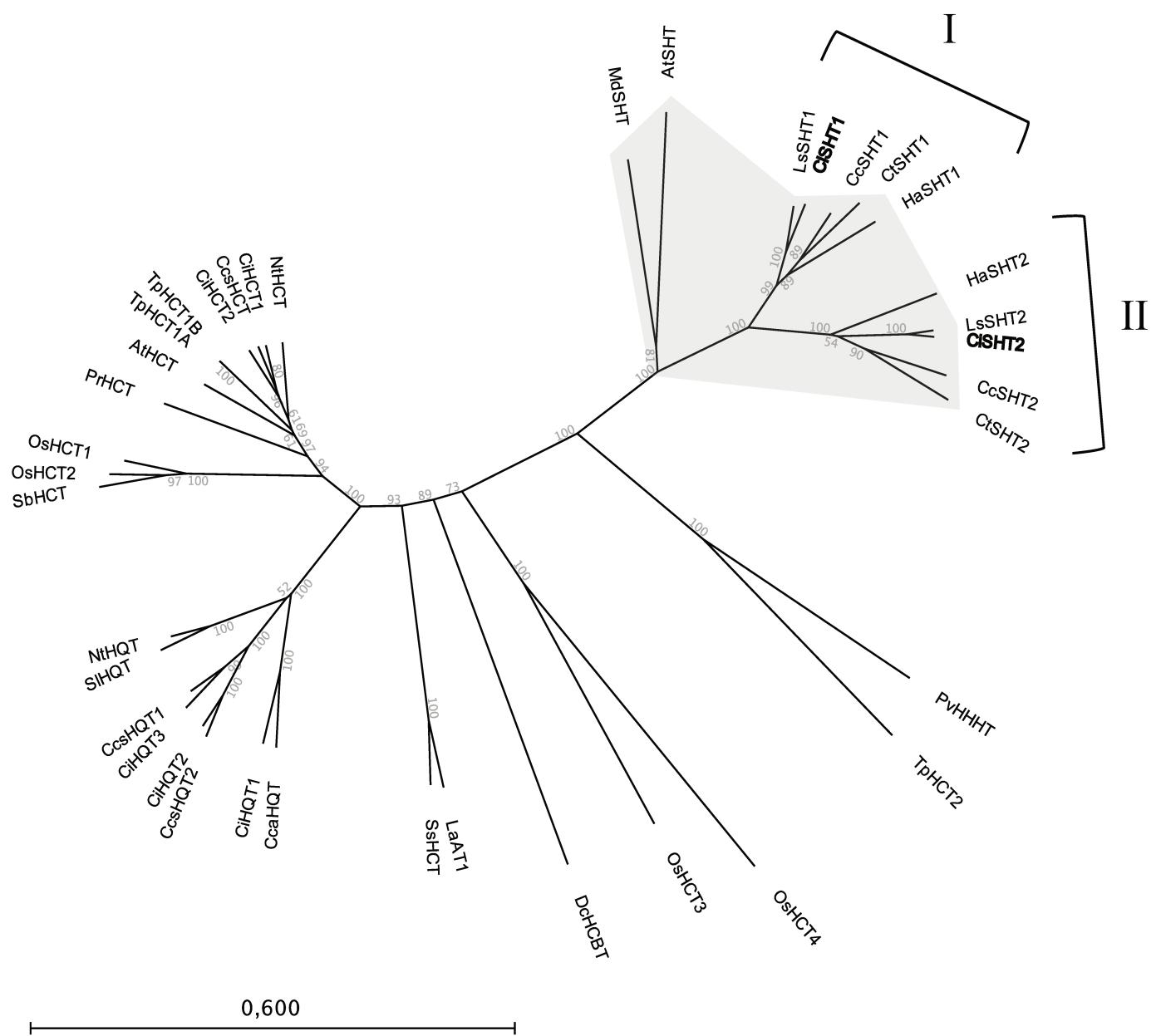


Figure 3: Phylogenetic analysis of the ten SHT-like amino acid sequences identified in five species of the Asteraceae family and other biochemically characterized members of the BAHD family clade Vb. Two clusters of SHT-like sequences are indicated for the Asteraceae sequences. See supplementary table S2, for details of protein names and accession numbers.

2.3. Identification of candidate genes for tetracoumaroyl 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*, *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 motives 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.

2.4. 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

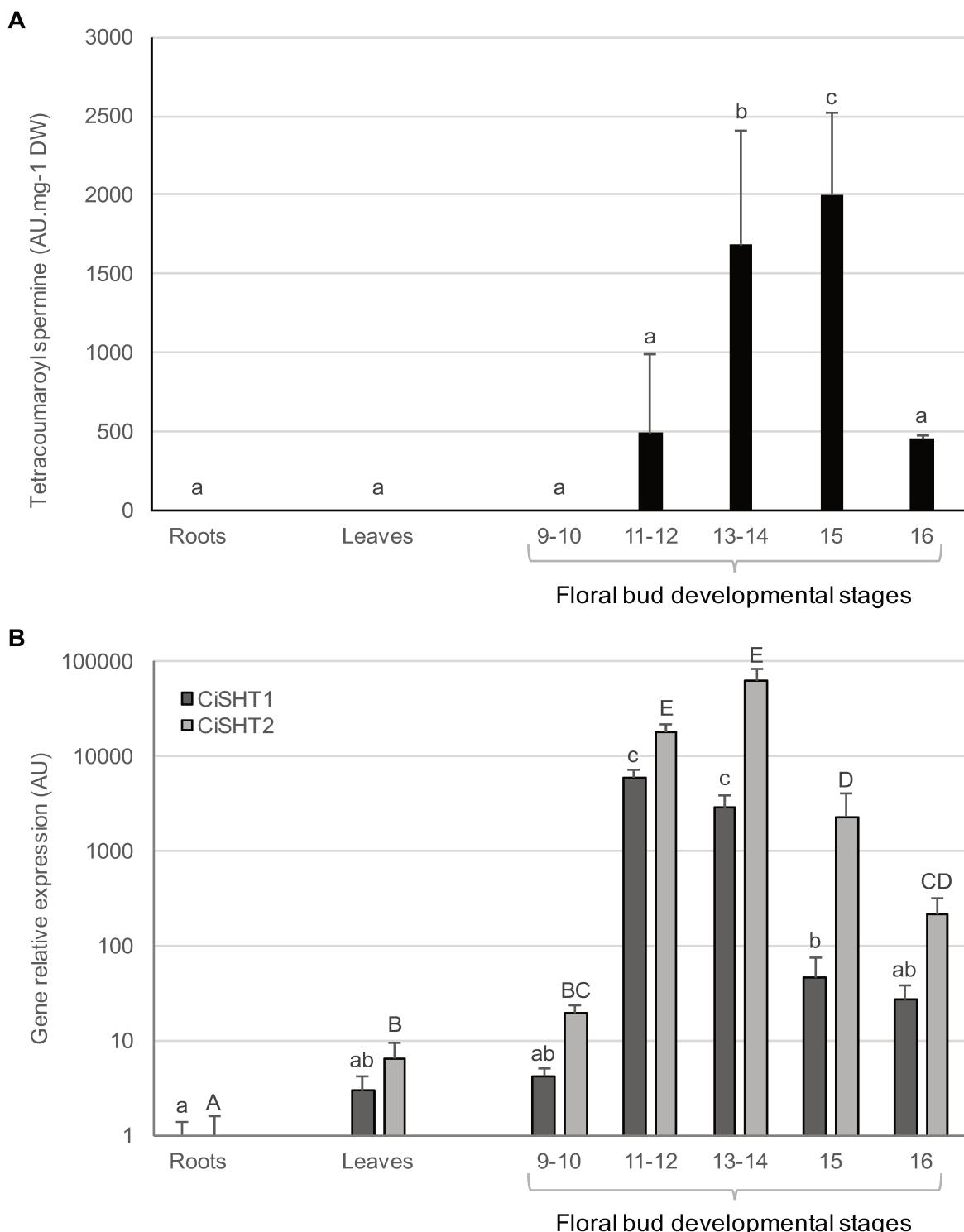


Figure 4: (A) Tetracoumaroyl spermine contents and (B) expression profiles of *CiSHT1* and *CiSHT2* in roots, leaves, and in floral buds at different developmental stages. Tetracoumaroyl spermine was not detected in roots and leaves. Gene expression was determined by qPCR: expression levels are normalized 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).

recently observed in *M. chamomilla* (Park *et al.*, 2017). In an attempt to assign 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 10000-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).

2.5. *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 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

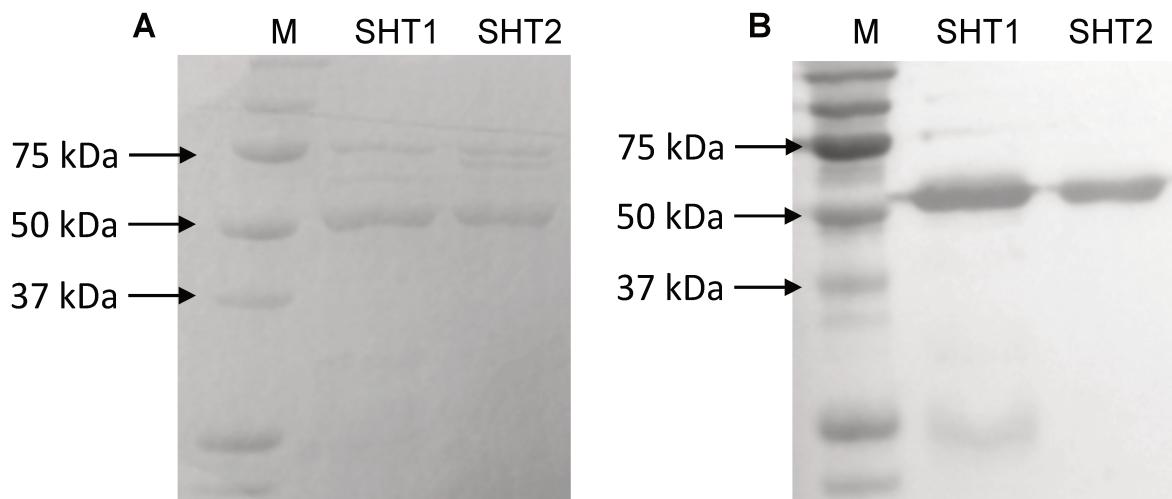


Figure 5: Expression and purification of recombinant CiSHT1 and CiSHT2. (A) SDS-PAGE separation of recombinant proteins. The gel was stained with Coomassie blue. (B) Immunoblot analysis confirming the expression of the recombinant proteins. The proteins were detected with an anti-His antibody. M, protein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

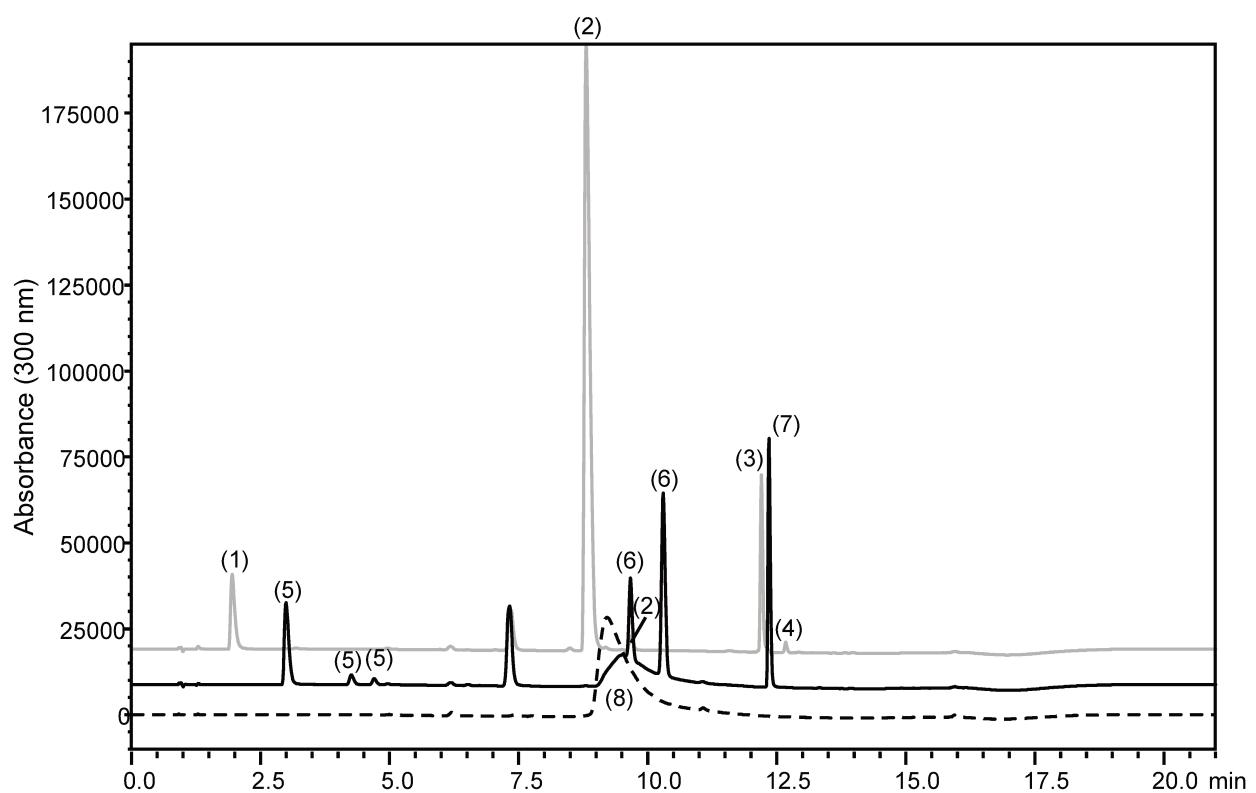


Figure 6: HPLC chromatograms of in vitro assays of recombinant CiSHT2 in the presence of coumaroyl-CoA and spermine (grey line) or spermidine (black line), or without an acyl acceptor (dotted line). The identities of the numbered peaks were confirmed by mass spectrometry: (1) monocoumaroyl spermine;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

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 (Grienberger *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.

2.6. Overexpression of CiSHT2 in chicory hairy roots leads to production of spermine derivates

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 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*.

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

Hairy root line	Phenolamide content (nmol mg ⁻¹ 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

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

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

The *Arabidopsis* sht T-DNA insertion mutant previously described as lacking trihydroxycinnamoyl spermidine conjugates in the pollen coat (Grienenerger 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 (Grienenerger 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* (Grienenerger 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 triferaloyl 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*.

2.8. CiSHT1 and CiSHT2 are targeted to the nucleo-cytoplasmic 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* (Grienenerger et al., 2009). The accumulation of phenolamides in the

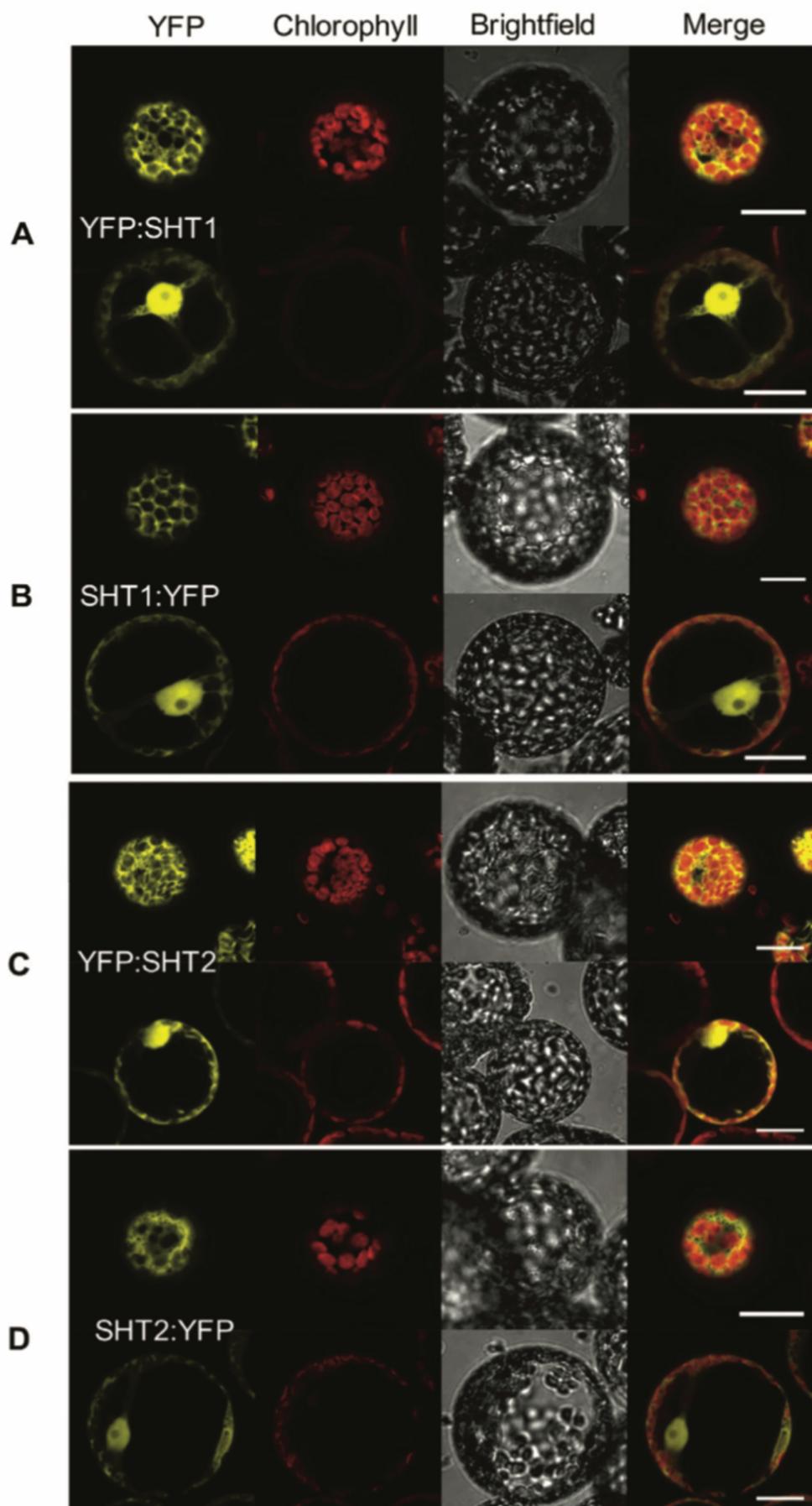


Figure 7: Subcellular imaging of yellow fluorescent protein (YFP)-tagged constructs transiently expressed in *N. benthamiana* leaves. The images of the leaf cell protoplasts using confocal microscopy show the fluorescence pattern for (A) YFP::SHT1, (B) SHT1::YFP, (C) YFP::SHT2, and (D) SHT2::YFP. Scale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

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 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).

2.9. 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

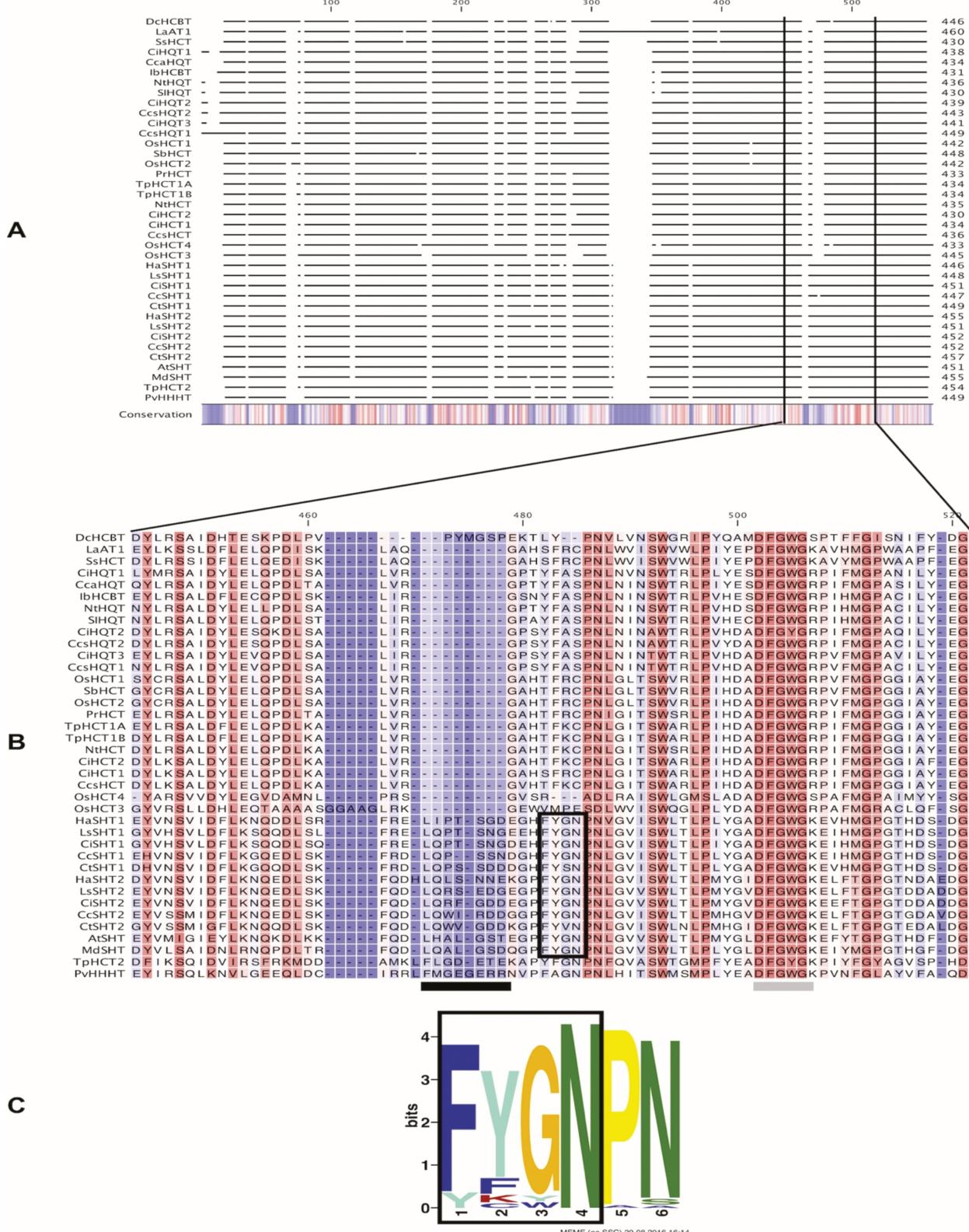


Figure 8 : Sequence analysis of the SHT-like enzymes. (A) Full alignment of the ten SHT-like sequences identified in Asteraceae species with other characterized members of the BAHD family clade Vb. (B) Detail of a particular portion of the alignment showing the BAHD motif DFGWG (grey bar), an amino acid 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.

emergence of 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).

3. 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 synthetized 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/quinate 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-likes 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 spermamines *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/CiSHT2* 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 spermamines 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:quinate 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 phaselic 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.

4. Materiel and methods

4.1. 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 *sht* mutant line SALK_055511, homozygous for T-DNA insertion, were used for complementation and overexpression experiments (Grienberger *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).

4.2. 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-µm 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-µm filter).

4.3. Metabolite analyses 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 × 4.6 mm Kinetex 2.6-µm 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⁻¹ 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). Separation was performed using a Waters ACQUITY HSS T3 C18 column (150×2.1 mm, $1.8\text{-}\mu\text{m}$) with a flow rate of 0.4 ml min^{-1} at 55°C . The injection volume was $5 \mu\text{l}$. 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^{-1} and 800 l h^{-1} , respectively. Targeted data collection was carried in selected ion monitoring (SIM) mode for *N*1,*N*5,*N*10-tricoumaroyl spermidine ($[\text{M}+\text{H}]^+$; m/z 584; RT=10.8 min) and *N*1,*N*5,*N*10,*N*14-tetracoumaroyl spermine ($[\text{M}+\text{H}]^+$; 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 quadruple 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 (100×2.1 mm, $1.7\text{-}\mu\text{m}$) 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 µl 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⁻¹ 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⁻¹ and 800 l h⁻¹, respectively. Typically, the capillary voltage was 3 kV, the sampling cone voltage was 20 V, and the source offset was 20 V. Lock mass corrections using [M+H]⁺ at *m/z* 556.2771 of a leucine-enkephalin solution (100 pg µl⁻¹ 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 40000 (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.

4.4. 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 µg of DNase-treated total RNA using an iScript cDNA synthesis kit (BioRad).

4.5. 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.

4.6. 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.ressources-electroniques.univ-lille.fr/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.

4.7. 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 µl containing 1 µg of purified protein, 50 mM phosphate buffer (pH 9.0), 400 µM acyl donor, and 2 mM acyl acceptor. After incubation at 28 °C for 1 h, the reaction was stopped by adding 15 µl acetonitrile and 1 µl 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 µg of protein were added in 50 mM phosphate buffer, pH 9.0, in a final volume of 50 µl. The production of coumaroyl-CoA was evaluated after incubation at 28 °C for 3 h. The reaction mixtures were then centrifuged at 14000 g for 5 min and filtered through a 0.45-µm 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).

4.8. 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⁻¹) in flasks on an orbital shaker at an agitation speed of 110 rpm with a light source providing about 5 W m⁻² 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.

4.9. 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 Spel 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* *sht* 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.

4.10. 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 phosphoinotricine. 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 form *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.

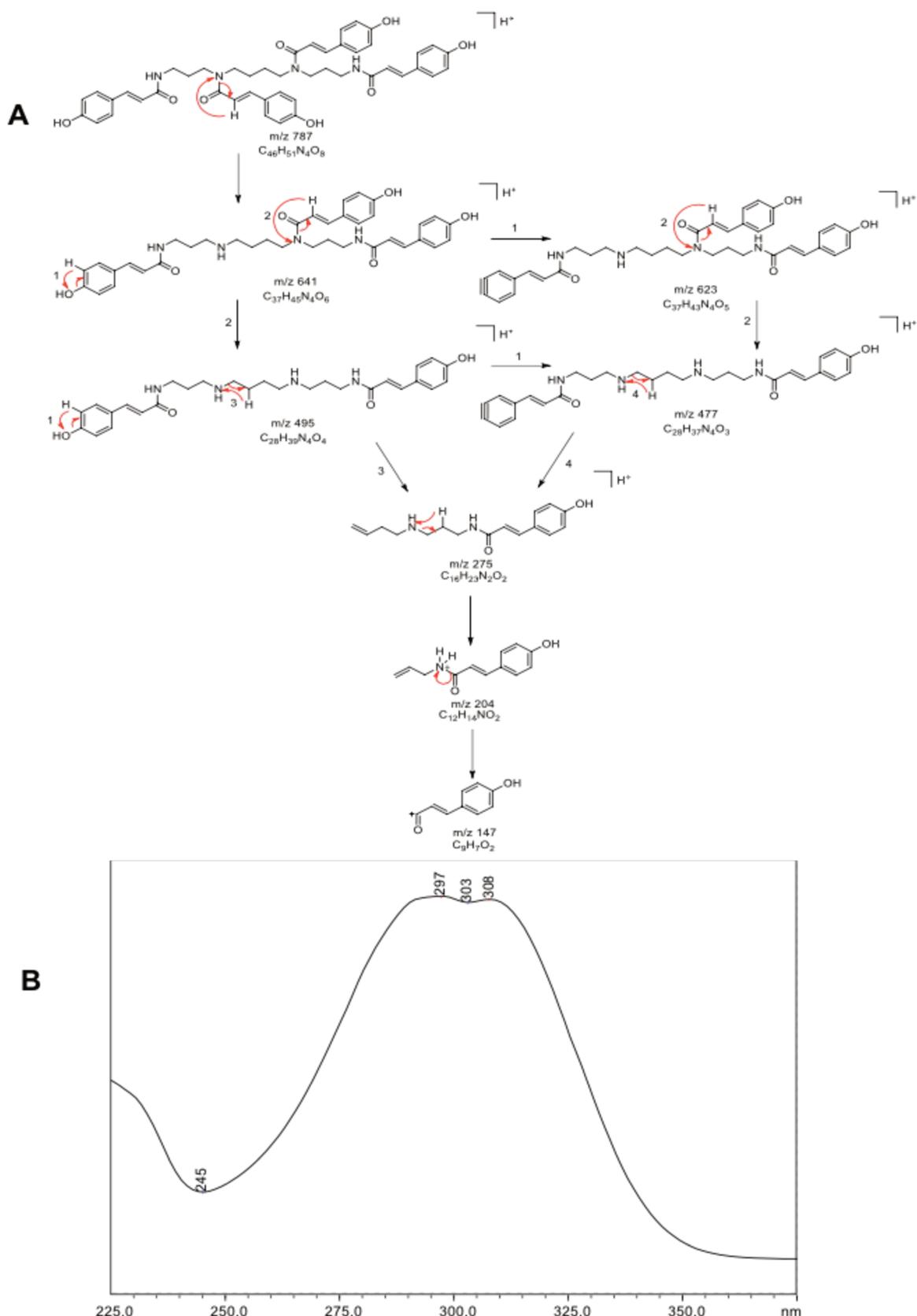
4.11. Sequence analysis

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).

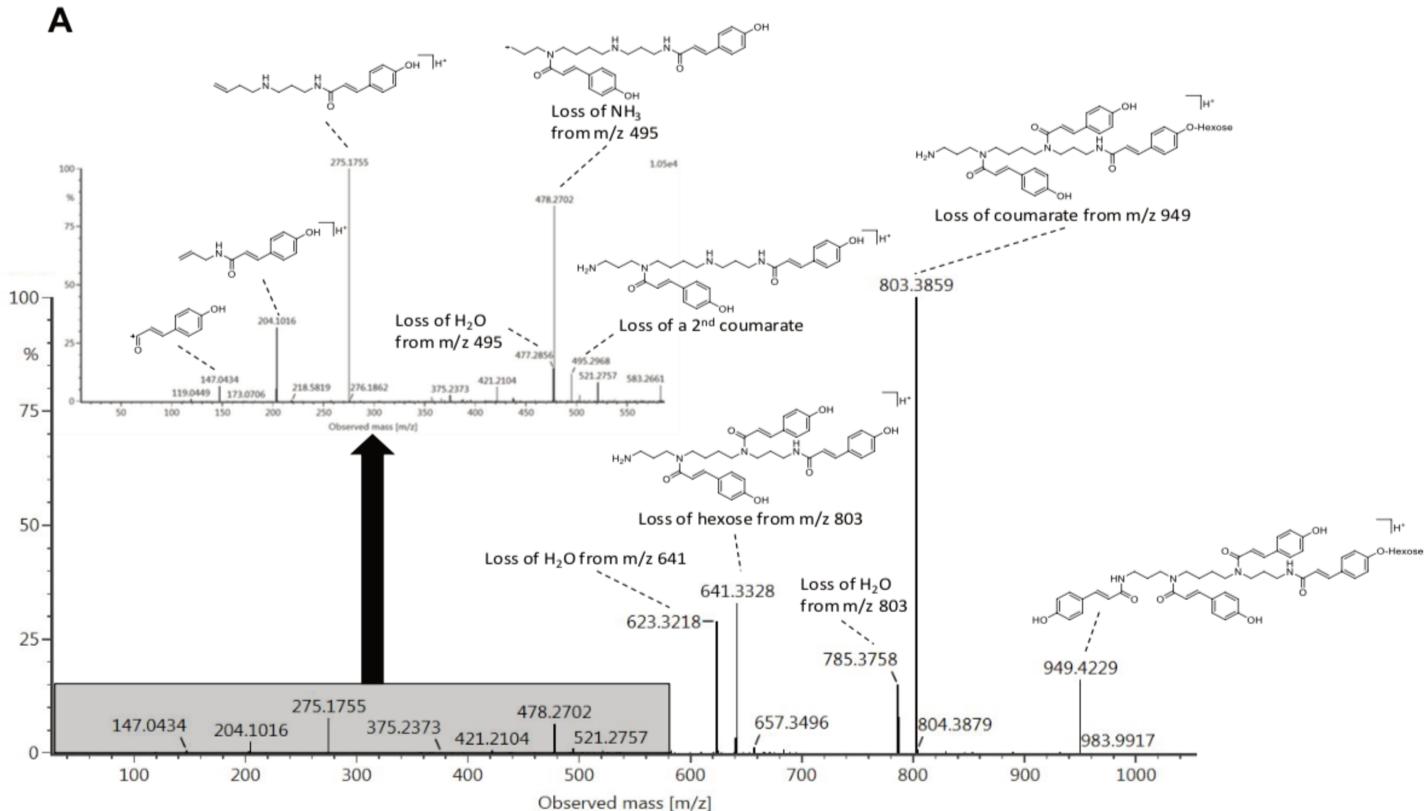
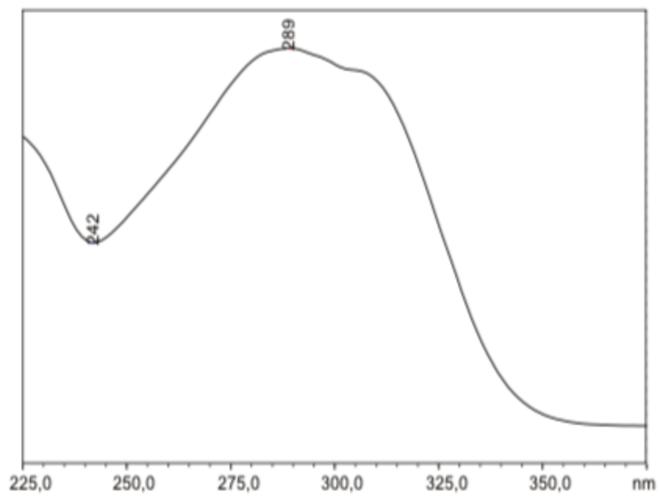
Acknowledgements

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 Albiotech grant (2016–2020) obtained from the CPER/FEDER program.

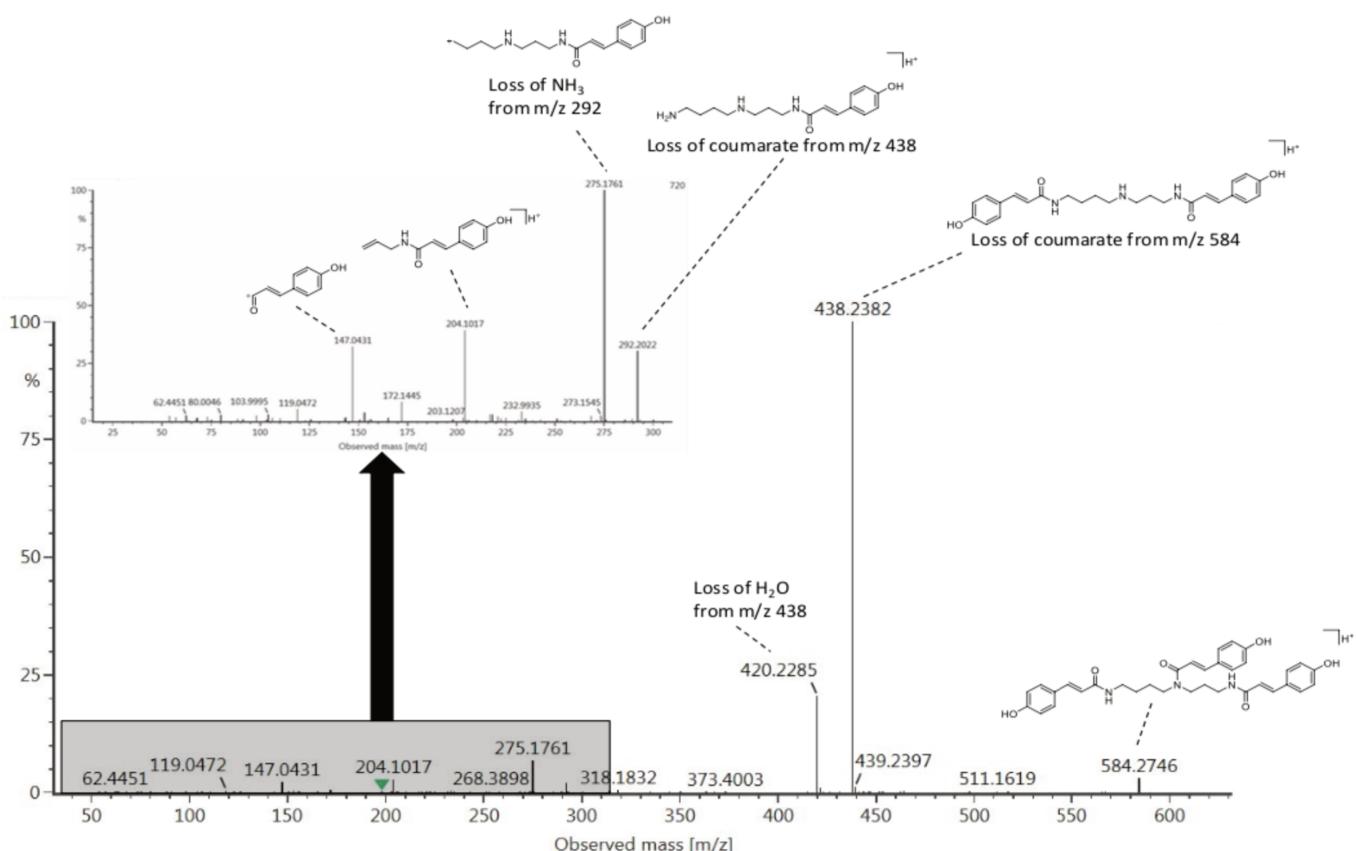
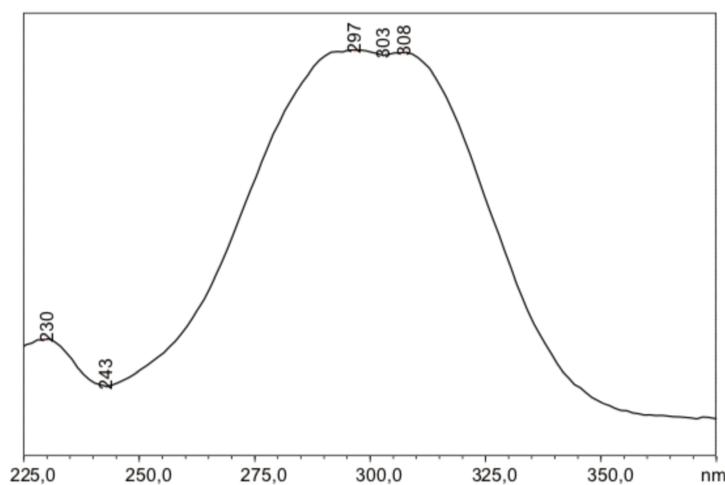
Supplementary data



Supplementary Figure S1: Proposed fragmentation scheme of the compound designated as C1 (A) and its UV spectra (B). C1 is the main phenolamide present in chicory floral buds. Fragmentation scheme (A) was suggested according to LC/MS/MS data presented in Fig. 1. UV spectra was recorded using a DAD (B).

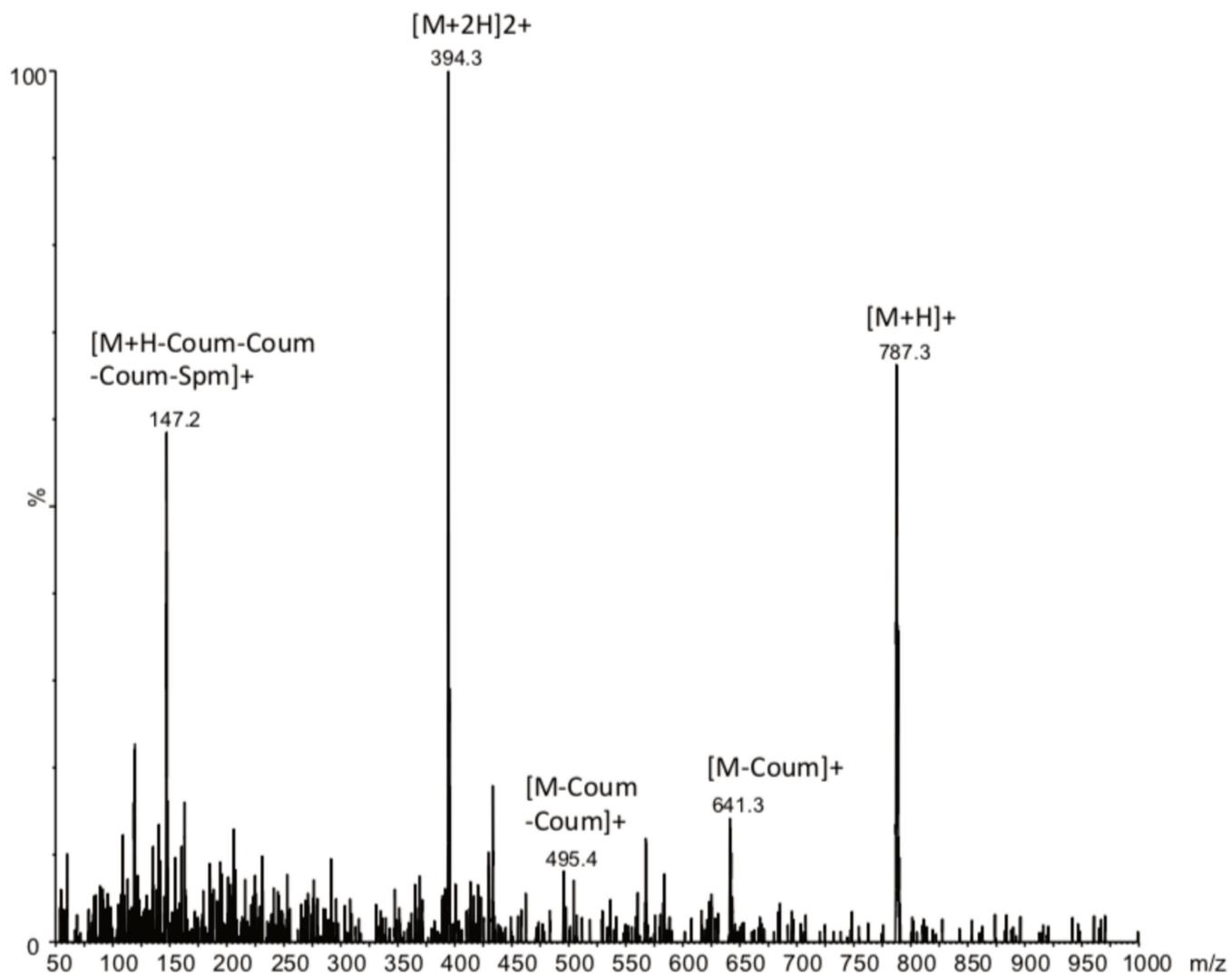
A**B**

Supplementary Figure S2: LC/MS/MS fragmentation of the compound designated as C2 (A) and its UV spectra (B). C2 is a minor phenolamide (together with C3) found in chicory floral buds. UV spectra was recorded using a DAD (B).

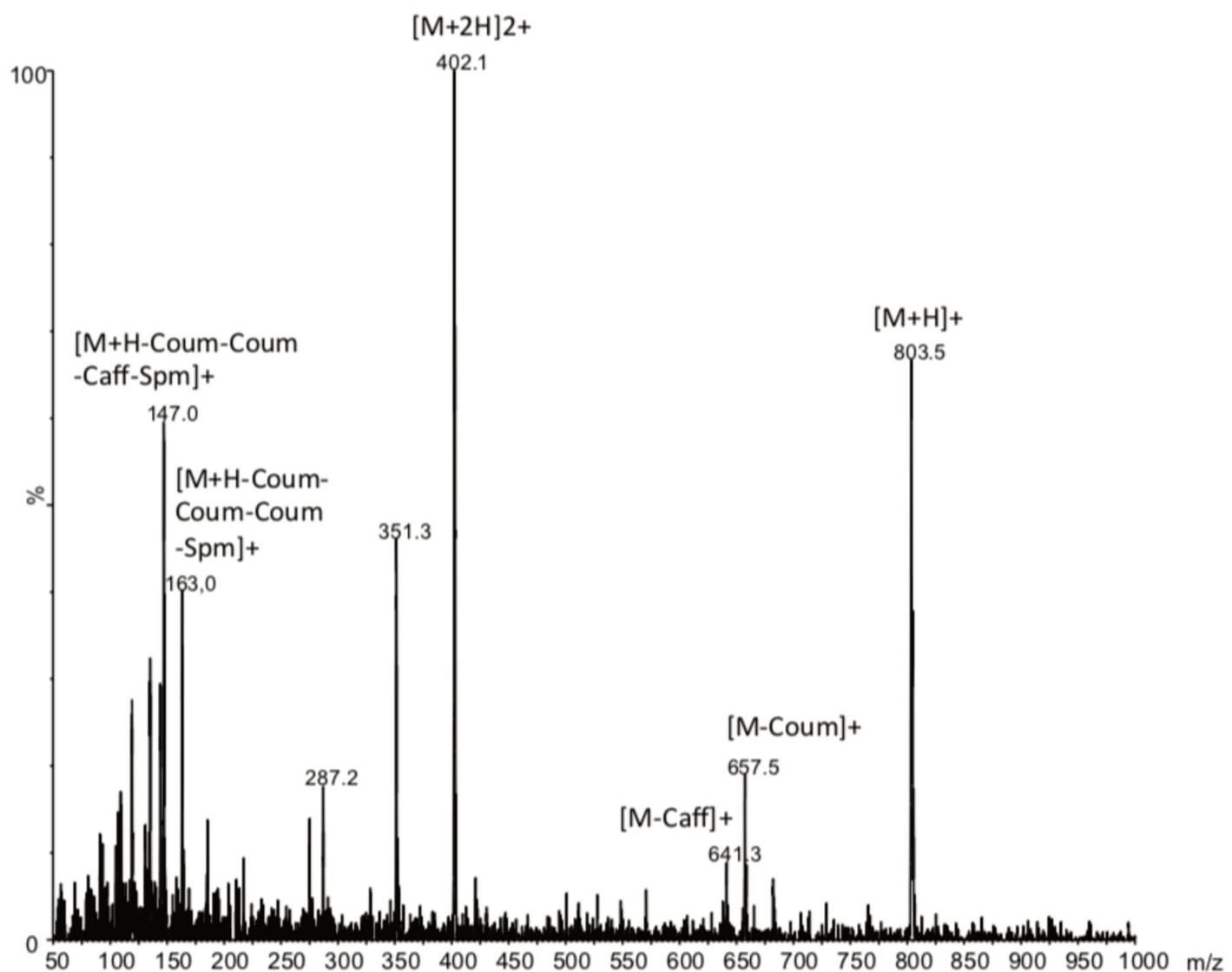
A**B**

Supplementary Figure S3: LC/MS/MS fragmentation of the compound designated as C3 (A) and its UV spectra (B). C3 is a minor phenolamide (together with C2) found in chicory floral buds. UV spectra was recorded using a DAD (B).

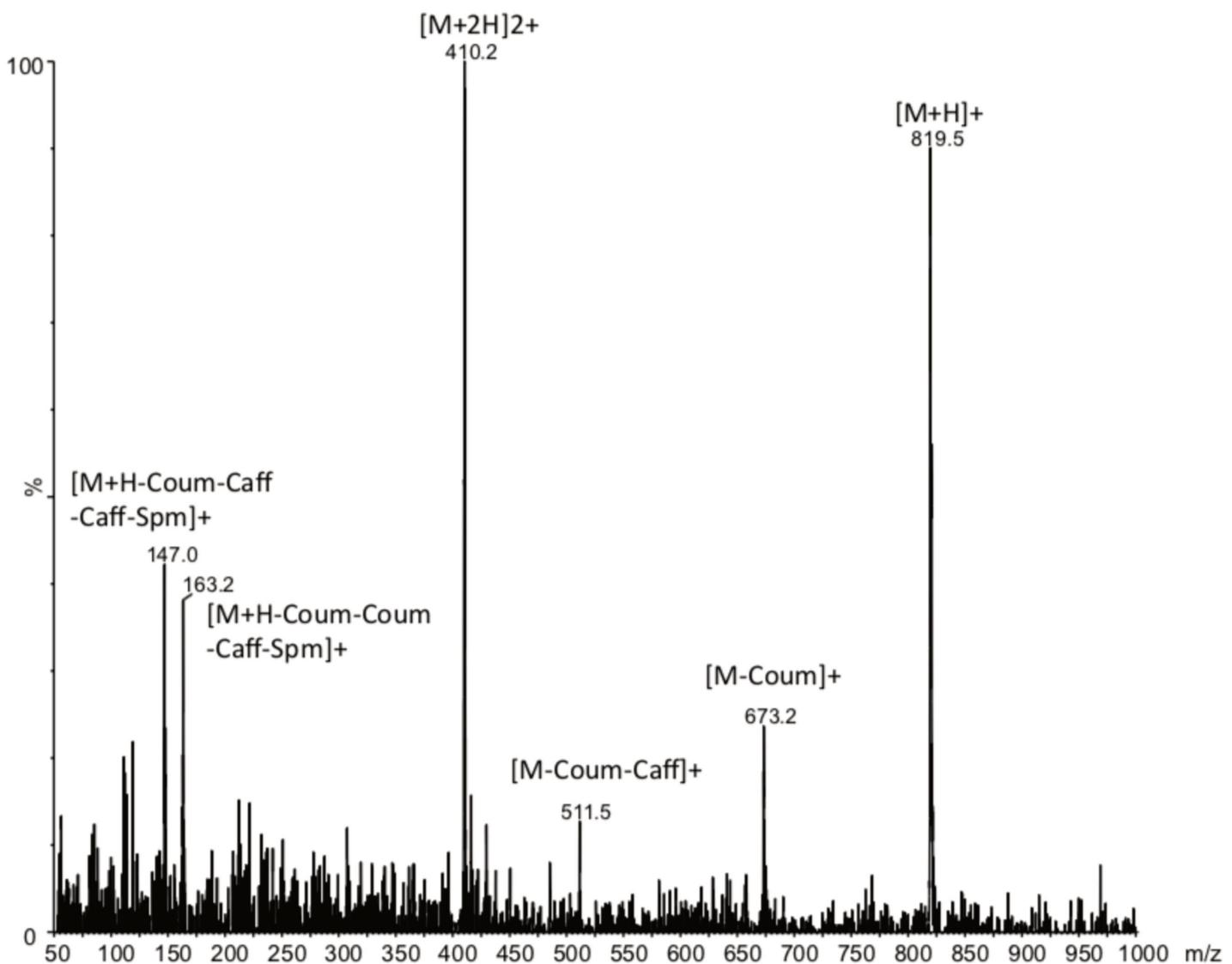
Supplementary Figure S4: LC/MS/MS fragmentation of spermine derivatives found in Asteraceae flower buds. Members of the Asteraceae family were screened for the presence of spermine derivatives in extract of floral buds or pollen methanolic wash (see Fig. 2). By using LC/MS/MS, four different compounds were identified: tetracoumaroyl spermine (1), tricoumaroyl caffeoyl spermine (2), dicoumaroyl dicaffeoyl spermine (3) and tetracaffeoyl spermine (4).



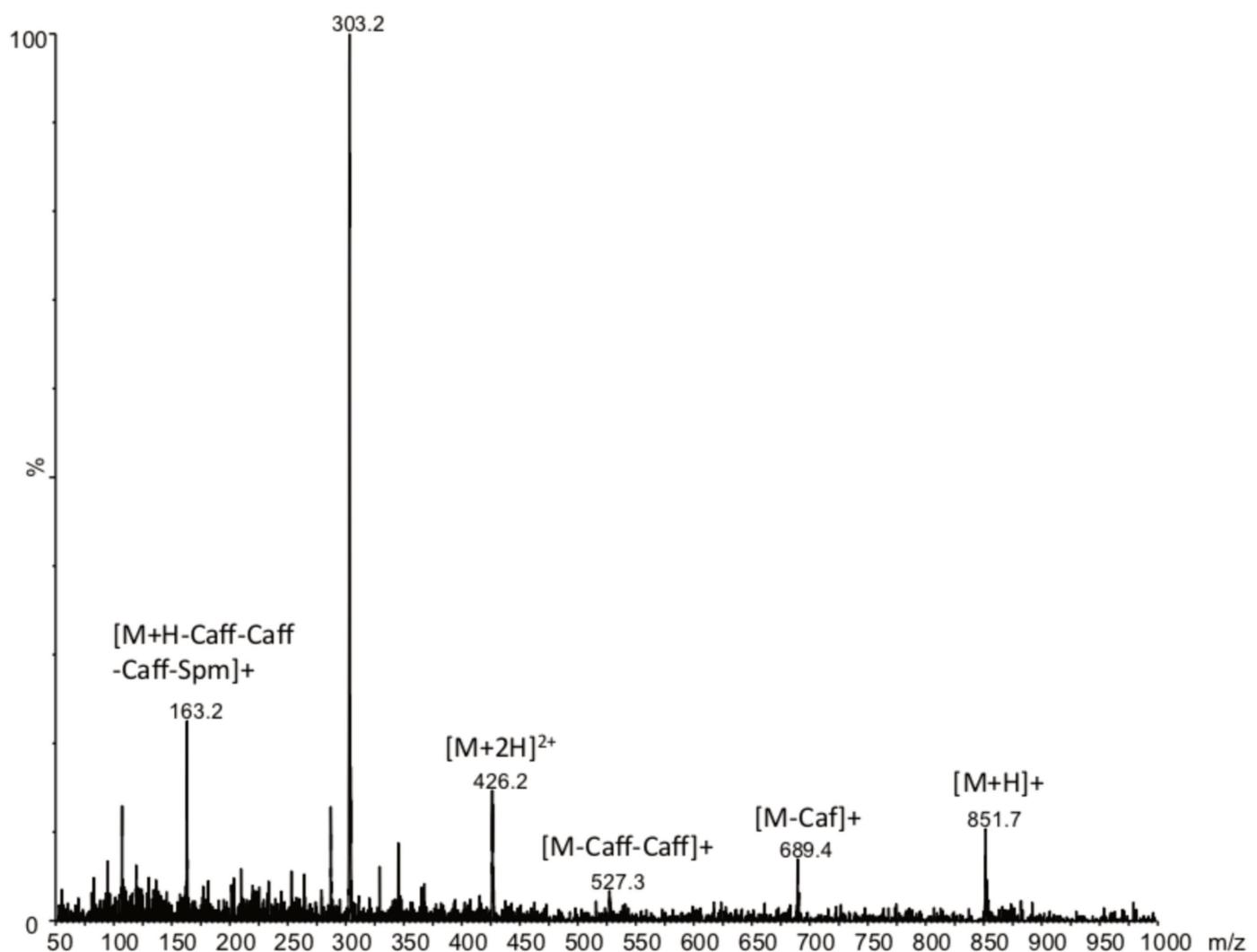
2: Tricoumaroyl caffeoyl spermine

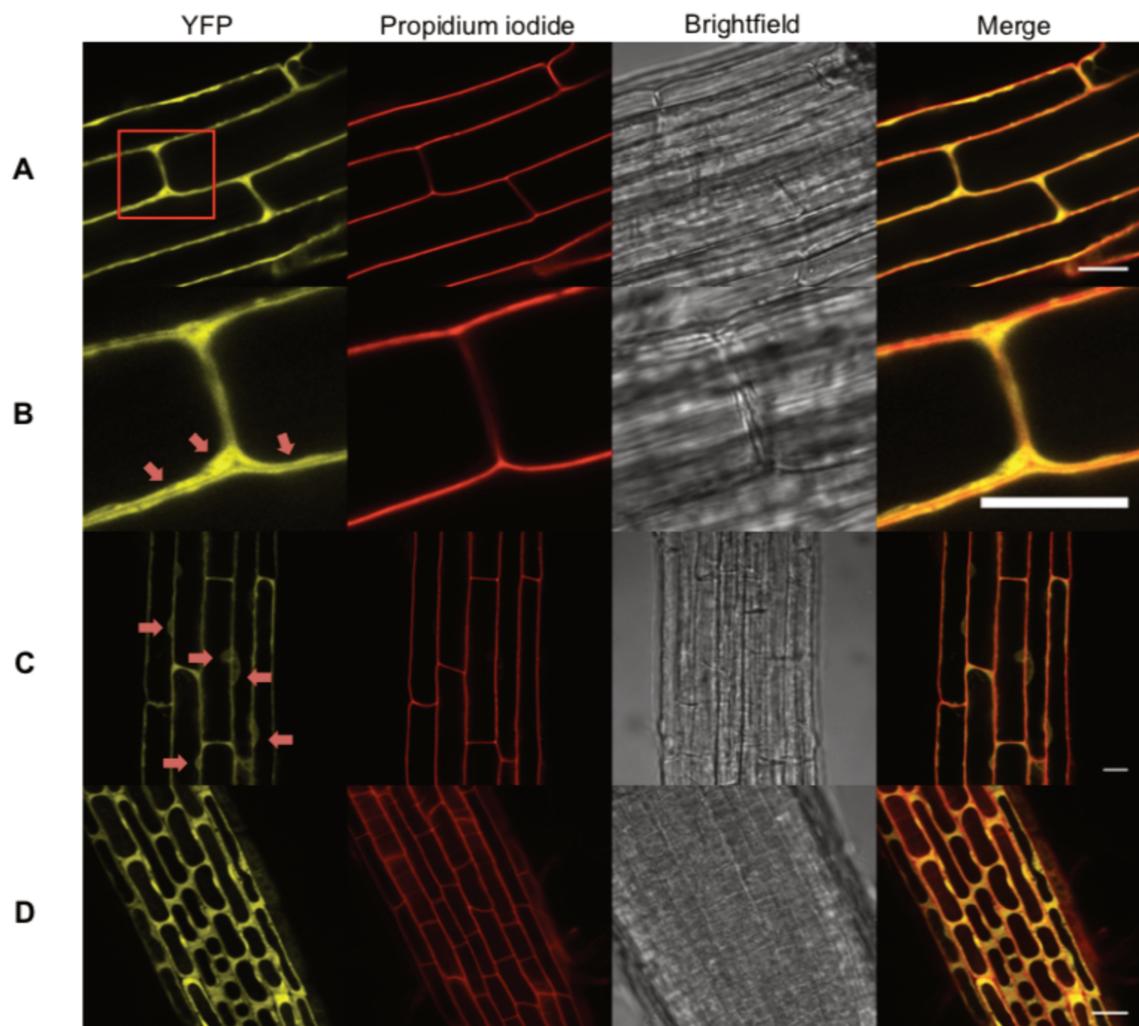


3: Dicoumaroyl dicaffeoyl spermine



4: Tetracaffeoyl spermine





Supplementary Figure S5: Subcellular imaging of YFP-tagged constructs stably expressed in *A. thaliana*. Confocal microscopy observations of 7 days old seedlings root cells show the fluorescence pattern for YFP: SHT1. A&B: Highlighting of a fluorescent signal distinct from the cell wall; C: Observation of fluorescent nucleus-like structures; D: Signal localization in the cytoplasm after vacuole retraction by plasmolysis. Bars = 20 μ m

Supplementary Table S1 : Sequence of primers used in this study. CLATH: Clathrin adaptor complex subunit, TIP41: TIP41 like protein

Primer sequence 5' to 3'		Target	Application
TTACCATGCCCATGTATCA CAAGGCTAACCCCTCCACAA	Fwr Rev	CiSHT1	
GAATGGGCTAGAGCTTGTGCG TCGATGCTTGAACATGTGGT	Fwr Rev	CiSHT2	
GTTGGGTGCGCATCTCTAAT AGCTCCGGCAGCTTTACTT	Fwr Rev	TIP41	qRT-PCR
TGCTTCCGCCATCTACTTT TCCCAAGTTCCCTTGTTGC	Fwr Rev	Clath	
GGGGACAAGTTGTACAAAAAAGCAGGCTCCATGAAGGTTCCAACAAGACTTCT GGGGACCACTTGTACAAGAAAGCTGGGTCTAACATTCCATATCAATATCTTGGT	Fwr Rev	CiSHT1	Cloning for expression in <i>E.Coli</i>
GGGGACAAGTTGTACAAAAAAGCAGGCTCCATGAGAGTCACCGTGAAATCTC GGGGACCACTTGTACAAGAAAGCTGGGTCTAACATTCCATATCAATGCTTGATAGAAC	Fwr Rev	CiSHT2	
GGGGACAAGTTGTACAAAAAAGCAGGCTCCACCATGAAGGTTCCAACAAGACTTCT GGGGACCACTTGTACAAGAAAGCTGGGTCTAACATTCCATATCAATATCTTGGT	Fwr Rev	CiSHT1	
GGGGACAAGTTGTACAAAAAAGCAGGCTCCACCATGAAGGTTCCAACAAGACTTCT GGGGACCACTTGTACAAGAAAGCTGGGTACATTCCATATCAATATCTTGGTAGAA	Fwr Rev	Without STOP	Cloning for expression in <i>planta</i>
GGGGACAAGTTGTACAAAAAAGCAGGCTCCACCATGAGAGTCACCGTGAAATCTC GGGGACCACTTGTACAAGAAAGCTGGGTCTAACATTCCATATCAATGCTTGATAGAAC	Fwr Rev	CiSHT2	
GGGGACAAGTTGTACAAAAAAGCAGGCTCCACCATGAGAGTCACCGTGAAATCTC GGGGACCACTTGTACAAGAAAGCTGGGTATCTCAATGCTTGATAGAACACC	Fwr Rev	CiSHT2	
ATGTATGATAATTGAGCTCTTTGGTAGCGAATGCA TTGTGATATCACTAGTTAGCTAATTCTTAAGAAAAAC	Fwr Rev	TA29 promoter	pB2GW7-TA29 construction

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, Glevarec 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.
- Eliašová A, Poracká V, Pal'ove-Balang P, Imrich J, Repcák M.** 2012. Accumulation of tetracoumaroyl spermine in *Matricaria chamomilla* during floral development and nitrogen deficiency. *Zeitschrift fur 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.
- Grienenerger E, Besseau S, Geoffroy P, Debayle D, Heintz D, Lapierre 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, Vuylsteke C, Mörcchen 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, Knevitt 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 dicaffeoylquinic 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, Yukiyoshi 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 NK1 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, Grienberger 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, Jirschitzka J, Porta T, et al. 2015. The last step in cocaine biosynthesis is catalyzed by a BAHD acyltransferase. *Plant Physiology* 167, 89–101.

- Schmidt GW, Jirschitzka 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 phasic 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 NK1 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.

Éléments complémentaires

Les résultats publiés dans l'article ont montré que lorsque le gène *CiSHT2* est surexprimé dans des chevelus racinaires de chicorée, ces chevelus racinaires sont capables de produire des dérivées de spermine (mono, di, tri coumaroyl spermine) avec une production de dicoumaroyl spermine majoritaire, mais ils sont aussi capables de produire des dérivés de spermidine (mono, di, tri coumaroyl spermidine). Ces résultats sont concordants avec les tests d'activités *in vitro*, pour lesquels il a été montré que *CiSHT2* était capable de produire tous les dérivés de spermine et tous les dérivés de spermidine. Cependant, le composé majoritaire produit par les chevelus racinaires transformés est la dicoumaroyl spermine lorsqu'on fournit aux cultures 1,5 mM de spermine pendant 6 jours. Les résultats obtenus n'étant pas ceux espérés, à savoir la production de tetracoumaroyl spermine, cela nous a incité à mener d'autres expériences afin de déterminer si l'accessibilité des substrats n'était pas la raison de cette non-production. D'autres expériences ont donc été réalisées sur les chevelus racinaires. L'activité des enzymes étant dépendante de la disponibilité des substrats et les substrats étant spermine, spermidine et coumaroyl-CoA, les cultures ont été supplémentées avec 1,5 mM de spermidine, 1,5 mM de spermine et 1 g/L de phénylalanine, 1,5 mM de spermidine et 1 g/L de phénylalanine et simplement avec 1 g/L de phénylalanine, durant 6 jours. La biosynthèse du coumaroyl-CoA étant initiée par la phénylalanine. Les résultats des analyses HPLC sont représentés dans la Figure 1.5A, pour faciliter la compréhension des chromatogrammes les aires de chaque dérivé de spermine et de spermidine ont été reportés dans la Figure 1.5B.

La spermine est une molécule symétrique possédant 4 azotes, il existe donc 2 types de mono coumaroyl spermine (Figure 1.4), 4 types de dicoumaroyl spermine, et 2 types de tricoumaroyl spermine. En revanche la spermidine n'est pas symétrique et possède 3 azotes, il existe donc 3 types de monocoumaroyl spermidine (Figure 1.4), 3 types de di coumaroyl spermidine et 1 type de tricoumaroyl spermidine.

Lors de ces expériences une des deux monocoumaroyl spermine a été détectée (mon spm1), deux des trois formes de monocoumaroyl spermidine (mon spd1 et 2) ont été détectées, et une dicoumaroyl spermine (di spm), une dicoumaroyl spermidine (di spd), une tricoumaroyl spermine (tri spm) et une tricoumaroyl spermidine (tri spd) ont

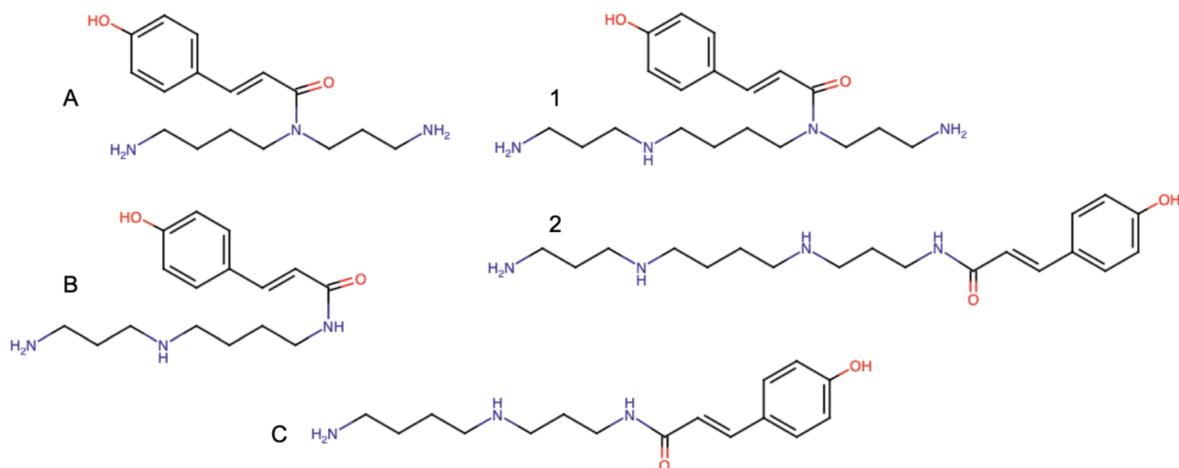


Figure 1.4 : représentation des différentes types monocoumaroyl spermidines (A ; B ; C) et monocoumaroyl speramines (1 ; 2).

été détectées. En revanche ces molécules n'ont pas encore été caractérisées, il n'est donc pas possible de savoir sur quels azotes les acylations ont été réalisées.

Les chevelus racinaires cultivés en conditions normales produisent peu de dérivés de spermidine et de spermine. Les dérivés majoritaires retrouvés sont la monocoumaroyl spermidine et la dicoumaroyl spermine, les dérivés tri-substitués étant très peu produits. La disponibilité des polyamines au sein des chevelus racinaires est sûrement la cause de cette très faible production. Après supplémentation avec de la spermine, les chevelus racinaires produisent environ 4 fois plus de dicoumaroyl spermine et un peu plus de tricoumaroyl spermine que les chevelus racinaires en conditions normales (sans supplémentation). Le contenu en dérivés de spermidine quant à lui ne change pas par rapport aux chevelus racinaires en conditions normales. Les chevelus racinaires sont donc capables de puiser la spermine dans le milieu, puis cette spermine est utilisée par la SHT2. Le pool de spermidine semble être consommé de la même façon que dans les conditions normales.

Lorsqu'on cultive les chevelus racinaires en présence de spermidine nous avons constaté que le contenu en dérivés de spermine et de spermidine ne changeait pas par rapport aux chevelus racinaires cultivés en conditions normales. Il est possible que les chevelus racinaires n'aient pas la capacité de puiser la spermidine dans le milieu, ou que la disponibilité du coumaroyl-CoA soit un facteur limitant.

En effet, les chevelus racinaires de chicorée produisent de grandes quantités d'acides chlorogénique (CQA) et isochlorogénique (diCQA) (Figure 1.5), dont la production nécessite l'utilisation de coumaroyl-CoA. La CiSHT2 surexprimée au sein des chevelus racinaires entre donc directement en compétition avec les HCTs/HQTs pour

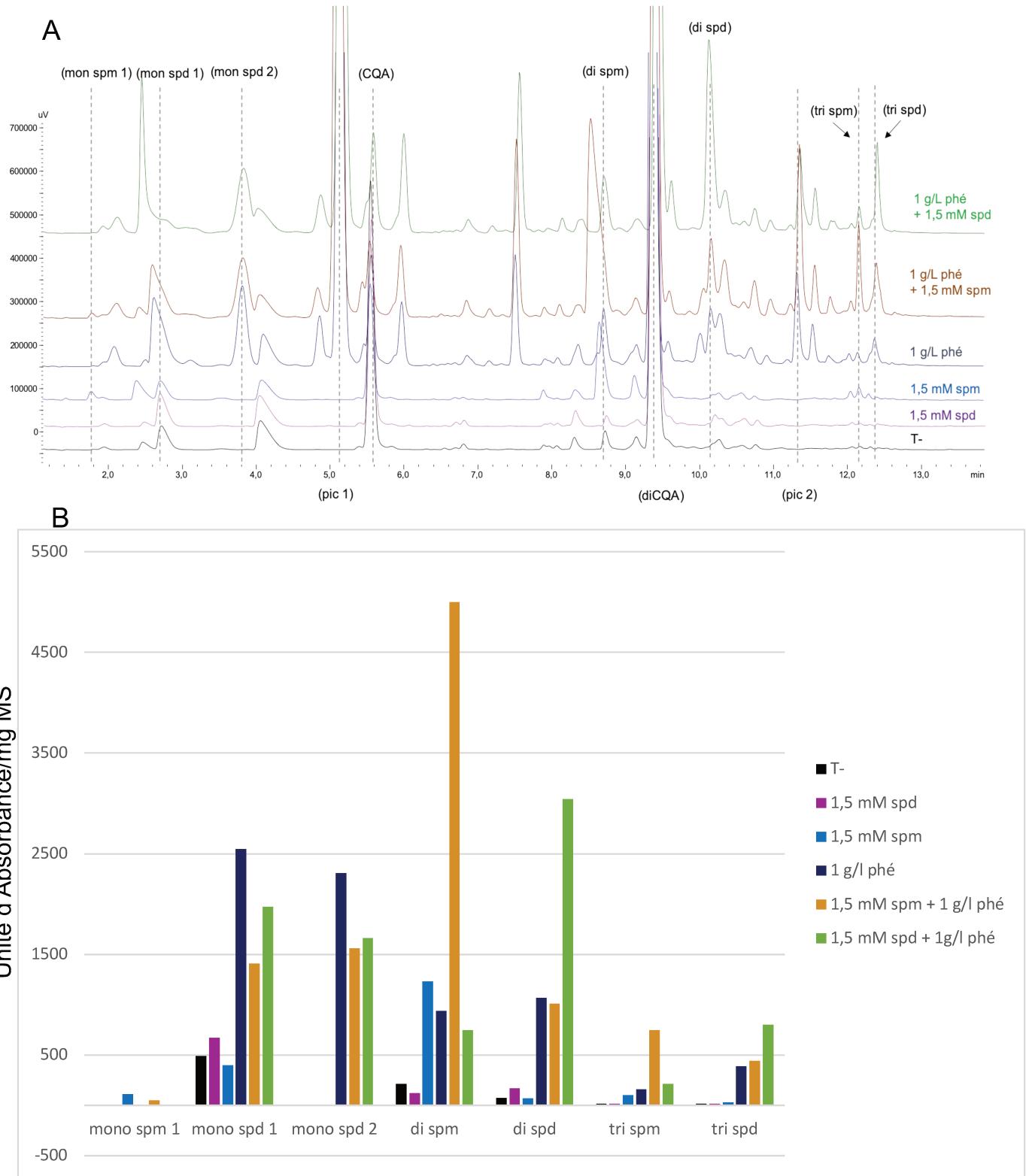


Figure 1.5 : (A) Analyses HPLC des chevelus racinaires sur-exprimant SHT2 après différentes conditions de cultures. (B) Evaluation des quantités de dérivés de spermine et de spermidine accumulés dans les chevelus racinaires de chicorée dans différentes conditions de culture. Les supplémentations ont été réalisées après 12 jours de croissance et la récolte 6 jours plus tard. En noir : sans supplémentation ; en violet : supplémentation avec 1,5 mM de spermidine ; en bleu : supplémentation avec 1,5 mM de spermine. En bleu foncé : supplémentation avec 1 g/l de phénylalanine ; en orange : supplémentation avec 1,5 mM de spermine et 1 g/l de phénylalanine ; en vert : supplémentation avec 1,5 mM de spermidine et 1 g/l de phénylalanine. Mono spm : monocoumaroyl spermine ; mono spd : monocoumaroyl spermidine ; di spm : dicoumaroyl spermine ; di spd : dicoumaroyl spermidine ; tri spm : tricoumaroyl spermine ; tri spd : tricoumaroyl spermidine.

l'utilisation du coumaroyl-CoA lors de la production des phénolamides dérivant de la spermine et de la spermidine. La SHT2 étant un peu plus active avec la spermine que la spermidine *in vitro*, peut expliquer le fait de l'augmentation de dicoumaroyl spermine lors de la culture avec de la spermine alors que la production de dicoumaroyl spermidine ne change pas lors de la culture avec de la spermidine. En effet cette forte activité permettrait à SHT2 d'être plus compétitive avec les HCTs/HQTs pour l'acquisition du coumaroyl-CoA.

Après supplémentation du milieu avec 1 g/L de phénylalanine, nous avons constaté une nette augmentation de la production de dérivés de spermidine par rapport au chevelus racinaires cultivés en conditions normales. En ce qui concerne la production de dérivés de spermine, elle est aussi augmentée, mais le contenu en dérivés de spermidine semble être supérieur. Un pool de spermidine accessible supérieur au pool de spermine au sein des chevelus racinaires peut expliquer cette différence.

Afin que les substrats soient le moins limitants possible nous avons donc supplémenté les chevelus racinaires avec spermine et phénylalanine ou spermidine et phénylalanine. Dans les deux cas les dérivés de spermine et de spermidine sont « boostés ». La production en dicoumaroyl spermine est augmentée d'un facteur proche de 20 par rapport aux conditions normales, de même pour la production en dicoumaroyl spermidine.

Lors de l'ajout de phénylalanine dans le milieu de culture, que ce soit seul ou avec spermine ou spermidine, nous avons constaté que de nouveaux composés (noté pic 1 et pic 2 sur la Figure 1) étaient produits en grande quantité. Ces composés ont été analysés en spectrométrie de masse et la masse de 349.0895 a été détectée pour le pic 1 et 364.2219 pour le pic 2. En enlevant la masse de 23 qui correspond à un adduit sodium on trouve les masses de 326.0895 et 341,2219 qui pourraient correspondre au coumaroyl-glucose et au cafféoyl-glucose. L'ajout d'une grande quantité de phénylalanine doit très fortement accroître la production d'acide coumarique et d'acide caféïque. Dans le but de retrouver une homéostasie normale les chevelus racinaires doivent donc glycosyler ces deux acides hydroxycinnamiques. En effet un des mécanismes de détoxicification chez les plantes pour réguler l'homéostasie cellulaire est la glycosylation (Lim et al. 2003).

La disponibilité des substrats, que ce soit le coumaroyl-CoA ou les polyamines, est donc un facteur très important pour la production des dérivés de spermine ou de spermidine. Au vu de ces résultats, nous pouvons penser que le ratio

spermine/spermidine présent au niveau des boutons inflorescentiels est un paramètre important pour l'orientation de la production de phénolamides chez les plantes au cours de l'évolution. Chez *Arabidopsis*, la quantité de spermidine au sein des anthères est environ dix fois supérieure à celle de la spermine (Fellenberg et al. 2012). Cela peut expliquer la spécialisation d'AtSHT dans la production de dérivés de spermidine totalement substitués. Chez la chicorée, et peut être chez la grande majorité des Astéracées, ce ratio pourrait être inversé, ce qui aurait engendré la néofonctionnalisation d'une SHT pouvant produire des dérivés de spermine totalement substitués.

Au cours de ces expériences nous avons largement augmenté la production de phénolamides par des cultures de chevelus racinaires en leur ajoutant les substrats adaptés, notamment nous avons montré l'implication de CiSHT2 dans la synthèse de phénolamides, cependant la présence de tétracoumaroyl spermine n'a jamais été détectée et le composé majoritaire produit est la dicoumaroyl spermine. Il semblerait que la présence d'une deuxième enzyme soit nécessaire pour produire des dérivés de spermine totalement substitués (tétracoumaroyl spermine) à partir des dérivés mono, di- ou tri-substitués. Bien que la surexpression de CiSHT1 dans les chevelus racinaires, n'ait pas permis la synthèse de phénolamides, la SHT1 découverte chez la chicorée pourrait quand même être une bonne candidate, en intervenant en fin de synthèse après l'action de la SHT2. Afin de déterminer si l'enzyme SHT1 pourrait être impliquée dans la synthèse des dérivés de spermine totalement substitués (terracoumaroyl spermine) nous avons donc décidé de mettre en place la technologie d'édition ciblée du génome, CRISPR/Cas 9 chez la chicorée, afin de créer des mutants SHT1 et SHT2, et de mieux caractériser le rôle de chacun de ces gènes.

Chapitre II : Mise en place de la technologie CRISPR/Cas 9 chez la chicorée

Avant-propos

Comprendre la fonction des gènes ainsi que leur régulation pour utiliser ces connaissances dans le but d'améliorer les plantes cultivées est une des principales missions pour les spécialistes en biologie végétale. Grâce à la remarquable évolution des techniques de séquençage ainsi que de celles du stockage des données, la bio-informatique est devenue une étape obligatoire pour l'étude fonctionnelle des gènes et l'ingénierie génétique chez de nombreuses plantes. Elle permet d'induire des mutations ciblées de gènes afin de comprendre leur fonction. L'objectif de la mutagenèse ciblée est de modifier la séquence nucléotidique d'un gène, de créer donc une mutation, afin que la séquence protéique issue du gène muté soit modifiée et qu'un nouveau phénotype soit observé. Ce nouveau phénotype pourra donc être attribué à l'inhibition du gène ciblé, ce qui permettra de potentiellement impliquer le gène dans la mise en place de ce phénotype.

Pendant de nombreuses années, des plantes mutantes ont été générées par mutagenèse aléatoire utilisant des agents mutagènes, des radiations ou l'insertion aléatoire d'un T-DNA médiée par *Agrobacterium*. Au cours de ces dernières années, trois techniques d'édition ciblée du génome ont été développées, à savoir la technique des nucléases à doigts de zinc (ZFNs) (Kim, Cha, and Chandrasegaran 1996), les nucléases effectrices de type activateur de transcription (TALENs) (Christian et al. 2010) et CRISPR (clustered regularly interspaced short palindromic repeats) associé avec la protéine Cas9 (Deltcheva et al. 2011). Ces trois méthodes induisent une coupure double brin sur l'ADN génomique cible. Les techniques ZFNs et TALENs ciblent l'ADN génomique grâce à des interactions protéines/ADN alors que le système CRISPR/Cas9 est basé sur de courtes hybridations ARN/ADN. Quoi qu'il en soit l'utilisation d'une technologie ciblée d'édition d'un gène nécessite la connaissance de la séquence du gène ainsi que sa structure et donc l'accessibilité à des données génomiques. Le système CRISPR/Cas9 étant beaucoup plus simple, rapide et efficace, il a été très utilisé ces 5 dernières années pour l'édition du génome des cellules eucaryotes.

Dans la nature, le système CRISPR/Cas9 est un système d'immunité adaptative contre des ADN étrangers, mis en place par les procaryotes. Il est utilisé par de nombreuses bactéries et par beaucoup d'archées. Le système a été classé en trois

types. Les types I et III possèdent de nombreuses protéines Cas qui forment un complexe pour la dégradation d'ADN et d'ARN étrangers. Le type II, quant à lui, dirige le clivage de l'ADN étranger en utilisant une unique protéine Cas, nommée Cas9. Cette particularité fait de ce système, un système de choix pour la mutation ciblée des génomes.

Le système CRISPR/Cas9 utilise un ARN CRISPR (crRNA) qui possède une séquence de 20 nucléotides complémentaire de l'ADN étranger ainsi qu'une partie complémentaire d'un petit ARN trans-activateur (tracrRNA). Le tracrRNA possède en plus de la séquence complémentaire du crRNA, une séquence capable d'interagir avec la protéine CAS9, mais ce n'est qu'une fois hybride avec le crRNA qu'il a la capacité de former un complexe fonctionnel avec la protéine CAS9. Quand le complexe fonctionnel reconnaît par liaisons ARN/ADN, la séquence complémentaire de 20 nucléotides sur l'ADN étranger en amont d'une séquence PAM (protospacer-adjacent motif) 5'-NGG-3', alors la Cas9 coupe le double brin d'ADN étranger. Cette coupure a lieu 3 nucléotides en aval de la séquence PAM. Cette coupure double brin est dans la plupart des cas réparée par la voie de réparation d'assemblage des extrémités non homologues (NHEJ) qui génère souvent l'apparition d'indels pouvant produire une mutation dans le cadre de lecture et donc une inhibition du gène. En revanche, si lors de cette coupure il y a présence d'une matrice d'ADN ayant des

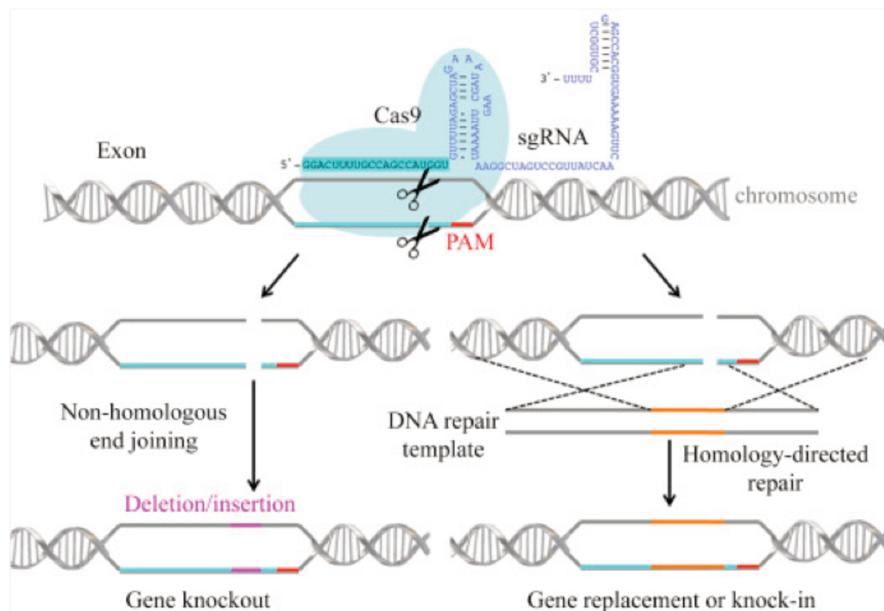


Figure 2.1 : Représentation schématique de l'édition d'un gène grâce à la technologie CRISPR/Cas 9. (D'après Zhang et al 2016)

régions homologues avec les extrémités limitrophes à la coupure, cela peut augmenter la probabilité de réparation *via* le mécanisme de recombinaisons homologues (HDR). De ce fait si cette matrice d'ADN contient une copie du gène modifié ou un gène étranger, il sera donc possible de l'insérer au niveau du site de coupure double brin généré par la technologie CRISPR/Cas9 (Figure 2.1).

L'ingénierie moléculaire a modifié ce système en remplaçant le complexe crRNA/tracrRNA par un simple sgRNA (single guide RNA). De ce fait, seuls le sgRNA et la protéine Cas9 sont nécessaires pour une mutation ciblée du génome. L'édition ciblée du génome nécessite donc d'exprimer un sgRNA complémentaire à 20 nucleotides de l'ADN cible ainsi que la protéine Cas9, tout cela dans une unique cellule indifférenciée. Pour l'édition ciblée chez les végétaux, de nombreux vecteurs ont été mis en place, ces vecteurs possèdent des promoteurs de l'ARN polymérase III dans le but de transcrire le sgRNA, ainsi que des promoteurs comme par exemple le CaMV 35S ou le promoteur Ubiquitine, pour la synthèse de la Cas9. Depuis l'apparition de la technologie CRISPR/Cas9, l'édition ciblée du génome de nombreuses espèces comme *Arabidopsis thaliana*, *Nicotiana benthamiana*, le riz, le tabac, le sorgho et de nombreuses autres espèces a été réalisée avec succès. En revanche, cette technologie n'a pas encore été utilisée chez la chicorée.

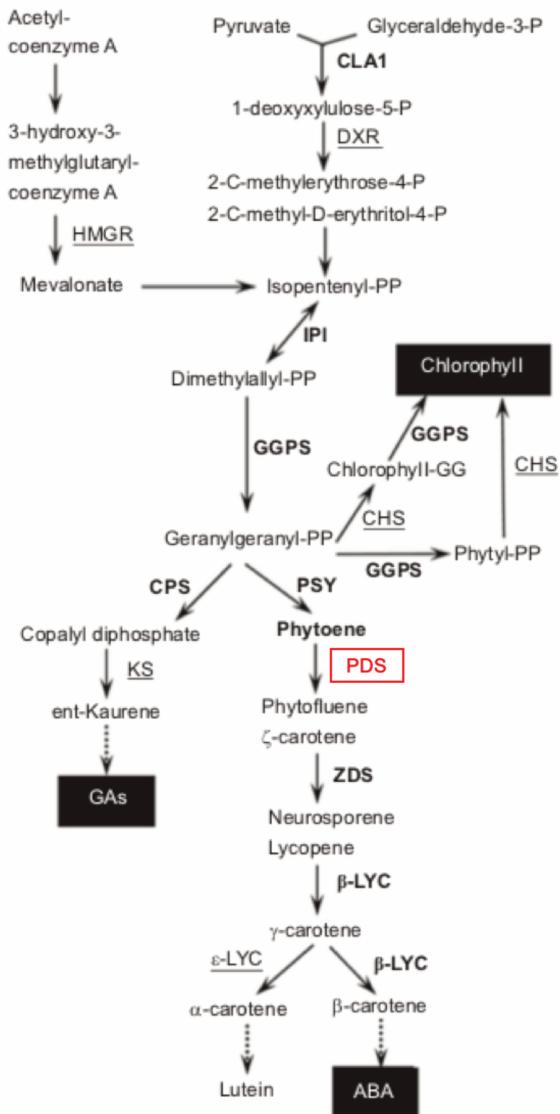


Figure 2.2: Représentation des voies de biosynthèse des caroténoïdes, de la chlorophylle et des gibberellines (GAs). La phytoène désaturase est encadrée en rouge. (D'après Qin et al. 2007).

La chicorée est une plante agronomique utilisée pour la production de substitut du café mais plus majoritairement pour la

production d'inuline qui est une fibre alimentaire digestible très utilisée en industrie agro-alimentaire. De plus la chicorée possède de nombreuses vertus thérapeutiques qui lui sont conférées par sa forte production de métabolites spécialisés. La mise au point de la technologie CRISPR/Cas9, chez la chicorée, pourrait donc permettre d'améliorer de façon très ciblée les caractères agronomiques de cette plante ainsi qu'êtrent utilisée en recherche fondamentale pour la compréhension des voies de biosynthèse qu'elle utilise. Grâce à son intérêt agronomique, la chicorée a vu son génome séquencé par un consortium privé, et notre laboratoire a un accès privilégié à cette séquence, ce qui est un atout majeur pour la mise en place de la technologie CRISPR/Cas9.

Dans le but de réaliser la validation fonctionnelle de cet outil chez la chicorée, nous avons décidé de cibler le gène de la *phytoène désaturase*, *PDS*, chez la chicorée. En effet cette enzyme intervient dans la voie de biosynthèse des caroténoïdes. La mutation de cette enzyme occasionne une augmentation de la quantité de phytoène dans les feuilles ce qui a pour conséquence l'inhibition des enzymes impliquées dans la biosynthèse de la chlorophylle (Figure 2.2) (Qin et al. 2007). La mutation de cette enzyme conduit donc indirectement à une plante sans chlorophylle, qui possèdera un

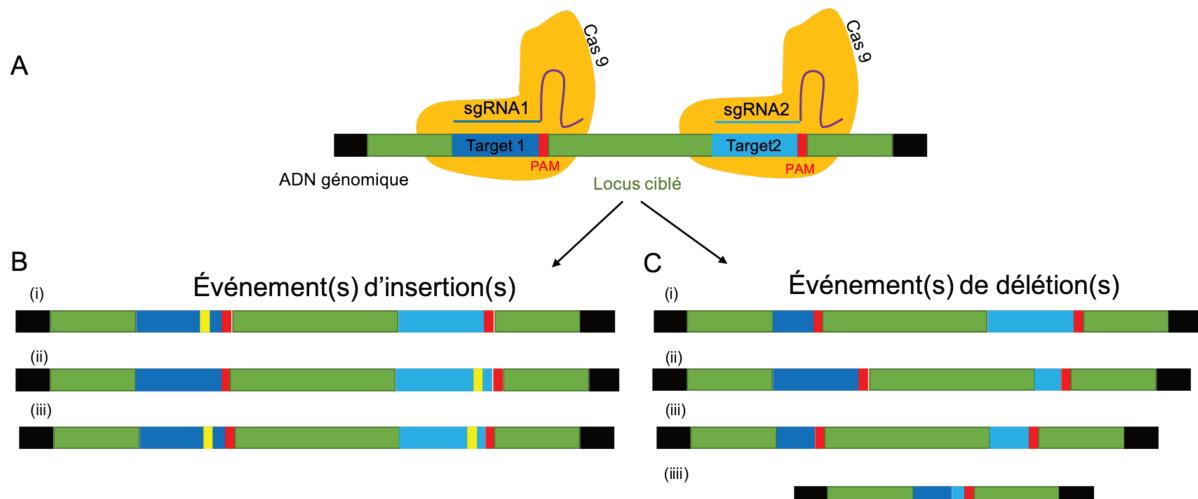


Figure 2.3 : Représentation schématique de la stratégie d'édition d'un gène en utilisant 2 sgRNAs et quelques événements de mutation possibles. (A) Disposition des sgRNAs sur le gène ciblé. (B) Événements d'insertion possibles au niveau des deux targets ; (i) insertion au niveau de la target 1, pas de mutation au niveau de la target 2, (ii) pas de mutation au niveau de la target 1, insertion au niveau de la target 2 ; (iii) insertion au niveau des deux targets. (C) Événements de délétion possibles au niveau des deux targets ; (i) délétion au niveau de la target 1, pas de mutation au niveau de la target 2 ; (ii) pas de délétion au niveau de la target 1, délétion au niveau de la target 2 ; (iii) délétion au niveau des deux targets ; (iv) délétion de l'inter-guide.

phénotype albinos, ce qui permettra de déceler facilement les plantes dont le gène a été édité.

Pour atteindre nos objectifs, nous avons décidé d'utiliser deux types de transformation : la transformation de protoplastes médiée par le PEG ainsi que la transformation médiée par *Agrobacterium rhizogenes*. La méthode de transformation par *Agrobacterium rhizogenes* parce que nous maîtrisons la technologie au laboratoire et parce que les chevelus racinaires de chicorée permettent facilement la régénération de plantes. Cela nous permettra de valider nos cibles et d'établir une preuve de concept, cependant le génome des plantes produites renfermera l'ADN-T d'*A. rhizogenes*. D'où l'idée d'utiliser la transfection de protoplastes qui permet une expression transitoire de l'ADN introduit dans la cellule et pour laquelle, la culture de protoplastes est aussi maîtrisée au laboratoire. De plus, dans le but d'augmenter nos chances d'édition, nous avons décidé d'utiliser une stratégie utilisant 2 sgRNAs pour éditer le gène d'intérêt (Figure 2.3). Ce travail a été publié dans la revue *International Journal of Molecular Sciences* en 2019 et constitue ce second chapitre.

Efficient Genome Editing Using CRISPR/cas9 Technology in chicory

Guillaume Bernard, David Gagneul, Harmony Alves Dos Santos, Audrey Etienne, Jean-Louis Hilbert and Caroline Rambaud *

EA 7394, 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

Correspondence: caroline.rambaud@univ-lille.fr

Keywords: CRISPR/Cas9; multiplex genome editing; *Cichorium intybus*; phytoene desaturase; *Agrobacterium rhizogenes*-mediated transformation; protoplast transformation

Abstract

CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated with protein CAS9) is a genome-editing tool that has been extensively used in the last five years because of its novelty, affordability, and feasibility. This technology has been developed in many plant species for gene function analysis and crop improvement but has never been used in chicory (*Cichorium intybus L.*). In this study, we successfully applied CRISPR/Cas9-mediated targeted mutagenesis to chicory using *Agrobacterium rhizogenes*-mediated transformation and protoplast transfection methods. A *U6* promoter (*CiU6-1p*) among eight predicted *U6* promoters in chicory was selected to drive sgRNA expression. A binary vector designed to induce targeted mutations in the fifth exon of the chicory *phytoene desaturase* gene (*CiPDS*) was then constructed and used to transform chicory. The mutation frequency was 4.5% with the protoplast transient expression system and 31.25% with *A. rhizogenes*-mediated stable transformation. Biallelic mutations were detected in all the mutant plants. The use of *A. rhizogenes*-mediated transformation seems preferable as the regeneration of plants is faster and the mutation frequency was shown to be higher. With both transformation methods, foreign DNA was integrated in the plant genome. Hence, selection of vector (transgene)-free segregants is required. Our results showed that genome editing with CRISPR/Cas9 system can be efficiently used with chicory, which should facilitate and accelerate genetic improvement and functional biology.

1. Introduction

Genome editing, which consists of targeting and digesting DNA at a specific site in the genome, is an important tool for gene function analysis and crop improvement [1]. To date, three specific genome-editing technologies have been developed: zinc finger nucleases (ZFNs) [2], transcription activator-like effector nucleases (TALENs) [3], and clustered regularly interspaced short palindromic repeat (CRISPR) associated protein (CAS) system. These methods induce double strand-breaks (DSBs) in the targeted DNA. In eukaryotic cells, these breaks can be repaired in two different pathways: the non-homologous end-joining (NHEJ) and homologous recombination (HR). NHEJ is the most commonly used DSB repair mechanism in many organisms and can cause insertions or deletions that can potentially produce a gene knockout [4].

In natural conditions, the CRISPR/Cas9 system uses a CRISPR RNA (crRNA) and a small trans-activating CRISPR RNA (trancrRNA), which can hybridize to form the mature dual crRNA. The CAS9 endonuclease forms a ribonucleoprotein complex with the crRNA, which guides the endonuclease to a specific target DNA. Cleavage is only possible if the complex recognizes a short seed sequence upstream of a protospacer-adjacent motif (PAM) 5'-NGG-3' [5]. Thanks to molecular biology engineering, crRNA has been replaced by a chimeric single guide RNA (sgRNA). Consequently, only sgRNA and endonuclease CAS9 are needed to cut and create DSBs at specific loci in the genome of the cells. Because of its simplicity, low cost, versatility, and high efficiency, the CRISPR/Cas9 system has become the most widely used technology for genome editing in many organisms such as bacteria, yeasts, animals, and plants.

In plants, CRISPR/Cas9 system is often inserted in a vector containing the CAS9 upstream of a strong ubiquitous promoter such as ubiquitin or Cauliflower Mosaic Virus (CaMV) 35S promoters. Moreover, germ-line specific promoters, such as *CDC45*, which have the advantage of being active in the meiotic/zygotic phase, have recently been successfully used [6]. To deliver the CAS9 protein to the nuclear genome, a nuclear localization signal (NLS) sequence is fused to the CAS9 gene. To express the sgRNA, RNA polymerase-III promoters such as *U6* and *U3* are used. The transcript initiation nucleotide is G and A in the *U6* promoter and in the *U3* promoter, respectively [7]. Using Golden Gate cloning technology, which uses second-generation restriction

enzymes [8], it is possible to easily insert numerous sgRNA and the CAS9 gene in a single expression vector, in order to simultaneously target several DNA sites [9].

To introduce vectors into plants, two main transformation methods are commonly used in plants: protoplast transformation, where plasmids can be delivered into protoplasts using various methods, such as polyethylene glycol (PEG) fusion, electroporation and microinjection; and leaf tissue transformation using *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* (*Rhizobium rhizogenes*). Both methods have been successfully applied to express CRISPR/Cas9 system in plants for genome editing.

Protoplasts have been used in *Arabidopsis* and tobacco [10] but also in crop species such as maize [11], rice, wheat [12], soybean [13] and tomato [14]. However, the regeneration of plants from mutated protoplasts is uncommon because many plant species cannot be regenerated from protoplast cultures, particularly the monocots. In addition, the regeneration process is long and subject to contamination. This approach is therefore often used as a preliminary step in evaluating the functionality of the CRISPR/Cas9 system before using it with a more appropriate transformation method such as *A. tumefaciens*-mediated transformation [15].

However, not all plant species can be transformed by these methods. In this case, other techniques may be used such as biolistic or *A. rhizogenes*-mediated transformation. *A. rhizogenes* has been widely used for plant transformation [16] to study rhizosphere or metabolic pathways. Genome editing using *A. rhizogenes*-mediated transformation has been successfully carried out in tomato (*Solanum spp*) [17], soybean (*Glycine max*) [13], *Brassica carinata* [18], *Salvia miltiorrhiza* [19]. However only a few plants have the ability to regenerate from hairy root lines [16], which hampers its use in genome editing.

Chicory (*Cichorium intybus L.*), an Asteraceae, is well known in traditional medicine. More than 100 individual compounds have been isolated from this plant, especially from its roots which are rich in specialized metabolites (also known as secondary metabolites or plant natural products) such as polyphenols and terpenoids [20–22]. Due to their ability to produce valuable compounds, hairy roots and cell cultures have been set up by biotechnologists [23]. It has been reported that specialized metabolites present in chicory exert many biological activities such as antioxidant, anti-cancer, anti-inflammatory or anti-hepatotoxic activities with beneficial effects on health in humans and livestock [24]. Chicory is also used as a vegetable crop, for leaf salad called chicon or witloof or as a coffee substitute. In addition, chicory roots have an inulin content

which can be up to 40%. Inulin is a polysaccharide used as a sugar substitute due to its low caloric value [25]. Chicory is used for multiple applications. This has led to large scale cultivation and therefore to an interest in genetically modified chicory to improve yields and nutritional properties and thus extend its potential [26]. Targeted mutagenesis using CRISPR/Cas9 technology could potentially be used in chicory as an efficient tool for studying gene function and generating new varieties.

The success of targeted mutagenesis with CRISPR/Cas9 method depends on the efficiency of the different molecular components (sgRNA design, CAS9, promoters...) but mainly on the ability of the species to be transformed and to regenerate plants. *A. tumefaciens*-mediated transformation of chicory is feasible but genotype dependent [27], whereas its transformation using *A. rhizogenes* is quite efficient and the regeneration of plants from hairy roots is possible [28,29]. This may provide a fast approach for chicory genome editing. This strategy has already been used in *Taraxacum kok-saghyz* [30]. Otherwise PEG-mediated protoplast fusion has already been applied to induce male sterility in chicory and fused protoplasts are able to regenerate [31]. The use of protoplasts for genome editing could be an alternative in the case of chicory.

In this study, we performed an efficient multiplex genome editing in chicory in the first generation (T0), using two sgRNAs and CAS9. Both *A. rhizogenes*-mediated stable transformation and protoplast transient expression system were successfully applied to engineer chicory genome. As a proof of principle, we succeeded in knocking out the *CiPDS* gene (*phytoene desaturase*) in plants regenerated from hairy roots and from protoplasts. With *CiPDS* specific sgRNAs, several independent albino lines were generated. An efficient method that can be used for chicory mutagenesis is thus provided.

2. Results

2.1. Prediction of U6 Promoter Sequences in Chicory

The U6 RNA polymerase-III promoters are typically used to drive the sgRNA expression in Dicots [7]. The *Arabidopsis* U6-26 promoter has been used to generate sgRNA in numerous plant species, such as in soybean [13], cotton [32], or *Salvia miltiorrhiza* [19]. To identify the U6 promoters in chicory, the *Arabidopsis* U6-26 small nuclear RNA (snRNA) sequence was used to search the chicory genome (unpublished database from Florimond Desprez SA). Eight U6 sequences (named *U6-1* to *U6-8*) with the highest similarity to the *Arabidopsis* sequence were selected for further analysis (Figure 1).

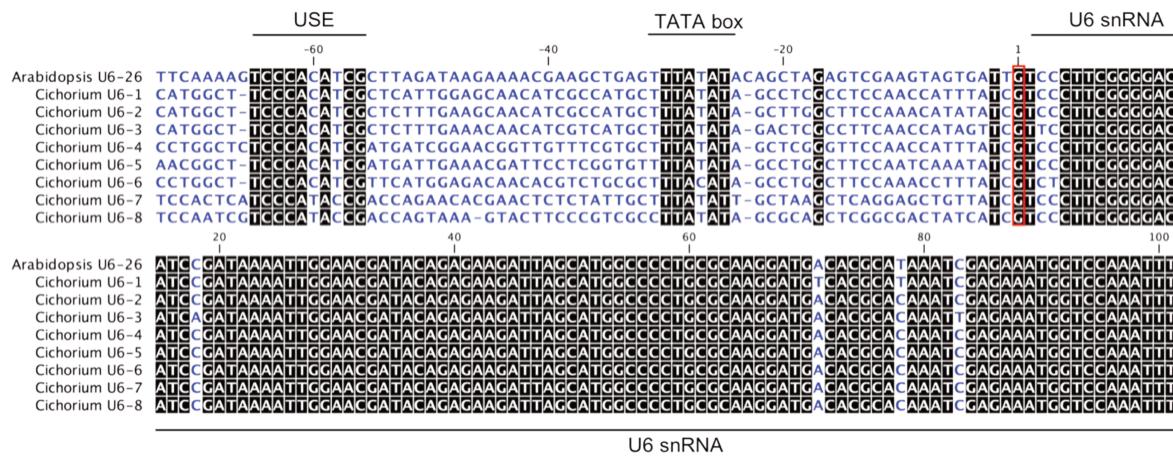


Figure 1: Chicory U6 and *Arabidopsis* U6-26 gene sequence partial alignment. The highly conserved parts of the sequences are in black while the Upstream Sequence Elements (USE), the TATA-boxes and the U6 small nuclear (snRNA) genes are marked with black bars. The transcription start sites are indicated (red frame).

Polymerase-III promoters express transcripts with purine, i.e., a G, as the initiation nucleotide. As expected, a G was located just in front of all chicory snRNA sequences. All snRNA, including that of *Arabidopsis*, were highly similar. The U6 promoter sequences contain two conserved elements essential for the transcription initiation by the RNA polymerase-III [33]: the upstream sequence element (USE; consensus sequence RTCCCCACATCG) and a TATA-like box (consensus sequence TTTATATA). The USE sequence of *U6-7* (GenBank accession MK455778) and *U6-8* (GenBank accession MK455779) and the TATA-like box of *U6-6* (GenBank accession MK455777) are slightly different from the others. *U6-1* (GenBank accession MK455772), *U6-2* (GenBank accession MK455773), *U6-3* (GenBank accession MK455774), *U6-4* (GenBank accession MK455775) and *U6-5* (GenBank accession MK455776) could be good candidates for driving the expression of sgRNAs in chicory.

U6-1 promoter (*CiU6-1p*) was randomly chosen and used for the following experiments.

2.2. Target Selection and Plasmid Construction for the CRISPR/Cas9 System

A proof-of-principle experiment showed that *CiPDS* gene (*phytoene desaturase*) was successfully knocked out in plants regenerated from hairy roots and from protoplasts. PDS is a key enzyme for chlorophyll biosynthesis and *PDS* mutants are easily detectable due to their albino phenotype [34]. To identify putative *PDS* genes, a tBLASTn search of the chicory genome (unpublished database of Florimond Desprez SA) was conducted with *Helianthus annuus* (Asteraceae) PDS protein sequence (GenBank accession AHA36971.1) as a query. The gene with the highest similarity was selected. Only one copy of the gene was predicted. *CiPDS* (GenBank accession MK455771) is a 2514 pb in length gene and has 10 exons, with 1746 nucleotides of transcript sequence coding 581 amino acids. To knock out the gene, we designed 2 target sequences (double mutations) in the fifth exon (Figure 2A; Supplementary Figure S1). Using crispor.tefor.net software [35], 2 guide RNAs were selected according to the position of the target sequence and to their GC content. The CAS9 nuclease was expected to cut twice in exon 5 upstream of the PAM.

Our goal was to produce a binary vector harboring 2 sgRNAs and the CAS9 for gene editing in chicory. The first step was to construct a plasmid with the *CiU6-1p* and the sgRNA scaffold, in which a guide RNA (gRNA) can be easily inserted. For this, we designed a sub-cloning vector containing the *CiU6-1p*, followed by a counter selection marker (*lacZ-ccdB*, *ccdB* under the control of the Isopropyl β-D-1-Thiogalactopyranoside (IPTG) inducible LacZ promoter) flanked by 2 *BbsI* recognition sites and the sgRNA scaffold (Figure 2C). This sequence was synthesized and inserted in pUCIDT-Kan to give pKanCiU6-1p-sgRNA. *BbsI* is a second-generation enzyme that produces distinct, non-palindromic sticky ends outside of its recognition site (Figure 2B). This feature is convenient in this case. Indeed, a G must be located directly

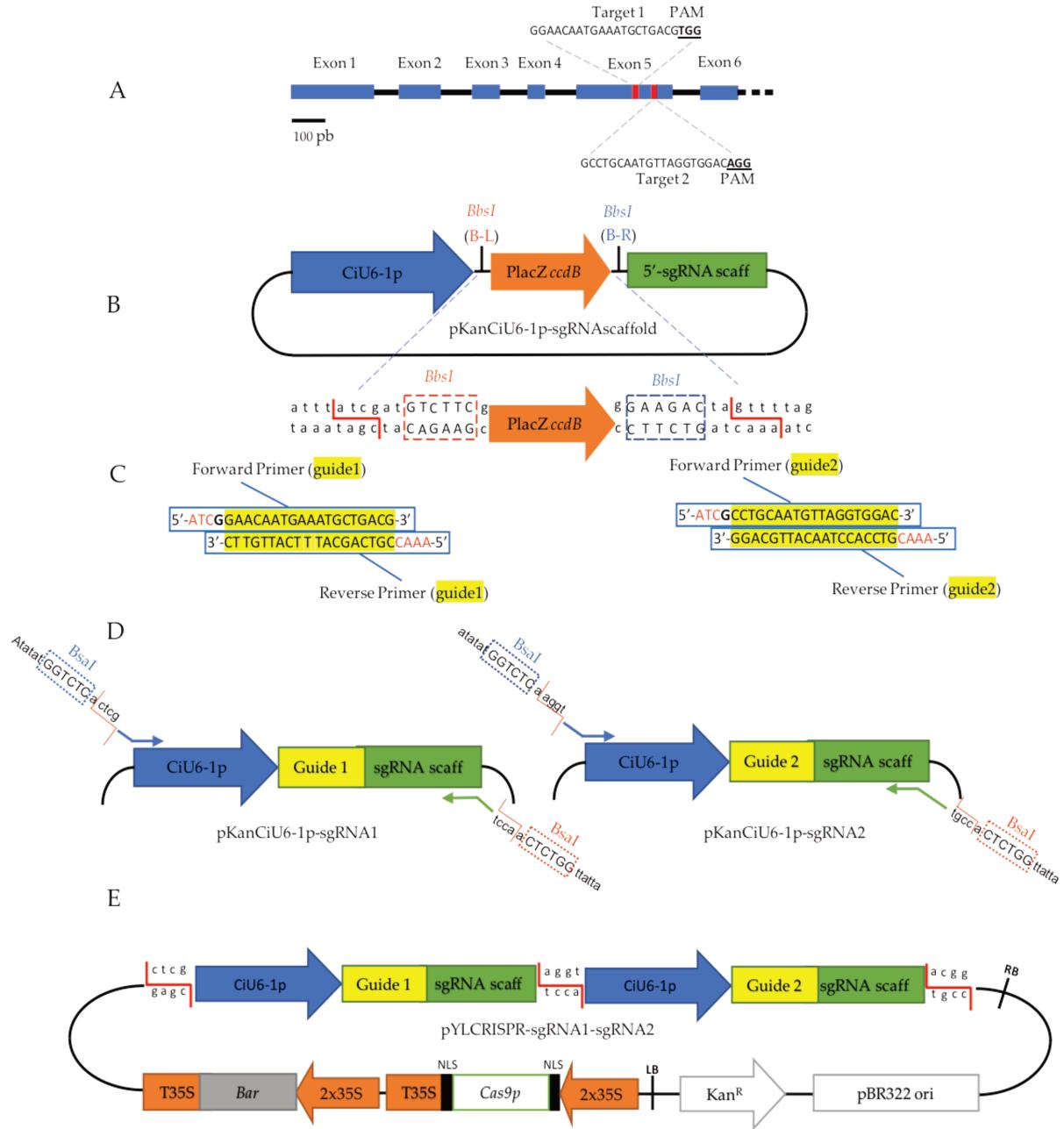


Figure 2: Overview of the experimental design for CiPDS disruption. (A) Schematic position of the two guide RNAs (red boxes) targeting the fifth exon of *CiPDS*. The blue boxes indicate exons; the black lines indicate introns. (B) Schematic view of pKanCiU6-1p-sgRNAscaffold. Between the *CiU6-1p* and the sgRNA scaffold (scaff), a *ccdB* gene driven by the LacZ promoter (counter selection marker) is surrounded by *BbsI* (second-generation enzyme) recognition sites, which allows the ligation of the hybridized guide adaptor shown in C. Red B-L depicts the orientation of the *BbsI* recognition site which achieves the cleavage on the left side. Blue B-R depicts the orientation of the *BbsI* recognition site which achieves cleavage on the right side. (C) Sequences of the complementary guide adaptors with the 19 pb guide sequence (yellow), the transcription initiator nucleotide G (bold), and the binding sites necessary to insert the hybridized guide adaptor into the pKanCiU6-1p-sgRNAscaffold (red). (D) Representation of the vector resulting from the insertion of the hybridized guide adaptors in the pKanCiU6-1p-sgRNAscaffold, (note that one vector is constructed for each guide). The primers used for preparation of CiU6-1p-guide-sgRNA cassettes for Golden Gate Cloning are also shown. (E) Representation of the final plasmid pYLCRISPR-sgRNA1-sgRNA2 resulting from the cloning of the two cassettes into the pYLCRISPR/Cas9P35S-B [36]. RB = Right Border, LB = Left Border, NLS = Nuclear Localization Signal.

in front of the sgRNA scaffold. The 19 pb of the gRNAs with specific adaptors were synthetized as two complementary pairs of oligonucleotides (Table A1) as described by Ma and Liu [36] (Figure 2C) and inserted into pKanCiU6-1p-sgRNA by digestion/ligation to produce the plasmid pKanCiU6-sgRNA. For each guide, one plasmid was generated (pKanCiU6-1p-sgRNA1 and pKanCiU6-1p-sgRNA2; Figure 2D). The complete cassettes were amplified with specific primer pairs (GG1-F, GG1-R, and GG2-F, GG2-R) designed for Golden Gate cloning [8,37] and inserted into the pYLCRISPR/Cas9P35S-B expression vector [36]. Using this protocol, vector pYLCRISPR-sgRNA1-sgRNA2 (Figure 2E) engineered to co-express two sgRNAs and the CAS9 nuclease in chicory was constructed within a week.

2.3. Regeneration of CRISPR-Edited plants from protoplasts and from Hairy Roots lines

Two transformation methods were used to express the CRISPR/Cas9 system in chicory: protoplast transformation mediated with PEG and *A. rhizogenes*-mediated stable transformation. The protoplasts were transformed with the pYLCRISPR-sgRNA1-sgRNA2, transferred to growth media to obtain calli, and then plant regeneration was induced on a regeneration medium. As control, plants regenerated from protoplasts following the same protocols but without plasmid addition were obtained. Albino and green plants were obtained 5 months after transformation (Figure 3A,B). From 198 calli resulting from one transformation experiment, 9 calli developed albino shoots (4.5%, see below for genetic confirmation) (Table 1). These data show that it is possible to obtain mutated chicory plants from protoplasts transformed by PEG-mediated method.

Table 1: Mutation analysis in the *CiPDS* gene. Analysis were performed on plants regenerated from protoplasts or hairy roots. ¹ HM : homozygous. ² HZ : heterozygous

Transformation type	N° of lines	N° of albinos lines	Mutation frequency (%)	Target 1					Target 2					
				Bi allelic mutation			Mono allelic mutation	No mutation	Mutation efficiency	Bi allelic mutation			Mono allelic mutation	No mutation
				HM ¹	HZ ²	%				HM ¹	HZ ²	%		
Protoplast	198	9	4.5	2	3	55	1	3	66.6%	5	4	100	0	0
Hairy root	32	10	31.25	1	5	60	3	1	90%	1	7	80	1	1

Regarding the stable transformation with *A. rhizogenes* and to show that transformed hairy roots were able to regenerate plants with edited genomes, chicory was infected with *A. rhizogenes* strain 15834 transformed with the pYLCRISPR-sgRNA1-sgRNA2 binary plasmid or not (used as a control). Two weeks after *A.*

rhizogenes transformation, 32 hairy root lines were generated from young leaves wounded with bacteria harboring the vector pYLCRISPR-sgRNA1-sgRNA2. The lines were cultivated on MS/2 medium with 300 mg/L ampicillin to remove the agrobacteria. From the 32 selected lines, 10 albino and 22 green lines (Figure 3C, D) were generated. These results were further confirmed at genetic level (see below). Non-albino lines grew more quickly than albino lines. From the 10 albino lines, 4 lines spontaneously regenerated albino plants. The others have regenerated albino plants after cultivation on bud induction medium MC4 typically used for chicory regeneration [38].

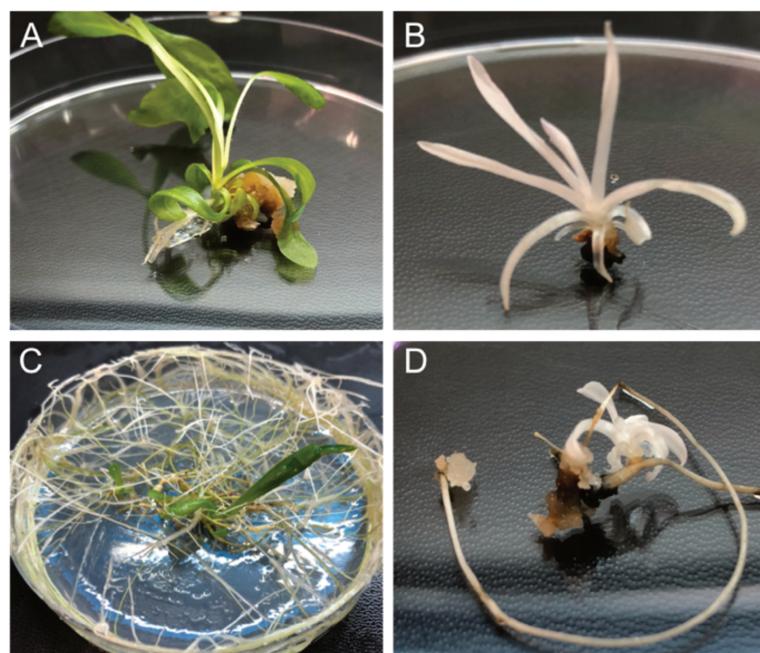


Figure 3: Phenotype of genome edited plants. (A) Wild-type (WT) shoot emerging from a WT callus. (B) CiPDS edited albino shoots emerging from a callus. (C) Hairy root line transformed with wild-type *A. rhizogenes* strain 15834 with emerging shoot. (D) Albino shoot emerging from hairy root line engineered to knock out CiPDS.

2.4. Identification of mutation events in genome edited plants

To confirm the genome editing at the molecular level and to characterize the mutations, we extracted genomic DNA from all the different albino materials (calli, hairy roots, regenerated plants). The DNA from the hairy root lines was extracted before the regeneration of the first albino shoots. Chicory is a diploid species and genes have one

copy in each homologous chromosome. Hence, the CRISPR/Cas9 system can induce two types of mutations in the targeted gene. On the one hand, only one of the alleles can be mutated, leading to a so-called monoallelic mutation. This type of mutation does not give a full knock out of the targeted gene because the other allele remains wild-type. On the other hand, the two alleles can be mutated which produces a biallelic mutation. In this case, if the two alleles share the same mutation, the biallelic mutation is homozygous, but if mutations are different, biallelic mutation is heterozygous.

	Target 1	PAM	Target 2	PAM	T1/T2	AA			
WT	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	WT/WT	581			
L1.1	GAGGAACAATGAAA-----	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-5/WT	232	
L1.2	GAGGAACAATGAAATGCT -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-1/WT	236	
L2.1	GAGGAACAATGAAATGC -	GAC	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-1/-	236	
L2.2	GAGGAACAATGAAATGCT -	CGT	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-2/+1	233	
L3.1	GAGGAACAATGAAATGCT -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-1/-19	236	
L3.2	GAGGAACAATGAAATGCT -	-	-	-	GAC	AGG	CTTATGTTGAGG	-56	235
L4.1	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTTGAGG	WT/+1	255		
L4.2	GAGGAACAATGAAATGC -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGAGAC	AGG	CTTATGTTGAGG	-2/+1	233	
L5.1(R)	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	WT/WT	581		
L5.2(R)	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	WT/+1	254		
L5.1(P)	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTTGAGG	WT/+1	254		
L5.2(P)	GAGGAACAATGAA-----	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTTGAGG	-6/+2	258	
L6.1	GAGGAACAATGAAATG-----	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTTGAGG	-3/+1	253	
L6.2	GAGGAACAATGAAATGCT -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-1/+1	236	
L7.1	GAGGAACAATGAAATGC -	GAC	TGG	CCT.....TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTTGAGG	-1/WT	236	
L7.2	GAGGAACAATGAAATGCT -	CGT	TGG	CCT.....TTTGCCTGCAATGTTAGGTGAAAC	AGG	CTTATGTTGAGG	-2/+2	233	
L8.1	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGT -	GAC	AGG	CTTATGTTGAGG	WT/-1	259	
L8.2	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGAGAC	AGG	CTTATGTTGAGG	WT/+1	254		
L9.1	GAGGAACAATGAAATGCTGACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGAGAC	AGG	CTTATGTTGAGG	WT/+1	254		
L9.2	GAGGAACAATGAAATGCT -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTAC	AGG	CTTATGTTGAGG	-1/+1	256	
L10.1	GAGGAACAATGAAATGCT -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTTGAGG	-1/+1	256	
L10.2	GAGGAACAATGAAATGCT -	ACG	TGG	CCT.....TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTTGAGG	-1/+1	256	

Figure 4: . Sequence analysis of the two alleles of the 10 albino hairy root lines. The target sequences are depicted in blue and the PAM sequences in red. The change in the number of nucleotides is shown on the right of each allele sequence. A + indicates an insertion and a – stands for a deletion. WT indicates that there is no mutation. Example: -5/WT = deletion of 5 nucleotides on the first target site, no mutation on the second target site. For the L5: (R) = Hairy root stage sequence, (P) = plant stage sequence. The length of the amino acid chain in the truncated protein is shown on the right. Lx.x: line x allele x; AA: Amino Acid chain length; T1: Target1; T2: Target2.

To determine and characterize the different types of mutations obtained in our experiments, we amplified the *PDS* region including the two target sequences, using exon 5 specific primer pairs (E5-F and E5-R), and sequenced the polymerase chain reaction (PCR) products (around 420 bp). All types of mutations were detected. The monoallelic and heterozygous biallelic mutations producing superimposed sequence chromatograms from direct sequencing were decoded using Degenerate Sequence Decoding method [39]. Mutation analysis confirmed that all T0 albino plants contained mutant alleles in the *PDS* gene for at least one of the two targeted sites (Figures 4 and

5). The mutation efficiency of the first target (target 1) differs between the two types of transformation: 66.6% with protoplast transformation and 90% with *A. rhizogenes*-mediated transformation. No differences were observed for target 2 (Table 1). Most of the mutation events were small insertions, (one or two bp) or small deletions (less than 19 pb) most probably because of the NHEJ repair mechanism (Figures 4 and 5). However, in some cases, the inter-guide fragment was completely deleted, resulting in a larger deletion, as shown for L3 (Figure 4) and for C2, C4 and C7 (Figure 5). This inter-guide deletion can be detected before sequencing by electrophoresis in agarose gel, (Supplementary Figures S2 and S3). In all the albino plants analyzed, biallelic mutations responsible for gene knock out were found for at least one of the two target sites. In most cases, the insertion or deletion of a nucleotide in the target sequence induced an early termination of protein translation (Figures 4 and 5, and Supplementary Figures S4 and S5). However, a deletion of the inter-guide fragment could just create a gap in the protein as shown for C4 and C7 (Supplementary Figure S5). In these examples, this gap is sufficient to induce the knock out of the gene.

	Target 1	PAM		Target 2	PAM	T1/T2	AA
WT	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGGAC	AGG	CTTATGTT	WT/WT 581
C1.1	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTT	WT/+1 254
C1.2	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTT	WT/+1 254
C2.1	GAGGAACAATGA-----			GAC	AGG CTTATGTT	-62 233
C2.2	GAGGAACAATGA-----			GAC	AGG CTTATGTT	-62 233
C3.1	GAGGAACAATGAAATGCT - ACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTT	-1/+1 236
C3.2	GAGGAACAATGAAATGCTG	TACG	TGG	CC.....	TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTT +1/+1 234
C4.1	GAGGAACAATGA-----	CG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTT -8/+1 231
C4.2	GAGGAACA-----			-AGG	CTTATGTT	-69 558
C5.1	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGG --	GAC	AGG CTTATGTT	WT/-2 253
C5.2	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGAGAC	AGG	CTTATGTT	WT/+1 254
C6.1	GAGGAACAATGAAATGCT - ACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGAC	AGG	CTTATGTT -1/S1 236	
C6.2	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGGGAC	AGG	CTTATGTT	WT/+1 254
C7.1	GAGGAACAATGA-----	ACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGTGAC	AGG	CTTATGTT -7/+1 234
C7.2	GAGGAACA-----			-AGG	CTTATGTT	-69 558
C8.1	GAGGAACAATGAAATGCT - ACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTG - (+ npb) -	TAT	GTT	-1/+n 236
C8.2	GAGGAACAATGAAATGCT - ACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTG - (+ npb) -	TAT	GTT	-1/+n 236
C9.1	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGGGAC	AGG	CTTATGTT	WT/+1 254
C9.2	GAGGAACAATGAAATGCTGACG	TGG	CCT.....	TTTGCCTGCAATGTTAGGTGGGAC	AGG	CTTATGTT	WT/+1 254

Figure 5: Sequence analysis of the two alleles of 9 albino calli obtained after protoplast transformation. The target sequences are depicted in blue and the PAM sequences in red. The change in the number of nucleotides is shown on the right of each allele sequence. A + or a – indicate an insertion or a deletion, respectively. A S means that there is a substitution. WT indicates that there is no mutation. Example: -5/WT = deletion of 5 nucleotides on the first target site, no mutation on the second target site. The length of the amino acid chain in the truncated protein is shown on the right. Cx.x: Callus x allele x; AA: Amino Acid chain length; T1: Target1; T2: Target2. (+ npb) = n base pairs resulting from insertion of plasmid parts.

Genetic analysis of either albino plants or the calli from which they originated showed that they shared the same mutation. No mosaic plants were found, showing that the mutation event was achieved at the protoplast stage. One hairy root line, the line L5 (Figure 4), did not carry the same mutation at hairy root stage than one month later at plant stage, suggesting that the CRISPR/Cas9 system may continue to modify the genes during the hairy root stage. This has been previously shown in soybean hairy roots [13]. All other albino plants have the same mutation than the hairy root line from which they originated, suggesting that mutations were stable in these lines.

All the lines analyzed gave a PCR product at the expected size, i.e., about 420 pb, except the C8 line (Supplementary Figure S3). PCR yielded a product of about 1500 pb, suggesting that during the repair mechanism, a large DNA fragment may also be inserted at the DNA cleavage site of the CAS9 protein. This phenomenon has already been observed in *Gossypium hirsutum* in which a large endogenous fragment (99 pb) had been inserted at the target site of the CAS9 protein [32]. Sequencing results of C8 line (Figure 5) showed that a DNA fragment was inserted at the target 2 site, in both alleles. This fragment is composed of multiple pYLCRISPR-sgRNA1-sgRNA2 fragments (Supplementary Figure S6).

2.5. *Transgene detection in albino plants*

During the *A. rhizogenes*-mediated transformation, the T-DNA of the binary vector (pYLCRISPR- sgRNA1-sgRNA2) is inserted into the genome of the plant, as well as the T-DNA of the Ri (Root Inducing) plasmid containing the *rol* genes responsible for the hairy root phenotype. Indeed, due to the expression of the *virulence* (*vir*) genes present on the RI plasmid, the Right Border (RB) and Left Border (LB) of the two T-DNAs are recognized and the two T-DNAs are excised, transported, and integrated into the nuclear DNA and therefore a double transformation event occurs. PCRs with *rolB* gene specific primer pair and with CAS9 specific primer pair showed that in all the albino plant lines regenerated from hairy root, the two T-DNAs were integrated into their genomes (data not shown). Further analysis was performed to determine whether mutated protoplast-derived plants had integrated the CAS9 sequence. PCR with CAS9-specific primers showed that 3 plants out of 9 albino plants regenerated from protoplasts contained at least one fragment of the CAS9 sequence (Supplementary Figure S7). This shows that vector fragments are inserted into the

plant genome during the transformation procedure. Moreover, sequencing of the C8 plant, having inserted a fragment at the guide 2 target site, showed that this fragment was composed of many parts of the vector used for the transformation (Supplementary Figure S6).

3. Discussion

In this work, we successfully applied the CRISPR/Cas9 technology for the first time in chicory. CRISPR/Cas9 technology was already successfully used to knock out genes in other plants of the Asteraceae family such as salsify [40], lettuce [41] and dandelion [30]. Asteraceae as well as members of the Brassicaceae and Solanaceae are amenable to this biotechnological approach because of their high ability to regenerate [42]. Many successes in genome editing have also been achieved with plants belonging to the Poaceae family, not because of their great ability to regenerate but because they are certainly more studied.

The success of genome editing depends on the activity of the promoter used to drive the expression of the sgRNA. *Arabidopsis U6-26* promoter is frequently used in plants; however, in some plant species, such as in wheat and rice, it has been proven ineffective [43]. A study on soybean hairy roots (*Glycine max*) has also shown that endogenous *Gm-U6* promoter produced a mutation efficiency 2 to 6 times higher than that of *Arabidopsis U6-26* [13]. The chicory genome was screened for *U6* genes and 8 were identified based on their similarity with *Arabidopsis U6-26*. *CiU6-1p* was randomly selected among the six most conserved and was used to promote the expression of two sgRNAs targeting the fifth exon of the *CiPDS* gene. The *PDS* gene is often used as a target for the development of the method. Indeed, the mutant plants are albinos, which provides an efficient visual screen. Since 2013, it has been used in many species such as wheat [44], rice [43], populus [45], apple [46], watermelon [47], cassava [48] and banana [49,50]. In our study, the selection marker i.e., Basta resistance present on pYLCRISPR-sgRNA1-sgRNA2 was not used for selection of mutated plants. Nevertheless, this can be used to efficiently select transformants in other studies.

The CRISPR/Cas9 system was inserted in a unique vector and the *CiPDS* gene was knocked out by two different transformation methods: *A. rhizogenes* stable transformation and protoplast transient transformation with PEG. With both

transformation methods, albino plants were regenerated, with mutation frequencies of 31.3% for *A. rhizogenes*-mediated transformation and 4.5% for protoplast transformation. The mutation frequencies obtained in other species are highly variable because of the intervention of numerous parameters, such as the ability of the species to be transformed, the transformation method used, the target design, the targeted gene. For example, mutation frequency in *Arabidopsis PDS* with PEG-protoplast transfection was 1.1% whereas in tobacco it was around 37% [1]. On the other hand, in soybean hairy roots the mutation frequency ranged from 14.7% to 20.2% depending on the mutated gene [13]. Many CRISPR/Cas9 studies have been summarized in [1]. In this study, the *A. rhizogenes*-mediated transformation has given a better mutation frequency (31.25%) than the protoplast transformation (4.5%). This better mutation rate may be due to the stable transformation, which allows the CRISPR/Cas9 system to continue the gene modification during hairy root development. This feature can give mosaic plants and if the target has some potential off-target sites, this can increase the unwanted mutations. Target design is a crucial step to avoid off-target cleavage in the genome. As a result, access to genome databases is very important for genome editing using CRISPR/Cas9 technology. However, not all plant species have complete or fully accessible genomic databases. Access to full genome sequence may facilitate the design of reliable targets for fine-tuning mutations. Nevertheless, it is essential and meaningful to confirm whether CRISPR/Cas9 can be applied, to gene function studies, in less studied plants.

Characterization of the chicory mutants resulted in the observation of mono- and biallelic mutants in plant regenerated from both kinds of transformation. In biallelic mutants, the ideal material for studying gene function, mutations were simultaneously detected at the desired positions in both targeted sites. This shows that it is possible to knock out several genes at the same time using both kinds of transformation methods in chicory. This multiplex genome editing is really convenient for gene function analysis, and with the *A. rhizogenes* transformation method, especially genes involved in specialized metabolism [51]. Indeed, in hairy roots, this metabolism is known to be very active.

In comparison with transformation and plant regeneration from protoplasts, it is quick and easy to establish hairy root lines in chicory. Plant regeneration is possible within 6 weeks after transformation and vernalization is not required for flowering [28,29]. Flowering occurred as early as 1 week after regeneration and this can be a real

advantage to study genes involved in the flower development or in reproduction. The drawbacks with using the *A. rhizogenes* method to produce CRISPR-mutant plants is that the T-DNAs are stably inserted into the genome. This is not a problem for many gene function analysis, but if the objective is to produce commercial varieties, it would be an inconvenient, because T-DNAs must be removed by self-pollination and this could be a long process in chicory. Indeed, chicory is a biennial plant and is often self-incompatible [52]. In this study, we showed that foreign DNA arising from the vector containing the CRISPR/Cas9 cassette can be inserted at a non-negligible frequency (3 plants out of 9 mutant plants) into the plant genome, after transient transformation in chicory. This phenomenon has also been shown in tobacco, where Cas9 DNA has been found in 17.2 % of mutant plants [53]. This could be caused by HR events promoted by the large number of plasmids present in the cells. In addition, once in the transfected cell, the plasmids are digested by endogenous nucleases and the resulting fragments could also be inserted inside the plant genome. This could explain the insertion of multiple vector DNA fragments at “on-target” site in the C8 mutant plant (Supplementary Figure S7). These results showed that after PEG-mediated transformation, it is extremely difficult to ascertain the location and the number of foreign DNA fragments inserted inside the plant genome. To obtain foreign DNA free plants would then be really challenging. To really set up a DNA free genome editing, it will be interesting to use pre-assembled ribonucleoprotein complexes (RNPs) instead of plasmid DNA to transform protoplasts in chicory. RNPs have been used in *Arabidopsis*, tobacco, lettuce, and rice protoplasts, through PEG-mediated transformation with the same conditions than when plasmid DNA are used, and the mutation frequencies seemed to be better [54]. Therefore, having regenerated mutant plants from protoplasts transformed with PEG-mediated transformation, is a very encouraging result for the use of RNPs in chicory.

In this study, we demonstrated that CRISPR/Cas9 technology could be used to generate mutant plants in chicory. Both methods of transformation tested were effective, but *A. rhizogenes*-mediated transformation was more efficient and faster. We showed that the CRISPR/cas9 system can be used to introduce small mutations at specific loci in the chicory genome. This method constitutes a powerful and easy-to-use tool for investigating gene function in chicory. To prevent foreign DNA integration into the genome that hinders the possibility of using this method to create commercial

varieties with new agronomical traits, new editing systems such as RNPs could be assessed in chicory.

4. Materials and methods

4.1. Plant material

In vitro plants of *C. intybus* L. clone 17 and *C. intybus* L. var. Orchies were used in this study. Plants or seeds were provided by Florimond Desprez SA (Capelle-en-Pévèle, France). Seeds of *C. intybus* var. Orchies were surface sterilized and cultivated as described in Rambaud et al. [31]. Clone 17 plants were maintained on H10 solid medium (macro and micro-elements from Heller [55], Fe-Edta [56], 10 g/L sucrose, 6 g/L agar, pH 5.5) on a 16/8-h light/dark cycle with a temperature of 25 ± 2 °C in the light phase (37.6 ± 10.1 $\mu\text{mol}/\text{m}^2/\text{s}$) and at 20 °C ± 2 °C in the dark phase. Clone 17 *in vitro* plants were subcultured every 8 weeks in H10 solid medium. The subcultures consisted to remove the old H10 solid medium, to cut the old leaves at around 1.5 cm of the basis of leaves, to carve it in V form to remove the roots and to insert the refreshed plants in H10 culture medium.

4.2. Sequence identification and gRNA design

The *CiPDS* and *CiU6* promoter sequences were identified by searching the chicory genome (Unpublished database, Florimond Desprez SA, Cappelle-en-Pévèle, France) with the *Helianthus annuus* PDS (GeneBank accession AHA36971.1) and *Arabidopsis* *U6-26* genes, respectively. To mutate *CiPDS*, guide RNAs were designed based on their location in the gene and their GC content using the software crispor.tefor.net [35]. Guide1 and guide2 sequences were 5'-GGAACAATGAAATGCTGACG-3' and 5'-GCCTGCAATGTTAGGTGGAC-3', respectively.

4.3. Vector construction

This protocol is adapted from [36] and binary vector pYLCRISPR/Cas9P_{35S}-B (Addgene plasmid #66190, Cambridge, MA, USA) was used as expression vector. The *C. intybus* U6-1 promoter (*CiU6-1p*) and the sgRNA scaffold were synthesized (IDT, Leuven, Belgium) (Supplementary Figure S8) and cloned into commercial plasmid pUCIDT-Kan resulting in the vector pKanCiU6-1p-sgRNAscaffold. The primers shown

in Figure 2C have been synthetized by SIGMA-ALDRICH (Gillingham, UK). For guide adaptor hybridation, 1 µM of each primer was mixed in 0.5× Tris-EDTA buffer, pH 8.0 (10 µL) then heated to 90 °C for 10 s and returned to room temperature. To insert the target inside pKanCiU6-1p-sgRNAscaffold, 1.5 µL of cutsmart buffer (NEB, Ipswich, MA, USA), 100 ng of pKanCiU6-1p-sgRNAscaffold, 4.2 ng of target preparation, 10 U of BbsI (NEB, Ipswich, MA, USA), 40 U of T4 DNA ligase (NEB, Ipswich, MA, USA), and deionized water up to 15 µL were mixed and the digestion/ligation reaction was performed in a thermocycler using the following parameters: 3 cycles including 37 °C for 10 min, 10 °C for 5 min and 20 °C for 5 min and then 10 cycles including 37 °C for 10 min, 10 °C for 5 min, 20 °C for 10 min. Then *Escherichia coli* TOP 10 thermo-competent bacteria (In Vitrogen, Waltham, MA, USA) were transformed with 5 µL of the digestion/ligation product, and selected on LB agar plate containing 25 µg/mL kanamycin and 1 mM IPTG overnight at 37 °C. The colonies were picked and grown in 5 mL LB liquid medium supplemented with 25 µg/mL kanamycin, overnight at 37 °C. Plasmid was purified with the nucleospin plasmid purification kit (Macherey-Nagel, Düren, Germany). Two vectors, pKanCiU6-1p-sgRNA1 and pKanCiU6-1p-sgRNA2, containing the cassette “CiU6-1p-Guide-sgRNAscaffold” were obtained (Figure 2D). To add the *Bsa*I restriction sites and the 4 nucleotides recombination sites for Golden Gate restriction/ligation method, at the ends of each cassette, PCR using a proofreading polymerase (PrimeStar HS DNA polymerase; TAKARA, Kasatsu, Japan) with specific primers were run: GG1-F ; GG1-R (Table A1): with pKanCiU6-1p-sgRNA1 and GG2-F ; GG2-R (Table A1) with pKanCiU6-1p-sgRNA2, as shown in Figure 2D. PCR products were purified with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) and were assembled with Golden Gate restriction/ligation method in pYLCRISPR/Cas9P_{35S}-B as described by [36] to create pYLCRISPR-sgRNA1-sgRNA2 (Figure 2E).

4.4. *rhizogenes* transformation

The *A. rhizogenes* strain 15834 (kindly provided by Marc Buée, INRA, Nancy, France) was transformed with pYLCRISPR-sgRNA1-sgRNA2 by electroporation, as described in Nagel et al. [57]. Following transfection, 950 µL of Yeast Extract Beef (YEB) solution were added in the cuvette, transferred to a 15 mL Falcon tube and incubated for 2–3

h at 28 °C at 200 rpm. After incubation, 100 µL were streaked on YEB agar plates with kanamycin (100 µg/mL) and incubated for 3 days at 28 °C.

4.5. Hairy root induction and maintenance

Hairy root induction in *C. intybus* clone 17 using *A. rhizogenes* strain 15834 (Empty strain or transformed with pYLCRISPR-sgRNA1-sgRNA2) was performed on leaves, 7 days after subculture. The leaves were cut at petiole level and wounded with a scalpel previously soaked in a 3-day-old *A. rhizogenes* culture. The wounds were made perpendicularly to the central nervure, at 2 mm intervals, and the infected leaves were placed on Murashige and Scoog Hairy Roots (MSHR) solid medium (macro and micro-elements, Fe-Edta, vitamins from [56], 30 g/L sucrose, 6 g/L agar, pH 5.6) in Petri dishes, and grown at 25 °C ± 2 °C under 16 h light (37.6 ± 10.1 µmol/m²/s) and at 20 °C ± 2 °C under 8 h dark cycles. The first roots appeared after 12 to 13 days. Individual putative transgenic hairy roots of approximately 3 cm long, emerging from leaf explants transformed with *A. rhizogenes* were excised and immediately transferred to Petri dishes with MSHR solid medium containing antibiotic (ampicillin 300 mg/L). All the transformed roots were maintained separately as independent lines. The roots were subcultured every 7 to 10 days in new MSHR solid medium supplemented with antibiotics until the complete elimination of bacteria.

When the hairy root lines were found to be bacteria free, they were maintained in MSHR/2 solid medium (0.5 × macro and micro-elements, 0.5 × Fe-Edta, 0.5 × vitamins from [56], 30 g/L sucrose, 6 g/L agar, pH 5.6). When a hairy root line spontaneously made a shoot, it was transferred on H10 medium, for development and flowering. The hairy root lines, which did not regenerated spontaneously shoots after 4 weeks, were transferred to a C4 bud induction medium [38].

4.6. Protoplast isolation, transformation, and culture

Leaves of chicory var. Orchies were cut from 12–14-day-old plants, the external epidermis was removed from each leaf and the leaves were incubated in the washing solution [38], added with 10 g/L cellulase Onozuka R-10 (Duchefa Biochemie, Haarlem, Nederland) and 5 g/L Macerozyme R-10 (Duchefa Biochemie, Haarlem, Nederland), overnight, in the dark, at room temperature. Then protoplasts were

cleaned by filtration through a 50 µm steel mesh, collected, and washed three times with washing solution, by low speed centrifugation (150 g for 10 min). The protoplasts were resuspended in 10 mL MMG medium (100 mM MgCl₂) [58]. Protoplasts were counted in Nageotte hemocytometer and incubated for 30 min on ice. After a low speed centrifugation (5 min at 150 g), MMG (100 mM MgCl₂) was added to reach a final protoplast concentration of 6.10⁵ protoplasts/mL.

Thirty µg of pYLCRISPR-sgRNA1-sgRNA2 were added to 300 µL of protoplast solution and gently mixed. Then 330 µL of PEG4000 (30%, *m/v*) were added and mixed by gently tapping the tube. The transfection mixture was incubated at room temperature for 15 min. The transfection process was stopped by slowly adding 10 mL of washing medium. The diluted solution was centrifugated for 5 min at 150 g. The supernatant was removed, and 10 mL of MC1 culture medium [38] were added. Then, the protoplasts were partitioned in Petri dishes (3 mL per dishes) and incubated for 7 days at 30 °C in the dark. Callus proliferation and plant regeneration were achieved as described in [38].

4.7. DNA extraction, detection of mutation, and CAS9 amplification

Genomic DNA from calli, hairy roots and plants were extracted using the NucleoSpin Plant II kit (Macherey-Nagel, Düren, Germany) according to the manufacturer instructions. The genomic DNA extracted was used as a template to amplify the fifth exon of *CiPDS* by PCR. PCR was performed using a specific primer pair hybridizing between the two gRNAs (E5-F and E5-R) and the PCR products (around 420 bp) were separated in a 1% agarose gel. To analyze the mutations, the PCR products were sequenced with Sanger sequencing method using the same primers (E5-F and E5-R, Table A1) and the chromatograms were analyzed with CodonCodeAligner software [39]. The CAS9 was amplified with specific primer pair (C9-F and C9-R, Table A1) and PCR products (around 410 bp) were separated in a 1% agarose gel. For the analysis of the hairy root lines, control assays were carried with DNA extracted from plants transformed with the wild-type *A. rhizogenes* strain 15834.

Appendix A

Used software “crispor-tefor.net” [35] and “codoncode.com” [39].

Accession numbers: U6-1: MK455772, U6-2: MK455773, U6-3: MK455774, U6-4: MK455775, U6-5: MK455776), U6-6: MK455777, U6-7: MK455778, U6-8: MK455779, CiPDS: MK455771, HaPDS: AHA36971.1.

Appendix B

Tableau A1. Sequences of primers used in this study.

Name	Sequences (5'-3')
Forward primer (guide1)	ATCGGAACAATGAAATGCTGACG
Reverse primer (guide1)	AAACCGTCAGCATTTCATTGTTC
Forward primer (guide2)	ATCGCCTGCAATGTTAGGTGGAC
Reverse primer (guide2)	AAACGTCCACCTAACATTGCAGG
GG1-F	ATATATGGTCTCACTCGAAAGAACCAACCTGTTTCATAGC
GG1-R	ATTATTGGTCTCAACCTAAAAAAAAGCACCGACTCGGTG
GG2-R	ATATATGGTCTCAAGGTAAAGAACCAACCTGTTTCATAGC
GG2-F	ATTATTGGTCTCAACCGAAAAAAAAGCACCGACTCGGTG
E5-F	TTTCTTGAAGTTGGAGCTTACC
E5-R	AAACCTCAGTAGTAACCTCGATCTG
C9-F	AAGCACGTTGCTCAGATCCT
C9-R	CCGTTCGTCTCGATAAGAGG

Author Contributions: conceptualization, G.B., D.G. and C.R.; methodology, G.B., D.G., H.A.D.S., A.E and C.R.; validation, G.B., D.G. and C.R.; writing—original draft preparation, G.B. and C.R.; writing—review and editing, G.B., D.G and C.R.; visualization, G.B., D.G., J-L.H. and C.R.; supervision, C.R.; project administration, J-L.H. and C.R.; funding acquisition, J-L.H. and C.R.

Funding: G.B. was supported by a doctoral fellowship from University of Lille and the Région Hauts de France. Experiments performed at Charles Viollette Institute, were supported by an Alibiotech grant (2016–2020) obtained from the CPER/FEDER program.

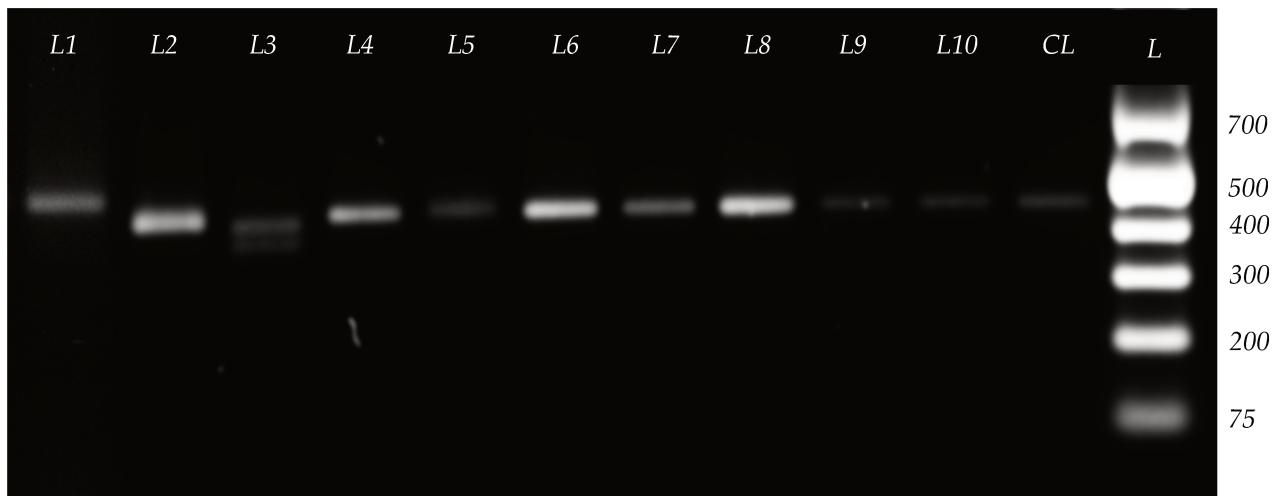
Acknowledgments: We would like to thank Florimond Desprez SA for giving us partial access to chicory genome database, we acknowledge Fabien Nogué for his wise and kind advice and we thank Pauline Letierce, and Dominic Ockenden, who provided independent mentoring to improve the language used in this paper.

Conflicts of Interest: The authors declare no conflict of interest.

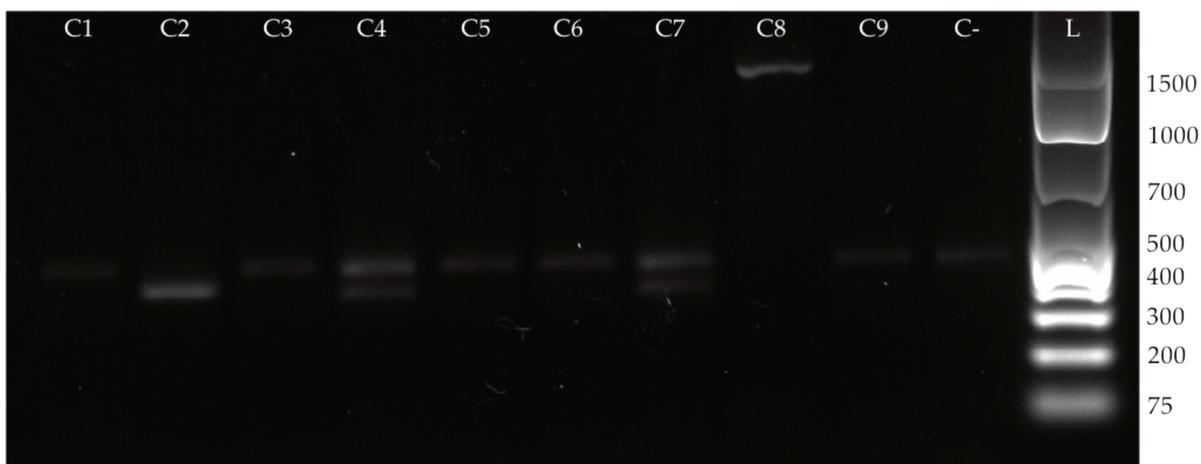
Supplementary materials

ttaggatattcggtgaggtgat**atg**tctctgttggaaatgtctccgcattaactcaag
 tggaaagtgtataatgaatcttcaga~~cacacac~~agtcattcaaaaagactgtttcaa
 catcacccaggacaaaaagatgtactctcattggatgctgtatgcacgggtcacag
 attgcaattcccaagtgc~~ccc~~gttctttacacaaaaatcaaagaatatactcccctt
 aaaggtaccttatatgctctaattattgttcatagtatcaagttaatgaacaaatgta
 tctctaataatcaaata~~at~~gtagttgtgtgactatccaagaccagatctgacaacacc
 tctaatttctt~~g~~gaagctg~~t~~ttacttgttcaac~~c~~ttcacaactccctgcccattt
 aagccattgaagattgttaattgctggcag~~g~~taaaattaaaaacttaattttaaattt
 tatttttactaac~~c~~ttactttgattatctgttatcaaactaacttttcaaattttgca
 tttcag~~g~~tttggctgg~~t~~atcaactg~~c~~aaagtattggctgatgcaggtcacaagcca
 atttgttagaagaagagacg~~t~~ttgg~~g~~aaag~~g~~tttataaaactataaagttctat
 ct~~t~~tattttcatttatctcaaata~~c~~tttatttgc~~c~~ttctaaattattacatag
 gtggcagctggaaagatgatgatggagattgg~~t~~acagacaggtt~~t~~acatattctgt
 aagtttaac~~c~~tttattatc~~t~~caagttcaaaaaatata~~g~~tcttttg~~g~~atacc
E5-F →
 taatgctgaatgaacttttaattt~~c~~ttgaag~~tt~~ggag~~c~~ttacc~~c~~aaatatacagaatt
 tg~~t~~ttggagagtt~~g~~agat~~t~~taatgatgat~~t~~g~~c~~aaatgg~~a~~agg~~a~~acattctatgatattg
 caatgc~~c~~aaataagc~~c~~ctggagaatttagt~~c~~gg~~t~~ttgactttc~~c~~agatgtttgc~~c~~ac
 cattaaatg~~g~~aaattttgc~~t~~atattg~~g~~agaacaatg~~a~~atg~~c~~tgac~~g~~**tgg**cctgaaaaaag
 taaaatttgc~~a~~attgg~~g~~cttt~~g~~c~~t~~gc~~a~~atgtt~~g~~agg~~g~~ac~~g~~**agg**cttatgtt~~g~~agg~~g~~ccc
 aagatgg~~t~~tgag~~t~~gtcaagactggatg~~g~~aaaaca~~a~~gtatg~~t~~taatg~~t~~aatcattaaa
 gtaatattttatctaaaacttt~~g~~ataatt~~g~~attacta~~a~~atcaatcttttaattttt
 tagggcata~~c~~ccagatcgagttactactgagg~~t~~ttattg~~c~~catgt~~c~~aaagg~~c~~attaaac
 ← E5-R
 tttataatcc~~c~~agat~~g~~aactt~~c~~atgc~~a~~gtgtattctcattg~~c~~tctcaaccgg~~t~~ttctt
 c~~g~~gtaaaattttat~~a~~atcc~~g~~gaagaaacaatatttttttcaattattgtt~~g~~
 ttgattaatataaaataatcaac~~g~~gaaaacatgg~~t~~ccaagatgg~~c~~atttttagatg
 gaagcccacccgaaagacttgc~~c~~ca~~g~~gtt~~g~~accatattg~~g~~agt~~c~~acttgg~~t~~ggcc
 aagt~~c~~agactcaattcaagaataaaaaattgaattaaacaagac~~g~~gaactgtcc~~g~~ga
 actttctactgaatg~~c~~acgg~~a~~atgttctagaag~~c~~tgatg~~c~~tatgtattgccacccccg
 g~~t~~atg~~t~~ttcactatttat~~t~~tatct~~t~~at~~t~~g~~t~~ataattctgatttcttaggg~~t~~ccaaa
 ttattgaattattttactttttttgtgtataaaac~~a~~gt~~t~~gacattttaaag~~c~~ttctt
 ttac~~c~~tgaagaatggaaaccaattccatattcaaaaaagtttagagaaacttagt~~c~~gg~~t~~gtt
 cc~~g~~gttataaac~~g~~ttcatat~~t~~g~~t~~tagt~~c~~gtt~~g~~actt~~c~~t~~g~~actttgacttta
 atcataacc~~g~~ttt~~t~~g~~t~~tattatctt~~a~~attttc~~g~~at~~t~~ttg~~t~~ataggaaactgaaaaa
 acacat~~a~~atgatcac~~c~~tactattc~~g~~at~~t~~ttc~~g~~at~~t~~ttg~~t~~at~~t~~g~~t~~acat~~g~~t
 c~~g~~t~~g~~acat~~t~~g~~t~~aa~~g~~gaatattatgat~~c~~gaat~~c~~gg~~t~~caat~~t~~g~~t~~g~~t~~aa~~t~~ttg~~t~~g~~c~~ac
 ctgc~~c~~agaagaatggattcaagaag~~t~~gacactg~~t~~atattgat~~c~~acaat~~g~~at~~g~~taac
 tttcaaggctttc~~c~~t~~g~~at~~g~~aaattt~~c~~agc~~g~~at~~g~~aa~~t~~g~~t~~aa~~a~~at~~c~~ct~~g~~aaat
 ataaagt~~t~~gtt~~a~~acaccaagat~~t~~gtttataaaaccataccagattgt~~g~~aa~~c~~ctt~~g~~ta
 gaccactacaaagat~~c~~t~~c~~aaattg~~g~~agg~~g~~attttat~~t~~g~~t~~g~~t~~attata~~a~~acaga
 agtatttag~~c~~t~~a~~atg~~g~~agg~~g~~tg~~c~~g~~t~~ttat~~c~~g~~t~~aa~~g~~ttt~~t~~gt~~g~~cacaag~~c~~tattg
 taaag~~g~~taagtt~~a~~agtatataaaagtag~~g~~ttgttataatt~~g~~ttaatt~~g~~tt~~a~~at~~g~~tt~~a~~at~~g~~tt~~a~~
 ttggattttgttttaggattatg~~g~~at~~g~~ttg~~c~~t~~g~~g~~c~~ac~~g~~t~~a~~agg~~g~~at~~g~~ttg~~c~~ta
 gg~~c~~aa~~g~~c~~t~~gt~~g~~**taa**gttaataattatggat~~g~~aa~~t~~tt~~a~~agg~~g~~at~~c~~t~~t~~acaag~~t~~ca
 E1 E1
 E2 E2
 E3 E3
 E4 E4
 E5 E5
 E6 E6
 E7 E7
 E8 E8
 E9 E9
 E10 E10

Supplementary Figure S1. CiPDS DNA sequence, sgRNA target sites and the primer sites for PCR amplification of the fifth exon. The exons are depicted in red and introns in dark. Blue nucleotides indicate the target sites, bold red nucleotides indicate the PAM sequences. Nucleotides highlighted in blue indicate start and stop codons, respectively.



Supplementary Figure S2: DNA amplification of albino plant regenerated from hairy root lines by PCR with specific primer pair (E5-F and E5-R) surrounding the two target sites in the fifth exon of the *CiPDS*. L = Line, CL = Control Line (DNA from plant transformed with the wild type *A.rhizogenes* strain 15834), L = Ladder.



Supplementary Figure S3: Amplification by PCR of DNA extracted from albino plants regenerated from calli with specific primer pair (E5-F and E5-R) surrounding the two target sites in the fifth exon of the *CiPDS*. C = Callus, C- = DNA from plant regenerated from non-mutated callus, L = Ladder.

WT-KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGGQAYVEAQDGL...
 L1.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNE^{NVA}*
 L1.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L2.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L2.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLV^A*
 L3.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L3.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLT^{GLC}*
 L4.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGV^{TGLC}*
 L4.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLH^VA*
 L5.1 (R) - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGGQAYVEAQDGL...
 L5.2 (R) - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGG^{TGLC}*
 L5.1 (P) - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGV^{TGLC}*
 L5.2 (P) - KPGEFSRDFDPDVLPA^PLNGIFAILRNNETWPEKVKAIGLLPAMLGV^{DRLMLRPKMV}*
 L6.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGV^{TGLC}*
 L6.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L7.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L7.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLV^A*
 L8.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGD^R^{LMLRPKMV}*
 L8.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGET^{GLC}*
 L9.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGR^T^{GLC}*
 L9.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L10.1- KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 L10.2- KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*

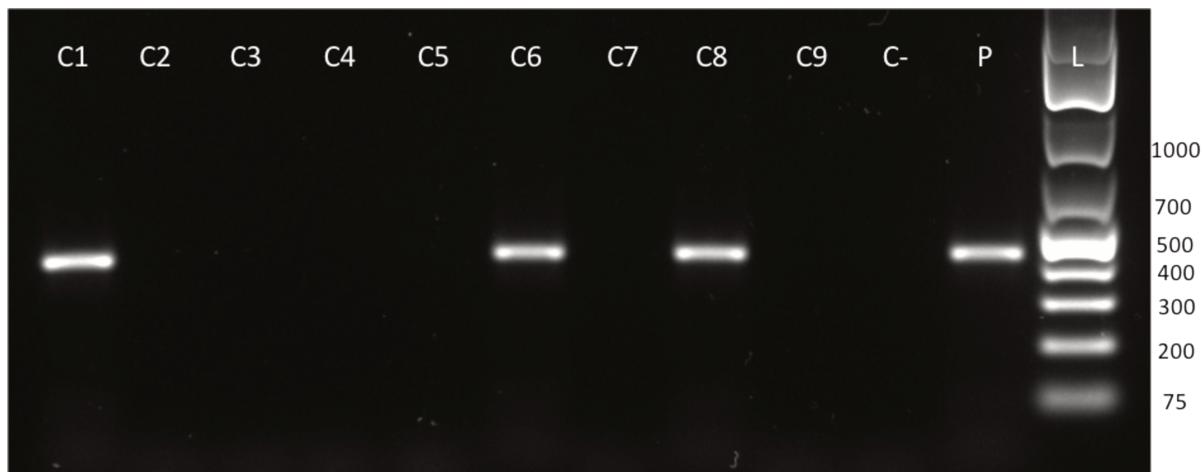
Supplementary Figure S4: Predicted amino acid sequences of the CiPDS from the 10 albino hairy root lines and their shoots. Lx.x = Line x. allele x. For L5 (R) = Hairy root stage sequence, (P) = Plant stage sequence. In blue, the amino acids that differ from the wild type sequences. (-) shows the deleted amino acids and (*) shows the end of the amino acid sequence.

WT- KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGGQAYVEAQDGL...
 C1.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGV^{TGLC}*
 C1.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGV^{TGLC}*
 C2.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNET^{GLC}*
 C2.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNET^{GLC}*
 C3.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 C3.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLY^VA*
 C4.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNND^VA*
 C4.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNK-----AYVEAQDGL...
 C5.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLG^{TGLC}*
 C5.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGET^{GLC}*
 C6.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 C6.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGG^{TGLC}*
 C7.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 C7.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNK-----AYVEAQDGL...
 C8.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 C8.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLR^GLKK*
 C9.1 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGG^{TGLC}*
 C9.2 - KPGEFSRDFDPDVLPA^PLNGIFAILRNNEMLTWPEKVKAIGLLPAMLGG^{TGLC}*

Supplementary Figure S5: Predicted amino acid sequence of the CiPDS from the 9 albino calli and their shoots. Cx.x = Callus x. allele x. In blue, the amino acids that differ from the wild type sequences, (-) shows the deleted amino acids and (*) shows the end of the amino acid sequence.

(TGAGTAGTCCCAGATAAGGGAATTAGGGTTCTTATAGGGTTCGCTCATGTGTTGAGCATATAAGAAACCCTAGTATGTATT
TGTATTTGAAAATCTTCTATCAATAAAATTCTAATTCTAAAACCAAAATCCAGTGACGCCGCGCATAAGGCCGCCATCTC
 GATGACGATGTTCTCAGGCTTGTGACGCCGTAAGCGGCTGGGTCCTGCCGCCCTGCAATGGCACTGGAAACCCCCAAGC
 CCGAGGAATCGCGTGAGCGTCGCAAACCATCCGGCCCGTACAAATCGCGCGCTGGGTGATGACCTGGTGGAGAA
 GTGAAGGCCGCAGGCCAGCGCAACGCGTACAGGCAAGCAGCAGCCCCGGTAATCGTGGCAAGCGGCCGCTGA
 TCGAATTGCGCAACCGGAACCGTACATTGGGAA(nn...nnn)TTCCCAATGTACGTGCTATCCACAGGAAAGAGAACTTTG
ACCTTTCCCCTGCTAGGGCAATTGCCCTAGCATCTGCTCCGTACATTAGGAACCGGCGATGCTCGCCCTGATCAGGTTG

Supplementary Figure S6: Partial sequence of the DNA fragments inserted at the target 2 site of the C8 mutant. The insert is composed of more than 6 fragments (depicted in red, black, green, purple or orange) of the pYLCRISPR-sgRNA1-sgRNA2. (nnn...nnn) is a part of the fragment, which was not sequenced.



Supplementary Figure S7: DNA amplification of albino plants regenerated from calli by PCR with specific primer pair (C9-F and C9-R) hybridizing in CAS9 gene. C = Callus, C- = DNA from plant regenerated from non-mutated callus, P = Plasmid pYLCRISPR-sgRNA1-sgRNA2 DNA, L = Ladder.

GGAATTGGCTGGCTCTCTCGAAAGAACCAACCTGTTTCTAGCTCTGTTGACTGATAGAGTTTGATA
 TGTTCAACTTCAGGTAAAACCAAGAGGATAGTAAAATTGAGGGTTAGCTTACAAATATTATCGGTTCCCTA
 GCTATCGTAATGCCATGGCTTCCCACATCGCTCATTGGAGCAACATGCCATGCTTATATAGCCTCGCCT
 CCAACCATTATCGATGTCTTCgAGTTAGCTCACTCATTAGGCACCCAGGCTTACACTTATGCTTCCG
 GCTCGTATGTTGTGGAATTGTGAGCGATAACAATTACACAAGAAACAGCTATGACCATGATTACGC
 CAAGCTATTAGGTGACACTATAGAATACTCAAGCTATGCATCAAGCTCAATGGGTCTAGTCTGTAGATAC
 CCATCACACTGGCGACCGCTCGAACATCAGTTAAGGTTACACCTATAAAAGAGAGAGGCCGTATCGTC
 TGTTTGTGGATGTACAGAGTGTATTATTGACACGCCGGGCGACGGATGGTGTACCCCTGGCCAGTG
 CACGTCTGCTGTCAGATAAAGTCTCCCGTGAACCTTACCCGGTGGTGCATATCGGGGATGGAAGCTGGC
 GCATGATGACCACCGATATGCCAGTGTGCTGCTCCGTATCGGGGAAGAAGTGGCTGATCTCAGCC
 ACCCGCGAAAATGACATAAAAACGCCATTAAACCTGATGTTCTGGGGATATAATGTCAGGCCTGAATGG
 CGAATGGACCGCCCTGTAGCGCGCATTAGCCGGCGGGTGAGCGTgGAAGACTAGTTTAGAGCTA
 GAAATAGCAAGTTAAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT
 TGAGACCGACCGGATCCC

Supplementary Figure S8: Sequence of the cassette “CiU6-PlacZ ccdb-sgRNAscaffold” harboring by the plasmid pKanCiU6-1p-sgRNAscaffold. The CiU6-1p, the selectable marker PlacZ-ccdB and the sgRNA scaffold are written in blue, yellow and green respectively. In red is highlighted the *BbsI* recognition site which achieves a cleavage on the left side (shown in underlined and bold). Note that the G present in the 4 nucleotides of the cutting site, is the transcription initiator. Highlighted in blue is the *BbsI* recognition site which achieves a cleavage on the right side (shown in underlined and bold).

References

- Bortesi, L.; Fischer, R. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* **2015**, *33*, 41–52; doi:10.1016/j.biotechadv.2014.12.006.
- Kim, Y.G.; Cha, J.; Chandrasegaran, S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156–1160; doi:10.1073/pnas.93.3.1156.
- Christian, M.; Cermak, T.; Doyle, E.L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A.J.; Voytas, D.F. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **2010**, *186*, 757–761; doi:10.1534/genetics.110.120717.
- Cao, H.X.; Wang, W.; Le, H.T.T.; Vu, G.T.H. The Power of CRISPR-Cas9-Induced Genome Editing to Speed Up Plant Breeding. *Int. J. Genom.* **2016**, *2016*, 1–10; doi:10.1155/2016/5078796.
- Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–609; doi:10.1038/nature09886.
- Feng, Z.; Zhang, Z.; Hua, K.; Gao, X.; Mao, Y.; Botella, J.R.; Zhu, J.K. A Highly Efficient Cell Division-Specific CRISPR/Cas9 System Generates Homozygous Mutants for Multiple Genes in *Arabidopsis*. *Int. J. Sci.* **2018**, *19*, 3925; doi:10.3390/ijms19123925.
- Belhaj, K.; Chaparro-Garcia, A.; Kamoun, S.; Nekrasov, V. Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **2013**, *9*, 39; doi:10.1186/1746-4811-9-39.
- Engler, C.; Gruetzner, R.; Kandzia, R.; Marillonnet, S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIIs Restriction Enzymes. *PLoS ONE* **2009**, *4*, e5553, doi:10.1371/journal.pone.0005553.
- Xing, H.L.; Dong, L.; Wang, Z.P.; Zhang, H.Y.; Han, C.H.; Liu, B.; Wang, X.C.; Chen, Q.J. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* **2014**, *14*, 327, doi:10.1186/s12870-014-0327-y.
- Li, J.F.; Norville, J.E.; Aach, J.; McCormack, M.; Zhang, D.; Bush, J.; Church, G.M.; Sheen, J. Multiplex and homologous recombination—Mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **2013**, *31*, 688–691; doi:10.1038/nbt.2654.
- Liang, Z.; Zhang, K.; Chen, K.L.; Gao, C.X. Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *J. Genet. Genom.* **2014**, *41*, 63–68; doi:10.1016/j.jgg.2013.12.001.
- Shan, Q.W.; Wang, Y.P.; Li, J.; Gao, C.C. Genome editing in rice and wheat using the CRISPR/Cas system. *Nat. Protoc.* **2014**, *9*, 2395–2410; doi:10.1038/nprot.2014.157.

Sun, X.; Hu, Z.; Chen, R.; Jiang, Q.; Song, G.; Zhang, H.; Xi, Y. Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci. Rep.* **2015**, *5*, 10342; doi:10.1038/srep10342.

Cermak, T.; Baltes, N.J.; Cegan, R.; Zhang, Y.; Voytas, D.F. High-frequency, precise modification of the tomato genome. *Genome Biol.* **2015**, *16*, 232; doi:10.1186/s13059-015-0796-9.

Tzfira, T.; Li, J.; Lacroix, B.; Citovsky, V. *Agrobacterium* T-DNA integration: Molecules and models. *Trends Genet.* **2004**, *20*, 375–383; doi:10.1016/j.tig.2004.06.004.

Tepfer, D. Genetic transformation using *Agrobacterium rhizogenes*. *Physiol. Plant.* **1990**, *79*, 140–146; doi:10.1111/j.1399-3054.1990.tb05876.x.

Ron, M.; Kajala, K.; Pauluzzi, G.; Wang, D.; Reynoso, M.A.; Zumstein, K.; Garcha, J.; Winte, S.; Masson, H.; Inagaki, S.; Federici, F.; Sinha, N.; Deal, R.B.; Bailey-Serres, J.; Brady, S.M. Hairy Root Transformation Using *Agrobacterium rhizogenes* as a Tool for Exploring Cell Type-Specific Gene Expression and Function Using Tomato as a Model. *Plant Physiol.* **2014**, *166*, 455–469; doi:10.1104/pp.114.239392.

Kirchner, T.W.; Niehaus, M.; Debener, T.; Schenk, M.K.; Herde, M. Efficient generation of mutations mediated by CRISPR/Cas9 in the hairy root transformation system of *Brassica carinata*. *PLoS ONE* **2017**, *12*, e0185429. doi:10.1371/journal.pone.0185429.

Zhou, Z.; Tan, H.; Li, Q.; Chen, J.; Gao, S.; Wang, Y.; Chen, W.; Zhang, L. CRISPR/Cas9-mediated efficient targeted mutagenesis of RAS in *Salvia miltiorrhiza*. *Phytochemistry* **2018**, *148*, 63–70; doi:10.1016/j.phytochem.2018.01.015.

Malarz, J.; Stojakowska, A.; Kisiel, W. Long-term cultured hairy roots of chicory—A rich source of hydroxycinnamates and 8-deoxylactucin glucoside. *Appl. Biochem. Biotechnol.* **2013**, *171*, 1589–1601; doi:10.1007/s12010-013-0446-1.

Legrand, G.; Delporte, M.; Khelifi, C.; Harant, A.; Vuylsteke, C.; Mörchen, M.; Hance, P.; Hilbert, J.L.; Gagneul, D. Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.* **2016**, *7*, 741; doi:10.3389/fpls.2016.00741.

Delporte, M.; Bernard, G.; Legrand, G.; Hielscher, B.; Lanoue, A.; Moliné, R.; Rambaud, C.; Mathiron, D.; Besseau, S.; Linka, N.; Hilbert, J.L.; Gagneul, D. A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in pollen coat of the Asteraceae family. *J. Exp. Bot.* **2018**, *69*, 5355–5371; doi:10.1093/jxb/ery320/5088093.

Bais, H.P.; Suresh, B.; Raghavarao, K.S.M.S.; Ravishankar, G.A. Performance of hairy root cultures of *Cichorium intybus* L. in bioreactors of different configurations. *In Vitro Cell. Dev. Biol.* **2002**, *38*, 573–580; doi:10.1079/IVP2002334.

Das, S.; Vasudeva, N.; Sharma, S. *Cichorium intybus*: A concise report on its ethnomedicinal, botanical, and phytopharmacological aspects. *Drug Dev. Ther.* **2016**, *7*, 1–12; doi:10.4103/2394-6555.180157.

Street, R.A.; Sidana, J.; Prinsloo, G. *Cichorium intybus*: Traditional Uses, Phytochemistry, Pharmacology, and Toxicology. *Evid.-Based Complementary Altern. Med.* **2013**, doi:10.1155/2013/579319.

Bais, H.P.; Ravishankar, G.A. *Cichorium intybus* L.—Cultivation, processing, utility, value addition and biotechnology, with an emphasis on current status and future prospects. *J. Sci. Food Agric.* **2001**, *81*, 467–484; doi.org/10.1002/jsfa.817.

Maroufi, A.; Karimi, M.; Mehdikhanlou, K.; Bockstaele, E.V.; De Loose, M. Regeneration ability and genetic transformation of root type chicory (*Cichorium intybus* var. *sativum*). *Afr. J. Biotechnol.* **2012**, *11*, 11874–11886; doi:10.5897/AJB11.438.

Bais, H.P.; Venkatesh, R.T.; Chandrashekhar, A.; Ravishankar, G.A. *Agrobacterium rhizogenes*-mediated transformation of Witloof chicory—*In vitro* shoot regeneration and induction of flowering. *Curr. Sci.* **2001**, *80*, 83–87.

Bogdanović, M.D.; Todorović, S.I.; Banjanac, T.; Dragićević, M.B.; Verstappen, F.W.A.; Bouwmeester, H.J.; Simonović, A.D. Production of guaianolides in *Agrobacterium rhizogenes*—Transformed chicory regenerants flowering in vitro. *Ind. Crops Prod.* **2014**, *60*, 52–59; doi:10.1016/j.indcrop.2014.05.054.

Iaffaldano, B.; Cornish, K.; Zhang, Y. CRISPR/Cas9 genome editing of rubber producing dandelion *Taraxacum kok-saghyz* using *Agrobacterium rhizogenes* without selection. *Ind. Crops Prod.* **2016**, *89*, 356–362; doi:10.1016/j.indcrop.2016.05.029.

Rambaud, C.; Dubois, J.; Vasseur, J. Male-sterile chicory cybrids obtained by intergeneric protoplast fusion. *Theor. Appl. Genet.* **1993**, *87*, 347–352; doi:10.1007/BF01184922.

Zhu, S.; Yu, X.; Li, Y.; Sun, Y.; Zhu, Q.; Sun, J. Highly Efficient Targeted Gene Editing in Upland Cotton Using the CRISPR/Cas9 System. *Int. J. Mol. Sci.* **2018**, *19*, 3000; doi:10.3390/ijms19103000.

Waibel, F.; Filipowicz, W. U6 snRNA genes of *Arabidopsis* are transcribed by RNA polymerase III but contain the same two upstream promoters as RNA polymerase II-transcribed U-snRNA genes. *Nucleic Acids Res.* **1990**, *18*, 3451–3458.

Qin, G.; Gu, H.; Ma, L.; Peng, Y.; Deng, X.W.; Chen, Z.; Qu, L.J. Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res.* **2007**, *17*, 471–482; doi:10.1038/cr.2007.40.

Haeussler, M.; Schönig, K.; Eckert, H.; Eschstruth, A.; Mianné, J.; Renaud, J.B.; Schneider-Maunoury, S.; Shkumatava, A.; Teboul, L.; Kent, J.; et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **2016**, *17*, 148; doi:10.1186/s13059-016-1012-2.

Ma, X.; Liu, Y.G. CRISPR/Cas9-Based Multiplex Genome Editing in Monocot and Dicot Plants. *Curr. Protoc. Mol. Biol.* **2016**, *115*, 31.6.1–31.6.21; doi:10.1002/cpmb.10.

Weber, E.; Engler, C.; Gruetzner, R.; Werner, S.; Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **2011**, *6*, e16765; doi:10.1371/journal.pone.0016765.

Rambaud, C.; Vasseur, J. Somatic Hybridization in *Cichorium intybus* L. (Chicory). In *Biotechnology in Agriculture and Forestry*; Nagata, T., Bajaj, Y.P.S., Eds.; Springer: Berlin, Germany, 2001; Volume 49, pp. 112–123.

Ma, X.; Chen, L.; Zhu, Q.; Chen, Y.; Liu, Y.G. Rapid Decoding of Sequence-Specific Nuclease-Induced Heterozygous and Biallelic Mutations by Direct Sequencing of PCR Products. *Mol. Plant* **2015**, *8*, 1285–1287; doi:10.1016/j.molp.2015.02.012.

Shan, S.; Mavrodiev, E.; Li, R.; Zhang, Z.; Hauser, B.A.; Soltis, P.S.; Soltis, D.E.; Yang, B. Application of CRISPR/Cas9 to *Tragopogon* (Asteraceae), an evolutionary model for the study of polyploidy. *Mol. Ecol. Resour.* **2018**, *18*, 1427–1443; doi:10.1111/1755-0998.12935.

Bertier, L.D.; Ron, M.; Huo, H.; Bradford, K.J.; Britt, A.B.; Michelmore, R.W. High-Resolution Analysis of the Efficiency, Heritability, and Editing Outcomes of CRISPR/Cas9-Induced Modifications of NCED4 in Lettuce (*Lactuca sativa*). *G3* **2018**, *8*, 1513–1521; doi:10.1534/g3.117.300396.

Hill, K.; Schaller, G.E. Enhancing plant regeneration in tissue culture A molecular approach through manipulation of cytokinin sensitivity. *Plant Signal. Behav.* **2013**, *8*, e25709; doi:10.4161/psb.25709.

Shan, Q.; Wang, Y.; Li, J.; Zhang, Y.; Chen, K.; Ling, Z.; Zhang, K.; Liu, J.; Xi, J.J.; Qiu, J.L.; et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 684–686; doi:10.1038/nbt.2650.

Upadhyay, S.K.; Kumar, J.; Alok, A.; Tuli, R. RNA-Guided Genome Editing for Target Gene Mutations in Wheat. *G3* **2013**, *3*, 2233–2238; doi:10.1534/g3.113.008847.

Fan, D.; Liu, T.; Li, C.; Jiao, B.; Li, S.; Hou, Y.; Luo, K. Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in *Populus* in the First Generation. *Sci. Rep.* **2015**, *5*, 12217; doi:10.1038/srep12217.

Nishitani, C.; Hirai, N.; Komori, S.; Wada, M.; Okada, K.; Osakabe, K.; Yamamoto, T.; Osakabe, Y. Efficient Genome Editing in Apple Using a CRISPR/Cas9 system. *Sci. Rep.* **2016**, *6*, 31481; doi:10.1038/srep31481.

Tian, S.; Jiang, L.; Gao, Q.; Zhang, J.; Zong, M.; Zhang, H.; Ren, Y.; Guo, S.; Gong, G.; Liu, F.; et al. Efficient CRISPR/Cas9-based gene knockout in watermelon. *Plant Cell Rep.* **2017**, *36*, 399–406; doi:10.1007/s00299-016-2089-5.

Odipio, J.; Alicai, T.; Ingelbrecht, I.; Nusinow, D.; Bart, R.; Taylor, N.J. Efficient CRISPR/Cas9 Genome Editing of *Phytoene desaturase* in Cassava. *Front. Plant Sci.* **2017**, *8*, 1780; doi:10.3389/fpls.2017.01780.

Naim, F.; Dugdale, B.; Kleidon, J.; Brinin, A.; Shand, K.; Waterhouse, P.; Dale, J. Gene editing the phytoene desaturase alleles of Cavendish banana using CRISPR/Cas9. *Transgenic Res.* **2018**, *27*, 451–460; doi:10.1007/s11248-018-0083-0.

Kaur, N.; Alok, A.; Shivani.; Kaur, N.; Pandey, P.; Awasthi, P.; Tiwari, S. CRISPR/Cas9-mediated efficient editing in *phytoene desaturase* (PDS) demonstrates precise manipulation in banana cv. Rasthali genome. *Funct. Integr. Genom.* **2018**, *18*, 89–99; doi:10.1007/s10142-017-0577-5.

Sharma, P.; Padh, H.; Srivastava, N. Hairy root cultures : A suitable biological system for studying secondary metabolic pathways in plants. *Eng. Life Sci.* **2013**, *13*, 62–75; doi:10.1002/elsc.201200030.

Gontier, L.; Blassiau, C.; Mörchen, M.; Cadalen, T.; Poiret, M.; Hendriks, T.; Quillet, M.C. High-density genetic maps for loci involved in nuclear male sterility (NMS1) and sporophytic self-incompatibility (S-locus) in chicory (*Cichorium intybus* L., Asteraceae). *Theor. Appl. Genet.* **2013**, *126*, 2103–2121; doi:10.1007/s00122-013-2122-9.

Lin, C.S.; Hsu, C.T.; Yang, L.H.; Lee, L.Y.; Fu, J.Y.; Cheng, Q.W.; Wu, F.H.; Hsiao, H.C.W.; Zhang, Y.; Zhang, R.; et al. Application of protoplast technology to CRISPR/Cas9 mutagenesis: From single-cell mutation detection to mutant plant regeneration. *Plant Biotechnol. J.* **2018**, *16*, 1295–1310; doi:10.1111/pbi.12870.

Woo, J.W.; Kim, J.; Kwon, S.I.; Corvalan, C.; Cho, S.W.; Kim, H.; Kim, S.G.; Kim, S.T.; Choe, S.; Kim, J.S. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* **2015**, *33*, 1162–1165; doi:10.1038/nbt.3389.

Heller, R. Researches on the mineral nutrition of plant development. *Ann. Sci. Nat. Bot. Biol. Vg.* **1953**, *14*, 1–223.

Murashige, T.; Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **1962**, *15*, 473–497.

Nagel, R.; Elliott, A.; Masel, A.; Birch, R.G.; Manners, M.J. Electroporation of binary Ti plasmid vector into *Agrobacterium tumefaciens*, and *Agrobacterium rhizogenes*. *FEMS Microbiol. Lett.* **1990**, *67*, 325–328; doi.org/01.1111/j.1574-6968.1990.tb04041.x.

Yoo, S.D.; Cho, Y.H.; Sheen, J. *Arabidopsis* mesophyll protoplasts: A versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2007**, *2*, 1565–1572; doi:10.1038/nprot.2007.199.

***Chapitre III : Mise en place de
mutants sht1 et sht2 grâce à
l'utilisation de la technologie
CRISPR/Cas 9***

Avant-propos

Depuis sa validation fonctionnelle chez de nombreuses espèces végétales, l'outil d'édition du génome CRISPR/Cas9 a beaucoup été utilisé pour améliorer les caractéristiques de production des plantes cultivées. Par exemple, de nombreuses composantes agronomiques du riz, une des denrées les plus consommées dans le monde, ont été améliorées par l'utilisation de cet outil. En effet le poids du grain de riz a été augmenté grâce à l'édition par CRISPR/Cas9 de 3 gènes régulant négativement le poids du grain (Xu et al. 2016). Le potentiel nutritif du grain de riz a aussi été amélioré, il a été démontré que le Knock Out (KO) du gène *SBEIb* (start branching enzymes iso-forme b) provoquait une augmentation de 25 % du contenu en amylose, ce qui induisait une augmentation du contenu en amidon de 9,8 % (Y. Sun et al. 2017). La mise en place de techniques permettant de synthétiser plusieurs sgRNAs ainsi que la Cas9 grâce à un plasmide unique (Xingliang Ma and Liu 2016) a permis l'amélioration de plantes ayant des génomes plus complexes, comme le blé qui a un génome hexaploïde. Chez le blé, la mutation d'environ 40 gènes dans la même lignée a permis de produire un blé avec un contenu en gluten diminué (Sánchez-León et al. 2018). Outre l'utilisation de cet outil pour l'amélioration des caractères agronomiques, CRISPR/Cas9 a aussi été utilisé pour améliorer la résistance des plantes aux stress biotiques ou abiotiques. L'attaque des plantes cultivées par des agents pathogènes (champignons, bactéries, virus) est la principale cause de perte de rendement. Pour lutter contre ces agents pathogènes de nombreuses molécules chimiques sont utilisées, molécules très efficaces, mais ayant des répercussions néfastes sur l'environnement et la santé humaine. Dans le but de développer des variétés résistantes et donc de diminuer l'utilisation de produits phytosanitaires, les spécialistes de la lutte contre les agents pathogènes ont réussi à cibler de nombreux gènes grâce à la technologie CRISPR/Cas9. Des variétés de tomate plus résistantes au mildiou ont été générées grâce à l'édition du gène *Mildew Locus O*, d'autres exemples sont répertoriés dans le tableau 3.1. La technologie CRISPR/Cas9 a aussi été utilisée pour augmenter la tolérance aux stress abiotiques. En effet, au vu des conditions climatiques changeantes allant vers un réchauffement global à la surface de la Terre, de nombreuses études ont porté sur la tolérance à la sécheresse des plantes cultivées. Chez le maïs, grâce à l'utilisation de la technologie CRISPR/Cas9 l'expression du gène *ARGOS8* a été augmentée, grâce à l'insertion d'un promoteur *GOS2* plus fort que le

promoteur naturel. Dans ce cas-là, l'ajout d'ADN homologue possédant le nouveau promoteur a permis l'insertion de celui-ci grâce au mécanisme de réparation haute-fidélité dirigé par l'homologie (HDR). L'augmentation de l'expression de ce gène a produit un maïs capable de produire la même quantité de grains qu'un maïs sauvage tout en étant exposé à des conditions de sécheresse (Shi et al. 2017).

Target gene	Type of modification	Phenotype/disease resistance	Encoded protein	Plants species
OSERF922	Indels in ORF	Enhanced resistance against <i>Magnaporthe oryzae</i>	ERF family transcription factor	<i>Oryza sativa</i>
SIMlo1	Deletion in ORF	Resistance against <i>Oidium neolycopersici</i> (tomato powdery mildew)	Transmembrane protein	<i>Solanum lycopersicum</i>
TaEDR1	Deletion in Exon4	Resistance against <i>Blumeria graminis</i> f. sp. <i>tritici</i> (powdery mildew)	Raf-like mitogen-activated protein kinase kinase kinase (MAPKKK)	<i>Triticum aestivum</i>
SIDmr6	Deletion in Exon3	Enhanced resistance against <i>Xanthomonas</i> , <i>Pseudomonas</i> and <i>Phytophtora capsici</i>	2-oxoglutarate (2OG) Fe(II)-dependent oxygenases	<i>Solanum lycopersicum</i>
Cs-eIFE	Deletion in Exon 1	Resistance against <i>Cucumber vein yellowing virus</i> , <i>Zucchini yellow mosaic virus</i> , and <i>papaya rings spot virus</i> ; no resistance against CMV or CGMMV	Eukaryotic translation initiator factor	<i>Cucumis sativus</i>
CsiLOB1	Small deletions in promoter region	Enhanced resistance against <i>Xanthomonas citri</i> subsp. <i>citri</i> (citrus canker)	LOB domain regulation of plant organ development	<i>Citrus sinensis</i>

Tableau 3.1 : Exemples d'espèces végétales dont la résistance aux pathogènes a été améliorée grâce à l'utilisation de la technologie CRISPR/Cas9 (D'après Langner, Kamoun, et Belhaj 2018).

En plus d'être utilisée pour l'amélioration des plantes agronomiques, la technologie CRISPR/Cas9 peut aussi être utilisée comme outil pour l'ingénierie métabolique des plantes. En effet, les données accumulées depuis des décennies sur les voies de biosynthèse des métabolites spécialisés permettent maintenant de créer des plantes avec des productions de biomolécules orientées. Par exemple chez *Panicum virgatum*, la technologie CRISPR/Cas9 a été utilisée pour créer des lignées plus propices à la production de biocarburant en diminuant la production des lignines. En effet, la lignine est un composant majeur des parois cellulaires et constitue un obstacle majeur à la production de biocarburant en limitant l'accès des enzymes aux polymères de la paroi cellulaire pendant le processus de fermentation. Grâce à la mutation de trois gènes *4CL*, enzymes impliquées dans les premières étapes de la voie de biosynthèse des lignines, les scientifiques ont donc réussi à produire des lignées produisant jusqu'à 30% de lignine en moins, les rendant donc plus fermentescibles (Park et al. 2017). Chez les plantes produisant des molécules à forte valeur ajoutée, la manipulation des voies de biosynthèse en utilisant la technologie CRISPR/Cas9 est un peu plus délicate du fait de la faiblesse des données génomiques accessibles. Cependant quelques études existent, comme chez le pavot à opium (*Papaver somniferum L.*) ou la mutation d'une *3'-hydroxyl-N-methylcochlaurine 4'-O-metyltransferase (4'OMT)* a permis la modulation de la production d'alcaloïdes chez les lignées mutées (Alagoz et al. 2016). Enfin, depuis son apparition, cet outil a aussi été très utile en recherche fondamentale, pour la validation fonctionnelle de gènes chez les plantes. De nombreuses études ont porté sur le développement des plantes en utilisant *Arabidopsis* comme modèle expérimental. La caractérisation de gènes impliqués, par exemple, dans la fermeture des stomates, la différenciation des tissus vasculaires ou dans les étapes de la photorespiration, a été réalisée grâce à l'utilisation de cette technologie (Osakabe et al. 2016, Saito, Kondo, et Fukuda 2018, Yang et al. 2018). Avec le séquençage de nombreux génomes, et l'accessibilité aux données, la recherche fondamentale sur *Arabidopsis* a peu à peu diminué au profit de celle sur des plantes cultivées, mais la plupart du temps cette recherche est axée sur des gènes impliqués dans le métabolisme primaire des plantes. Concernant la recherche sur le métabolisme spécialisé chez les plantes et l'élucidation de nouvelles voies de biosynthèse, aucune étude utilisant CRISPR/Cas9 n'a encore montré la caractérisation de nouveaux gènes impliqués dans ce métabolisme. En effet, les plantes intéressantes dans la production de métabolites spécialisés ont moins été étudiées, et peuvent être récalcitrantes à la

transformation ou à la régénération en culture *in vitro*, de plus les données génomiques de ces plantes sont souvent inexistantes ou peu accessibles.

Cichorium intybus L., qui est une plante agronomique, et qui est aussi connue pour ses vertus thérapeutiques et la production de métabolites spécialisés (Street, Sidana, et Prinsloo 2013) a fait l'objet du séquençage de son génome par un consortium privé, dont fait partie l'entreprise Florimond Desprez (FD). Grâce à ce séquençage, l'entreprise FD a pu réaliser le re-séquençage d'un génotype de chicorée industrielle (le clone 17) et du fait de notre étroite collaboration avec l'entreprise FD, nous avons un accès privilégié à la séquence du génome du clone 17. De plus, notre laboratoire possède une grande expertise en culture *in vitro* de la chicorée, depuis de nombreuses années (Vasseur 1990, Rambaud et al. 1992, Cappelle et al. 2007), ce qui nous a permis de mettre en place l'outil CRISPR/Cas9 via deux techniques de transformation, à savoir la transformation stable médiaée par *A. rhizogenes* et la transformation de protoplastes médiaée par le PEG comme montré précédemment dans le chapitre 2 (Bernard et al. 2019). Ces deux techniques de transformation permettent la régénération de plantes mutées, mais il s'est avéré que l'utilisation d'*A. rhizogenes* est plus efficace et la régénération de plantes, plus rapide, ce qui est un énorme avantage en recherche fondamentale. La régénération de plantes est possible en environ 6 semaines et les plantes ne nécessitent pas de vernalisation pour fleurir (Harsh Pal Bais and Venkatesh 2001). La floraison peut intervenir une semaine après la régénération de la plante ce qui est un réel avantage pour l'étude de gènes impliqués dans la floraison ou la reproduction comme les gènes *CiSHT1* et *CiSHT2* étudiés dans le chapitre 1, et impliqués dans la voie de biosynthèse de spermine totalement substituée, au niveau des anthères de la chicorée.

Au cours du chapitre précédent, nous avons pu constater que l'enzyme *CiSHT2* était capable de produire des dérivés de spermine et de spermidine *in vitro*. *In planta* (chevelus racinaires), la présence de mono-, di- et tricoumaroyl spermine/spermidine a été détectée mais aucune trace de tetracoumaroyl spermine n'a été mise en évidence, or dans les boutons inflorescentiels de chicorée, seules la tricoumaroyl spermidine et la tétracoumaroyl spermine sont détectées. Les conditions de réaction (pH, biodisponibilité des substrats etc ...) différentes *in vitro* et dans les racines, de celles du tapetum des anthères peuvent expliquer cette différence mais pour l'instant, ces résultats ne nous permettent pas d'affirmer avec certitude que le gène, *CiSHT2*, mis en évidence chez la chicorée est le seul acteur dans la production de la

tricoumaroyl spermidine et de la tetracoumaroyl spermine dans les boutons floraux. D'autant plus que l'enzyme CiSHT1, est aussi capable de produire des phénolamides *in vitro*, mais elle possède une très faible activité (Delporte et al. 2018). D'ailleurs, la sur-expression de *CiSHT1* *in planta* ne permet pas la production de phénolamides, ce qui pourrait être corrélé avec la faible activité de l'enzyme *in vitro*. Cependant, encore une fois, avec ces résultats, nous ne pouvons pas exclure l'implication potentielle de cette enzyme (CiSHT1) dans la biosynthèse de spermine totalement substituée. Au vu de ces résultats nous avons décidé de muter les deux gènes *CiSHT1* et *CiSHT2* grâce à la technologie CRISPR/Cas 9 en utilisant la transformation stable médiée par *A. rhizogenes* mise en place dans le chapitre 2. L'édition de ces gènes permettra, si une différence avec les lignées sauvages est détectée, de valider l'implication de ces gènes dans la voie de biosynthèse de la tetracoumaroyl spermine. De plus, les lignées mutantes pourraient aider à la compréhension du rôle de ces molécules. La mise en place et l'étude des mutants sont rapportés sous forme d'un article dans ce chapitre.

Fully substituted spermine-derived phenolamide synthesis requires sequential action of two BAHD acyltransferases in chicory flower buds

Guillaume Bernard¹, Julie Buges¹, Marianne Delporte¹, Roland Molinié², Alain Bouchereau³, Amandine Watrin¹, David Mathiron⁴, Jean-Louis Hilbert¹, Caroline Rambaud¹, David Gagneul^{1\$}

\$. Corresponding author. Email: david.gagneul@univ-lille.fr. Phone: +33 (0)3 20 43 40 21

1. Univ. Lille, INRA, ISA, Univ. Artois, Univ. Littoral Côte d'Opale, EA 7394-ICV-Institut Charles Viollette, F-59000 Lille, France.
2. Biologie des Plantes & Innovation (EA 3900 BIOP), Laboratoire de Phytotechnologie et Pharmacognosie, Faculté de Pharmacie, Université de Picardie Jules Verne, 80037 Amiens Cedex 1, France.
3. INRA, UMR 1349 Institut de Génétique, Environnement et Protection des Plantes, INRA, Agrocampus Ouest, Université de Rennes 1, Rennes, France.
4. Plateforme Analytique (PFA), Université de Picardie Jules Verne, 80039 Amiens Cedex 1, France.

1. Introduction

Plants accumulate a wide array of specialized metabolites. Historically defined as opposed to primary metabolites, these molecules have long been called secondary metabolites because they were often considered as evolutionary errors or waste products. Primary metabolites i.e. central metabolites referred to molecules that are required for the survival of the host and which production is conserved in all living organisms. Specialized metabolites include molecules that are dispensable for survival in an optimal environment but confer a great advantage in specific ecological niches. Therefore, their synthesis and accumulation are often distributed in a taxonomically restricted manner with few exceptions. These molecules are responsible for the tremendous adaptability of plants to their environment. Terrestrialization of the green lineage is often associated with the emergence of these new molecules¹. They are involved in various ecological functions notably plant interaction with its environment, for example, pollinator attraction, symbiotic microorganism attraction, defense against pathogens and herbivores or environmental stresses².

At the molecular level, the emergence of new metabolic pathways during plant evolution relies on gene duplication followed by random mutations. These genetic events can promote the appearance of new enzymes with novel catalytic properties and therefore new functions can arise. Furthermore, in contrast to primary metabolism, in which selection results in enzyme with high efficiency and strict substrate specificity, enzymes from specialized metabolism have usually broad substrate specificity but low catalytic specificity, which favors the emergence of new pathways during evolution. This enzymatic promiscuity is thought to play a central role in evolution of specialized metabolic enzymes and generation of chemical diversity³.

According to their chemical structures, specialized metabolites are ranked in three main classes: nitrogen containing compounds including alkaloids, terpenoids (isoprenoids) and phenolics. These latter are widely distributed in the plant kingdom and they contribute to all aspects of plant life. Their basic structure is a six-carbon aromatic ring with one or more hydroxyl groups attached. The aromatic part of most phenolic compounds is synthesized by the phenylpropanoid pathway⁴. The entry point of this path is the aromatic amino acid phenylalanine arising from the plastidic-located shikimate pathway. Deamination of phenylalanine by phenylalanine ammonia lyase

(PAL) gives rise to cinnamic acid, which in turn can be converted into basic phenylpropanoid backbones (C6-C3). All enzymes involved in this pathway are known for years⁵. These hydroxycinnamic acids are recruited to form majority of phenolic compounds including lignin, suberin and flavonoids. Hydroxycinnamic acid moiety, for example, coumaric acid, caffeic acid or ferulic acid are rarely accumulated under free forms but are often conjugated to acids, sugars or amines to form esters, glycosides or phenolamides, respectively. For instance, caffeic acid can be combined to tartaric acid to form caftaric acid (caffeoyleltartaric acid) as found in several members of the Asteraceae like chicory, in arachid and in horsetail^{6,7}, to quinic acid to form chlorogenic acid (caffeoylequinic acid) as found in tobacco, tomato or chicory^{6,8} or isochlorogenic acid (dicaffeoylquinic acid) as in coffee, sweet potato or chicory^{6,9-11}. Many polyamines are conjugated with hydroxycinnamic acids to form phenolamides also termed phenylamides or hydroxycinnamic acid amides¹². These polyamine-derived molecules are mostly known for their involvement in floral induction or as defense compounds¹³. Phenolic acids can be conjugated to either aromatic amines like tyramine, tryptamine, octopamine or anthranilate or aliphatic polyamines like putrescine, spermidine or spermine¹³.

To date, all the enzymes involved in *N*-acylation of aliphatic polyamines belong to the superfamily of the BAHD acyltransferases. For instance, agmatine coumaroyl transferase (ACT) catalyzes acylation of agmatine in barley and *Arabidopsis*^{14,15}, and hydroxycinnamoyl-CoA:putrescine acyltransferase (AT1) that of putrescine in *Nicotiana attenuata*¹⁶. The triamine spermidine was shown to be acylated by spermidine disinapoyl transferase (SDT) and spermidine dicoumaroyltransferase (SCT) in *Arabidopsis*¹⁷, by hydroxycinnamoyl-CoA:spermidine acyltransferase (DH29) in *N. attenuata*¹⁶ and by spermidine hydroxycinnamoyl transferase (SHT) in the tapetal cells of *Arabidopsis* and *Malus domestica*^{18,19}. We have previously identified and characterized, for the first time, two enzymes, called CiSHT1 and CiSHT2 due to their sequence homology with AtSHT and MdSHT, able to fully acylate spermine to produce tetracoumaroyl spermine in chicory²⁰. Thereafter, an enzyme, SrSpmHT, able to acylate the tetraamine to produced monohydroxycinnamoyl spermine, was identified in *Solanum richardii*²¹. The only enzymes characterized so far that have the capability to acylate secondary amino groups to synthesize fully substituted polyamines are SHTs (AtSHT, MdSHT, CiSHT1 and CiSHT2). *AtSHT* was shown to be specifically expressed in the tapetal cells of the anthers¹⁸. The accumulation of fully substituted

spermidine derivatives was shown to be restricted to the pollen coat of the *Arabidopsis* and apple tree^{18,19}. Therefore, accumulation of fully substituted spermidine was suggested as a marker of the pollen exine of Eudicotyledons¹⁹. The 2 proteins of chicory (*Cichorium intybus*) CiSHT1 and CiSHT2, were shown to be involved in unusual spermine derivative synthesis and accumulation in pollen coat²⁰. The fully substituted spermine derived phenolamide, tetracoumaroyl spermine, was shown to be the main phenolamide accumulated in the pollen cell wall of chicory in replacement of trihydroxycinnamoyl spermidine found in the Eudicotyledon species. We have extended this chemical particularity to the whole Asteraceae families and have shown that this molecule was absent in members of close families such as species belonging to the Calyceraceae, the Stylidiaceae and the Argophyllaceae. Both enzymes identified in chicory acylated either spermine or spermidine *in vitro* and all intermediates were detected, but, surprisingly the fully substituted polyamines were not the major products. Additionally, overexpression of *CiSHT2*, in chicory hairy roots that did not accumulate phenolamides naturally, or heterologous expression in *sht* *Arabidopsis* mutants devoid of fully substituted phenolamide accumulation in the pollen coat, led to the same inconsistency. Moreover heterologous and homologous expression of *CiSHT1* gave no metabolic modification in both plant systems. In order to know if both enzymes were required for the production of tetracoumaroyl spermine, we generated chicory plants silenced in the expression of *CiSHT1* or *CiSHT2* using the recently optimized CRISPR/Cas9 technology²² with *Agrobacterium rhizogenes*-mediated stable transformation. Indeed, the regenerated plants from hairy root lines don't require vernalization for flowering, and as flowering occurs as early as 1 week after regeneration, this is a real advantage for this study, in which both genes are expressed in flower buds. Interestingly both *Cisht1* and *Cisht2* mutant plants have a different flower bud phenolamide pattern than the wild type and data suggest that both enzymes act sequentially *in vivo* to synthesize tetracoumaroyl spermine. Heterologous expression of these genes in either *Nicotiana benthamiana* or *Saccharomyces cerevisiae* confirmed this assumption. This enzymatic system is able to acylate spermine or spermidine *in vivo* and *in vitro*, and further investigations demonstrated that free polyamine availability directs the flux toward either spermidine or spermine derivate production.

2. Results

2.1. Generation of *Cisht1* and *Cisht2* chicory mutants

In order to decipher the functional roles of *CiSHT1* and *CiSHT2* and try to understand the meaning of this apparent biochemical redundancy, single mutants for these genes have been generated. The genes *CiSHT1* and *CiSHT2* are 1356-bp and 1359-bp long, respectively and do not contain any intron²⁰. The target sequences were selected to induce breaks at position 550 and 663 for *CiSHT1* and at position 235 and 351 for *CiSHT2* (Figure 1A). Two binary vectors were prepared. The first named pYLCRISPR-sgRNA1-sgRNA2-CiSHT1 was designed to generate a double mutation in *CiSHT1* and the second termed pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 to generate a double mutation in both genes, thus expressing four sgRNAs. The schematic diagrams of the resulting vectors are shown in Figures 1B and C. Through

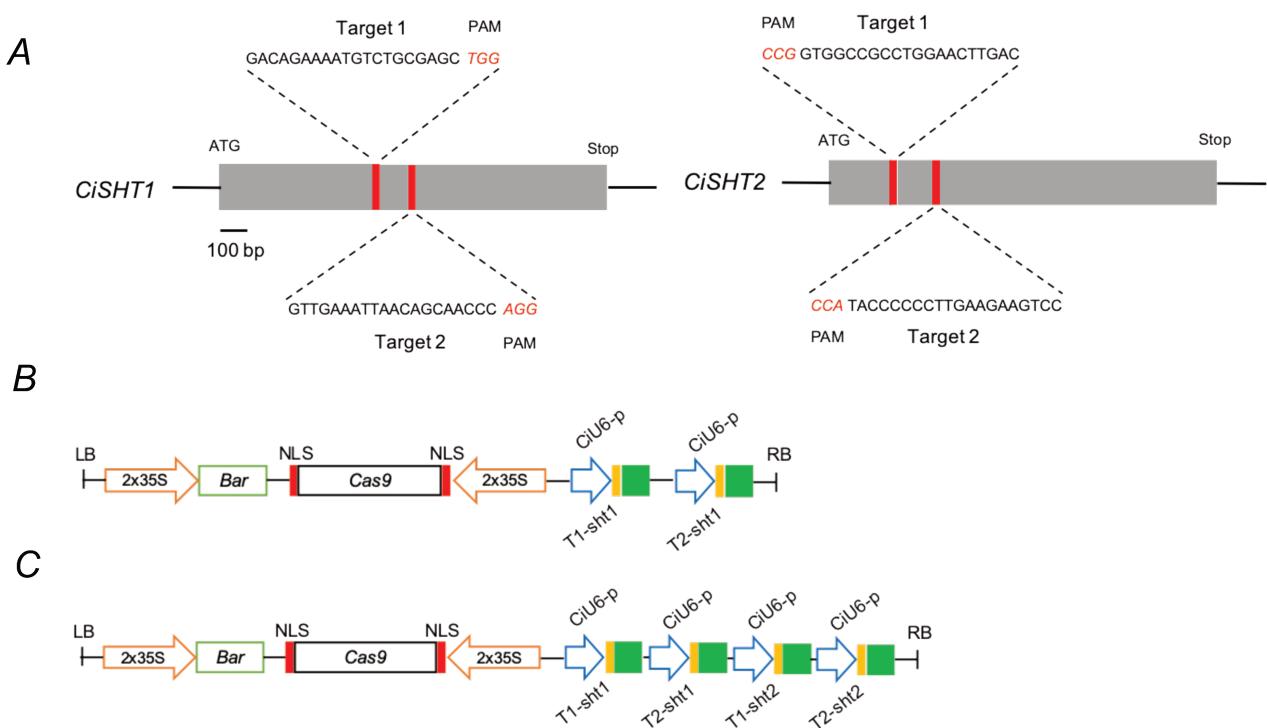


Figure 1: Schematic diagram of *CiSHT1* and *CiSHT2* gene structures and of vectors used to generate mutants. (A) Schematic diagram of *CiSHT1* and *CiSHT2* gene structures (grey rectangles) and target sequences (red lines). The translation initiation codon (ATG) and termination codon (stop) are shown. The target sequence is shown in capital letters, and the protospacer adjacent motif (PAM) sequence is marked in red. (B) and (C) Schematic diagram of the pYLCRISPR-sgRNA1-sgRNA2-CiSHT1 and pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 vectors, respectively. Bar, Phosphinothrin acetyltransferase; 35S, cauliflower mosaic virus promoter; NLS, nuclear localization sequence; Tnos, gene terminator; CiU6-1P, *C. intybus* U6-1p promoter.

A. rhizogenes-mediated transformation, 12 and 7 Basta-resistant hairy root lines were obtained with pYLCRISPR-sgRNA1-sgRNA2-CiSHT1 and pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 vectors, respectively. Genomic DNA was extracted from each resistant line in order to determine the presence of mutations in the target loci by High Resolution Melting-Curve (HRM) analysis (data not shown). PCR products from plants with divergent melting-curves compared to the wild type, were further sequenced in order to confirm mutations and to characterize them. As a result, three lines with biallelic mutations for each transformation event, have been identified (Figure 2). Most of the mutations occurred at a position about 3 pb upstream the PAM sequence, as described by²³. Editing types are either small deletion or insertions, but sometimes the inter-guide fragment was completely deleted as for the line SHT2 (23) (Figure 2B), which is consistent with what was shown in our previous study²². The target 1 of SHT1 was not effective in editing the target sequence with both binary vectors. The fact that some targets are less effective than others was already observed in apple²⁴. For hairy root lines transformed with the vector pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2, the target 2 designed to target

A Plant line		Target 2	PAM	mutation events			
WT		TCACAAAT	GTTGAAATTAAACAGCAACCC	AGG TCGAGAAATTGCGAACAAAGCTAATAAAAG			
SHT1 (8)	A1	TCACAAAT	GTTGAAATTAAACAGC - ACCCAGG	TCGAGAAATTGCGAACAAAGCTAATAAAAG -1			
SHT1 (8)	A2	TCACAAAT	GTTGAAATTAAACAGC - ACCCAGG	TCGAGAAATTGCGAACAAAGCTAATAAAAG -1			
SHT1 (17)	A1	TCACAAAT	GTTGAAATTAAACAGCAA	CCCC AGG TCGAGAAATTGCGAACAAAGCTAATAAAAG +1			
SHT1 (17)	A2	TCACAAAT	(+46)-TGAAATTAAACAGCA	GAGGACCO AGG TCGAGAAATTGCGAACAAAGCTAATAAAAG +50 -1			
SHT1 (19)	A1	TCACAAAT	GTTGAAATTAAACAGCAA	TCCC AGG TCGAGAAATTGCGAACAAAGCTAATAAAAG +1			
SHT1 (19)	A2	TCACAAAT	GTTGAAATTAAACAGCAA	TCCC AGG TCGAGAAATTGCGAACAAAGCTAATAAAAG +1			
B Plant line		PAM	Target 1	PAM	Target 2	Mutation events T1/T2	
WT		CG	CCGGTGGCCGCCTGGA	ACTTGAC	TG...CG	CCA TACCCCCCTGAAGAAGTCC CATTAT	
SHT2 (17)	A1	CG	CCG	GTGGCCGCCTGGA	ACTTGACTG...CG	CCA TAC TCCCCCTGAAGAAGTCC CATT	wt / +1
SHT2 (17)	A2	CG	CCG	GTGG -- GCCTGGA	ACTTGACTG...CG	CCA TAC ----- GAAGAAGTCC CATT	-2 / -7
SHT2 (22)	A1	CG	CCG	GTGGCCGCCTGGA	ACTTGACTG...CG	CCA TAC - CCCCTGAAGAAGTCC CATTAT	wt / -1
SHT2 (22)	A2	CG	CCG	GTGGCCGCCTGGA	ACTTGACTG...CG	CCA TAC CCCCCCTGAAGAAGTCC CATT	wt / +1
SHT2 (23)	A1	CG	CCG	----- ACTTGACTG...CG	CCA ----- AAGAAGTCC CATT	----- AAGAAGTCC CATT -13 / -10	
SHT2 (23)	A2	CG	CCG	GTG -----	CCCCCTGAAGAAGTCC CATT	----- CCCCCCTGAAGAAGTCC CATT -139	

Figure 2: Representative genotypes of *sht1* and *sht2* mutants. (A) mutation events induced on CiSHT1 after transformation with pYLCRISPR-sgRNA1-sgRNA2-CiSHT1. (B) mutation events induced on CiSHT2 after transformation with pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2. Target sequences are shown in blue letters, with PAM sequences in red. The change in the number of nucleotides is shown on the right.

SHT1 has generated biallelic mutation in only one line (Line 7) (supplementary Figure S3) and as the target designed to edit *SHT2* has been more efficient, three *sht2* mutants have been regenerated, but only one *sht1/sht2* double mutant (supplementary data S4). Ultimately, four independent *sht1* mutant lines arising from transformation with pYLCRISPR-sgRNA1-sgRNA2-CiSHT1, three independent *sht2* mutant lines arising from transformation with pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 and one *sht1/sht2* double mutant line (supplementary data S4) have been obtained. The frame shifts induced by the genome editing introduced an early stop on mRNA translation in all the mutated hairy root lines (supplementary Figure S5). The 8 lines generated by CRISPR-Cas9 genome editing were used for further experiments.

2.2. Metabolic analysis of genome edited plants

To investigate whether mutations generated by the CRISPR/Cas9 system had an impact on the phenolamide content of chicory flower buds, plants have been regenerated from mutated hairy root lines. Indeed, chicory has the capacity to regenerate plants after having been transformed with *A. rhizogenes*²⁵ and these plants do not need vernalization for flowering^{26,27}. Flowering occurs as early as 1 week after regeneration. Flower buds from *sht1* mutant lines, *sht2* mutant lines and from plants transformed with *A. rhizogenes* 15834 wild type strain (thereafter these lines are called WT) were collected. Flower buds were harvested at developmental stage 13-15 when the phenolamide content was shown to be the highest²⁰. As described earlier²⁰, tetracoumaroyl spermine was the main phenolamide in WT flower buds and tricoumaroyl spermidine was accumulated at a lower level (Figure 3). Tricoumaroyl spermine was also detected at low level. Identification of the different phenolamides was confirmed by mass spectrometry (not shown). In the flower buds of *sht2* mutant lines, no phenolamide was detected. The same metabolic profile was observed with the *sht1/sht2* double mutant line (not shown). In the flower buds of the *sht1* mutant lines, no fully substituted phenolamide i.e. tetracoumaroyl spermine or tricoumaroyl spermidine, was detected. Only tricoumaroyl spermine and dicoumaroyl spermine were accumulated. All together, these data suggest that both SHT1 and SHT2 are required to promote the synthesis of the fully substituted phenolamides in chicory

flower buds. SHT2 would catalyze the first acylation steps whereas SHT1 would fulfil the acylations to produce the fully substituted phenolamides.

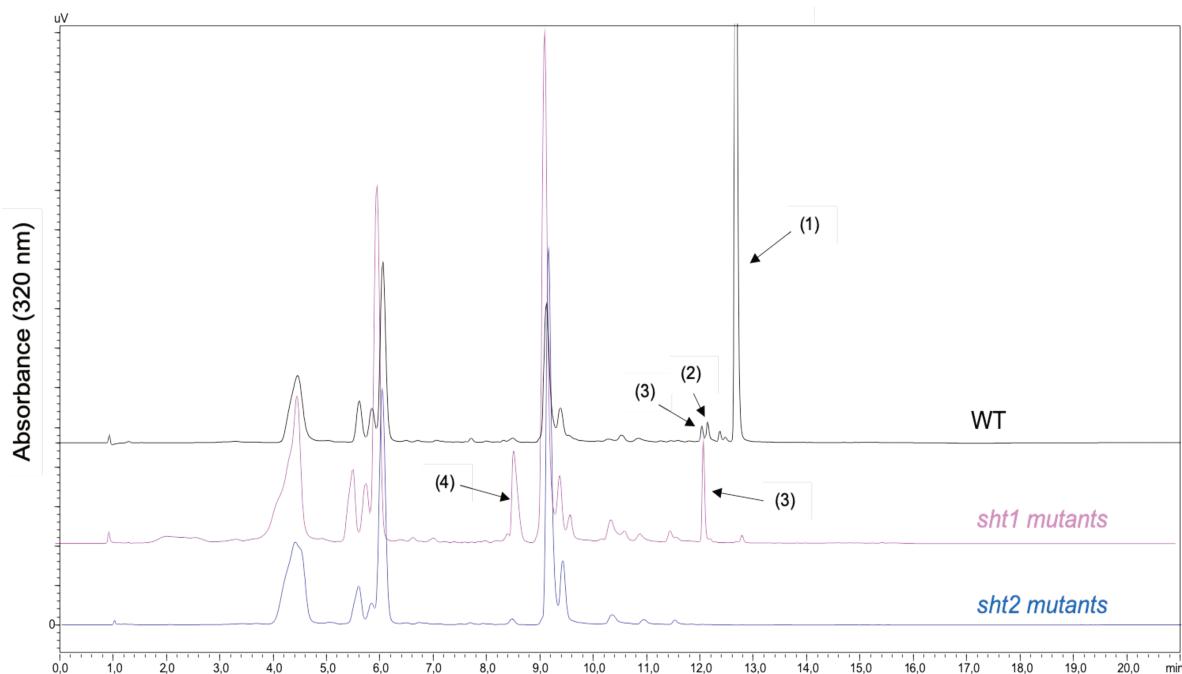


Figure 3: Stacked HPLC chromatograms of extracts obtained from WT, *sht1* or *sht2* flower buds at stage 13-15. The identity of the numbered peaks was confirmed by mass spectrometry: (1) tetracoumaroyl spermine; (2) tricoumaroyl spermidine, (3) tricoumaroyl spermine, (4) dicoumaroyl spermine.

2.3. Transient expression of CiSHT1 and CiSHT2 in tobacco leaves

The analyses of *sht1* and *sht2* mutants have shown that both enzymes are needed to synthesize tetracoumaroyl spermine. In order to confirm these results, *CiSHT1* and *CiSHT2* were transiently expressed in *N. benthamiana*. Leaves of *N. benthamiana* were infiltrated with *Agrobacterium tumefaciens* containing constructs of *CiSHT1* or *CiSHT2* or with a mixture of both strains. As a control, leaves were infiltrated with the void vector, pEAQ-HT. Four days after infiltration, leaves were collected, and the phenolic compounds were extracted before analysis. No spermine or spermidine derivatives were detected in tobacco leaves infiltrated with bacteria containing either the empty vector, pEAQ-HT or the pEAQ-HT-DEST1-SHT1 vector. Metabolic profiles were identical (Figure 4). In leaves transiently expressing SHT2, a mixture of phenolamides was detected. Notably, monocoumaroyl spermine, monocoumaroyl spermidine, dicoumaroyl spermine and dicoumaroyl spermidine were present. No fully

substituted phenolamide could be detected. The identities of all peaks were confirmed by mass spectrometry. Leaves agroinfiltrated with the vector pEAQ-HT-DEST1-SHT1 together with the vector pEAQ-HT-DEST1-SHT2 accumulated tricoumaroyl spermidine, tricoumaroyl spermine and tetracoumaroyl spermine. The main fully substituted phenolamide was tricoumaroyl spermidine. Interestingly, the peaks present in the chromatogram obtained with an extract of leaf infiltrated with pEAQ-HT-DEST1-SHT2 alone disappeared. These data are in agreement with those obtained with the chicory mutants and confirm that both enzymes could act consecutively to synthesize the fully substituted phenolamides.

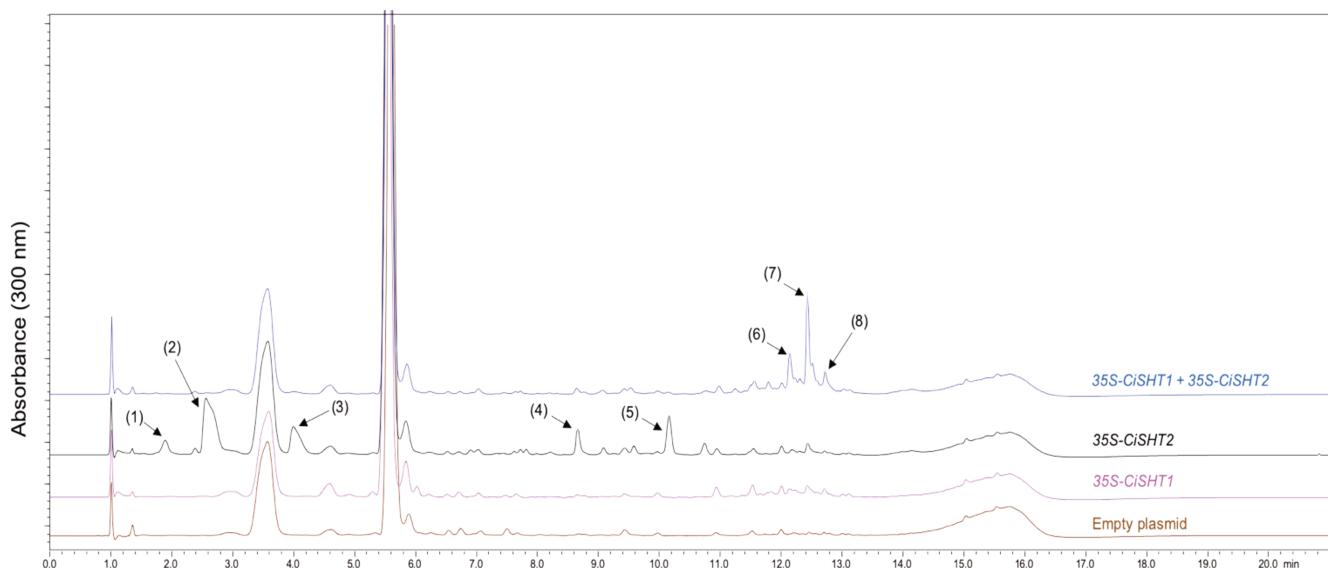


Figure 4: Stacked HPLC chromatograms of extracts obtained from agroinfiltrated tobacco leaves. Plants were infiltrated with either the empty plasmid (brown), CiSHT1 (35S-CiSHT1; pink), CiSHT2 (35S-CiSHT2; black), or CiSHT1 plus CiSHT2 (35S-CiSHT1+35S-CiSHT2; blue). The identity of the numbered peaks was confirmed by mass spectrometry: (1) monocoumaroyl spermine; (2)(3) monocoumaroyl spermidine; (4) dicoumaroyl spermine; (5) dicoumaroyl spermidine; (6) tricoumaroyl spermine; (7) tricoumaroyl spermidine; (8) tetracoumaroyl spermine.

2.4. Reconstitution of the tetracoumaroyl spermine biochemical pathway in *S. cerevisiae*

S. cerevisiae was used as an alternative system to prove the *in vivo* function of the chicory genes. Yeasts were transformed with *At4CL5* alone, with *At4CL5* and either *CiSHT1* or *CiSHT2*, with *At4CL5* and *CiSHT2* plus *CiSHT1*. The engineered yeasts were incubated on medium supplemented with coumarate and added or not with spermidine or spermine. After 48h, yeasts were pelleted, phenolic compounds

extracted and analyzed by LC-UV (Figure 5). As expected, no phenolamide could be detected when yeasts were transformed with *At4CL5* alone. The same results were obtained in yeasts coexpressing *At4CL5* and *CiSHT1* (not shown). In yeast transformed with the three genes and incubated with coumarate alone, accumulations of monocoumaroyl spermidine and tricoumaroyl spermidine were evidenced. This suggests that endogenous pools of spermidine were sufficient to promote spermidine-derived phenolamide production. When *At4CL5* and *CiSHT2* were coexpressed and yeasts incubated on medium supplemented with coumarate and spermine, accumulation of monocoumaroyl spermine, dicoumaroyl spermine and tricoumaroyl spermine occurred. In these conditions, yeasts accumulated also the spermidine derivatives. Interestingly, expression of *At4CL5*, *CiSHT2* and *CiSHT1* together, promotes accumulation of tetracoumaroyl spermine when coumarate and spermine were added in the medium. Moreover, in these conditions partially acylated spermine derivative accumulation decreased accordingly, and accumulation of spermidine

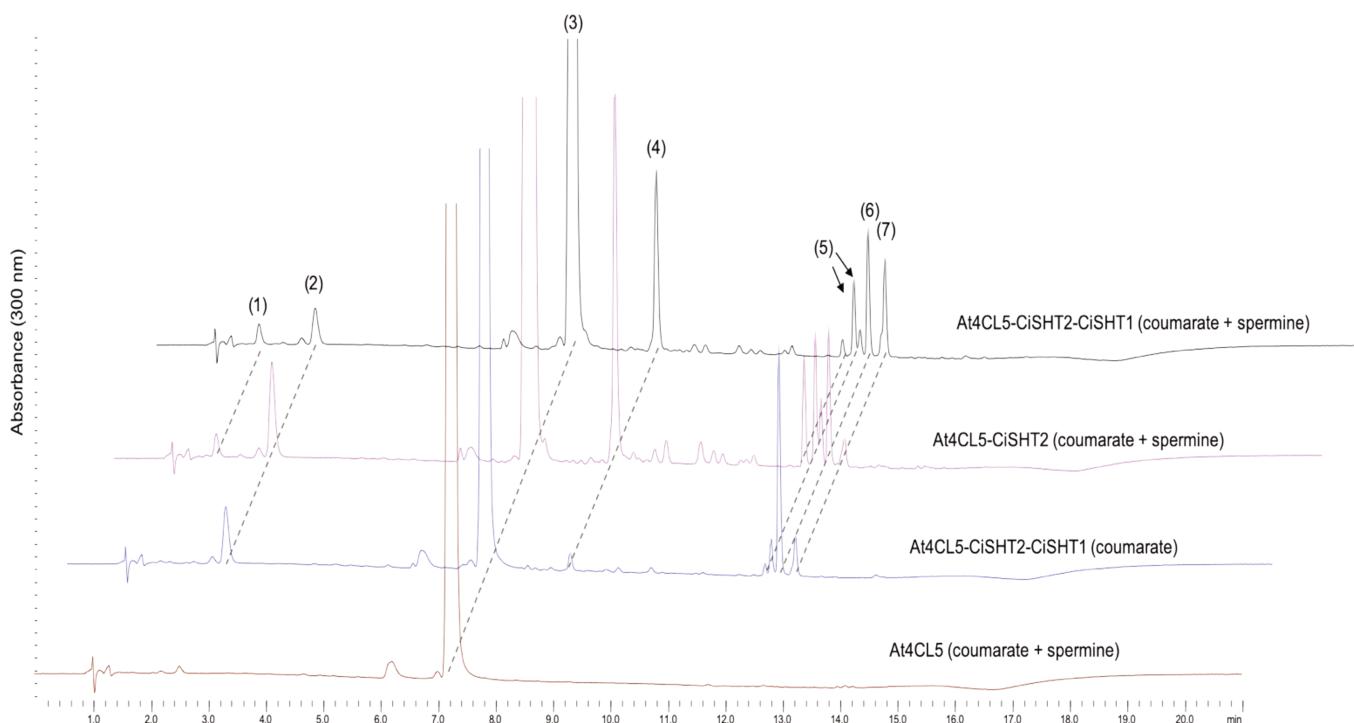


Figure 5: Stacked HPLC chromatograms of extracts from transgenic yeast expressing different combination of genes. The compounds added to the external medium are indicated in brackets. The identity of the numbered peaks was confirmed by mass spectrometry: (1) monocoumaroyl spermine; (2) monocoumaroyl spermidine; (3) coumarate; (4) dicoumaroyl spermine; (5) tricoumaroyl spermidine; (6) tricoumaroyl spermidine; (7) tetracoumaroyl spermine.

derivatives was also effective. Phenolamides were absent in the external medium in all experimental conditions. These data confirm our findings i.e. both enzymes are needed for tetracoumaroyl spermine accumulation *in vivo*. Surprisingly, in this system, tricoumaroyl spermidine production is efficient with only one gene (*CiSHT2*). This heterologous system could provide an interesting tool for bioproduction of phenolamides either derived from spermidine or from spermine.

2.5. Polyamine availability controls the nature of the final product

In experiments done with tobacco, yeast or *Arabidopsis*, the main products accumulated when transformed with the two chicory genes were spermidine derivatives, while the main compound found in chicory flower buds was tetracoumaroyl spermine. Moreover, *in vitro* CiSHTs were shown to be able to catalyze spermidine and spermine acylation without apparent substrate preferences²⁰. This prompted us to determine whether the pools of available polyamines in a specific tissue, and not the specificity of the enzymes, could be responsible for the production of a particular compound. For instance, it is known that in *Arabidopsis* flower buds, the ratio spermidine over spermine is clearly in favour of spermidine (around 5 times more spermidine than spermine)²⁸. We quantify free polyamines in chicory flower buds as well as in tobacco leaves. As expected, in tobacco leaves, spermidine level (1 nmol.g⁻¹ DW) was higher than that of spermine (0.12 nmol.g⁻¹ DW). The ratio spermidine/spermine was close to 8. In contrast in chicory flower buds, this ratio was close to 0.8. Spermidine and spermine levels were 0.5 and 0.7 nmol.g⁻¹ DW, respectively. Thus, the relative amount of aliphatic polyamines, and not the enzyme specificity, may direct the flux towards production of fully substituted spermine or spermidine.

2.6. The main CiSHT2 products is N^{5,N¹⁰}-dicoumaroyl spermine

In order to get more insights concerning the catalytic specificity of both enzymes, we purified one of the main product synthesized by SHT2 *in vitro* as well as accumulated in chicory hairy roots overexpressing *CiSHT2* or in chicory *sht1* mutants, a dicoumaroyl spermine isomer as was confirmed by mass spectrometry²⁰. This compound eluted at 8.5 min in our chromatographic conditions. This compound was purified from hairy root

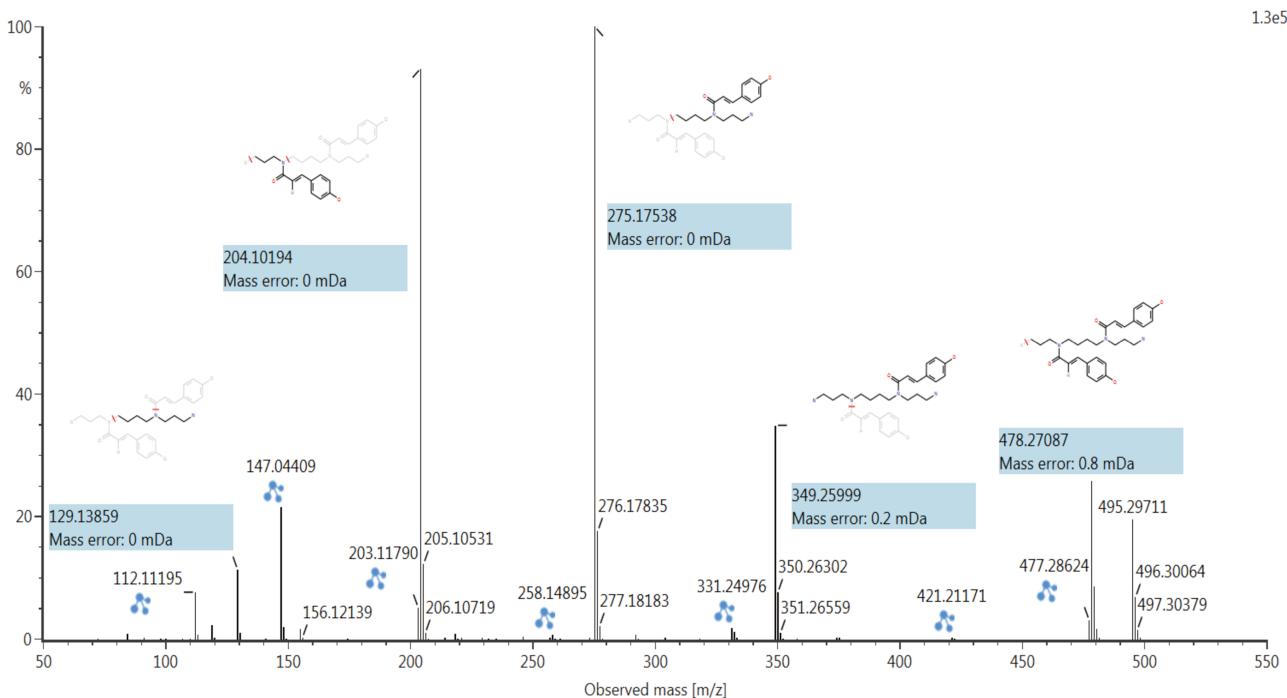


Figure 6: LC/MS/MS spectra of purified N⁵, N¹⁰-dicoumaroyl spermine. The position of the coumaroyl moieties was deduced from RMN data (see Table 1).

overexpressing *CiSHT2*. The purified compound gave the expected [M+H]⁺ ion at *m/z* of 495.2971 and the molecular formula was predicted to be C₂₈H₃₈N₄O₄. In the ESI-HRMS/MS spectra, typical fragments were observed at *m/z* 147.0441, 204.1019 and 275.1754 confirming the identity of this compound (Figure 6).

To determine the position of the coumaroyl moieties, the purified compound was analyzed by NMR spectroscopy. The characteristic NMR signals of the spermine derivative were observed (Table 1 and supplementary Figure S6). Correlation peaks were observed between the carbons of the carboxylic functions (δ 169.9 ppm) and the protons H-4 / H-11 (δ 3.57 ppm) and H-6 / H-9 (δ 3.59 ppm). These connectivity and the chemical displacement values of these protons suggest the grafting of coumaroyl units on secondary amines (Figure 7). A minor compound was also detected in our mixture. MS/MS and NMR analyses suggested that this molecule is an (E, Z) isomer of N⁵,N¹⁰-dicoumaroyl spermine. These data indicate that SHT2 efficiently catalyzes the acylation of secondary amines while SHT1 acylates primary amines of N⁵,N¹⁰-dicoumaroyl spermine

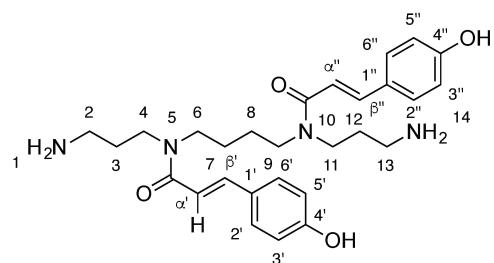


Figure 7: Chemical structure of N⁵, N¹⁰-dicoumaroyl spermine as deduced from NMR data.

Table 3.1: ^1H and ^{13}C -NMR data of $\text{N}^5, \text{N}^{11}$ -dicoumaroyl spermine in methanol-d₄.

Position	δ_{C}	δ_{H} (J in Hz)	HMBC
1	-	-	-
2	37.6	2.88 (t, J = 6.7)	3, 4
3	26.7	1.94 (t, J = 6.7)	2, 4
4	43.7	3.57 (m)	2, 3, 6
5	-	-	-
6	48.6	3.59 (m)	4, 8, CONH'
7	27.6	1.77 (m)	8, 9
8	27.6	1.77 (m)	4, 7
9	48.6	3.59 (m)	7, 11, CONH''
10	-	-	-
11	43.7	3.57 (m)	9, 12, 13
12	26.7	1.94 (t, J = 6.7)	11, 13
13	37.6	2.88 (t, J = 6.7)	11, 13
14	-	-	-
1'	127.4	-	3', 5', α'
2'	130.8	7.47 (d, J = 8.4)	4', 6', β'
3'	116.6	6.79 (d, J = 8.4)	1', 4', 5'
4'	160.8	-	
5'	116.6	6.79 (d, J = 8.4)	1', 3', 4'
6'	130.8	7.47 (d, J = 8.4)	2' 4', β'
α'	113.8	6.87 (d, J = 15.7)	1', CONH'
β'	145.0	7.59 (d, J = 15.7)	2', 6', α' , CONH'
CONH'	169.9	-	
1''	127.4	-	3'', 5'', α''
2''	130.8	7.47 (d, J = 8.4)	4'', 6'', β''
3''	116.6	6.79 (d, J = 8.4)	1'', 4'', 5''
4''	160.8	-	
5''	116.6	6.79 (d, J = 8.4)	1'', 3'', 4''
6''	130.8	7.47 (d, J = 8.4)	2'', 4'', β''
α''	113.8	6.87 (d, J = 15.7)	1'', CONH''
β''	145.0	7.59 (d, J = 15.7)	2'', 6'', α'' , CONH''
CONH''	169.9	-	

3. Discussion

The presence of fully substituted phenolamides in the outer surface of male gametophytes is a common feature in Angiosperms¹⁹. Trisubstituted spermidines were shown to be synthetized in the tapetum of anthers¹⁸. Two enzymes AtSHT and MdSHT belonging to the BAHD acyltransferases family were characterized and shown to be directly responsible for the synthesis of these compounds^{18,19}. The expression of these genes is restricted to the tapetal cells. Thereafter the synthesized compounds have to be exported out of the tapetal cells to be deposited on the pollen surface. The mechanism responsible for this deposition still need to be described. A recent study has shown that in the Asteraceae family the major phenolamides found in the pollen wall are fully substituted spermine conjugates, and that the presence of these unusual compounds is linked to the activity of two new members of the SHT family in chicory: CiSHT1 and CiSHT2²⁰. The biochemical and functional characterization of these two enzymes has evidenced their involvement in the biosynthesis of tetrahydroxycinnamoyl spermine. Nevertheless, *in vitro* and *in vivo* experiments suggested that there might be some particular features in the tapetum of chicory that act in favor of tetracoumaroyl spermine accumulation. Here, additional experiments were conducted to deepen our knowledge of this biochemical pathway in chicory.

Gene edition is a powerful tool to study gene function. Recently several new methods of genetic engineering using endonucleases have been developed. The CRISPR/Cas9 system has proven to be the most effective method for gene editing. In this study, in order to knock out CiSHT1 and CiSHT2, the protocol set up by Bernard et al (2019) was used. Because plants regenerated from hairy root lines don't need vernalization for flowering, results have been obtained in around 8 weeks using *A. rhizogenes* transformation. The metabolic analysis of mutant plants (figure 3) have highlighted that both CiSHT1 and CiSHT2 are required to promote tetracoumaroyl spermine biosynthesis in chicory. The *sht2* mutants did not produce phenolamides whereas the *sht1* mutants only produced dicoumaroyl and tricoumaroyl spermine but not the fully acylated spermine. These data indicate a role for CiSHT2 in the first acylation steps and for CiSHT1 in the last steps.

The results obtained with these mutants are in accordance with the experiments done with chicory hairy roots overexpressing either *CiSHT1* and *CiSHT2*²⁰. The requirement

of both enzymes to synthesize the fully substituted phenolamides was further confirmed using yeast and tobacco heterologous systems. Thus, no phenolamide was produced when expressing only *CiSHT1* while *CiSHT2* expression only promoted partially substituted accumulation. The co-expression of both genes led to the expected fully acylated phenolamide synthesis and accumulation. The main molecule produced *in vitro* or *in vivo* by *CiSHT2* is a dicoumaroyl spermine as determined by mass spectrometry. In order to decipher the reaction mechanism, this compound was purified and analyzed by NMR. It was identified as N⁵,N¹⁰-dicoumaroyl spermine. These data suggest that *CiSHT2* has a higher affinity for secondary amines of spermine and N⁵- or N¹⁰-monocoumaroyl spermine whereas *CiSHT1* prefers primary amines of di- or tri-coumaroyl spermine. The data obtained with spermine may be extended to spermidine. When dealing with *CiSHT2*, the main product obtained *in vivo* and *in vitro* was a dicoumaroyl spermidine. According to the structural homology of spermine and spermidine, it might be N⁵,N¹⁰-dicoumaroyl spermidine. Further experiments are needed to confirm this structure.

Taking together these data suggest that *CiSHT2* would promote dicoumaroyl spermine and dicoumaroyl spermidine synthesis, whereas *CiSHT1* would perform the last acylation steps toward fully substituted amide biosynthesis. More detailed *in vitro* biochemical experiments are necessary to assess these findings, but this task is hampered by the lack of commercial standards. A similar biochemical mechanism was suggested in *N. attenuata*¹⁶. An acyltransferase (DH29) was shown to catalyze the first acylation step in dicaffeoyl spermidine formation whereas a second enzyme named CV86 was suggested to act on monocaffeoyl spermidine to promote diacylated spermidine synthesis. Position of caffeoyl moieties was not further investigated but both enzymes belong to the BAHD acyltransferase superfamily.

Recently an enzyme able to acylate spermine has been identified in *S. richardii*, a wild relative of eggplant²¹. This enzyme named SrSpmHT (Spermine hydroxycinnamoyl transferase) was shown to use exclusively spermine as the acyl acceptor substrate and promote monohydroxycinnamoyl spermine synthesis. The product of the reaction was tentatively identified as N¹-coumaroyl spermine based on comparison with an authentic standard with LC-UV. Thus, SrSpmHT would acylate primary amine of spermine. SrSpmHT belongs to the BAHD family and is phylogenetically close to AtSDT and AtSCT. Like AtSDT and AtSCT, SrSpmHT lacks the FYGN motif, which was shown to be a specific motif in amino acid sequences of the enzymes belonging

to the SHT family such as AtSHT, MdSHT, CiSHT1 and CiSHT2²⁰. Taking together, our data suggest that in chicory fully acylated spermine synthesis relies on a two-step mechanism. Furthermore, they confirm that the enzymes belonging to the SHT subfamily of the BAHD acyltransferase superfamily are the only enzymes identified to date, that are capable of acylating secondary amino groups of polyamines.

CiSHT1 and CiSHT2 are both able to use spermine or spermidine as acyl acceptors *in vitro* and *in vivo*. Nonetheless in chicory flower buds, the main phenolamide is tetracoumaroyl spermine. With no apparent preference for the substrate, it was tempting to assume that the supply of free polyamines could determine which phenolamide would be the main accumulated molecule. It was shown that in *Arabidopsis* flower buds, the content of spermidine was 5 times higher than that of spermine²⁸. We have previously shown that in *sht* *Arabidopsis*, mutant expressing CiSHT2 in the tapetal cells, spermidine derivates were the main phenolamides²⁰. Furthermore, in tobacco leaves transiently expressing CiSHT1 and CiSHT2, tricoumaroyl spermidine was the main product (this study). The measurement of free polyamines has shown that the ratio spermidine/spermine were 8 and 0.8 in tobacco leaves and in chicory flower buds, respectively. This shows that substrate availability or substrate relative availability, and not the substrate specificity determines the identity of the final product. This observation suggests that in the Asteraceae there are at least two evolutionary events that led to the accumulation of these unusual spermine derivatives: modification of polyamine metabolism and neo-functionalization of BAHD-acyltransferase. An extended analysis of polyamine contents in species belonging to the Asteraceae family and close families may help to strengthen this hypothesis. It was previously shown that spermine conjugates were absent in Stylidiaceae, Argophyllaceae, and Calyceraceae²⁰. This additional analysis should provide an overview of the chronology of the evolutionary events mentioned above. It will be interesting to decipher if modification of polyamine metabolism drove neo-functionalization of SHTs, if SHTs neo-functionalization led to polyamine metabolism changes or if both events occurred independently. Ultimately whatever the chain of events, this has led to metabolic diversification that could be responsible for an evolutionary advantage that still deserves to be uncovered.

In any case, the functions of these fully acylated polyamines, deposited on the pollen surface, are still under debate, as is the benefit of accumulating spermine derivatives instead of spermidine derivatives.

Many hypotheses have been made about their biological roles^{18,19,29}. The most obvious function of these compounds is the protection of male gametophytes harboring sperm cells, i.e. the cells responsible for transmitting half part of the genetic material to offspring against UV radiation. It is possible that the advantage of switching from spermidine- to spermine- conjugates is the decoration of all the four nitrogen atoms of the spermine which increases the UV absorbance of a single molecule by about 30% compared to spermidine derivatives, without affecting the concentration of the metabolites. In *Arabidopsis*, the lack of phenolamides in the pollen coat was shown to have no effect on pollen viability or fertility¹⁸. Nevertheless, the impact of UV radiation on DNA has not been studied and it seems relevant to investigate UV damage at the molecular level. In this context, the analysis of pollen collected from the *Arabidopsis sht* mutant and the chicory *sht* double mutant compared to pollen collected from their WT counterparts could provide interesting clues. These compounds could also have some effect on plant interaction with pollinator³⁰. Insect feeding studies using pollen collected from the mutants or from the WT could be considered. Moreover, the conjugates of hydroxycinnamic acids with polyamines have been proposed to have numerous biological activities such as, antifungal, antimicrobial and anti-viral effects³¹⁻³³. The evaluation of biological activity is a difficult task due to the availability of pure compounds. The elucidation of the biochemical pathway of tetracoumaroyl spermine and its reconstruction in the heterologous system *S. cerevisiae*, implemented in this work, is an important step toward providing sufficient amounts of molecules to support these biological tests. Nevertheless, this bioproduction system still needs to be optimized to increase the production of either spermidine or spermine derivatives. Indeed, main compounds produced were not the fully substituted amines. The expected compounds were retained in the yeast intracellular medium and catalytic constraints, such as feedback inhibition, may not favor the synthesis of the fully acylated compounds. The engineering of a strain secreting the final products in the culture media should be considered. Finally, the catalytic versatility of the enzymatic system characterized in this study may be used to synthetize related bioactive compounds such as the kukoamines widely studied, among other, that possess numerous pharmacological effects, such as antihypertensive, anti-inflammatory, anti-analgesic, anti-sepsis, autoimmune-enhancing and neuroprotective activities³⁴⁻³⁶. Feeding an optimized strain with dihydrocaffeic acid could result to kukoamines

production and maybe tetradihydrocaffeic spermine, a compound still never synthetized and biologically assessed.

4. Materials and methods

4.1. Plant material

Cichorium intybus (L.) line 'clone 17' (Florimond Desprez, Cappelle-en-Pévèle, France) was used in this study. Hairy root lines overexpressing *CiSHT2* as previously described and plants regenerated from them were also used in this study²⁰. The *Arabidopsis* ecotype Columbia 0 (Col-0), the *sht* mutant line SALK_055511, the *sht* mutant line SALK_055511 overexpressing either *CiSHT1* or *CiSHT2* were used for complementation and overexpression experiments^{18,20}. Plants were grown in a greenhouse under a 16/8h light/dark cycle. *N. benthamiana* plants were grown in the same conditions.

4.2. Design of sgRNA, construction of CRISPR/Cas9 vectors and *A. rhizogenes* transformation

To delete *CiSHT1* (GenBank accession MG457243.1) and *CiSHT2* (GenBank accession MG457244.1) genes, two guides RNAs have been designed for each gene according to their location in the gene and their GC content using the software crispor.tefor.net³⁷. Target were chemically synthesized by SIGMA-ALDRICH (Gillingham, UK)

Primary vectors pKanCiU6-1p-sgRNA1-SHT1, pKanCiU6-1p-sgRNA2-SHT1, pKanCiU6-1p-sgRNA1-SHT2 and pKanCiU6-1p-sgRNA1-SHT2 containing the cassette "CiU6-1p-Guide-sgRNAs scaffold were constructed as previously described²². Binary vector pLYCRISPR-sgRNA1-sgRNA2-CiSHT1 harboring two sgRNAs and CAS9 was constructed as described in²². To construct binary vectors expressing four sgRNAs (two sgRNAs targeting *SHT1* and two sgRNAs targeting *SHT2*), PCR using proofreading polymerase (PrimeStar HS DNA polymerase; TAKARA, Kasatu, Japan) with primary vectors as templates were run using the following primer pairs: GG1-F and GG1-R with pKanCiU6-1p-sgRNA1-SHT1, GG2-F and GG3-R with pKanCiU6-1p-sgRNA2-SHT1, GG3-F and GG4-R with pKanCiU6-1p-sgRNA1-SHT2, GG4-F and

GG2-R with pKanCiU6-1p-sgRNA2-SHT2. The primer sequences are listed in Supplementary Table S1. The *Bsa*/ restriction site and the 4 nucleotides recombination sites for Golden Gate restriction/ligation method was added in the 5' extremity of each primer. Then PCR products were purified with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) and were assembled with Golden Gate restriction/ligation method in pYLCRISPR/Cas9P35S-B as described by Ma and Liu³⁸ to generate pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 and pYLCRISPR-sgRNA1-sgRNA2-CiSHT1.

The *Agrobacterium rhizogenes* strain 15834 (kindly provided by Marc Buée, INRA, Nancy, France) was independently transformed by electroporation with the binary vectors pYLCRISPR-sgRNA1-sgRNA2-CiSHT1 or pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 as described in Nagel et al³⁹.

4.3. Hairy root culture and selection

Hairy root induction was performed on leaves of 'clone 17' using *A. rhizogenes* strains (wild-type or transformed with either pYLCRISPR-sgRNA1-sgRNA2-CiSHT1 or pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2) as described in Bernard et al.²² and the infected leaves were placed on Murashige and Skoog Hairy Root (MSHR) solid medium (macro and micro-elements, Fe-Edta, vitamins from⁴⁰, added with 30 g.L⁻¹ sucrose, 6 g.L⁻¹ agar, pH 5.6) in Petri dishes, and grown at 25°C ± 2°C under 16 h light (37.6 ± 10.1 µmol.m⁻².s⁻¹) and at 20°C ± 2°C under 8 h dark cycles. Individual putative transgenic hairy roots of approximately 3 cm long, emerging from transformed leaf explants transformed with *A. rhizogenes* harboring pYLCRISPR-sgRNA1-sgRNA2-CiSHT1 or pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2 were excised and immediately transferred on Petri dishes with MSHR/2 solid medium (0.5 × macro and micro-elements, 0.5 × Fe-Edta, 0.5 × vitamins from⁴⁰, 30 g.L⁻¹ sucrose, 6 g.L⁻¹ agar, pH 5.6) containing ampicillin 300 mg.L⁻¹ to remove Agrobacteria and glucosinate-ammonium (10 mg.L⁻¹) to select transformants.

4.4. Genomic DNA extraction and High-Resolution Melting Temperature analysis

Genomic DNA from hairy roots was extracted using the NucleoSpin Plant II kit (Macherey-Nagel, Düren, Germany) according to the manufacturer instructions. PCR was performed using Type-it HRM™ PCR (QIAGEN) according to the manufacturer instructions. Twenty ng genomic DNA were used as template for PCR. The primers used for the detection of gene editing events are listed in Supplementary Table 1 and were designed using Primer3 and Netprimer softwares. The PCR products were analyzed by High-Resolution Melting (HRM) using Rotor-Gene 6000 cycler. Data were analyzed with Rotor Gene Q Serie software. Edited DNA fragments were amplified with primer (SHT1seq-fwr and SHT1seq-Rev) for *sht1* and with (SHT2seq-fwr and SHT2seq-rev) for *sht2* (Supplementary data S1). PCR products were analyzed by Sanger sequencing method and the chromatograms were analyzed with CodonCodeAligner software⁴¹.

4.5. Phenolic acid extraction

Lyophilized plant material was powdered, and 25 mg were resuspended in 1 mL of 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 (14 000 g, 4 °C, 10 min) and passed through a 0.45-µm filters.

4.6. Metabolite analyses by HPLC/DAD

Metabolite analysis was carried out using 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 × 4.6 mm Kinetex 2.6-µm 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⁻¹ and the oven temperature was set at 45 °C. The HPLC conditions were as follow: start at 5% solvent B, 6 min gradient to 20% solvent B followed by 1 min

isocratic 20% solvent B, and 1 min gradient to 25% solvent B were applied before returning to 5% solvent B and 7 min re-equilibration. Ten μ L of the extracts were injected on the column.

4.7. Extraction and analysis of free polyamines

Polyamines were extracted and analyzed as described in Duhazé et al. (2002)⁴²

4.8. Transient expression in *Nicotiana benthamiana*

Previously reported entry vectors pDONR207_SHT1 and pDONR207_SHT2 were used to generate expression vector for transient expression in tobacco²⁰. The open reading frames of *CiSHT1* and *CiSHT2* were introduced in pEAQ-HT-DEST1 vector by LR recombination to generate pEAQ-HT-DEST1-SHT1 and pEAQ-HT-DEST1-SHT2⁴³. The empty pEAQ-HT vector was used as a negative control⁴³. These three vectors were independently introduced in *A. tumefaciens* GV2260 by electroporation. Agrobacteria were grown and prepared for agroinfiltration as described in Legrand et al. (2016)⁶. For transformation, agrobacteria strains were prepared alone or mixed to reach a final OD₆₀₀ of 0.8 (equal concentration of each strain). Tobacco was then infiltrated with strains harboring pEAQ-HT as a control, with pEAQ-HT-DEST1-SHT1 alone, with pEAQ-HT-DEST1-SHT2 alone or with a mixture of strains containing pEAQ-HT-DEST1-SHT1 or pEAQ-HT-DEST1-SHT2. Four 4 to 6 weeks-old-plants per combination were used. After 4 days, the infiltrated leaves were collected, immediately frozen in liquid nitrogen and stored at -80°C until use.

4.9. Stable expression in *Saccharomyces cerevisiae*

The vectors pDRf1-4CL5-GW, pDRf1-4CL5 and pRS423-RjTAL were provided by Dominique Loqué^{44,45}. pRS423-GW was generated by BP reaction with pRS423-RJTAL and pDONR221. pDONR-SHT1 and pDONR-SHT2 were used to generate pDRf1-4CL5-SHT2, pDRf1-4CL5-SHT1 and pRS423-SHT1 by LR recombination with pDRF1-4CL5-GW or pRS423-GW. The *S. cerevisiae* pad1 knockout (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Δpad1*, ATCC 4005833)⁴⁶ were transformed using the freeze-thaw method⁴⁷. Yeasts were selected on solid medium containing yeast nitrogen base

(YNB) without amino acids supplemented with 2% glucose and 1X dropout-uracil, 1X dropout-histidine or 1X dropout-uracil-histidine. Overnight cultures from 4 independent single colonies of recombinant yeast harbouring pDRf1-4CL5, pDRf1-4CL5-SHT2, pDRf1-4CL5-SHT1 or pDRf1-4CL5-SHT2 plus pRS423-SHT1 constructs were mixed and grown in 1X YNB medium supplemented with 2% glucose and the required yeast synthetic drop-out medium supplement. These cultures (4 independent colonies per construct) were used to inoculate 2 mL of fresh medium to reach an OD₆₀₀ of 0.15 and incubate at 30°C at 200 rpm. Substrates were added 5 h after inoculation. Coumarate, spermidine and spermine were added at a final concentration of 1 mM. The cultures were further grown for 24 h at 200 rpm at 30°C in the presence or absence of the different precursors. For metabolite measurement, supernatant and pellet were separately collected by centrifugation (20000 g, 5 min, 4°C). The supernatants were passed through a 0.45-µm filter prior to HPLC-UV analysis. The pellets were ground with glass beads in a mixture of methanol/water/acetic acid (75/23/2, v/v/v). Homogenates were clarified by centrifugation (14 000 g, 4 °C, 10 min) and passed through a 0.45-µm filter before HPLC-UV analysis.

4.10. Dicoumaroyl spermine purification

Hairy roots overexpressing *CiSHT2* generated in our previous study were used²⁰. These lines were shown to accumulate high amounts of dicoumaroyl spermine when fed with exogenous spermine. Five g of lyophilized material, cultivated as described, were powdered and extracted with 200 mL of a methanol/water/acetic acid mixture (75/23/2, v/v/v) for 2h at 4 °C in the dark under agitation. Homogenate was then clarified by centrifugation (5 min at 6000 rpm) and the supernatant was passed through a 0.45-µm filter. The pellet was re-extracted a second time following the same procedure. Then, the filtrates were pooled and washed with ethyl-acetate (v/v) and the aqueous phase was collected and passed through a cation exchanger column (50 mg, CM-Sephadex C-50, PHARMACIA, Uppsala, Sweden) equilibrated with water. The resin was washed with 2L of water and eluted with 20 ml of a mixture of methanol/ 8M acetic acid (v/v). The eluted fraction was concentrated under rotary evaporation to get a final volume of 1 mL. Dicoumaroyl spermine was purified after separation by HPLC using the protocol described above. The collected fractions containing dicoumaroyl

spermine were then pooled and concentrated under rotary evaporation to reach a final volume of 2ml. This concentrated fraction was further purified using a purify flash 4250₂₅₀ (Interchim). Purification was performed on a 250x21 mm Interchrom 2mm preparative column (RP18). The chromatographic separation was performed using water (solvent A) and methanol (solvent B), both acidified with 0.1% formic acid. The solvents were delivered at a flow rate of 22 ml min⁻¹. The column was equilibrated with 20% solvent B. The separation conditions were as followed: start at 20% solvent B, 10 min gradient to 70% solvent B followed by 2 min gradient to 100%, and 3 min isocratic 100% solvent B, then a 2 min gradient to return to 20% solvent B and 5 min of isocratic re-equilibration at 20% solvent B. The collected fractions were concentrated under rotary evaporation. A 1 ml concentrated fraction was frozen and lyophilized in order to obtain a dried powder. Around 2.5 mg of dicoumaroyl spermine were purified following this procedure. The purified compound was analyzed by NMR and LC/MS/MS.

4.11. ESI-HRMS and NMR Analysis

Electrospray high resolution mass spectrometry (ESI-HRMS) analysis, was performed according to the procedure described in our previous work ²⁰. For nuclear magnetic resonance (NMR) analysis, pure compounds were dissolved in 0.75 mL of solvent (methanol-d₄ or DMSO-d₆). Six hundred μ L of the solution were then transferred into a 5mm NMR tube. NMR spectra were acquired at 300 K on a Bruker Avance III 600 spectrometer (600.13 MHz for proton frequencies, Wissembourg, France) equipped with a z-gradient inverse probe head (TXI, 5-mm tube). The TOPSPIN (V3.2, Bruker) software was used. The 1D proton spectra were acquired using 32 scans of 128 K data points, using spectral widths of 8403 Hz. The 1D carbon spectra were acquired using 24 K scans of 64 K data points, using spectral widths of 37,878 Hz. The 2D COSY and 2D TOCSY spectra were acquired using 8 scans per 256 increments that were collected into 2 K data points, using spectral widths of 12019 Hz in both dimensions. For the TOCSY, a mixing time of 100 ms were employed. The NUS based 2D HSQC spectra were acquired using 16 scans per 8 K increments that were collected into 4K data points, using spectral widths of 12019 Hz in F2 and 26412 Hz in F1. The NUS based 2D HMBC spectra were acquired using 64 scans per 8 K increments that were collected into 4K data points, using spectral widths of 12019 Hz in F2 and 37732 Hz in

F1. The number of NUS sampling points was 256 complex points (3.125% sampling density of 8 K points⁴⁸). All non-zero filled obtained spectra were manually phased and baseline-corrected, calibrated with solvent signal (¹H (3.31) and ¹³C (49.0) for methanol-d₄⁴⁹).

Funding

GB and JB were supported by a doctoral fellowship from University of Lille and the Région Hauts-de-France, while MD was supported by a doctoral fellowship from the doctoral school 104 SMRE,. This research was funded by an Albiotech grant (2016-2020) obtained from the CPER/FEDER program.

Acknowledgments

Authors would like to thank Professor George Iomonossof (John Innes Center) for providing the pEAQ vectors (pEAQ-HT and pEAQ-DEST1)

Supplementary materials

CiSHT1

ATGAGGTTCCAACAAGACTTCTCCATGGTAAACCAGCTGAACCAACATGGTCGGTAAGCTAGCCTATCGAACT
 TGATCAAACGGTGTCAACACTCATGTTCCAACCATTACTCTATGCACAATCCCCGAGGACTGGGTCACTGTTCTAC
 ATACCCCTAAAAACTCCCTAAGTTGACCTAGTTCTATCCATTGGCTGGCCGGCTGTCTTCTATTGCTGGAGGC
 CGTTGGAGCTTGATTGCAATGCGGCCGGCGTGCATTCACAGAACGCTACGCTGACAAGAAACTACTGATCTGGACAG
 TCTGTTACCATGCCCATGTATCACCAACTTATCCTCAATAGATTACAAAGTACACTTCTGAAGAAATCCATTGT
 TAGTTTGCAAGTCACAAGGTTCATTTGTGGAGGGTTAGCCTGGTAAACATATCGCATGTTGCTGACGGACAA
Seq-SHT1-F →
 AGTGCCTTCATTTGTGACCGAATGGGGCTGGATTTCACAGAGGTGAGTTATTAGAGTTACCACTTATCTT**GACAGAAA**
HRM-T1-SHT1-F →
TGGGGATCCATCAAGGGCAAGCTGAGGTTGAACATGCAGAGTTCGATCCTCCACCAGTCCTTTAG → **HRM-T2-SHT1-F**
 ATCAATCTGGAATGAAAGAGAATGTGAGCAGGAGACCGCGGTACAAT**GTTGAAATTAAACAGCAACCCAGTCGAGAAA**
 TTGCGAAACAAAGCTAATAAAAGTCAGGAGAGTGGAGATGAGCCGTGGATTACACGGTATGAGGGTGTAACTGCACACAT → **HRM-T1-SHT1-R**
 ATGGAGAAAGTCATGCAAAGCGCTAATCACAAACCAGAGCAGCAGCAACTGCCTGGTATTGATATTGAAAGTA → **HRM-T2-SHT1-R**
Seq-SHT1-F →
 AGATGACGCCACATTGCAAGAAAGTACTTTGAAACCGAATTGCAATGCGATTGCCATCGGCTGCTCCGGTGGAGATA
 ATTCAAAACCTTAGGATACGTTCCGGTAAACTAAGGGAGGTGATTGAAAAGGTTGACGATGGCTACGTGCACCGT
 GTGGACTTCTGAAGAGTCACAAGATTACAGTTGGAGAGTGGCAACCAACAAGTAACAGTGTGAGCATT
 ACGGCAATCCAATCTGGGTGATAAGTTGTTAACACTACCAATCTATGGTCCGGATTGGATGGGAAAGGAGATT
 CATATGGGCCCGAACACATGATTGGATGGTATTCTTGTATTCATGGCATAGATGGAGATGGATCGTGTGATAGT
 TGCATTATGTTACAGTCAGGGCATATGAAAGATTCCAGAAGGTTCTACCAAGATATTGATATGAAATGT**TGA**

Supplementary figure S1 : *CiSHT1* DNA sequence, sgRNA target sites and the primer sites for sequencing and HRM analysis. Blue nucleotides indicate the target sites, bold red nucleotides indicate the PAM sequences. Nucleotides highlighted in blue indicate start and stop codons, respectively. Dark arrows represent primers used for sequencing. Green arrows represent primers for HRM analysis of target 1. Yellow arrows represent primers used for HRM analysis of target 2.

CiSHT2

ATGAGAGTCACCGTAAATCTCTCCATGATCAAACCCAGCCGAACCAACATGGCTGGTAGGCTCCACTATCCGAACCT
 AGACCAACAGGAATGAACACACATGTTCAACCATAATTCTACACAAAACCTCTCAGGACTGGACACTGTCCTGC
Seq-SHT2-F →
 AAACCTTAAAGCTCCCTCACCACCACTTGGTCTATTCTATCCTTAGCCAGCCGGTATCTTCCATCG**CCGGTGGC**
HRM-T1-SHT2-F →
CGCCTGGAACCTGACTGCAATGCGGATGGTTCAGTTCATCGAAGCTTACGCCGACATGAAACTAGTTGACTTAGATGA
HRM-T1-SHT2-R ← **HRM-T2-SHT2-F** →
 TTCATGCAATCTCAATATTTCACCACTTATCCTCCGTAGATTATCG**CCATACCCCCCTTGAAAGAAGTCCCATTAT**
 TGCTTTGCAAGCTACAAATTGTTGAGGTTGGCTTTAGGGTTTGATATCACATGTTGTCGACGGTCAA
HRM-T2-SHT2-R ← **Seq-SHT2-R** →
 AGCGCACTTCATTTCTGCTGAATGGGCTAGAGCTTGTGCGGTGAGCCGCTGGATTCAACCACCTCATATCGATAGGAA
 ATGTTTGACCGCTGGTGAGGCCACCGACCACATGTTCAAGCATCGAACAAACATGTCAGTTTCGTCCTCCACCGATCC
 TTTAGGTCAATCAAGCAATGAAAGTACCGCAGGAAACCACTGAGCAATGCTGAAACTAACAGAAACTCTAGTC
 GAAAAACTCGAAATAAGCTAATAAAAGCGGAAACTGAAACCGACCGTGGTTTACACGGTATGAGGGTTGAACTGC
 ACATGCATGGCAACAGCATGCAAAGCTCGTAATCTAGAACAGCAGGCCACTGCACTGGAGTAGGCATATGTATTGATGTT
 GAACCGAAATGAAGCCGGTTGCGGAAAGTACTTTGGAAACGCCATTAGACGGTATGAGGGTTGAACTGAGTATGTGAA
 GAGTTACTTCGGAGACATTAGGGTATGTTGACTAAGATAAGGAACCGCAGTGGGAGACTTTGATTCTACGTGGGAGGGAGGTGGATC
 CGTTTACGGGAATCGAATCTGGAGTGGTAAGCTGGTTAACATTGCCATGTACGGTGTGATTGGGAGGGTAAG
 GAGGAATTACAGGTCCAGGAACCGATGATGATGGAGACTTTGATTCTACGTGGGAGAGGGAGGTGGATC
 TCTGGTTGTTGAGACATGGAGGATTCAAAGGTGTCTATCAAAGCATTGAAAGAT**TGA**

Supplementary figure S2: *CiSHT2* DNA sequence, sgRNA target sites and the primer sites for sequencing and HRM analysis. Blue nucleotides indicate the target sites, bold red nucleotides indicate the PAM sequences. Nucleotides highlighted in blue indicate start and stop codons, respectively. Dark arrows represent primers used for sequencing. Green arrows represent primers for HRM analysis of target 1. Yellow arrows represent primers used for HRM analysis of target 2.

ATCGACAGAAAATGTCTGCGAGC	fwr	Target 1	CiSHT1	T1-SHT1-F
AAACGCTCGCAGACATTTCTGT	rev	Target 1	CiSHT1	T1-SHT1-R
ATCGTTGAAATTAAACAGCAACCC	fwr	Target 2	CiSHT1	T2-SHT1-F
AAACGGGTTGCTGTTAATTCAA	rev	Target 2	CiSHT1	T2-SHT1-R
ATCGTCAAGTTCCAGGCCGAC	fwr	Target 1	CiSHT2	T1-SHT2-F
AAACGTGGCCGCCTGGAACCTGA	rev	Target 1	CiSHT2	T1-SHT2-R
ATCGGACTTCTCAAGGGGGTA	fwr	Target 2	CiSHT2	T2-SHT2-F
AAAATACCCCCCTTGAAGAAGTC	rev	Target 2	CiSHT2	T2-SHT2-R
ATATATGGTCTCACTCGAAAGAACCAACCTGTTTCATAGC	fwr	Golden Gate		GG1-F
ATTATTGGTCTCACCTAAAAAAAGCACCAGACTCGGTG	rev	Golden Gate		GG1-R
ATATATGGTCTCAAGGTAAAGAACCAACCTGTTTCATAGC	fwr	Golden Gate		GG2-F
ATTATTGGTCTCACAGA- AAAAAAAGCACCAGACTCGGTG	rev	Golden Gate		GG3-R
ATATATGGTCTCATCTGAAAGAACCAACCTGTTTCATAGC	fwr	Golden Gate		GG3-F
ATTATTGGTCTCAGACTAAAAAAAGCACCAGACTCGGTG	rev	Golden Gate		GG4-R
ATATATGGTCTCAAGTCAAAGAACCAACCTGTTTCATAGC	fwr	Golden Gate		GG4-F
ATTATTGGTCTCAACCGAAAAAAAGCACCAGACTCGGTG	rev	Golden Gate		GG2-R
CGAATGGGCTCGGATTTAC	fwr	HRM-Target1	CiSHT1	HRM-T1-SHT1-F
TCGACCTGGGTTGCTGTTAA	rev	HRM-Target1	CiSHT1	HRM-T1-SHT1-R
CGATCCTCCACCAGTCCTT	fwr	HRM-Target2	CiSHT1	HRM-T2-SHT1-F
AATCCACGGCTCATCTCACT	rev	HRM-Target2	CiSHT1	HRM-T2-SHT1-R
CCTCACCAACCACACTTGTTC	fwr	HRM-Target1	CiSHT2	HRM-T1-SHT2-F
ACTGAACACCATCCGCATTG	rev	HRM-Target1	CiSHT2	HRM-T1-SHT2-R
GCTTACGCCGACATGAAACT	fwr	HRM-Target2	CiSHT2	HRM-T2-SHT2-F
AGACCAACCTCCACACACAA	rev	HRM-Target2	CiSHT2	HRM-T2-SHT2-R
AAGTCACAAGGTTCATTTGTGGAG	fwr	sequencing	CiSHT1	Seq-SHT1-F
TATCAATGCAAATAGCCAAGGCAG	rev	sequencing	CiSHT1	Seq-SHT1-R
AGACCAACAGGAATGAACAC	fwr	sequencing	CiSHT2	Seq-SHT2-F
ATCGATATGAGGTGGTGAATCC	rev	sequencing	CiSHT2	Seq-SHT2-R

Supplementary table T1: Sequences of primers used in this study

		Target 1	PAM
SHT1 WT		<u>ATGTTGAAATTAAACAGCAACCC</u> AGG TC	
SHT1 (17) A1		<u>ATGTTGAAATTAAACAGCAACCC</u> AGG TC	
SHT1 (17) A2		<u>ATGTTGAAATTAAACAGCAACCC</u> AGG TC	
SHT1 (22) A1		<u>ATGTTGAAATTAAACAGCAACCC</u> AGG TC	
SHT1 (22) A2		<u>ATGTTGAAATTAAACAGCAA</u> TCC AGG TC	
SHT1 (23) A1		<u>ATGTTGAAATTAAACAGCAACCC</u> AGG TC	
SHT1 (23) A2		<u>ATGTTGAAATTAAACAGCAA</u> TCCC AGG TC	

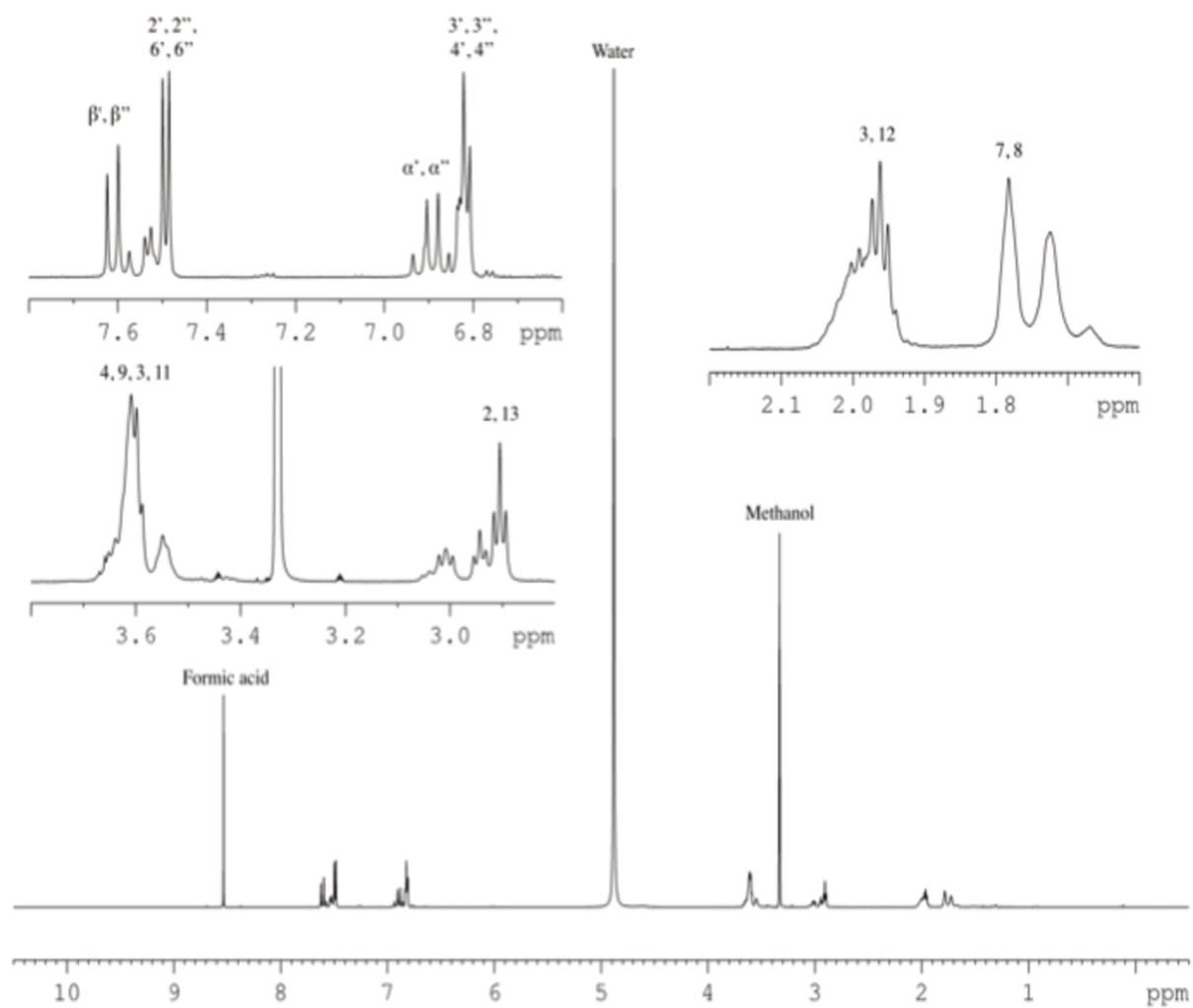
Supplementary Figure S3: Mutation events occur in *SHT1* gene at target 2 localisation in plants transformed with pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2. Blue nucleotides indicate the target site, bold red nucleotides indicate the PAM sequence.

Plant line	Target 1	PAM
SHT1 (WT)	ATGTTGAAATTAAACAGCAACCC AGG TC	
SHT1/SHT2 (7) A1	ATGTTGAAATTAAACAGCAATCCC AGG TC	
SHT1/SHT2 (7) A2	ATGTTGAAATTAAACAGCAATTCCC AGG TC	
SHT1 (WT)	...ECEQETAVTMLKLTATQVEKLRNKANKSQESEMSRGFTRYEVVTA...	451 aa
SHT1/SHT2 (7) A1	...ECEQETAVTMLKLTAPGREIAKQS*	245 aa
SHT1/SHT2 (7) A2	...ECEQETAVTMLKLTAPRSRNCETKLIKVRRVR*	253 aa
Target 2	PAM	Target 1
SHT2 (WT)	CG CCG GTGGCCGCCTGGAAC TG TG...CG CCA TACCCCCCTTGAGAAGT CC CATTAT	
SHT1/SHT2 (7) A1	CG CCG GTC --- GCCTGGAA CTTGACT TG ...CG CCA TAC-----TTGAAGAAGT CC CATTAT	
SHT1/SHT2 (7) A2	CG CCG GTG ----- ... CCCCTTGAGAAGT CC CATTAT	
SHT2 (WT)	...LSSIAGGRLELDCNADGVQFIEAYADMKLVDLDDFMQSPIFHHLIPS...	452 aa
SHT1/SHT2 (7) A2	...LSSIAGAPLKKSHYCFCKLQNLCVEVGL*	101 aa
SHT1/SHT2 (7) A1	...LSSIAGAPGT*	84 aa

Supplementary Figure S4: Representative genotype and predicted amino acid sequences of the double mutant sht1/sht2 obtained with pYLCRISPR-sgRNA1-sgRNA2-SHT1/sgRNA1-sgRNA2-SHT2. Blue nucleotides indicate the target site, bold red nucleotides indicate the PAM sequence. Lx.x = Line x. allele x. In grey, the amino acids that differ from the wild type sequences. (*) shows the end of the amino acid sequence, AA represent the amino acid chain length.

Plant line	AA
SHT1 (WT) ... NERECEQETAVTMLKLTATQVEKLRNKANKSQESEMSR...	451 aa
SHT1 (8) A1 NERECEQETAVTMLKLTAPRSRNCETKLIKVRRVR*	252 aa
SHT1 (17) A1 NERECEQETAVTMLKLTATPGREIAKQS*	245 aa
SHT1 (17) A1 NERECEQETAVTMLKLTAPFLFWKPWHILYYSF*	251 aa
SHT1 (19) A1 NERECEQETAVTMLKLTAPGREIAKQS*	245 aa
Plant line	AA
SHT2 (WT) ...LSSIAGGRLELDCNADGVQFIEAYADMKLVDLDDFMQSPIFHHLIPSVDYRHTPLEEVPLLLLQATKFVCGGW....	452 aa
SHT2 (17) A1 ... LSSIAGGRLELDCNADGVQFIEAYADMKLVDLDDFMQSPIFHHLIPSVDYRHTPP*	128 aa
SHT2 (17) A2 ... LSSIAGGRLELDCNADGVQFIEAYADMKLVDLDDFMQSPIFHHLIPSVDYRHTKKSHYCFCKLQNLCVEVGL*	145 aa
SHT2 (22) A1 ...LSSIAGGRLELDCNADGVQFIEAYADMKLVDLDDFMQSPIFHHLIPSVDYRHTPLKKSHYCFCKLQNLCVEVGL*	147 aa
SHT2 (22) A2 ...LSSIAGGRLELDCNADGVQFIEAYADMKLVDLDDFMQSPIFHHLIPSVDYRHTPP*	128 aa
SHT2 (23) A1 ...LSSIADLTAMRMVFSSSKLPT*	95 aa
SHT2 (23) A2 ...LSSIAGAPGT*	84 aa

Supplementary Figure S5: Predicted amino acid sequences of the SHT1 and SHT2 genes. Lx.x = Line x. allele x. In grey, the amino acids that differ from the wild type sequences. (*) shows the end of the amino acid sequence, AA represent the amino acid chain length.



Supplementary Figure S6: ¹H NMR spectra of *N*⁵, *N*¹⁰-dicoumaroyl spermine.

References

1. Weng, J. The evolutionary paths towards complexity: a metabolic perspective. *New Phytol.* **201**, 1141–1149 (2014).
2. Ahmed, E. et al. Secondary metabolites and their multidimensional prospective in plant life. *J. Pharmacogn. Phytochem.* **6**, 205–214 (2017).
3. Weng, J. K., Philippe, R. N. & Noel, J. P. The rise of chemodiversity in plants. *Science (80-.)* **336**, 1667–1670 (2012).
4. Cheynier, V., Comte, G., Davies, K. M., Lattanzio, V. & Martens, S. Plant phenolics: Recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiol. Biochem.* **72**, 1–20 (2013).
5. Vanholme, R. et al. Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in *Arabidopsis*. *Plant J.* **64**, 885–897 (2010).
6. Legrand, G. et al. Identification and characterization of five BAHD acyltransferases involved in hydroxycinnamoyl ester metabolism in chicory. *Front. Plant Sci.* **7**, 741 (2016).
7. 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**, 1091–100 (2014).
8. Niggegeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat. Biotechnol.* **22**, 746–754 (2004).
9. 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).
10. 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).
11. 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).
12. Edreva, A. M., Velikova, V. B. & Tsonev, T. D. Phenylamides in plants. *Russ. J. Plant Physiol.* **54**, 287–301 (2007).

13. Bassard, J. E., Ullmann, P., Bernier, F. & Werck-Reichhart, D. Phenolamides: Bridging polyamines to the phenolic metabolism. *Phytochemistry* **71**, 1808–1824 (2010).
14. Burhenne, K., Kristensen, B. K. & Rasmussen, S. K. A New Class of N-Hydroxycinnamoyltransferases: PURIFICATION, CLONING, AND EXPRESSION OF A BARLEY AGMATINE COUMAROYLTRANSFERASE (EC 2.3.1.64). *J. Biol. Chem.* **278**, 13919–13927 (2003).
15. Muroi, A. *et al.* Accumulation of hydroxycinnamic acid amides induced by pathogen infection and identification of agmatine coumaroyltransferase in *Arabidopsis thaliana*. *Planta* **230**, 517–527 (2009).
16. Onkokesung, N. *et al.* MYB8 Controls Inducible Phenolamide Levels by Activating Three Novel Hydroxycinnamoyl-Coenzyme A:Polyamine Transferases in *Nicotiana attenuata*. *Plant Physiol.* **158**, 389–407 (2012).
17. Luo, J. *et al.* A Novel Polyamine Acyltransferase Responsible for the Accumulation of Spermidine Conjugates in *Arabidopsis* Seed. *Plant Cell Online* **21**, 318–333 (2009).
18. Grienberger, E., Besseau, S. & Geoffroy, P. 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).
19. 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).
20. 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).
21. Peng, H. *et al.* A novel hydroxycinnamoyl transferase for synthesis of hydroxycinnamoyl spermine conjugates in plants. *BMC Plant Biol.* **19**, 1–13 (2019).
22. Bernard, G. *et al.* Efficient genome editing using CRISPR/Cas9 technology in chicory. *Int. J. Mol. Sci.* **20**, (2019).
23. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).

24. Nishitani, C., Hirai, N., Komori, S., Wada, M. & Kazuma, O. Efficient Genome Editing in Apple Using a CRISPR / Cas9 system. 1–8 (2013). doi:10.1038/srep31481
25. Tepfer, D. A. Genetic transformation using *Agrobacterium rhizogenes*. *Physiol. Plant.* **79**, 140–146 (1990).
26. Bais, H. P. & Venkatesh, R. T. *Agrobacterium rhizogenes* -mediated transformation of Witloof chicory – In vitro shoot regeneration and induction of flowering. *Current Sci* **83**, 83–87 (2001).
27. Bogdanovifá, M. D. et al. Production of guaianolides in *Agrobacterium rhizogenes* - transformed chicory regenerants flowering in vitro. *Ind. Crops Prod.* **60**, 52–59 (2014).
28. 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).
29. Vogt, T. Unusual spermine-conjugated hydroxycinnamic acids on polle: function and evolutionary advantage. *J. Exp. Bot.* **69**, 5311–5318 (2018).
30. Lin, S. & Mullin, C. A. Lipid, polyamide, and flavonol phagostimulants for adult western corn rootworm from sunflower (*Helianthus annuus* L.) pollen. *J. Agric. Food Chem.* **47**, 1223–1229 (1999).
31. Fixon-Owoo, S. et al. Preparation and biological assessment of hydroxycinnamic acid amides of polyamines. *Phytochemistry* **63**, 315–334 (2003).
32. Walters, D., Meurer-Grimes, B. & Rovira, I. Antifungal activity of three spermidine conjugates. *FEMS Microbiol. Lett.* **201**, 255–258 (2001).
33. Kyselka, J. et al. Antifungal Polyamides of Hydroxycinnamic Acids from Sunflower Bee Pollen. *J. Agric. Food Chem.* **66**, 11018–11026 (2018).
34. Chantrapromma, K. & Ganem, B. Total synthesis of kukoamine A, an antihypertensive constituent of *lycium chinense*. *Tetrahedron Lett.* **22**, 23–24 (1981).
35. Wang, Q. et al. Kukoamine A inhibits human glioblastoma cell growth and migration through apoptosis induction and epithelial-mesenchymal transition attenuation. *Sci. Rep.* **6**, 36543 (2016).
36. Li, G., Zhou, F., Chen, Y., Zhang, W. & Wang, N. Kukoamine A attenuates insulin resistance and fatty liver through downregulation of Srebp-1c. *Biomed. Pharmacother.* **89**, 536–543 (2017).

37. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 1–12 (2016). doi:10.1186/s13059-016-1012-2
38. Ma, X. & Liu, Y. CRISPR / Cas9-Based Multiplex Genome Editing in Monocot and Dicot Plants. *Curr. Protoc. Mol. Biol.* **115**, 1–21 (2016). doi:10.1002/cpmb.10
39. Nagel, R, Elliott. A, Masel. A, Birch. RG, M. J. . an of binary Ti plasmid ~ efaciens and Agrobacter. *FEMS Microbiology letters.* **67**, 1–4 (1990).
40. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497 (1962).
41. Xing, H.-L. *et al.* A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* **14**, 327 (2014).
42. Duhaze, 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**, 639–646 (2002).
43. Sainsbury, F., Thuenemann, E. C. & Lomonossoff, G. P. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693 (2009).
44. Eudes, A. *et al.* Exploiting members of the BAHD acyltransferase family to synthesize multiple hydroxycinnamate and benzoate conjugates in yeast. *Microb. Cell Fact.* 1–16 (2016). doi:10.1186/s12934-016-0593-5
45. Eudes, A. *et al.* Production of tranilast [N-(3',4'-dimethoxycinnamoyl)- anthranilic acid] and its analogs in yeast *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **89**, 989–1000 (2011).
46. Connelly, C. *et al.* Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science (80-.).* **285**, 901–906 (1999).
47. Gietz, R. D. & Schiestl, R. H. Quick and easy yeast transformation using the LiAc / SS carrier DNA / PEG method. *Nat protoc* **2**, 35–38 (2007).
48. Le Guennec, A., Dumez, J.-N., Giraudeau, P. & Caldarelli, S. Resolution-enhanced 2D NMR of complex mixtures by non-uniform sampling. *Magn. Reson. Chem.* **53**, 913–920 (2015).
49. Fulmer, G. R. *et al.* NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **29**, 2176–2179 (2010).

Éléments complémentaires

Dans cet article, l'analyse des mutants *sht1* et *sht2*, mise en place grâce à la technologie CRISPR/Cas9, a permis de mettre en évidence l'implication des gènes *CiSHT1* et *CiSHT2* dans la production de tetracoumaroyl spermine au niveau des boutons inflorescentiels de chicorée. L'analyse RMN de la dicoumaroyl spermine produite dans les plantes mutantes *sht1* ainsi que dans les feuilles de tabac et les hairy roots de chicorée surexprimant le gène *CiSHT2*, a montré la présence de N⁵, N¹⁰-dicoumaroyl spermine. L'enzyme *CiSHT2* serait donc responsable de l'acylation des amines secondaires de la spermine, puis l'enzyme *CiSHT1* réaliseraient l'acylation des amines primaires de la N⁵,N¹⁰-dicoumaroyl spermine, pour donner la tetracoumaroyl spermine. La surexpression de ces deux gènes dans des feuilles de tabac, a montré l'action en tandem de ces deux enzymes, pour la production de tricoumaroyl spermidine ainsi que pour la production de tetracoumaroyl spermine. Dans but de valider le fonctionnement en tandem de ces deux enzymes la double complémentation du mutant *sht* d'*Arabidopsis* avec *CiSHT1* et *CiSHT2* a été initiée. De plus le gène *CiSHT1* a été surexprimé dans les chevelus racinaires de chicorée surexprimant déjà *CiSHT2* et mis en place précédemment (Delporte et al. 2018)

1. Matériel et méthodes

1.1. Complementation du mutants *sht* d'*Arabidopsis* avec *CiSHT1* et *CiSHT2*

Pour les expériences complémentaires le vecteur d'expression pH2GW7 a été modifié en remplaçant le promoteur 35S par le promoteur TA29 pour l'expression spécifique dans les anthères (Koltunow et al. 2007; Karimi et al. 2002). Le vecteur d'expression pH2GW7(TA29) a été créé en suivant le protocole décrit par Delporte et al. (2018) pour la construction du vecteur pB2GW7(TA29). Comme précédemment décrit, les vecteurs d'entrée pDONR207_SHT1 et pDONR_SHT2 ont été utilisés pour produire les vecteurs pH2GW7(TA29)_SHT1 et pH2GW7(TA29)_SHT2 grâce à la recombinaison LR (Delporte et al. 2018). Les mutants *sht* d'*Arabidopsis*, lignées SALK_055511 surexprimant *CiSHT2* ou *CiSHT1* comme reporté dans Delporte et al. (2018) ont été transformés avec pH2GW7(TA29)_SHT1 ou pH2GW7(TA29)_SHT2, respectivement en utilisant *Agrobacterium tumefaciens* strain GV220 et en suivant le

protocole décrit par Delporte et al. (2018). Les plantes résistantes à l'hygromycine ont été sélectionnées.

1.2. Génération de chevelus racinaires de chicorée surexprimant CiSHT1 et CiSHT2

Le vecteur binaire d'expression pH2GW7 a été utilisé. Le pDONR207_SHT1 a été utilisé pour créer le vecteur pH2GW7_SHT1 grâce à la recombinaison LR (Delporte et al. 2018). La souche 15834 d'*A. rhizogenes* a été transformée par électroporation avec le vecteur pH2GW7_SHT1 (Nagel et al. 1990). L'induction de chevelus racinaires a été réalisée sur des feuilles de 'clone 17' provenant de plantes régénérées à partir d'une lignée de chevelu racinaire surexprimant SHT2 (Delporte et al. 2018). Les lignées ayant intégré le gène SHT1 ont été sélectionnées en ajoutant 10 mg/L d'hygromycine dans le milieu de culture solide MSHR.

1.3. Addition de spermine et de phénylalanine dans le milieu de culture

Les expériences ont été réalisées en triplicata dans 20 ml de milieu liquide MSHR/2 préalablement ensemencés avec environ 300 mg de chevelu racinaire provenant d'une pré-culture en milieu liquide MSHR/2. Douze jours plus tard, de la spermine et de la phénylalanine ont été ajoutées dans le milieu pour atteindre des concentrations de 1.5 mM et 6 mM, respectivement. Les racines ont été récoltées 6 jours après l'addition de spermine et de phénylalanine, en les plongeant immédiatement dans de l'azote liquide, puis conservées à -80°C jusqu'à l'extraction des métabolites.

2. Résultats attendus

Le mutant *sht* d'*Arabidopsis* surexprimant le gène *CiSHT2* produit des dérivés de spermine di- et tri-substitués (Delporte et al. 2018). La surexpression du gène *CiSHT1* dans ce mutant devrait donc permettre la production de spermine totalement substituée au niveau des anthères des plantes sélectionnées. Une production de spermine tetra-substituée permettrait de valider le fonctionnement en tandem des enzymes *CiSHT1* et *CiSHT2*.

De plus il a été montré que les chevelus racinaires de chicorée surexprimant le gène *CiSHT2* avaient aussi la capacité de produire des spermine mono-, di- et tri-

substituées, et que cette production était modulable par l'ajout de spermine et/ou phénylalanine dans le milieu de culture. Ces chevelus racinaires étant capables de régénérer des plantes, il est donc possible de retransformer ces plantes avec *A. rhizogenes*, et donc de surexprimer le gène *CiSHT1*. La production de tetracoumaroyl spermine au sein des chevelus racinaires surexprimant les deux gènes *CiSHT1* et *CiSHT2* permettrait aussi de valider le fonctionnement en tandem des deux enzymes, et si les quantités de spermine tetra-substituée sont assez conséquentes, de produire cette molécule dans le but de réaliser divers tests d'activités biologiques.

Conclusions et perspectives

La région Hauts de France a la chance d'abriter dans son terroir, la chicorée, *Cichorium intybus*. Cette plante qui est présente à l'état sauvage, le long des fossés ou des chemins des régions méditerranéennes, est retrouvée dans la région Hauts de France, à l'état domestiqué, dans les champs, où elle est cultivée en majorité pour la production d'inuline et en moindre mesure pour être utilisée comme substitut du café. Mais la chicorée, avant de devenir une plante emblématique de cette région, a marqué des temps très anciens, comme les civilisations Égyptienne, Grecque et Romaine. La chicorée sauvage, plante méditerranéenne, a fait partie des pharmacopées de ces grandes civilisations, et a perduré comme plante médicinale jusqu'à la renaissance. Puis, elle a peu à peu été délaissée, au profit sûrement, de plantes possédant plus d'activités biologiques et provenant de pharmacopées étrangères dont l'accessibilité fut facilitée par l'évolution du commerce maritime. C'est paradoxalement grâce à une bataille navale et la mise en place du blocus continental en 1806, par Napoléon, qu'elle retrouvera sa popularité et qu'elle deviendra une plante agronomique, cultivée non plus pour ses vertus médicinales, mais pour son utilisation comme substitut du café. Cette casquette agronomique va lui permettre de persister et d'être améliorée par les agronomes puis par les biotechnologistes. La chicorée va s'avérer comme étant une plante « docile » ayant la capacité d'être cultivée *in vitro*, et de régénérer des plantes entières, à partir de cultures cellulaires, protoplastes et chevelus racinaires. Les biotechnologies végétales ont, dans un premier temps, permis d'étudier le métabolisme primaire (croissance, reproduction) de la plante puis avec l'évolution de la biologie moléculaire, des techniques analytiques et de la génomique, se sont étendues au métabolisme spécialisé. De nombreuses molécules appartenant au métabolisme spécialisé ont été identifiées chez la chicorée et il a été montré que la plante accumule majoritairement de façon originale quatre polyphénols : les CQA, CTA, diCQA et diCTA. Ces quatre acides hydroxycinnamiques possèdent de nombreuses propriétés nutritionnelles et thérapeutiques. C'est encore grâce à sa casquette de plante agronomique que de nombreuses données génomiques ont été accumulées et que la chicorée a vu son génome séquencé. L'accumulation et l'accessibilité de ces données ont permis aux généticiens de détecter et de cartographier des gènes par homologie de séquences avec des gènes déjà caractérisés chez d'autres plantes beaucoup plus étudiées comme *Arabidopsis thaliana* ou la luzerne (*Medicago truncatula*). Au sein de notre laboratoire, l'analyse du génome de la chicorée a permis d'isoler 7 gènes appartenant au sous-clade Vb de la

famille des BAHD. Parmi ces séquences, 3 HQTs et 2 HCTs impliquées dans la voie de biosynthèse de l'acide chlorogénique ont été caractérisées par le Dr Legrand (Legrand et al. 2016). Mes travaux de thèse ont porté sur l'étude des deux séquences restantes au travers d'approches biotechnologiques et d'édition du génome, mises en place chez la chicorée.

Dans un premier temps la transformation de la chicorée médiée par *A. rhizogenes* a été mise en place au sein du laboratoire, puis a été utilisée pour surexprimer le gène CiSHT1 dans des chevelus racinaires de chicorée durant mon stage de fin d'étude (Master 2). La surexpression du gène avait été vérifiée par RT-PCR mais aucune phénolamide n'avait été détectée dans les chevelus racinaires. Pour l'étude de ce gène nous avions utilisé deux souches d'*A. rhizogenes*, la souche 15834 et la souche 2659, que nous avions utilisées sur deux génotypes de chicorée, la variété industrielle Orchies et le clone 17. L'étude des lignées mises en place, nous avait permis de montrer que lorsqu'on transformait le clone 17 avec la souche 15834, 100% des lignées produites avaient la capacité de régénérer des plantes en moins de deux mois après transformation et ne nécessitaient pas de vernalisation pour monter à fleur. Ces observations et la maîtrise de cette technique de transformation se sont avérées primordiales pour mes travaux de thèse.

Par la suite, la surexpression de CiSHT2 dans les chevelus racinaires de chicorée en tant que validation fonctionnelle *in planta* a permis de valider l'implication de ce gène dans la voie de biosynthèse des phénolamides chez la chicorée, en revanche aucune trace de tetracoumaroyl spermine n'a été détectée. Les expériences de supplémentation du milieu de culture avec de la spermine, de la spermidine et/ou de la phénylalanine ont montré que nous pouvions moduler la production de dérivés de spermine ou de spermidine mono-, di- ou tri-substitués, ce qui prouve bien la faible spécificité de substrat de cette enzyme. La présence de dérivés de spermine ou de spermidine totalement substitués pourraient donc directement être liée au pool de polyamines présentes au niveau des boutons floraux des différentes espèces. Chez *Arabidopsis*, qui accumule des spermidines totalement substituées, le ratio spermidine/spermine au niveau des boutons floraux est d'environ 5 (Fellenberg et al. 2012). D'après ces hypothèses la supplémentation du mutant *sht* d'*Arabidopsis* avec CiSHT2 aurait dû montrer la capacité de CiSHT2 à produire des spermidines totalement substituées, or la phénolamide majoritaire est la feruloyl spermidine. Ceci peut être expliqué par la présence d'une O-méthyltransférase chez les Brassicacées

(Grienenerger et al. 2009). En effet, il est fort probable que seules les SHTs présentes dans cette famille aient la capacité d'utiliser le feruloyl-CoA pour acyler la spermidine, d'ailleurs la MdSHT présente chez le pommier, n'a pas la capacité d'utiliser cet ester de coenzymeA (Elejalde-Palmett et al. 2015). Toutes les expériences menées sur la CiSHT2 ont montré son implication dans la voie de biosynthèse de phénolamides. Cependant, les activités mises en évidence, ont engendré principalement la production de spermine di-substituée et dans une moindre mesure de spermine tri-substituée, ce qui n'explique pas l'accumulation exclusive de spermine tetra-substituée au niveau du grain de pollen chez la chicorée. La CiSHT1 quant à elle, paraît très peu active mais son profil d'expression est concordant avec une activité de type SHT.

La création de mutants restant une des techniques les plus fiables pour valider la fonction d'un gène et la technologie CRISPR/Cas9 étant dans l'air du temps, nous avons donc décidé de mettre en place cet outil chez la chicorée.

La technologie CRISPR/Cas9 permettant d'éditer spécifiquement un gène, est un outil très puissant pour l'amélioration génétique des plantes et la validation fonctionnelle de gènes. Pour pouvoir utiliser efficacement cet outil chez les plantes, l'espèce étudiée doit avoir la capacité d'être transformée ainsi que la capacité à régénérer des plantes après transformation. De plus, des données génomiques conséquentes et accessibles doivent être disponibles chez l'espèce étudiée. Le fait que la chicorée soit étudiée depuis de nombreuses années au sein de notre laboratoire, que nous maîtrisions de nombreuses techniques de transformation et de régénération, et que les données génomiques de la chicorée soient accessibles grâce à notre étroite collaboration avec l'entreprise Florimond Desprez, nous a permis de mettre en place cet outil chez cette espèce. L'édition du gène de la *phytoène désaturase* (*pds*) chez la chicorée grâce à deux techniques de transformation : la transformation stable par *A. rhizogenes* et la transformation transitoire de protoplastes médiée par le polyéthylène glycol (PEG), nous a permis de montrer notre maîtrise des deux techniques, de dresser les avantages et les inconvénients de chacune, ainsi que de valider le fonctionnement du promoteur U6 sélectionné chez la chicorée. La transformation à partir de protoplastes est plus longue, le taux de réussite plus faible (4%), et il n'est pour l'instant pas possible de réaliser une sélection précoce des protoplastes transformés, le screening des plantes mutées est donc plus long. De plus nous avons montré que ce type de transformation, du fait d'une entrée massive d'ADN plasmidique dans la cellule, provoque l'insertion d'ADN étranger dans le génome de la plante. En revanche,

maitriser la transformation de protoplastes médiée par le PEG, nous permet d'entrevoir la possibilité d'utiliser le système ribonucléoproteique (RNP) chez la chicorée, pour lequel des complexes ribonucléoprotéiques sont utilisés à la place des plasmides, ce qui évite toute introduction d'ADN étranger dans le génome de la plante. La transformation avec *Agrobacterium rhizogenes* quant à elle, permet d'obtenir des plantes mutées plus rapidement car une sélection précoce est possible. Cette technique est donc plus adaptée à la recherche fondamentale. De plus, la capacité des plantes transformées *via A. rhizogenes*, à monter à fleur rapidement, sans avoir besoin de vernalisation, a été un énorme avantage pour l'étude des gènes *SHTs* exprimés au niveau des boutons inflorescentiels. Grâce à la maîtrise de la technique Golden gate, de clonage des targets, mise en place pour l'édition de la *phytoene desaturase*, et la validation fonctionnelle du promoteur CiU6 sélectionné dans le génome de la chicorée, nous étions donc capables de potentiellement muter n'importe quel gène d'intérêt chez la chicorée, ce qui nous a permis de produire des mutants pour les gènes *shts*.

Lors de l'étude des mutants pour les gènes *shts*, nous avons pu voir une différence de phénotypes chez toutes les plantes ayant subi une mutation bi-allélique ou homozygote, résultant en un arrêt précoce de la traduction de la protéine d'intérêt. Pour les plantes mutées au niveau du gène *sht1* nous avons pu constater une production de di- et de tricoumaroyl spermine au niveau des boutons inflorescentiels, alors que les mutants *sht2* et le double mutant *sht1sht2* ne produisent aucune phénolamide. L'enzyme SHT2 paraît donc responsable de la production de di et de tricoumaroyl spermine, ces résultats sont concordants avec ceux obtenus lors de la surexpression du gène *CiSHT2* dans les chevelus racinaires de chicorée, lors des tests enzymatiques, ainsi que dans les expériences menées sur le tabac. La présence majoritaire de dicoumaroyl spermine dans les mutants *sht2* ainsi que dans les expériences susmentionnées, portent à croire que la SHT2 est responsable de la production de dicoumaroyl spermine ; la présence de tricoumaroyl spermine serait due à la promiscuité catalytique de cette enzyme, c'est à dire qu'elle possède un large spectre de substrats. Le fait que les mutants *sht2* ne soient pas capables de produire de phénolamides nous fait penser que l'enzyme SHT1 a besoin de l'activité préalable de SHT2 pour pouvoir fonctionner ensuite, et serait donc la dernière enzyme de la voie

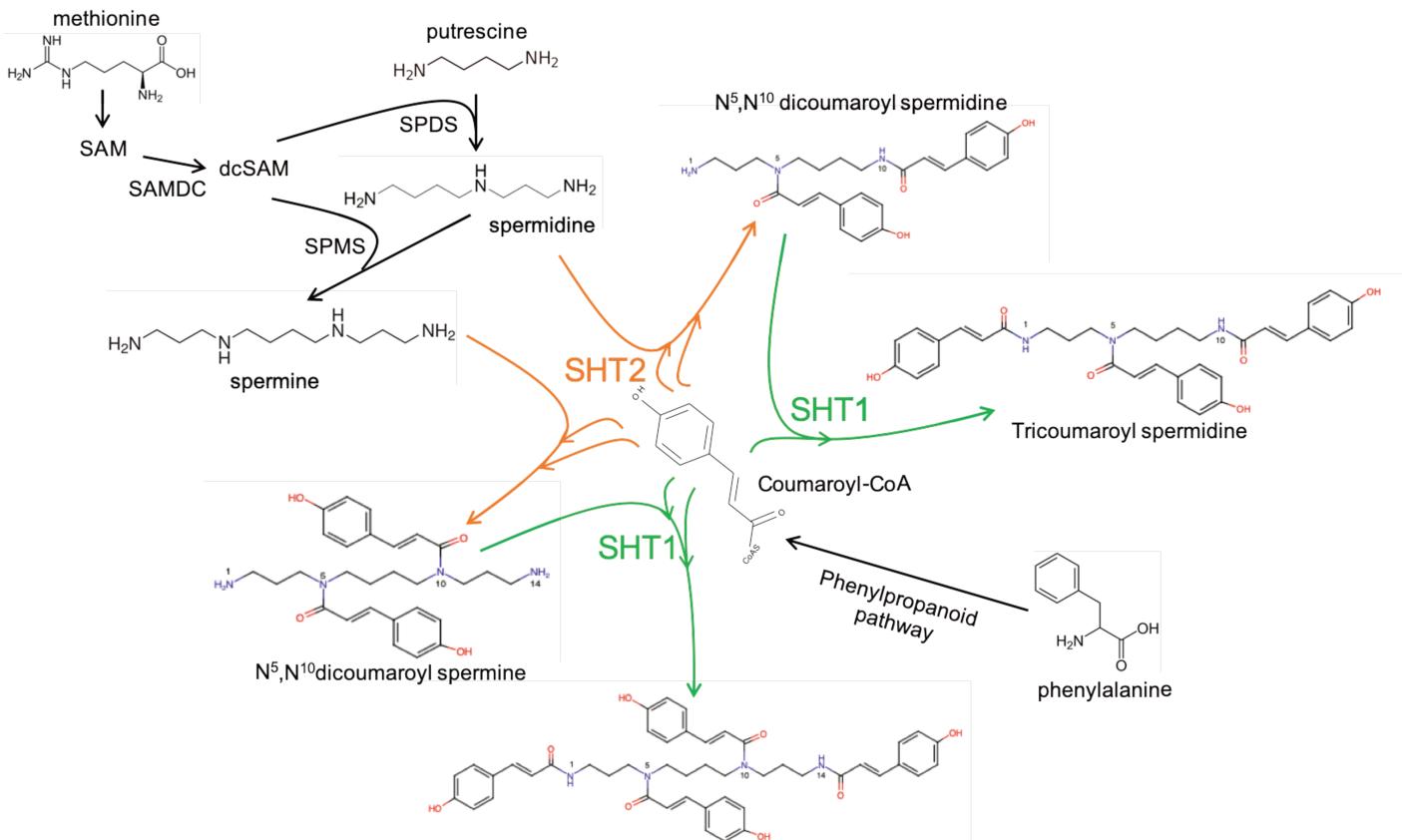


Figure 4.1 : Représentation schématique de la connexion entre la voie de biosynthèse des polyamines et le métabolisme des acides hydroxycinnamiques au niveau du manteau pollinique chez la chicorée. Abréviations : SAM, S-AdénosylMéthionine ; dcSAM, décarboxy-S-AdénosylMéthionine ; SAMDC, S-AdénosylMéthionine décarboxylase ; SPDS, Spermidine Synthase ; SPMS, Spermine Synthase ; SHT, Spermidine Hydroxycinnamoyl Transférase

de biosynthèse de la tetracoumaroyl spermine permettant la transformation de di- et de tricoumaroyl spermine en tetracoumaroyl spermine. Les analyses RMN de la dicoumaroyl spermine ont montré que la spermine était acylée sur les deux amines N⁵ et N¹⁰, la SHT2 serait donc responsable de l'acylation des amines secondaires, et la SHT1 réaliseraient, ensuite, l'acylation des amines primaires, après l'action de la SHT2 (Figure 4.1). Chez *N. attenuata*, deux enzymes structuralement très proches ont été caractérisées et sembleraient fonctionner en tandem comme CiSHT1 et CiSHT2. En effet, l'enzyme DH29 serait responsable de la production de spermidine monosubstituée, puis l'enzyme CV86 réaliserait l'acylation d'une seconde fonction amine pour produire de la spermidine disubstituée. De plus, comme chez les mutants *sht1*, l'inhibition de la CV86 causerait l'accumulation de spermidine monosubstituée, produit par la DH29 (Onkokesung et al. 2012).

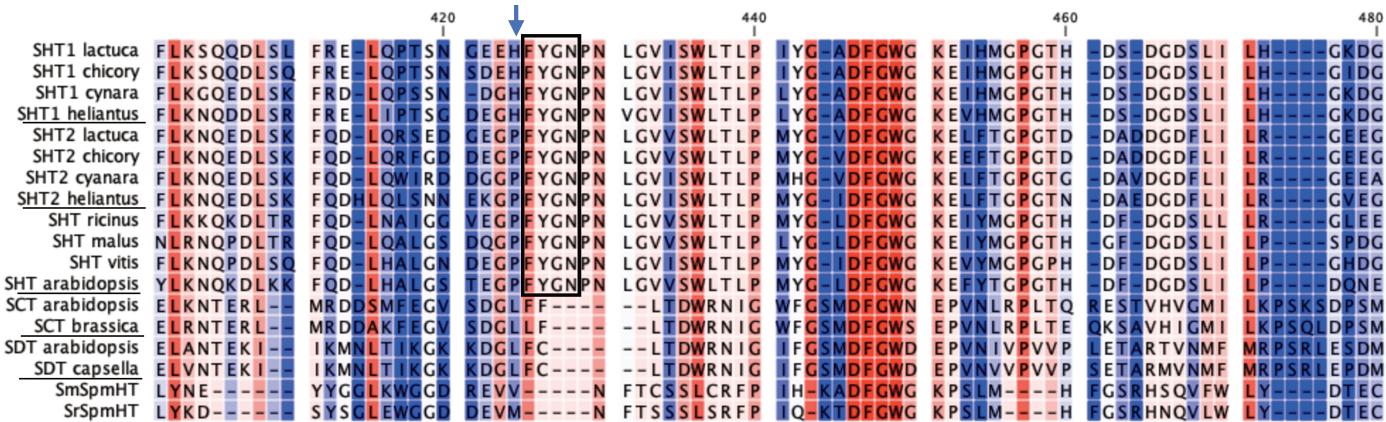


Figure 4.2: Analyse des séquences d'enzymes de la famille des BAHD capables d'acyler les polyamines. La flèche bleue montre le 424 ème AA ; le motif FYGN est encadré en noir ; le motif DFGWG caractéristique des enzymes appartenant à la famille des BAHD est souligné en rouge.

Une récente étude menée sur l'aubergine a aussi mis en évidence des enzymes capables d'acyler la spermine au niveau des fleurs et des fruits (Peng et al. 2019). Ces enzymes nommées spmHT (spermine Hydroxycinnamoyl Transférase) appartiennent elles aussi à la famille des BAHD, mais l'analyse phylogénétique montre que ces enzymes sont plus proches des SCT et SDT déjà caractérisées. Comme les SCT et SDT, les spmHT sont seulement capables d'acyler les fonctions amines primaires des polyamines. D'un point de vue structural, elles ne possèdent pas au niveau de leur séquence en acides aminés (AA), le motif FYGN retrouvé chez toutes les enzymes capables de réaliser au minimum des acylations doubles sur la spermidine et la spermine, mis en évidence par Delporte et al 2018 (Figure 4.2).

Les enzymes SHT1 et SHT2 étudiées lors de cette thèse sont donc les seules enzymes connues capables de réaliser plusieurs acylations sur la spermine. La forte homologie de séquence entre ces deux enzymes laisse penser qu'elles proviennent toutes deux de la duplication d'un même gène qui pourrait vraisemblablement être une SHT ancestrale. Cette duplication ayant sûrement été suivie par une néo-fonctionnalisation aléatoire d'un des deux gènes. Il est donc possible qu'entre la SHT1 et la SHT2, une modification d'un seul AA, ait permis la différence catalytique observée entre ces deux enzymes. Si on s'attarde sur la Figure 4.2 nous pouvons constater qu'au niveau du motif FYGN, le 424 ème AA, situé juste avant le motif FYGN, est une proline chez les SHT et les SHT2, alors que chez les SHT1 il est remplacé par une histidine, AA connu pour son intervention dans les sites actifs, du fait de sa capacité à capter un proton. Un changement de la sorte pourrait bien être à l'origine de la néo-

fonctionnalisation de la SHT1. De son côté, la SHT2 pourrait bien être la descendante de la SHT ancestrale. Pour tester cette hypothèse la création de chimères SHT1/SHT2 est en cours dans notre laboratoire. Concernant l'origine de la duplication de cette SHT ancestrale, il se pourrait qu'elle puisse être liée à la modification du pool de polyamines libres au niveau des organes reproducteurs mâles. En effet nous avons constaté que chez la chicorée le ratio spermidine/spermine était en faveur de la spermine, qu'il y avait présence de deux SHTs fonctionnant en tandem, et que la phénolamide majoritairement produite était un dérivé de spermine totalement substitué. Un dosage des polyamines libres au niveau des boutons floraux d'autres espèces de la famille des Astéracées ainsi que chez des espèces proches pourrait aller dans le sens de cette hypothèse. Récemment nous avons réalisé des tests d'activités *in vitro* sur l'AtSHT au sein de notre laboratoire. Ces résultats restent à confirmer mais il semblerait que l'AtSHT soit capable d'utiliser la spermine pour former des dérivés mono- di- et trisubstitués, un peu comme la CiSHT2. Nous pouvons supposer qu'au cours de l'évolution, une augmentation de l'activité de la spermine synthase (SPMS) due à un événement inconnu, soit responsable de la modification du ratio spermidine/spermine en faveur de la spermine. La qualité et la concentration des phénolamides présentes au niveau du grain de pollen auraient donc été modifiées, présence de moins de dérivés de spermidine totalement substitués et présence de dérivés de spermine di- ou tri- substitués. Les dérivés di- et tri- substitués de spermine étant potentiellement moins efficace que les spermidines et spermines totalement substituées et supposé que ces molécules soient vitales pour la conservation du grain de pollen ou dans les étapes de reproduction, il est envisageable de penser que la pression de sélection engendrée par cette modification ait pu causer la duplication de la SHT et la néo-fonctionnalisation de la SHT1 pour produire des dérivés de spermine totalement substitués.

La mise en place de plantes mutantes *Cisht1* et *Cisht2* nous a permis de montrer l'implication des deux enzymes dans la production de tetracoumaroyl spermine au niveau du tapetum des anthères de chicorée, de plus les expériences de dosage des polyamines libres ont montré que c'était la quantité relative de chaque polyamine et non la spécificité enzymatique qui dirigeait le flux vers la production de spermidine ou de spermine entièrement substituée. Maintenant, une étude approfondie des mutants mis en place permettra peut-être de répondre à la question « pourquoi ? », pourquoi ces molécules sont-elles produites, quels sont leurs rôles ? Et quels sont les avantages

ou les inconvénients de produire des dérivés de spermine à la place de dérivés de spermidine ? Pour cela des tests de viabilité du pollen sont en cours et des tests de résistances aux UV ainsi que d'attraction des pollinisateurs sont envisagés.

La connaissance des voies de biosynthèse et la caractérisation des gènes impliqués dans la synthèse de ces phénolamides est un atout pour la production de ces molécules qui sont susceptibles de présenter des activités biologiques intéressantes. En ce qui concerne les propriétés biologiques des dérivés de spermine seulement peu d'études ont été réalisées, du fait de la rareté de ces molécules. Une étude sur la N¹-coumaroyl spermine montre que cette molécule à la capacité d'inhiber efficacement les neurorécepteurs de mammifères et d'écrevisse *in vitro* (Fixon-Owoo et al. 2003) ainsi que la trypanothion réductase, une enzyme essentielle pour la survie de protozoaires pathogènes tels que les trypanosomes, ce qui représenterait un grand intérêt pour la gestion de la douleur et la lutte antiparasitaire (Peng et al. 2019). Par ailleurs, il a été montré que la N^{1,N¹⁴}-bis (dihydrocafféoyl) spermine, également nommée kukoamine A possédait des activités anticancéreuses, antidiabétiques et hépatoprotectrices (Li et al. 2017)

Dans le chapitre I nous avons vu que les chevelus racinaires de chicorée surexprimant le gène *CiSHT2* étaient capables de produire des dérivés de spermine et de spermidine, mono-, di-, et tris substitués, et que cette production était modulable, ce qui nous a permis de produire quelques milligrammes de N^{5,N¹⁰}-dicoumaroyl spermine. Il serait donc intéressant de mettre au point des protocoles de purification dans le but de purifier chaque dérivé, pour par la suite réaliser des tests d'activités biologiques. La surexpression du gène *CiSHT1* dans ces chevelus racinaires, va peut-être nous permettre de produire de la tetracoumaroyl spermine, et si cette production est conséquente, la purification suivie de la réalisation de tests d'activités sur cette molécule sera envisageable. De plus, nous avons pu constater que les chevelus racinaires étaient capables d'absorber la spermine et la phénylalanine lorsqu'elles étaient ajoutées dans le milieu de culture, il est donc envisageable de remplacer la phénylalanine par de l'acide phlorétique (acide hydrocoumarique), pour produire des molécules de la famille des kukoamines.

La production en utilisant le système levure a également été mis en place dans notre laboratoire, *via* l'introduction simultanée des gènes codant, *CiSHT2* et/ou *CiSHT1* et une 4CL (*At4CL5* d'*Arabidopsis*). La production de tetracoumaroyl spermine par la

souche exprimant CiSHT1 et CiSHT2 a été obtenue. Des expériences d'optimisation de la production sont toutefois à mettre en œuvre comme par exemple l'ajout de la tyrosine ammonia-lyase de *Flavobacterium johnsoniae* (*FjTAL*), capable de synthétiser le *p*-coumarate, dans le but d'éliminer le besoin de supplémenter le milieu de culture en ce dernier. L'ajout d'une spermine synthase est aussi envisageable. L'utilisation d'une souche capable d'excréter ces molécules serait aussi un bon moyen d'améliorer la production, en vue d'un éventuel « scale up » en bioréacteur. De l'acide dihydrocaféïque pourrait aussi être ajouté dans le milieu pour produire des kukoamines.

Enfin, outre l'étude des gènes *SHTs*, la maîtrise de l'outil CRISPR/Cas9 que j'ai pu mettre au point au cours de ma thèse, chez la chicorée, au sein du laboratoire, est un grand pas pour l'étude des voies de biosynthèse chez cette plante. De nombreux caractères de cette plante vont pouvoir être étudiés et améliorés rapidement chez la chicorée. Un projet ayant pour but de moduler l'amertume de la chicorée, en éditant des gènes de la voie de biosynthèse des lactones sesquiterpéniques a déjà débuté au laboratoire, en partenariat avec l'entreprise Florimond Desprez (Thèse CIFRE de Justine Domont, en cours).

Références bibliographiques

- Ahmed, E., Arshad, M., Zakriyya Khan, M., Shoaib Amjad, M., Mehreen Sadaf, H., Riaz, I., Sabir, S., Ahmad. N., and Ejaz Ahmed, C.** (2017). Secondary Metabolites and Their Multidimensional Prospective in Plant Life. *Journal of Pharmacognosy and Phytochemistry* **6** (2): 205–14.
- Alagoz, Y., Gurkok, T., Zhang, B and Unver, T.** (2016) Manipulating the Biosynthesis of Bioactive Compound Alkaloids for Next-Generation Metabolic Engineering in Opium Poppy Using CRISPR-Cas 9 Genome Editing Technology. *Scientific Reports* **6** (August): 30910. <https://doi.org/10.1038/srep30910>.
- B. Rees, S. and B. Harborne, J.** (1985). The Role of Sesquiterpene Lactones and Phenolics in the Chemical Defence of the Chicory Plant. *Phytochemistry* **24** : 2225–2231. [https://doi.org/https://doi.org/10.1016/S0031-9422\(00\)83015-0](https://doi.org/https://doi.org/10.1016/S0031-9422(00)83015-0).
- Bais, H. P., George, J. and Ravishankar. G. A.** (1999). Influence of Polyamines on Growth of Hairy Root Cultures of Witloof Chicory (*Cichorium Intybus L.* Cv. Lucknow Local) and Formation of Coumarins. *Journal of Plant Growth Regulation* **18**: 33–37. <https://doi.org/10.1007/PL00007043>.
- Bais, H. P., Venkatesh, R. T., Arun Chandrashekhar and Ravishankar, G. A.** (2001). Agrobacterium Rhizogenes -Mediated Transformation of Witloof Chicory – *In Vitro* Shoot Regeneration and Induction of Flowering. *Current science* **80** : 83–87.
- Bais, H. P., Govindaswamy, S and Ravishankar, G. A.** (2000). Enhancement of Growth and Coumarin Production in Hairy Root Cultures of Witloof Chicory (*Cichorium Intybus L.* . Cv . Lucknow Local) under the Influence of Fungal Elicitors. *Journal of Bioscience and Bioengineering*. **90** (6): 648–53.
- Barreda, V. D., Palazzi, L., Tellería, M. C., Olivero, E. B., Raine, J. I and Forest, F.** (2015). Early Evolution of the Angiosperm Clade Asteraceae in the Cretaceous of Antarctica. *Proc. natl. Acad. Sci. U. S. A.* **112** : 10989–94. <https://doi.org/10.1073/pnas.1423653112>.
- Beek, T. A., Maas, P., King, B. M., Leclercq, E., Voragen, A. G. J and Groot, A.** (1990). Bitter Sesquiterpene Lactones from Chicory Roots. *J. Agric. Food Chem.* **38**: 1035–1038. <https://doi.org/10.1021/jf00094a026>.
- Bernard, G., Gagneul, D., Alves Dos Santos, H., Etienne, A., Hilbert, J. L and Rambaud, C.** (2019). Efficient Genome Editing Using CRISPR / Cas9 Technology in Chicory. *Internationnal journal of molecular sciences.* <https://doi.org/10.3390/ijms20051155>.

Binding, H., Nehls, R., Kock, R., Finger, J and Mordhorst, G. (1980). Comparative Studies on Protoplast Regeneration in Herbaceous Species of the Dicotyledoneae Class. *Z. Pflanzenphysiol* **101** : 119–130. [https://doi.org/10.1016/s0044-328x\(81\)80046-3](https://doi.org/10.1016/s0044-328x(81)80046-3).

Bontpart, T., Cheynier, V., Ageorges, A and Terrier, N. (2015). BAHD or SCPL Acyltransferase? What a Dilemma for Acylation in the World of Plant Phenolic Compounds. *New Phytologist* **208** : 695–707. <https://doi.org/10.1111/nph.13498>.

Borrelli, G. M and Trono, D. (2016). Molecular Approaches to Genetically Improve the Accumulation of Health-Promoting Secondary Metabolites in Staple Crops-a Case Study: The Lipoxygenase-B1 Genes and Regulation of the Carotenoid Content in Pasta Products. *International Journal of Molecular Sciences* **17**, 1177. <https://doi.org/10.3390/ijms17071177>.

Bourgaud, F., Gravot, A., Milesi, S and Gontier, E. (2001). Production of Plant Secondary Metabolites: A Historical Perspective. *Plant Science* **161** : 839-851. [https://doi.org/10.1016/S0168-9452\(01\)00490-3](https://doi.org/10.1016/S0168-9452(01)00490-3).

Bulgakov, V. P. (2008). Functions of R_{ol} Genes in Plant Secondary Metabolism. *Biotechnology Advances* **26** : 318–24. <https://doi.org/10.1016/j.biotechadv.2008.03.001>.

Cappelle, C., Mörchen, M., Hilbert, J. L. and Rambaud, C. (2007). Regeneration and Molecular Characterization of a Male Sterile Interspecific Somatic Hybrid between *Cichorium Intybus* and *Cichorium Endivia*. *Plant Science* **172** : 596–603. <https://doi.org/https://doi.org/10.1016/j.plantsci.2006.11.010>.

Carazzone, C., Mascherpa, D., Gazzani, G and Papetti, A. (2013). Identification of Phenolic Constituents in Red Chicory Salads (*Cichorium Intybus*) by High-Performance Liquid Chromatography with Diode Array Detection and Electrospray Ionisation Tandem Mass Spectrometry. *Food Chemistry* **138** : 1062–71. <https://doi.org/https://doi.org/10.1016/j.foodchem.2012.11.060>.

Cartea, M. E., Velasco, P., Obregón, S., Padilla, G and de Haro, A. (2008). Seasonal Variation in Glucosinolate Content in *Brassica Oleracea* Crops Grown in Northwestern Spain. *Phytochemistry* **69**: 403–410. <https://doi.org/10.1016/j.phytochem.2007.08.014>.

Castelblanque, L., Balaguer, B., Martí, C., Rodríguez, J. J., Orozco, M and Vera, P. (2016). Novel Insights into the Organization of Laticifer Cells: A Cell Comprising a Unified Whole System. *Plant Physiology* **172** : 1032–1044. <https://doi.org/10.1104/pp.16.00954>.

Cheynier, V., Comte, G., Davies, K., Lattanzio,V and Martens, S. (2013). Plant Phenolics: Recent Advances on Their Biosynthesis, Genetics, and Ecophysiology. *Plant Physiology and Biochemistry* **72** : 1–20. <https://doi.org/https://doi.org/10.1016/j.plaphy.2013.05.009>.

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J. and Voytas, D. F. (2010). Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **186** : 757-761 <https://doi.org/10.1534/genetics.110.120717>.

Cocking, E.C. (1960). A Method for the Isolation of Plant Protoplasts and Vacuoles. *Nature* **187** : 962–63. <https://doi.org/10.1042/bj0260801>.

Cohen, S. N., Chang, A. C. Y., Boyer, H. B and Helling, R. B. (1973). Construction of Biologically Functional Bacterial Plasmids *In Vitro*. *Proc. Nat. Acad. Sci. USA* **70** : 3240–3244.

Comino, C., Hehn, A., Moglia, A., Menin, B., Bourgaud, F., Lanteri, S and Portis, E. (2009). The Isolation and Mapping of a Novel Hydroxycinnamoyltransferase in the Globe Artichoke Chlorogenic Acid Pathway. *BMC Plant Biology* **13**: 1–13. <https://doi.org/10.1186/1471-2229-9-30>.

Crepé, L., Chupeau, M. C and Chupeau, Y. (1982). The Isolation and Culture of Leaf Protoplasts of *Cichorium Intybus* and Their Regeneration into Plants. *Z. Pflanzenphysiol* **107**: 123–131. [https://doi.org/10.1016/s0044-328x\(82\)80098-6](https://doi.org/10.1016/s0044-328x(82)80098-6).

Cruz H. A and Guillén, J. C. (2012). How to Get Exogenous DNA to Cross the Cell Membrane of Plants: Comment on ‘Physical Methods for Genetic Transformation in Plants’ by Rivera et Al.” *Physics of Life Reviews* **9** : 348–349. <https://doi.org/https://doi.org/10.1016/j.plrev.2012.07.008>.

D' Auria, J. C. (2006). Acyltransferases in Plants : A Good Time to Be BAHD. *Current Opinion in Plant Biology* **9** : 331–34. <https://doi.org/10.1016/j.pbi.2006.03.016>.

Das, S., Vasudeva, N and Sharma, S. (2016). *Cichorium Intybus* : A Concise Report on Its Ethnomedicinal , Botanical , and Phytopharmacological Aspects. *Drug development and therapeutics* **7** : 1–12. <https://doi.org/10.4103/2394-6555.180157>.

Delporte, M., Bernard, G., Legrand, G., Hielscher, B., Lanoue, A., Molinié, R., Rambaud, C., Mathiron, D., Besseau, S., Linka, N., Hilbert, J. L and Gagneul, D. A BAHD Neofunctionalization Promotes Tetrahydroxycinnamoyl Spermine Accumulation in Pollen Coat of the Asteraceae Family. *Journal of Experimental Botany* **69**: 5365–5371. <https://doi.org/10.1093/jxb/ery320/5088093>.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J and Charpentier, E. (2011). CRISPR RNA Maturation by Trans-Encoded Small RNA and Host Factor RNase III, *Nature* **471** : 2–9. <https://doi.org/10.1038/nature09886>.

Demarly, Y, and Sibi, M. Amélioration de Plantes et Biotechnologies. Paris France, John Libbey Eurotext, 1996, 150

Eenink, A H. (1982). Compatibility and incompatibility in witloof-chicory (*cichorium Intybus*. L.) gametic competition after mixed pollination and double pollinations. *Euphytica* **31** : 773–786.

Eibl, R., Meier, P., Stutz, I., Schildberger, D., Huhn, T and Eibl, D. (2018). Plant Cell Culture Technology in the Cosmetics and Food Industries : Current State and Future Trends. *Appl Microbiol Biotechnol* **102** : 8661–8675.

El-Sayed, Y. S., Lebda, M. A., Hassinin, M and Neoman, S. A. (2015). Chicory (*Cichorium Intybus* L.) Root Extract Regulates the Oxidative Status and Antioxidant Gene Transcripts in CCl₄-Induced Hepatotoxicity. *PLOS ONE* **10** (3): e0121549. <https://doi.org/10.1371/journal.pone.0121549>.

Elejalde-Palmett, C., De Bernonville, T. D., Glevarec, G., Pichon, O., Papon, N., Courdavault, V., St-Pierre, B., Giglioli-Guivarc'H, N., Lanoue, A and 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** (22): 7271–85. <https://doi.org/10.1093/jxb/erv423>.

Fellenberg, C., Ziegler, J., Handrick, V and Vogt, T. (2012). Polyamine Homeostasis in Wild Type and Phenolamide Deficient *Arabidopsis Thaliana* Stamens. *Frontiers in Plant Science* **3** : 1–11. <https://doi.org/10.3389/fpls.2012.00180>.

Fixon-Owoo, S., Levasseur, F., Williams, K., Sabado, T. N., Lowe, M., Klose, M., Mercier, A. J., Fields, P and Atkinson, J. (2003). Preparation and Biological Assessment of Hydroxycinnamic Acid Amides of Polyamines. *Phytochemistry* **63** : 315–34. [https://doi.org/https://doi.org/10.1016/S0031-9422\(03\)00133-X](https://doi.org/10.1016/S0031-9422(03)00133-X).

Georgiev, M., Pavlov, A and Bley, T. (2007). Hairy root type plant *in vitro* systems as sources of bioactive substances. *Appl Microbiol Biotechnol* **74** : 1175–1185

Gershenzon, J and Dudareva, N. (2007). The Function of Terpene Natural Products in the Natural World. *Nature Chemical Biology* **3** : 408. <https://doi.org/10.1038/nchembio.2007.5>.

Gomez, C., Terrier, N., Torregrosa, L., Vialet, S., Fournier-Level, A., Verriès, C., Souquet, J. M., Mazauric, J. P., Klein, M., Cheynier, V and Ageorges, A. (2009). Grapevine MATE-Type Proteins Act as Vacuolar H⁺-Dependent Acylated Anthocyanin Transporters. *Plant Physiology* **150** : 402–15. <https://doi.org/10.1104/pp.109.135624>.

Gonthier, L., Blaissiau, C., Morchen, M., Cadalen, T., Poiret, M., Hendriks, T and Quillet, M. C. (2013). High-Density Genetic Maps for Loci Involved in Nuclear Male Sterility (NMS1) and Sporophytic Self-Incompatibility (S -Locus) in Chicory (*Cichorium Intybus* L ., Asteraceae) *Theor Appl Genet* **126** : 2103–2121.

Gramazio, P., Prohens, J., Plazas, M., Andújar, I., Herraiz, F. J., Castillo, E., Knapp, S., Meyer, R. S and Vilanova, S. (2014). Location of Chlorogenic Acid Biosynthesis Pathway and Polyphenol Oxidase Genes in a New Interspecific Anchored Linkage Map of Eggplant. *BMC Plant Biology* **14** : 350. <https://doi.org/10.1186/s12870-014-0350-z>.

Grienenerger, E., Besseau, S and Geoffroy, P. (2009). A BAHD Acyltransferase Is Expressed in the Tapetum of *Arabidopsis* Anthers and Is Involved in the Synthesis of Hydroxycinnamoyl Spermidines. *Plant Journal* **58** : 246–59. <https://doi.org/10.1111/j.1365-313X.2008.03773.x>.

Gwozdz, E. (1973). Effect of IAA on Growth, Organogenesis and RNA Metabolism during Development of *Cichorium Intybus* Explants Cultured *in Vitro*. *Acta Societatis Botanicorum Poloniae* **17** : 3

Haslam, T. M., Gerelle, W. K., Graham, S.W and Kunst, L. (2017). The Unique Role of the ECERIFERUM2-LIKE Clade of the BAHD Acyltransferase Superfamily in Cuticular Wax Metabolism. *Plants* **6** : 23. <https://doi.org/10.3390/plants6020023>.

Heller, R. (1953). Researches on the Mineral Nutrition of Plant Development. *Ann. Sci. Nat. Bot. Biol. Vg* **14** : 1–223.

Heyman, H. M., Senejoux, F., Seibert, I., Klimkait, T., Maharaj, V. J and Meyer, J. J. M. (2015). Identification of Anti-HIV Active Dicaffeoylquinic- and Tricaffeoylquinic Acids in *Helichrysum Populifolium* by NMR-Based Metabolomic Guided Fractionation. *Fitoterapia* **103** : 155–64. <https://doi.org/10.1016/j.fitote.2015.03.024>.

Hooykaas, P. J. J and Schilperoort, R. A. (1992). Agrobacterium and Plant Genetic Engineering. *Plant Molecular Biology* **19** : 15–38. <https://doi.org/10.1007/BF00015604>.

Huchelmann, A., Boutry, M and Hachez, C. (2017). Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. *Plant Physiology* **175** : 6–22. <https://doi.org/10.1104/pp.17.00727>.

Jackson, D. A., Symons, R. H and Berg, P. (1972). Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40 : Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of *Escherichia Coli*. *Proceedings of the National Academy of Sciences* **69** : 2904–9. <https://doi.org/10.1046/j.1341-8076.2003.00083.x>.

Kandeler, R and Ullrich, W. R. (2009). Symbolism of Plants : Examples from European-Mediterranean Culture Presented with Biology and History of Art. *Journal of Experimental Botany* **60** : 3973–74. <https://doi.org/10.1093/jxb/erp248>.

Karimi, M., Inze, D., Depicker, A., Inzé, D and Depicker, A. (2002). GATEWAY(TM) Vectors for Agrobacterium-Mediated Plant Transformation. *Trends in Plant Science* **7** : 193–95.

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 Physiology* **152** : 1731-1747. <https://doi.org/10.1104/pp.109.151738>.

Kaushik, P., Andújar, I., Vilanova, S., Plazas, M., Gramazio, P., Herraiz, F. J., Brar, N. S and Prohens, J. (2015). Breeding Vegetables with Increased Content in Bioactive Phenolic Acids. *Molecules* **20** 18464–81. <https://doi.org/10.3390/molecules201018464>.

Kim, B. G., Jung, W. D., Mok, H and Ahn, J. H. (2013). Production of Hydroxycinnamoyl-Shikimates and Chlorogenic Acid in *Escherichia Coli*: Production of Hydroxycinnamic Acid Conjugates. *Microbial Cell Factories* **12** 1–11.

Kim, Y. G., Cha, J and Chandrasegaran, S. (1996). Hybrid Restriction Enzymes: Zinc Finger Fusions to Fok I Cleavage Domain." *Proceedings of the National Academy of Sciences* **93** : 1156–60. <https://doi.org/10.1073/pnas.93.3.1156>.

Kim, Y. B., Thwe, A. A., Kim, Y. J., Li, X., Kim, H. H., Park, P. B., Suzuki, T., Kim, S. J and Park, S. U. (2013). Characterization of Genes for a Putative Hydroxycinnamoyl-coenzyme A Quinate transferase and *p*-Coumarate 3'-Hydroxylase and Chlorogenic Acid Accumulation in Tartary Buckwheat. *J. Agric. Food Chem* **61** : 4120-4126,

Koltunow, A. M., Truettner, J., Cox, K., Wallroth, M and Goldberg, R. B. (2007). Different Temporal and Spatial Gene Expression Patterns Occur during Anther Development. *The Plant Cell* **2** : 1201. <https://doi.org/10.2307/3869340>.

Kraker, J. W, Franssen, M. C. R., De Groot, A., Ko, W. A and Bouwmeester, H. J. (1998). Germacrene A Biosynthesis The Committed Step in the Biosynthesis of Bitter Sesquiterpene Lactones in Chicory. *Plant Physiol* **117** : 1381–92.

Lallemand, L. A., Zubieta, C., Lee, S. G., Wang, Y and McCarthy A. A. (2012). A Structural Basis for the Biosynthesis of the Major Chlorogenic Acids Found in Coffee. *Plant Physiology* **160** : 249–60. <https://doi.org/10.1104/pp.112.202051>.

Langner, T., Kamoun, S and Belhaj, K. (2018). CRISPR Crops: Plant Genome Editing Toward Disease Resistance.” *Annual Review of Phytopathology* **56** : 479-512. <https://doi.org/10.1146/annurev-phyto-080417-050158>.

Legrand, G., Delporte, M. 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 <https://doi.org/10.3389/fpls.2016.00741>.

Li, G., Zhou, F., Chen, Y., Zhang, W and Wang, N. (2017). Kukoamine A Attenuates Insulin Resistance and Fatty Liver through Downregulation of Srebp-1c. *Biomedicine & Pharmacotherapy* **89** : 536–43. <https://doi.org/https://doi.org/10.1016/j.biopha.2017.02.024>.

Li, J., Ou-Lee, T. M., Raba, R., Amundson, R. G and Last, R. L. (1993). *Arabidopsis* Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. *The Plant Cell* **5** : 171-79. <https://doi.org/10.1105/tpc.5.2.171>.

Lim, E. K., Higgins, G. S., Li, Y and Bowles, D. J. (2003). Regioselectivity of Glucosylation of Caffeic Acid by a UDP-Glucose:Glucosyltransferase Is Maintained in *Planta*. *The Biochemical Journal* **373** : 987–92. <https://doi.org/10.1042/BJ20021453>.

Limami, M. A., Sun, L. Y., Douat, C., Helgeson, J and Tepfer, D. (1998). Natural Genetic Transformation by Agrobacterium Rhizogenes . Annual Flowering in Two Biennials, Belgian Endive and Carrot. *Plant Physiology* **118** : 543–50. <https://doi.org/10.1104/pp.118.2.543>.

Lin, S and Mullin, C. A. (1999). Lipid, Polyamide, and Flavonol Phagostimulants for Adult Western Corn Rootworm from Sunflower (*Helianthus Annuus L.*) Pollen.” *Journal of Agricultural and Food Chemistry* **47** : 1223–29. <https://doi.org/10.1021/jf980858w>.

Luo, J., Fuell, C., Parr, A., Hill, L., Bailey, P., Elliott, K., Fairhurst, S. A., Martin, C and Michael, A. J. (2009). A Novel Polyamine Acyltransferase Responsible for the Accumulation of Spermidine Conjugates in *Arabidopsis* Seed. *The Plant cell* **21** : 318-333. <https://doi.org/10.1105/tpc.108.063511>.

Ma, C., Zhang, H., Li, J., Tao, S., Qiao, X., Korban, S. S., Zhang, S and Wu, J. (2017). Genome-Wide Analysis and Characterization of Molecular Evolution of the *HCT* Gene Family in Pear (*Pyrus Bretschneideri*). *Plant Syst Evol* **303** : 71–90. <https://doi.org/10.1007/s00606-016-1353-z>.

Ma, X and Liu, Y. G. (2016). CRISPR / Cas9-Based Multiplex Genome Editing in Monocot and Dicot Plants. *Curr. Protoc. Mol. Biol* **115** : 1–21. <https://doi.org/10.1002/cpmb.10>.

Ma, X., Koepke, J., Panjikar, S., Fritzsch, G and Stockigt, J. (2005). Crystal Structure of Vinorine Synthase , the First Representative of the BAHD Superfamily. *The Journal of Biological Chemistry* **280** : 13576–83.

Malarz, J., Stojakowska, A and Kisiel, W. (2013). Long-Term Cultured Hairy Roots of Chicory - A Rich Source of Hydroxycinnamates and 8-Deoxylactucin Glucoside. *Appl Biochem Biotechnol* **171** : 1589-1601. <https://doi.org/10.1007/s12010-013-0446-1>.

Marshall, W., Nirenberg, W and Matthaei, J. H. (1961). The Dependence of Cell Free Protein Synthesis in *E. Coli* upon Naturally Occurring or Synthetic Polyribonucleotides. *Biological Chemistry*, **47** : 1588-1602. <https://doi.org/10.5194/bg-10-399-2013>.

Martin-Tanguy, J. (1997). Conjugated Polyamines and Reproductive Development: Biochemical, Molecular and Physiological Approaches. *Physiologia Plantarum* **100** : 675–88. <https://doi.org/10.1111/j.1399-3054.1997.tb03074.x>.

Meurer, B., Wiermann, R and Strack, D. (1988). Phenylpropanoid Patterns in Fagales Pollen and Their Phylogenetic Relevance. *Phytochemistry* **27** : 823–28. [https://doi.org/https://doi.org/10.1016/0031-9422\(88\)84100-1](https://doi.org/https://doi.org/10.1016/0031-9422(88)84100-1).

Mugnier, J. (1988). Establishment of New Axenic Hairy Root Lines by Inoculation with Agrobacterium Rhizogenes. *Plant Cell Reports* **7** : 9–12. <https://doi.org/10.1007/BF00272966>.

Murashige, T and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol. Plant.* **15** : 473–97.

Muroi, A., Ishihara, A., Tanaka, C., Ishizuka, A., Takabayashi, J., Miyoshi, H and Nishioka, T. (2009). Accumulation of Hydroxycinnamic Acid Amides Induced by Pathogen Infection and Identification of Agmatine Coumaroyltransferase in *Arabidopsis Thaliana*. *Planta* **230** : 517. <https://doi.org/10.1007/s00425-009-0960-0>.

Nagel, R., Elliott, A., Masel, A., Birch, R. G and Manners, J.M. (1990). Electroporation of binary Ti plasmid vector in *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. *FEMS Microbiology letters* **67** : 325-328

Nirenberg, M., Caskey, T., Marshall, R., Brimacombe, R., Kellogg, D., Doctor, B., Hatfield, D., Levin, J., Rottman, F., Pestka, S., Wilcox, M and Anderson, F. (1966). The RNA Code and Protein Synthesis. *Cold Spring Harbor Symposia on Quantitative Biology* **31** : 11–24. <https://doi.org/10.1101/SQB.1966.031.01.008>.

Nitsch, J. P., Ha, D. B. D and Nitsch, C. (1966). Extraction d'un Facteur de Bourgeonnement de l'Endive (*Cichorium Intybus L.*). " *Bulletin de La Societe Botanique de France* **113** : 425–29. <https://doi.org/10.1080/00378941.1966.10835550>.

Nitsch, J P, and Nitsch, C. (1964). Néoformation de Boutons Floraux Sur Culture in Vitro de Feuilles et de Racines de *Cichorium Intybus L.* — Existence d'un État Vernalisé En l'Absence de Bourgeons. *Bulletin de la Société Botanique de France* **111** : 299-304. <https://doi.org/10.1080/00378941.1964.10835472>.

Onkokesung, N., Gaquerel, E., Kotkar, H., Kaur, H., Baldwin, L.T and 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. <https://doi.org/10.1104/pp.111.187229>.

Osakabe, Y., Watanabe, T., Sugano, S. S., Ueta, R., Ishihara, R., Shinozaki, K and Osakabe, K. (2016). Optimization of CRISPR/Cas9 Genome Editing to Modify Abiotic Stress Responses in Plants. *Scientific Reports* **6** : 26685. <https://doi.org/10.1038/srep26685>.

Ozyigit, I. I., Dogan, I and Tarhan, E. A. (2013). *Agrobacterium rhizogenes*-Mediated transformation and its biotechnologiacal Applications in Crops. *Crop Improvement* <https://doi.org/10.1007/978-1-4614-7028-1>.

Pál, M., Szalai, G and Janda, T. (2015). Speculation: Polyamines Are Important in Abiotic Stress Signaling. *Plant Science* **237** : 16–23. <https://doi.org/10.1016/j.plantsci.2015.05.003>.

Pandino, G., Lombardo, S., Moglia, A., Portis, E., Lanteri, S and Mauromicale, G. (2014). Leaf Polyphenol Profile and SSR-Based Fingerprinting of New Segregant *Cynara Cardunculus* Genotypes. *Frontiers in Plant Science* **5** : 800. <https://doi.org/10.3389/fpls.2014.00800>.

Park, J. J., Yoo, C. G., Flanagan, A., Pu, Y., Debnath, S., Ge, Y., Ragauskas, A. J and Wang, Z. Y. (2017). Defined Tetra-Allelic Gene Disruption of the 4-Coumarate:Coenzyme A Ligase 1 (Pv4CL1) Gene by CRISPR/Cas9 in Switchgrass Results in Lignin Reduction and Improved Sugar Release Mike Himmel. *Biotechnology for Biofuels* **10** : 1–11. <https://doi.org/10.1186/s13068-017-0972-0>.

Peng, H., Meyer, R. S., Yang, T., Whitaker, B. D., Trout, F., Shangguan, L., Huang, J., Litt, A., Little, D. P., Ke, H and Jurick, W. M. (2019). A Novel Hydroxycinnamoyl Transferase for Synthesis of Hydroxycinnamoyl Spermine Conjugates in Plants. *BMC Plant Biology* **19** : 1–13. <https://doi.org/10.1186/s12870-019-1846-3>.

Profumo, P., Gastaldo, P., Caffaro, L., Dameri, R. M., Roti Michelozzi, G and Bennici, A. (1985). Callus Induction and Plantlet Regeneration in *Cichorium Intybus* L.: II. Effect of Different Hormonal Treatments." *Protoplasma* **126** : 215–20. <https://doi.org/10.1007/BF01281797>.

Qin, G., Gu, H., Ma, L., Peng, Y., Deng, X. W., Chen, Z and Qu, L. J. (2007). Disruption of Phytoene Desaturase Gene Results in Albino and Dwarf Phenotypes in *Arabidopsis* by Impairing Chlorophyll, Carotenoid, and Gibberellin Biosynthesis. *Cell Research* **17** : 471–82. <https://doi.org/10.1038/cr.2007.40>.

Rambaud, C., Dubois, J and Vasseur, J. (1992). The Induction of Tetraploidy in Chicory (*Cichorium Intybus* L. Var. Magdebourg) by Protoplast Fusion. *Euphytica* **62** : 63–67. <https://doi.org/10.1007/BF00036089>.

Rambaud, C., Dubois, J and Vasseur, J. (1993). Male-Sterile Chicory Cybrids Obtained by Intergeneric Protoplast Fusion. *Theor Appl Genet* **87** : 347–52.

Rohmer, M. (1999). The Discovery of a Mevalonate-Independent Pathway for Isoprenoid Biosynthesis in Bacteria , Algae and Higher Plants. *Nat. Prod. Rep.* **16** : 565–74.

Saito, M., Kondo, Y and Fukuda, H. (2018). BES1 and BZR1 Redundantly Promote Phloem and Xylem Differentiation. *Plant and Cell Physiology* **59** : 590–600. <https://doi.org/10.1093/pcp/pcy012>.

Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F and Barro, F. (2018). Low-Gluten, Nontransgenic Wheat Engineered with CRISPR/Cas9. *Plant Biotechnology Journal* **16** : 902–10. <https://doi.org/10.1111/pbi.12837>.

Shi, J., Gao, H., Wang, H., Lafitte, H. R. Archibald, R. L. Yang, M., Hakimi, S. M., Mo, H and Habben, J. E. (2017). ARGOS8 Variants Generated by CRISPR-Cas9 Improve Maize Grain Yield under Field Drought Stress Conditions. *Plant Biotechnology Journal* **15** : 207–16. <https://doi.org/10.1111/pbi.12603>.

Short, L., Thoms, A. V., Cao, B., Sinyukov, A. M., Joshi, A., Scully, R and Voronine, D. V. (2015). Facile Residue Analysis of Recent and Prehistoric Cook Stones Using Handheld Raman Spectrometry. *J. Raman Spectrosc* **46** : 126–32. <https://doi.org/10.1002/jrs.4593>.

St-Pierre, B., Laflamme, P., Alarco, A. M., Vincenzo, D and Luca, E. (1998). The Terminal O-Acetyltransferase Involved in Vindoline Biosynthesis Defines a New Class of Proteins Responsible for Coenzyme A-Dependent Acyl Transfer. *The Plant Journal* **14** : 703–13. <https://doi.org/10.1046/j.1365-313x.1998.00174.x>.

Steingroewer, J., Bley, T., Ivanov, I., Lenk, F., Marchev, A and Pavlov, A. (2013). Bioprocessing of Differentiated Plant in Vitro Systems. *Eng. Life Sci* **13** : 26–38. <https://doi.org/10.1002/elsc.201100226>.

Street, R. A., Sidana, J and Prinsloo, G. (2013). *Cichorium Intybus* : Traditional Uses , Phytochemistry , Pharmacology , and Toxicology.” *Evidence-Based Complementary and Alternative Medicine* <http://dx.doi.org/10.1155/2013/579319>

Sun, L. Y., Touraud, G., Charbonnier, C and Tepfer, D. (1991). Modification of Phenotype in Belgian Endive (*Cichorium Intybus*) through Genetic Transformation by *Agrobacterium Rhizogenes*: Conversion from Biennial to Annual Flowering.” *Transgenic Research* **1** : 14–22. <https://doi.org/10.1007/BF02512992>.

Sun, Y., Jiao, G., Liu, Z., Zhang, X., Li, J., Guo, X., Du, W., Du, J., Francis, F., Zhao, Y and Xia, Lanqin. (2017). Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. *Frontiers in Plant Science* **8** : 1–15. <https://doi.org/10.3389/fpls.2017.00298>.

Suzuki, H., Nakayama, T and Nishino, T. (2003). 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** : 1764–71. <https://doi.org/10.1021/bi020618g>.

Tepfer, D. A. (1990). Genetic Transformation Using *Agrobacterium Rhizogenes*. *Physiol. Plant.* **79** : 140–46.

Tiburcio, A. F., Altabella, T., Bitrián, M and Alcázar, R. (2014). The Roles of Polyamines during the Lifespan of Plants: From Development to Stress. *Planta* **240** : 1–18. <https://doi.org/10.1007/s00425-014-2055-9>.

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. <https://doi.org/10.1186/1471-2164-12-236>.

Vasil, I. K and Hildebrandt A. C. (1965). Growth and Chlorophyll Production in Plant Callus Tissues Grown *in Vitro*. *Planta* **68** : 69–82.

Vasil, I. K and Hildebrandt A. C. (1966). Variation of Morphogenetic Behavior in Plant Tissue Cultures in *Chchorium Endivia*. *American Journal of Botany* **53** : 860–69.

Rambaud, C., Dubois, J and Vasseur, J. (1990). Some Factors Related to Protoplast Culture and Plant Regeneration from Leaf Mesophyll Protoplasts of Magdeburg Chicory (*Cichorium Intybus* L Var Magdeburg).*Agronomie* **10** : 767-772

Veena, V and Taylor, C. G. (2007). *Agrobacterium Rhizogenes* : Recent Developments and Promising Applications. *In Vitro Cellular and Developmental Biology - Plant* **43** : 383–403. <https://doi.org/10.1007/s11627-007-9096-8>.

Sinkovic, L., Demsar, L., Znidarcic, D., Vidrih, R., Hribar, J and Treutter, D. (2015). Phenolic Profiles in Leaves of Chicory Cultivars (*Cichorium Intybus* L.) as Influenced by Organic and Mineral Fertilizers ^ Nidarc. *Food Chemistry* **166** : 507–13. <https://doi.org/10.1016/j.foodchem.2014.06.024>.

Vladimirov, I.A., Matveeva, T. V and Lutova, L. A. (2015). Opine Biosynthesis and Catabolism Genes of *Agrobacterium Tumefaciens* and *Agrobacterium Rhizogenes*. *Russian Journal of Genetics* **51** : 121–29. <https://doi.org/10.1134/S1022795415020167>.

Walters, D., Meurer-Grimes, B and Rovira, I. (2001). Antifungal Activity of Three Spermidine Conjugates. *FEMS Microbiology Letters* **201** : 255–58. [https://doi.org/10.1016/S0378-1097\(01\)00278-6](https://doi.org/10.1016/S0378-1097(01)00278-6).

Willeman, H., Hance, P., Fertin, A., Voedts, N., Duhal, N., Goossens, J. F and Hilbert, J. L. (2014). A Method for the Simultaneous Determination of Chlorogenic Acid and Sesquiterpene Lactone Content in Industrial Chicory Root Foodstuffs. *The Scientific World Journal*. <https://doi.org/10.1155/2014/583180>.

Xu, J. G., Hu, Q. P and Liu, Y. (2012). Antioxidant and DNA-Protective Activities of Chlorogenic Acid Isomers. *Journal of Agricultural and Food Chemistry* **60** : 11625–30. <https://doi.org/10.1021/jf303771s>.

Xu, R., Yang, Y., Qin, R., Li, H., Qiu, C., Li, L., Wei, P and Yang, J. (2016). Rapid Improvement of Grain Weight via Highly Efficient CRISPR/Cas9-Mediated Multiplex Genome Editing in Rice. *Journal of Genetics and Genomics* **43** : 529–32. <https://doi.org/10.1016/j.jgg.2016.07.003>.

Yang, L., Wen, K. S., Ruan, X., Zhao, Y. X and Wang, Q. (2018). Response of Plant Secondary Metabolites To Environmental factors. *Molecules* **23** : 1–26. <https://doi.org/10.3390/molecules23040762>.

Yang, Y., Zhu, K., Li, H., Han, S., Meng, Q., Khan, S. U., Fan, C., Xie, K and Zhou, Y. (2018). Precise Editing of CLAVATA Genes in *Brassica Napus* L. Regulates Multilocular Siliques Development. *Plant Biotechnology Journal* **16** : 1322–35. <https://doi.org/10.1111/pbi.12872>.

Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A and Sawada, H. (2001). A Revision of *Rhizobium* Frank 1889, with an Emended Description of the Genus, and the Inclusion of All Species of *Agrobacterium* Conn 1942 and *Allorhizobium Undicola* de Lajudie et Al. 1998 as New Combinations: *Rhizobium Radiobacter*, *R. Rhizogenes*, *R. Rubi*, *R. Undicola* and *R. Vitis*. *International Journal of Systematic and Evolutionary Microbiology* **51** : 89–103. <https://doi.org/https://doi.org/10.1099/00207713-51-1-89>.

Yu, X. H., Gou, J. Y and Liu. (2009). “BAHD Superfamily of Acyl-CoA Dependent Acyltransferases in *Populus* and *Arabidopsis*: Bioinformatics and Gene Expression. *plant Mol Biol* **70** : 421–42. <https://doi.org/10.1007/s11103-009-9482-1>.

Zaenen, I., van Larebeke, N., Teuchy, H., van Montagu, M and Schell, J. (1974). Supercoiled Circular DNA in Crown-Gall Inducing *Agrobacterium* Strains. *Journal of Molecular Biology* **86** : 109–27. [https://doi.org/10.1016/S0022-2836\(74\)80011-2](https://doi.org/10.1016/S0022-2836(74)80011-2).

Zamble, A., Sahpaz, S., Hennebelle, T., Carato, P and Bailleul, F. (2006). N1,N5,N10-Tris(4-Hydroxycinnamoyl)Spermidines from *Microdesmis Keayana* Roots. *Chemistry & Biodiversity* **3** : 982–89. <https://doi.org/10.1002/cbdv.200690107>.

Zenk, M. H., El-Askary, H and Schulte, U. (1975). Anthraquinone Production by Cell Suspension Cultures of *Morinda Citrifolia*. *Planta Med* **28** : 79–101.

Zhao, Y., Wang, J., Ballevre, O., Luo, H and Zhang, W. (2011). Antihypertensive Effects and Mechanisms of Chlorogenic Acids. *Hypertension Research* **35** : 370–74. <https://doi.org/10.1038/hr.2011.195>.

Contribution à la caractérisation du métabolisme des phénolamides chez la chicorée : approches biotechnologiques et édition ciblée du génome

Cichorium intybus, communément nommée chicorée est une plante qui fut, dans un premier temps, consommée pour ses vertus thérapeutiques, puis sa culture s'est peu à peu industrialisée pour la production de substitut du café et la production d'inuline. Chez la chicorée, par une approche de génétique inverse, 2 gènes nommés *CiSHT1* et *CiSHT2*, proches des *Spermidine hydroxycinnamoyl transférases* (*SHTs*), ont été identifiés et semblent impliqués dans la voie de biosynthèse de la tetracoumaroyl spermine au niveau du tapetum des anthères de chicorée. L'expression de ces deux gènes et l'accumulation de tetrahydroxycinnamoyl spermine semblent spécifiques des plantes appartenant à la famille des Astéracées, mais le rôle de ces molécules est encore inexpliqué. Dans le but de valider la fonction de ces deux gènes, la technologie d'édition du génome : CRISPR/Cas 9 a été mise en place chez la chicorée. La validation de cette technologie a été réalisée grâce à la mutation de la *phytoène désaturase* (*CiPDS*) via deux techniques de transformation, la transformation stable par *A. rhizogenes* et la transformation transitoire de protoplastes médierée par le polyéthylène glycol. La transformation stable médierée par *A. rhizogenes* étant plus efficace et la régénération de plantes plus rapide, elle a été utilisée pour générer des plantes mutantes *sht1* et *sht2*. L'étude de ces mutants a permis de valider l'implication de ces deux gènes dans la production de tetracoumaroyl spermine chez la chicorée et de montrer l'action combinée et séquentielle des deux enzymes. La production de cette molécule en systèmes homologue et hétérologues ainsi qu'une étude plus approfondie des mutants permettra peut-être de faciliter la compréhension de l'accumulation originale de cette molécule au niveau du grain de pollen, ainsi que de la valoriser à des fins thérapeutiques. De plus, la mise en place et la maîtrise de la technologie CRISPR/Cas9 chez la chicorée permettra d'étudier et d'améliorer rapidement de nombreux caractères chez cette plante cultivée.

Mots clés : BAHD, acyltransférases, phénolamides, CRISPR/Cas9, *A. rhizogenes*, *C. intybus L.*

Contribution to the characterisation of phenolamide metabolism in chicory: Biotechnological approaches and targeted genome editing

Cichorium intybus, commonly known as chicory, is a plant that was first consumed for its therapeutic properties, then its cultivation was gradually industrialized for the production of coffee substitute and inulin production. In chicory, by an inverted genetic approach, 2 genes named *CiSHT1* and *CiSHT2*, closed to the *Spermidine hydroxycinnamoyl transferases* (*SHTs*), have been highlighted and seem to be involved in the biosynthesis pathway of tetracoumaroyl spermine in the chicory anther tapetum. The expression of these two genes and the accumulation of tetrahydroxycinnamoyl spermine appear to be specific to plants belonging to the Asteraceae family, but the role of these molecules is still unexplained. In order to validate the function of these two genes, genome editing technology: CRISPR/Cas9 was implemented in chicory. Validation of this technology was achieved through the edition of the *phytoene desaturase* gene (*CiPDS*), using *Agrobacterium rhizogenes*-mediated transformation and protoplast transfection methods. The *Agrobacterium rhizogenes*-mediated transformation being more effective and plant regeneration faster, it has been used to generate *sht1* and *sht2* mutant plants. The involvement of these two genes in the production of tetracoumaroyl spermine in chicory has been validated by the study of the mutant plants and has shown the combinatory and sequential action of these two enzymes. The production of this molecule in homologous and heterologous systems and further study of the mutants may facilitate the understanding of the original accumulation of these molecules in pollen grains, as well as its value for industrial applications. In addition, the implementation and the mastery of CRISPR/Cas9 technology in chicory will allow to study and to quickly improve many traits in this cultivated plant.

Key words: BAHD, acyltransferases, phenolamides, CRISPR/Cas9, *A. rhizogenes*, *C. intybus L*