Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

PROTEOMIC AND MOLECULAR STUDY OF SOMATIC EMBRYOGENESIS IN *Coffea* spp.

Tesis que presenta

ANA ODETTH QUINTANA ESCOBAR

En opción al título de

DOCTORA EN CIENCIAS
(Ciencias Biológicas: Biología integrativa)

Mérida, Yucatán, México

2023

CENTRO DE INVESTIGACIÓN CIENTÍFICA DE YUCATÁN, A. C.
POSGRADO EN CIENCIAS BIOLÓGICAS

RECONOCIMIENTO

Por medio de la presente, hago constar que el trabajo de tesis de **Ana Odetth Quintana Escobar** titulado **"Proteomic and molecular study of somatic embryogenesis in *Coffea* spp",** fue realizado en la Unidad de Biología Integrativa, en la línea de investigación de Biología celular y del desarrollo, en el laboratorio No. 24 del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección del **Dr. Víctor Manuel Loyola Vargas**, dentro de la opción de Bioquímica y Biología Molecular, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Atentamente

Dra. Cecilia Hernández Zepeda
Directora de Docencia

Mérida, Yucatán, México, a 04 de diciembre de 2023

DECLARACIÓN DE PROPIEDAD

Declaro que la información contenida en las secciones de: Materiales y Métodos, Resultados y Discusión de este documento, proviene de las actividades de experimentación realizadas durante el período que se me asignó para desarrollar mi trabajo de tesis, en las Unidades y Laboratorios del Centro de Investigación Científica de Yucatán, A.C., y que, a razón de lo anterior, y en contraprestación de los servicios educativos o de apoyo que me fueron brindados, dicha información, en términos de la Ley Federal del Derecho de Autor y la Ley Federal de Protección a la Propiedad Industrial, le pertenece patrimonialmente a dicho Centro de Investigación. Por otra parte, en virtud de lo ya manifestado, reconozco que de igual manera los productos intelectuales o desarrollos tecnológicos que deriven o pudieran derivar de lo correspondiente a dicha información le pertenecen patrimonialmente al Centro de Investigación Científica de Yucatán, A.C., y se regirán, en todo caso, por lo dispuesto por la Ley Federal del Derecho de Autor y la Ley de de Protección a la Propiedad Industrial, en el tenor de lo expuesto en la presente Declaración.

Firma: _____

Ana Odetth Quintana Escobar

Este trabajo se llevó a cabo en la Unidad de Biología Integrativa del Centro de Investigación Científica de Yucatán, y forma parte de los proyectos titulados MODIFICACIÓN DEL GENOMA DE PLANTAS SUPERIORES USANDO CRISPR/CAS9 PARA ESTUDIAR LA DIFERENCIACIÓN CELULAR, número 1515, financiado por el CONACyT, bajo la dirección del Dr. Víctor Manuel Loyola Vargas.

AGRADECIMIENTOS

Al **Centro de Investigación Científica de Yucatán**, por abrirme las puertas para realizar el posgrado; y al **CONAHCyT** por la beca otorgada, sin la cual no habría sido posible mi formación.

Al **Dr. Víctor Manuel Loyola Vargas**, por la confianza y la libertad que me brindó para proponer y ejecutar ideas. Gracias también por motivarme a menudo para continuar en esta carrera recordándome de lo que soy capaz; y por ser un excelente mentor. ¡Ah! Y gracias también por todos los libros.

A la **Dra. Rosa María Galaz Ávalos**, por ser un ser humano excepcional. En ella encontré no solo apoyo técnico y académico, sino una amiga. Cada laboratorio del mundo debería tener una "maestra Rosy".

A mi comité tutoral y de revisión de tesis, por sus valiosos aportes a esta investigación.

Al **Dr. Víctor Aguilar Hernández** y la **M.C. Ligia Brito Argáez**, por su disponibilidad para solventar mis dudas y por el apoyo brindado en la realización de una parte de este proyecto. A la Dra. **Angela Kú González**, por su apoyo ténico con los cortes histológicos y la microscopía confocal.

A mis mejores amigos del posgrado: **Hugo Méndez Hernández**, por su paciencia y apoyo técnico; y por todos los debates académicos (que a menudo terminaban en peleas), ya que me hicieron crecer personal y profesionalmente. **Evelyn Carrillo**, por ser una excelente compañera adoptiva; Eve: eres brillante, carismática, bondadosa y un gran ser humano (recuérdalo siempre). **Pedro Quiam**, quien llegó a sacarme de la rutina; gracias por tantos paseos y dulces.

A mis compañeros del laboratorio 24 con quienes compartí buenos momentos e hicieron más ligero el estrés del doctorado.

DEDICATORIAS

A mi pérdida más reciente, mi abuelo, papá **Nayo**. Su partida me enseñó que no existe el momento perfecto, que no todo saldrá como lo planeamos, y que hay llamadas que no se deben posponer. Fue un verdadero honor haber sido tu nieta; poder escuchar tu risa y tus anécdotas llenas de lucha y superación constante. A mi abuela, mamá **Oli**, que está presente en todos y cada uno de mis días desde el momento de su partida. Si existe la vida eterna quiero encontrarme de nuevo con ustedes y decirles lo mucho que los amo y los eché de menos.

A mis padres, **Ana** y **Juan**. Mamá, no te lo digo a menudo pero te amo con todo mi ser; todo lo que soy y lo que hago, es por y para ti. Papá, gracias por tu ejemplo y tus consejos. La distancia no es fácil, pero aún así me siguen demostrando su apoyo incondicional. Se que están orgullosos de mí, pero soy solo un reflejo de la excelente calidad de padres que tengo. Los amo.

A mi hermana, **Mariam**. Cada día me siento más orgullosa de ti, por tu valentía y tu carisma. Te amo y siempre voy a procurar tu bien.

A mi tío **Ricardo Vidal Gallegos** (QEPD), de quien no pude despedirme por última vez. El inicio de esta travesía en el posgrado fue posible gracias a ti. Hoy que culmino esta etapa te la dedico por haber confiado siempre en mi y en mi familia, por habernos apoyado incondicionalmente hasta el último de tus días. Que sepas que nuestra gratitud y lealtad hacia ti perdurará por esta y mil vidas más.

A **Mimí** y a **Volt**; porque más que mis mascotas, son mis angelitos de cuatro patas. El amor que siento por ellos es algo difícil de explicar. Hasta que no hayas amado a un animal, una parte de tu alma permanece sin despertar (*Anatole France*).

A **Dios**.

PRODUCTOS GENERADOS

I. Artículos Científicos

- ✓ PROTEOMIC APPROACH DURING THE INDUCTION OF SOMATIC EMBRYOGENESIS IN *Coffea canephora*. **Quintana-Escobar A. O,** Bojórquez-Velázquez E., Ruíz-May E., Loyola-Vargas V. M. (Accepted).
- ✓ *IN VITRO* CONVERSION OF *Coffea* spp. SOMATIC EMBRYOS IN SETIS™ BIOREACTOR SYSTEM (2023). Méndez-Hernández H. A., Galaz-Ávalos R. M., **Quintana-Escobar A. O.**, Pech-Hoil R., Collí-Rodríguez, A. M., Salas-Peraza I. Q., Loyola-Vargas V. M. Plants, 12(17), 3055. https://doi.org/10.3390/ plants12173055.
- ✓ CHARACTERIZATION OF THE PIN AUXIN EFFLUX CARRIER GENE FAMILY AND ITS EXPRESSION DURING ZYGOTIC EMBRYOGENESIS IN *Persea americana* (2023). Monroy-González Z., Uc-Chuc M. A., **Quintana-Escobar A. O.**, Duarte-Aké F., & Loyola-Vargas V. M. Plants, 12(12), 2280. https://doi.org/10.3390/plants12122280.
- ✓ AUXIN-CYTOKININ CROSS TALK IN SOMATIC EMBRYOGENESIS OF *Coffea canephora* (2022). Avilez-Montalvo J. R., **Quintana-Escobar A. O.**, Méndez-Hernández H. A., Aguilar-Hernández V., Brito-Argáez L., Galaz-Ávalos R. M., Uc-Chuc M. A., Loyola-Vargas V. M. Plants 11(15), 2013. https://doi.org/10.3390/plants11152013.
- ✓ DIFFERENCES IN THE ABUNDANCE OF AUXIN HOMEOSTASIS PROTEINS SUGGEST THEIR CENTRAL ROLES FOR *IN VITRO* TISSUE DIFFERENTIATION IN *Coffea arabica* (2021). **Quintana-Escobar A.O.,** Méndez-Hernández H.A., Galaz-Ávalos R.M., Elizalde-Contreras J.M., Reyes-Soria F.A., Aguilar-Hernández V., Ruíz-May E., Loyola-Vargas V.M. Plants 10, 2607. https://doi.org/10.3390/plants10122607.

- ✓ GENOME-WIDE ANALYSIS, MODELING, AND IDENTIFICATION OF AMINO ACID BINDING MOTIFS SUGGEST THE INVOLVEMENT OF GH3 GENES DURING SOMATIC EMBRYOGENESIS OF *Coffea canephora* (2021). Méndez-Hernández H.A., **Quintana-Escobar A.O.,** Uc-Chuc M.A., Loyola-Vargas V.M. Plants 10, 2034. https://doi.org/10.3390/plants10102034.
- ✓ CYTOKININS, THE CINDERELLA OF PLANT GROWTH REGULATORS (2019). Márquez-López R.E., **Quintana-Escobar A.O.**, Loyola-Vargas V.M. Phytochem Rev 18, 1387–1408. https://doi.org/10.1007/s11101-019-09656-6.

II. Capítulos de libro:

Quintana-Escobar A.O., Loyola-Vargas V.M., (2024). PROTEOMIC ANALYSIS OF *in vitro* PLANT TISSUES: *Coffea canephora* A CASE OF STUDY, in: Plant Cell Culture Protocols 5th edition, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. New York, NY. (In Press).

Quintana-Escobar A.O., Loyola-Vargas V.M., (2024). TRANSCRIPTOMIC ANALYSIS DURING THE INDUCTION OF SOMATIC EMBRYOGENESIS IN C*offea canephora,* in: Plant Cell Culture Protocols 5th edition, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. New York, NY. (In Press).

III. Actividades de divulgación

- ✓ 2023 Actividades recreativas de ciencia, correspondientes al verano de la ciencia: Mi currículum científico, Ciencia Cakotanú A. C. Voluntaria.
- ✓ 2023 Evaluación de proyectos de investigación de niñas, niños y adolescentes durante las Ferias de Ciencia del Programa Adopte un Talento (PAUTA) 2021-2022. Evaluadora.

- ✓ 2023 Programa Mentorías STEAM, realizando diversas actividades de mentoría dirigidas a mujeres jóvenes de nivel bachillerato del sur del país. Mentora.
- ✓ 2022 Semana Nacional del Conocimiento Yucatán 2022. Organizadora.
- ✓ 2022 Seminario de estudiantes de la unidad de Bioquímica y Biología Molecular de Plantas, CICY. Evaluadora.
- ✓ 2022 Evaluación de proyectos de investigación de niñas, niños y adolescentes durante las Ferias de Ciencia del Programa Adopte un Talento (PAUTA) 2021-2022. Evaluadora. 30 h.
- ✓ 2022 Taller "Vernos como hermanas". Facilitadora de taller. 2 horas. Programa Mujeres en STEAM.
- ✓ 2022 Programa de voluntariado "Día Internacional de la Mujer y la Niña en la Ciencia". Ciencia Cakotanú A. C. Voluntaria.
- ✓ 2021 Participación en el Programa Mujeres en STEAM. 120 horas. Mentora.
- ✓ 2021 Taller “Redefiniéndonos”. Facilitadora de taller. 4 horas. Programa Mujeres en STEAM.
- ✓ 2021 Taller “Heroínas de la ciencia… ¿para los hombres es igual?”. Facilitadora de taller. 4 horas. Programa Mujeres en STEAM.
- ✓ 2021 Taller "Vernos como hermanas". Facilitadora de taller. 2 horas. Programa Mujeres en STEAM.

IV. Congresos

- ✓ **Quintana-Escobar Ana O.**, Hugo A. Méndez-Hernández, Rosa M. Galaz-Ávalos, José M. Elizalde-Contreras, Francisco A. Reyes-Soria, Ligia Brito-Argaez, Víctor Aguilar-Hernández, Eliel Ruíz-May, and Víctor M. Loyola-Vargas, PROTEOMIC AND MOLECULAR STUDY OF SOMATIC EMBRYOGENESIS IN *Coffea* SPP., XXXIII National Meeting of

Biochemistry, Mérida, Yuc., México, 16-21 octubre de 2022. Cartel and Flash talk.

- ✓ **Quintana-Escobar A. O.**, Nic-Can G. I., Galaz-Ávalos R. M., Góngora-Castillo E. B., Loyola-Vargas V. M., TRANSCRIPTOME ANALYSIS OF THE INDUCTION OF SOMATIC EMBRYOGENESIS IN *Coffea canephora* AND THE PARTICIPATION OF ARF AND AUX/IAA GENES. XVlll National Plant Biochemistry and Molecular Biology Congress - XI Symposium México-USA & 1st. ASPB Mexico Section Meeting, November 2019, Mérida, Yuc., México.
- ✓ **Quintana-Escobar A. O.**, Reyes-Soria F. A., Elizalde-Contreras J. M., Ruíz-May E, and Loyola-Vargas V. M. COMPARISON OF THE BIOCHEMICAL STATE BETWEEN in vivo AND in vitro PLANT TISSUES IN *C. arabica* BY PROTEOMIC ANALYSIS, 8th Symposium of the Mexican Proteomics Society, 3rd PanAmerican-Human Proteome Organization (Pan-HUPO) Meeting and 2nd Ibero-American Symposium on Mass Spectrometry. Octubre 2019, Acapulco, Gro., México. 425.
- ✓ Loyola-Vargas Víctor M., Ruíz-May Eliel, Aguilar-Hernández Víctor, **Quintana-Escobar Ana**, Elizalde-Contreras José M., Galaz-Ávalos Rosa M., Brito-Argáez Ligia, Aragón Rodríguez Centeotl, Reyes Soria Francisco A., Avilez-Montalvo Johny. CHARACTERIZATION OF THE CHANGES IN AUXINS, CYTOKININS, AND PROTEINS DURING THE SOMATIC EMBRYOGENIC PROCESS IN COFFEA SPP., AVOCADO, AND COCONUT, 8th Symposium of the Mexican Proteomics Society, 3rd PanAmerican-Human Proteome Organization (Pan-HUPO) Meeting and 2nd Ibero-American Symposium on Mass Spectrometry. Octubre 2019, Acapulco, Gro., México.

V. Actividades de docencia

- ✓ 2023 Curso Tópicos selectos: Embriogénesis somática, impartido en del Posgrado en Ciencias Biológicas (CICY). Docente. 4 horas.
- ✓ 2022 Curso Tópicos selectos: Auxinas, impartido en del Posgrado en Ciencias Biológicas (CICY). Docente. 4 horas.
- ✓ 2021 Curso de Cultivo de Tejidos, impartido en del Posgrado en Ciencias Biológicas (CICY). Docente. 4 horas.

VI. Distinciones

- ✓ 2021 Best paper award for young investigator, by the Plants journal.

Index

INDEX I

FIGURES LIST VII

TABLES LIST XV

INTRODUCTION 1

CHAPTER I 3

BACKGROUND 3

1.1. COFFEE TREE: GENERALITIES 3

1.2. SOMATIC EMBRYOGENESIS IN COFFEE TREE AND THE IMPORTANCE OF GROWTH REGULATORS 4

1.3 AUXIN HOMEOSTASIS IN PLANT CELL DEVELOPMENT AND SOMATIC EMBRYOGENESIS 6

1.4. PROTEOMIC APPROACHES TO STUDY PLANT CELL DEVELOPMENT AND SOMATIC EMBRYOGENESIS 12

1.5. INTEGRATION OF OMIC SCIENCES FOR THE STUDY OF PLANT PROTEOMES . 15

1.6. RESEARCH QUESTION 16

1.7. GENERAL OBJECTIVE 16

1.8. SPECIFIC OBJECTIVES 17

1.9. JUSTIFICATION 17

1.10. EXPERIMENTAL STRATEGY 17

CHAPTER II 21

DIFFERENCES IN THE ABUNDANCE OF AUXIN HOMEOSTASIS PROTEINS SUGGEST A CENTRAL ROLE FOR *IN VITRO* TISSUE DIFFERENTIATION IN *COFFEA ARABICA* 21

ABSTRACT 22

2.1. INTRODUCTION 22

2.2. MATERIALS AND METHODS 24

2.2.1. *IN VITRO* TISSUE GROWTH CONDITIONS 25

2.2.2. PROTEIN EXTRACTION 25

2.2.3. PROTEIN REDUCTION, ALKYLATION, AND DIGESTION 25

2.2.4. PEPTIDE ISOBARIC LABELING WITH TANDEM MASS TAG (TMT) AND FRACTIONATION 26

2.2.5. NANO LC/MS-MS ANALYSIS 26

2.2.6. SYNCHRONOUS PRECURSOR SELECTION (SPS)-MS3 FOR TMT ANALYSIS 27

2.2.7. DATA PROCESSING 27

2.2.8. IDENTIFICATION OF AUXIN HOMEOSTASIS PROTEIN HOMOLOGS 28

2.2.9. PHYLOGENETIC ANALYSIS 28

2.3. RESULTS 29

2.4. DISCUSSION 41

2.5. CONCLUSIONS 46

CHAPTER III 47

PROTEOMIC APPROACH DURING THE INDUCTION OF SOMATIC EMBRYOGENESIS IN *COFFEA CANEPHORA* 47

ABSTRACT 47

3.1. INTRODUCTION 47

3.2. MATERIALS AND METHODS 49

3.2.1. BIOLOGICAL MATERIAL AND GROWTH CONDITIONS 49

3.2.2. PROTEIN EXTRACTION 50

3.2.3. REDUCTION, ALKYLATION, AND DIGESTION 51

3.2.4. NANO LC/MS-MS ANALYSIS 51

3.2.5. DATA PROCESSING 52

3.3. RESULTS 53

3.4 DISCUSION 64

3.5. CONCLUSIONS 69

CHAPTER IV 71

COMPLEMENTARY ANALYZES FOR THE STUDY OF SOMATIC EMBRYOGENESIS IN *C. CANEPHORA* 71

4.1. INTRODUCTION 71

4.2. MATERIALS AND METHODS 72

4.2.1. PROTEIN EXTRACTION BY THE TCA METHOD 72

4.2.2. PROTEIN EXTRACTION BY THE PHENOL-BASED METHOD 72

4.2.3. PROTEIN QUANTIFICATION (PETERSON, 1977) 73

4.2.4. ISOELECTROFOCUSING AND 2DE 73

4.2.5. GEL PROCESSING 74

4.2.6. SAMPLE PREPARATION FOR TRANSCRIPTOMIC ANALYSIS 74

4.2.7. BIOINFORMATIC ANALYSIS 74

4.2.8. AUXIN AND CYTOKININ EXTRACTION AND HPLC QUANTIFICATION 75

4.2.9. CLONING OF SYNTHETIC REPORTER DR5V2 IN *ESCHERICHIA COLI* 75

4.2.10. TRANSFORMATION OF *AGROBACTERIUM TUMEFACIENS* WITH THE SYNTHETIC REPORTER GENE DR5V2 76

4.2.11. TRANSFORMATION OF FOLIAR EXPLANTS OF *C. CANEPHORA* BY AGROINFILTRATION ... 77

4.3. RESULTS 78

4.3.1. PROTEOMIC ANALYSIS THROUGH 2DE 78

4.3.2. TRANSCRIPTOMIC ANALYSIS 84

4.3.3. ENDOGENOUS LEVELS OF PLANT GROWTH REGULATORS AND DR5V2 SYNTHETIC REPORTER ACTIVITY 91

4.4 DISCUSION 98

4.5 CONCLUSIONS 99

CHAPTER V 101

5. GENERAL DISCUSION 101

CHAPTER VI 103

6. CONCLUSIONS AND PERSPECTIVES 103

6.1 CONCLUSIONS 103

6.2 PERSPECTIVES 103

BIBLIOGRAPHY 105

ANNEXES 125

Figures list

Figure 1.1. Comparison between the somatic embryogenesis process in *C. arabica* and *C. canephora* (Nic-Can et al., 2015). **A**) Leaf explants of *C. arabica* during the SE induction. **B**) Leaf explants of *C. canephora* during the SE process. 6

Figure 1.2. Pathways to regulate auxin homeostasis by biosynthesis and conjugation (Salopek-Sondi et al., 2013)..... 7

Figure 1.3. Components in the perception and signaling of auxins and transcriptional regulation. **A**) Transcription activation when auxin levels are high. Ubiquitination of Aux/IAA transcriptional co-regulators and their degradation by the 26S proteasome. **B**) Transcription repression when auxin levels are low. Repression of the transcriptional co-regulators Aux/IAA with TPL proteins. 9

Figure 1.4. Synthetic reporter for auxin response: DR5v2 and DR5. **A**) Components of the plasmid. DR5v2 and DR5 activity in **B**) early globular and **C**) heart-stage embryos (Liao et al., 2015)..... 10

Figure 1.5. Endogenous IAA and IAA conjugate content before and during the induction of somatic embryogenesis in *C. canephora*. Endogenous free IAA (yellow bars), IAA-Ala (blue bars), IAA-Glu (purple bars) (Ayil-Gutiérrez et al., 2013). 11

Figure 1.6. IAA immunolocalization during the development of *C. canephora* somatic embryos (Márquez-López et al., 2018). **A-F**) globular stage. **G-L**) Heart stage. **M-R**) Torpedo stage. **S-X**) Cotiledonary stage. 12

Figure 1.7. Connected papers related to the proteomic studies of somatic embryogenesis and cellular growth in *Coffea* spp. Each node is an academic paper related to the origin paper, arranged according to similarity. Node size is the number of citations. Node color is the publishing year. Similar papers have strong connecting lines and cluster together. ... 14

Figure 1.8. New omics technologies and their applications (Bedia, 2018). 16

Figure 1.9. General diagram of the experimental strategy followed for both systems. 18

Figure 1.10. The experimental strategy followed for the quantitative proteomic analysis based on TMT for *in vitro* cultured tissues of *C. arabica*..... 19

Figure 1.11. The experimental strategy followed for the proteomic analysis of the induction of somatic embryogenesis in *C. canephora*. 20

Figure 2.1. *In vitro* tissues of *C. arabica*. **A**) Starting material: plantlet leaves (P), calli (C) and suspension cultures (S); arrows indicate the comparisons that were made between tissues. **B**) 1D-SDS-PAGE profile of *in vitro* tissues; the molecular mass standard is indicated on the left side of the gel. Two biological replicates were used for each tissue. 29

Figure 2.2. Proteome distribution. Heat map showing the distribution of 2,614 proteins among the different tissue comparisons. P: plantlets. C: calli. S: suspension cultures. Down-accumulated proteins are shown in green. Up-accumulated proteins are shown in red. 30

Figure 2.3. Differentially accumulated proteins. **A**) The number of differentially accumulated proteins (DAPs) among the different tissue comparisons. Up-accumulated proteins are shown in red >1.5; down-accumulated proteins are shown in green <0.66; *P*<0.05). **B**) Venn diagram of the differentially accumulated proteins (DAPs) shared between each comparison. The overlapping regions correspond to the number of shared DAPs. Red and green arrows correspond to the number of up-accumulated and down-accumulated proteins, respectively. P: plantlets. C: calli. S: suspension cultures..... 31

Figure 2.4. Gene ontology analyses of up-accumulated proteins in the CvsP comparison. The proteins identified when comparing calli vs. plantlets were grouped according to **A**) GO enrichment and **B**) hierarchical grouping of the most significant routes based on KEGG. 32

Figure 2.5. Gene ontology analyses of up-accumulated proteins in the SvsP comparison. The proteins identified when comparing suspension cultures vs. plantlets were grouped according to **A**) GO enrichment and **B**) hierarchical grouping of the most significant routes based on KEGG..... 33

Figure 2.6. Gene ontology analysis of up-accumulated proteins in the SvsC comparison.

The proteins identified when comparing suspension cultures vs. calli were grouped according to GO enrichment. Hierarchical grouping of the most significant routes based on KEGG could not be performed. 34

Figure 2.7. Hierarchical grouping tree of the relationship between significantly enriched routes in constitutive proteins. Clustering of the 126 proteins found consistently across the three tissue comparisons. Pathways with many shared genes are clustered. More prominent points indicate more significant *P* values. 35

Figure 2.8. Phylogenetic analysis of the GH3 family involved in auxin conjugation. A phylogenetic tree was constructed to study the phylogenetic relationship of the alignments of 47 GH3 sequences. The red, green, and blue branches represent groups I, II, and III, respectively. GH3 sequences were aligned in MUSCLE. Subsequently, a phylogenetic tree was created using the MEGA7 software. The evolutionary history was inferred using the maximum likelihood method. At: *A. thaliana*, Ca: *C. arabica*, Os: *O. sativa*, Sl: *S. lycopersicum*. 36

Figure 2.9. Phylogenetic analysis of the ILR1 family involved in auxin hydrolysis. A phylogenetic tree was constructed to study the phylogenetic relationship of the alignments of 24 ILR1 (ILL) sequences. The red, green, and blue branches represent groups I, II, and III, respectively. GH3 sequences were aligned in MUSCLE. Subsequently, a phylogenetic tree was created using the MEGA7 software. The evolutionary history was inferred using the maximum likelihood method. At: *A. thaliana*, Ca: *C. arabica*, Os: *O. sativa*, Sl: *S. lycopersicum*. 38

Figure 2.10. Phylogenetic analysis of the ABCB subfamily involved in auxin transporter. A phylogenetic tree was constructed to study the phylogenetic relationship of the alignments of 31 ABCB sequences. The blue, red, and green branches represent the clades I, II, and III, respectively. ABCB sequences were aligned in MUSCLE. Subsequently, a phylogenetic tree was created using the MEGA7 software. The evolutionary history was inferred using the maximum likelihood method. At: *A. thaliana*, Ca: *C. arabica*, Os: *O. sativa*, Sl: *S. lycopersicum.* 39

Figure 2.11. Model of auxin homeostasis during cell differentiation in *C. arabica in vitro*

culture. Participation of proteins identified in this study (marked in red) involved in signaling, transport, conjugation, and β-oxidation of auxin. 44

Figure 3.1. Somatic embryogenesis in *C. canephora*. **A**) From left to right: plantlet and leaf explant 14 and 0 days before induction (dbi); leaf explant at 7 and 21 days after induction (dai). **B**) 1D-SDS-PAGE visualization of the protein profiles from 14 dbi, 0 dbi, 7 dai, and 21 dai samples (10 µg protein). **C**) Flask with somatic embryos released into the culture medium. **D**) Somatic embryos germinated and converted to complete seedlings. 54

Figure 3.2. Proteomic distribution of somatic mbryogenesis induction in *C. canephora*. Heatmap depicting the abundance of the 1630 proteins found among different sampling days of the process. 55

Figure 3.3. Comparison of total proteins among the different sampling days of somatic embryogenesis induction in C. canephora. **A**) Venn diagram showing the distribution of 1630 proteins shared among the different days. **B**) UpsetR plot. The overlapping regions correspond to the number of shared proteins between conditions. Dbi: days before induction. Dai: days after induction. 56

Figure 3.4. Differentially accumulated proteins (DAPs). **A**) Venn diagram showing the total number of DAPs among the different days compared with 14dbi (control): 0dbi/14dbi, 7dai/14dbi, and 21dai/14dbi. The overlapping regions correspond to the number of shared DAPs. **B**) Volcano plots showing the distribution of DAPs among the different sampling days comparisons: 0dbi/14dbi, 7dai/14dbi, and 21dai/14dbi (panels from left to right). Down-accumulated proteins are indicated with blue dots on the left side of the plots. Up-accumulated proteins are indicated with orange dots on the right side of the plots. **C**) Number of DAPs in the different comparisons. **D**) Venn diagram of the up-accumulated DAPs (LFC≥1; P ≤ 0.05). **E**) Venn diagram of the down-accumulated DAPs (LFC≤1; P ≤ 0.05). 57

Figure 3.5. Gene enrichment analysis of up-accumulated DAPs. Gene ontology and KEGG enrichment of the most significant terms and pathways during **A**) 0dbi/14dbi, **B**) 7dai/14dbi and **C**) 21dai/14dbi comparison. 59

Figure 3.6. Protein-protein interaction network of SE-related proteins in *C. canephora*.

Kmeans clustering against *Arabidopsis thaliana* database. Edges represent protein-protein associations. Filled nodes mean that a 3D structure is known or predicted. The names and descriptions provided by STRING for each protein in the network are specified in Annexed 3. 60

Figure 3.7. Accumulation profile of auxin and cytokinin related proteins during SE in *C. canephora*. **A**) Auxin and **B**) Cytokinin-related proteins, among the different sampling days of the SE induction. Blue, black, and pink gradients represent low, intermediate, and high accumulation, respectively. 61

Figure 3.8. Accumulation profile of other SE-related proteins during SE in *C. canephora*. Blue, black, and pink gradients represent low, intermediate, and high accumulation, respectively. 62

Figure 3.9. Accumulation profile of other proteins involved in plant development. **A**) 14-3-3 and **B**) PP2A proteins among the different sampling days of the SE induction. Blue, black, and pink gradients represent low, intermediate, and high accumulation, respectively. 63

Figure 3.10. Participation of some SE-related proteins in different metabolic pathways. Abundances of LOG3, adenine phosphoribosyltransferase 1, ILR1 and tryptophan synthase proteins. 64

Figure 3.11. Model of SE induction in *C. canephora* and the main proteins found to be involved during the process. 67

Figure 4.1 Somatic embryogenesis induction process in *C. canephora*. **A**) Induction. **B**) Development. **C**) Protein extraction. **D**) Protein pattern visualized with 1D- SDS-PAGE (10 µg). 79

Figure 4.2. Electrophoretic pattern of the 2DE analysis of the SE induction process in *C. canephora* at **A**) 14 dbi, **B**) 0 dbi, **C**) 7 dai, and **D**) 21 dai. Red arrows indicate the location of the large (Ls) and small (Ss) subunits of RuBisCO. 80

Figure 4.3. Plot for the 497 proteins of *C. canephora*. Distribution of the proteins identified in the different sampling points of SE induction, according to their molecular mass and isoelectric point. 81

Figure 4.4. Melanie's display area compares the pattern of differentially accumulated proteins (enclosed in green boxes) between the different days of SE induction. 81

Figure 4.5. Global expression levels of 25,574 genes of *C. canephora* during different stages of the process of SE induction. Dbi=days before induction; dai=days after induction. 84

Figure 4.6. Venn diagram of the total number of genes expressed at each stage of the SE induction process. The overlapping regions correspond to the number of genes shared between each point. Dbi=days before induction; dai=days after induction..... 85

Figure 4.7. **A**) Number of differentially expressed genes at different stages of the SE induction. Different comparisons were made by using 14 dbi as control. **B**) Venn diagram of differentially expressed genes between each comparison. The overlapping regions correspond to the number of DEG shared between each condition. Dbi=days before induction; dai=days after induction. 86

Figure 4.8. Gene ontology analysis of up-regulated differentially expressed genes found in **A**) 0 dbi/14 dbi, **B**) 3 dai/14 dbi, **C**) 14 dai/14 dbi, **D**) 21 dai/14 dbi conditions. MF=Molecular function, BP=biological process, CC=cellular component. 88

Figure 4.9. Number of DEGs related to Aux, Ck, and SE. **A**) Venn diagram. **B**) UpsetR plot. The overlapping regions correspond to the number of shared proteins between conditions. Dbi: days before induction. Dai: days after induction. 89

Figure 4.10. Free and conjugated auxin quantification during SE induction in *C. canephora*..... 91

Figure 4.11. Cytokinin quantification during SE induction in *C. canephora*. *tZ*: *trans*-zeatin. KIN: Kinetin. IPA: isopentenyladenine. 92

Figure 4.12. Visualization of the auxin response site by confocal microscopy from transformed embryos of *C. canephora* with DR5v2. **A**) Control embryo. **B-E**) Transformed embryos. 94

Figure 4.13. Embryos transformed with DR5v2 auxin response synthetic reporter placed in

the medium for germination (upper panel). Regenerated seedlings from the transformed embryos (bottom panel). 95

Figure 4.14. PCR confirmation of plasmids in three transformed seedlings: eGFP, ntdTomato, and the internal reference gene (actin). 95

Figure 4.15. Transverse sections of leaves *C. canephora*. **A**) Control. **B**) Transformed. **C**) Transformed, 21 dai. **D**) Transformed, 35 dai. 96

Figure 4.16. Transverse section of somatic embryos. **A** and **B**) Torpedo embryo, control. **C-G**) Cotyledonary embryo, transformed. 98

Tables list

Table 2.1. Identification of cell wall proteins. 40

Table 4.1. Standard curve preparation for protein quantification. 73

Table 4.2. Spots corresponding to differentially accumulated proteins at different points of SE induction were detected with Melanie. 82

Table 4.3. Candidate proteins found in the SWISS-2DPAGE repository. 84

Table 4.4. List of DEGs at different points of SE induction of *C. canephora*, related to Aux, Cks and SE. 90

ABBREVIATIONS

2DE	Two-dimensional gel electrophoresis
ABC	ATP-binding cassette
ABP1	Auxin Binding Protein 1
APTR	Adenine phosphoribosyl transferase
ARF	Auxin Response Factor
ARR	Arabidopsis response regulator
Aux/IAA	Auxin/Indole-3-acetic acid
AuxRE	Auxin-response element
BA	Benzyl adenine
DAI	Days after induction
DBI	Days before induction
GH3	Gretchen Hagen 3
HPLC	High-performance liquid chromatography
IAA	Indole-3-acetic acid
IAM	Iodoacetamide
IBA	Indole-3-butyric acid
IBR	Indole Butyric Response
ILR	IAA leucine resistant
KIN	Kinetin
LC	Liquid chromatography
LEC1	Leafy cotyledon 1
LFC	Log2 fold change
LOG	Lonely guy
MS	Mass spectrometry
NAA	1-naphthaleneacetic acid
PBS	Phosphate-buffered saline

PGR	Plant growth regulators Reguladores de crecimiento vegetal
RPM	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	Somatic embryogenesis
SERK1	Somatic embryogenesis receptor kinase 1
TIR1/AFB	Transport Inhibitor Resistant 1/Auxin signaling F-Box
TMT	Tandem mass tag
TPL	Topless protein

RESUMEN

El género *Coffea* contiene más de 127 especies, de las cuales *C. arabica* y *C. canephora* son las más importantes económicamente a nivel mundial. El cultivo de tejidos es una herramienta que permite estudiar y comprender el mecanismo de diferenciación celular de diversos modelos biológicos. La embriogénesis somática (ES) es una herramienta biotecnológica eficaz para el estudio de los procesos bioquímicos y moleculares llevados a cabo durante el desarrollo de diversas especies vegetales. Este proceso consiste en que las células somáticas, al ser cultivadas en las condiciones adecuadas, dan origen a células embriogénicas que al pasar por procesos morfo-fisiológicos originan embriones somáticos y, posteriormente, plantas completas. En nuestro laboratorio se está optimizando un protocolo para inducir ES de *Coffea arabica* por vía indirecta, es decir, a partir de callo y suspensiones celulares; mientras que para *Coffea canephora*, el proceso estandarizado para la inducción de la ES se realiza de forma directa en los explantes. El proceso de inducción de la ES consta de dos etapas cruciales: preacondicionamiento e inducción. En ambos sistemas, los reguladores del crecimiento vegetal juegan un papel importante para dar lugar a la ES. Por lo anterior, el objetivo general es estudiar el proteoma e identificar las proteínas que se acumulan diferencialmente a lo largo del proceso de inducción y, con base en los resultados, interpretar su posible participación durante la diferenciación celular en la ES de *Coffea* spp. Se encontraron proteínas relacionadas a la homeostasis de auxinas y citocininas, como las ABC, BIG, GH3, ILR, ARR, LOG, SERK1, LEC1, durante diferentes puntos del proceso de ES. El acercamiento transcriptómico y proteómico permitió ampliar el conocimiento sobre el desarrollo celular y embriogénesis somática en nuestros modelos de estudios.

ABSTRACT

The genus *Coffea* includes more than 127 species, of which *C. arabica* and *C. canephora* are the most economically important worldwide. Tissue culture is a tool that allows studying and understanding the mechanism of cell differentiation of different biological models. One of the objectives of this work is to analyze the proteome of different types of *C. arabica* tissues grown *in vitro*. On the other hand, somatic embryogenesis (SE) is a useful biotechnological tool for studying biochemical and molecular processes carried out during the development of various plant species. When cultivated under the right conditions, somatic cells give rise to embryogenic cells that, when they go through morpho-physiological processes, produce somatic embryos and subsequently complete plants. In our laboratory, a protocol is being optimized to induce indirect SE in *Coffea arabica*, that is, from callus and cell suspensions. At the same time, for *Coffea canephora*, the standardized process for SE induction is performed directly on the explants. SE induction consists of two crucial stages: preconditioning and induction. Plant growth regulators (PGR) are essential in developing SE in both biological systems. Thus, this work aims to study the proteome, identify the proteins that accumulate differentially throughout the induction process, and based on the results, analyze their possible participation during cell differentiation and SE induction in *Coffea* spp. Proteins related to auxin and cytokinin homeostasis, such as ABC, BIG, GH3, ILR, ARR, LOG, SERK1, LEC1, among others, were found to be accumulated during different points of the SE process. The transcriptomic and proteomic approach allowed us to expand knowledge about cell development and somatic embryogenesis in our study models.

INTRODUCTION

Coffee is one of the most important crops in world agriculture, and several researches have been conducted to generate knowledge about its genetic improvement. Among the total number of species that compose the genus *Coffea*, only two of them are the most economically important: *C. arabica* and *C. canephora*. For this reason, biotechnological techniques such as plant tissue culture represent an important alternative. Plant tissue culture allows the study, multiplication, and regeneration of plants or parts of them, having the control of different factors.

To promote the different responses in in vitro culture, it is essential to have control of the components of the culture medium, mainly of the plant growth regulator (PGRs). PGRs can generate signaling between cells, tissues and/or organs in the plant and trigger a series of morpho-physiological changes. The most studied PGRs are auxins and cytokinins due to their critical role in plant cell cycle control.

Auxins play a central role in many processes involved in plant growth, such as division, elongation, meristem activity, and root and embryo formation, among others. The most important step for cell growth and differentiation, such as in the somatic embryogenesis (SE) induction, is the increase in the endogenous concentration of auxin and its signaling. This triggers a series of reactions and molecular changes that lead to modifying the genetic program of a somatic cell to become an embryo. These changes are complex, ranging from gene expression and the production of new proteins to changes in the concentration of endogenous PGRs. Therefore, omics sciences are a powerful tool for knowledge about the mechanisms that operate during cell differentiation.

Functional genomics involves different levels: the genome, the transcriptome, the proteome, and the metabolome. However, instead of using them separately as before, it is necessary to combine and integrate them to reinforce the results obtained and the knowledge generated more robustly. In this sense, Proteomics helps elucidate the biochemical and molecular processes that occur during cellular differentiation, thus identifying and quantifying the proteins with differential abundances during the process under study.

CHAPTER I

BACKGROUND

1.1. COFFEE TREE: GENERALITIES

Coffee, after oil, represents the second commercial product with the highest demand worldwide. It is cultivated in over 80 countries on around 11 million hectares in Africa, Asia, and America; sales are estimated at 173,000 million dollars and the intervention of approximately 100 million people during its cultivation and management. Mexico ranks eighth as a producer worldwide, and the primary production is centered in rural areas, which provides livelihoods and a way of life for hundreds of families (Martins et al., 2017; Denoeud et al., 2014; Marraccini et al., 2012).

The genus *Coffea* comprises more than 130 species of evergreen woody trees with differences in size, morphology, and adaptation conditions. Among these species, *C. arabica* and *C. canephora* are the most cultivated (Ferreira et al., 2019) as they represent approximately 60 and 40 % of production (ICO, 2022), respectively.

C. canephora (2n = 2x = 22 chromosomes) is a highly heterozygous allogamous diploid species. It is one of the parents of autogamous *C. arabica* (2n = 4x = 44 chromosomes), the only genus tetraploid, and was derived from hybridization between *C. canephora* and *C. eugenioides. C. canephora* has a higher content of caffeine (but with poor quality) and is more resistant to pests and diseases than *C. arabica* (Etienne, 2005). It develops well in areas with temperatures varying from 22 to 30 °C, 800 meters above sea level, and under precipitation regimes of 1,200 to 2,000 mm per year. In addition, it requires a short dry period to stimulate flowering. For its part, *C. arabica* grows best at higher altitudes, at temperatures between 18 to 21 °C (Denoeud et al., 2014; DaMatta and Ramalho, 2006). Due to the strictest cultivation conditions of *C. arabica*, its production costs are higher than those of *C. canephora*, coupled with a greater risk of pests and diseases, such as coffee leaf rust and coffee berry disease (Van der Vossen, 2015).

Given the importance of this crop, various investigations have been developed related to the increase in mass propagation, its genetic improvement, and improvement in agronomic

qualities such as productivity, grain quality, physicochemical processes, and resistance to biotic and abiotic factors, among others.

Biotechnological techniques, such as *in vitro* plant tissue culture, are an alternative to achieve the purposes mentioned in the previous paragraph. *In vitro* culture allows the aseptic culture of cells, tissues, organs, and their components under controlled conditions (Thorpe, 2012), to produce a large number of uniform plants with characteristics identical to the mother plant and free of diseases in less time than conventional methods (Vinod et al., 2006).

Due to the commercial importance of *C. arabica* and *C. canephora*, several investigations have been carried out worldwide to optimize the production of somatic embryos. However, the SE induction system of *C. arabica* is longer than that of *C. canephora*.

1.2. SOMATIC EMBRYOGENESIS IN COFFEE TREE AND THE IMPORTANCE OF GROWTH REGULATORS

SE is a process that can occur both in nature and in the laboratory under controlled conditions (Loyola-Vargas and Ochoa-Alejo, 2016). When cultivated in the right conditions, somatic cells undergo various biochemical and molecular processes, originating in somatic embryos and whole plants. SE allows easy manipulation and control of culture conditions for the study of morpho-physiological, biochemical, and molecular processes that take place during the development of higher plants, unlike zygotic embryogenesis, in which the embryo is found inside the seed, making it challenging to study (Fehér, 2019; Loyola-Vargas and Ochoa-Alejo, 2016; Winkelmann, 2016; Vondráková et al., 2016). SE is useful for the clonal propagation of various species of commercial importance and as a model for studying the mechanisms that cause embryo formation in plant seeds (Wójcikowska and Gaj, 2017).

SE can be direct or indirect. In direct SE, embryos are formed directly on the explant, while in indirect SE, a set of undifferentiated cells called callus is formed before the formation of embryos (Santana-Buzzy et al., 2007; Söndahl et al., 1985). SE starts from a cell or a group of cells that respond to external stimuli, such as the culture medium, to form a polar structure (Méndez-Hernández et al., 2019; Nic-Can et al., 2016; Nic-Can and Loyola-Vargas, 2016) with two growing points. An essential factor for SE to be carried out is the

culture medium and its components, such as the concentration and type of plant growth regulator (PGR) (Loyola-Vargas et al., 2016). PGRs are compounds in low concentrations that act locally, near or at the site of synthesis, and even in distant tissues (Santner et al., 2009). These can generate a signaling cascade that gives rise to various physiological responses. Among the PGRs known today we can mention indole-3-acetic acid (IAA), cytokinins, abscisic acid (ABA), brassinosteroids, gibberellic acid (GA), ethylene, polyamines, jasmonic acid, nitric oxide, strigolactones and salicylic acid (Vondráková et al., 2016; Santner et al., 2009). However, auxins and cytokinins are the PGRs that mainly, depending on their balance, determine the state of differentiation or de-differentiation of cells.

SE in coffee dates back to the 1970s, when the micropropagation of three species of the genus *Coffea* and the formation of embryos and seedlings from callus were first reported by Staristky (1970). Later, Herman and Haas (1975) induced SE in *C. arabica* from calli generated from foliar explants. Söndahl and Sharp (1977) generated a two-phase protocol for SE from *C. arabica* leaves. Dublin (1981) reported the use of culture medium without auxins but with the addition of cytokinins [benzyladenine (BA) and kinetin (KIN)] for the induction of SE from leaf explants in the Arabusta variety; while Yasuda et al. (1985) obtained embryogenic callus and somatic embryos from *C. arabica* leaves by adding only BA in the culture medium.

Our laboratory has made several efforts to obtain SE in *Coffea* spp. We now have an efficient SE process in *C. canephora* by direct route (Quiroz-Figueroa et al., 2006). This same protocol was used for *C. arabica*; however, obtaining somatic embryos was unsuccessful because the tissue began to oxidize, and cell proliferation stopped after 14 d (Figure 1.1). It is presumed that the accumulation of phenolic compounds in the culture medium could interfere with the development of SE (Nic-Can et al., 2015). Due to the above, it was decided to use a different protocol currently being optimized, which involves the induction of SE through the indirect pathway (Méndez-Hernández et al., 2023) .

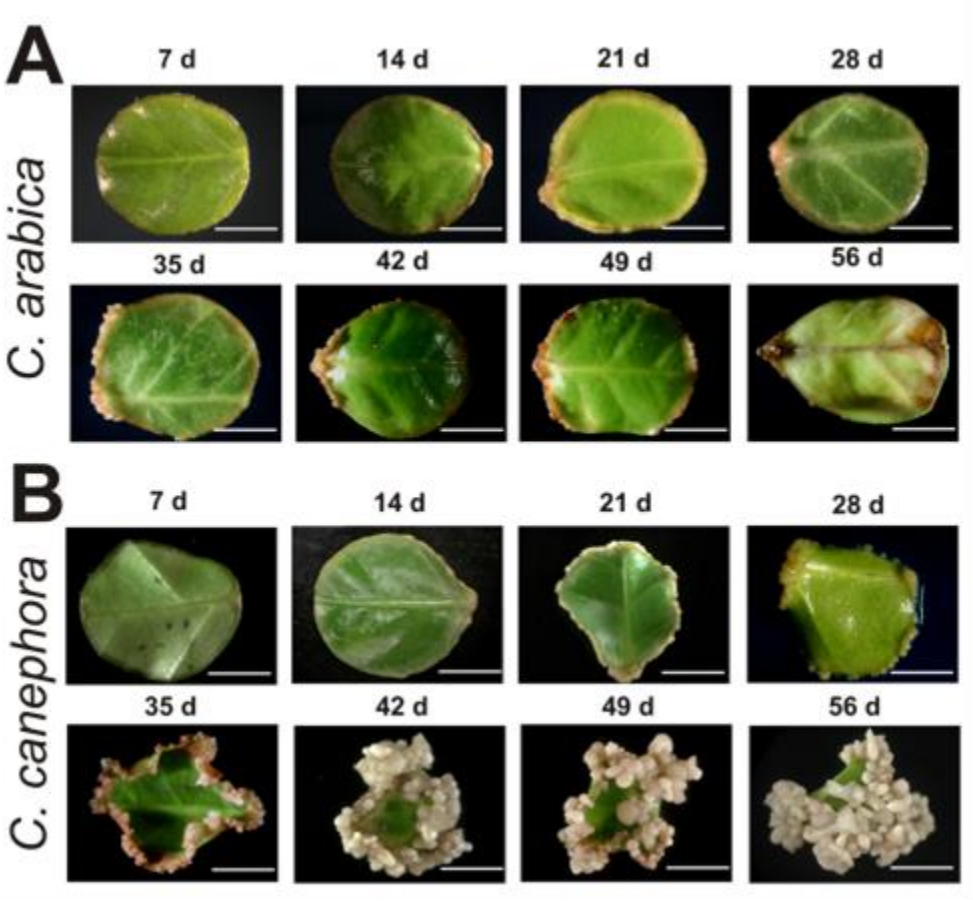


Figure 1.1. Comparison between the somatic embryogenesis process in *C. arabica* and *C. canephora* (Nic-Can et al., 2015). **A**) Leaf explants of *C. arabica* during the SE induction. **B**) Leaf explants of *C. canephora* during the SE process.

1.3 AUXIN HOMEOSTASIS IN PLANT CELL DEVELOPMENT AND SOMATIC EMBRYOGENESIS

Genetic analyses of the biology of auxin have revealed that its synthesis, as well as its transport, signaling, and response, are essential for SE to be carried out (Figure 1.2) (Méndez-Hernández et al., 2019; Nic-Can and Loyola-Vargas, 2016; Weijers and Wagner, 2016). The polar form in which they are transported differentiates auxins from other PGRs, among other properties. Transport can be active or passive. Passive transport occurs through vascular tissue, and the mass flow determines the direction of its movement. Active transport occurs from cell to cell in the vascular cambium and in the xylem of parenchymal cells (Zazimalová et al., 2007).

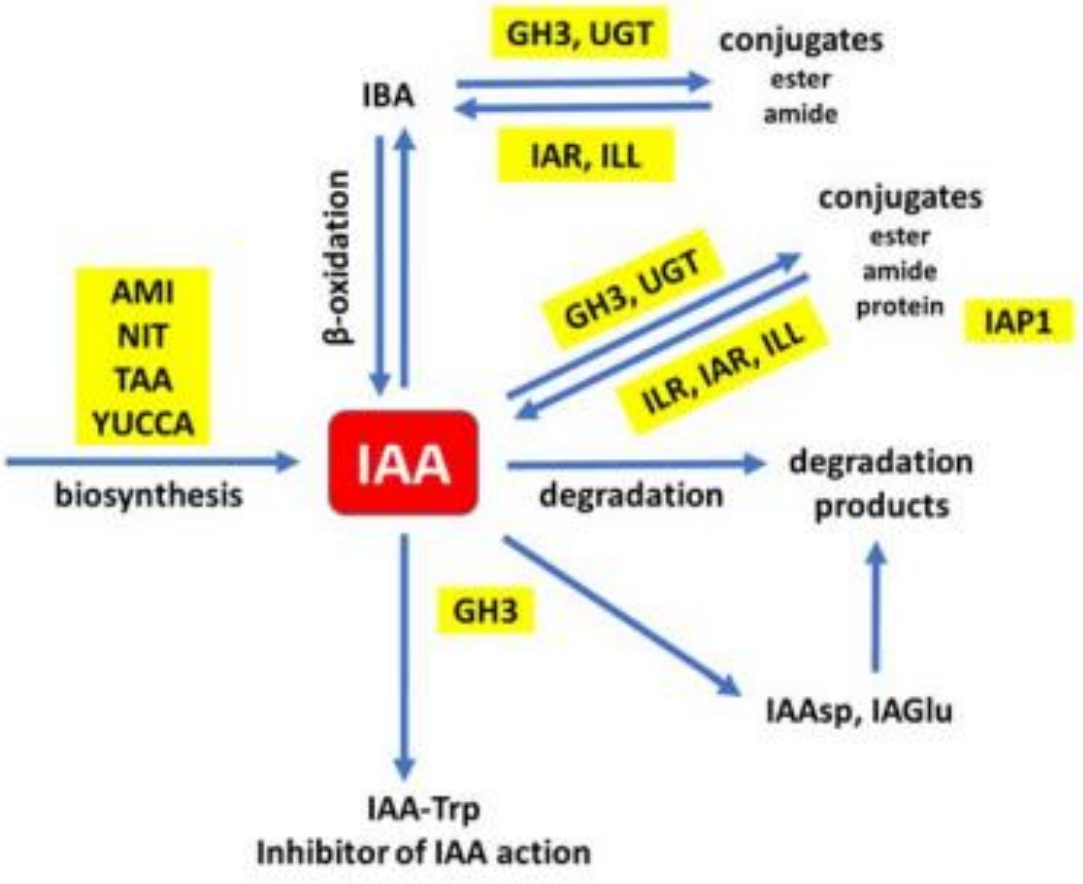


Figure 1.2. Pathways to regulate auxin homeostasis by biosynthesis and conjugation (Salopek-Sondi et al., 2013).

For a biological response to be carried out by the action of endogenous auxins or synthetic compounds, these must be perceived by the plant and converted to a signal. At least two auxin transporters play an essential role in the cellular response identified in Arabidopsis: the PIN protein family and the ABCs (ATP-binding cassette) (Ljung, 2013). Other auxin transporters are PGPs (phosphoglycoproteins), AUX1/LAX, and PIN-like (PILS) proteins (Schaller et al., 2015; Zazimalová et al., 2007).

Characterizing the PIN1 *Arabidopsis thaliana* mutants leads to elucidating the importance of PIN proteins in the auxin flow in plants. However, ABC transporters and phosphoglycoproteins (PGPs) are also involved in auxin flow and are thought to act in conjunction with PINs to regulate the distribution (Schaller et al., 2015; Zazimalová et al.,

2007). The function of the PINs seems to be more directly connected with the specific physiological effects of auxin. In contrast, the function of the PGPs appears to be more general and in areas with a high concentration of auxin. The net effect of the distribution of these transporters is to generate an asymmetry of the auxin that regulates the differentiation and cell division during growth and development of the plant (Schaller et al., 2015; Zazimalová et al., 2007).

Likewise, there are specific carriers of the IAA, such as the AUX1/LAX family and the most recently identified PIN-likes (PILS), which are supposed to intervene in the transport of IAA between the cytosol and the endoplasmic reticulum (ER) (Ljung, 2013).

It is considered that there are at least two auxin receptors: the nuclear co-receptor Transport Inhibitor Resistant 1 (TIR1) and the Auxin Binding Protein 1 (ABP1). During transport, the Transport Inhibitor Resistant 1 (TIR1/AFB Aux/AIA) and Auxin Binding Protein 1 (ABP1) receptors must perceive auxin.

ABP1 is located in the endoplasmic reticulum but is partially secreted into the apoplast, where it activates and binds two auxin molecules as a dimer. At the same time, TIR1 forms a complex with the Aux/IAA protein and auxin. However, the finding of the protein S-Phase Kinase-Associated Protein 2A (SKP2A), which binds to auxin, could act as another receptor of the nucleus (Fendrych et al., 2016; Barbez and Kleine-Vehn, 2013).

The route of TIR1 is carried out by the de-repression of genes induced by auxin. At the same time, it is proposed that the pathway of ABP1 directly regulates protein activities and can act immediately after the perception of auxin. Although the cell growth induced by auxins is a question that has been done for years within the branch of plant physiology, there needs to be accurate data on how auxin is perceived during this process (Fendrych et al., 2016).

The main components of the signaling and regulation mechanism of the expression of the auxin response genes are the TIR1/AFB proteins, the Aux/IAA transcriptional co-regulators, and the ARF binding proteins (Weijers and Wagner, 2016).

When auxin levels are low, Aux/IAA proteins dimer with ARFs to inhibit their activity by binding to the TPL co-repressor (TOPLESS), repressing auxin-responsive genes (Figure

1.3). When high auxin levels, it binds to its receptor and induces the transcriptional regulators' Aux/IAA to attach to the SCFTIR1/AFB complex. Therefore, they are ubiquitinated and degraded by the 26S proteasome (Figure 1.3). In this way, the ARF proteins are free to regulate the transcription of auxin-responsive genes (Li et al., 2016). Given the importance of ARF proteins and Aux/IAA transcriptional regulators in cell development, there is particular interest in determining how and when these genes are expressed. However, the knowledge about ARF proteins in plant models could be improved since practically all this knowledge has been generated in the Arabidopsis model plant.

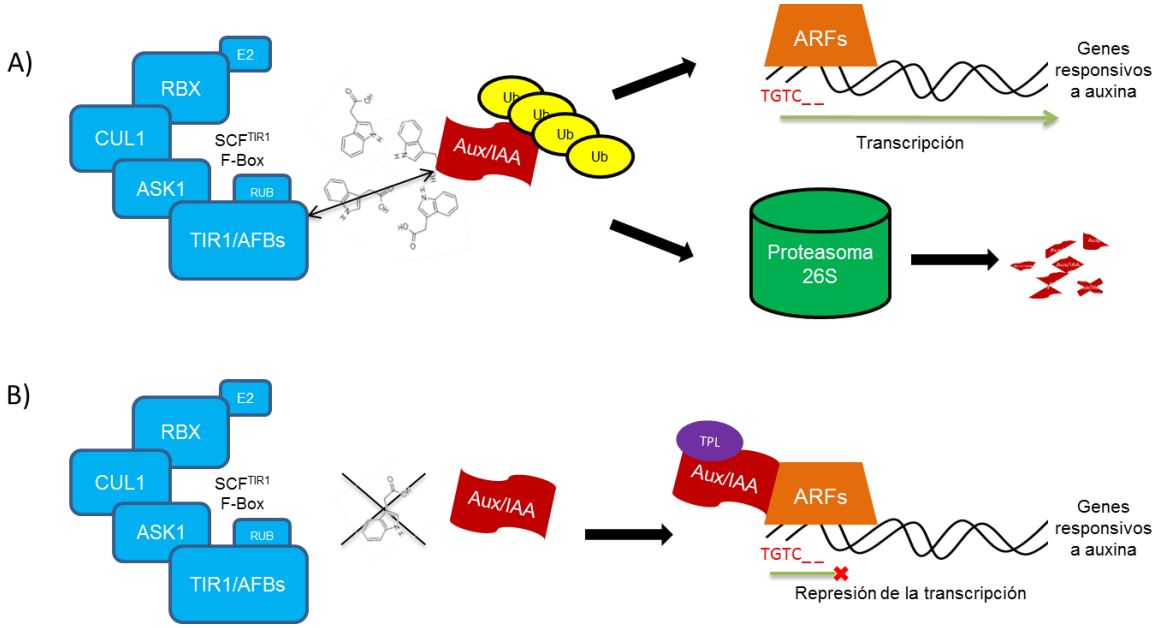


Figure 1.3. Components in the perception and signaling of auxins and transcriptional regulation. **A**) Transcription activation when auxin levels are high. Ubiquitination of Aux/IAA transcriptional co-regulators and their degradation by the 26S proteasome. **B**) Transcription repression when auxin levels are low. Repression of the transcriptional co-regulators Aux/IAA with TPL proteins.

With the generation of synthetic reporters (such as DR5 and DR5v2), it has been possible to study the distribution, transport, and action of auxins in transgenic embryos and seedlings of *A. thaliana* (Dubas et al., 2014; Ni et al., 2001). These reporters act as substrates for auxin receptors and allow the visualization of the transcriptional response (Hayashi et al., 2014; Bargmann et al., 2013). The synthetic DR5 promoter consists of seven to nine TGTCTC AuxRE repeats and marks the transcriptional response sites to

auxin (Figure 1.4) by activating reporter genes such as β-glucuronidase, fluorescent proteins or the gene that codes for luciferase (Liao et al., 2015). The DR5v2 reporter gene is a variant of the DR5 reporter gene, in which a modification in the last two bases increases the affinity of the reporter, as it has shown to be highly sensitive to low concentrations of auxin.

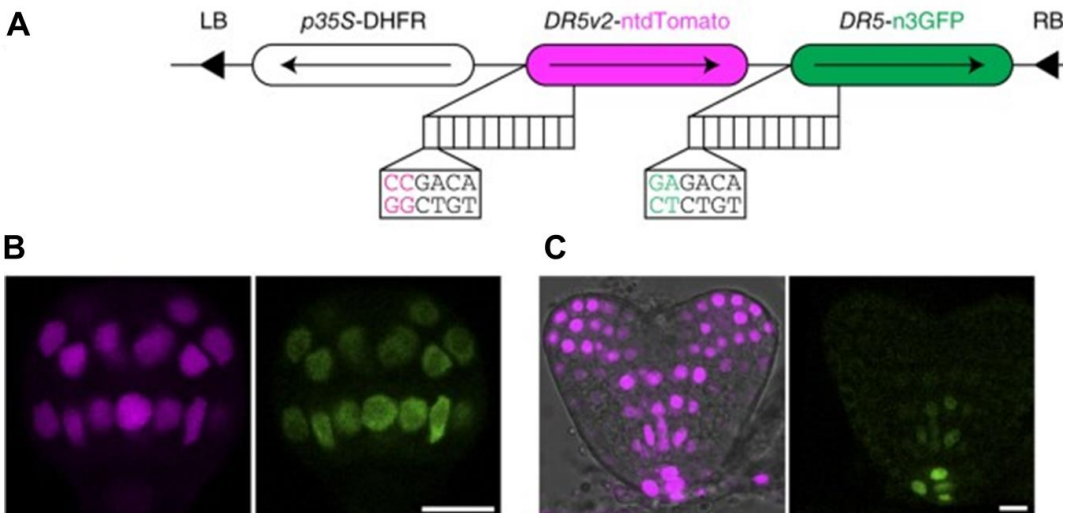


Figure 1.4. Synthetic reporter for auxin response: DR5v2 and DR5. **A**) Components of the plasmid. DR5v2 and DR5 activity in **B**) early globular and **C**) heart-stage embryos (Liao et al., 2015).

In our laboratory, it has been shown that during the preconditioning stage in *C. canephora*, there is an increase in the amount of free and conjugated IAA, specifically in its IAA-Ala and IAA-Glu forms (Figure 1.5). This increase is essential for changing the genetic program of the cells and preparing them for the stage of SE induction.

The increase in IAA before induction suggests that it may be synthesized *de novo* or comes from hydrolysis of the conjugates (Ayil-Gutiérrez et al., 2013). In later studies, the activity of the YUC enzymes was inhibited by the use of yucasin, responsible for the conversion of IPA to IAA. As a result, a decrease in IAA levels was obtained, and the embryogenic response after induction. With this, it was concluded that the biosynthesis of IAA during the SE of *C. canephora* occurs de novo (Uc-Chuc et al., 2020). Similarly, the increase in auxin concentration is accompanied by increased expression of genes related to auxin homeostasis, such as YUCCA, TAA1, GH3, ARF, and Aux/IAA. The GH3.17 and GH3.6 genes are differentially expressed 24 h after induction. This indicates that they can

influence the induction of SE through the conjugation of auxin with acidic amino acids as a substrate, which leads to the degradation of excess auxin (Méndez-Hernández et al., 2021).

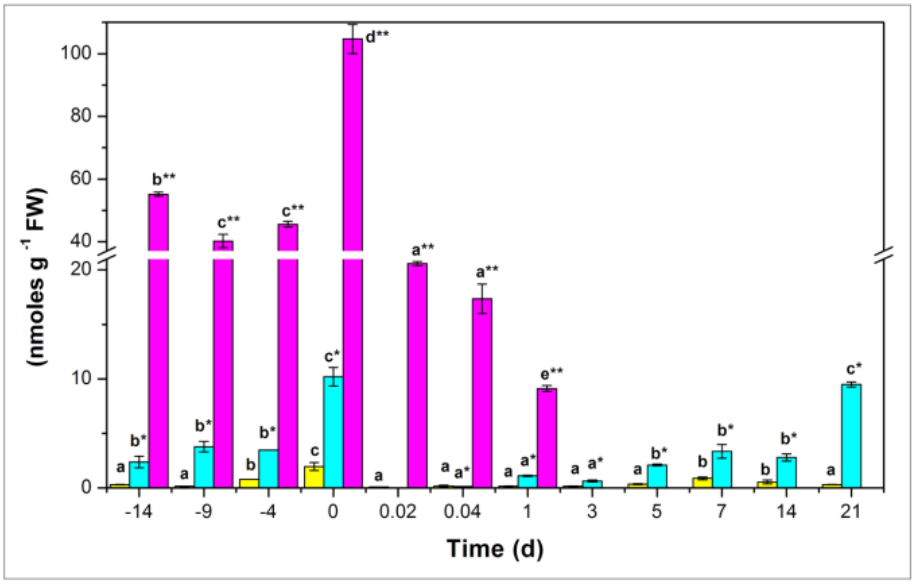


Figure 1.5. Endogenous IAA and IAA conjugate content before and during the induction of somatic embryogenesis in *C. canephora*. Endogenous free IAA (yellow bars), IAA-Ala (blue bars), IAA-Glu (purple bars) (Ayil-Gutiérrez et al., 2013).

Likewise, it was determined that IAA levels gradually increased during pretreatment. Later, during the development of the embryos, the auxin is mobilized towards the growth zones, such as the cotyledons and the root meristem (Figure 1.6). This movement is given utilizing the PIN1 transporter (Márquez-López et al., 2018). However, we still do not know if these accumulation sites are the same as those where auxin is being perceived. Around day 21, when we observe the appearance of the first proembryogenic structures, an increase in the expression of genes involved in auxin perception and signaling, specifically ARF5 and ARF18, is observed, followed by a decrease in Aux/IAA7 and Aux/IAA12 (Quintana-Escobar et al., 2019).

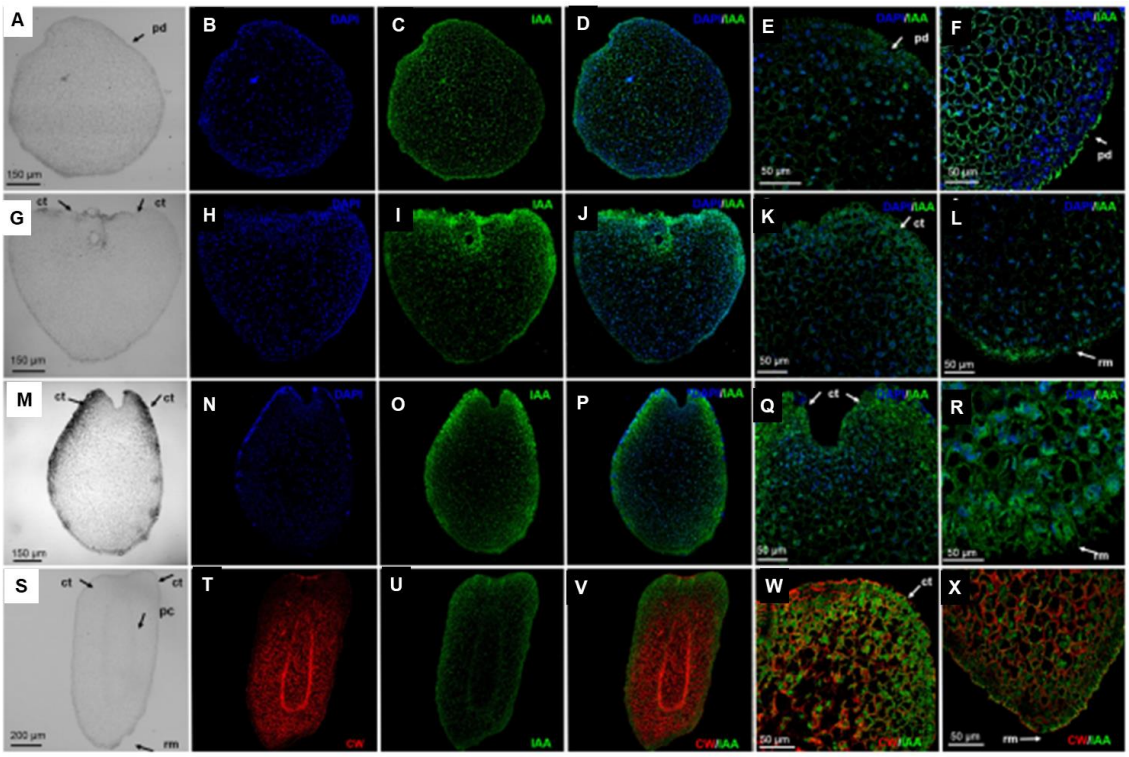


Figure 1.6. IAA immunolocalization during the development of *C. canephora* somatic embryos (Márquez-López et al., 2018). **A-F**) globular stage. **G-L**) Heart stage. **M-R**) Torpedo stage. **S-X**) Cotiledonary stage.

1.4. PROTEOMIC APPROACHES TO STUDY PLANT CELL DEVELOPMENT AND SOMATIC EMBRYOGENESIS

The tools used to study plant secretions and the general strategy adopted depends on the biological model (plant species and/or tissues) and the biological questions to be answered. Nevertheless, the essential steps are sample preparation, protein extraction, reduction, alkylation, digestion, and further analyses.

Plant tissues have specific contaminants that could interfere with proteomic analyses, such as high levels of proteases and secondary metabolites (Wang 2008). Due to this, it is important to mention that the extraction of proteins is a critical step in proteomic studies. In addition, because specific proteins are accumulated in a specialized way in certain types of cells or stages of development, compartmentalized proteins can be extracted through various differential extraction techniques (Agrawal et al., 2010).

Once the proteins are extracted, they are reduced and alkylated. Then, they can be separated by gel-free or gel-based strategies combined with chromatography and MS. For separation by two-dimensional gel electrophoresis (2DE), an isoelectric focusing of the proteins is first made, thereby increasing the resolution power of spots. Subsequently, the gels are digitized to assess the intensity of the points. Another alternative is the two-dimensional difference gel electrophoresis (2D-DIGE), where pre-electrophoretic labeling of samples is carried out. After selecting target spots, peptide separation is achieved by enzymatic action. In the case of a gel-free based strategy, peptide preparation must be performed before the liquid chromatography-mass spectrometry (LC-MS) analysis.

With the improvement of MS technology, gel-free methods such as shotgun have received particular attention. The shotgun method allows the identification of complex peptide fractions after proteolytic digestion. However, this gel-free method should be considered as a complement of 2DE rather than a replacement since there are several points of contrast between these two techniques, such as sample consumption and depth of proteome coverage, analyses of isoforms, and quantitative statistical power, among others (Jorrín-Novo, 2021). The quantification can be done following label-free or label-based methods. Each technique has advantages and disadvantages, so it could be challenging to make a single choice. Therefore, both forms are complementary and must be used in parallel to obtain a more excellent picture of protein expression in a system under specific conditions (Jorrín-Novo et al., 2019; Baggerman et al., 2005).

Post LC-MS/MS analysis, the samples generate thousands of MS/MS spectra, which are submitted to several search engines (MASCOT, Sequest, Phenyx, Sonar, X!Tandem) to predict and compare the relative intensities of the fragment ions (Cottrell, 2011) against different proteome databases such as SwissProt, Trembl, Ensembl, and RefSeq, among others.

There are several reports of studies carried out to elucidate the proteome of *Coffea* spp. during somatic embryogenesis or cell development through different proteomic strategies (Figure 1.7).

In one of the first reported works on proteomics in *C. arabica*, three stages of somatic embryo development were evaluated through 2DE and mass spectrometry (Tonietto et al.,

2012). Specific proteins were identified at each stage. It was proposed that enolase and 11S storage globulin proteins could serve as molecular markers for embryo development and for the differentiation between embryogenic and non-embryogenic lines.

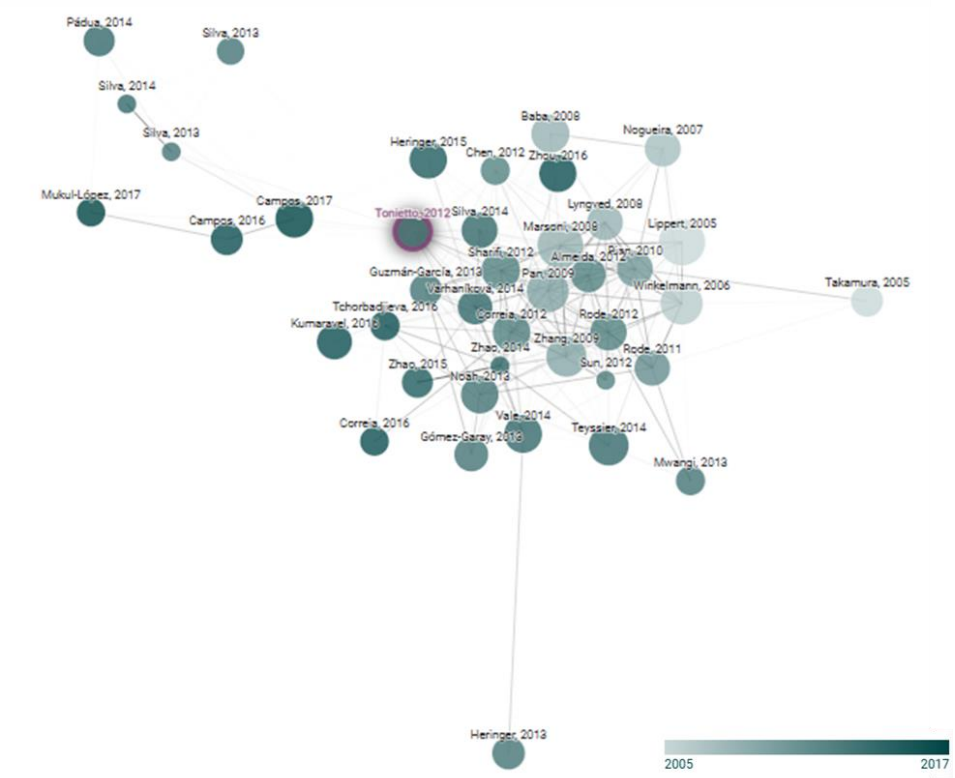


Figure 1.7. Connected papers related to the proteomic studies of somatic embryogenesis and cellular growth in *Coffea* spp. Each node is an academic paper related to the origin paper, arranged according to similarity. Node size is the number of citations. Node color is the publishing year. Similar papers have strong connecting lines and cluster together.

Another work evaluated two embryogenic suspensions through mass spectrometry and shotgun (Campos et al., 2016). The proteome was characterized, and proteins related to stress and energy production was identified. Although the analysis is very descriptive, it lays the groundwork for considering possible molecular markers for SE.

On the other hand, in our laboratory, the first proteomic study was carried out on *Coffea* spp. It compared, through 2DE, the extracellular proteome of embryogenic and non-embryogenic cell suspensions of *C. arabica* and *C. canephora* (Mukul-López et al., 2012). It was found that a larger population of proteins secreted into the medium of *C. canephora*

compared to *C. arabica* and those proteins are secreted exclusively under embryogenic conditions.

1.5. INTEGRATION OF OMIC SCIENCES FOR THE STUDY OF PLANT PROTEOMES

At present, substantial advances in the study of the complexity and diversity of genomes thanks to generating biological data with high-performance technologies and developing computational tools have been made (Goodwin et al., 2016; Iquebal et al., 2015). There is a growing interest in the study and generation of transcriptomic data of the genus *Coffea* (Nic-Can and De-la-Peña, 2014), specifically related to changes in the genetic program that allow a somatic cell to develop into an embryo.

New technologies allow the carry out of studies at the genomic, transcriptomic, proteomic, and metabolomic levels (Figure 1.8) and thus deepen research topics aimed at improving agriculture, the environment, human health, and biotechnology, among others (Goodwin et al., 2016; Wickham, 2016; Iquebal et al., 2015). Nowadays, many studies use transcriptomics to answer various biological questions. It allows for identifying the changes in the expression level of genes of interest in a given condition and, therefore, understanding how the changes in the abundance of the transcripts control the growth and development of an organism (Goodwin et al., 2016; Rhee et al., 2006).

Not all transcribed genes are indeed translated into their functional products. In addition, translation can be affected by many factors, which interfere with the interpretation of results obtained at a precise time (Feussner and Polle, 2015). The proteomic approaches are the best techniques to study global plant proteome on a large scale and with high throughput. Although transcriptomics is widely used and complementary to proteomics, these two approaches sometimes only match. This is because mRNA stability, translational and post-translational regulation, splicing, protein degradation, new protein formation, or a combination of these factors creates complexity (Wu et al., 2019). Because of the increasing number of biological questions and data obtained, it has been necessary to improve biochemical, proteomics, mass spectrometry (MS), and bioinformatics techniques (Yadav et al., 2015).

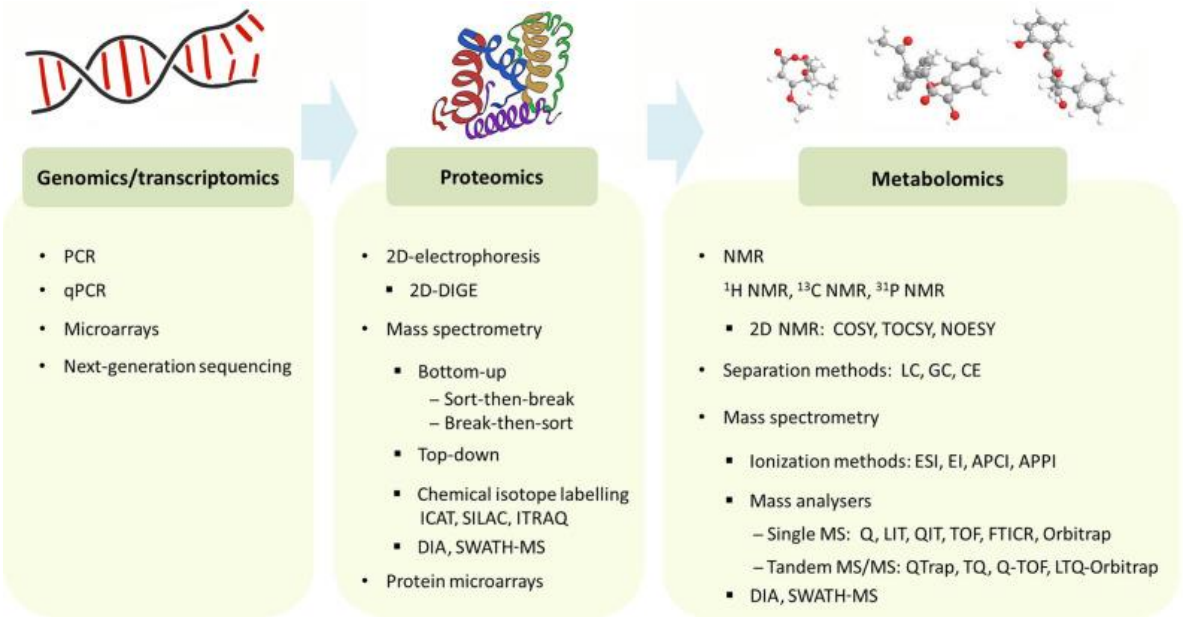


Figure 1.8. New omics technologies and their applications (Bedia, 2018).

However, proteomics allows the quantitative and qualitative analysis of proteins involved in cell development, providing a more precise biochemical state of cells and the changes that occur during their development.

Despite the increased use of proteomic strategies, more is needed to know about the study of *Coffea* spp. A few studies focus on the comparison of embryogenic versus non-embryogenic lines. An officially sequenced genome of *C. arabica* has yet to be made available. Although the genome of *C. canephora* is very useful, many proteins still need to be annotated or characterized, denoting the poor molecular knowledge that there is still about SE in *Coffea* spp. (Campos et al., 2017).

1.6. RESEARCH QUESTION

What signals determine the change of the genetic program of a somatic cell to give rise to an embryogenic cell?

1.7. GENERAL OBJECTIVE

Contribute to understanding the main cellular mechanisms involved in cell growth, differentiation, and somatic embryogenesis in our two standarized *Coffea* spp. systems, through a proteomic approach.

1.8. SPECIFIC OBJECTIVES

Stablishment of a suitable protocol for proteomic studies in different plant tissues of *Coffea* spp.

To identify the differentially accumulated proteins during different *in vitro* tissues in *C. arabica*.

To identify the differentially accumulated proteins during the SE induction process in *C. canephora*.

1.9. JUSTIFICATION

SE is a biotechnological tool within plant tissue culture whose main objectives are crop propagation and genetic improvement. Likewise, SE is an effective model for studying cell differentiation and development and the processes during SE induction. Despite the multiple studies carried out in different biological models, there are still many unanswered questions and parts to be integrated to understand SE fully.

Due to the importance of the coffee tree worldwide, it is a biological model of great interest to study embryogenic transition. New sequencing technologies, quantitative analysis of transcripts, and bioinformatics have been integrated to study SE. However, proteomics offers a more precise approach to what happens in time and space since it evaluates and quantifies the functional products of genes: proteins.

1.10. EXPERIMENTAL STRATEGY

The general strategy for both models consisted of performing the phenolic extraction of proteins and their subsequent processing (reduction, alkylation and digestion) for identification by LC-MS/MS (Figure 1.9). In *C. arabica* the proteomic strategy was based on isobaric labeling by TMT (Figure 1.10); while for *C. canephora* a label-free strategy was used, using the total protein extract and, on the other hand, their separation by 2DE (Figure 1.11).

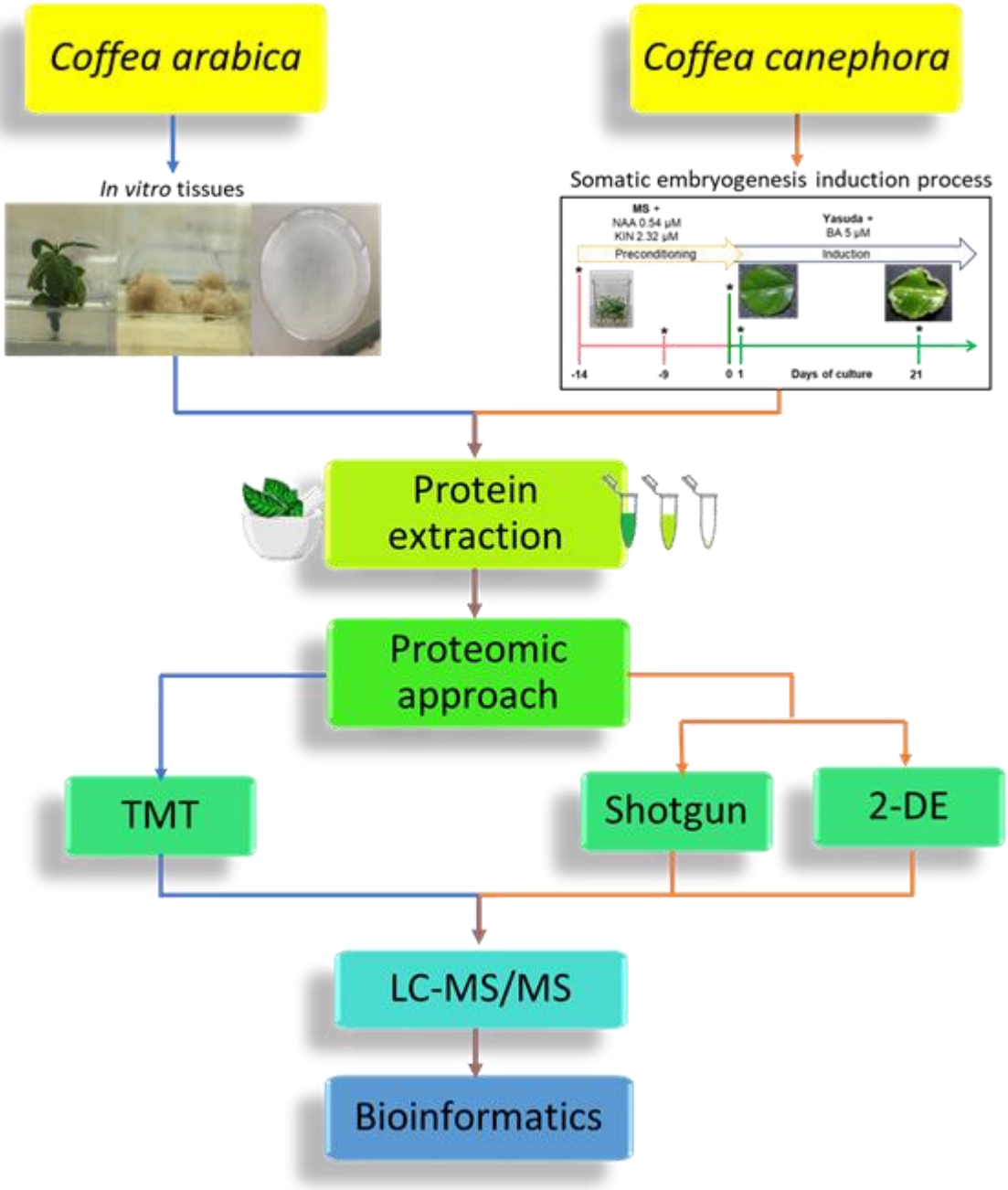


Figure 1.9. General diagram of the experimental strategy followed for both systems.

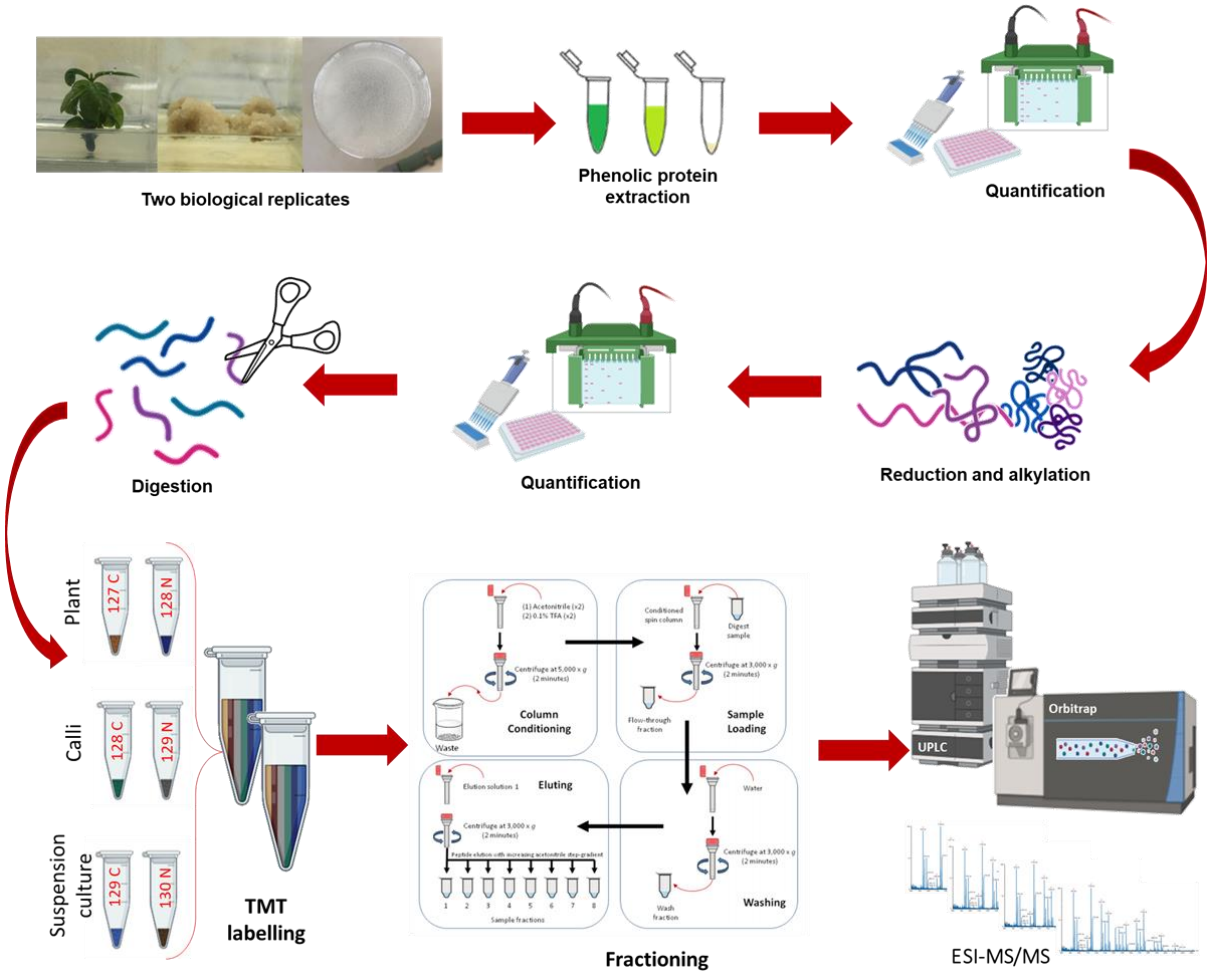


Figure 1.10. The experimental strategy followed for the quantitative proteomic analysis based on TMT for *in vitro* cultured tissues of *C. arabica*.

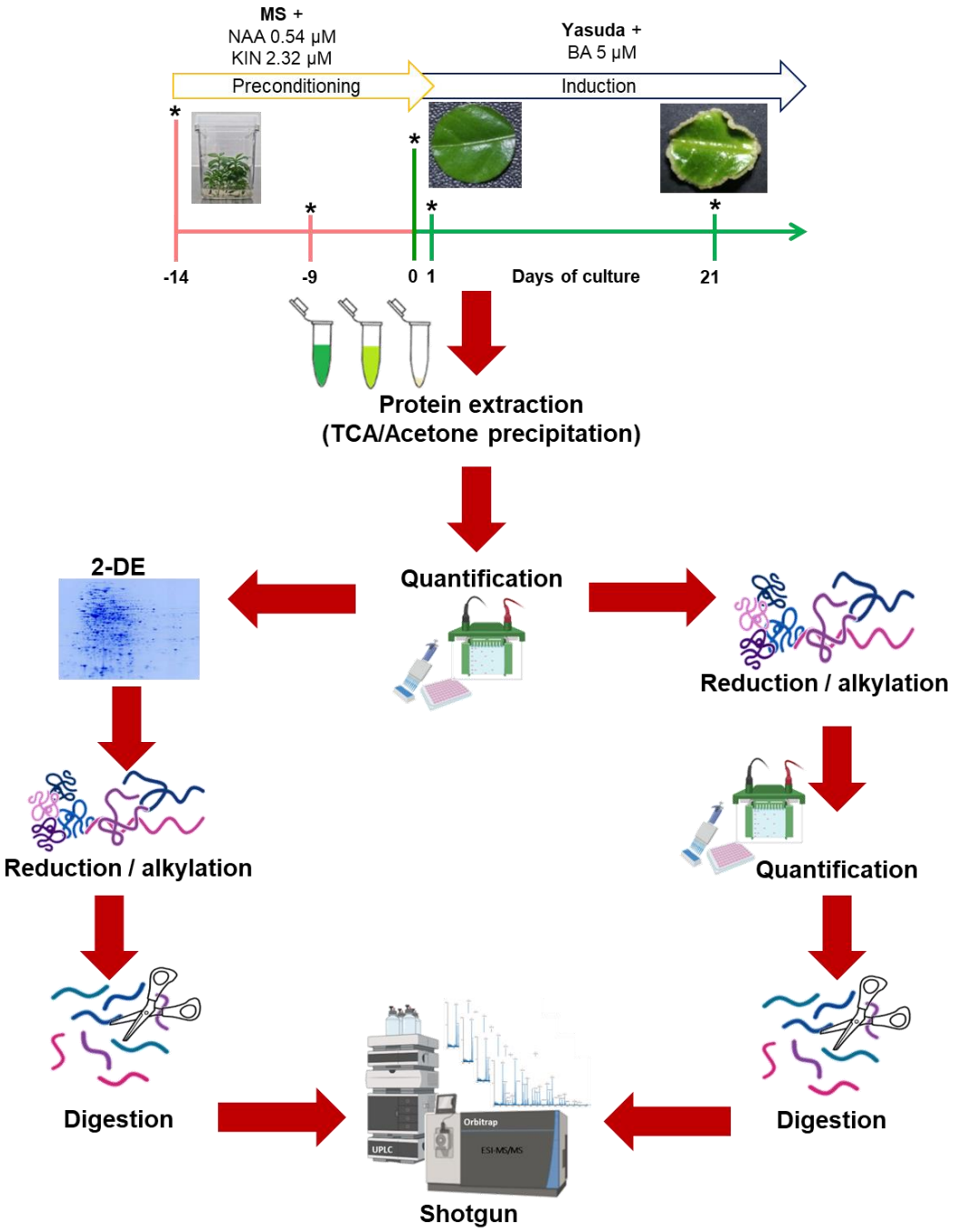


Figure 1.11. The experimental strategy followed for the proteomic analysis of the induction of somatic embryogenesis in *C. canephora*.

CHAPTER II

DIFFERENCES IN THE ABUNDANCE OF AUXIN HOMEOSTASIS PROTEINS SUGGEST A CENTRAL ROLE FOR *in vitro* TISSUE DIFFERENTIATION IN *Coffea arabica*

This chapter was published as a research article and was selected as the **winner** of the **"2021 best paper award for young investigator"** contest carried out by the Plants journal. **Quintana-Escobar, A.O.**; Méndez-Hernández, H.A.; Galaz-Ávalos, R.M.; Elizalde-Contreras, J.M.; Reyes-Soria, F.A.; Aguilar-Hernández, V.; Ruíz-May, E.; Loyola-Vargas, V.M. Plants 10, 2607, (2021). https://doi.org/10.3390/plants10122607.

plants

Differences in the Abundance of Auxin Homeostasis Proteins Suggest Their Central Roles for In Vitro Tissue Differentiation in *Coffea arabica*

by Ana O. Quintana-Escobar [1], Hugo A. Méndez-Hernández [1], Rosa M. Galaz-Ávalos [1], José M. Elizalde-Contreras [2], Francisco A. Reyes-Soria [2], Víctor Aguilar-Hernández [3], Eliel Ruíz-May [2,*] and Víctor M. Loyola-Vargas [1,*]

[1] Centro de Investigación Científica de Yucatán, Unidad de Bioquímica y Biología Molecular de Plantas, Calle 43, No. 130 x 32 y 34, CP, Mérida 97205, Mexico

[2] Instituto de Ecología A.C. (INECOL), Red de Estudios Moleculares Avanzados, Clúster Científico y Tecnológico BioMimic®, Carretera Antigua a Coatepec No. 351, Congregación el Haya, CP, Xalapa 91070, Mexico

[3] Centro de Investigación Científica de Yucatán, Catedrático CONACYT, Unidad de Bioquímica y Biología Molecular de Plantas, Mérida 97205, Mexico

* Authors to whom correspondence should be addressed.

Academic Editor: Shri Mohan Jain

Plants **2021**, *10*(12), 2607; https://doi.org/10.3390/plants10122607

Received: 28 October 2021 / Revised: 16 November 2021 / Accepted: 17 November 2021 / Published: 27 November 2021

(This article belongs to the Special Issue Young Investigators in Plant Sciences)

plants
an Open Access Journal by MDPI

IMPACT FACTOR 4.658 — Indexed in: PubMed

CERTIFICATE

2021 BEST PAPER AWARD FOR YOUNG INVESTIGATOR

This certificate is given to

Ana O. Quintana-Escobar, Hugo A. Méndez-Hernández, Rosa M. Galaz-Ávalos, José M. Elizalde-Contreras, Francisco A. Reyes-Soria, Victor Aguilar-Hernández, Eliel Ruíz-May and Víctor M. Loyola-Vargas

"Differences in the Abundance of Auxin Homeostasis Proteins Suggest Their Central Roles for In Vitro Tissue Differentiation in Coffea arabica"
Plants **2021**, *10*(12), 2607; doi:10.3390/plants10122607

MDPI Academic Open Access Publishing since 1996

Basel, April 2023

Prof. Dr. Dilantha Fernando
Editor-in-Chief

Abstract

Coffea arabica is one of the most important crops worldwide. *In vitro* culture is an alternative for achieving *Coffea* regeneration, propagation, conservation, genetic improvement, and genome editing. The aim of this work is to identify proteins involved in auxin homeostasis by isobaric tandem mass tag (TMT) and the synchronous precursor selection (SPS)-based MS3 technology on the Orbitrap Fusion™ Tribrid mass spectrometer™ in three types of biological material corre-sponding to *C. arabica*: plantlet leaves, calli, and suspension culture. Proteins included in the β-oxidation of indole butyric acid, and in the signaling, transport, and conjugation of in-dole-3-acetic acid were identified, such as the indole butyric response (IBR), the auxin binding protein (ABP), the ATP-binding cassette transporters (ABC), the Gretchen-Hagen 3 proteins (GH3), and the indole-3-acetic-leucine resistant proteins (ILR). A more significant accumulation of proteins involved in auxin homeostasis was found in the suspension cultures vs. the plantlet comparison, followed by callus vs plantlet and suspension culture vs. callus, suggesting greater participation of these proteins as cell differentiation increases.

Keywords: *Coffea arabica*, cellular differentiation, mass spectrometry analysis, plant tissue culture, quantitative proteomics, tandem mass tag.

2.1. INTRODUCTION

Coffee is one of the most important crops worldwide. The genus *Coffea* is composed of more than 130 species, of which *Coffea arabica* and *C. canephora* are the most economically important (Campos et al., 2017), as they represent around 60 and 40% of world coffee production, respectively (ICO, 2022). The market has a high demand for *Coffea* spp. plants to meet the growing demand for coffee production. These plants must be of high quality and resistant to the range of diseases that affect the genus. Vegetative propagation is preferred for large-scale production of elite plants (Devasia et al., 2020) to ensure the maintenance of the desired characteristics of the mother plant. This type of propagation also allows the development of high homogenous resistance to pests and diseases in a short time and limited space (Ebrahim et al., 2007). Therefore, plant tissue culture has long been a preferred alternative for the conservation (Pereira-Dias et al., 2020), propagation, and genetic improvement of recalcitrant cultures.

Somatic embryogenesis (SE) has been positioned as an alternative and effective method for propagation compared to conventional methods either by seed or by cuttings (Mwaniki et al., 2019), whether it is for commercial or research purposes. SE can be achieved directly on the explant or indirectly through disorganized tissue. The first route is of low frequency since the number of embryos obtained is lower, while the second is of high frequency, and is preferred for achieving mass propagation (Ducos et al., 2007). Currently, SE is a useful biotechnological tool for propagation, genetic transformation, and genetic improvement, as well as for basic research on the molecular mechanisms underlying SE (Loyola-Vargas and Ochoa-Alejo, 2016). The study of SE in coffee has been carried out since 1970 (Staritsky, 1970); the two main foci of research since then have been to improve the methodology and to understand the mechanism by which somatic embryos are obtained (Loyola-Vargas et al., 2016).

Because SE in coffee can be started from different tissues, such as suspension cultures (Campos et al., 2016; Landey et al., 2013; Quiroz-Figueroa et al., 2002b), calli (Wang et al., 2018), leaves (Quiroz-Figueroa et al., 2006), and others (Loyola-Vargas et al., 2016), one critical element to understand is the SE induction mechanism. Suspension cultures are an effective substrate for metabolism research since they can be synchronized (Zhang et al., 2019). Calli are a group of disorganized cells, useful for genetic transformation studies (Zhang et al., 2019). Differentiated organs, such as leaves, contain different types of cells and are preferred for starting SE due to their availability and also because the embryos are obtained faster, although in smaller number (Devasia et al., 2020; Etienne et al., 2018; Campos et al., 2017).

Plant growth regulators play an essential role in all aspects of plant growth and development (Santner et al., 2009), of which auxins are one of the most important. Auxins have a primary role in cell division, elongation, differentiation, organogenesis, embryogenesis, and response to external stimuli, as well as in the formation of cells and tissues (Ljung, 2013). For those processes to be carried out, regulation of biosynthesis, conjugation, transport, and signaling is required, which integrates auxin homeostasis (Cazzonelli et al., 2013). Biological processes such as directional transport and the formation of auxin gradients are achieved by the different input transporters such as the AUXIN RESISTANT1/LIKE AUX1 (AUX1/LAX) (Péret et al., 2012), and output transporters

such as the PINs (Adamowski and Friml, 2015) and the ABCB (ATP-BINDING CASSETTE subfamily B) (Jenness et al., 2019). Regarding their conjugation, several genes that code for indole-3-acetic acid (IAA) amido synthetases have been identified, such as those belonging to group II of the *GH3* family (Westfall et al., 2012).

Different conjugates of IAA have been implicated in various biological processes; for example, IAA-Glu is synthesized during induction of SE in *Coffea canephora* and is considered a precursor of auxin degradation (Méndez-Hernández et al., 2021). However, conjugates such as IAA-Leu and IAA-Ala can be hydrolyzed to release free IAA through amido hydrolase enzymes encoded by genes of the ILR1 family (LeClere et al., 2002).

The use of omics sciences has been crucial for the study of plant development in recent years (Campos et al., 2017). With modern omics tools, it is possible to understand the molecular mechanism that leads to the formation of embryos, starting from somatic cells. Proteomics is a valuable tool for studying protein levels during plant development (Zhang et al., 2019) and somatic embryogenesis (Aguilar-Hernández and Loyola-Vargas, 2018). Various proteomic studies on differentiation in *Coffea* species have been carried out using different types of initial explants, such as suspension cultures (Campos et al., 2016), calli (Mwaniki et al., 2019) and leaves (Tonietto et al., 2012). In proteomic techniques, the differential labeling of peptides with isobaric tags, such as the tandem mass tag (TMT), reduces handling and analysis time, allowing the quantification of peptides by measuring the intensity of the reporter ion (Bindschedler and Cramer, 2011). Furthermore, the application of synchronous precursor selection (SPS)-MS^3 technology, available as a hybrid platform in the Orbitrap Fusion™ Tribid, provides the means of eliminating contaminants by isolating near-isobaric ions that fragment together with the target ions (Ting et al., 2011). This was the approach used in this study, as it yields robust comparative proteomics data without ratio distortion in isobaric multiplexed quantitative proteomics. Likewise, proteomic analysis allows the identification of molecular markers associated with *in vitro* morphogenesis (Wendt dos Santos et al., 2018). This work aims to identify which proteins involved in auxin homeostasis may be involved in the process of cell differentiation by comparing calli, suspension cultures, and plantlet leaves from *C. arabica*.

2.2. MATERIALS AND METHODS

2.2.1. *In vitro* tissue growth conditions

The samples used for proteomic analysis consist of *in vitro* plantlet leaves (P), calli (C), and suspension cultures (S) of *C. arabica*. The establishment and maintenance of the biological material were carried out according to the methodology previously reported by Quiroz-Figueroa (2002a).

2.2.2. Protein extraction

In order to extract the proteins, the plant tissue was triturated to obtain a fine powder, using a mortar and liquid nitrogen. The extraction buffer included 0.5 M Trizma base (pH 8; Sigma, T1503), 50 mM EDTA (pH 8; Sigma, EDS), 700 mM sucrose, 100 mM KCl (Sigma, P9541), 2% β-mercaptoethanol (Sigma, M6250), 1 mM PMSF (Sigma, 78830), 1% SDS (Sigma, L3771) and a protease inhibitor cocktail (Sigma, P9599). For each 100 mg of sample in Eppendorf tubes, 1 mL of extraction solution was added, and it was vortexed for 5 min, with 1-min rest intervals. Subsequently, 1 mL of phenol solution (Sigma, P4557) was added in the fume hood to tubes sealed with parafilm to avoid spillage. The tubes were briefly vortexed and placed on ice with shaking for 20 min. Then they were centrifuged at 4 °C and 15,000 x g for 30 min. The upper phenolic phase was recovered. The volume of each tube was increased to 2 mL with acetone supplemented with 0.07% β-mercaptoethanol (Sigma, M6250), and they were allowed to precipitate overnight at -20 °C. The next day, the tubes were centrifuged at 4 °C and 3,000 x g for 30 min. The supernatant was discarded, and the samples were allowed to dry in the vacuum concentrator. The pellet was resuspended in 300 µL of 1x PBS (Sigma, P5493) supplemented with 1% SDS (Sigma, L3771) by vortexing for 15 min. The tubes were centrifuged at 24 °C, and 15,000 x g for 10 min, and the supernatant was recovered in new tubes. The quantification of the total protein was carried out with the BCA Protein Assay Kit (Thermo Scientific, 23227), and the quality of the extract was verified by SDS-PAGE. Samples were stored at -80 °C until use.

2.2.3. Protein reduction, alkylation, and digestion

One hundred µg of protein was taken from the previous extract, and the volume was increased to 100 µL with PBS supplemented with 1% SDS solution. For the protein

reduction, 10 mM TCEP (Sigma, 68957) was added and incubated for 45 min at 60 °C. Subsequently, the proteins were alkylated for 60 min with 30 mM IAM (Sigma, A3221) in the dark at room temperature. Then 30 mM DTT (Sigma, D9779) was added, and it was incubated for 10 min at room temperature. Cold acetone was added to the tubes and incubated overnight at -20 °C to precipitate the proteins. Next, the tubes were centrifuged at 10,000 x g for 15 minutes at 4 °C. The supernatant was discarded, and the pellet was dried in a vacuum concentrator. The dry pellet was resuspended with 50 mM TEAB (Sigma, T7408) supplemented with 0.1% SDS (Sigma, L3771). Finally, the protein content was quantified again with the BCA Protein Assay Kit (Thermo Scientific, 23227) and visualized on SDS-PAGE. Samples were stored at -80 °C until use. Proteins were digested with trypsin (Thermo Scientific, 90058) 1:30 (trypsin:protein) overnight at 37 °C, followed by incubation with trypsin 1:60 at 37 °C for 4 h. Afterward, samples were vacuum-dried.

2.2.4. Peptide isobaric labeling with tandem mass tag (TMT) and fractionation

The TMT Isobaric Label Reagent Set plus TMT11-131C kit (Thermo Scientific, A34808) was used to perform the isobaric labeling. Two biological replicates were used for each tissue. Peptides were labeled with 127C and 128N tags for peptides from leaves; 128C and 129N tags for peptides from callus; and 129C and 130N tags for peptides from suspension cultures. After protein labeling, peptide fractionation was carried out with Pierce™ High pH Reversed-Phase Peptide Fractionation Kit (Thermo Scientific, 84868).

2.2.5. Nano LC/MS-MS analysis

Samples were analyzed by nano LC-MS/MS analysis using an Orbitrap Fusion Tribid (Thermo-Fisher Scientific, San Jose, CA) mass spectrometer equipped with an “EASY spray” nano ion source (Thermo-Fisher Scientific, San Jose, CA). The Orbitrap Fusion Tribid (Thermo Scientific, San Jose, California, USA) was interfaced with an UltiMate 3000 RSLC system (Dionex, Sunnyvale, CA). Each sample was reconstituted with 0.1% formic acid in LC-MS grade water (solvent A; Thermo Scientific, 85178), and 5 µL was injected into a nanoviper C_{18} trap column (3µm, 75 µm X 2 cm, Dionex) at 3 µL min^{-1} flow rate, and then separated with a 100 min gradient on an EASY spray C_{18} RSLC column (2 µm, 75 µm x 25 cm), with a flow rate of 300 nL min^{-1}, and using solvent A and 0.1% formic acid in

90% acetonitrile (solvent B). The gradient was as follows: 10 min solvent A, 7-20% solvent B for 25 min, 20% solvent B for 15 min, 20-25% solvent B for 15 min, 25-95% solvent B for 20 min, and eight min solvent A. The mass spectrometer was operated in positive ion mode with nanospray voltage set at 3.5 kV and source temperature at 280 °C. External calibrants included caffeine, Met-Arg-Phe-Ala (MRFA) and Ultramark 1621 (88323, Thermo Fisher ScientificTM PierceTM).

2.2.6. Synchronous precursor selection (SPS)-MS3 for TMT analysis

Full MS scans were run in the Orbitrap analyzer with 120,000 (FWHM) resolution, scan range 350-1500 m/z, AGC of 2.0e5, maximum injection time of 50 ms, intensity threshold 5.0e3, dynamic exclusion one at 70s and 10 ppm mass tolerance. For MS2 analysis, the 20 most abundant MS1s were isolated with charge states set to 2–7. Frag-mentation parameters included collision-induced dissociation with collision energy set to 35% and an activation Q of 0.25, an AGC of 1.0e4 with a maximum injection time of 50 ms, a precursor selection mass range of 400-1200 m/z, a precursor ion exclusion width of a low of 18 m/z and a high of 5 m/z, isobaric tag loss TMT and detection run in the ion trap. Afterward, MS3 spectra were acquired as previously described (McAlister et al., 2014) using synchronous precursor selection (SPS) of 10 isolation notches. MS3 precursors were fragmented by HCD with 65% of collision energy and analyzed using the Orbitrap with 60,000 resolution power at 120-500 m/z scan range, a two m/z isolation window, 1.0e5 AGC, and a maximum injection time of 120 ms with one microscan.

2.2.7. Data processing

The resulting MS/MS data were processed using the MASCOT (v.2.4.1, Matrix, Science, Boston, MA) search engine implemented in Proteome Discoverer 2.1 (Thermo Fisher Scientific). *C. canephora* and viridiplantae Swiss-Prot databases were used for MS/MS data analysis. Search parameters included 10 ppm and 0.6 Da mass tolerance, trypsin digestion with two missed cleavages allowed. Static modifications included cysteine carbamidomethylation, N-terminal TMT6plex, and lysine TMT6plex. Dynamic modifications included methionine oxidation and arginine/glutamine deamidation. The global false discovery rate (FDR) was up to 1%. Differentially abundant proteins were determined as a fold change ≥1.5 for those up-accumulated or ≤0.66 for those down-accumulated, and $P <$

0.05 was used to identify statistical significance. Functional annotation and GO classification of all identified proteins were determined with Blast2GO software against the viridiplantae NCBI Swiss-Prot database, with a default functional annotation pipeline (Götz et al., 2008). The heatmaps were generated using the ggplot2 package for R (Wickham, 2016). The InteractiveVenn tool was used to create the Venn diagram (Heberle et al., 2015). The hierarchical grouping of the KEGG enrichment was carried out with the ShinyGO V0.66 online platform (http://bioinformatics.sdstate.edu/go/).

2.2.8. Identification of auxin homeostasis protein homologs

The identification of possible homologs of genes that encode proteins involved in the transport and catabolism of auxins was carried out by *in silico* analysis. The families of genes and proteins selected for this study were: ABCB (ATP-BINDING CASSETTE subfamily B) involved in auxin transport; GRETCHEN HAGEN 3 (GH3), and IAA LEUCINE RESISTANT1 (ILR1) involved in auxin catabolism. We used each of our *C. arabica* protein sequences to perform a BLAST analysis against The Arabidopsis Information Resource database (TAIR; accessed on September 7th, 2021) to determine possible homologs. The sequences with the highest percentage of identity and similarity were considered homologous with Arabidopsis.

2.2.9. Phylogenetic analysis

For the phylogenetic analysis, the sequences of each protein family were aligned using the "MUSCLE" tool within the MEGA7 software (http://www.megasoftware.net/; accessed on September 10th, 2021). The aligned sequences were trimmed for non-aligned residues within regions of more significant variability. The best evolutionary model was determined in each protein family using the tool "Find Best DNA/Protein Models." Phylogenetic trees were constructed using the Maximum Likelihood method based on the JTT matrix-based model, with a bootstrap analysis of 100 replicates. A phylogenetic tree was built using the MEGA 7 software (http://www.megasoftware.net/; accessed on September 1st, 2021). The sequences of rice were obtained from http://riceplantbiology.msu.edu and NCBI (https://www.ncbi.nlm.nih.gov/; accessed on September 9th, 2021). Tomato sequences were obtained from https://solgenomics.net and NCBI; accessed on September 9th, 2021. Arabidopsis sequences were obtained from https://www.arabidopsis.org/ and NCBI; accessed on September 9th, 2021. The amino acid sequences of the *C. arabica* proteins

identified in this study were used. The accession numbers for each of the sequences used are listed in Annexed 1.

2.3. RESULTS

Three types of plant tissues grown *in vitro* were used to identify the differentially abundant proteins between each: calli vs. plantlet leaves (CvsP comparison), suspension culture vs. plantlet leaves (SvsP comparison), and suspension culture vs. calli (SvsC comparison) (Figure 2.1A). In the 1D-SDS-PAGE, it is possible to observe the banding pattern of the total protein extract of each tissue, which highlights notable differences between them (Figure 2.1B). A total of 2,614 proteins were identified among the three comparisons.

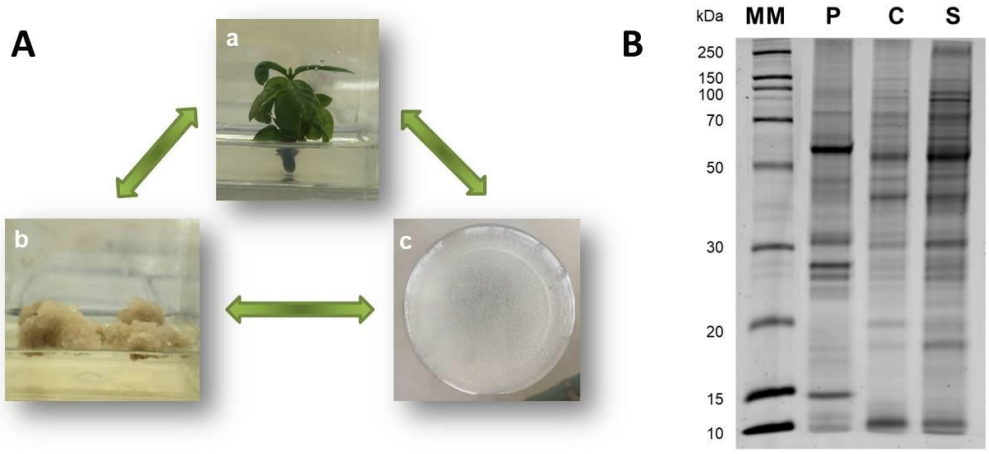


Figure 2.1. *In vitro* tissues of *C. arabica*. **A**) Starting material: plantlet leaves (P), calli (C) and suspension cultures (S); arrows indicate the comparisons that were made between tissues. **B**) 1D-SDS-PAGE profile of *in vitro* tissues; the molecular mass standard is indicated on the left side of the gel. Two biological replicates were used for each tissue.

The heat map (Figure 2.2) shows the difference between the distributions of the protein abundances of each tissue. Differences in the distribution of proteins were found. Two clusters were formed under the tissue comparisons. The accumulation of proteins in the suspension cultures was more similar to that in the calli than to that in the plantlets. In contrast, the distribution of the proteins in the SvsC comparison had a starker contrast than the other two. There are slightly more down-accumulated proteins (green) than up-accumulated (red) when comparing suspension cultures against calli. On the other hand,

more up- than down-accumulated proteins are seen when suspensions and calli are compared against plantlets.

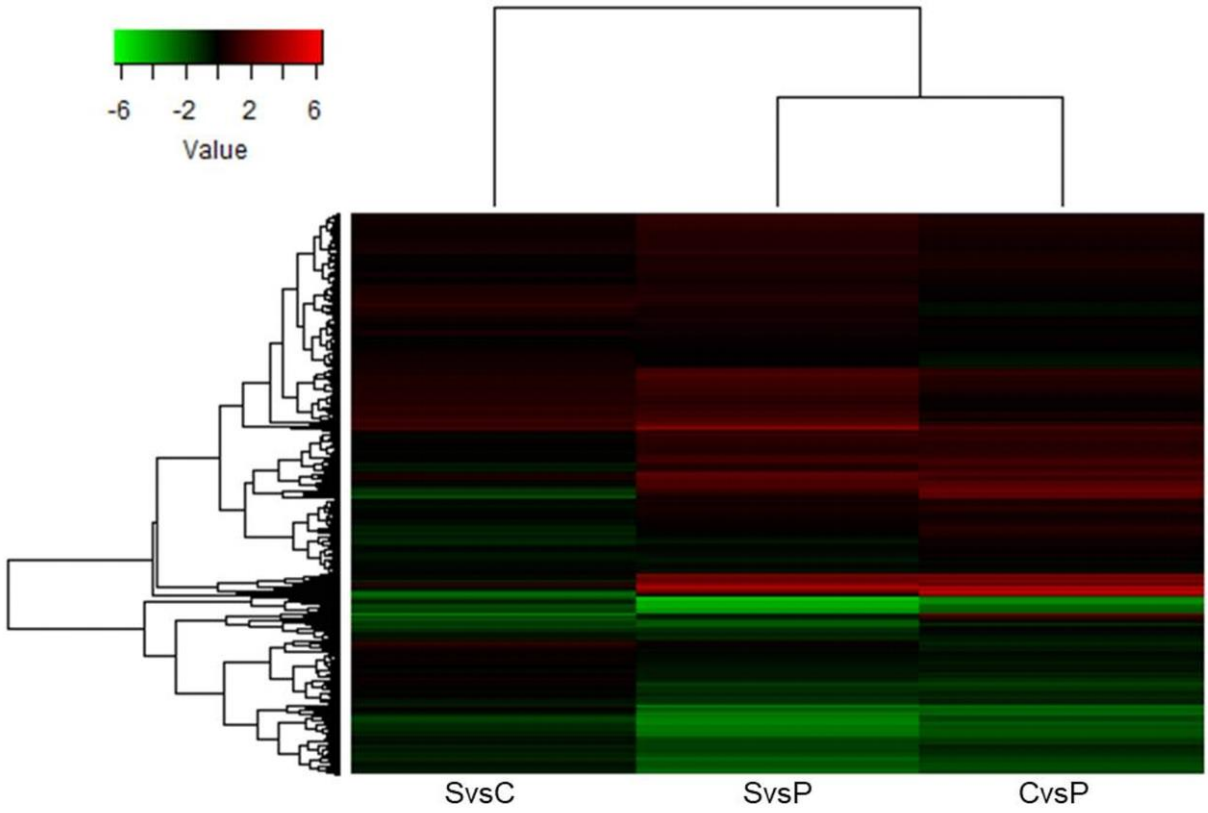


Figure 2.2. Proteome distribution. Heat map showing the distribution of 2,614 proteins among the different tissue comparisons. P: plantlets. C: calli. S: suspension cultures. Down-accumulated proteins are shown in green. Up-accumulated proteins are shown in red.

The total of differentially accumulated proteins (DAPs) was 744, 982, and 295 for the CvsP, SvsP, and SvsC comparisons, respectively (Figure 2.3). Of these, the number of highly accumulated proteins (up-accumulated) was 414, 541, and 120, respectively, while the number of proteins less accumulated (down-accumulated) was 330, 441, and 175 (Figure 2.3).

From the total proteins, the separation of those differentially accumulated was carried out according to their fold change in relative abundance (up >1.5; down <0.66; P<0.05). In accordance with what was observed from the global protein abundance presented in Figure 2.2, in the CvsP and SvsP comparisons, the number of up-accumulated proteins

was higher than those down-accumulated. In contrast, in the SvsC comparison, the opposite occurred, where there were a higher number of down-accumulated proteins (Figure 2.3A). When comparing SvsC, the lowest number of DAPs was found, suggesting that both tissues are at a similar level of differentiation. Therefore, significant changes at the proteomic level would not be expected. On the contrary, more DAPs were found in the SvsP comparison, as they are remarkably distinct at the differentiation stage.

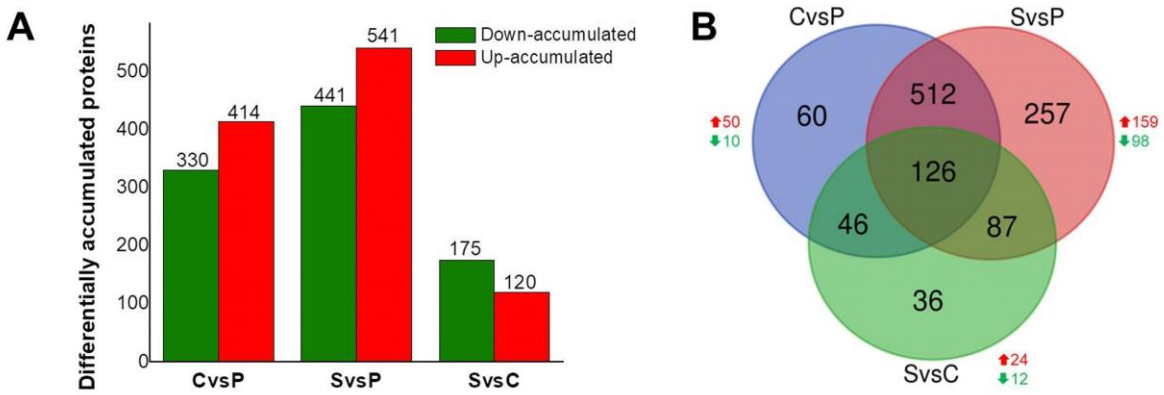


Figure 2.3. Differentially accumulated proteins. **A**) The number of differentially accumulated proteins (DAPs) among the different tissue comparisons. Up-accumulated proteins are shown in red >1.5; down-accumulated proteins are shown in green <0.66; *P*<0.05). **B**) Venn diagram of the differentially accumulated proteins (DAPs) shared between each comparison. The overlapping regions correspond to the number of shared DAPs. Red and green arrows correspond to the number of up-accumulated and down-accumulated proteins, respectively. P: plantlets. C: calli. S: suspension cultures.

A Venn diagram was generated to visualize specific and shared DAPs between the different sets of samples (Figure 2.3B). The highest number of unique DAPs was found in the SvsP comparison, followed by CvsP and SvsC, with 257, 60, and 36 proteins, respectively. Out of the 257 DAPs in SvsP, 159 were up-, and 98 were down-accumulated; and out of the 60 DAPs in CvsP, 50 were up-and ten down-accumulated. On the other hand, with only 36 unique proteins, the SvsC comparison presents the lowest number of unique DAPs, of which 24 were up- and 12 down-accumulated (Figure 2.3B).

The gene ontology analysis for each tissue comparison was performed to classify up-accumulated DAPs according to the biological processes, molecular functions, and cellular components. In the CvsP comparison (Figure 2.4A), the most enriched biological

processes were the catabolic process, the biosynthetic process, carbohydrate and nitrogen metabolism, response to stress, precursor metabolites, and energy. The most enriched molecular functions were ion binding, oxidoreductase, hydrolase and kinase activity, and transmembrane transporter activity. The cytosol, plasma membrane, extracellular region, plastid, and mitochondrion were the most enriched cellular components. According to the hierarchical grouping, the most significant routes correspond to the response to toxic substances and antioxidant activity (Figure 2.4B).

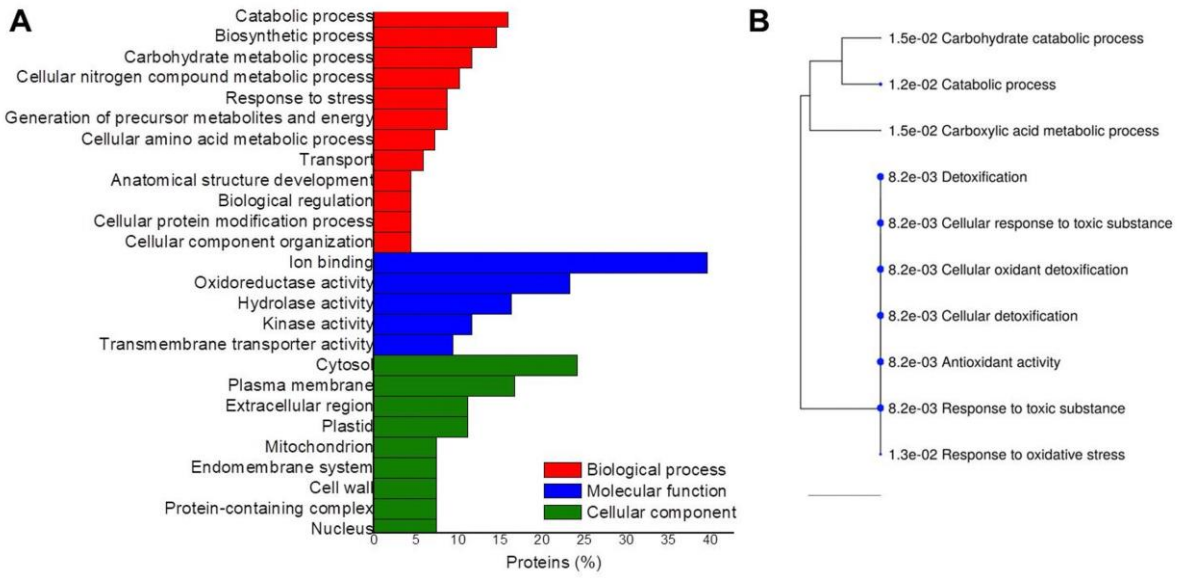


Figure 2.4. Gene ontology analyses of up-accumulated proteins in the CvsP comparison. The proteins identified when comparing calli vs. plantlets were grouped according to **A**) GO enrichment and **B**) hierarchical grouping of the most significant routes based on KEGG.

In addition to biological processes found in CvsP, the SvsP comparison included translation, and cellular component organization, among others (Figure 2.5A). The most enriched cellular components were the cytosol, nucleus, mitochondrion, ribosome, and endomembrane system. Some of the previous molecular functions were also enriched, in addition to those associated with protein binding. The most notable routes correspond to those involved in peptide metabolism (Figure 2.5B).

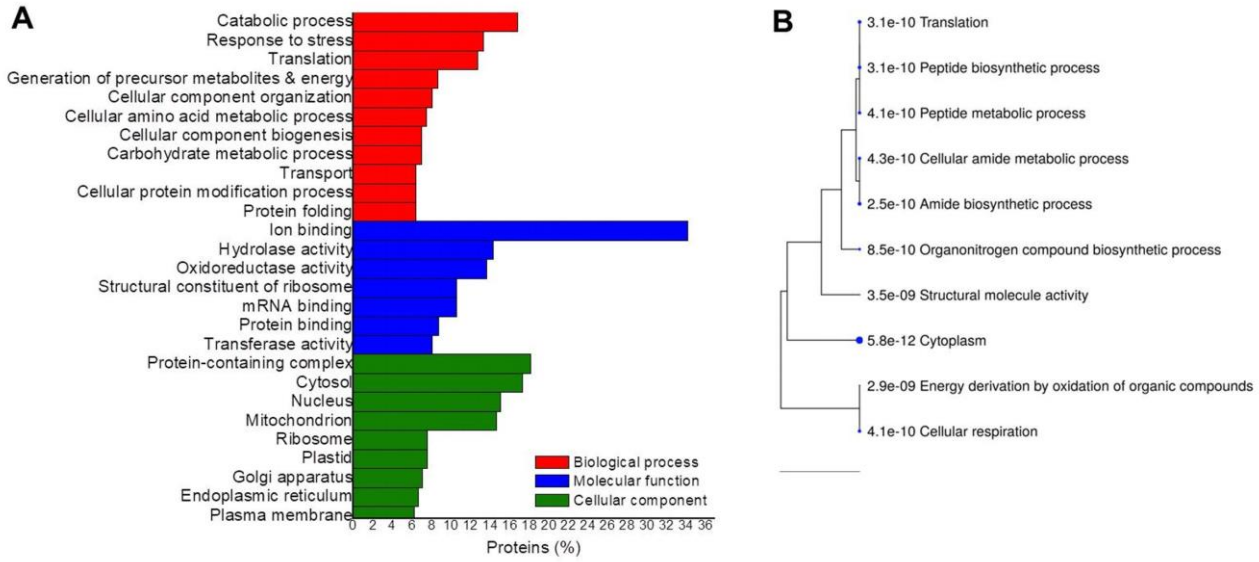


Figure 2.5. Gene ontology analyses of up-accumulated proteins in the SvsP comparison. The proteins identified when comparing suspension cultures vs. plantlets were grouped according to **A**) GO enrichment and **B**) hierarchical grouping of the most significant routes based on KEGG.

In the SvsC comparison (Figure 2.6), the most enriched biological processes were mRNA processing, ribosome biogenesis, protein folding, and the biosynthetic process, among others. The most enriched molecular functions were mRNA and ion binding, and some belonging to protein metabolism. The protein-containing complex, cytosol, plastid, Golgi apparatus, and ribosome were the most enriched cellular components. In this comparison, it was impossible to identify the most significant routes to perform the hierarchical grouping due to the small number of identified DAPs.

In addition, we identified 126 DAPs shared among the three comparisons (Figure 2.3B). These proteins are involved in essential functions, such as the metabolism of energy, carbohydrates, lipids, amino acids, and biosynthesis of other secondary metabolites. Also, a hierarchical GO enrichment grouping was carried out (Figure 2.7). The correlation between the functional categories of the significantly enriched routes of the 126 proteins found to accumulate continuously is summarized. The functions related to photosynthesis and energy generation were the most active.

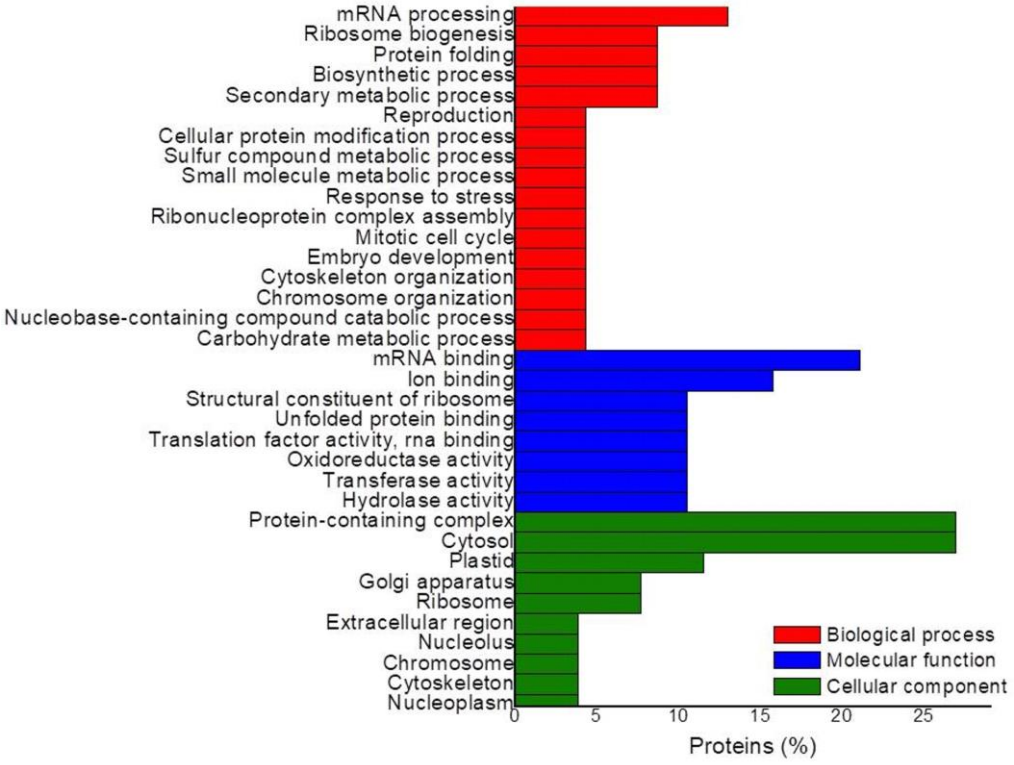


Figure 2.6. Gene ontology analysis of up-accumulated proteins in the SvsC comparison. The proteins identified when comparing suspension cultures vs. calli were grouped according to GO enrichment. Hierarchical grouping of the most significant routes based on KEGG could not be performed.

A manual search was carried out for the proteins involved in auxin homeostasis. Members of families responsible for signaling, transport, conjugation, hydrolysis, and β-oxidation were found, such as ABP, ABC, BIG, GH3, ILR, IBR, and UGT. GH3 proteins play a crucial role in auxin homeostasis through the conjugation of IAA with various amino acids (Westfall et al., 2010). For example, conjugates such as IAA-Asp and IAA-Glu are considered precursors of an irreversible degradation pathway for IAA (Ludwig-Müller, 2011). We found a pair of GH3.17 proteins (CaGH3.17a and CaGH3.17 b) within the *C. arabica* genome. In *A. thaliana*, AtGH3.17 has been reported to correspond to group II. This group is involved in the conjugation of auxin with amino acids (Staswick et al., 2005).

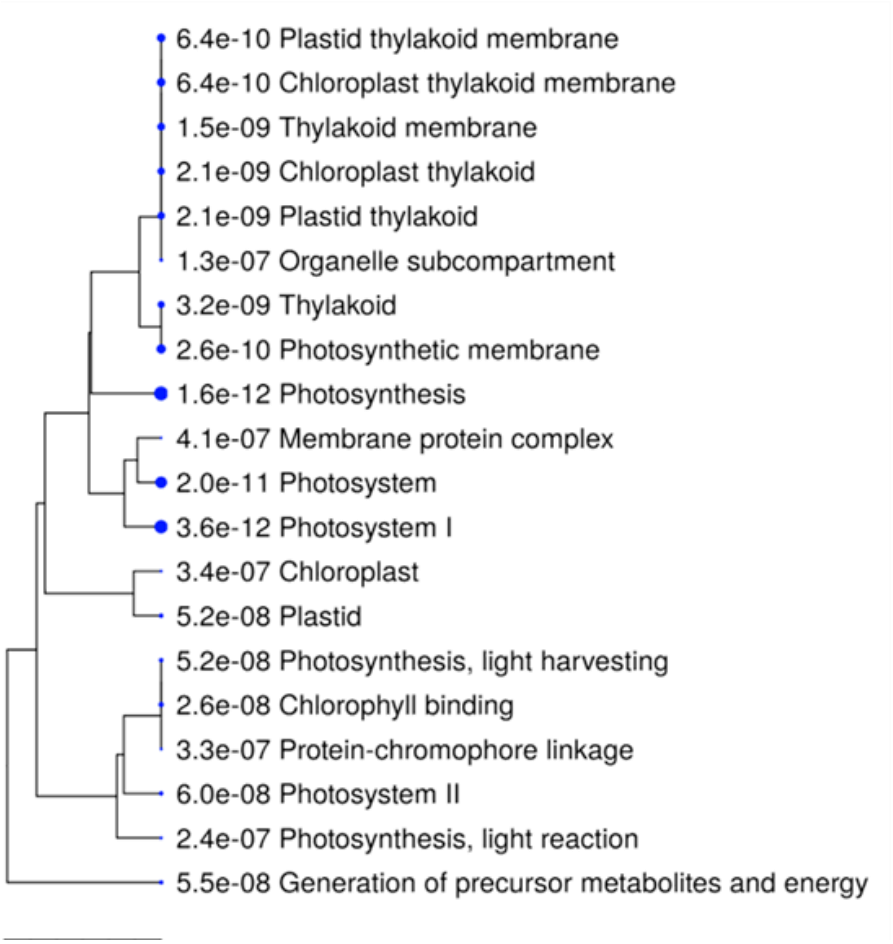


Figure 2.7. Hierarchical grouping tree of the relationship between significantly enriched routes in constitutive proteins. Clustering of the 126 proteins found consistently across the three tissue comparisons. Pathways with many shared genes are clustered. More prominent points indicate more significant *P* values.

We performed a BLAST analysis with our GH3.17a and GH3.17b sequences against The Arabidopsis Information Resource database (TAIR). The sequence GH3.17a from *C. arabica* shared 79% identity and 91% similarity with the GH3.17 protein from Arabidopsis (Annexed 1), while the sequence GH3.17b shared 61% identity and 75% similarity with the Arabidopsis GH3.17 protein (Annexed 1). We subsequently analyzed 47 GH3 protein sequences to build a phylogenetic tree (Annexed 2) using the GH3 proteins of *O. sativa* as an outer group (Figure 2.8). In the phylogenetic tree, the GH3 proteins were grouped into four clades. We observed the three groups previously reported in Arabidopsis (I, II, III). Group I proteins consist only of AtGH3.10 and AtGH3.11 (Staswick et al., 2002). Group II enzymes catalyze the formation of conjugates between auxins (mainly IAA) and amino

acids, which function as a regulatory mechanism to maintain auxin homeostasis (Staswick et al., 2005). The CaGH3.17a and CaGH3.17b proteins from *C. arabica* clustered with the AtGH3.17 protein from *A. thaliana* (Figure 2.8) and possibly fulfill the same function as amido synthetases. They were also grouped with other group II proteins related to the conjugation of IAA with amino acids (Figure 2.8).

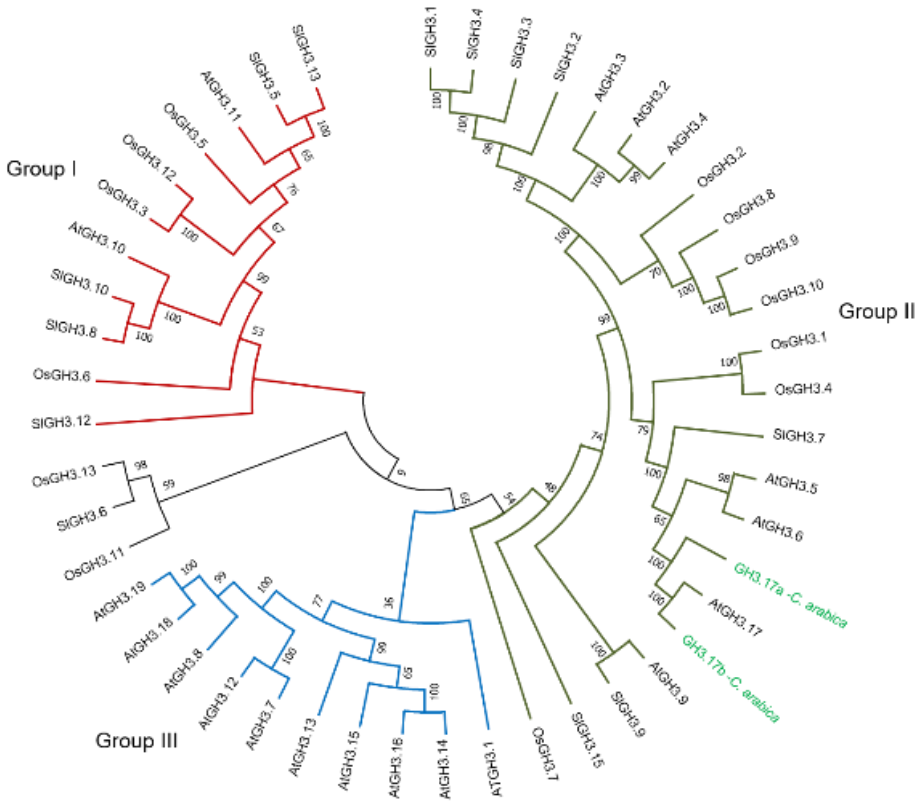


Figure 2.8. Phylogenetic analysis of the GH3 family involved in auxin conjugation. A phylogenetic tree was constructed to study the phylogenetic relationship of the alignments of 47 GH3 sequences. The red, green, and blue branches represent groups I, II, and III, respectively. GH3 sequences were aligned in MUSCLE. Subsequently, a phylogenetic tree was created using the MEGA7 software. The evolutionary history was inferred using the maximum likelihood method. At: *A. thaliana*, Ca: *C. arabica*, Os: *O. sativa*, Sl: S*. lycopersicum*.

Endogenous conjugates such as IAA-Ala, IAA-Leu, IAA-Phe have been reported to appear to be biologically active. They probably provide an easily accessible temporary storage form of auxin (Ljung et al., 2002). A family of amidohydrolases hydrolyzes these conjugates (Fu et al., 2019). Currently, in Arabidopsis, the ILR1-like family consists of seven members: ILR1, ILL1, ILL2, ILL3, IAR3, ILL5, and ILL6 (Rampey et al., 2004). The

best-characterized are ILL1, ILL2, and IAR3 that show more significant catalytic activity with IAA-Ala, while ILR1 prefers IAA-Leu and IAA-Phe as substrates (LeClere et al., 2002). ILL3 and ILL6 show no activity on IAA conjugates *in vitro* (Widemann et al., 2013).

In this work, we found in the *C. arabica* genome proteins related to the hydrolysis of conjugated auxin. The sequences correspond to the proteins ILR1-like 1, ILR1-like 2, and ILR1-like 4. We carried out a BLAST analysis with each of our *C. arabica* ILR1 sequences against The Arabidopsis Information Resource database (TAIR). The CaILR1-like one sequence of *C. arabica* shared 76% identity and 90% similarity with the Arabidopsis AtILR1-like1 protein (Anneced 1), the ILR2-like two sequence shared 55% identity and 73% of similarity with the Arabidopsis AtILL2 protein (Annexed 1). In comparison, the *C. arabica* ILR4-like four sequence shared 73% similarity and 90% identity with the Arabidopsis AtIAR3 (ILL4) protein (Annexed 1). We constructed a phylogenetic tree using 24 protein sequences considered amidohydrolases (Annexed 2), using ILR proteins from *O. sativa* as an outer group (Figure 2.9). We observed four clades in the phylogenetic tree, we found that CaILR1-like 1 is an ortholog of AtILL1 from Arabidopsis, but it is also found in the clade with AtILR1, CaILR-like 2 is grouped with AtILL2. Ilr1 mutants have been reported in Arabidopsis that show reduced sensitivity with biologically active amide conjugates (Davies et al., 1999), while those that overexpress ilr1 and ill2 show greater sensitivity to conjugates with non-polar amino acids (Rampey et al., 2004). CaILR1-like 4 is orthologous with AtIAR3 (ILL4) from Arabidopsis and is also in the same clade as the previous ones, possibly participating in the hydrolysis of conjugates to maintain auxin homeostasis and release it when the cell requires it.

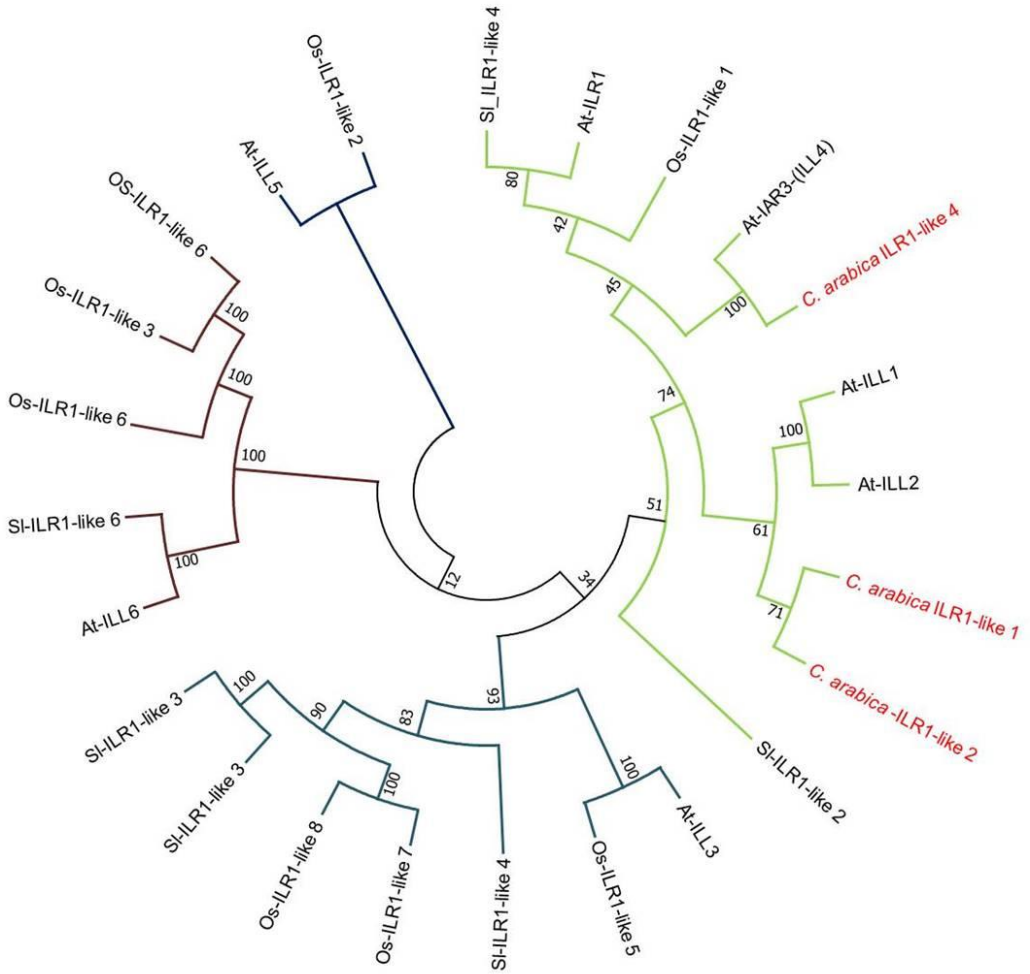


Figure 2.9. Phylogenetic analysis of the ILR1 family involved in auxin hydrolysis. A phylogenetic tree was constructed to study the phylogenetic relationship of the alignments of 24 ILR1 (ILL) sequences. The red, green, and blue branches represent groups I, II, and III, respectively. GH3 sequences were aligned in MUSCLE. Subsequently, a phylogenetic tree was created using the MEGA7 software. The evolutionary history was inferred using the maximum likelihood method. At: *A. thaliana*, Ca: *C. arabica*, Os: *O. sativa*, Sl: *S. lycopersicum*.

In *A. thaliana*, the ABCB subfamily includes 21 members distributed in three clades: ABCB1, ABCB14, and ABCB19 in clade I, ABCB4 in clade II, and ABCB15 in clade III (Geisler and Murphy, 2006). The ABCB1, ABCB4, and ABCB19 proteins have been well characterized as auxin transporters. However other ABC proteins such as ABCB14, ABCB15, and ABCB21 are linked to auxin transport (Jenness et al., 2019; Kaneda et al., 2011; Titapiwatanakun and Murphy, 2009).

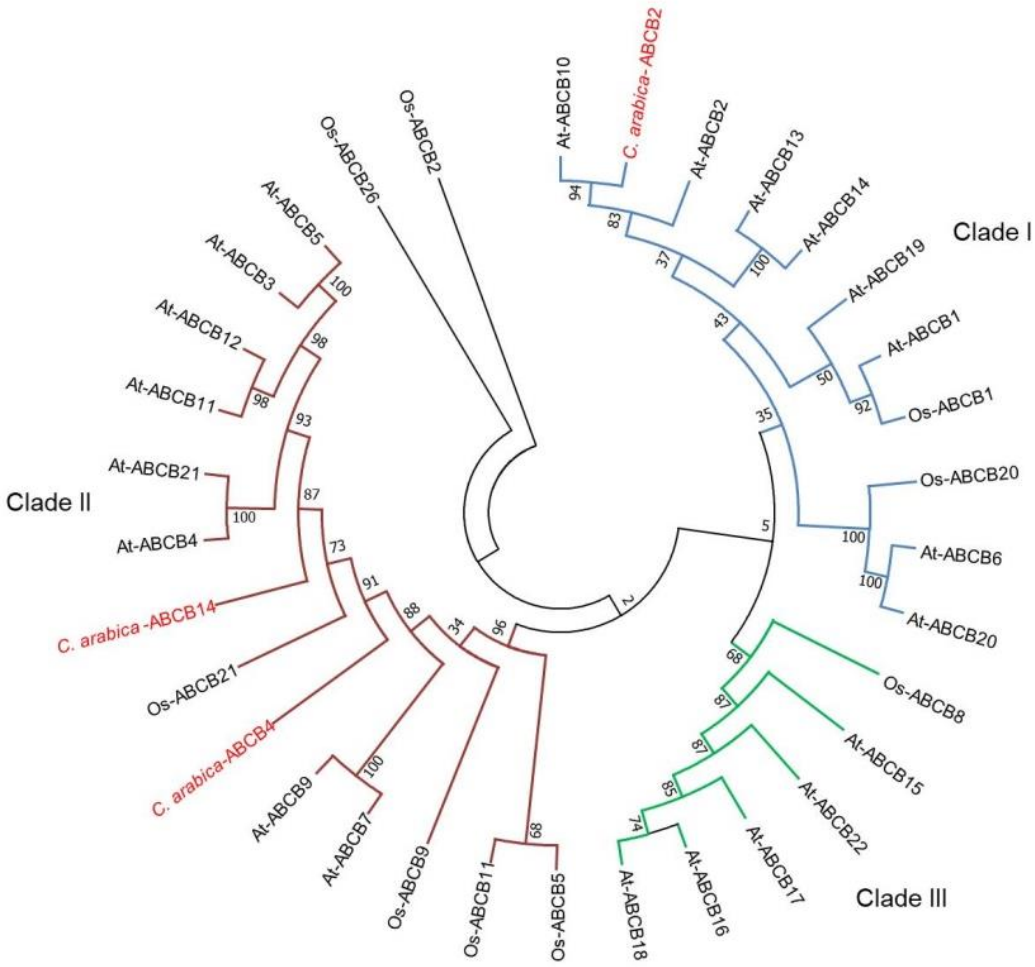


Figure 2.10. Phylogenetic analysis of the ABCB subfamily involved in auxin transporter. A phylogenetic tree was constructed to study the phylogenetic relationship of the alignments of 31 ABCB sequences. The blue, red, and green branches represent the clades I, II, and III, respectively. ABCB sequences were aligned in MUSCLE. Subsequently, a phylogenetic tree was created using the MEGA7 software. The evolutionary history was inferred using the maximum likelihood method. At: *A. thaliana*, Ca: *C. arabica*, Os: *O. sativa*, Sl: *S. lycopersicum.*

In our work, we found eight proteins of the ABC family; however, we will focus the analysis on the members of the ABCB subfamily (ABCB2, ABCB4, and ABCB14). A BLASTp analysis of each of our sequences (ABCB2, ABCB4, and ABCB14) was carried out against TAIR to determine the possible homologs with *A. thaliana*. The ABCB2 sequence of *C. arabica* shared 78% identity and 90% similarity with the Arabidopsis ABCB2 protein. The

C. arabica ABCB4 sequence shared 65% identity and 79% similarity with the Arabidopsis ABCB11 protein. In comparison, the *C. arabica* ABCB14 sequence shared 74% identity and 86% similarity with the Arabidopsis ABCB21 protein (Annexed 1). We also performed a phylogenetic analysis of members of the ABCB subfamily, selecting a total of 31 protein sequences, the outer group being the ABCB proteins from *O. sativa*. The tree included the three groups described above (Figure 2.10).

Because this work aimed to analyze the proteins involved in IAA homeostasis, they were manually selected from the global proteome to compare abundance among the different tissues. One protein involved in IAA signaling was found, the Auxin Binding Protein 20 (ABP20). Eight proteins belonging to the family of ABC transporters and an Auxin transport protein, BIG, were also identified. Two proteins of the GH3 family responsible for conjugation were found, and three ILR1 amidohydrolases that participate in the conjugates' hydrolysis. It was also possible to identify two proteins belonging to the Indole Butyric Response (IBR) family involved in the β-oxidation of indole butyric acid (IBA), and one UGT protein, which is likely to be participating in the conjugation of IBA with sugars. As additional information, some cell wall proteins were identified (Table 2.1). Our findings provide an overview of auxin homeostasis in *C. arabica* and provide a solid basis for further experiments investigating the role of auxin homeostasis in regulating callus and suspension formation in *Coffea.*

Table 2.1. Identification of cell wall proteins.

ID	Protein
CDP00178	Probable xyloglucan endotransglucosylase/hydrolase protein 33
CDO99717	Pectinesterase/pectinesterase inhibitor 18
CDP05792	Pectinesterase/pectinesterase inhibitor PPE8B; I
CDO99718	Pectinesterase/pectinesterase inhibitor 18
CDP21422	Probable pectinesterase/pectinesterase inhibitor 7
CDP19933	Probable pectinesterase/pectinesterase inhibitor 44
CDP02802	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit
CDP06414	D-alanine-D-alanine ligase family
CDP14176	Protein FATTY ACID EXPORT 1, chloroplastic
CDP13911	Expansin-A4
CDP09064	Expansin-A6
CDP07782	Cellulose synthase A catalytic subunit 3 [UDP-forming]

CDP05491 Xyloglucan endotransglucosylase/hydrolase protein 28
CDP15123 Probable xyloglucan endotransglucosylase/hydrolase protein 16
CDP15161 Pectin acetylesterase 2
CDO97787 Pectinesterase/pectinesterase inhibitor 3
CDP20867 SH3 domain-containing protein 2
CDP12145 Probable polygalacturonase
CDP20181 UDP-arabinopyranose mutase 3
CDP11417 Cellulose synthase A catalytic subunit 6 [UDP-forming]
CDP11096 Callose synthase 9; AltName: Full=1,3-beta-glucan synthase
CDP05068 Probable pectinesterase/pectinesterase inhibitor 34
CDP15269 Pectinesterase 3
CDP17099 Probable polygalacturonase
CDP00815 Pectinesterase 2
CDP15690 Pectin acetylesterase 8
CDP19988 Basic endochitinase
CDP15270 Probable pectinesterase/pectinesterase inhibitor 40
CDP10829 Endochitinase EP3
CDP02651 Probable xyloglucan endotransglucosylase/hydrolase protein 6

2.4. DISCUSSION

Growth regulators are vital factors for plant development. Auxins in particular are involved in a large number of processes during tissue development. The responses are dependent on the concentration of auxin and its homeostasis, which in turn depends on its homeostasis. Homeostasis is controlled by several mechanisms, such as auxin biosynthesis, degradation, transport, and conjugation (Ludwig-Müller, 2011).

For IAA signaling to be carried out, the participation of several groups of proteins is required, including the ABP. Within this family, ABP1 is the most studied because it is involved in auxin perception and binding with high specificity and affinity, and has an essential role in several processes such as cell division and cell expansion (Tromas et al., 2013; Teale et al., 2006). We found an ABP20, although its abundance was low in all three comparisons. An ABP20 protein was identified in *Prunus persicaria*, which has an auxin binding motif homologous to the ABP1 protein, but with different specificity (Ades et al., 2014; Ohmiya, 2002). ABP20 is involved in the perception of auxins and it has superoxide dismutase (SOD) activity (Ohmiya, 2002). Subsequently, auxin transport is carried out by

other large groups of proteins, such as ATP-binding cassette (ABC), PIN-FORMED (PIN), and BIG proteins.

The ABC proteins comprise one of the largest families of plant proteins, and participate in the transport of various molecules across the membrane, such as mineral ions, lipids, peptides, metals, secondary metabolites, and growth regulators such as auxin (Geisler et al., 2017; Verrier et al., 2008; Rea, 2007). They also have a primary role in the cellular detoxification mechanism (Kang et al., 2011). This family of transporters in plants is divided into eight subfamilies: A, B, C, D, E, F, G, and I. However, even though it is a very pervasive family of transporters, studies on individual members are scarce (Lane et al., 2016).

Members of five different subfamilies were identified in this study: B, C, E, F, and I. Proteins ABCB4 and ABCB14 showed a notable difference in their abundance between the three different tissue comparisons. In heterologous systems, it was identified that ABCB4 from Arabidopsis participates as an exporter or importer of auxin, depending on its concentration (Yang and Murphy, 2009). On the other hand, it is known that the ABCB14 regulates stomatal activity in the face of changes in CO_2 concentration by importing apoplastic malate (Lee et al., 2008). The earlier research supports what was observed in this study, as the high accumulation of both proteins was more significant in the comparisons against plantlets. In contrast, when comparing calli versus suspensions, the accumulation of ABCB4/14 was negative because both are tissues that lack chloroplasts, and, therefore, there are no stomatal cells to regulate. The ABCs of subfamily C are involved in the vacuolar transport of some compounds, as well as in the compartmentation of anthocyanins, detoxification, heavy metal sequestration, chlorophyll catabolite transport, and ion channel regulation (Wanke and Üner Kolukisaoglu, 2010; Verrier et al., 2008). We also identified ABCC2. However, its abundance was similar between the three comparisons. This protein has been reported to give Arabidopsis resistance to heavy metals such as cadmium and arsenic, along with ABCC1 (Song et al., 2014). Subfamily E is highly conserved in archaea, bacteria, and eukaryotes, which is why these family members are considered necessary for essential functions (Lane et al., 2016).

In this study, we identified the ABCE2 protein, which has been identified as participating in RNA silencing in Arabidopsis (Mõttus et al., 2021). We also found two members of subfamily F, which is not yet well characterized in plants (Verrier et al., 2008); however,

some studies suggest that ABCF3 and ABCF5 could be related to the response to stress. It has been demonstrated that ABCF3 is also involved in the control of protein translation, defense against pathogen infection, and regulation of H_2O_2 uptake by modulating the expression of aquaporin genes (Faus et al., 2021; Dong et al., 2019; Li et al., 2018). Subfamily I is found exclusively in plants (Wanke and Üner Kolukisaoglu, 2010), and they are involved in primary metabolism and responses to stress. It was recently determined that certain members of this subfamily, including ABCI21 found in our study, are involved in modulating cytokinin responses during seedling growth and development (Kim et al., 2020). However, we identified that this protein is not abundant in comparing calli and suspensions against plants.

More than 90% of the auxin in plants is in conjugated form; that is, inactive (Méndez-Hernández et al., 2021; Pencik et al., 2018). When auxin levels are high, a more significant induction of *GRETCHEN-HAGEN 3* (GH3) genes is observed; these genes catalyze the formation of auxin-amino acid conjugates dependent on ATP (Chen et al., 2010). Some records show an increased expression of *GH3.17* genes during cell differentiation (Nic-Can and Loyola-Vargas, 2016). On the other hand, some auxin conjugates (IAA–Ala, IAA–Leu, and IAA–Phe) can be hydrolyzed to return to their active form through the action of hydrolase enzymes such as ILR1. Conjugates with Asp and Glu belong to the degradation pathway (Ludwig-Müller, 2011). ILR1s have been shown to reside in the endoplasmic reticulum, where hydrolases regulate the rates of amido-IAA hydrolysis resulting in activation of the auxin signal (Campanella et al., 2008). Auxins can also be conjugated to sugars using the uridine diphosphate (UDP) glycosyltransferases (UGTs) (Casanova-Sáez and Voß, 2019; Jin et al., 2013; Ludwig-Müller, 2011). In this study, UGT75C1 was identified as highly abundant in the SvsP comparison. Various UGTs have been characterized as participants in the control of the metabolism of different plant growth regulators (Mateo-Bonmatí et al., 2021); However, there are only two reports of UGT75C1 found in plants: one in Arabidopsis (Gachon et al., 2005) and another recently in *Lonicera japonica* (Xia et al., 2021), where it is thought to function as an anthocyanin-5-*O*-glucosyltransferase *in planta*. Nevertheless, more studies should be done to confirm the role of UGTs in plant development. In addition to the hydrolysis of the conjugates, another way to obtain free IAA is from IBA, through the elimination of two side-chain methylene in

a β-oxidation process catalyzed by the indole butyric acid response enzyme (IBR) (Zolman et al., 2008).

A model was made with the set of identified proteins, which summarizes each of them in the different tissue comparisons (Figure 2.11).

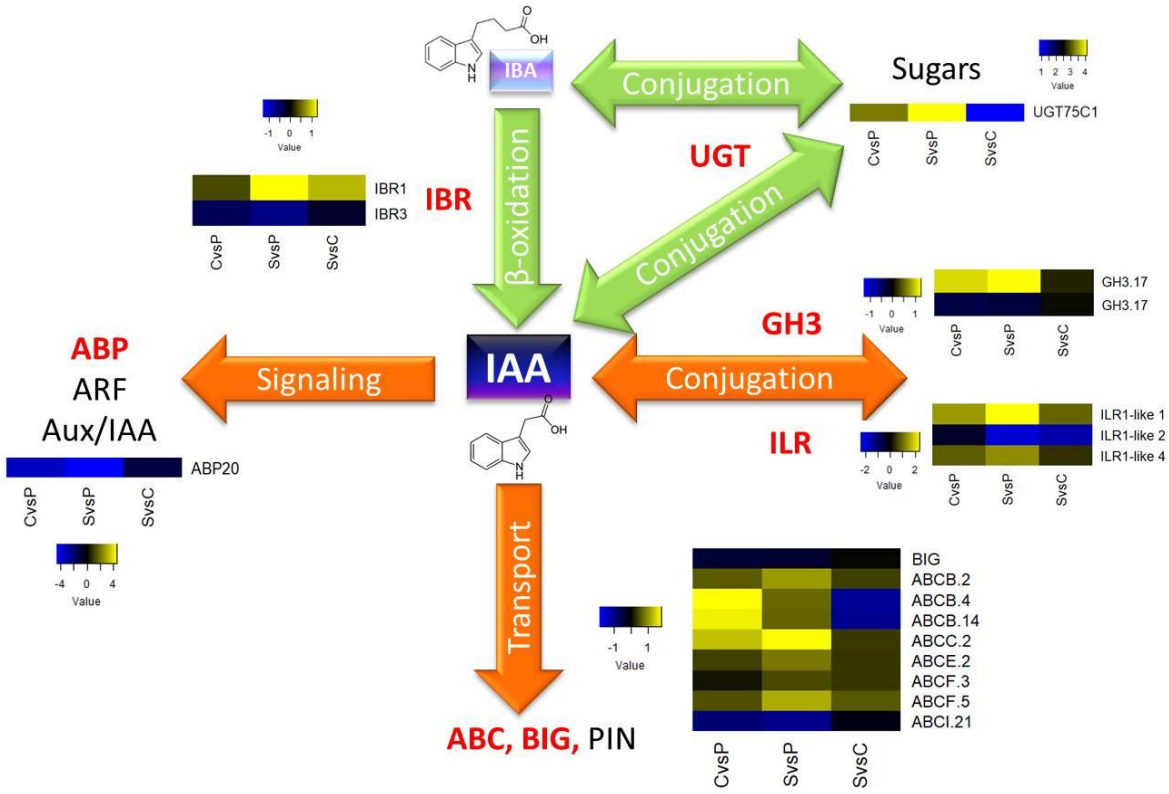


Figure 2.11. Model of auxin homeostasis during cell differentiation in *C. arabica in vitro* culture. Participation of proteins identified in this study (marked in red) involved in signaling, transport, conjugation, and β-oxidation of auxin.

Of the eight proteins identified from the ABC family, three were type B, one was type C, one was type E, two were type F, and one was type I, which remained in low abundance in the three comparisons. The ABP20 protein involved in IAA transport had a low accumulation in all three tissue comparisons; likewise, the accumulation of the auxin transport protein BIG remained unchanged in all three comparisons. Another five proteins had moderately high accumulation in all three tissue comparisons (ABCB.2, ABCC.2, ABCE.2, ABCF.3 and ABCF.5); meanwhile, the abundance of ABCB.4 and ABCB.14 was much lower in the SvsC comparison only. Of the two GH3.17 proteins, one had high accumulation compared to calli and suspension cultures against plantlets, while the other,

GH3.17, remained in low abundance in all three comparisons. Three ILR1 were found, of which the ILR1-like 2 remained with low abundance in every comparison.

On the other hand, the ILR1-like 1 and 4 had similar behavior with a relatively high accumulation in all three comparisons. The IBR1 protein was highly abundant in the SvsP comparison, while the abundance of the IBR3 was lower in all comparisons. Also, the UGT75C1 accumulation was very high in the SvsP comparison, while in the SvsC comparison, the accumulation was low. According to the model, it is inferred that the most significant changes in terms of the abundance of most of the proteins involved in auxin homeostasis occurred when the cell suspensions were compared with the plantlets (SvsP). On the contrary, fewer differences were observed when comparing suspension cultures with calli (SvsC).

The pant cell wall serves as dynamic physical barrier, consisting of interconnected layers that contains cellulose, hemicellulose, pectin, lignin, and proteins. Proteins associated with plant cell assembly and biogenesis were accumulated in callus. Protein with the role in the cleavage and build of polysaccharides including two xyloglucan endotransglucosylase, two cellulose synthases, and a callose synthase. A Remarkable feature of callus is the up-accumulation of seven pectin methylesterase inhibitors (PMEI) that can modulate the de-methylesterification of homogalacturonan by the inhibition of pectin methylesterase. Biochemical studies have found highly methylesterified pectin in callus and variation during the somatic embryo formation from callus.

The liquid chromatography-mass spectrometry (LC-MS) tool has been gaining attention in recent years for proteomic studies, because it offers a quantitative approach to the proteome (Liu et al., 2019). Quantitative proteomics provides information on the molecular mechanisms that operate in the cell under various study conditions (Jamet and Santoni, 2018). In this sense, the Tandem Mass Tag allows a precise identification and quantification of proteins (Xinxin et al., 2020). Few proteomic studies using gel-based techniques coupled with mass spectrometry have been performed in *C. arabica* suspension cultures and embryos (Campos et al., 2016; Tonietto et al., 2012; Mukul-López et al., 2012). Up to now, there are no reports of the use of TMT in *in vitro* tissue studies of this species so that this study can serve as a reference for the characterization of the proteome of *C. arabica* suspension cultures, calli, and plantlets.

2.5. CONCLUSIONS

Tissue culture represents an alternative for the conservation, propagation, and genetic improvement of *C. arabica*. Different tissues are used; however, few proteomic studies have been performed to identify the proteins involved in cell differentiation. Auxin plays a fundamental role in the maintenance and development of *in vitro* plant tissue culture. The most significant difference was found when comparing the proteins accumulated in the suspensions with the plants. The more significant accumulation of proteins, such as some ABCs, GH3.17, UGT75C1, and IBR1, suggests auxin's control in its active and inactive form through the mechanisms of homeostasis given by signaling, transport, conjugation, and hydrolysis as tissue differentiation increases.

CHAPTER III

PROTEOMIC APPROACH DURING THE INDUCTION OF SOMATIC EMBRYOGENESIS IN *Coffea canephora*

This chapter was accepted for publication as a research article in the Plants journal.

Abstract

Plant growth regulators (PGR) are essential for somatic embryogenesis (SE) in different species, and *Coffea canephora* is no exception. In our study model, previously, we have been able to elucidate the participation of various genes involved in SE by using different strategies; however, until now, we have not used a proteomic approach. Therefore, in this study, we analyzed different days of the SE induction process using shotgun label-free proteomics. We found that some of the most enriched pathways during the process were the biosynthesis of amino acids and secondary metabolites. Eighteen proteins were found related to auxin homeostasis and two to cytokinin metabolism, such as ABC, BIG, ILR, LOG, and ARR. Ten proteins and transcription factors related to SE were also identified, and 19 related to other processes of plant development, among which the 14-3-3 and PP2A proteins stand out. This is the first report on the proteomic approach to elucidate the mechanisms that operate during the induction of SE in *C. canephora*. So our findings provide the groundwork for future, more in-depth research. Data are available via ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD047172 and DOI: 10.6019/PXD047172.

Keywords: *Coffea canephora*, mass spectrometry, plant tissue culture, somatic embryogenesis.

3.1. INTRODUCTION

Somatic embryogenesis (SE) is an effective biotechnological tool for studying the morpho-physiological, biochemical, and molecular processes that develop crops of interest, such as the coffee plant. Also, SE represents a viable alternative that allows the mass production of homogeneous plants, free of diseases and with desirable agronomic characteristics, in less time and space.

The genus *Coffea* comprises more than 127 species of evergreen woody trees with differences in size, morphology, and adaptation conditions. Among these species, *C. canephora* and *C. arabica* are the most cultivated worldwide (Ferreira et al., 2019) with a worldwide production of 40 and 60%, respectively (ICO, 2022). Among them, *C. canephora* has a higher caffeine content and is more resistant to pests, diseases (Etienne, 2005), and extreme climatic changes than *C. arabica*. Given the importance of this crop, various investigations have been developed related to the increase in mass propagation, its genetic improvement, and improvement in agronomic qualities such as productivity, grain quality, physicochemical processes, and resistance to biotic and abiotic factors, among others.

SE has currently been used to study cell differentiation in the *C. canephora* system, and it has been shown that plant growth regulators (PGRs) play a primary role in cell signaling and differentiation (Méndez-Hernández et al., 2019). It has been found that during the SE process in *C. canephora*, the most important auxin (indole3-acetic acid/IAA) plays a fundamental role. It has become the object of study to elucidate the mechanisms involved in its biosynthesis, transport, signaling, accumulation, and homeostasis (Nic-Can and Loyola-Vargas, 2016).

Although transcriptomics has broadened the panorama for the correct understanding of the first stages of the SE process by studying the genes involved in the embryogenic response (Chen et al., 2020; Guo et al., 2020; Enríquez-Valencia et al., 2019; Quintana-Escobar et al., 2019; Góngora-Castillo et al., 2018), there is another area that has gained great scope in recent years: proteomics. Proteomics offers a closer approach to the state of the cell in a specific space and time (Kumaravel et al., 2020; Gulzar et al., 2019; Aguilar-Hernández and Loyola-Vargas, 2018).

Despite the valuable results obtained using transcriptomics regarding the genes directly involved in SE, the lack of correlation between the number of transcripts and the abundance of proteins may be a limiting factor. Among the reasons that cause this lack of correlation are variations in mRNA stability, its translation, protein stability, changes in structure, activity, and function due to its cellular location, interaction with other molecules, or post-translational modifications (Tchorbadjieva, 2016). For the above, it is increasingly a requirement to complement the information of the transcripts with the final functional products of the genes: the proteins. Proteomics helps elucidate the biochemical and

molecular processes necessary to carry out SE through the identification and/or quantification of proteins with differential abundances (Heringer et al., 2018; Ge et al., 2014).

In spite of the increased use of proteomic strategies, more is needed to know about the study of *Coffea* spp. A few studies focus on the comparison of embryogenic versus non-embryogenic lines. In one of the first reported works on proteomics in *C. arabica*, three stages of somatic embryo development were evaluated through 2DE and mass spectrometry (Tonietto et al., 2012). Specific proteins were identified at each stage. It was proposed that enolase and 11S storage globulin proteins could serve as molecular markers for embryo development and for the differentiation between embryogenic and non-embryogenic lines. Another work evaluated two embryogenic suspensions through mass spectrometry and shotgun (Campos et al., 2016). The proteome was characterized, and proteins related to stress and energy production was identified. Although the analysis is very descriptive, it lays the groundwork for considering possible molecular markers for SE. Another investigation focused in the comparison, through 2DE, the extracellular proteome of embryogenic and non-embryogenic cell suspensions of *C. arabica* and *C. canephora* (Mukul-López et al., 2012). It was found that a larger population of proteins secreted into the medium of *C. canephora* compared to *C. arabica* and those proteins are secreted exclusively under embryogenic conditions. However, the identification of these proteins was not possible. So far, to our knowledge, there are no reports on the study of the proteome during the somatic embryogenesis of *C. canephora*. Thus, this work aims to contribute to understanding the main cellular mechanisms involved during the induction of SE in *C. canephora* through a proteomic approach.

3.2. MATERIALS AND METHODS

3.2.1. Biological material and growth conditions

Plantlets of *C. canephora* grown *in vitro* were used as initial biological material. These plantlets were subcultured every six weeks in a semi-solid maintenance medium without PGR [MS salts (PhytoTechnology Laboratories, M524), 11.85 µM thiamine-HCl (Sigma, T3902), 550 µM myo-inositol (Sigma, I5125), 158 µM cysteine hydrochloride hydrate (Sigma, C121800), 16.24 µM nicotinic acid (Sigma, N4126), 9.72 µM pyridoxine-HCl

(Sigma, P9755), 87.64 mM sucrose (Sigma, S539) and 0.285% (w/v) Gellan gum (PhytoTechnology Laboratories, G434), adjusted to pH 5.8]. Plantlets were incubated under a photoperiod of 16 h light and 8 h dark at 25 ± 2 °C.

For the SE induction process, the seedlings were previously incubated for 14 d in a preconditioning semi-solid medium [same composition of the maintenance medium, added with 0.54 µM NAA (Sigma, N1145; St. Louis, MO, USA) and 2.32 µM KIN (Sigma, K0753; St. Louis, MO, USA) adjusted to pH 5.8]. After 14 d of preconditioning, we cut circular explants from the second and third pair of leaves in a basipetal direction with a sterile punch of 0.8 mm in diameter.

Five explants were placed in 50 mL induction liquid culture medium [Yasuda salts (1985) supplemented with 5 µM BA (PhytoTechnology Laboratories, B800; Kansas, USA) and adjusted to pH 5.8] in 250 mL flasks. The explants were incubated in the dark at 25 ± 2 °C and shaking (60 rpm). Samples were taken for subsequent analysis during preconditioning (14 and 0 dbi) and after induction of SE (7 and 21 dai). For sampling, sections of leaf discs were used in all samples, including the control day, so that the comparison between them was homogeneous. The leaf discs were briefly rinsed with distilled water to remove excess culture medium, and placed on absorbent paper to remove liquid on the explant surfaces. They were subsequently weighed into 100 mg packets per triplicate, frozen in liquid nitrogen, and then stored at -80 °C until use.

3.2.2. Protein extraction

One hundred mg of plantlet leaves were pulverized in a mortar with liquid nitrogen until a fine powder, avoiding thawing. The sample powder was transferred to a 2 mL microcentrifuge tube, and 1 mL of extraction buffer was added, containing 0.5 M Trizma base (pH 8; Sigma, T1503; St. Louis, MO, USA), 50 mM EDTA (Sigma, EDS; St. Louis, MO, USA), 0.7 M sucrose, 0.1 M KCl (Sigma, P9541; St. Louis, MO, USA), 50 mM DTT (Sigma, D5545; St. Louis, MO, USA), 1% SDS (Sigma, L3771; St. Louis, MO, USA), 1 mM PMSF (Sigma, 78830; St. Louis, MO, USA) and a protease inhibitor cocktail (Sigma, P9599; St. Louis, MO, USA). The extract was vigorously vortexed for 2 min, followed by 15 min ice incubation with gentle shaking. An equivalent volume of phenol solution (Sigma, P4557; St. Louis, MO, USA) was added and vigorously vortexed, followed by incubation on ice with gentle shaking for 30 min. The tubes were centrifuged at 15,000 x g for 30 min

at 4 °C. The upper phase was recovered in a new tube, avoiding carrying cell debris. Proteins were precipitated overnight at -20 °C with five volumes of 0.1 M ammonium acetate (CTR, 00140; Monterrey N. L., México) dissolved in methanol with 5 mM DTT. The following day, tubes were centrifuged at 15,000 x g for 30 min at 4 °C, and the supernatant was discarded. The protein pellet was washed once with 1 mL 0.1 M ammonium acetate/methanol/5 mM DTT and twice with 80% acetone/5 mM DTT. The pellet was allowed to dry in an extraction hood for 3-5 minutes and then resuspended in 50 mM ammonium bicarbonate (Sigma, A6141; St. Louis, MO, USA) supplemented with 0.1% SDS (Sigma, L3771; St. Louis, MO, USA). Protein quantitation was determined by the Peterson method (Peterson, 1977), and the quality was visualized on SDS-PAGE.

3.2.3. Reduction, alkylation, and digestion

For sample reduction, 10 mM TCEP (Sigma, 68957; St. Louis, MO, USA) was added and incubated at 60 °C for 45 min. For alkylation, 30 mM IAM (Sigma, A3221; St. Louis, MO, USA) was added and incubated for 60 min in the dark at room temperature. Then, 30 mM DTT was added and incubated for 10 min. Proteins were precipitated with cold acetone overnight at -20 °C. The next day, the tubes were centrifuged at 10,000 x g for 15 minutes at 4 °C, and the supernatant was discarded. The pellet was allowed to dry in an extraction hood for 3-5 minutes and resuspended in 50 mM ammonium bicarbonate with 0.1% SDS. Protein concentration was quantified, and quality was visualized on SDS-PAGE. Digestion was carried out with trypsin (Thermo Scientific, 90058; Rockford, IL, USA) in 150 µg of protein at a 1:60 ratio (trypsin:protein) overnight at 37 °C. The next day, more trypsin was added at 1:100 for 4 h.

3.2.4. Nano LC/MS-MS analysis

The preparation of the samples and subsequent analysis was carried out as reported by Bautista (2023). The fractionation was carried out off line before LC-MS/MS analysis with High-pH reversed-phase liquid chromatography spin columns (Pierce High pH Reversed-Phase Cat No. 84868). Three fractions were obtained after elution with increasing concentrations of acetonitrile (15, 17, and 20%). Then, fractions were desalted using ZipTip-C_{18} tips (Merck Millipore, Darmstadt, Germany) and dried in a vacuum concentrator. An Orbitrap Fusion Tribrid (Thermo-Fisher Scientific, San Jose, CA, USA) mass spectrometer equipped with an “EASY spray” nano ion source (Thermo-Fisher Scientific,

San Jose, CA, USA) and interfaced with an UltiMate 3000 RSLC system (Dionex, Sunnyvale, CA, USA) was used to analyze the samples. Each sample was reconstituted with 0.1% formic acid in LC-MS grade water (solvent A). Five µL were injected into a nanoviper C_{18} trap column (3 µm, 75 µm × 2 cm, Dionex) at a flow rate of 3 µL min^{-1} and separated on an EASY spray C_{-18} RSLC column (2 µm, 75 µm × 25 cm). A 100 min gradient of Solvent A and 0.1% formic acid in 90% acetonitrile (Solvent B) with a 300 nL min^{-1} flow rate was used as follows: 10 min with 100% solvent A, 25 min with 7%–20% solvent B, 15 min with 20% solvent B, 15 min with 20%–25% solvent B, 20 min with 25%–95% solvent B, 8 min with solvent A. The mass spectrometer was set to positive ion mode, with a nanospray voltage of 3.5 kV and a source temperature of 280 °C; precursor selection mass range of 400–1200 m/z, precursor ion exclusion width of low 18 m/z and high 5 m/z, The external calibrants were Caffeine, Met-Arg-Phe-Ala (MRFA), and Ultramark 1621 (Cat No. 88323, Thermo-Fisher Scientific Pierce). MIAPE Reporting guidelines for mass spectrometry are described in Annexed 4. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2022) partner repository (http://www.ebi.ac.uk/pride) with the dataset identifier PXD047172 and DOI: 10.6019/PXD047172.

3.2.5. Data processing

The resulting MS/MS data were processed using the MASCOT (v.2.4.1, Matrix, Science, Boston, MA) search engine implemented in Proteome Discoverer 2.2 (PD, Thermo Fisher Scientific, San Jose, CA, USA). The PD parameters were set as follows: *viridiplantae* Swiss-prot database, mass tolerance of 10 ppm and 0.6 Da, two missed cleavages allowed, 0.01 FDR, cysteine carbamidomethylation as fixed modification, methionine oxidation and N-terminal acetylation as dynamic modification. For label-free quantification, the Minora node was incorporated into the Processing workflow. The output data was filtered, eliminating those rows with empty abundance values. The filtered results in the spreadsheets were uploaded to the Galaxy platform for the differential analysis, and the Limma tool and normalization by TMM were used. Differentially abundant proteins were selected according to a fold change of 1 and $P < 0.05$. Annotation was carried out in KOBAS against the *Arabidopsis* database. Protein IDs were loaded on g:Profiler (https://biit.cs.ut.ee/gprofiler/gost; accessed on August 2023) to perform GO enrichment. ShinyGO v0.77 (http://bioinformatics.sdstate.edu/go/; accessed on August 2023) was used

to perform KEGG enrichment. The protein-protein interaction network of SE-related proteins was carried out using STRING (https://string-db.org/; accessed on August 2023). PlantCyc (https://pmn.plantcyc.org/; accessed on September 2023) was used to visualize SE-related protein abundances in the different biochemical pathways.

3.3. RESULTS

The induction of SE in *C. canephora* was carried out successfully. This process consisted of cultivating *C. canephora* seedlings for 14 days in a medium added with plant growth regulators (PGR) that favor the subsequent embryogenesis induction. After preconditioning, the circular explants were placed in a liquid induction medium with benzyladenine (BA) as the only PGR. At 7 d after induction, a slight change was observed in the periphery of the explants due to an increase in the size of the explants. By 21 d, the growth of proembryogenic structures was already evident throughout the periphery of the explants. 14 dbi (control), 0dbi (seedlings at the end of 14 d in preconditioning medium), 7 and 21 d after induction were selected as sampling days (Figure 3.1A). Day 14 dbi was selected as a control to make comparisons in subsequent analyses because the seedlings were under maintenance in a culture medium without PGR.

The quality of the extracted proteins for each sampling day was verified on 1D-SDS-PAGE gels (Figure 3.1B), in which no difference was visually observed between the different sampling points. It is worth mentioning that the development process of the somatic embryos was carried out to the end to verify the protocol's effectiveness. At 56 d, embryos released into the culture medium could already be observed at different stages of development (Figure 3.1C). These embryos were collected and placed in a semisolid culture medium without PGR for germination and conversion to complete seedlings (Figure 3.1D), which can be used to start a new SE process.

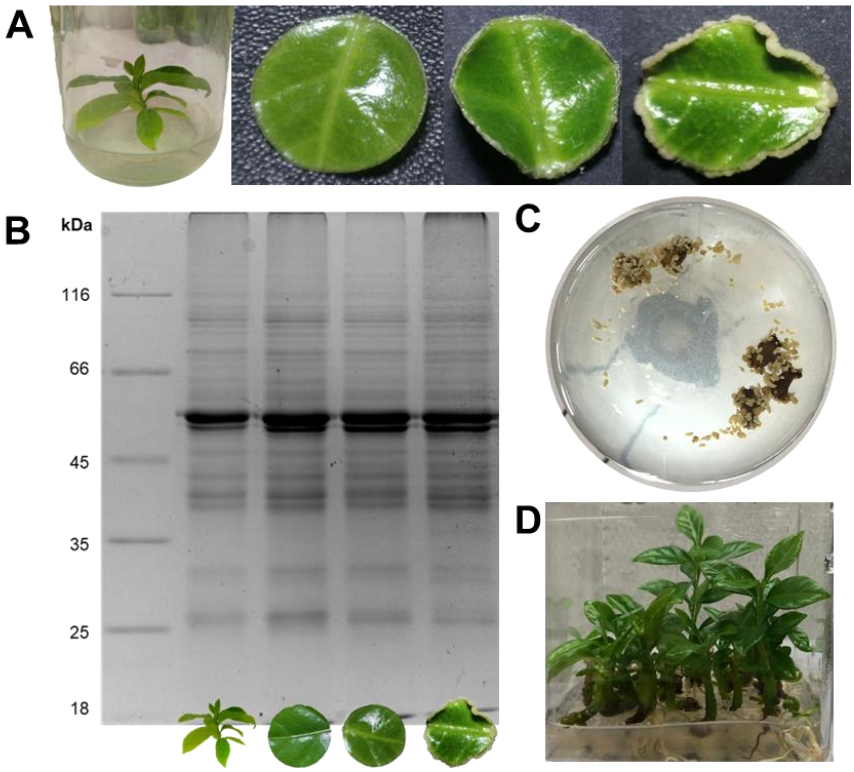


Figure 3.1. Somatic embryogenesis in *C. canephora*. **A**) From left to right: plantlet and leaf explant 14 and 0 days before induction (dbi); leaf explant at 7 and 21 days after induction (dai). **B**) 1D-SDS-PAGE visualization of the protein profiles from 14 dbi, 0 dbi, 7 dai, and 21 dai samples (10 µg protein). **C**) Flask with somatic embryos released into the culture medium. **D**) Somatic embryos germinated and converted to complete seedlings.

After proteomic analysis by LC-MS/MS, 1,630 proteins accumulated on the different sampling days were identified. When performing the heatmap (Figure 3.2), two different clusters were observed: in the first cluster, day 14 dbi was grouped with 0 dbi. In the second cluster, the 7 dai were grouped with the 21 dai. The different samples observed a differential pattern in global protein abundance levels. In this grouping, it can be seen that there is a greater tendency for poorly accumulated proteins in the first days. In contrast, highly accumulated proteins are more significant in the days after induction.

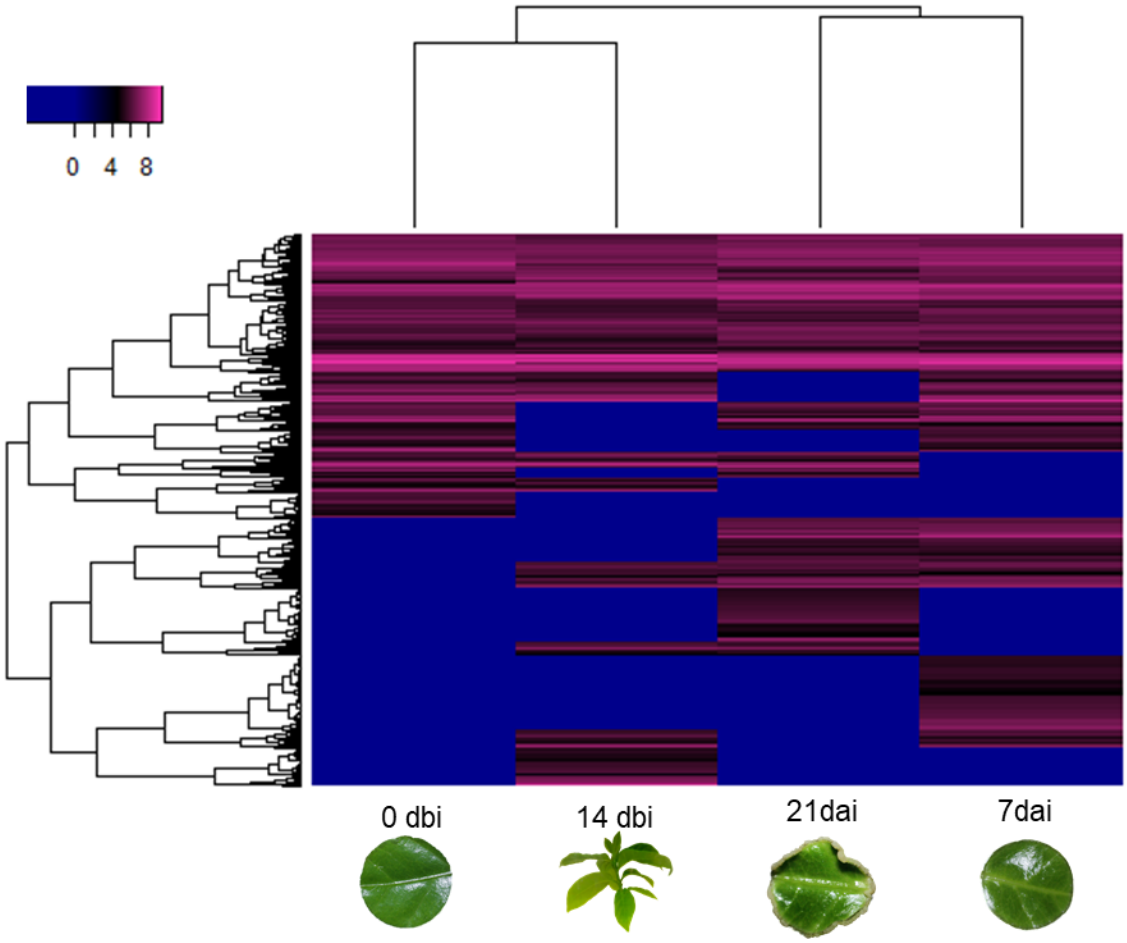


Figure 3.2. Proteomic distribution of somatic mbryogenesis induction in *C. canephora*. Heatmap depicting the abundance of the 1630 proteins found among different sampling days of the process.

Among the 1,630 proteins identified, the highest number of accumulated proteins was found in the 7 dai samples, followed by 21 dai, 14 dbi, and 0 dbi (Figure 3.3). In the Venn diagram (Figure 3.3A) and the UpsetR plot (Figure 3.3B) you can see the comparison of proteins between the different sampling days, as well as those proteins that are unique to each condition. On day 7 dai, there was a greater number of proteins exclusive to this day (224 proteins), while the smallest number of unique proteins was concentrated on day 0 dbi (74 proteins). On the other hand, 407 proteins were found accumulated on all sampling days, which can be considered constituent proteins of the SE process.

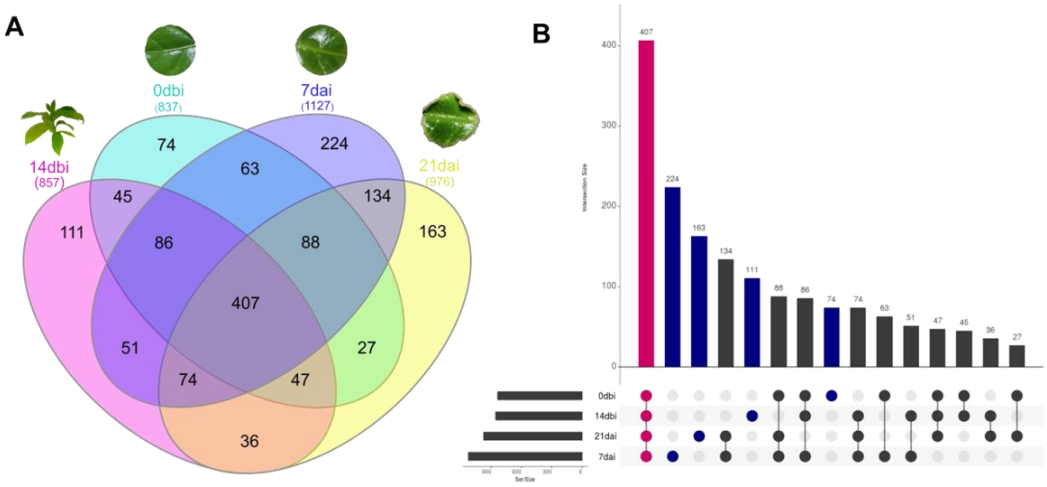


Figure 3.3. Comparison of total proteins among the different sampling days of somatic embryogenesis induction in C. canephora. **A**) Venn diagram showing the distribution of 1630 proteins shared among the different days. **B**) UpsetR plot. The overlapping regions correspond to the number of shared proteins between conditions. Dbi: days before induction. Dai: days after induction.

To perform the studies of differential accumulation analysis, the abundances of all proteins from days 0 dbi, 7 dai, and 21 dai were compared against those from 14 dbi. In this way, a total of 3 compared conditions were obtained: 0dbi/14dbi, 7dai/14dbi, and 21dai/14dbi.

The lowest number of differentially accumulated proteins (DAPs) was found in the 0dbi/14dbi condition, with 77 proteins (Figure 3.4A), followed by the 21dai/14dbi and 7dai/14dbi conditions, with 162 and 163 differential proteins, respectively. Of these proteins, 39 accumulated only in the 0dbi/14dbi condition, 92 in the 7dai/14dbi condition, and 100 in the 21dai/14dbi condition (Figure 3.4A).

Those differential proteins (P = 0.05) with a log fold change (LFC) greater than one were selected as up-accumulated, and those with a LFC greater than one were selected as down-accumulated (Figure 3.4B and C). In the 0dbi/14dbi condition, 43 up-accumulated and 34 down-accumulated proteins were found, with 20 (up-) and 19 (down-) proteins unique to this condition, respectively (Figure 3.4D and 4E). In the 7dai/14dbi condition, 124 up-accumulated and 39 down-accumulated proteins were found, of which 80 (up-) and 12 (down-) were unique to this comparison. On the other hand, in the 21dai/14dbi condition, 92 and 70 up and down-accumulated proteins were found, of which 58 (up-) and 42

(down-) were unique to this condition. The up-accumulated proteins from each comparison were selected for gene ontology analysis and KEGG pathway enrichment.

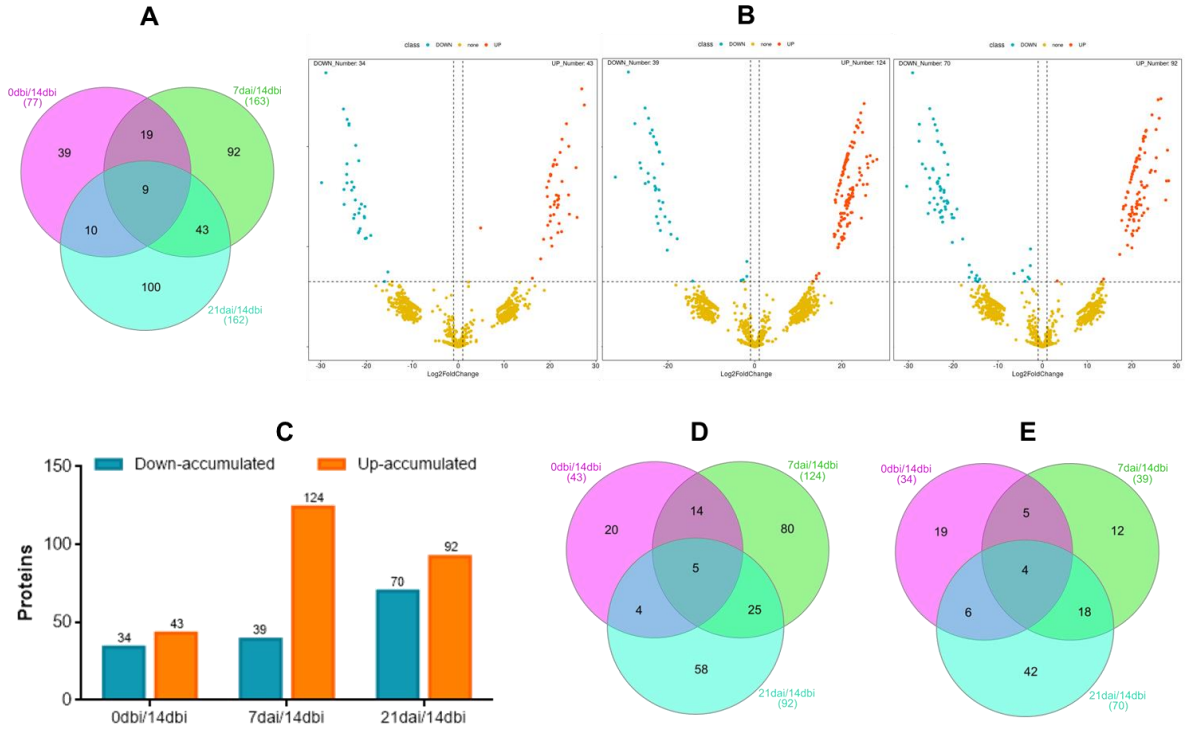


Figure 3.4. Differentially accumulated proteins (DAPs). **A**) Venn diagram showing the total number of DAPs among the different days compared with 14dbi (control): 0dbi/14dbi, 7dai/14dbi, and 21dai/14dbi. The overlapping regions correspond to the number of shared DAPs. **B**) Volcano plots showing the distribution of DAPs among the different sampling days comparisons: 0dbi/14dbi, 7dai/14dbi, and 21dai/14dbi (panels from left to right). Down-accumulated proteins are indicated with blue dots on the left side of the plots. Up-accumulated proteins are indicated with orange dots on the right side of the plots. **C**) Number of DAPs in the different comparisons. **D**) Venn diagram of the up-accumulated DAPs (LFC≥1; P ≤ 0.05). **E**) Venn diagram of the down-accumulated DAPs (LFC≤1; P ≤ 0.05).

When comparing 0 dbi versus 14 dbi (Figure 3.5A), the most enriched biological processes were gluconeogenesis, glycolytic process, fructose 1,6-bisphosphate metabolic process, carboxylic acid process, and response to toxic substances. The most significant cellular components were the plastid, apoplast, cell-cell junction, and photosystem I. The most enriched molecular functions were ion binding, fructose-bisphosphate aldolase activity, peroxidase activity, and organic cyclic compound binding. As for the KEGG pathways,

some of the most significant routes were those related to carbon metabolism and biosynthesis of amino acids and secondary metabolites.

In the 7 dai/14 dbi comparison (Figure 3.5B), the most enriched terms corresponding to biological processes were small molecule metabolic process, lignin biosynthetic process and cellular process. The most significant cellular components were the cytoplasm, cell-cell junction, and plant cell wall. The most enriched molecular functions were metal ion binding, oxidoreductase activity, and coumarate hydroxylase activity. Among the different enriched KEGG pathways, we can highlight different amino acids biosynthesis, metabolism, and biosynthesis of secondary metabolites.

In the 21 dai/14 dbi condition (Figure 3.5C), the most important biological processes were gluconate and carboxylic acid metabolic processes, precursor metabolites, and energy generation. The most significant cellular components were the cytoplasm, catalytic complex, plant cell wall, and membrane protein complex. The small molecule binding, ATP hydrolysis, phosphogluconate dehydrogenase, and catalytic activity were the most enriched molecular functions. Once again, it was found that the biosynthesis of amino acids and secondary metabolites were some of the most enriched KEGG pathways, including autophagy and starch, sucrose, and carbon metabolism.

Once the global panorama of the biological processes carried out at the different points of the SE induction process has been described, we focused on performing a manual search for all those proteins closely related to SE and plant development. We were able to identify 49 proteins accumulated on the four different sampling days, of which 18 are related to auxin, 2 to cytokinin (CK), 10 to SE process, 10 are 14-3-3 proteins, and 9 are serine/threonine protein phosphatases 2A (PP2A).

An interaction network was carried out (Figure 3.6) in which the grouping of 3 large clusters was observed: one composed mainly of the ATP-binding cassette (ABC) transporters followed by another including the PP2A proteins and the third cluster with the 14-3-3 proteins.

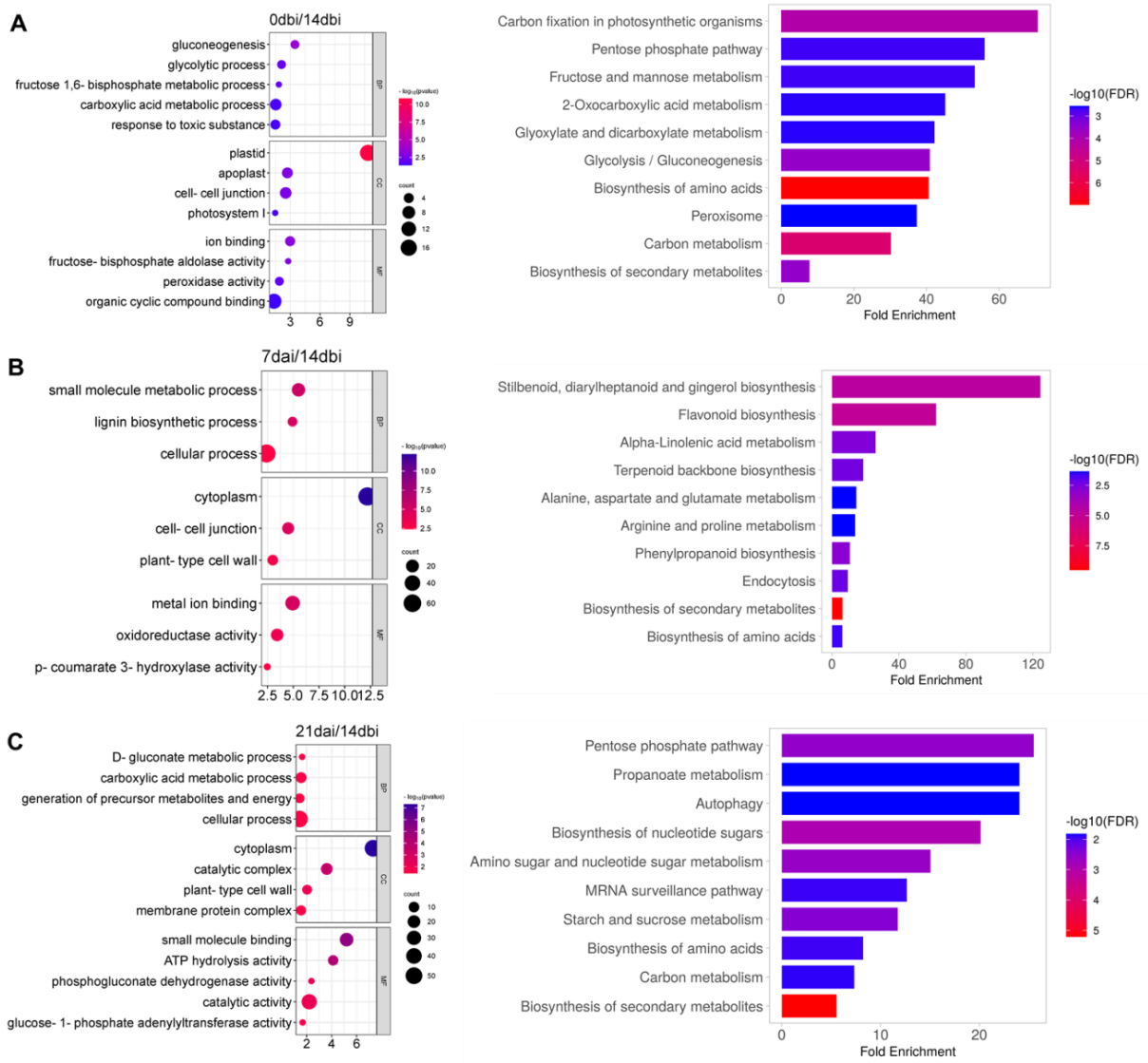


Figure 3.5. Gene enrichment analysis of up-accumulated DAPs. Gene ontology and KEGG enrichment of the most significant terms and pathways during **A**) 0dbi/14dbi, **B**) 7dai/14dbi and **C**) 21dai/14dbi comparison.

The ABC transporters of the different subfamilies interact with each other and are closely related. ABCB1 serves as a binding node with the cluster of PP2A proteins, which are connected to 14-3-3. Other essential proteins related to the metabolism of CK (CK) did not show a direct connection in this interactome.

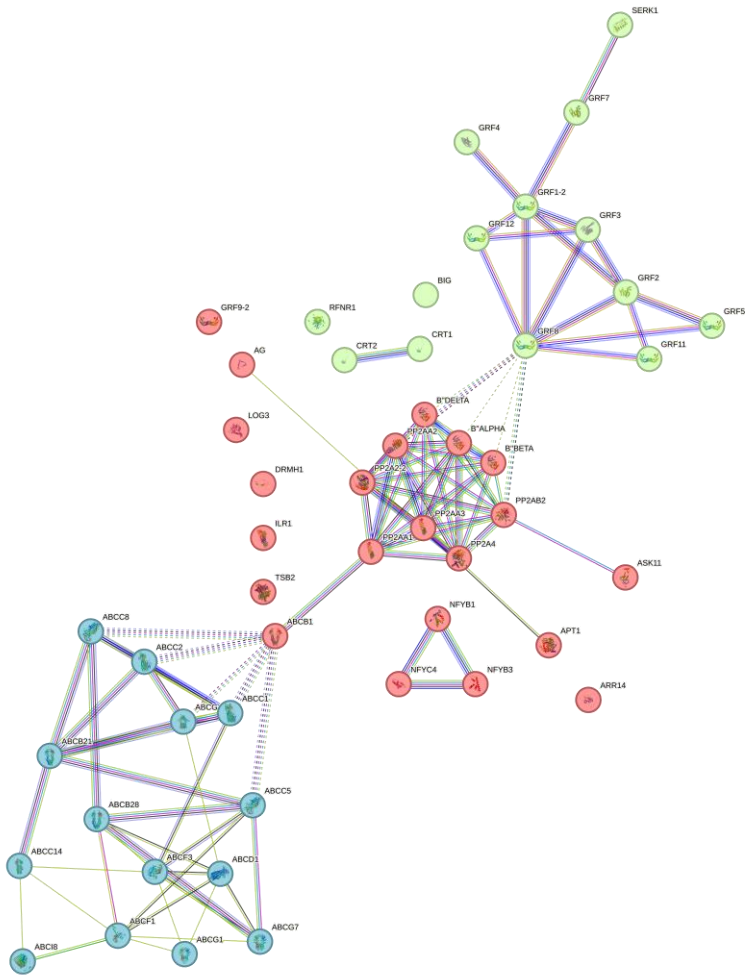


Figure 3.6. Protein-protein interaction network of SE-related proteins in *C. canephora*. Kmeans clustering against *Arabidopsis thaliana* database. Edges represent protein-protein associations. Filled nodes mean that a 3D structure is known or predicted. The names and descriptions provided by STRING for each protein in the network are specified in Annexed 3.

Proteins related to auxin homeostasis were selected (Figure 7A), which is crucial for the SE process. Within these, 15 proteins of the ABC transporter family were identified, belonging to the different subfamilies B, C, D, F, G, and I. Of these, ABCI8, ABCB21, and ABCG7 maintained a constantly high accumulation every sampling day. ABCB28 and ABCC1 were only accumulated on the control day (14dbi). ABCB1 was only accumulated on day 7dai.

On the other hand, ABCC14, ABCF3, and ABCF1 were the only ones highly accumulated on day 21dai. Another transporter known as BIG protein was also observed, the only accumulation of which occurred on day 7 dai, as well as the ILR1-like 7 protein. This last one involved the hydrolysis of auxin conjugates with amino acids.

We could also identify two proteins closely related to the metabolism of CK (Figure 3.7B): the cytokinin-riboside 5'-monophosphate phosphoribohydrolase (LOG3) and the two-component response regulator (ARR14). LOG3 accumulated only on day 21 dai, while ARR14 did so on day 7 dai; both after induction with BA.

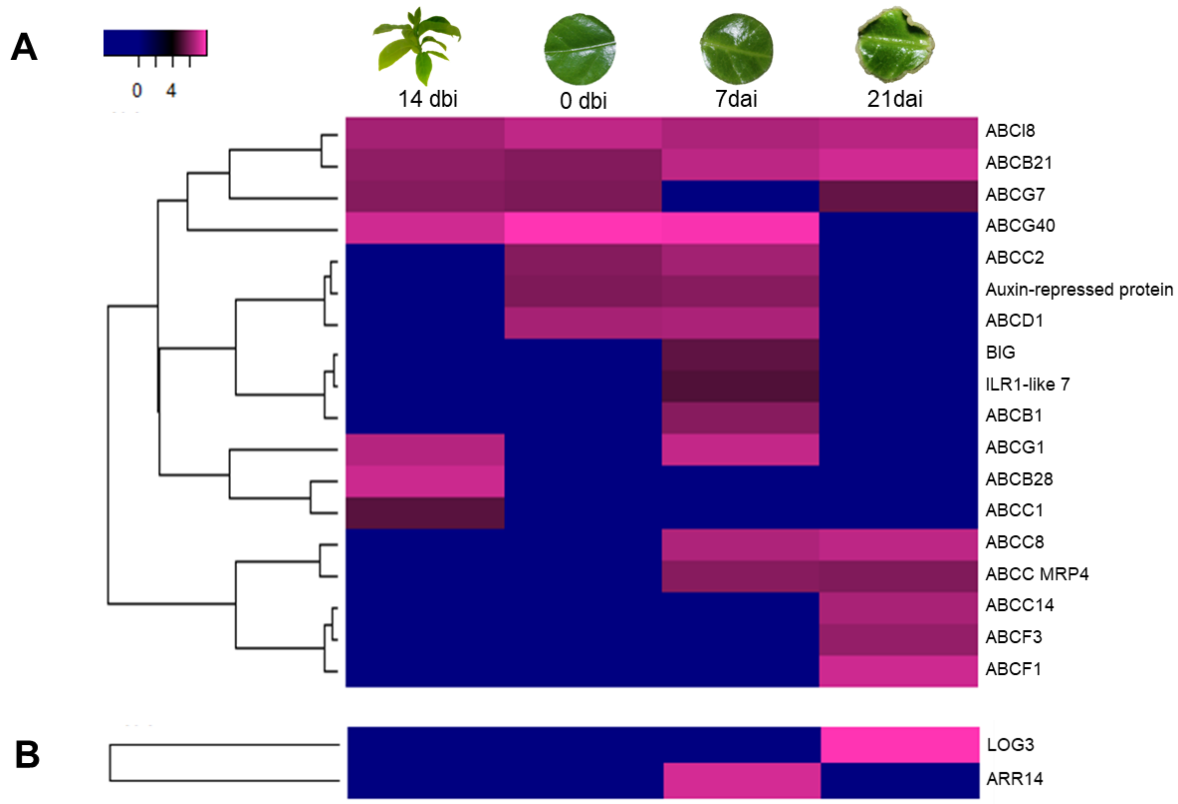


Figure 3.7. Accumulation profile of auxin and cytokinin related proteins during SE in *C. canephora*. **A**) Auxin and **B**) Cytokinin-related proteins, among the different sampling days of the SE induction. Blue, black, and pink gradients represent low, intermediate, and high accumulation, respectively.

Other SE-related proteins (Figure 3.8), such as the adenine phosphoribosyl transferase 1 and SE receptor kinase 1 (SERK1), were found to be accumulated throughout the entire process. The tryptophan synthase β-chain 2 showed no accumulation at 14 dbi. The

SKP1-like protein 11 and the nuclear transcription factor Y subunit C-4 shared a similar pattern, with no accumulation at day 0 dbi. Calreticulins were observed at 0 dbi, 7 dai, and 21 dai. The nuclear transcription factor Y subunit B-1 was found only days after the induction. The MADS-box transcription factor 58 and the nuclear transcription factor Y subunit B-10 shared a similar pattern, being accumulated only on day 14 dbi.

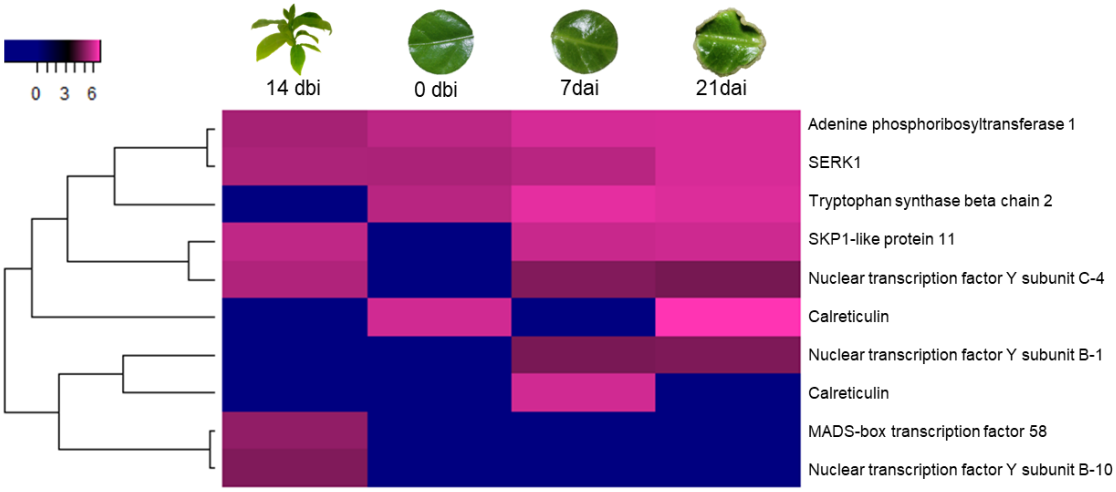


Figure 3.8. Accumulation profile of other SE-related proteins during SE in *C. canephora*. Blue, black, and pink gradients represent low, intermediate, and high accumulation, respectively.

Some 14-3-3 proteins are present during the induction of SE in different plant species, as it was in our study model (Figure 3.9A). These proteins are strongly related to PP2A proteins (Figure 3.9B), which in turn are also linked to different processes of plant development which will be discussed later. Two of the ten proteins in the 14-3-3 family found in our work (Figure 3.9A) were constantly accumulated throughout the process. The 14-3-3-like GF14 υ and 14-3-3-like-A shared a similar pattern, with no accumulation at 7 dai. 14-3-3-like GF14 κ and 14-3-3 7 showed no accumulation at 0 dbi. 14-3-3-like GF14 X protein was only accumulated at 14 dbi, while 14-3-3-like GF14 ι did so on the last day (21 dai).

Of the 9 PP2As (Figure 3.9B), those of the 65 kDa A γ and a β isoforms were highly abundant during the entire process. Both of the PP2A-2 catalytic subunits showed no accumulation at 21 dai. Those with the regulatory subunits B β and δ were only present at 21 dai, while the regulatory subunit B α was only found at 7 dai.

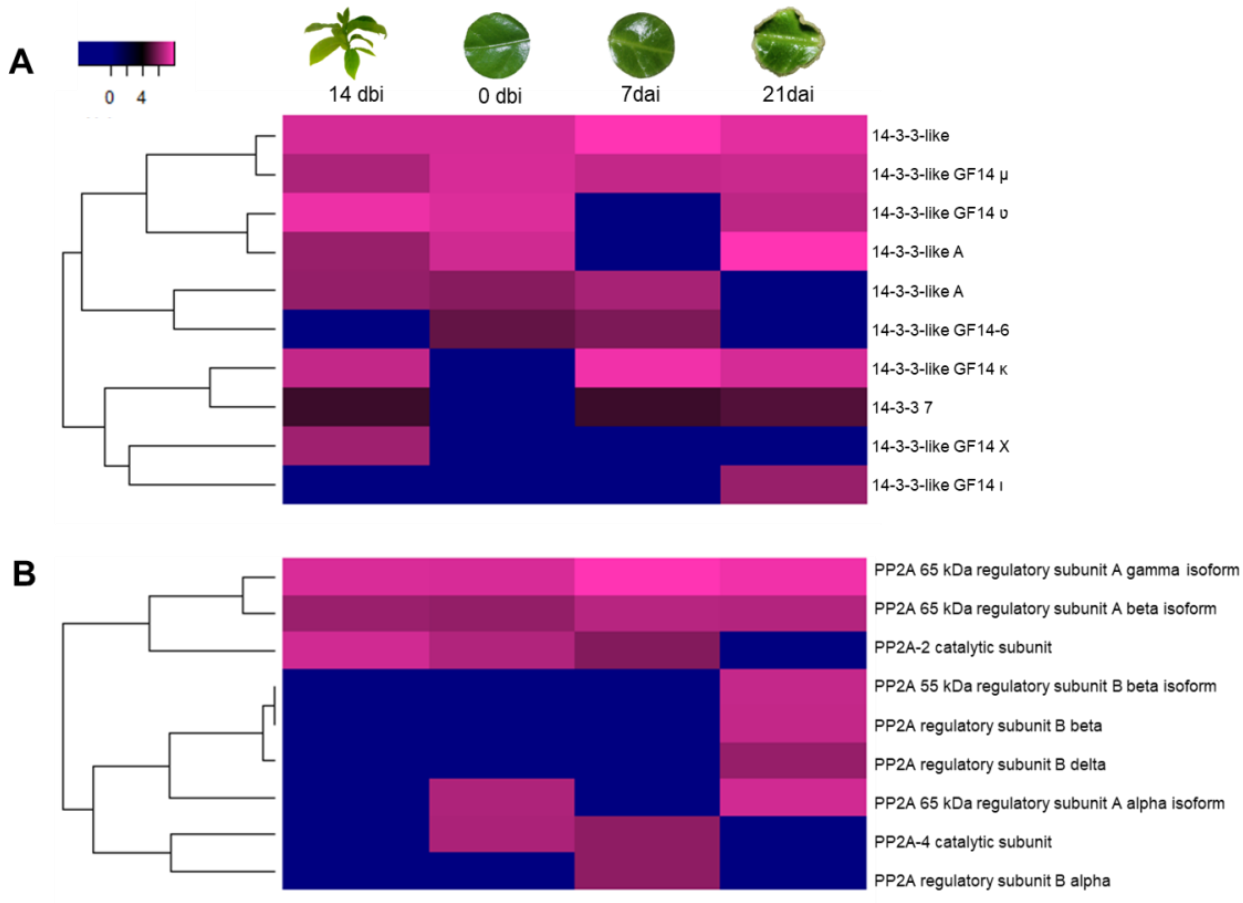


Figure 3.9. Accumulation profile of other proteins involved in plant development. **A**) 14-3-3 and **B**) PP2A proteins among the different sampling days of the SE induction. Blue, black, and pink gradients represent low, intermediate, and high accumulation, respectively.

The abundances of proteins that seem to be related to SE were used to visualize the most significant metabolic pathways in which they intervene (Figure 3.10). The LOG3 protein intervenes in the trans-zeatin biosynthesis pathway during the conversion of N^6-isopentenyl-adenosine-5'-monophosphate to N^6-dimethylallyadenine; and from *trans*-zeatin riboside monophosphate to the final product: *trans*-zeatin. LOG3 was highly accumulated at the end of induction, where the first well-differentiated embryogenic structures were seen.

The adenine phosphoribosyltransferase 1 plays a significant role in the purine nucleosides salvage pathway. This protein catalyzes a salvage reaction involving adenine, resulting in the formation of AMP, and was highly accumulated throughout the entire process. In the

indole-3-acetic acid pathway, ILR1 converted IAA-Leu conjugate to free IAA, its active form, and was highly accumulated in 7 dai. In the tryptophan biosynthesis pathway, we found the tryptophan synthase β chain 2, mediating the reaction from indole-3-glycerol phosphate to indole and tryptophan. This protein was accumulated on all days except for 14 dbi.

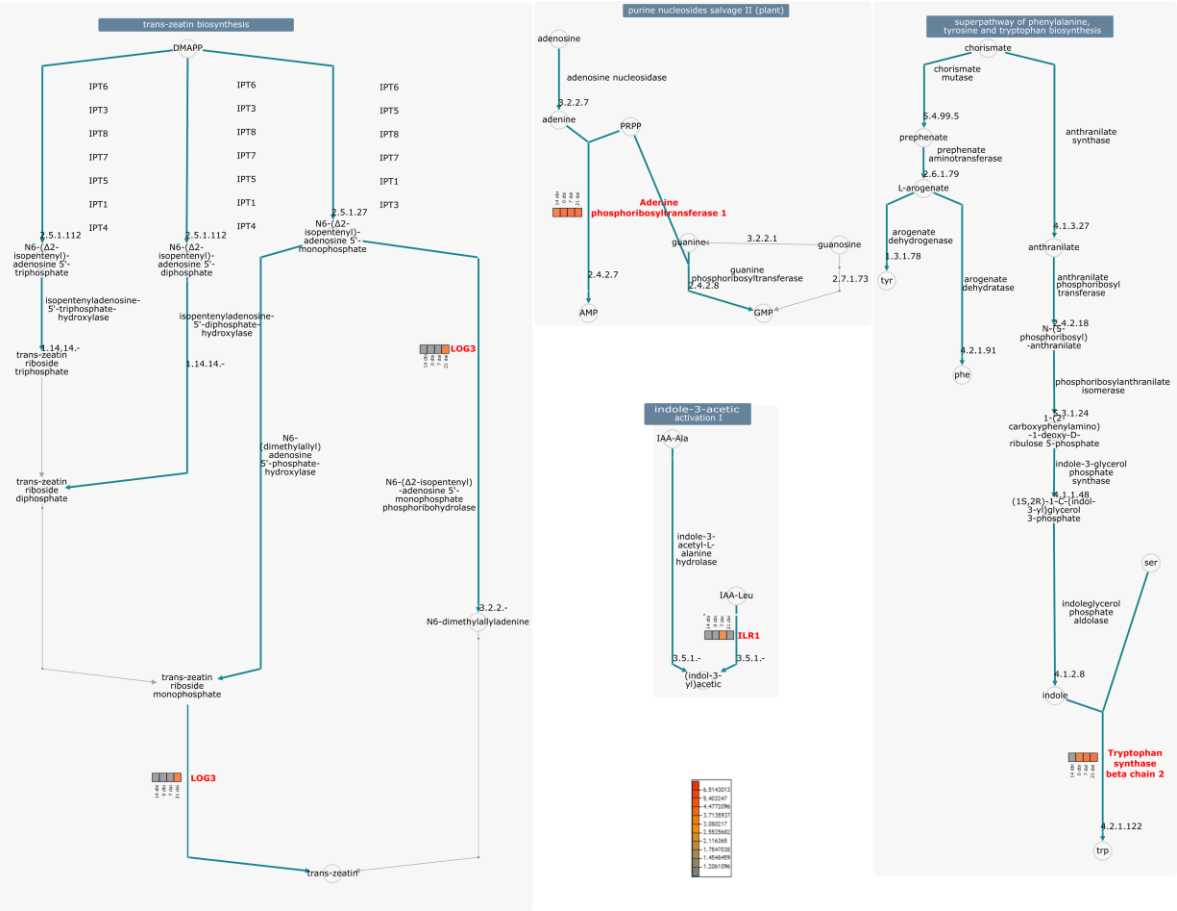


Figure 3.10. Participation of some SE-related proteins in different metabolic pathways. Abundances of LOG3, adenine phosphoribosyltransferase 1, ILR1 and tryptophan synthase proteins.

3.4 DISCUSION

C. canephora is a crop of great economic and cultural importance worldwide. SE has been a powerful biotechnology tool used in the *Coffea* genus, useful for carrying out large-scale

genetic improvement and micropropagation studies to increase material in the field. However, although SE is a tissue culture tool widely used and studied for several decades, there are still many questions regarding its regulation, which would greatly help to manipulate and optimize the process and even to understand zygotic embryogenesis. New technologies allow the carry out of studies at the genomic, transcriptomic, proteomic, and metabolomic levels and thus deepen research topics aimed at improving agriculture, the environment, human health, and biotechnology, among others (Goodwin et al., 2016; Wickham, 2016; Iquebal et al., 2015). Nowadays, many studies use transcriptomics to answer various biological questions. It allows for identifying the changes in the expression level of genes of interest in each condition and, therefore, understanding how the changes in the abundance of the transcripts control the growth and development of an organism (Goodwin et al., 2016; Rhee et al., 2006).

Due to the global importance of the genus *Coffea*, there is a growing interest in the study and generation of transcriptomic data, specifically related to changes in the genetic program that allow a somatic cell to develop into an embryo (Nic-Can and De-la-Peña, 2014). A wide repertoire of work is aimed at unraveling somatic embryogenesis in *C. canephora*, addressing different strategies. However, the use of proteomic tools, such as shotgun and mass spectrometry, in this model is still scarce.

The closest proteomic study of SE induction in *C. canephora* was that carried out by Mukul (2012). They found proteins that are secreted exclusively in the embryogenic condition and other proteins in the non-embryogenic condition. However, the identity of these proteins was not established.

During SE of *C. canephora*, it has been previously shown that the exogenous addition of PGR is crucial for forming the first embryogenic structures. During this process, there are dynamic changes in the endogenous pools of auxin (Ayil-Gutiérrez et al., 2013) and CK (Avilez-Montalvo et al., 2022). The content of auxin and its conjugates increases during pretreatment, while the expression of different genes involved in auxin homeostasis increases, such as YUCCA (Ayil-Gutiérrez et al., 2013), GH3 (Méndez-Hernández et al., 2019), PIN (Márquez-López et al., 2018), ARF and Aux/IAA (Quintana-Escobar et al., 2019). However, during the induction stage, IAA levels decrease while an increase in the expression of CK signaling genes is observed. There is a mobilization of auxin from the chloroplast to the growing areas PIN (Márquez-López et al., 2018). Now, by implementing

proteomics in this same study model, we were able to identify and quantify some of the proteins that could be vigorously participating in the induction of SE. Thus, a model was proposed where the proteins most related to the SE induction process in *C. canephora* are summarized (Figure 3.11) and later discussed.

One of the crucial processes during SE is the transport of auxin. This transport is mediated by different families of proteins, among which the ABC transporters stand out. ABC transporters' family is ubiquitous and divided into eight groups: A-G and I (Geisler et al., 2017). They regulate the transport of auxin and other molecules, such as lipids, sugars, and polysaccharides. This work identified ABCs belonging to families B, C, D, F, G, and I.

The ABCs in which the relationship with IAA has been demonstrated belong to subfamily B: ABCB1, ABCB4, ABCB19, and ABCB21 (Xu et al., 2014). We found that ABCB21 was accumulated on each sampling day, while ABCB1 was only accumulated seven days after induction. These two proteins were a crucial player in the SE of *Lilium pumilum* (Song et al., 2020). It is also essential in cell differentiation, as found in *in vitro* tissues of *C. arabica* (Quintana-Escobar et al., 2021). The only ABC proteins detected on day 21 were ABCC14, ABCF3, and ABCF1. However, in plants, there are only records of the function of the first two. ABCC14 has been found in response to stressful conditions and possibly involved in transporting heavy metals such as tomatoes (Su et al., 2021) and peanuts (Yu et al., 2019). For its part, ABCF3 seems to be involved in the development of thylakoids in chloroplasts (Faus et al., 2021). CK, like auxins, can be transported by large families of proteins, including PUP, ENT, and ABC subfamily G. In our model, we identified some of them, but there are no previous reports of their participation in other plant species.

There are few studies where the BIG protein has been identified in different processes of plant development (López-Bucio et al., 2005), tissue differentiation (Quintana-Escobar et al., 2021) and in the coordination of some PGR pathways (Kanyuka et al., 2003). Its function is associated with the polar transport of auxin, in addition to participating in vesicular trafficking and targeting of auxin transporters such as PINs in the endocytic pathway cycling, and as a mediator of auxin in pericycle cell activation promoting root hair elongation (Zazimalová et al., 2007; López-Bucio et al., 2005; Blakeslee et al., 2005). However, we have not found any study that demonstrates its effect on SE (Casanova-Sáez et al., 2021). One way to regulate endogenous IAA levels is through conjugates with amino acids. When the cell requires free auxin, the conjugates can be hydrolyzed by

amidohydrolases such as IAA-LEUCINE RESISTANT1 (ILR1) or ILR1-LIKE (ILL) to return to their active form. In cotton, a decrease in ILR/ILL expression was observed as the SE induction process progressed and until the development of the embryos, as occurred in our model (Yang et al., 2012). It is also worth mentioning that during the induction stage, proteins related to auxin signaling, such as SKP1 and the 26S proteasome, were found (data not shown). The SKP1 protein is part of the SCF complex where the auxin is perceived, and the Aux/IAA proteins are ubiquitinated and degraded via the 26S proteasome. This way, the transcription of auxin-responsive genes that participate in SE occurs. All of the above could suggest that the ABC, BIG, and ILR1 proteins found on 7 dai are important in regulating auxin flow, leading to the future formation of embryogenic structures at 21 dai.

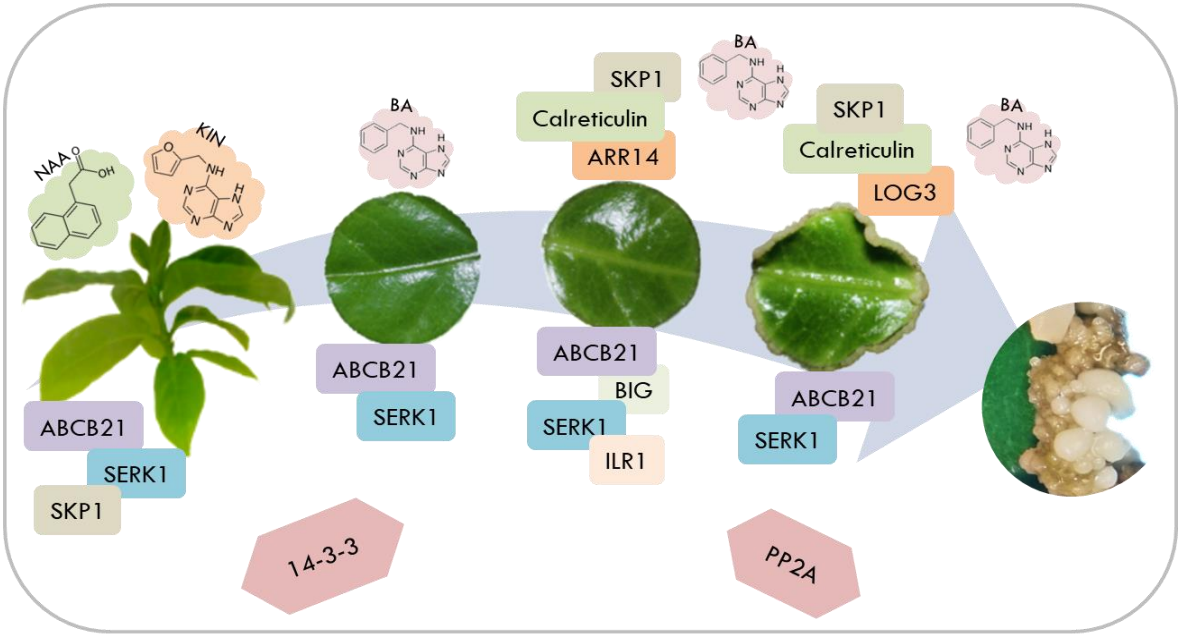


Figure 3.11. Model of SE induction in *C. canephora* and the main proteins found to be involved during the process.

In the case of auxins, endogenous IAA increases in response to adding exogenous PGR to the culture medium. However, in CK, this panorama has not been fully demonstrated (Avilez-Montalvo et al., 2022; Saptari and Susila, 2018). Identifying the accumulation of proteins related to CK metabolism during the induction stage, in which BA is added to the

culture medium to initiate SE, could suggest that endogenous changes of these PGRs also occur in response to their exogenous addition.

The genes LONELY GUY (LOG) and ARABIDOPSIS RESPONSE REGULATOR (ARR) are within the components of CK homeostasis. The first ones are involved in nucleotide activation. That is, in the same way that occurs in auxins, the active form of CK can be obtained through the synthesis of cytokinin-ribotides mediated by the cytokinin-riboside 5'-monophosphate, also known as LOG (Kieber and Schaller, 2018). In *A. thaliana*, LOG3 expression levels gradually increased towards the last day of induction, as occurred in our model at day 21 dai (Saptari and Susila, 2018). On the other hand, ARRs participate in CK signaling. In previous work in our laboratory (Avilez-Montalvo et al., 2022), various member genes of the ARR family were detected, the majority of which had a high expression the first days after induction, while on the last day (21 days), this expression was considerably reduced. Again, we detected the same pattern in this study, but now with a proteomic approach.

Although a differential accumulation of ADENINE PHOSPHORIBOSYL TRANSFERASE (APRT) was not observed, it is important to mention its role in other plant development processes and SE. This enzyme converts adenine to AMP in a single step and is part of plants' purine nucleosides salvage pathway. APRTs recycle adenine into adenylate nucleotides. They can also use CK as substrates since they are adenine derivatives with N^6 substitutions (side chains of different lengths and structures), which control their biological action. It has been hypothesized that CK biosynthesis and their interconversion depend on APT activity (Allen et al., 2002). As in the case of auxins, CK also seems to regulate their active form (Chen et al., 1982). In *Picea glauca*, fluctuations in the components of the purine salvage pathway were found (Ashihara et al., 2001). The above suggests a key point of regulation that determines the end of cell proliferation of the proembryogenic tissue and the beginning of embryo development.

We found a highly accumulated calreticulin on day 21 dai. Calreticulin was discovered by analyzing Ca^{2+}-associated proteins in spinach, and a high homology was subsequently found between its counterparts in mammals. The biological function of these proteins is inferred to be the regulation of Ca^{2+} signaling, modulation of gene expression, and as molecular chaperones (Jia et al., 2009). There are no current reports that report its participation in the SE. However, studies carried out on *Nicotiana plumbaginifolia* (Borisjuk

et al., 1998) and *Dacus carota* (Libik and Przywara, 2000) a few decades ago reported an essential activity of calreticulin in zygotic (ZE) and SE. In *Nicotiana*, maximum calreticulin activity was obtained during the early stages of the SE and in response to auxin, while in the ZE it accumulated in the embryo proper. In *Dacus*, the accumulation of this and other Ca^{2+}-associated proteins was localized in the protoderm of somatic embryos. With the above, it was concluded that the function of calmodulin in embryogenesis is to bind Ca2+ and store it for the correct development of the embryos.

One of the main transcription factors (TF) determining SE is somatic embryogenesis receptor kinase 1 (SERK1) (Méndez-Hernández et al., 2019). In Araucaria angustifolia, this TF was expressed in the periphery of the embryogenic callus and later in the embryo proper (Steiner et al., 2012). Previous studies carried out in *C. canephora* (Pérez-Pascual et al., 2018) reported that overexpression of SERK1 caused an increase in the number of somatic embryos, concluding that this TF regulates the induction of SE through activation of auxin homeostasis genes. 14-3-3 proteins interact with SERK1 to enhance embryogenic competence (Rienties et al., 2005). Using quantitative proteomics, it is possible to determine that 14-3-3 proteins, through the regulation of ATP synthases, participate in the first stages of SE in response to exogenous PGR (Zhao et al., 2015). PP2A proteins regulate histone modifications and gene expression, which are essential for forming and developing somatic embryos (Méndez-Hernández et al., 2019).

We also identified a nuclear transcription factor Y subunit B-C on induction days. However, no reports of this particular TF and its relationship with SE exist. On the contrary, it has been shown that nuclear transcription factor Y subunit alpha (NFYA) participates in embryogenesis. This family of TFs is not well characterized in plants, but a few studies suggest that it is a stress- and PGR-responsive TF closely linked to ES and embryo development (Zhang et al., 2014).

Proteomics, in conjunction with other omics and molecular tools, can provide novel information for understanding the functioning of the SE process.

3.5. CONCLUSIONS

The addition of growth regulators exogenously favors the induction of somatic embryos through a series of response reactions to this stimulus. In our model, proteins involved in the metabolism of auxin and CK were observed throughout the process. The above is an

indication of a crosstalk between both regulators. Proteins of the ABC and BIG family indicate active mobilization of IAA, while ILR1 would participate in the hydrolysis of IAA conjugates. On the other hand, the ARR and LOG proteins demonstrate that there are active CK signaling and activation pathways. Other proteins involved in SE and diverse processes of plant development were also confirmed, such as 14-3-3, PP2A, SKP1, calreticulin; as well as some transcription factors like SERK1. . There is no previous record of the proteomic study of SE induction in *C. canephora*. Hence, our results provide basic information to better understand the SE mechanism in *C. canephora* using proteomic tools, and lay the foundations for future more in-depth work. Proteomics, in conjunction with other omics and molecular tools, can provide novel information for understanding the functioning of the SE process.

CHAPTER IV

COMPLEMENTARY ANALYZES FOR THE STUDY OF SOMATIC EMBRYOGENESIS IN *C. canephora*

This chapter addresses complementary experiments for the study of SE and auxin homeostasis, including 2DE, transcriptomics, and the use of the synthetic auxin reporter DR5v2. The information presented here will form part of a third manuscript from this doctoral thesis, which will be published later.

4.1. INTRODUCTION

Two routes can be followed for proteomic studies: gel-based or gel-free. Both have advantages and disadvantages, and their use will depend on the scope of research and the biological question. One of the fundamental differences is the low cost of the analysis based on 2DE gels. There are more than 100 works related to SE (Aguilar-Hernández and Loyola-Vargas, 2018). It can be time-consuming, and many low-abundant proteins can be lost or masked. Gel-free proteomic techniques via LC-MS are becoming one of the most reliable and sensitive for identifying and quantifying proteins since it would allow the identification of less abundant proteins that would be very difficult to identify in 2DE gels (Gulzar et al., 2019). To carry out this technique, we start with a complex mixture of total proteins, digested to obtain peptides that will be ionized and subsequently identified based on the charge/mass ratio against a database. The best alternative is to carry out partial purifications. Proteins can be separated by charge or mass. This allows the enrichment of low abundant proteins and the separating of plentiful proteins such as Rubisco, producing a richer diversity mixture to be analyzed. The growing development and optimization of omics sciences allow a complete study of SE. In this sense, it has been shown that an intricate system determining the fate of the somatic cell has been identified by a comparison of transcriptome data from various stages of the growing somatic embryo, indicating that an interconnected network functions at the protein level (Aguilar-Hernández and Loyola-Vargas, 2018).

Because growth regulators play a predominant role in the induction of SE, studying their distribution, accumulation, and signaling dynamics is particularly interesting. As is known, growth regulators can act at the site of synthesis, near it, or even in distant tissues. Although there is already a record of the sites where IAA accumulates during the SE

process, it is still being determined if these are the same sites where auxin acts. That is why a synthetic auxin response reporter could provide complementary information to understand SE.

4.2. MATERIALS AND METHODS

4.2.1. Protein extraction by the TCA method

To extract the proteins, 100 mg of plantlet leaves were triturated to obtain a fine powder using a mortar and liquid nitrogen. The extraction buffer included 0.5 M Trizma base (pH 8; Sigma, T1503), 0.05 M EDTA (pH 8; Sigma, EDS), 0.7 M sucrose, 0.1 M KCl (Sigma, P9541), 0.05 M DTT (Sigma, D5545), 1 mM PMSF (Sigma, 78830) and a protease inhibitor cocktail (Sigma, P9599). For each 100 mg sample in a 2 mL Eppendorf tube, 1 mL of extraction solution was added and vortexed for 2 min. Then, they were centrifuged at 4 °C and 15,000 x g for 20 min. The upper phase was recovered in a new Eppendorf tube. The proteins were precipitated 1:1 with 20% TCA (Sigma, T6399)/acetone supplemented with 5 mM DTT, and they were allowed to pour on ice for 5 min. Then, the tubes were centrifuged at 4 °C and 15,000 x g for 5 min. The supernatant was discarded, and the pellet was washed with 80% cold acetone supplemented with 5 mM DTT three times. The pellet was allowed to dry for 3-5 minutes and then resuspended in 300 µL of 50 mM ammonium bicarbonate (Sigma, A6141) supplemented with 0.1% SDS (Sigma, L3771) by vortexing.

4.2.2. Protein extraction by the phenol-based method

With the help of a mortar and liquid nitrogen, 500 mg of plantlet leaves were ground to obtain a fine powder. The same extraction buffer of the previous method was added with 1% SDS. For each 100 mg sample, 1 mL of extraction solution was added to a 15 mL conical tube. Samples were vortexed for 5 min, followed by 30 min ice incubation and gentle shaking. An equivalent volume of phenol solution (Sigma, P4557) was added, followed by incubation on ice with gentle shaking for 30 min. Then, they were centrifuged at 4 °C and 6,000 x g for 30 min. The upper phase was recovered in a new 15 mL conical tube, and proteins were precipitated overnight at -20 °C, with five volumes of 0.1 M ammonium acetate/methanol/5 mM DTT. Tubes were centrifuged at 15,000 x g for 30 min at 4 °C, and the supernatant was discarded. The protein pellet was washed once with 4 mL 0.1 M ammonium acetate/methanol/5 mM DTT and twice with 80% acetone/5 mM

DTT. The pellet was allowed to dry for 3-5 minutes and then resuspended in the rehydration buffer [8 M urea (Sigma, U4884), 1.5 M thiourea (Sigma, T7875), 1.5% CHAPS (Sigma, C5070) and 0.05 M DTT].

4.2.3. Protein quantification (Peterson, 1977)

The standard curve was prepared according to Table 4.1. For sample quantification, 20 µL of the protein extract was used.

Table 4.1. Standard curve preparation for protein quantification.

µg mL^{-1}	H_2O (µL)	BSA (µL)	Sample (µL)
Blank	1000	0	
10	900	100	
20	800	200	
40	600	400	
60	400	600	
80	200	800	
100	0	1000	
Sample	980		20

One hundred µL of 0.15% sodium deoxycholate (DOC) was added and vortexed to each curve tube and sample. Subsequently, they were incubated for 10 min at room temperature and then placed on ice for 5 min. One hundred µL of 72% TCA was added and incubated on ice for 20 min. They were centrifuged at 3000 rpm and 4 °C for 15 min. The supernatant was quickly discarded, and the tubes were inverted to expel the remaining liquid. After this, one mL of H_2O was added and vortexed. One mL of reagent A was added [mix of equal parts of CTC (10% Na_2CO_3, 0.2% sodium potassium tartrate, 0.1% $CuSO_4$), 0.8 N NaOH, 10% SDS and H_2O. The tubes were vortexed and incubated for 10 min at room temperature. 0.5 mL of reagent B was added (a mix of one volume of Folin Ciocalteu reagent and five volumes of distilled H_2O). It was vortexed and incubated for 30 min at room temperature. The absorbance was measured in a spectrophotometer at 750 nm. The concentration of the samples was calculated according to the calibration curve.

4.2.4. Isoelectrofocusing and 2DE

After quantitation, 0.2% of ampholytes and bromophenol blue were added to 2300 µg of protein in a final volume of 450 µL. This sample was loaded on a 24 cm IPG-strip (Bio-rad), pH 3-10, and passively hydrated overnight (16-18 h) covered with mineral oil.

Afterward, isoelectrofocusing (IEF) was performed in a PROTEAN IFE Cell (Bio-rad) at 25 °C as follows: 100 V for 4 h, 250 V for 1 h, 1000 V for 1 h, 10000 V for 2 h, 10000 V until it reaches 80000 V, and 100 V for 4 h. The IPG-strip was first equilibrated for 15 min with solution I (0.375 M Tris-HCl, 6 M urea, 20% glycerol, 2% SDS, and 2% DTT) and then for another 15 min with solution II (same as solution I but substituting the DTT for 2.5% IAM) in darkness. The equilibrated strips were rinsed with running buffer and then placed on 12% acrylamide gels to perform electrophoresis at 200 V for around 10 h in a PROTEAN Plus Dodeca Cell (Bio-rad). The gels were stained with 0.1% Coomassie Blue (dissolved in 45% methanol, 5% acetic acid, and 50% H_2O).

4.2.5. Gel processing

Gels were revealed in a ChemiDoc MP Imaging system (Bio-rad). Images were captured and adjusted to 16 bits and 300 dpi with the Image Lab Software (6.0.1 Bio-rad). Visualization, matching, detection, and quantitation were performed in the Melanie software (v9, Genebio). The 14 dbi was selected as control, and the proteins with a Fold change of ≥2 and $P<0.05$ were chosen as differentials. A search for candidate proteins was carried out with the pI and MM in the SWISS-2DPAGE (https://world-2dpage.expasy.org/swiss-2dpage/) repository against the Arabidopsis database.

4.2.6. Sample preparation for transcriptomic analysis

70 mg tissues per sample and plant/fungi total RNA purification kit (Norgene, 25800) were used for RNA extraction. The quality of total RNA was verified on agarose gel at 1.5% and quantified in a Nanodrop (Thermo Fischer Scientific). Two biological replicates per sampling day (14 dbi, 0 dbi, 3 dai, 14 dai, and 21 dai) were sequenced in Novogene (Sacramento, CA). The sequencing platform and strategy used was NovaSeqPE150, producing 10 paired-end reads libraries and 12G of raw data per sample.

4.2.7. Bioinformatic analysis

Subsequent bioinformatic analyses were performed in the Galaxy platform with default options unless otherwise mentioned. Reads quality was corroborated with FastQC and then aligned with Bowtie to *C. canephora* reference genome preloaded in the platform. Read count was performed with htseq-count, where the stranded parameter was set to “no,” feature type: “gene,” and ID attribute: “name.” Raw counts of every sample were

concentrated in a data sheet and then used to estimate differential expression via DESeq2 R-package in Rstudio. Values with a fold change ≥1.5 or ≤-1.5 ($P < 0.05$) were considered up- and down-regulated, respectively..

4.2.8. Auxin and cytokinin extraction and HPLC quantification

One hundred mg of fresh-weight tissue was ground until a fine powder using a mortar and liquid nitrogen. One mg of 2,6-Di-tert-butyl-4-methylphenol (BHT) and 500 µL of extraction buffer (methanol 14/water 4/formic acid 1 v/v) were added to each sample and shaken vigorously. Samples were centrifuged at 15000 x g, 4 °C for 15 min. The supernatant was recovered in a new microcentrifuge tube and dried in a CentriVap at 40 °C until it reached ¼ of the initial volume. The Oasis MCX columns were equilibrated with one mL methanol, followed by one mL of SPE load solvent (1 M formic acid). The samples previously evaporated were diluted in 0.5 mL of SPE load solvent. This sample was applied to the column, and the flow-through was discarded. The columns were washed with 0.5 mL of SPE load solvent and 1 mL of water. The flow-through was discarded. 0.5 mL of elution solvent 1 (methanol) was applied to the columns, and the flow-through was collected in a new microcentrifuge tube (Fraction A). 0.5 mL of elution solvent 2 (0.35 M ammonium hydroxide in 70% methanol) was applied to the columns, and the flow-through was collected in a new microcentrifuge tube (Fraction B). The fractions were evaporated until dryness until further analysis.

4.2.9. Cloning of synthetic reporter DR5v2 in *Escherichia coli*

The amount of the auxin response plasmid DR5v2 (donated by the laboratory of Dr. Dolf Weijers) was increased in *E. coli* cells by heat shock (Top10 cells). For this, the preparation of calcium-competent cells was first carried out. An isolated colony of Top10 cells was picked and cultured in a test tube with 3 mL of LB medium and incubated for 12 to 16 h at 37 °C and 200 rpm. One mL was taken and inoculated from this pre-inoculum into 250 mL flasks with 50 mL LB medium. The flasks were incubated under the same conditions for 6 h. Subsequently, the culture was cooled on ice for 15 min, and the volume of each flask was placed in a 50 mL conical centrifuge tube and centrifuged at 7,000 rpm for 5 min at 4 °C. The supernatant was discarded, the pellet was resuspended in 20 mL of 0.1 M $CaCl_2$, and the centrifugation procedure was repeated. The previous step was repeated with volumes of 10 and 7.5 mL of 0.1 M $CaCl_2$. The pellet was resuspended in 2

mL of $CaCl_2$:15% glycerol and the content was distributed in 50 µL aliquots in Eppendorf tubes for storage at -80 °C. A 50 µL aliquot of competent cells was taken for its transformation, and one µL of the DR5v2 plasmid was added and mixed slightly by pipetting. It was incubated for 5 to 10 min on ice, and then placed in a water bath at 42 °C for 60 s, and, as time passed, it was cooled on ice for 5 min. Subsequently, 950 µL of LB medium without antibiotics was added, and it was incubated at 37 °C and 200 rpm for 60 min. After that, the tube was centrifuged at 5,000 rpm for 3 min, and approximately half of the supernatant was discarded.

To select the transformed bacteria, the pellet was resuspended and inoculated in Petri dishes with semisolid LB culture medium, added with kanamycin (50 µg mL^{-1}), and distributed throughout the plate. The Petri dishes were incubated inverted at 37 °C for 12 to 16 h. To carry out the purification and extraction of the plasmids, isolated colonies were taken with sterile wooden sticks and inoculated in glass tubes with a screw cap with 5 mL of LB medium, added with the selection antibiotic kanamycin. The medium was incubated for 12 to 16 h at 37 °C and 200 rpm. The Invitrogen Quick Plasmid Miniprep kit was used to purify the plasmids, with the contents of the tubes inoculated with the transformed colonies. In the case of home lysis extraction, the bacterial culture in the glass tubes was centrifuged in Eppendorf tubes at 5,000 rpm for 3 min, discarding the supernatant. The pellet was resuspended in 150 µL of solution I [50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, pH 8.0]. Subsequently, 200 µL of solution II [0.2 N NaOH, 1% SDS] were added and mixed by inversion without sudden movements. One hundred and fifty µL of solution III [60 mL of 5 M potassium acetate, 11.5 mL acetic acid, 28.5 mL H_2O, pH 6.0] were added and mixed by inversion without sudden movements. It was incubated on ice for 5 min and centrifuged at 4 °C at 15,000 rpm for 5 min. The supernatant was collected in a new tube, and one mL of absolute ethanol was added. It was slowly mixed by inversion, the centrifugation step was repeated for 15 min, and the supernatant was discarded. One mL of 70% ethanol was added to wash the bar. The centrifugation step was repeated, the supernatant was discarded, and the ethanol residues were allowed to evaporate. The pellet was resuspended in 30 to 50 µL of nuclease-free water and stored at -20 °C.

4.2.10. Transformation of *Agrobacterium tumefaciens* with the synthetic reporter gene DR5v2

First, competent cells of *Agrobacterium tumefaciens* GV3301 were obtained containing the helper plasmid pSOUP (strain provided by Dr. Dolf Weijers), for which striatum was made in Petri dishes with YEP medium added with rifampicin (100 µg mL^{-1}) and it was incubated for 48 h at 28 °C. Subsequently, isolated colonies were selected with sterile wooden sticks, inoculated into glass tubes with a screw cap containing 3 mL of YEP medium, and added with rifampicin. They were incubated at 28 °C and 200 rpm for 48 h. After the incubation time, 3 mL of the culture were transferred to a flask with 50 mL of YEP medium, and it was incubated under the same conditions until reaching an OD_{600} of 0.5. The culture was chilled on ice for 15 min and then placed in a 50 mL Falcon tube to centrifuge at 5,000 rpm for 5 min at 4 °C. The resulting pellet was resuspended in 10 mL of cold 0.15 M NaCl and incubated on ice for 20 min. Then, the centrifugation step was repeated, and the pellet was resuspended in 1 mL of 20 mM CaCl2:10% glycerol. The total volume was distributed into Eppendorf tubes in 100 µL aliquots, frozen with liquid nitrogen, and stored at -80 °C. *Agrobacterium* transformation was carried out by heat shock. In 100 µL of competent cells, approximately 10 ng of plasmid DR5v2 were added and incubated on ice for 30 min. Subsequently, they were frozen in liquid nitrogen for one min and then in a water bath at 37 °C until thawing. Five hundred µL of YEP medium was added and incubated at 28 °C for 2 to 4 h with gentle shaking. Afterward, it was centrifuged at 14,000 x g for one min, and the pellet was resuspended in 50 µL of YEP medium. The total volume was plated in 20 µL aliquots in a semi-solid YEP medium with kanamycin (50 µg mL^{-1}) as a selection antibiotic for DR5v2 until growth was observed in isolated colonies.

4.2.11. Transformation of foliar explants of *C. canephora* by agroinfiltration

The *A. tumefaciens* strain GV3301, containing the helper plasmid pSOUP (cell guard in glycerol at -80 °C or a colony), was refreshed by inoculating with a toothpick in 5 mL of YEP medium with rifampicin (100 µg mL^{-1}). It was incubated in the dark at 28 °C, shaking at 200 rpm for 48 h. Subsequently, this bacterial culture was subcultured in a 250 mL Erlenmeyer flask with 50 mL YEP medium, antibiotic, and 100 µM acetosyringone. The bacterial culture was incubated under the same conditions until reaching an optimum density of OD_{600} = 0.5. Subsequently, it was centrifuged at 2,500 x g for 10 min (Sorvall Legend Mach 1.6 R Centrifuge). The pellet was resuspended in 25 mL of MS medium supplemented with acetosyringone at a final concentration of 200 µM and 0.05% Silwet L-77.

From the seedlings grown in a preconditioning medium for 14 d, under aseptic conditions, the first and second pair of leaves were selected, and circular explants were cut with the help of a 1 cm diameter punch. The explants were transferred to the culture with competent cells of *A. tumefaciens* transformed with the plasmid of interest and subjected to infiltration under a vacuum of 400 mm Hg for 15 min. After infiltration, the explants were allowed to stand in the bacterial suspension for one h. The explants were placed in an MS medium without antibiotics or growth regulators. They were incubated at 100 rpm and 28 ºC in the dark for 48 h. Subsequently, the bacteria were eliminated with a washing solution composed of MS salts, thiamine-HCl (11.86 µM), myo-inositol (550 µM), cysteine (158 µM), sucrose (87.64 mM), NAA (0.54 µM), KIN (2.32 µM) adjusted to pH 5.8, added with the antibiotics cefotaxime (1 g L^{-1}) and timentin (0.4 g L^{-1}). The washing process was carried out thrice, with 12 h between each one. After the last wash, the transformed explants were reduced in size with the help of a smaller punch (0.8 cm in diameter). Five explants were placed per 250 mL flask containing 50 mL of modified Yasuda culture medium for SE induction and incubated in the dark at 100 rpm and 26 °C.

4.3. RESULTS

4.3.1. Proteomic analysis through 2DE

For the proteomic analysis by 2DE, another SE induction experiment was established. During the first hours of initiation of the induction process, no changes were observed in the explants at first sight. After 14 d, it was possible to observe the formation of a pro-embryogenic mass.

Representative samples of the induction were selected to perform the analysis: 14 dbi (control), 0 dbi, 7 dai, and 21 dai (Figure 4.1 A). However, the experiment continued under evaluation until the development of the embryos was observed (Figure 4.1 B).

Two methods were tested for protein extraction and precipitation: TCA and phenol. After evaluating its performance, it was decided to continue the experimental strategy using the phenol method due to its advantages of eliminating contaminants in the samples (Figure 4.1 C). The quality of the proteins was evaluated in a 1D-SDS-PAGE (Figure 4.1 D), where no notable differences were observed.

To ensure good visualization of the gels with the Coomassie stain, the protein concentration was decided to increase to 2.3 mg, starting with 500 mg of leaf tissue from

days 14 and 0 dbi and 7 and 21 dai to carry out protein extraction (Figure 4.2). The long subunit of RuBisCO was identified at a pI of approximately 6 and an MM of 52 kDa, while the small subunit was located at a pI of 6 and MM of 20 kDa. It is essential to mention that strict control was carried out on the times and volumes used during the staining and destaining of the gels so that the intensity was equal between replicates and sampling days. With the above, it was guaranteed that the intensity of the spots was due to the protein concentration and not to the effect of the staining. It was observed that the identified proteins were concentrated in an approximate pH range of 4-10 and a MM of 17-113 kDa (Figure 4.3).

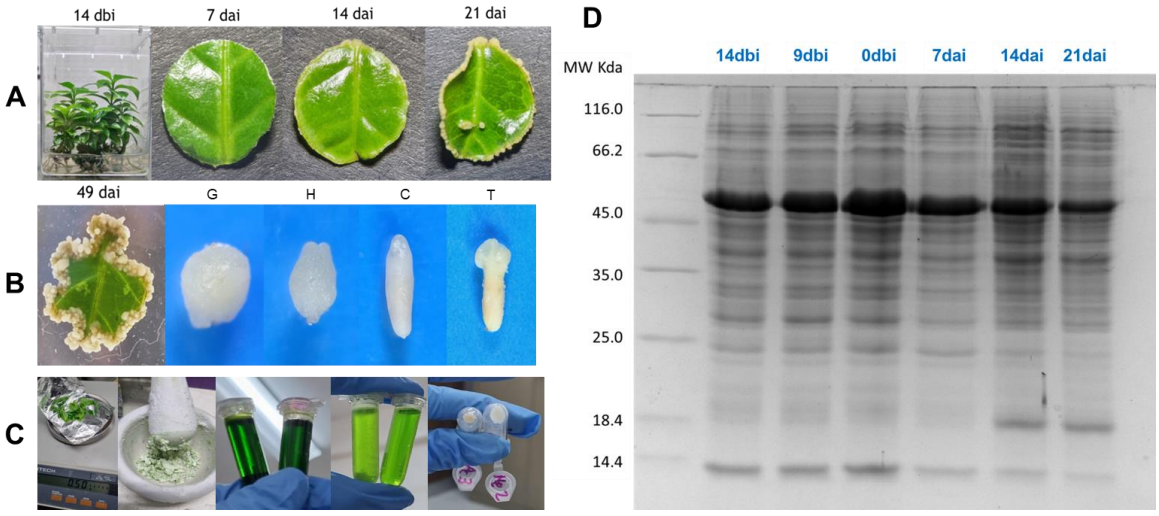


Figure 4.1 Somatic embryogenesis induction process in *C. canephora*. **A**) Induction. **B**) Development. **C**) Protein extraction. **D**) Protein pattern visualized with 1D- SDS-PAGE (10 µg).

After obtaining the gels, they were digitalized and subsequently analyzed in the Melanie v9 software. Two replicates were used for each sampling point, using the 14 dbi gels as a control to make the comparisons. Among them, 497 proteins were found, filtered with a fold change cutoff of 2 and a $P<0.05$ to select those differentially accumulated. The 23 differentially accumulated proteins are marked in green in Figure 4.4 and described in Table 4.2.

The fold change range in the differential proteins varied from 1.8 to 7.8. Spots 432, 84, and 82 showed a similar pattern, where the highest abundance was found in 21 dai, while the abundance in the first days was considerably low. These proteins could be potential

candidates as markers of embryogenesis, so it will be interesting to carry out their identification by LC-MS/MS.

On the contrary, there were proteins whose most significant accumulation occurred in the days of pretreatment before induction. Such is the case of spots 305, 112 and 294.

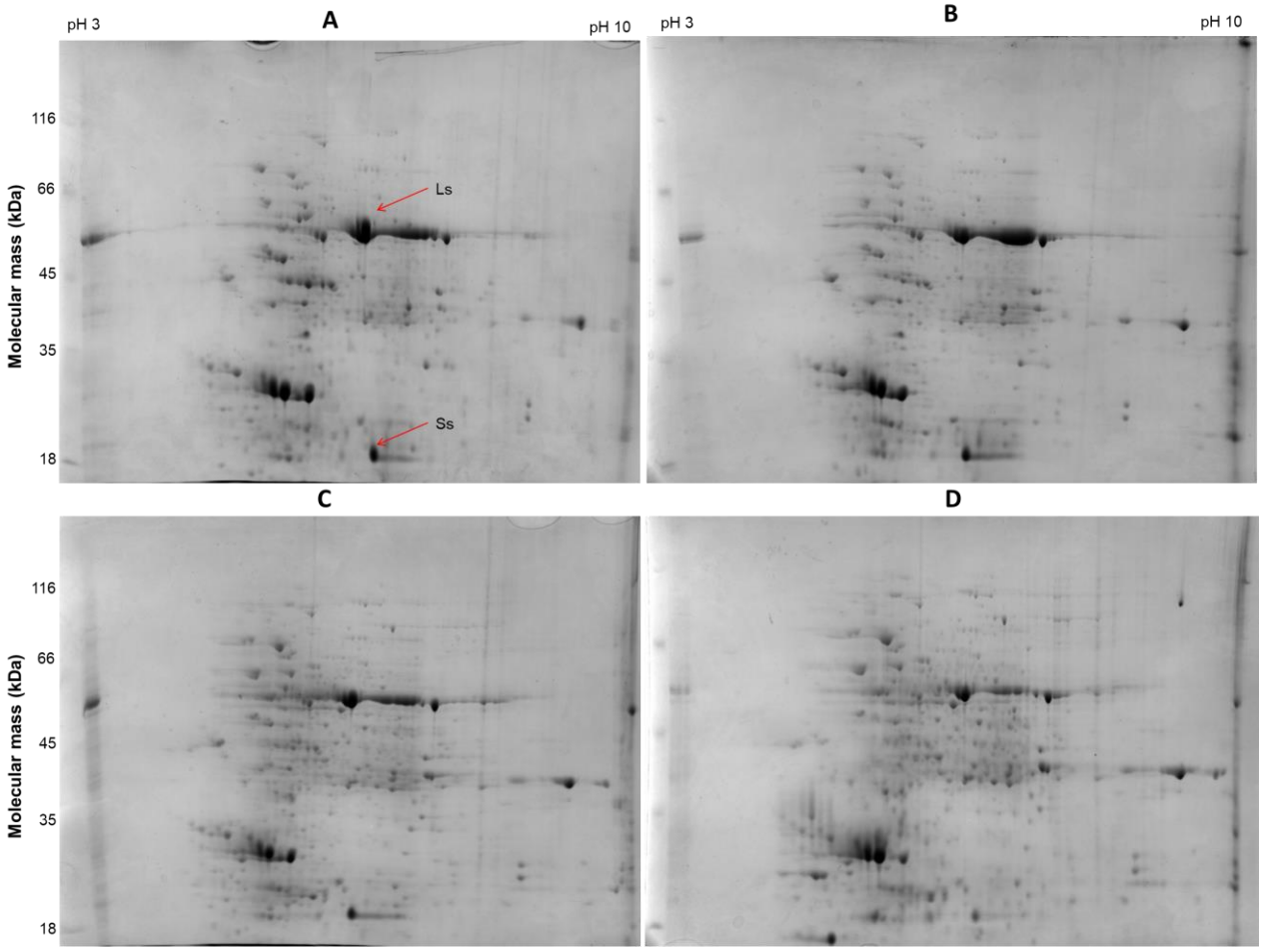


Figure 4.2. Electrophoretic pattern of the 2DE analysis of the SE induction process in *C. canephora* at **A**) 14 dbi, **B**) 0 dbi, **C**) 7 dai, and **D**) 21 dai. Red arrows indicate the location of the large (Ls) and small (Ss) subunits of RuBisCO.

After identifying the differential proteins, the pI and MM values were used to search and find the possible identity of candidate proteins whose pI and MM match or are similar in the database. It should be noted that the SWISS-2DPAGE repository is no longer maintained, so there is no update of the database, and it was only possible to make a list of the possible identity of 4 of the 23 proteins (Table 4.3).

For spot 462, the malate dehydrogenase and quinone oxidoreductase-like proteins were selected, although the difference between pI and MM was slightly different. For spot 112, the closest values were those belonging to the oxygen-evolving enhancer protein 2-1. For spot 468, two candidates with similar values were found: phosphoribulokinase and sedoheptulose-1,7-bisphosphatase. Finally, spot 458 was identified as belonging to the protein L-ascorbate peroxidase 1.

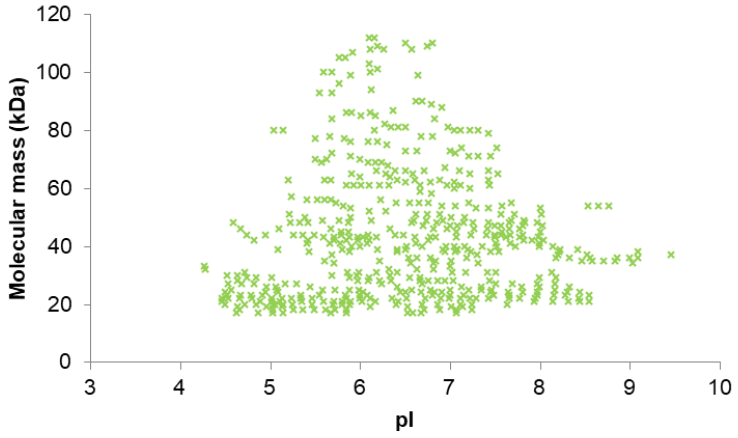


Figure 4.3. Plot for the 497 proteins of *C. canephora*. Distribution of the proteins identified in the different sampling points of SE induction, according to their molecular mass and isoelectric point.

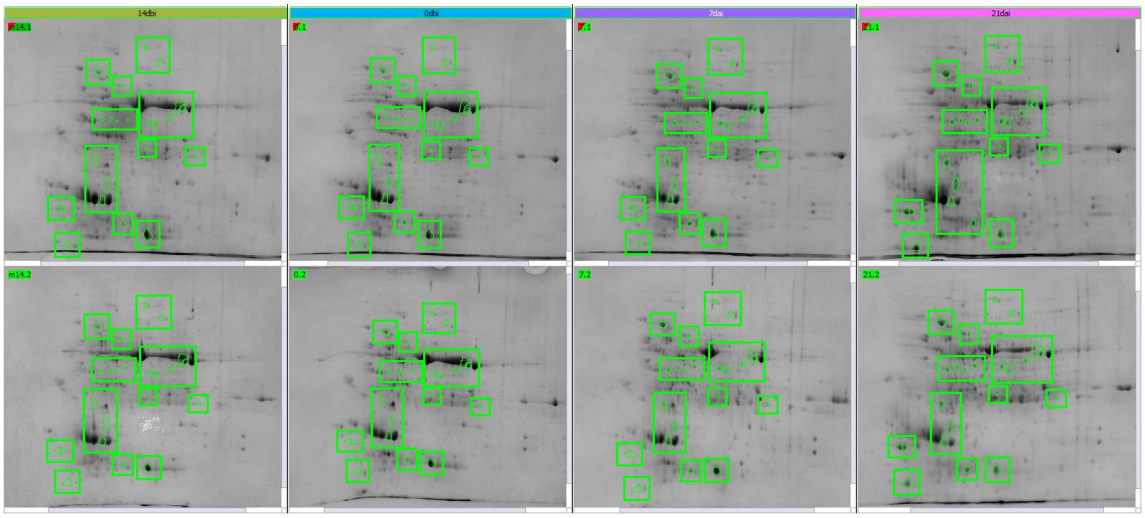
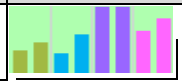
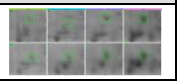
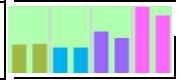
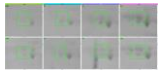
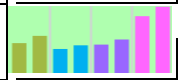
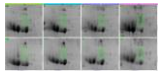
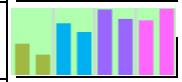
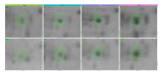
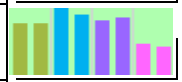
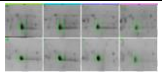
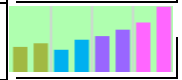
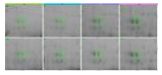
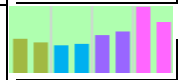
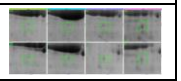
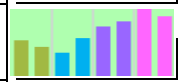
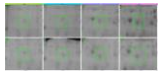
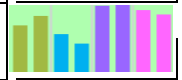
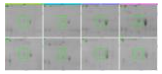
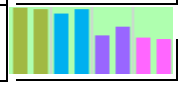
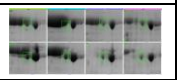


Figure 4.4. Melanie's display area compares the pattern of differentially accumulated proteins (enclosed in green boxes) between the different days of SE induction.

Table 4.2. Spots corresponding to differentially accumulated proteins at different points of SE induction were detected with Melanie.

ID	Bars	Fold	Anova (p)	14dbi	0dbi	7dai	21dai	pI	MW	Spot
432		7.80925	0.00723304	1,277,763	1,153,807	1,116,050	8,715,512	4.9	17	
84		7.5497	0.00343955	580,898	364,291	483,334	2,750,286	4.8	23	
82		7.14241	0.0174272	1,219,983	690,925	996,705	4,934,868	4.7	23	
266		6.00912	6.77E-05	556,818	398,434	1,640,529	2,394,239	6.6	48	
277		5.7845	0.00128924	388,720	349,224	1,072,398	2,020,086	5.6	49	
462		4.81313	2.08E-04	930,446	853,651	2,026,405	4,108,729	5.4	38	
449		4.34664	0.00245204	294,455	343,738	818,662	1,279,892	6.1	49	
474		4.19444	0.0398788	129,567	80,519	213,735	337,734	5.9	48	
451		4.1584	0.0283805	162,966	357,098	439,032	677,679	5.8	49	
305		3.48406	0.016842	16,581,627	19,367,872	7,943,064	5,558,988	7.2	55	
112		2.84091	0.00174505	6,567,990	8,904,596	6,257,842	3,134,413	5.6	25	
377		2.63755	0.0096621	3,348,396	4,702,288	7,255,153	8,831,577	5.5	77	
452		2.56751	0.0212418	334,866	416,355	859,773	808,041	7.6	38	

(Continuation)

ID	Bars	Fold	Anova (p)	14dbi	0dbi	7dai	21dai	pI	MW	Spot
468		2.55473	0.0279839	225,282	250,798	575,534	417,782	6.5	40	
423		2.46303	0.00244907	455,313	410,237	614,627	1,010,424	6.5	110	
458		2.40182	0.00609606	2,427,127	1,864,557	2,239,576	4,478,331	5.7	30	
480		2.37259	0.0247656	1,050,345	1,929,566	2,492,035	2,480,729	6.0	21	
9		2.15819	9.02E-04	12,851,050	15,972,770	13,979,626	7,401,002	6.5	19	
338		2.15449	0.0366073	521,764	533,219	770,298	1,124,136	6.0	64	
283		2.09524	0.0201414	934,651	819,532	1,154,526	1,717,116	7.0	50	
260		2.09238	0.0170479	1,023,937	964,962	1,658,181	2,019,069	6.7	47	
400		2.03697	0.00857957	711,073	453,925	924,631	827,934	6.8	89	
294		1.84637	0.00210888	5,845,226	5,607,078	3,831,574	3,165,799	7.3	53	

DAPs were selected according to a Fold change value of 2 and $P<0.05$. The list is sorted in descending order according to the fold change.

Table 4.3. Candidate proteins found in the SWISS-2DPAGE repository.

ID	pI	MW	Accesion	ID	pI	MW (Da)	
462	5.4	38	P93819	(MDHC1_ARATH)	5.9	38674	Malate dehydrogenase
			Q9ZUC1	(QORL_ARATH)	5.37	36133	Quinone oxidoreductase-like protein At1g23740
112	5.6	25	Q42029	(PSBP1_ARATH)	5.24	25036	Oxygen-evolving enhancer protein 2-1
468	6.5	40	P25697	(KPPR_ARATH)	6.51	40135	Phosphoribulokinase
			P46283	(S17P_ARATH)	6.58	39317	Sedoheptulose-1,7-bisphosphatase
458	5.7	30	Q05431	(APX1_ARATH)	5.73	29134	L-ascorbate peroxidase 1

Subsequent analysis needs to be carried out to confirm the identity of the proteins.

4.3.2. Transcriptomic analysis

We found 25,574 genes expressed in at least one point of the SE induction process. A notable difference was observed between 1) the maintenance stage (14 dbi), where the plants are cultivated without growth regulators, 2) the preconditioning stage with NAA and KIN (0 dbi), 3) and the induction stage (3, 14, and 21 dai) where the explants are grown in liquid medium with BA (Figure 4.5).

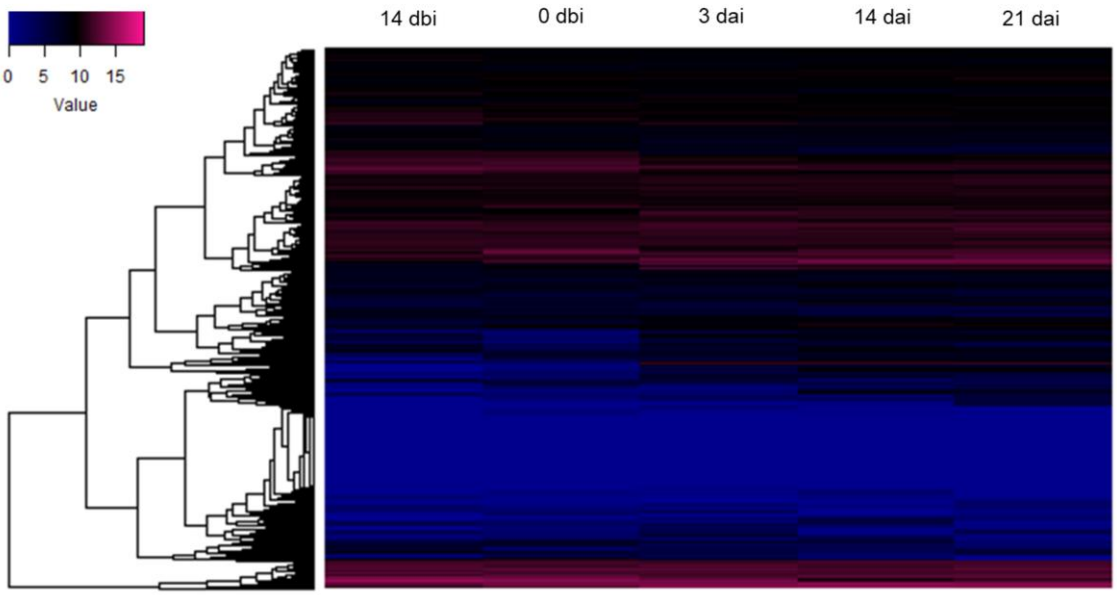


Figure 4.5. Global expression levels of 25,574 genes of *C. canephora* during different stages of the process of SE induction. Dbi=days before induction; dai=days after induction.

Of these genes, 18,315 were constitutively expressed on all sampling days. In addition, 102 genes were expressed only on day 14 dbi, 234 on day 0 dbi, 192 on day 3 dai, 171 on day 14 dai, and 190 on day 21 dai (Figure 4.6). Subsequently, differential expression analysis was performed, selecting those genes whose expression was >2 (up-regulated) or <2 (down-regulated).

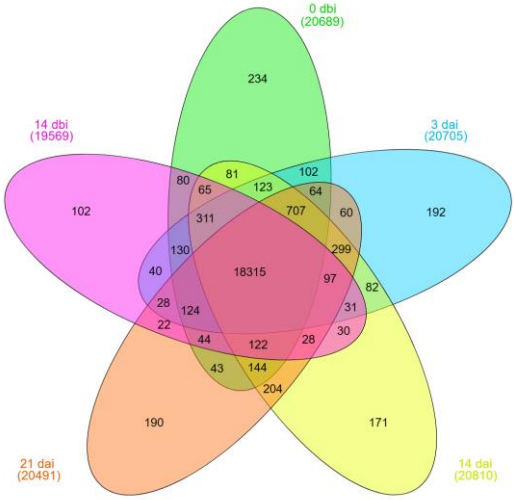


Figure 4.6. Venn diagram of the total number of genes expressed at each stage of the SE induction process. The overlapping regions correspond to the number of genes shared between each point. Dbi=days before induction; dai=days after induction.

In general, there were a more significant number of up-regulated than down-regulated genes when comparing every day against 14 dbi (control with no growth regulators). The comparison with day 0dbi was the one where fewer differential genes were found. In contrast, in comparison with day 21 dai, where the appearance of the first embryogenic structures was observed, the highest number of DEGs was found (Figure 4.7 A). Two hundred eighty-nine shared genes were identified among all comparisons, and the highest number of unique genes expressed was found in the comparison 21 dai vs 14 dbi (Figure 4.7 B).

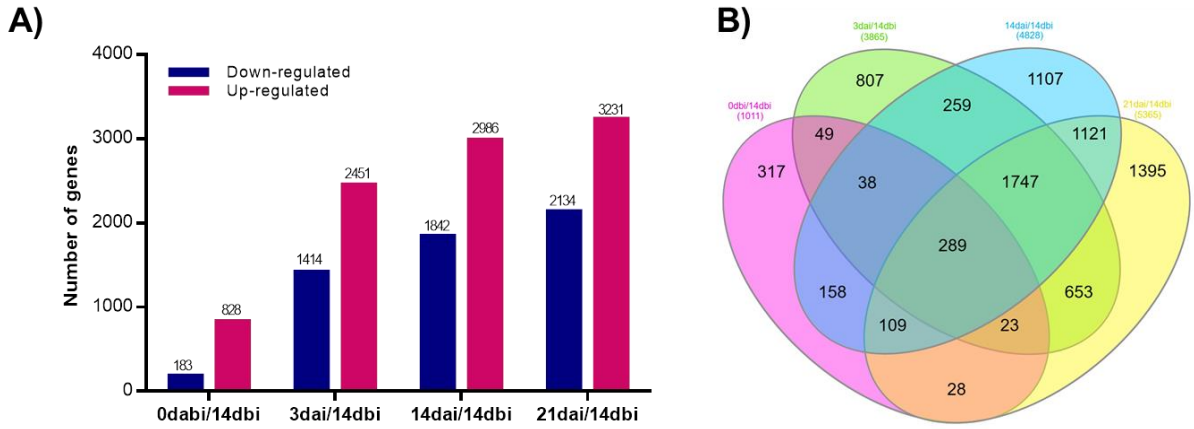


Figure 4.7. **A**) Number of differentially expressed genes at different stages of the SE induction. Different comparisons were made by using 14 dbi as control. **B**) Venn diagram of differentially expressed genes between each comparison. The overlapping regions correspond to the number of DEG shared between each condition. Dbi=days before induction; dai=days after induction.

To perform the GO analysis, the differentially expressed genes of each comparison were selected. In comparing the day 0 dbi/14 dbi, three terms corresponding to the molecular function were significant: DNA-binding transcription factor activity, phosphoenolpyruvate carboxylase activity, and transcription regulator activity. Among the biological processes, the most enriched term was a response to a stimulus. The cellular components that were more significant were membrane and extracellular matrix (Figure 4.8 A).

In the comparison of day 3 dai/14 dbi, 15 terms corresponding to molecular function were found, of which the most significant were oxidoreductase activity, enzyme inhibitor activity, molecular function inhibitor activity, catalytic activity, and chlorophyll-binding. Regarding the category of biological process, 12 enriched terms were found, of which the most significant were photosynthesis, polysaccharide catabolic process, cell wall modification, response to oxidative stress, and response to chemical and homeostatic processes. In the cellular component category, 18 terms were found, of which the most significant were thylakoid, photosynthetic membrane, photosystem, plastid, photosystem I, chloroplast, photosystem II, outer membrane, and extracellular region (Figure 4.8 B).

In the comparison of day 14 dai/14 dbi, 22 terms of molecular function were found, of which the most significant were oxidoreductase activity, catalytic activity, enzyme inhibitor activity, molecular function inhibitor activity, DNA-binding transcription factor activity, and

transcription regulator activity. Twenty-six biological process terms were enriched, of which the most significant were photosynthesis, generation of precursor metabolites and energy, cell cycle, mitotic cell cycle process, and DNA replication initiation. In the cellular component category, 27 terms were found, of which the most significant were thylakoid, photosynthetic membrane, photosystem, thylakoid membrane, and chloroplast (Figure 4.8 C).

Five terms of molecular function were found in comparing day 21 dai/14 dbi, but the most significant were oxidoreductase activity, chlorophyll-binding, FAD binding, and sigma factor activity. As for the biological process category, 14 terms were found, but the most significant were photosynthesis, plastid organization, response to stimulus, cell cycle, hydrogen peroxide catabolic process, and response to oxidative stress. In the cellular component category, 24 terms were found, of which the most significant were thylakoid, chloroplast, extracellular region and extrinsic component of membrane (Figure 4.8 D).

A manual search was carried out in the DEGs results files to identify the expression of several genes related to auxins (Aux), cytokinins (Ck), and somatic embryogenesis (SE).

As a result of this search, 18 DEGs were identified in the 0 dbi/14 dbi comparison, of which 10 were unique to this condition. In the 3 dai/14 dbi comparison, 57 genes were found, of which 15 were unique. In the 14dai/21dbi comparison, the highest number of total DEGs related to Aux, Ck, and SE was found, followed by the 21 dai/14 dbi condition, with 88 and 75, respectively. Of these last two, 25 and 13 were unique to these comparisons, respectively (Figure 4.9 A, B). It is important to note that only 2 DEGs were expressed in all conditions and were identified as GH3.1. The identity of these genes is shown in Table 4.4, as well as their fold change and *P* value.

At day 0 dbi/14 dai, a few differential genes were shared, of which almost half belonged to the ABC transporter family. Three members of the GH3 family were also identified. Those with the highest expression were GH3.3, cytokinin dehydrogenase (CKX) 9, and tryptophan synthase. Other genes include an IAA 16, a purine permease transporter (PUP) 3, the Somatic embryogenesis receptor kinase 1, a calreticulin-3, and a MADS-box (Table 4.4).

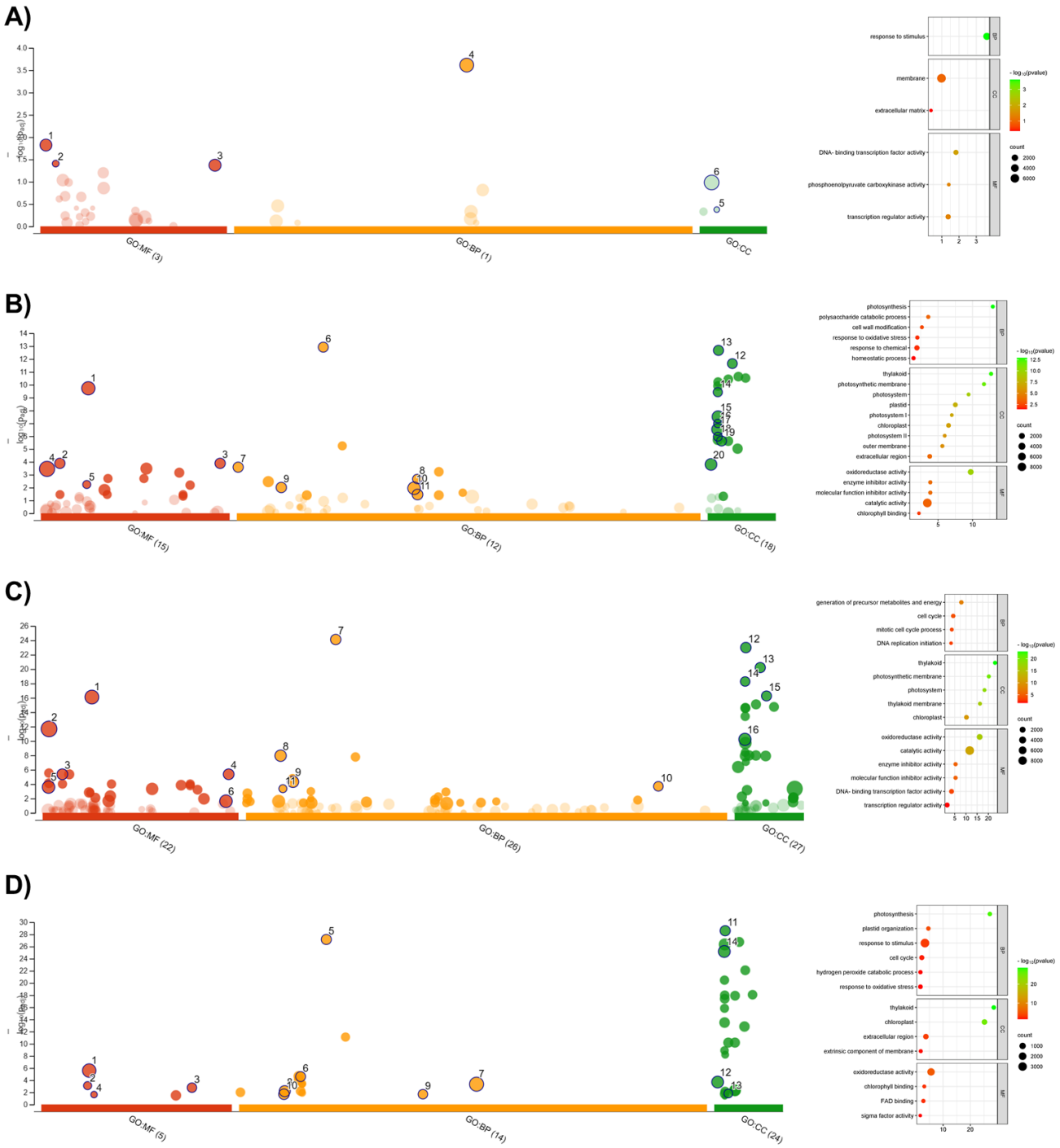


Figure 4.8. Gene ontology analysis of up-regulated differentially expressed genes found in **A**) 0 dbi/14 dbi, **B**) 3 dai/14 dbi, **C**) 14 dai/14 dbi, **D**) 21 dai/14 dbi conditions. MF=Molecular function, BP=biological process, CC=cellular component.

In the 3 dai/14 dbi condition, members of the ABC, GH3, IAA-amino acid hydrolases (ILR), IAA, Flavin-containing monooxygenases (YUCCA), auxin response factors (ARF), and Tryptophan aminotransferase-related (TAR) families were identified. , Cytokinin riboside 5'-monophosphate phosphoribohydrolases (Lonely guy/LOG), Two-component response regulators (ARR/PRR), WUSCHEL-related homeobox, CLAVATA, purine permeases (PUP), serine/threonine protein phosphatases 2A (PP2A), 14-3-3, nuclear transcription factor Y (NF-Y) and MADS-box were present. High expression of LOG5, ABCG11, ABCC10, YUCCA2, LOG1, and ARF9 was observed (Table 4.4).

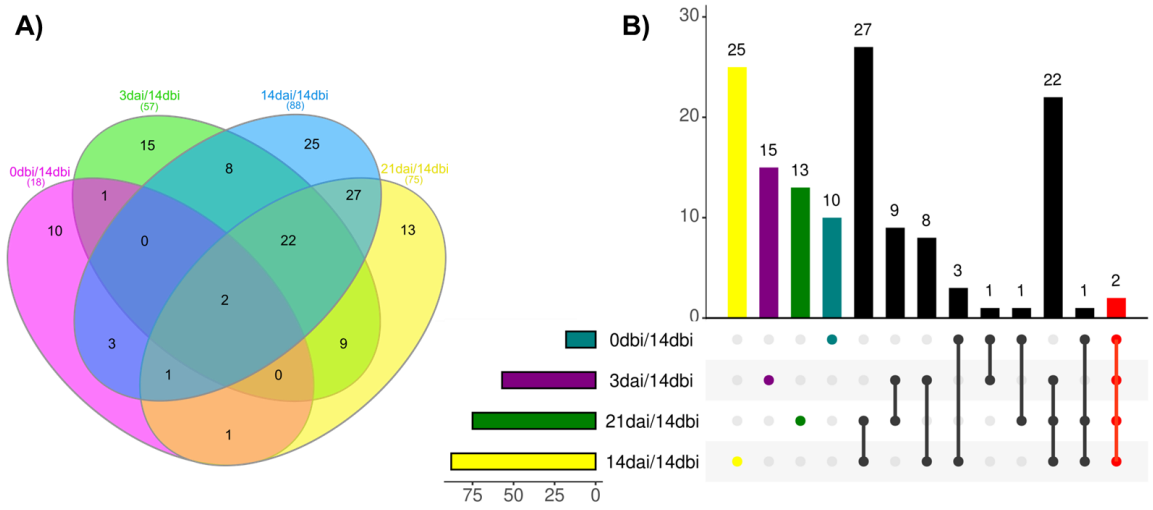


Figure 4.9. Number of DEGs related to Aux, Ck, and SE. **A**) Venn diagram. **B**) UpsetR plot. The overlapping regions correspond to the number of shared proteins between conditions. Dbi: days before induction. Dai: days after induction.

In the comparison where the largest number of differential genes was found, 14 dai/14 dbi, genes from the ABC, ARF, TAR, GH3, YUCCA, ILR, CKX, LOG, PUP, ARR/PRR, adenine phosphoribosyltransferases (APTR), SERK1, Leafy cotyledon 1 (LEC1), WUSCHEL, CLAVATA, 14-3-3, NF-Y and MADS-box families were identified. Of the above, the up-regulated ones with the highest expression were CKX3, LEC1, WUSCHEL, ILR1, and PUP1 (Table 4.4). The same gene families mentioned above participate in the 21dai/14dbi comparison. However, the up-regulated DEGs with the highest expression were ABCG11, GH3.1, PUP9, NF-Y, MADS-box and LOG1.

Table 4.4. List of DEGs at different points of SE induction of *C. canephora*, related to Aux, Cks and SE.

0dbi/14dai				3dai/14dbi				14dai/14dbi				21dai/14dbi			
Id	LFC	*P* value	Name	Id	LFC	*P* value	Name	Id	LFC	*P* value	Name	Id	LFC	*P* value	Name
Cc04_g15820	2.21	0.0077	ABCC3	Cc02_g23880	15.74	8.98E-27	ABCG11	Cc09_g08420	8.43	0.0025	ABCG23	Cc02_g23880	16.91	2.04E-29	ABCG11
Cc01_g21810	1.81	0.0157	ABCB11	Cc05_g02690	12.37	4.20E-16	ABCC10	Cc02_g08610	4.03	0.0126	ABCG1	Cc11_g16490	9.25	3.47E-04	Putative ABCC3
Cc06_g10430	1.57	0.0455	Putative ABCB8	Cc02_g21280	10.48	6.40E-10	ABCG11	Cc01_g21810	3.48	0.0000	ABCB11	Cc02_g03450	8.52	8.91E-04	ABCB28
Cc08_g03030	-1.72	0.0188	ABCB2	Cc10_g12380	9.58	3.52E-07	Putative ABCC9	Cc04_g17200	2.96	0.0000	ABCA7	Cc07_g16310	8.34	2.32E-03	ABCF5
Cc06_g09210	-2.17	0.0229	ABCG32	Cc00_g35360	8.48	3.93E-20	Putative ABCA7	Cc00_g33050	2.49	0.0000	ABCC14	Cc01_g04960	8.25	4.03E-03	ABCG15
Cc08_g08030	-2.43	0.0049	ABCG32	Cc11_g16490	8.43	7.89E-04	Putative ABCC3	Cc00_g22210	2.43	0.0001	ABCC4	Cc04_g15240	8.12	3.84E-03	Putative ABCA12
Cc07_g04380	-2.83	0.0072	ABCB25	Cc04_g15240	6.82	4.19E-02	Putative ABCA12	Cc04_g15820	2.40	0.0000	ABCC3	Cc10_g01440	7.87	1.40E-02	ABCB13
Cc00_g20000	-6.84	0.0246	ABCG8	Cc08_g08030	4.41	4.12E-14	ABCG32	Cc05_g02690	2.23	0.0073	ABCC10	Cc06_g05430	7.85	1.00E-02	ABCG39
Cc02_g19470	4.96	0.0051	GH3.3	Cc02_g03450	2.39	3.75E-02	ABCB28	Cc00_g12830	2.12	0.0009	ABCB9	Cc09_g08460	6.52	1.42E-06	ABCC1
Cc07_g06610	3.12	0.0090	Probable GH3.1	Cc00_g09060	2.28	1.04E-03	ABCD2	Cc04_g02100	2.12	0.0016	ABCC8	Cc04_g05140	5.33	4.19E-13	ABCC13
Cc00_g01360	-3.50	0.0004	Probable GH3.1	Cc07_g16310	1.86	1.89E-15	ABCF5	Cc01_g10540	1.98	0.0000	ABCF1	Cc00_g09060	3.78	2.77E-17	ABCD2
Cc02_g11550	4.14	0.0006	Trp synthase β chain 2	Cc06_g04260	-1.53	1.21E-04	ABCB1	Cc10_g12380	1.65	0.0029	Putative ABCC9	Cc02_g35820	3.74	4.88E-06	ABCI11
Cc04_g03620	1.59	0.0351	IAA16	Cc11_g05750	-2.16	9.95E-04	Putative ABCB12	Cc07_g02790	-2.00	0.0006	ABCB15	Cc01_g10590	1.81	2.84E-07	ABCG3
Cc02_g30100	4.89	0.0230	CKX9	Cc00_g33060	-2.23	6.80E-05	ABCC14	Cc01_g04960	-2.30	0.0000	ABCG15	Cc06_g17650	1.59	4.28E-05	ABCA1
Cc10_g06800	2.71	0.0001	Putative PUP3	Cc04_g10800	-2.96	3.07E-03	Putative ABCB8	Cc08_g05640	-2.45	0.0154	ABCG14	Cc02_g00550	1.56	5.02E-03	Putative ABCB19
Cc02_g19620	1.67	0.0142	SERK1	Cc00_g04550	7.10	2.44E-02	GH3.17	Cc06_g06960	-2.68	0.0037	ABCG22	Cc05_g02690	-1.72	1.38E-02	ABCC10
Cc00_g15850	1.83	0.0016	Calreticulin-3	Cc00_g01360	5.76	1.43E-42	Probable GH3.1	Cc10_g10680	-2.78	0.0022	ABCG25	Cc01_g21810	-1.72	2.13E-03	ABCB11
Cc02_g37000	-2.53	0.0001	MADS-box	Cc05_g06700	3.41	3.33E-10	Probable GH3.5	Cc10_g01440	-2.85	0.0003	ABCB13	Cc05_g11300	-2.29	3.14E-04	ABCG10
				Cc10_g16320	2.12	1.43E-02	GH3.17	Cc07_g16310	-2.87	0.0000	ABCF5	Cc09_g08420	-3.50	1.38E-23	ABCG23
				Cc07_g06610	-1.66	4.75E-06	Probable GH3.1	Cc00_g35360	-3.23	0.0000	Putative ABCA7	Cc06_g03950	3.46	2.44E-15	ARF18
				Cc10_g12030	-3.76	1.70E-03	ILR1-like 6	Cc11_g16490	-5.57	0.0000	Putative ABCC3	Cc06_g12540	2.64	3.16E-08	ARF
				Cc01_g17790	3.27	1.44E-10	IAA12	Cc06_g05430	-9.05	0.0001	ABCG39	Cc00_g00210	2.12	1.02E-04	ARF19
				Cc06_g09670	12.22	1.46E-15	YUCCA2	Cc02_g21280	-5.21	0.0000	ABCG11	Cc00_g12260	1.70	1.00E-06	ARF2
				Cc01_g20210	3.74	3.80E-02	Putative YUCCA10	Cc10_g01900	3.58	0.0000	ARF5	Cc10_g01900	-1.99	3.14E-02	ARF5
				Cc08_g08920	3.50	3.95E-05	YUCCA6	Cc00_g12260	-1.55	0.0003	ARF2	Cc03_g04670	9.94	8.00E-05	IAA7
				Cc08_g16330	9.57	3.07E-07	ARF9	Cc02_g11300	-1.60	0.0014	ARF18	Cc02_g30730	7.51	1.75E-03	IAA16
				Cc00_g12260	4.19	6.43E-18	ARF2	Cc00_g00210	-1.95	0.0000	ARF19	Cc01_g17790	2.45	1.34E-09	IAA12
				Cc06_g03950	3.23	3.35E-05	ARF18	Cc02_g39520	-2.39	0.0255	ARF9	Cc07_g07780	2.10	1.41E-05	IAA9
				Cc01_g11410	1.74	3.89E-02	ARF4	Cc08_g16330	-2.87	0.0000	ARF9	Cc06_g13230	-3.25	6.94E-10	IAA33
				Cc02_g11300	1.63	3.83E-06	ARF18	Cc06_g13230	8.40	0.0019	IAA33	Cc10_g12030	-4.29	1.57E-02	ILR1-like 6
				Cc04_g11870	6.85	1.83E-02	TAR2	Cc01_g17790	-1.85	0.0000	IAA12	Cc00_g01360	16.07	1.32E-26	Probable GH3.1
				Cc09_g10550	16.51	2.11E-29	LOG5	Cc02_g30730	-2.56	0.0068	IAA16	Cc00_g22520	13.47	5.19E-19	Probable GH3.1
				Cc10_g01830	10.13	9.32E-09	LOG1	Cc03_g04670	-3.05	0.0139	IAA7	Cc00_g04550	9.61	5.65E-06	GH3.17
				Cc01_g10950	8.59	2.09E-04	LOG1	Cc10_g15290	2.76	0.0010	TAR4	Cc05_g06700	9.00	5.97E-22	Probable GH3.5
				Cc02_g28100	-2.94	2.58E-02	LOG1	Cc05_g13460	1.66	0.0070	Trp synthase β chain 2	Cc00_g04540	3.61	3.20E-04	GH3.17
				Cc02_g00820	8.41	6.45E-04	Putative APRR5	Cc04_g11870	-1.59	0.0162	TAR2	Cc02_g19470	-1.86	2.89E-02	GH3.3
				Cc09_g02330	3.50	4.95E-02	PRR73	Cc07_g06610	4.06	0.0005	Probable GH3.1	Cc07_g06610	-2.36	2.77E-05	Probable GH3.1
				Cc06_g03460	2.68	1.88E-06	Putative APRR7	Cc00_g04530	1.57	0.0076	GH3.17	Cc08_g08920	9.42	0.0139	YUCCA6
				Cc11_g16330	2.45	3.35E-05	Putative ARR11	Cc00_g04540	-1.98	0.0000	GH3.17	Cc01_g20210	1.86	0.0127	Putative YUCCA10
				Cc11_g14050	-1.82	1.58E-19	ARR3	Cc00_g04550	-2.46	0.0003	GH3.17	Cc11_g01360	-4.41	0.0000	YUCCA4
				Cc06_g18920	-1.92	5.50E-08	ARR9	Cc05_g06700	-4.09	0.0000	Probable GH3.5	Cc04_g11870	2.30	0.0017	TAR2
				Cc09_g09080	-1.61	3.34E-05	Probable PUP9	Cc00_g01360	-6.54	0.0000	Probable GH3.1	Cc10_g02380	-2.21	0.0141	CKX5
				Cc03_g13540	-2.88	2.05E-24	PUP1	Cc11_g01360	7.56	0.0199	YUCCA4	Cc06_g11480	-8.68	0.0021	CKX3
				Cc00_g07320	-1.98	1.55E-02	SERK1	Cc01_g20250	1.59	0.0220	YUCCA10	Cc01_g10950	10.60	0.0000	LOG1
				Cc10_g04700	-1.88	1.50E-04	WUSCHEL 4	Cc08_g08920	-2.61	0.0103	YUCCA6	Cc01_g00760	-2.63	0.0000	Probable LOGL1
				Cc00_g05100	-2.17	9.55E-07	WUSCHEL 11	Cc06_g09670	-8.10	0.0476	YUCCA2	Cc02_g00820	7.16	0.0000	Putative APRR5
				Cc10_g01480	4.19	2.92E-09	CLAVATA1	Cc10_g12030	9.87	0.0000	ILR1-like 6	Cc09_g02330	3.67	0.0000	PRR73
				Cc05_g03200	6.78	4.41E-02	PP2A γ	Cc06_g11480	14.55	0.0000	CKX3	Cc06_g03460	2.90	0.0000	Putative APRR7
				Cc04_g12250	2.29	6.84E-05	PP2A β	Cc10_g02380	5.06	0.0010	CKX5	Cc11_g16330	2.48	0.0002	Putative ARR11
				Cc00_g17460	4.98	1.09E-03	14-3-3-like	Cc02_g30100	4.77	0.0098	CKX9	Cc08_g02100	1.97	0.0000	APRR2
				Cc10_g07120	1.53	1.64E-05	14-3-3	Cc08_g01180	1.52	0.0037	CKX7	Cc06_g18920	-1.66	0.0076	ARR9
				Cc06_g06710	9.42	8.43E-07	Putative NF-Y sub A-2	Cc01_g00760	7.62	0.0207	Probable LOGL1	Cc11_g14050	-1.80	0.0000	ARR3
				Cc05_g14900	7.47	8.57E-03	NF-Y sub B	Cc10_g01830	-3.45	0.0001	LOG1	Cc09_g09160	12.44	0.0000	Probable PUP9
				Cc04_g01480	5.73	1.12E-63	Putative NF-Y sub A-3	Cc01_g10950	-4.96	0.0132	LOG1	Cc10_g15400	8.61	0.0007	PUP3
				Cc02_g28060	2.35	4.09E-03	Putative NF-Y sub A-1	Cc03_g13540	9.32	0.0000	PUP1	Cc10_g06500	-1.76	0.0011	PUP3
				Cc03_g03220	1.56	6.44E-03	NF-Y sub C-9	Cc10_g06500	2.99	0.0001	PUP3	Cc03_g11350	-1.83	0.0024	PUP3
				Cc00_g02800	9.50	1.01E-06	Agamous-like MADS-box	Cc03_g11350	2.47	0.0049	PUP3	Cc03_g13540	-3.23	0.0040	PUP1
								Cc10_g15400	-2.79	0.0000	PUP3	Cc04_g04460	-3.93	0.0000	APTR 2
								Cc09_g04610	-7.00	0.0462	Probable PUP4	Cc08_g00490	4.24	0.0000	APTR 2
								Cc11_g14050	4.62	0.0000	ARR3	Cc09_g00330	-8.69	0.0014	LEC1-like protein
								Cc06_g18920	4.56	0.0000	ARR9	Cc00_g07320	-2.44	0.0000	SERK1
								Cc08_g02100	-1.55	0.0000	APRR2	Cc04_g06330	5.81	0.0003	Putative WUSCHEL 1
								Cc06_g03460	-1.63	0.0001	Putative APRR7	Cc07_g11890	-1.57	0.0007	WUSCHEL 8
								Cc02_g00820	-2.22	0.0001	Putative APRR5	Cc10_g04700	-2.40	0.0000	WUSCHEL 4
								Cc02_g02330	-2.69	0.0001	PRR73	Cc00_g05100	-3.21	0.0000	WUSCHEL 11
								Cc04_g04460	8.06	0.0046	APTR 2	Cc02_g14220	-3.89	0.0178	Putative WUSCHEL 8
								Cc08_g12840	2.95	0.0000	APTR 1	Cc01_g12690	-3.90	0.0000	Putative WUSCHEL 2
								Cc10_g05670	1.75	0.0000	APTR 1	Cc02_g06840	-4.11	0.0000	WUSCHEL 5
								Cc08_g00490	-1.61	0.0373	APTR 2	Cc00_g17460	2.92	0.0004	14-3-3-like
								Cc00_g07310	7.19	0.0375	SERK1	Cc07_g16060	12.44	0.0000	NF-Y-alpha
								Cc00_g07320	6.93	0.0000	SERK1	Cc06_g06710	10.29	0.0000	Putative NF-Y sub A-2
								Cc10_g06160	2.65	0.0000	SERK1	Cc05_g14900	5.72	0.0001	NF-Y sub B
								Cc02_g19620	1.88	0.0000	SERK1	Cc04_g01480	1.89	0.0000	Putative NF-Y sub A-3
								Cc09_g00330	12.73	0.0000	LEC1-like	Cc07_g07970	11.38	0.0000	MADS-box 27
								Cc02_g06840	10.39	0.0000	WUSCHEL 5	Cc10_g08300	-2.73	0.0000	MADS-box
								Cc01_g12690	8.88	0.0001	Putative WUSCHEL 2				
								Cc02_g14220	7.94	0.0057	Putative WUSCHEL 8				
								Cc10_g04700	4.86	0.0024	WUSCHEL 4				
								Cc07_g11890	2.12	0.0000	WUSCHEL 8				
								Cc04_g06330	-2.05	0.0439	Putative WUSCHEL 1				
								Cc10_g01470	-2.32	0.0157	CLAVATA1				
								Cc10_g01480	-2.66	0.0014	CLAVATA1				
								Cc11_g08430	-24.22	0.0000	Putative CLAVATA3				
								Cc07_g14680	1.91	0.0000	14-3-3-like				
								Cc00_g17460	-2.16	0.0000	14-3-3-like				
								Cc05_g14900	-2.83	0.0000	NF-Y sub B				
								Cc06_g06710	-3.54	0.0000	Putative NF-Y sub A-2				
								Cc07_g07970	-7.41	0.0195	MADS-box 27				

Those genes unique to each condition are shown in red.

4.3.3. Endogenous levels of plant growth regulators and DR5v2 synthetic reporter activity

The extraction and purification of auxins and cytokinins from all sampling days was performed according to Dobrev (Dobrev et al., 2017). The IAA conjugates were found in the highest concentration (Figure 4.9), as reported in the literature and previous works in our laboratory. On the control day, 14 dbi, it was possible to detect both free and conjugated IAA (Figure 4.10). Of the above, the free IAA was the one with the lowest concentration, and the IAA-Glu conjugate was the one that was abundantly found, which gradually increased by the end of the pretreatment. After induction, only IAA-Glu and IAA-Leu conjugates, involved in degradation and storage, respectively, were detected. From day 7 to day 21, after induction, the concentration of both free and conjugated auxin increased again.

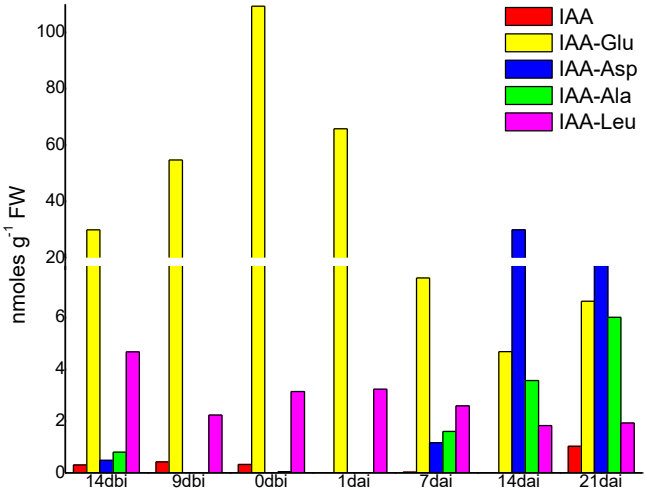


Figure 4.10. Free and conjugated auxin quantification during SE induction in *C. canephora*.

The Cks were found in smaller quantities than the Aux, as expected. The concentration of Kin was high during the pretreatment when the plants were incubated in a culture medium with added Kin. The above indicates that the plant absorbs Kin. Kin levels gradually decreased throughout the entire process. A similar behavior was observed for the

concentration of *trans*-zeatin. Unlike the previous ones, isopentenyladenine was detected at very low concentrations, with almost imperceptible changes throughout the entire SE induction process (Figure 4.11).

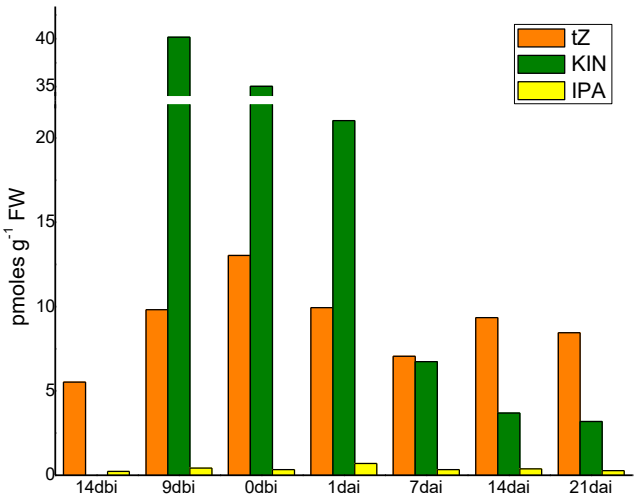


Figure 4.11. Cytokinin quantification during SE induction in *C. canephora*. *t*Z: *trans*-zeatin. KIN: Kinetin. IPA: isopentenyladenine.

The foregoing denotes the importance of the regulation of auxin and cytokinin concentration through conjugation in order to give rise to SE.

Although recordings of the auxin accumulation sites during SE already exist, a thorough knowledge of the processes and dynamics of auxins depends on the ability to subjectively and quantitatively visualize the response sites. By acting as auxin receptor substrates and enabling the observation of the transcriptional response, synthetic reporters, such as the DR5v2, have made it feasible to examine the distribution, transport, and activity of auxins in embryos and seedlings of several species.

The transformation of foliar explants of *C. canephora* was carried out by agroinfiltration with the synthetic reporter DR5v2. The process of agroinfiltration and disinfection of the explants was successful. The explants become necrotic due to the antibiotics. However, they continued with their response and the formation of embryos, although slower than the control explants.

To corroborate the transformation of the embryos, we proceeded to make histological sections of the embryos in the torpedo stage. The sections were analyzed by confocal microscopy before DNA extraction to verify the transformation by PCR.

Figure 4.12 shows the bright field images, DAPI signal, DR5v2-ntdTomato signal, DR5-eGFP signal, and the superposition of the previous ones obtained from a control sample (A) and four samples from transformed plants (B-E). In the control embryo (untransformed), only the DAPI signal was observed. For the other samples, the conditions were established to evaluate them and avoid confusing the vector signal with tissue autofluorescence.

In samples B) and C), it was possible to observe the signal of both DR5v2 (ntdTomato) and DR5 (GFP), although the latter was much less intense, as expected. DR5 (GFP) signal was focused on the nuclei while that of DR5v2 (ntdTomato) around them. In sample D) only the DR5v2 signal was observed around the nuclei and in the cytoplasm. In sample E) only a signal from DR5 was obtained. However, Liao (2015) showed that only some of the two signals will be obtained together due to the difference in the specificity of the promoter. When they appear together, DR5 tends to be less intense than DR5v2.

The embryos were placed in the germination medium (Figure 4.13). After approximately six months, it was possible to obtain fully regenerated seedlings from which a small sample was taken to perform DNA extraction. Subsequently, the transformation was confirmed by PCR, using specific primers for the synthetic reporter DR5v2: eGPF and ntdTomato (Figure 4.14).

Once the transformed plants were selected, the SE induction process was repeated to generate a time course of the auxin response site at the different points.

Transverse cuts were made to the leaves in pretreatment. In the control samples (untransformed), only the DAPI signal was observed, as expected (Figure 4.15 A). While in the transformed sample, we could detect both the GFP reporter signal and the ntdTomato (Figure 4.15 B). According to these images, IAA is perceived in the chloroplasts, in some cases in the cytoplasm, and even in nuclei (Figure 4.15 C). It is important to remember that the auxin perception and signaling machinery is located in the nucleus. Now, towards 35 dai where a well-defined embryogenic mass is already observed, we can observe a

close-up of one of the edges of the leaf, where cells with dense cytoplasm and actively dividing can be seen (Figure 4.15 D).

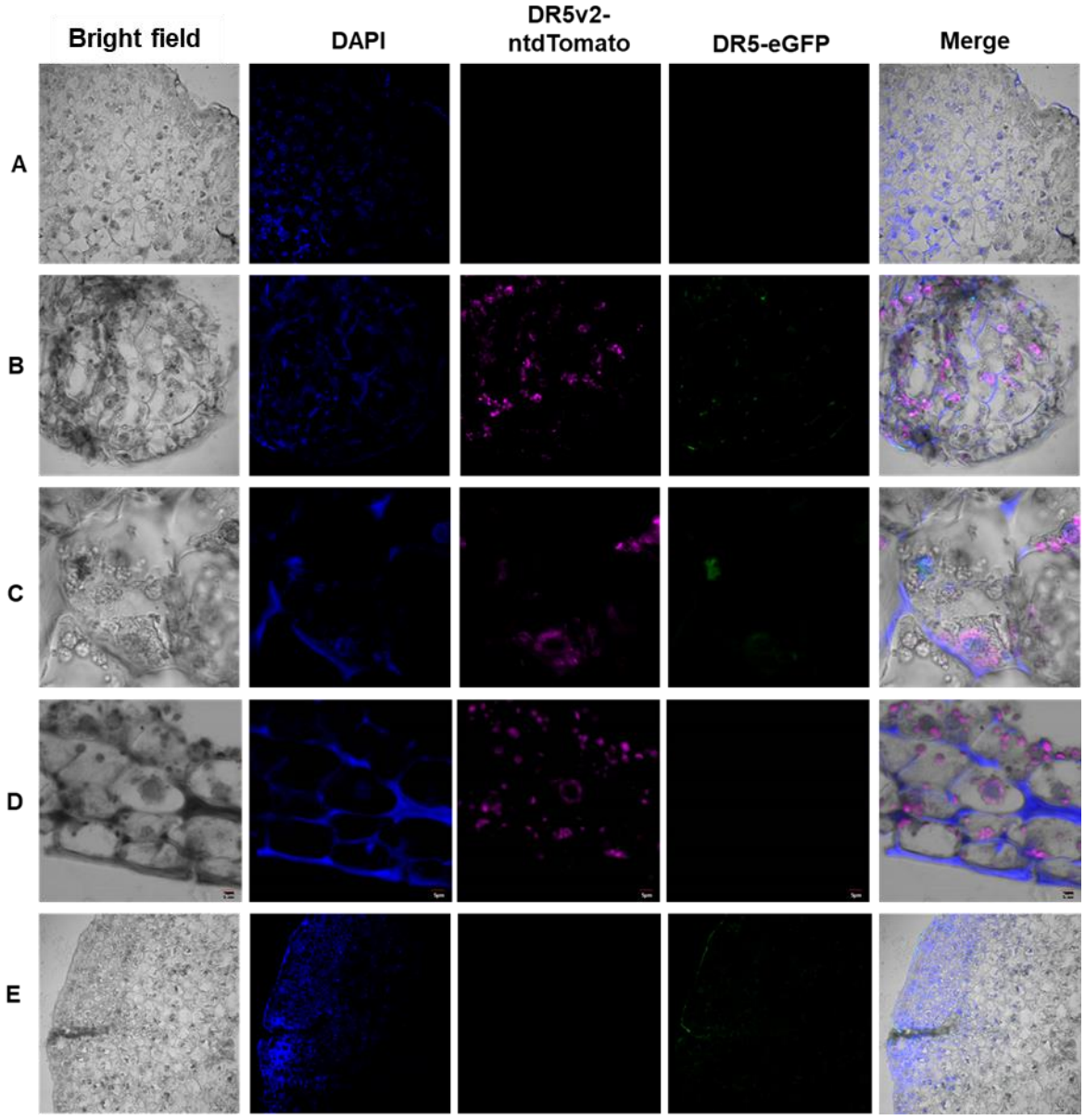


Figure 4.12. Visualization of the auxin response site by confocal microscopy from transformed embryos of *C. canephora* with DR5v2. **A**) Control embryo. **B-E**) Transformed embryos.

Figure 4.13. Embryos transformed with DR5v2 auxin response synthetic reporter placed in the medium for germination (upper panel). Regenerated seedlings from the transformed embryos (bottom panel).

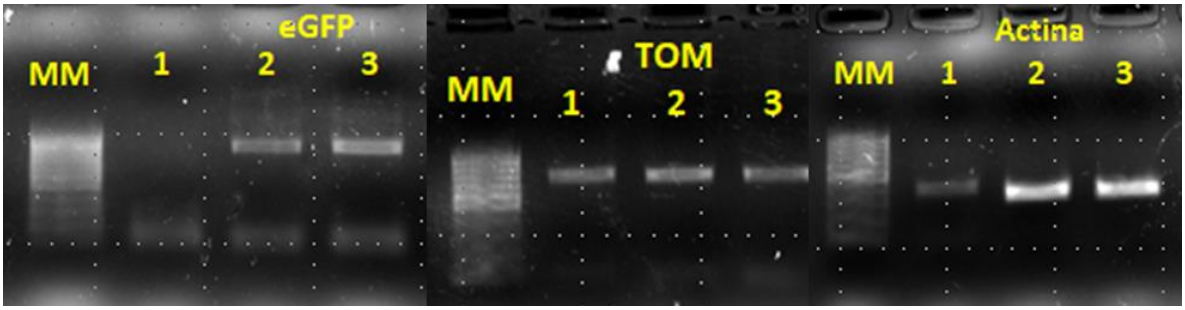


Figure 4.14. PCR confirmation of plasmids in three transformed seedlings: eGFP, ntdTomato, and the internal reference gene (actin).

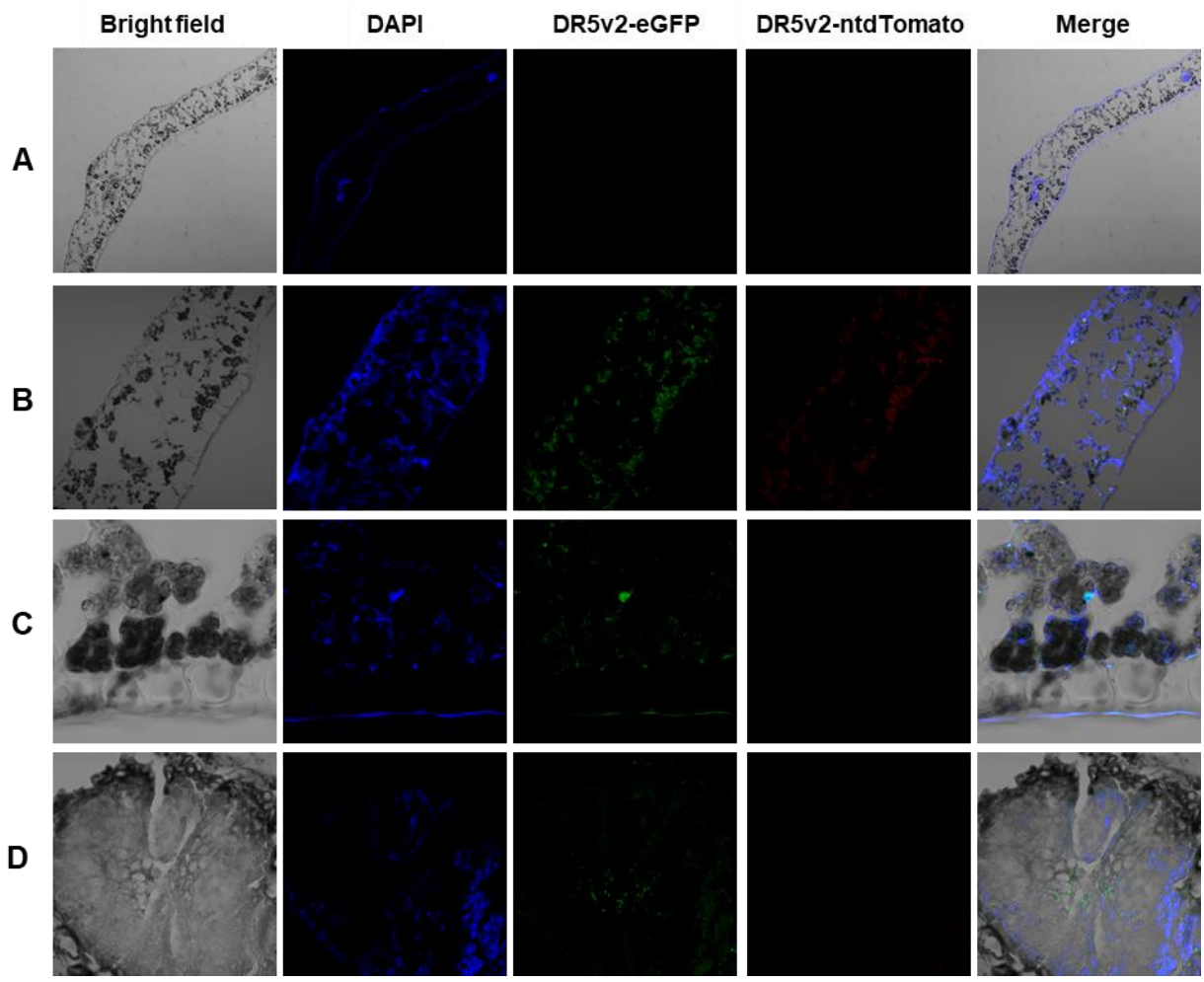


Figure 4.15. Transverse sections of leaves *C. canephora*. **A**) Control. **B**) Transformed. **C**) Transformed, 21 dai. **D**) Transformed, 35 dai.

A first approach was made to samples of embryos in the last stages of development. In the torpedo embryo from the control treatment (untransformed), we only observed the DAPI signal (Figure 4.16 A and B). We could detect the signal from both reporters in a transformed cotyledonary embryo. This signal was focused on the edges of the embryo and the procambium (Figure 4.16 C-E), but with greater intensity at the ends of the cotyledons (Figure 4.16 F) and in the basal part of the embryo where the root meristem is located (Figure 4.16 G).

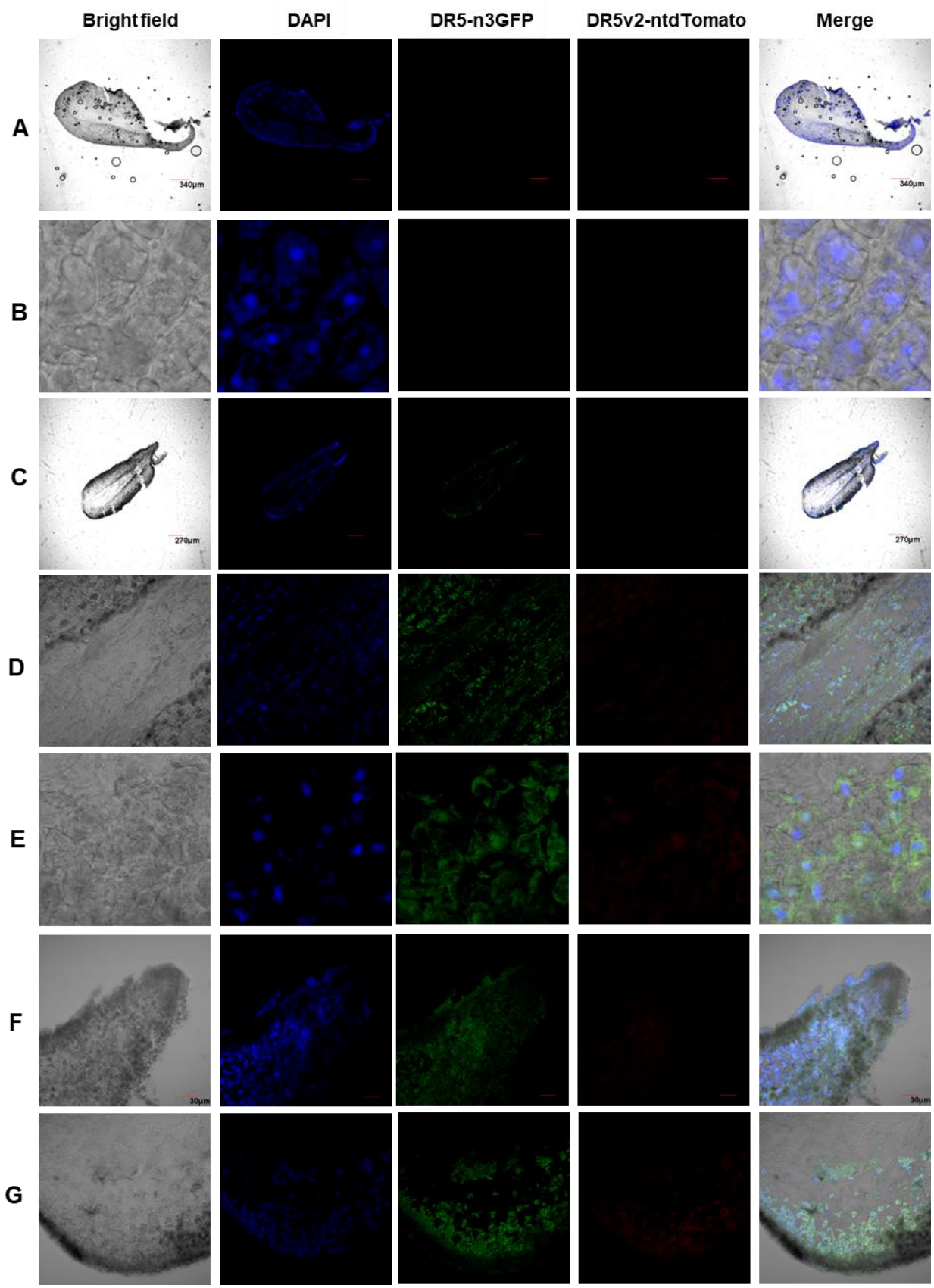

Bright field
DAPI
DR5-n3GFP
DR5v2-ntdTomato
Merge
A
B
C
D
E
F
G
340µm
340µm
270µm
270µm
30µm
30µm

Figure 4.16. Transverse section of somatic embryos. **A** and **B**) Torpedo embryo, control. **C-G**) Cotyledonary embryo, transformed.

4.4 DISCUSION

SE is a clear example of cellular differentiation and does not depend on the action of a single molecule but on a complex signaling network (Méndez-Hernández et al., 2019). Transcriptional regulation is essential in the induction of SE (Horstman et al., 2017). For example, the ARF genes show specific expressions during the induction of somatic embryogenesis in Arabidopsis (Wójcikowska and Gaj, 2017). It has also been reported that the YUCCA and AUX/IAA genes involved in auxin biosynthesis are transcriptionally regulated during embryogenesis, which suggests that auxin signaling is crucial in the embryogenic process (Uc-Chuc et al., 2020). In *C. canephora,* it has been reported that the addition of exogenous auxin induces the biosynthesis of indole-3-acetic acid, which correlates with the expression of biosynthesis genes such as YUCCA and TAA1 (Ayil-Gutiérrez et al., 2013). Several transcription factors have been reported in hormonal signaling, cell differentiation, and organogenesis (Salaün et al., 2021). Among the transcription factors involved during the induction of SE in different species are BABY BOOM (BBM) (Boutilier et al., 2002), ABAINSENSITIVE 3 (ABI3) (Shiota et al., 1998), WUSCHEL (WUS) (Xiao et al., 2018), AGAMOUS LIKE (AGL) (Thakare et al., 2008), LEAFY COTYLEDON LIKE (LIL) (Kwong et al., 2003), LEAFY COTYLEDON (LEC) (Gaj et al., 2005), VIVIPAROUS1 (VP1) (Footitt et al., 2003) and SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 (SERK1) (Pérez-Pascual et al., 2018).

The LEC family of transcription factors plays an essential role in regulating somatic embryogenesis. For example, in Arabidopsis, the loss of LEC function drastically affects embryonic development (Gaj et al., 2005). On the other hand, the ectopic expression of specific transcription factors, such as LEC, BBM, or WUS, can increase the efficiency of SE induction (Tian et al., 2020). LEC1 has been reported to have an essential role in zygotic embryogenesis and has been suggested to control various processes in seed development (Tvorogova and Lutova, 2018). On the other hand, evidence of the participation of another transcription factor, such as SERK, has emerged through gene expression analysis. In *A. thaliana*, SERK1 is expressed in cells that develop in somatic

embryos until the heart stage; after this stage, SERK1 expression is no longer detectable in the embryo (Schmidt et al., 1997).

It has been reported that SERK overexpression is observed in the globular stage along with other genes such as BBM and LEC, which promotes the transition from non-embryogenic tissues to embryogenic cells (Tvorogova et al., 2019). On the other hand, WUSCHEL regulates the development of the shoot apical meristem (SAM) (Laux et al., 1996). It has been reported that WUS is associated with shoot regeneration and somatic embryogenesis in different species. A characteristic of WUS is the ability to move from its central site of biosynthesis to the periphery cells, which activates the transcription of *CLAVATA3* (*CVL3*) (Yadav et al., 2011). In different species, genes related to WUS have been observed during SE. For example, in *G. hirsitum*, ectopic expression of AtWUS promotes the proliferation and differentiation of callus, in addition to positively regulating *LEC1*, *LEC2*, and *FUS3* (Zheng et al., 2014). Also, *WUS* overexpression improves the induction of embryogenesis and can improve regeneration in cotton (Bouchabke-Coussa et al., 2013).

4.5 CONCLUSIONS

It was possible to identify 23 differential proteins using 2DE; although their identity still needs to be verified, they could be specific potential candidates for developing SE. Of the 25,574 genes of *C. canephora*, several genes involved in auxin homeostasis belonging to the ABC, ARF, IAA, GH3, YUCCA, and ILR families were identified as differentially expressed, as well as genes related to cytokinin metabolism, such as CKS, LOG, PUP, APR, APTR. The possible participation of other genes involved in SE was also confirmed, such as SERK1, LEC1, CLAVATA, and WUSCHEL, among others.

There is a correlation between certain families of proteins and genes identified as potential markers of the SE induction process in *C. canephora*, identified through proteomic and transcriptomic tools.

Also, there is a marked dynamic in the accumulation of both IAA and its conjugates and the different cytokinins, showing a similar pattern where the highest concentration is located during pretreatment. Then, these levels gradually decrease towards induction.

It was possible to establish a *C. canephora* transformation protocol with the synthetic auxin reporter DR5v2. So far, the visualization of the IAA with the synthetic reporter DR5v2 coincides with what was previously reported by Márquez-López et al. (2018). The above could suggest that the sites where auxin accumulates are the same as where its action is carried out.

CHAPTER V

5. GENERAL DISCUSION

Coffea arabica and *C. canephora* are two of the most important crops worldwide. Various strategies have been implemented to increase production and meet global demand. To this end, somatic embryogenesis represents a viable alternative that allows the mass production of homogeneous plants, free of diseases and with desirable agronomic characteristics, in less time and space. In our laboratory, efficient protocols have been generated for both species (Méndez-Hernández et al., 2023), which have been used not only for scaling and covering the demand of producers but also to generate basic science that provides information that allows us to understand the mechanisms that operate in the SE, being highly reproducible and controllable systems.

We recently approached the study of SE from the transcriptomic point of view; however, it was possible to complement the information with powerful proteomic tools on this occasion. Although there will not always be a correlation between the results obtained by both methods, they will always offer significant advantages for exploring various research questions.

Auxins and cytokinins are two of the most significant plant growth regulators, and they play a crucial role in all areas of plant growth and development (Santner, 2009 33231 /id). These two PGRs are involved in cell division, elongation, differentiation, vascular and flower development, nutrient homeostasis, leaf expansion, and stress responses, among others (Márquez-López et al., 2019; Ljung, 2013).

Various transcription factors and other genes are commonly called "master regulators" of the SE process. Among them, we can mention LEC1, the nuclear transcription factor Y, WUSCHEL, TAR, and SERK1, among others, involved in the induction of SE and in the maturation of embryos. It is known that these regulate common metabolic pathways and that there is even a connection between them (Horstman et al., 2017).

The content of PGR increases during pretreatment, while the expression of different genes involved in auxin and cytokinin homeostasis increases, such as YUCCA (Ayil-Gutiérrez et al., 2013), GH3 (Méndez-Hernández et al., 2019), PIN (Márquez-López et al., 2018), ARF

and Aux/IAA (Quintana-Escobar et al., 2019), CKK and PUP (Avilez-Montalvo et al., 2022).

Our results correspond with previous reports where only a molecular approach was made. During SE, it is clear that the exogenous addition of PGRs is crucial for developing the first embryogenic structures.

CHAPTER VI

6. CONCLUSIONS AND PERSPECTIVES

6.1 Conclusions

- ✓ Identifying genes and proteins associated with cellular differentiation and somatic embryogenesis of *C. arabica* and *C. canephora* was possible.
- ✓ In both systems, common proteins/genes were observed, although with slight changes in their accumulation or expression.
- ✓ Auxins and cytokinins are detected in higher concentrations during the pretreatment stage. Towards the induction stage, these concentrations gradually decrease.
- ✓ Combining the data obtained from the transcriptomic and proteomic analysis, we can conclude that there is a regulation of auxin and cytokinin homeostasis, which triggers a series of reactions that will subsequently give rise to somatic embryogenesis. The ABC and BIG proteins could be playing an important role in the transport and mobilization of auxin, while the PUP in the transport of cytokinins. GH3 and ILR/ILL could regulate auxin levels.
- ✓ The ARR and LOG proteins show that CK signaling and activation are active.
- ✓ The master regulators SERK1, LEC1, and WUSCHEL are important for SE.
- ✓ The location of the auxin response site was detected in the chloroplasts, then in the cytoplasm and nuclei.
- ✓ Proteomics, in conjunction with other omics and molecular tools, can provide novel information for understanding the functioning of the SE process.

6.2 Perspectives

- ✓ Now that 2DE has detected the differential proteins, identifying their identity by picking the spots on the gels and identifying them by LC/MS-MS would be of particular interest. In this way, a comparison and complementation could be carried out with the results obtained from the gel-free shotgun carried out in this work.

- ✓ Another important approach to enrich the information obtained so far would be the analysis of the proteins secreted into the culture medium during the different points of the induction process and in each of the various stages of development of the somatic embryos.
- ✓ The generation of a transcriptome of the SE induction process in *C. canephora* represents a valuable source of information, available to be consulted at any time to search for any gene of interest, whether related to SE or other processes of the SE. Plant growth and development are even approached from an epigenetic vision. In this sense, a poorly studied aspect worth scrutinizing further is homeostasis and cytokinin metabolism dynamics.
- ✓ The transgenic line generated with the synthetic reporter DR5v2 could complement future works in this system.
- ✓ On the contrary, we can mention a reporter similar to DR5v2 for cytokinins: TCS (Two-component system). This synthetic sensor allows for visualization of the transcriptional output of the cytokinin signaling network. Therefore, implementing its use in our system would allow us to expand the information we have until now on cytokinin homeostasis

BIBLIOGRAPHY

Adamowski, M. y J. Friml (2015). PIN-dependent auxin transport: Action, regulation, and evolution. The Plant Cell, 27, 20-32.

Ades, L. M., D. M. Morré y D. J. Morré (2014). Cloning and characterization of a candidate auxin plant growth regulator-activated cell surface hydroquinone (NADH) oxidase. Advances in Biological Chemistry, 4, 415.

Agrawal, G. K., N.-S. Jwa, M.-H. Lebrun, D. Job y R. Rakwal (2010). Plant secretome: unlocking secrets of the secreted proteins. Proteomics, 10, 799-827.

Aguilar-Hernández, V. y V. M. Loyola-Vargas (2018). Advanced proteomic approaches to elucidate somatic embryogenesis. Frontiers in Plant Science, 9, 1658.

Allen, M., W. Qin, F. Moreau y B. Moffatt (2002). Adenine phosphoribosyltransferase isoforms of Arabidopsis and their potential contributions to adenine and cytokinin metabolism. Physiologia Plantarum, 115, 56-68.

Ashihara, H., C. Stasolla, N. Loukanina y T. A. Thorpe (2001). Purine metabolism during white spruce somatic embryo development: salvage of adenine, adenosine, and inosine. Plant Science, 160, 647-657.

Avilez-Montalvo, J., A. O. Quintana-Escobar, H. A. Méndez-Hernández, M. Á. Uc-Chuc, L. Brito-Argáez, V. Aguilar-Hernández, R. M. Galaz-Ávalos y V. M. Loyola-Vargas (2022). Auxin-cytokinin cross talk in somatic embryogenesis of *Coffea canephora*. Plants, 11, 2013.

Ayil-Gutiérrez, B. A., R. M. Galaz-Ávalos, E. Peña-Cabrera y V. M. Loyola-Vargas (2013). Dynamics of the concentration of IAA and some of its conjugates during the induction of somatic embryogenesis in *Coffea canephora*. Plant Signaling & Behavior, 8, e26998.

Baggerman, G., E. Vierstraete, A. De Loof y L. Schoofs (2005). Gel-based versus gel-free proteomics: A review. Combinatorial Chemnistry & High Throughput Screening, 8, 669-677.

Barbez, E. y J. Kleine-Vehn (2013). Divide *Et Impera*-cellular auxin compartmentalization. Current Opinion in Plant Biology, 16, 78-84.

Bargmann, B. O. R., S. Vanneste, G. Krouk, T. Nawy, I. Efroni, E. Shani, G. Choe, J. Friml, D. C. Bergmann, M. Estelle y K. D. Birnbaum (2013). A map of cell type-specific auxin responses. Molecular Systems Biology, 9, 688.

Bautista-Valle, M. V., C. Camacho–Vazquez, J. M. Elizalde–Contreras, J. L. Monribot–Villanueva, A. M. V. Limón, E. Bojórquez–Velázquez, A. Zamora–Briseño, J. V. Jorrín-Novo y E. Ruiz–May (2023). Comparing and integrating TMT-SPS-MS3 and label-free quantitative approaches for proteomics scrutiny in recalcitrant mango (*Mangifera indica* L.) peel tissue during postharvest period. Proteomics, 2300239.

Bedia, C. (2018). Experimental approaches in omic sciences. Comprehensive Analytical Chemistry, 82, 13-36.

Bindschedler, L. V. y R. Cramer (2011). Quantitative plant proteomics. Proteomics, 11, 756-775.

Blakeslee, J. J., W. A. Peer y A. S. Murphy (2005). Auxin transport. Current Opinion in Plant Biology, 8, 494-500.

Borisjuk, N., L. Sitailo, K. Adler, L. Malysheva, A. Tewes, L. Borisjuk y R. Manteuffel (1998). Calreticulin expression in plant cells: developmental regulation, tissue specificity and intracellular distribution. Planta, 206, 504-514.

Bouchabke-Coussa, O., M. Obellianne, D. Linderme, E. Montes, A. Maia-Grondard, F. Vilaine y C. Pannetier (2013). Wuschel overexpression promotes somatic embryogenesis and induces organogenesis in cotton (*Gossypium hirsutum* L.) tissues cultured *in vitro*. Plant Cell Reports, 32, 675-686.

Boutilier, K., R. Offringa, V. K. Sharma, H. Kieft, T. Ouellet, L. Zhang, J. Hattori, C. Liu, A. A. M. Van Lammeren, B. L. A. Miki, J. B. M. Custers y M. M. Van Lookeren-Campagne (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. The Plant Cell, 14, 1737-1749.

Campanella, J. J., S. M. Smith, D. Leibu, S. Wexler y J. Ludwig-Müller (2008). The auxin conjugate hydrolase family of *Medicago truncatula* and their expression during the interaction with two symbionts. Journal of Plant Growth Regulation, 27, 26-38.

Campos, N. A., L. V. Paiva, B. Panis y S. C. Carpentier (2016). The proteome profile of embryogenic cell suspensions of *Coffea arabica* L. Proteomics, 16, 1001-1005.

Campos, N. A., B. Panis y S. C. Carpentier (2017). Somatic embryogenesis in coffee: The evolution of biotechnology and the integration of omics technologies offer great opportunities. Frontiers in Plant Science, 8, 1460.

Casanova-Sáez, R. y U. Voß (2019). Auxin metabolism controls developmental decisions in land plants. Trends in Plant Science, 24, 741-754.

Casanova-Sáez, R., E. Mateo-Bonmatí y K. Ljung (2021). Auxin metabolism in plants. Cold Spring Harbor Perspectives in Biology, 13, a039867.

Cazzonelli, C. I., M. Vanstraelen, S. Simon, K. Yin, A. Carron-Arthur, N. Nisar, G. Tarle, A. J. Cuttriss, I. R. Searle y E. Benková (2013). Role of the Arabidopsis PIN6 auxin transporter in auxin homeostasis and auxin-mediated development. PLoS ONE, 8, e70069.

Chen, C. M., D. K. Melitz y F. W. Clough (1982). Metabolism of cytokinin: phosphoribosylation of cytokinin bases by adenine phosphoribosyltransferase from wheat germ. Archives of Biochemistry and Biophysics, 214, 634-641.

Chen, Q., C. S. Westfall, L. M. Hicks, S. Wang y J. M. Jez (2010). Kinetic basis for the conjugation of auxin by a GH3 family indole-acetic acid-amido synthetase. Journal of Biological Chemistry, 285, 29780-29786.

Chen, Y., X. Xu, Z. Liu, Z. Zhang, X. XuHan, Y. Lin y Z. Lai (2020). Global scale transcriptome analysis reveals differentially expressed genes involve in early somatic embryogenesis in *Dimocarpus longan* Lour. BMC Genomics, 21, 4.

Cottrell, J. S. (2011). Protein identification using MS/MS data. Journal of Proteomics, 74, 1842-1851.

DaMatta, F. M. y J. D. C. Ramalho (2006). Impacts of drought and temperature stress on coffee physiology and production: a review. Brazilian Journal of Plant Physiology, 18, 55-81.

Davies, R. T., D. H. Goetz, J. Lasswell, M. N. Anderson y B. Bartel (1999). IAR3 encodes an auxin conjugate hydrolase from Arabidopsis. The Plant Cell, 11, 365-376.

Denoeud, F., L. Carretero-Paulet, A. Dereeper, G. Droc, R. Guyot, M. Pietrella, C. Zheng, A. Alberti, F. Anthony, G. Aprea, J. M. Aury, P. Bento, M. Bernard, S. Bocs, C. Campa, A. Cenci, M. C. Combes, D. Crouzillat, C. Da Silva, L. Daddiego, F. De Bellis, S. Dussert, O. Garsmeur, T. Gayraud, V. Guignon, K. Jahn, V. Jamilloux, T. Joët, K. Labadie, T. Lan, J. Leclercq, M. Lepelley, T. Leroy, L. T. Li, P. Librado, L. Lopez, A. Muñoz, B. Noel, A. Pallavicini, G.

Perrotta, V. Poncet, D. Pot, Priyono, M. Rigoreau, M. Rouard, J. Rozas, C. Tranchant-Dubreuil, R. VanBuren, Q. Zhang, A. C. Andrade, X. Argout, B. Bertrand, A. De Kochko, G. Graziosi, R. J. Henry, Jayarama, R. Ming, C. Nagai, S. Rounsley, D. Sankoff, G. Giuliano, V. A. Albert, P. Wincker y P. Lashermes (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Science, 345, 1181-1184.

Devasia, J., B. Muniswamy y M. K. Mishra (2020). Investigation of ZnO nanoparticles on *in vitro* cultures of Coffee (*Coffea Arabica* L.). International Journal of Nanoscience and Nanotechnology, 16, 271-277.

Dobrev, P. I., K. Hoyerová y J. Petrášek, (2017). Analytical determination of auxins and cytokinins, en: Auxins and Cytokinins in Plant Biology: Methods and Protocols, Dandekar, T. y M. Naseem, (eds). Springer. New York. pp. 31-39.

Dong, Q., R. O. Magwanga, X. Cai, P. Lu, J. Nyangasi Kirungu, Z. Zhou, X. Wang, X. Wang, Y. Xu y Y. Hou (2019). RNA-sequencing, physiological and RNAi analyses provide insights into the response mechanism of the ABC-mediated resistance to Verticillium dahliae infection in cotton. Genes, 10, 110.

Dubas, E., J. Moravciková, J. Libantová, I. Matusiková, E. Benková, I. Zur y M. Krzewska (2014). The influence of heat stress on auxin distribution in transgenic *B. napus* microspores and microspore-derived embryos. Protoplasma, 251, 1077-1087.

Dublin, P. (1981). Embryogenèse somatique directe sur fragments de feuilles de caféier Arabusta. Café Cacao Thé, 25, 237-242.

Ducos, J. P., C. Lambot y V. Pétiard (2007). Bioreactors for coffee mass propagation by somatic embryogenesis. International Journal of Plant Developmental Biology, 1, 1-12.

Ebrahim, N., R. Shibli, I. Makhadmeh, M. Shatnawi y A. Abu-Ein (2007). *In vitro* propagation and *in vivo* acclimatization of three coffee cultivars (*Coffea arabica* L.) from Yemen. World Appl. Sci. J, 2, 142-150.

Enríquez-Valencia, A. J., F. A. Vázquez-Flota, J. R. Ku-Cauich y R. M. Escobedo-GraciaMedrano (2019). Differentially expressed genes during the transition from early to late development phases in somatic embryo of banana (Musa spp. AAB group, Silk subgroup) cv. Manzano. Plant Cell, Tissue and Organ Culture, 136, 289-302.

Etienne, H., (2005). Somatic embryogenesis protocol: Coffee (*Coffea arabica* L. and *C. canephora* P.), en: Protocol for Somatic Embryogenesis in Woody Plants, Jain, S. y P. Gupta, (eds). Springer. Netherlands. pp. 167-179.

Etienne, H., D. Breton, J. C. Breitler, B. Bertrand, E. Déchamp, R. Awada, P. Marraccini, S. Léran, E. Alpizar, C. Campa, P. Courtel, F. Georget y J.-P. Ducos (2018). Coffee somatic embryogenesis: How did research, experience gained and innovations promote the commercial propagation of elite clones from the two cultivated species. Frontiers in Plant Science, 9, 1630.

Faus, I., R. Niñoles, V. Kesari y J. Gadea (2021). The ABCF3 gene of Arabidopsis is functionally linked with GCN1 but not with GCN2 during stress and development. Plant Molecular Biology Reporter, 1-10.

Fehér, A. (2019). Callus, dedifferentiation, totipotency, somatic embryogenesis: What these terms mean in the era of molecular plant biology? Frontiers in Plant Science, 10, 536.

Fendrych, M., J. Leung y J. Friml (2016). TIR1/AFB-Aux/IAA auxin perception mediates rapid cell wall acidification and growth of Arabidopsis hypocotyls. eLife, 5, e19048.

Ferreira, T., J. Shuler, R. Guimardes y A. Farah, (2019). Introduction to coffee plant and genetics, en: Coffee: Production, Quality and Chemistry, Farah, A., (ed). The Royal Society of Chemistry. Croydon, UK. pp. 3-25.

Feussner, I. y A. Polle (2015). What the transcriptome does not tell - proteomics and metabolomics are closer to the plants' patho-phenotype. Current Opinion in Plant Biology, 26, 26-31.

Footitt, S., M. Ingouff, D. Clapham y S. von Arnold (2003). Expression of the viviparous 1 (Pavp1) and p34cdc2 protein kinase (cdc2Pa) genes during somatic embryogenesis in Norway spruce (*Picea abies* [L.] Karst). Journal of Experimental Botany, 54, 1711-1719.

Fu, X., Z. Shi, Y. Jiang, L. Jiang, M. Qi, T. Xu y T. Li (2019). A family of auxin conjugate hydrolases from *Solanum lycopersicum* and analysis of their roles in flower pedicel abscission. BMC Plant Biology, 19, 233.

Gachon, C. M., M. Langlois-Meurinne, Y. Henry y P. Saindrenan (2005). Transcriptional co-regulation of secondary metabolism enzymes in Arabidopsis: functional and evolutionary implications. Plant Molecular Biology, 58, 229-245.

Gaj, M., S. Zhang, J. Harada y P. Lemaux (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of *Arabidopsis*. Planta, 222, 977-988.

Ge, X., C. Zhang, Q. Wang, Z. Yang, Y. Wang, X. Zhang, Z. Wu, Y. Hou, J. Wu y F. Li (2014). iTRAQ protein profile differential analysis between somatic globular and cotyledonary embryos reveals stress, hormone, and respiration involved in increasing plantlet regeneration of *Gossypium hirsutum* L. Journal of Proteome Research, 14, 268-278.

Geisler, M. y A. S. Murphy (2006). The ABC of auxin transport: The role of p-glycoproteins in plant development. FEBS Letters, 580, 1094-1102.

Geisler, M., B. Aryal, M. di Donato y P. Hao (2017). A critical view on ABC transporters and their interacting partners in auxin transport. Plant and Cell Physiology, 58, 1601-1614.

Góngora-Castillo, E., G. I. Nic-Can, R. M. Galaz-Ávalos y V. M. Loyola-Vargas, (2018). Elaboration of transcriptome during the induction of somatic embryogenesis, en: Plant Cell Culture Protocols, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer New York. New York, NY. pp. 411-427.

Goodwin, S., J. D. McPherson y W. R. McCombie (2016). Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17, 333-351.

Götz, S., J. M. García-Gómez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talón, J. Dopazo y A. Conesa (2008). High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research, 36, 3420-3435.

Gulzar, B., A. Mujib, M. V. Rajam, A. Frukh y N. Zafar (2019). Identification of somatic embryogenesis (SE) related proteins through label-free shotgun proteomic method and cellular role in *Catharanthus roseus* (L.) G. Don. Plant Cell, Tissue and Organ Culture, 137, 225-237.

Guo, H., H. Guo, L. Zhang, Y. Fan, J. Wu, Z. Tang, Y. Zhang, Y. Fan y F. Zeng (2020). Dynamic transcriptome analysis reveals uncharacterized complex regulatory pathway underlying genotype-recalcitrant somatic embryogenesis transdifferentiation in cotton. Genes, 11, 519.

Hayashi, K. i., S. Nakamura, S. Fukunaga, T. Nishimura, M. K. Jenness, S. MurphyAngus, H. Motose, H. Nozaki, M. Furutani y T. Aoyama (2014). Auxin transport sites are visualized in planta using fluorescent auxin analogs. Proceedings of the National Academy of Sciences of the United States of America, 111, 11557-11562.

Heberle, H., G. V. Meirelles, F. R. da Silva, G. P. Telles y R. Minghim (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics, 16, 169.

Heringer, A. S., C. Santa-Catarina y V. Silveira (2018). Insights from proteomic studies into plant somatic embryogenesis. Proteomics, 18, 1700265.

Herman, E. B. y G. J. Haas (1975). Clonal propagation of *Coffea arabica* L. from callus culture. HortScience, 10, 588-589.

Horstman, A., M. Bemer y K. Boutilier (2017). A transcriptional view on somatic embryogenesis. Regeneration, 4, 201-216.

ICO (2022). Trade Statistics. Available online at: http://www.ico.org/trade_statistics.asp (Accessed January 2022). ICO,

Iquebal, M. A., S. Jaiswal, C. S. Mukhopadhyay, C. Sarkar, A. Rai y D. Kumar, (2015). Applications of bioinformatics in plant and agriculture, en: PlantOmics: The Omics of Plant Science, Barh, D., M. S. Khan y E. Davies, (eds). Springer. India. pp. 755-789.

Jamet, E. y V. Santoni (2018). Editorial for Special Issue: 2017 Plant Proteomics. Proteomes, 6, 28.

Jenness, M. K., N. Carraro, C. A. Pritchard y A. S. Murphy (2019). The Arabidopsis ATP-BINDING CASSETTE transporter ABCB21 regulates auxin levels in cotyledons, the root pericycle, and leaves. Frontiers in Plant Science, 10, 806.

Jia, X.-Y., L.–H. He, R.–L. Jing y R.–Z. Li (2009). Calreticulin: conserved protein and diverse functions in plants. Physiologia Plantarum, 136, 127-138.

Jin, S. H., X. M. Ma, P. Han, B. Wang, Y. G. Sun, G. Z. Zhang, Y. j. Li y B. K. Hou (2013). UGT74D1 is a novel auxin gluycosyltransferase from *Arabidopsis thaliana*. PLoS ONE, 8, e61705.

Jorrín-Novo, J. V., S. Komatsu, R. Sanchez-Lucas y L. E. Rodríguez de Francisco (2019). Gel electrophoresis-based plant proteomics: Past, present, and future. Happy 10th anniversary Journal of Proteomics! Journal of Proteomics, 198, 1-10.

Jorrín-Novo, J. V. (2021). Proteomics and plant biology: Contributions to date and a look towards the next decade. Expert Review of Proteomics, 18, 93-103.

Kaneda, M., M. Schuetz, B. S. P. Lin, C. Chanis, B. Hamberger, T. L. Western, J. Ehlting y A. L. Samuels (2011). ABC transporters coordinately expressed

during lignification of Arabidopsis stems include a set of ABCBs associated with auxin transport. Journal of Experimental Botany, 62, 2063-2077.

Kang, J., J. Park, H. Choi, B. Burla, T. Kretzschmar, Y. Lee y E. Martinoia (2011). Plant ABC transporters. The Arabidopsis Book, 9, e0153.

Kanyuka, K., U. Praekelt, K. A. Franklin, O. E. Billingham, R. Hooley, G. C. Whitelam y K. J. Halliday (2003). Mutations in the huge Arabidopsis gene BIG affect a range of hormone and light responses. The Plant Journal, 35, 57-70.

Kieber, J. J. y G. E. Schaller (2018). Cytokinin signaling in plant development. Development, 145, dev149344.

Kim, A., J. Chen, D. Khare, J. Y. Jin, Y. Yamaoka, M. Maeshima, Y. Zhao, E. Martinoia, J. U. Hwang y Y. Lee (2020). Non-intrinsic ATP-binding cassette proteins ABCI19, ABCI20 and ABCI21 modulate cytokinin response at the endoplasmic reticulum in *Arabidopsis thaliana*. Plant Cell Reports, 39, 473-487.

Kumaravel, M., S. Uma, S. Backiyarani y M. S. Saraswathi (2020). Proteomic analysis of somatic embryo development in *Musa* spp. cv. Grand Naine (AAA). Scientific Reports, 10, 4501.

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

Landey, R. B., A. Cenci, F. Georget, B. Bertrand, G. Camayo, E. Dechamp, J. C. Herrera, S. Santoni, P. Lashermes, J. Simpson y H. Etienne (2013). High genetic and epigenetic stability in *Coffea arabica* plants derived from embryogenic suspensions and secondary embryogenesis as revealed by AFLP, MSAP and the phenotypic variation rate. PLoS ONE, 8, e56372.

Lane, T. S., C. S. Rempe, J. Davitt, M. E. Staton, Y. Peng, D. E. Soltis, M. Melkonian, M. Deyholos, J. H. Leebens-Mack y M. Chase (2016). Diversity of ABC transporter genes across the plant kingdom and their potential utility in biotechnology. BMC Biotechnology, 16, 47.

Laux, T., K. F. Mayer, J. Berger y G. Jurgens (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. Development, 122, 87-96.

LeClere, S., R. Tellez, R. A. Rampey, S. P. T. Matsuda y B. Bartel (2002). Characterization of a family of IAA-amino acid conjugate hydrolases from Arabidopsis. Journal of Biological Chemistry, 277, 20446-20452.

Lee, M., Y. Choi, B. Burla, Y. Y. Kim, B. Jeon, M. Maeshima, J. Y. Yoo, E. Martinoia y Y. Lee (2008). The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO_2. Nature Cell Biology, 10, 1217-1223.

Li, S., D. Li, P. Zhang, R. Wang, L. Sun, J. Wan y J. Xu (2018). ABCF3 regulates the expression of aquaporin genes and endoplasmic reticulum stress-related genes in Arabidopsis. Theoretical and Experimental Plant Physiology, 30, 215-222.

Li, S.-B., Z. Z. Xie, C. G. Hu y J. Z. Zhang (2016). A review of auxin response factors (ARF) in plants. Frontiers in Plant Science, 7, 47.

Liao, C. Y., W. Smet, G. Brunoud, S. Yoshida, T. Vernoux y D. Weijers (2015). Reporters for sensitive and quantitative measurement of auxin response. Nature Methods, 12, 207-210.

Libik, M. y L. Przywara (2000). Immunolocalization of calreticulin in protoplasts and somatic embryos of *Daucus carota* L. grown in suspension culture. Acta Biologica Cracoviensia Series Botanica, 42, 87-92.

Liu, Y., S. Lu, K. Liu, S. Wang, L. Huang y L. Guo (2019). Proteomics: a powerful tool to study plant responses to biotic stress. Plant Methods, 15, 135.

Ljung, K., A. K. Hull, M. Kowalczyk, A. Marchant, J. Celenza, J. D. Cohen y G. Sandberg (2002). Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in *Arabidopsis thaliana*. Plant Molecular Biology, 50, 309-332.

Ljung, K. (2013). Auxin metabolism and homeostasis during plant development. Development, 140, 943-950.

López-Bucio, J., E. Hernández-Abreu, L. Sánchez-Calderón, A. Pérez-Torres, R. A. Rampey, B. Bartel y L. Herrera-Estrella (2005). An auxin transport independent pathway is involved in phosphate stress-induced root architectural alterations in Arabidopsis. Identification of BIG as a mediator of auxin in pericycle cell Activation. Plant Physiology, 137, 681-691.

Loyola-Vargas, V. M. y N. Ochoa-Alejo, (2016). Somatic embryogenesis. An overview, en: Somatic Embryogenesis. Fundamental Aspects and Applications, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. Switzerland. pp. 1-10.

Loyola-Vargas, V. M., J. R. Avilez-Montalvo, R. N. Avilez-Montalvo, R. E. Márquez-López, R. M. Galaz-Ávalos y E. Mellado-Mojica, (2016). Somatic embryogenesis in *Coffea* spp., en: Somatic Embryogenesis. Fundamental

Aspects and Applications, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. Handel, Switzerland. pp. 241-266.

Ludwig-Müller, J. (2011). Auxin conjugates: their role for plant development and in the evolution of land plants. Journal of Experimental Botany, 62, 1757-1773.

Márquez-López, R. E., C. A. Pérez-Hernández, Á. Kú-González, R. M. Galaz-Ávalos y V. M. Loyola-Vargas (2018). Localization and transport of indole-3-acetic acid during somatic embryogenesis in *Coffea canephora*. Protoplasma, 255, 695-708.

Márquez-López, R. E., A. O. Quintana-Escobar y V. M. Loyola-Vargas (2019). Cytokinins, the Cinderella of plant growth regulators. Phytochemistry Reviews, 18, 1387-1408.

Marraccini, P., F. Vinecky, G. S. C. Alves, H. J. O. Ramos, S. Elbelt, N. G. Vieira, F. A. Carneiro, P. S. Sujii, J. C. Alekcevetch, V. A. Silva, F. M. DaMatta, M. A. G. Ferrao, T. Leroy, D. Pot, L. G. E. Vieira, F. R. da Silva y A. C. Andrade (2012). Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. Journal of Experimental Botany, 63, 4191-4212.

Martins, M. Q., A. S. Fortunato, W. P. Rodrigues, F. L. Partelli, E. Campostrini, F. C. Lidon, F. M. DaMatta, J. C. Ramalho y A. I. Ribeiro-Barros (2017). Selection and validation of reference genes for accurate RT-qPCR data normalization in Coffea spp. under a climate changes context of interacting elevated [CO_2] and temperature. Frontiers in Plant Science, 8, 307.

Mateo-Bonmatí, E., R. Casanova-Sáez, J. Simura y K. Ljung (2021). Broadening the roles of UDP–glycosyltransferases in auxin homeostasis and plant development. New Phytologist, 232, 642-654.

McAlister, G. C., D. P. Nusinow, M. P. Jedrychowski, M. Wu¦êhr, E. L. Huttlin, B. K. Erickson, R. Rad, W. Haas y S. P. Gygi (2014). MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Analytical Chemistry, 86, 7150-7158.

Méndez-Hernández, H. A., M. Ledezma-Rodríguez, R. N. Avilez-Montalvo, Y. L. Juárez-Gómez, A. Skeete, J. Avilez-Montalvo, C. De-la-Peña y V. M. Loyola-Vargas (2019). Signaling overview of plant somatic embryogenesis. Frontiers in Plant Science, 10, 77.

Méndez-Hernández, H. A., A. O. Quintana-Escobar, M. Á. Uc-Chuc y V. M. Loyola-Vargas (2021). Genome-wide analysis, modeling, and identification of amino acid binding motifs suggest the involvement of *GH3* genes during somatic embryogenesis of *Coffea canephora*. Plants, 10, 2034.

Méndez-Hernández, H. A., R. M. Galaz-Ávalos, A. O. Quintana-Escobar, R. Pech-Hoil, A. M. Collí-Rodríguez, I. Q. Salas-Peraza y V. M. Loyola-Vargas (2023). *In vitro* conversion of *Coffea* spp. somatic embryos in SETIS™ bioreactor system. Plants, 12, 3055.

Mõttus, J., S. Maiste, P. Eek, E. Truve y C. Sarmiento (2021). Mutational analysis of Arabidopsis thaliana ABCE2 identifies important motifs for its RNA silencing suppressor function. Plant Biology, 23, 21.

Mukul-López, H. G., C. De-la-Peña, R. M. Galaz-Ávalos y V. M. Loyola-Vargas (2012). Evaluation of the extracellular proteome profile during the somatic embryogenesis process of *Coffea spp*. Journal of the Mexican Chemical Society, 56, 72-79.

Mwaniki, W. I., A. H. Lubabali, K. K. Asava, C. O. Agwanda y S. E. Anami (2019). Effects of genotype and plant growth regulators on callus induction in leaf cultures of *Coffea arabica* L. F1 hybrid. African Journal of Biotechnology, 18, 1004-1015.

Ni, D. A., L. J. Wang, C. H. Ding y Z. H. Xu (2001). Auxin distribution and transport during embryogenesis and seed germination of Arabidopsis. Cell Research, 11, 273-278.

Nic-Can, G. I. y C. De-la-Peña, (2014). Epigenetic advances on somatic embryogenesis of agronomical and important crops, en: Epigenetics in Plants of Agronomic Importance: Fundamentals and Applications, Álvarez-Venegas, R., C. De-la-Peña y J. A. Casas-Mollano, (eds). Springer. Cham Heidelberg New York Dordrecht London. pp. 91-109.

Nic-Can, G. I., R. M. Galaz-Ávalos, C. De-la-Peña, A. Alcazar Magana, K. Wrobel y V. M. Loyola-Vargas (2015). Somatic embryogenesis: Identified factors that lead to embryogenic repression. A case of species of the same genus. PLoS ONE, 10, e0126414.

Nic-Can, G. I. y V. M. Loyola-Vargas, (2016). The role of the auxins during somatic embryogenesis, en: Somatic Embryogenesis. Fundamental Aspects and Applications, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. Handel, Switzerland. pp. 171-181.

Nic-Can, G. I., J. R. Avilez-Montalvo, R. N. Avilez-Montalvo, R. E. Márquez-López, E. Mellado-Mojica, R. M. Galaz-Ávalos y V. M. Loyola-Vargas, (2016). The relationship between stress and somatic embryogenesis, en: Somatic Embryogenesis. Fundamental Aspects and Applications, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. Switzerland. pp. 151-170.

Ohmiya, A. (2002). Characterization of ABP19/20, sequence homologues of germin-like protein in Prunus persica L. Plant Science, 163, 683-689.

Pencik, A., R. Casanova-Sáez, V. Pilarová, A. Zukauskaité, R. Pinto, J. Luis Micol, K. Ljung y O. Novák (2018). Ultra-rapid auxin metabolite profiling for high-throughput Arabidopsis mutant screening. Journal of Experimental Botany, 69, 2569-2579.

Pereira-Dias, F., N. Steiner, G. C. Cangahaula-Inocente, A. P. Lando, M. Santos y M. P. Guerra (2020). Integrated proteomics and histochemical analysis of *Araucaria angustifolia* (Bertol.) Kuntze (Araucariaceae) in embryogenic suspension culture. Annals of Forest Research, 63, 27-43.

Péret, B., K. Swarup, A. Ferguson, M. Seth, Y. Yang, S. Dhondt, N. James, I. Casimiro, P. Perry, A. Syed, H. Yang, J. Reemmer, E. Venison, C. Howells, M. A. Perez-Amador, J. Yun, J. Alonso, G. T. S. Beemster, L. Laplaze, A. Murphy, M. J. Bennett, E. Nielsen y R. Swarup (2012). AUX/LAX genes encode a family of auxin influx transporters that perform distinct functions during Arabidopsis development. The Plant Cell, 24, 2874-2885.

Pérez-Pascual, D., D. Jiménez-Guillen, H. Villanueva-Alonzo, R. Souza-Perera, G. Godoy-Hernández y J. J. Zúñiga-Aguilar (2018). Ectopic expression of the *Coffea canephora* SERK1 homologue induced differential transcription of genes involved in auxin metabolism and in the developmental control of embryogenesis. Physiologia Plantarum, 163, 530-551.

Perez-Riverol, Y., J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D. J. Kundu, A. Prakash, A. Frericks-Zipper y M. Eisenacher (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Research, 50, D543-D552.

Peterson, G. L. (1977). A simplification of the protein assay method of Lowry et al. which is more generally applicable. Analytical Biochemistry, 83, 346-356.

Quintana-Escobar, A. O., G. I. Nic-Can, R. M. Galaz-Ávalos, V. M. Loyola-Vargas y E. Góngora-Castillo (2019). Transcriptome analysis of the induction of somatic embryogenesis in *Coffea canephora* and the participation of arf and AUX/IAA genes. PeerJ, 7, e7752.

Quintana-Escobar, A. O., R. M. Galaz-Ávalos, J. M. Elizalde-Contreras, F. A. Reyes-Soria, V. Aguilar-Hernández, E. Ruíz-May y V. M. Loyola-Vargas (2021). Differences in the abundance of auxin homeostasis proteins suggest their central roles for *in vitro* tissue differentiation in *Coffea arabica*. Plants, 10, 2607.

Quiroz-Figueroa, F. R., C. F. J. Fuentes-Cerda, R. Rojas-Herrera y V. M. Loyola-Vargas (2002a). Histological studies on the developmental stages and differentiation of two different somatic embryogenesis systems of *Coffea arabica*. Plant Cell Reports, 20, 1141-1149.

Quiroz-Figueroa, F. R., M. Méndez-Zeel, F. Sánchez-Teyer, R. Rojas-Herrera y V. M. Loyola-Vargas (2002b). Differential gene expression in embryogenic and non-embryogenic clusters from cell suspension cultures of *Coffea arabica* L. Journal of Plant Physiology, 159, 1267-1270.

Quiroz-Figueroa, F. R., M. Monforte-González, R. M. Galaz-Ávalos y V. M. Loyola-Vargas, (2006). Direct somatic embryogenesis in *Coffea canephora*, en: Plant Cell Culture Protocols, Loyola-Vargas, V. M. y F. A. Vázquez-Flota, (eds). Humana Press. Totowa, NJ, USA. pp. 111-117.

Rampey, R. A., S. LeClere, M. Kowalczyk, K. Ljung, G. Sandberg y B. Bartel (2004). A family of auxin-conjugate hydrolases that contributes to free indole-3-acetic acid levels during Arabidopsis germination. Plant Physiology, 135, 978-988.

Rea, P. A. (2007). Plant ATP-binding cassette transporters. Annual Review of Plant Biology, 58, 347-375.

Rhee, S. Y., J. Dickerson y D. Xu (2006). Bioinformatics and its applications in plant biology. Annual Review of Plant Biology, 57, 335-360.

Rienties, I. M., J. Vink, J. W. Borst, E. Russinova y S. C. Vries (2005). The Arabidopsis SERK1 protein interacts with the AAA-ATPase AtCDC48, the 14-3-3 protein GF14 and the PP2C phosphatase KAPP. Planta, 221, 394-405.

Salaün, C., L. Lepiniec y B. Dubreucq (2021). Genetic and molecular control of somatic embryogenesis. Plants, 10, 1467.

Salopek-Sondi, B., D. Samec, S. Mihaljevic, A. Smolko, I. Pavlovic, I. Jankovic y J. Ludwig-Müller (2013). Influence of stress hormones on the auxin homeostasis in *Brassica rapa* seedlings. Plant Cell Reports, 32, 1031-1042.

Santana-Buzzy, N., R. Rojas-Herrera, R. M. Galaz-Ávalos, R. Ku-Cauich, J. O. Mijangos Cortés, L. C. Gutiérrez-Pacheco, A. Canto-Flick, F. R. Quiroz-Figueroa y V. M. Loyola-Vargas (2007). Advances in coffee tissue culture and its practical applications. In Vitro Cellular & Developmental Biology - Plant, 43, 507-520.

Santner, A., L. I. Calderon-Villalobos y M. Estelle (2009). Plant hormones are versatile chemical regulators of plant growth. Nature Chemical Biology, 5, 301-307.

Saptari, R. T. y H. Susila (2018). Data mining study of hormone biosynthesis gene expression reveals new aspects of somatic embryogenesis regulation. In Vitro Cellular & Developmental Biology - Plant, 55, 139-152.

Schaller, G. E., A. Bishopp y J. J. Kieber (2015). The Yin-Yang of hormones: Cytokinin and auxin interactions in plant development. The Plant Cell, 27, 44-63.

Schmidt, E. D. L., F. Guzzo, M. A. J. Toonen y S. C. De Vries (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. Development, 124, 2049-2062.

Shiota, H., R. Satoh, K. Watabe, H. Harada y H. Kamada (1998). *C-AB*13, the carrot homologue of the *Arabidopsis AB*13, is expressed during both zygotic and somatic embryogenesis and functions in the regulation of embryo-specific ABA-inducible genes. Plant and Cell Physiology, 39, 1184-1193.

Söndahl, M. R. y W. R. Sharp (1977). High frequency induction of somatic embryos in cultured leaf explants of *Coffea arabica* L. Zeitschrift für Pflanzenphysiologie, 81, 395-408.

Söndahl, M. R., T. Nakamura y W. R. Sharp (1985). Propagation of coffee. Basic Life Sciences, 32, 215-232.

Song, S., Z. Wang, Y. Ren y H. Sun (2020). Full-length transcriptome analysis of the ABCB, PIN/PIN-LIKES, and AUX/LAX families involved in somatic embryogenesis of *Lilium pumilum* DC. Fisch. International Journal of Molecular Sciences, 21, 453.

Song, W.-Y., D. G. Mendoza-Cózatl, Y. Lee, J. I. Schroeder, A. Sang-Nag, H.-S. Lee, T. Wicker y E. Martinoia (2014). Phytochelatin–metal(loid) transport into vacuoles shows different substrate preferences in barley and Arabidopsis. Plant, Cell and Environment, 37, 1192-1201.

Staritsky, G. (1970). Embryoid formation in callus tissues of coffee. Acta Botanica Neederlandica, 19, 509-514.

Staswick, P. E., I. Tiryaki y M. Rowe (2002). Jasmonate response locus *JAR1* and several related *Arabidopsis* genes encode enzymes of the firefly luciferase superfamily that show activity on jasmonic, salicylic, and indole-3-acetic acids in an assay for adenylation. The Plant Cell, 14, 1405-1415.

Staswick, P. E., B. Serban, M. Rowe, I. Tiryaki, M. T. Maldonado, M. C. Maldonado y W. Suza (2005). Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. The Plant Cell, 17, 616-627.

Steiner, N., C. Santa-Catarina, M. Guerra, L. Cutri, M. Carnier Dornelas y E. Floh (2012). A gymnosperm homolog of SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE-1 (*SERK1*) is expressed during somatic embryogenesis. Plant Cell, Tissue and Organ Culture, 109, 41-50.

Su, L., Y. Xie, Z. He, J. Zhang, Y. Tang y X. Zhou (2021). Network response of two cherry tomato (*Lycopersicon esculentum*) cultivars to cadmium stress as revealed by transcriptome analysis. Ecotoxicology and Environmental Safety, 222, 112473.

Tchorbadjieva, M., (2016). Advances in proteomics of somatic embryogenesis, en: Somatic Embryogenesis in Ornamentals and Its Applications, Mujib, A., (ed). Springer. India. pp. 67-90.

Teale, W. D., I. A. Paponov y K. Palme (2006). Auxin in action: signalling, transport and the control of plant growth and development. Nature Reviews Molecular Cell Biology, 7, 847-859.

Thakare, D., W. Tang, K. Hill y S. E. Perry (2008). The MADS-domain transcriptional regulator AGAMOUS-Like 15 promotes somatic embryo development in *Arabidopsis* and soybean. Plant Physiology, 146, 1663-1672.

Thorpe, T., (2012). History of plant tissue culture, en: Plant Tissue Culture Protocols, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Humana Press. Heidelberg. pp. 9-27.

Tian, R., P. Paul, S. Joshi y S. E. Perry (2020). Genetic activity during early plant embryogenesis. Biochemical Journal, 477, 3743-3767.

Ting, L., R. Rad, S. P. Gygi y W. Haas (2011). MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nature Methods, 8, 937-940.

Titapiwatanakun, B. y A. S. Murphy (2009). Post-transcriptional regulation of auxin transport proteins: cellular trafficking, protein phosphorylation, protein maturation, ubiquitination, and membrane composition. Journal of Experimental Botany, 60, 1093-1107.

Tonietto, Â., J. Hiromi Sato, J. Batista Teixeira, E. M. de Souza, F. O. Pedrosa, O. L. Franco y A. Metha (2012). Proteomic analysis of developing somatic embryos of *Coffea arabica*. Plant Molecular Biology Reporter, 30, 1393-1399.

Tromas, A., S. Paque, V. Stierlé, A. L. Quettier, P. Muller, E. Lechner, P. Genschik y C. Perrot-Rechenmann (2013). Auxin-Binding Protein 1 is a negative regulator of the SCF TIR1/AFB pathway. Nature Communications, 4, 2496.

Tvorogova, V. E. y L. A. Lutova (2018). Genetic regulation of zygotic embryogenesis in angiosperm plants. Russian Journal of Plant Physiology, 65, 1-14.

Tvorogova, V. E., Y. A. Fedorova, E. A. Potsenkovskaya, A. A. Kudriashov, E. P. Efremova, V. A. Kvitkovskaya, T. W. Wolabu, F. Zhang, M. Tadege y L. A. Lutova (2019). The WUSCHEL-related homeobox transcription factor MtWOX9-1 stimulates somatic embryogenesis in *Medicago truncatula*. Plant Cell, Tissue and Organ Culture, 138, 517-527.

Uc-Chuc, M. Á., C. A. Pérez-Hernández, R. M. Galaz-Ávalos, L. Brito-Argáez, V. Aguilar-Hernández y V. M. Loyola-Vargas (2020). YUCCA-mediated biosynthesis of the auxin IAA is required during the somatic embryogenic induction process in *Coffea canephora*. International Journal of Molecular Sciences, 21, 4751.

Verrier, P. J., D. Bird, B. Burla, E. Dassa, C. Forestier, M. Geisler, M. Klein, U. Kolukisaoglu, Y. Lee, E. Martinoia, A. Murphy, P. A. Rea, L. Samuels, B. Schulz, E. J. Spalding, K. Yazaki y F. L. Theodoulou (2008). Plant ABC proteins - a unified nomenclature and updated inventory. Trends in Plant Science, 13, 151-195.

Vinod, K., M. M. Naidu y G. A. Ravishankar (2006). Developments in coffee biotechnology-*in vitro* plant propagation and crop improvement. Plant Cell, Tissue and Organ Culture, 87, 49-65.

Vondráková, Z., J. Krajnáková, L. Fischerová, M. Vágner y K. Eliášová, (2016). Physiology and role of plant growth regulators in somatic embryogenesis, en: Vegetative Propagation of Forest Trees, Park, S., J. M. Bonga y H.-K. Moon, (eds). National Institute of Forest Science. Seoul, Korea. pp. 123-169.

Wang, Y. C., M. Z. Lin, B. Huang, H. H. Chung y J. T. Chen (2018). Thidiazuron enhanced somatic embryogenes is from callus lines of Arabica coffee and subsequent plant regeneration. Acta Biologica Cracoviensia Series Botanica, 60, 35-44.

Wanke, D. y H. Üner Kolukisaoglu (2010). An update on the ABCC transporter family in plants: many genes, many proteins, but how many functions? Plant Biology, 12, 15-25.

Weijers, D. y D. Wagner (2016). Transcriptional responses to the auxin hormone. Annual Review of Plant Biology, 67, 539-574.

Wendt dos Santos, A. L., R. Souza Reis, A. Schuabb Heringer, E. I. Segal Floh, C. Santa-Catarina y V. Silveira, (2018). Proteomics as a tool to study molecular changes during plant morphogenesis *in vitro*, en: Plant Cell Culture Protocols, Loyola-Vargas, V. M. y N. Ochoa-Alejo, (eds). Springer. New York, NY. pp. 339-349.

Westfall, C. S., J. Herrmann, Q. Chen, S. Wang y J. M. Jez (2010). Modulating plant hormones by enzyme action. The GH3 family of acyl acid amido synthetases. Plant Signaling & Behavior, 5, 1607-1612.

Westfall, C. S., C. Zubieta, J. Herrmann, U. Kapp, M. H. Nanao y J. M. Jez (2012). Structural basis for prereceptor modulation of plant hormones by GH3 proteins. Science, 336, 1708-1711.

Wickham, H. (2016). Ggplot2: elegant graphics for data analysis, Springer, New York, NY. 260 p.

Widemann, E., L. Miesch, R. Lugan, E. Holder, C. Heinrich, Y. Aubert, M. Miesch, F. Pinot y T. Heitz (2013). The amidohydrolases IAR3 and ILL6 contribute to jasmonoyl-isoleucine hormone turnover and generate 12-hydroxyjasmonic acid upon wounding in Arabidopsis leaves. Journal of Biological Chemistry, 288, 31701-31714.

Winkelmann, T., (2016). Somatic versus zygotic embryogenesis: Learning from seeds, en: In Vitro Embryogenesis in Higher Plants, Germanà, M. A. y M. Lambardi, (eds). Springer. New York. pp. 25-46.

Wójcikowska, B. y M. D. Gaj (2017). Expression profiling of *AUXIN RESPONSE FACTOR* genes during somatic embryogenesis induction in Arabidopsis. Plant Cell Reports, 36, 843-858.

Wu, X., J. Yan, Y. Wu, H. Zhang, S. Mo, X. Xu, F. Zhou y H. Ding (2019). Proteomic analysis by iTRAQ-PRM provides integrated insight into mechanisms of resistance in pepper to *Bemisia tabaci* (Gennadius). BMC Plant Biology, 19, 270.

Xia, Y., W. Chen, W. Xiang, D. Wang, B. Xue, X. Liu, L. Xing, D. Wu, S. Wang y Q. Guo (2021). Integrated metabolic profiling and transcriptome analysis of pigment accumulation in *Lonicera japonica* flower petals during colour-transition. BMC Plant Biology, 21, 98.

Xiao, Y., Y. Chen, Y. Ding, J. Wu, P. Wang, Y. Yu, X. Wei, Y. Wang, C. Zhang y F. Li (2018). Effects of *GhWUS* from upland cotton (*Gossypium hirsutum* L.) on

somatic embryogenesis and shoot regeneration. Plant Science, 270, 157-165.

Xinxin, Z., Y. Shuang, Z. Xunming, W. Shang, Z. Juhong y X. Jinghui (2020). TMT-Based quantitative proteomic profiling of overwintering Lissorhoptrus oryzophilus. Frontiers in Physiology, 10, 1623.

Xu, Y. X., Y. Liu, S. T. Chen, X. Q. Li, L. G. Xu, Y. H. Qi, D. A. Jiang y S. H. Jin (2014). The B subfamily of plant ATP binding cassette transporters and their roles in auxin transport. Biologia Plantarum, 58, 401-410.

Yadav, N., S. M. P. Khurana y D. Yadav, (2015). Plant secretomics: Unique initiatives, en: PlantOmics: The Omics of Plant Science, Barh, D., M. S. Khan y E. Davies, (eds). Springer. India. pp. 357-384.

Yadav, R. K., M. Perales, J. Gruel, T. Girke, H. Jonsson y G. V. Reddy (2011). WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. Genes & Development, 25, 2025-2030.

Yang, H. y A. S. Murphy (2009). Functional expression and characterization of Arabidopsis ABCB, AUX 1 and PIN auxin transporters in Schizosaccharomyces pombe. The Plant Journal, 59, 179-191.

Yang, X., X. Zhang, D. Yuan, F. Jin, Y. Zhang y J. Xu (2012). Transcript profiling reveals complex auxin signalling pathway and transcription regulation involved in dedifferentiation and redifferentiation during somatic embryogenesis in cotton. BMC Plant Biology, 12, 110.

Yasuda, T., Y. Fujii y T. Yamaguchi (1985). Embryogenic callus induction from *Coffea arabica* leaf explants by benzyladenine. Plant and Cell Physiology, 26, 595-597.

Yu, R., Q. Jiang, C. Xv, L. Li, S. Bu y G. Shi (2019). Comparative proteomics analysis of peanut roots reveals differential mechanisms of cadmium detoxification and translocation between two cultivars differing in cadmium accumulation. BMC Plant Biology, 19, 1-15.

Zazimalová, E., P. Krecek, P. Skúpa, K. Hoyerová y J. Petrášek (2007). Polar transport of the plant hormone auxin – the role of PIN-FORMED (PIN) proteins. Cellular and Molecular Life Sciences, 64, 1621-1637.

Zhang, L. f., W. F. Li, H. y. Xu, L. w. Qi y S. y. Han (2014). Cloning and characterization of four differentially expressed cDNAs encoding NFYA homologs involved in responses to ABA during somatic embryogenesis in Japanese larch (*Larix leptolepis*). Plant Cell, Tissue and Organ Culture, 117, 293-304.

Zhang, Y., Y. Zhang, J. Yu, H. Zhang, L. Wang, S. Wang, S. Guo, Y. Miao, S. Chen y Y. Li (2019). NaCl-responsive ROS scavenging and energy supply in alkaligrass callus revealed from proteomic analysis. BMC Genomics, 20, 990.

Zhao, J., B. Wang, X. Wang, Y. Zhang, M. Dong y J. Zhang (2015). iTRAQ-based comparative proteomic analysis of embryogenic and non-embryogenic tissues of Prince Rupprecht's larch (*Larix principis-rupprechtii* Mayr). Plant Cell, Tissue and Organ Culture, 120, 655-669.

Zheng, W., X. Zhang, Z. Yang, J. Wu, F. Li, L. Duan, C. Liu, L. Lu, C. Zhang y F. Li (2014). AtWuschel promotes formation of the embryogenic callus in *Gossypium hirsutum*. PLoS ONE, 9, e87502.

Zolman, B. K., N. Martinez, A. Millius, A. R. Adham y B. Bartel (2008). Identification and characterization of Arabidopsis indole-3-butyric acid response mutants defective in novel peroxisomal enzymes. Genetics, 180, 237-251.

ANNEXES

Annexed 1. Several protein families involved in auxin homeostasis in *C. arabica*.

Protein GH3 family (Group II)

% identity *% similarity	AtGH3.1	AtGH3.2	AtGH3.3	AtGH3.4	AtGH3.5	AtGH3.6	AtGH3.9	AtGH3.17
CaGH3.17a	47 *66	47 *68	48 *68	45 *63	51 *61	50*68	49 *67	79 *91
CaGH3.17b	46 *64	27 *32	22 *47	48 *66	48 *79	48*67	45 *64	61 *75

Amidohydrolases ILR1

% identity *% similarity	ILR1	ILL1	ILL2	ILL3	IAR3	ILL5	ILL6
CaILR1-Like-1	76 *90	72 *85	69 *84	58 *83	75 *88	51 *70	59 *77
CaILR1-Like-2	51*70	53 *72	55 *73	50 *69	50 *71	48 *59	51 *71
CaILR1-Like-4	54 *72	64 *80	66 *80	53 *70	73 *90	22 *41	49 *69

Protein ABC family (subfamily ABCB)

% identity *% similarity	ABCB1	ABCB2	ABCB11	ABCB19	ABCB21
ABCB2	51 *69	78 *90	45 *65	53 *73	45*66
ABCB4	43 *64	43 *62	65 *79	44 *64	45*67
ABCB14	44 *65	45 *64	65 *75	46 *66	74 *86

The highest identity value that *C. arabica* proteins share with respect to *A. thaliana* is shown in bold and the highest similarity value is shown with an asterisk. Data were obtained by BLAST analysis. For GH3, the analysis was performed with group II proteins involved in auxin conjugation. For amidohydrolases, an analysis was carried out with the proteins of the seven members of the ILR1 family involved in the hydrolysis of auxin. For the ABCB subfamily, the analysis was carried out against some of the proteins related to auxin transport.

Annexed 2. Selected sequences for phylogenetic analysis.

Specie	Protein family	ID	mRNA ID	Protein ID
Arabidopsis thaliana	GH3	AT2G14960	NM_127059.3	NP_179101.1
		AT4G37390	NM_119902.4	NP_195455.1
		AT2G23170	NM_127881.3	NP_179898.1
		AT1G59500	NM_104643.1	NP_176159.1
		AT4G27260	NM_118860.5	NP_194456.1
		AT5G54510	NM_124831.3	NP_200262.1
		AT1G28130	NM_102578.4	NP_174134.1
		AT5G51470	NM_124526.2	NP_199960.1
		AT2G47750	NM_130342.3	NP_182296.1
		AT4G03400	NM_001340446.1	NP_001319858.1
		AT2G46370	NM_180122.4	NP_850453.1
		AT5G13320	NM_121335.4	NP_196836.1
		AT5G13350	NM_001343270.1	NP_001318554.1
		AT5G13360	NM_001203372.2	NP_001190301.1
		AT5G13370	NM_121340.3	NP_196841.2
		AT5G13360	NM_001203372.2	NP_001190301.1
		AT1G28130	NM_102578.4	NP_174134.1
		AT1G48670	NM_001333347.1	NP_001321105.1
		AT1G48660	NM_001333346.1	P_001322825.1
Solanum lycopersicum	GH3	101251833	XM_004231843.4	XP_004231891.1
		101246970	NM_001321687.1	P_001308616.1
		101268544	XM_004233398.3	XP_004233446.1
		101251833	XM_004231843.4	XP_004231891.1
		101251682	M_004240036.4	XP_004240084.1
		101262636	XM_004244120.4	XP_004244168.1
		101258198	XM_010328855.3	XP_010327157.1
		101258277	XM_004243271.4	XP_004243319.1
		101262636	XM_004244120.4	XP_004244168.1
		101258495	NM_001368308.1	NP_001355237.1
		101258495	XM_004248101.4	XP_004248149.1
		1262663	XM_004248029.4	XP_004248077.1
		101262357	XM_004248028.3	XP_004248076.1
		101266841	M_004243380.4	XP_004243428.1
		101264094	XM_004251437.4	XP_004251485.1
Oryza sativa	GH3	NC_029256.1	XM_015762657.1	XP_015618143.1
		NC_029256.1	XM_015787988.1	XP_015643474.1
		NC_029260.1	XM_015782624.1	XP_015638110.1
		NC_029262.1	XM_015790552.1	XP_015646038.1
		NC_029262.1	XM_015790835.1	XP_015646321.1
		NC_029262.1	XM_015792311.1	XP_015647797.1
		NC_029262.1	XM_015792333.1	XP_015647819.1
		NC_029266.1	XM_015760301.1	XP_015615787.1
		LOC4343704		
		LOC_O38860		
		LOC4344247		
		LOC107275638		
		Os11g0528700		
Arabidopsis	ILR1	AT3G02875	NM_001337440.1	NP_001325526.1

thaliana		AT5G56650	NM_001345211.1	NP_001332665.1
		AT5G56660	NM_125049.3	NP_200477.1
		AT5G54140	NM_124794.3	NP_200225.1
		AT1G51760	NM_104055.4	NP_175587.1
		AT1G51780	NM_104057.1	NP_175589.1
		AT1G44350	NM_103546.4	NP_175086.1
Oryza sativa	ILR1	OSNPB_010560000	XM_015766297.1	XP_015621783.1
		OSNPB_010706900	XM_015768847.2	XP_015624333.1
		OSNPB_030836800	XM_015773443.2	XP_015628929.1
		OSNPB_040521800	XM_015780885.2	XP_015636371.1
		OSNPB_010917300	XM_015776151.2	XP_015631637.1
		OSNPB_070249700	XM_015792387.1	XP_015647873.1
		OSNPB_070249800	XM_015790315.2	XP_015645801.1
		OSNPB_060691400	XM_015785503.2	P_015640989.1
Solanum lycopersicum	ILR1	LOC101267174	XM_019216147.2	XP_019071692.1
		LOC101264404	XM_004251605.4	XP_004251653.1
		LOC101264129	XM_004241011.4	XP_004241059.1
		LOC101254822	XM_004228599.4	XP_004228647.1
		LOC101266415	XM_004238668.4	XP_004238716.1
		LOC101249161	XM_004235576.4	P_004235624.1
Arabidopsis thaliana	ABCB	AT4G25960	NM_118729.4	NP_194326.2
		AT2G36910	NM_129247.3	NP_181228.1
		AT4G01820	NM_116412.2	NP_192091.1
		AT2G47000	NM_130268.4	NP_182223.1
		AT4G01830	NM_001340339.1	NP_001329375.1
		AT2G39480	NM_129506.2	NP_181480.1
		AT5G46540	NM_124024.2	NP_199466.1
		AT4G18050	NM_001341257.1	NP_001328203.1
		AT1G10680	NM_001331921.1	NP_001320481.1
		AT3G55320	NM_115390.3	NP_191092.1
		AT1G02520	NM_100133.3	NP_171753.1
		AT1G02530	NM_001331344.1	NP_001320752.1
		AT1G27940	NM_102559.2	NP_174115.1
		AT1G28010	NM_102566.3	NP_174122.1
		AT3G28345	NM_113754.3	NP_189475.1
		AT3G28360	NM_113756.5	NP_189477.4
		AT3G28390	NM_113759.1	NP_189480.1
		AT3G28380	NM_113758.2	NP_189479.1
		AT3G28860	NM_113807.3	NP_189528.1
		AT3G28415	NM_148757.1	NP_683599.1
		AT3G62150	NM_001340142.1	NP_001327193.1
Oryza sativa	ABCB	OSNPB_080564300	XM_015794906.2	P_015650392.1
		OSNPB_020693700	XM_015768291.2	XP_015623777.1
		OSNPB_010290700	XM_015756580.2	XP_015612066.1
		OSNPB_050548500	XM_015782329.2	XP_015637815.1
		OSNPB_010534700	XM_015786358.2	XP_015641844.1
		OSNPB_030181675	XM_026024189.1	XP_025879974.1
		OSNPB_020323000	XM_015771498.2	XP_015626984.1
		OSNPB_030280000	XM_015775083.2	XP_015630569.1
		OSNPB_010911300	XM_015758160.2	XP_015613646.1
		OSNPB_010723800	XM_015758719.2	XP_015614205.1
Solanum lycopersicum	ABCB	NC_015444.2	XM_004243033.1	XP_004243081.1
		NC_015446.2	XM_004246576.3	XP_004246624.1

NC_015442.2	XM_004239442.1	XP_004239490.1
NC_015448.2	XM_004251332.3	XP_004251380.1
NC_015449.2	XM_004253109.3	XP_004253157.2
NC_015443.2	XM_010323557.2	XP_010321859.1
NC_015439.2	NM_001247280.1	NP_001234209.1
NC_015444.2	XM_004244308.3	XP_004244356.1
NC_015448.2	XM_010315050.2	XP_010313352.1
NC_015440.2	XM_010319878.2	XP_010318180.1
NC_015449.2	XM_019211312.1	XP_019066857.1
NC_015439.2	XM_004232205.3	XP_004232253.1

The sequences of the proteins GH3, ILR1 and ABCB of C. arabica used for the phylogenetic trees were obtained in this study.

Annexed 3. Description of protein-protein interaction network of SE-related proteins.

Node	Domain summary
ABCB1	ABC transporter B family member 1: Auxin efflux transporter that acts as a negative regulator of light signaling to promote hypocotyl elongation. This transporter mediates the accumulation of chlorophyll and anthocyanin and the expression of genes in response to light. Participates directly in auxin efflux and thus regulates the polar (presumably basipetal) auxin transport (from root tips to root elongating zone). Transports also some auxin metabolites such as oxindoleacetic acid and indole acetaldehyde. It involves diverse auxin-mediated responses, including gravitropism, phototropism, etc.
ABCB21	ABC transporter B family member 21.
ABCB28	ABC transporter B family member 28 belongs to the ABC transporter superfamily. ABCB family. Multidrug resistance exporter (TC 3.A.1.201) subfamily.
ABCC1	ABC transporter C family member 1; Pump for glutathione S-conjugates. Mediates the transport of S-(2,4-dinitrophenyl)-glutathione (DNP-GS), GSSG, cyanidin 3-glucoside- GS (C3G-GS), and metolachlor-GS (MOC-GS); Belongs to the ABC transporter superfamily. ABCC family. Conjugate transporter (TC 3.A.1.208) subfamily.
ABCC14	ABC transporter C family member 14; Pump for glutathione S-conjugates; Belongs to the ABC transporter superfamily. ABCC family. Conjugate transporter (TC 3.A.1.208) subfamily.
ABCC2	ABC transporter C family member 2; Pump for glutathione S-conjugates. Mediates the transport of S-conjugates such as GSH, S-(2,4-dinitrophenyl)-glutathione (DNP-GS), GSSG, cyanidin 3-glucoside-GS (C3G-GS) and metolachlor-GS (MOC-GS), glucuronides such as 17-beta-estradiol 17-(beta-D-glucuronide) (E(2)17betaG), and of the chlorophyll catabolite such as B.napus nonfluorescent chlorophyll catabolite (Bn-NCC-1). This protein belongs to the ABC transporter superfamily. ABCC family. Conjugate transporter (TC 3.A.1.208) subfamily.
ABCC5	ABC transporter C family member 5; Pump for glutathione S-conjugates. It regulates K(+) and Na(+) cell content. Mediates resistance to NaCl and Li(+), confers sensitivity to sulfonylurea drugs such as glibenclamide (inducer of stomatal opening) and is required for stomatal opening regulation by auxin, abscisic acid (ABA), and external Ca(2+). Transports oestradiol-17-(beta-D-glucuronide) (E(2)17G). Involved in the root auxin content regulation that controls the transition from primary root elongation to lateral root formation. Plays a role in ABA-mediated germination inhibit [...]
ABCC8	ABC transporter C family member 8; Pump for glutathione S-conjugates.
ABCD1	ABC transporter D family member 1: Contributes to transporting fatty acids and their derivatives (acyl CoAs) across the peroxisomal membrane. Provides acetate to the glyoxylate cycle in developing seedlings. It is involved in pollen tube elongation, ovule fertilization, and seed germination after imbibition (controls the switch between the opposing developmental programs of dormancy and germination), probably by promoting beta-oxidation of storage lipids during gluconeogenesis. It is required to synthesize jasmonic acid and convert indole butyric acid to indole acetic acid.
ABCF1	ABC transporter F family member 1 belongs to the ABC transporter superfamily. ABCF family. EF3 (TC 3.A.1.121) subfamily.
ABCF3	ABC transporter F family member 3 belongs to the ABC transporter superfamily. ABCF family. EF3 (TC 3.A.1.121) subfamily.
ABCG1	ABC transporter G family member 1.

ABCG40	ABC transporter G family member 40; Maybe a general defense protein (By similarity). Functions as a pump to exclude Pb(2+) ions and/or Pb(2+)-containing toxic compounds from the cytoplasm. This transporter contributes to Pb(2+) ions resistance. Confers some resistance to the terpene sclareol.
ABCG7	ABC transporter G family member 7 belongs to the ABC transporter superfamily. ABCG family. Eye pigment precursor importer (TC 3.A.1.204) subfamily.
ABCI8	UPF0051 protein ABCI8, chloroplastic; involved in light signaling, probably by mediating the transport and correct distribution of protoporphyrin IX, a chlorophyll precursor, in response to far-red light; Belongs to the UPF0051 (ycf24) family.
AG	Floral homeotic protein AGAMOUS: Probable transcription factor involved in the control of organ identity during the early development of flowers. It is required for normal development of stamens and carpels in the wild-type flower. It plays a role in maintaining the determinacy of the floral meristem. Acts as a C-class cadastral protein by repressing the A-class floral homeotic genes like APETALA1. Forms a heterodimer via the K-box domain with either SEPALATTA1/AGL2, SEPALATTA2/AGL4, SEPALLATA3/AGL9 or AGL6 that could be involved in genes regulation during floral meristem development.
APT1	Adenine phosphoribosyltransferase 1, chloroplastic: Catalyzes a salvage reaction resulting in the formation of AMP that is energetically less costly than de novo synthesis. It contributes primarily to the recycling of adenine into adenylate nucleotides but is also involved in the inactivation of cytokinins by phosphoribosylation. This enzyme catalyzes the conversion of cytokinins from free bases (active form) to the corresponding nucleotides (inactive form). This protein belongs to the purine/pyrimidine phosphoribosyltransferase family.
ARR14	Two-component response regulator ARR14; Transcriptional activator that binds specifically to the DNA sequence 5'-[AG]GATT-3'. Functions as a response regulator involved in the His-to-Asp phosphorelay signal transduction system. Phosphorylation of the Asp residue in the receiver domain activates the ability of the protein to promote the transcription of target genes. Could directly activate some type-A response regulators in response to cytokinins (By similarity); Belongs to the ARR family—Type-B subfamily.
ASK11	SKP1-like protein 11; Involved in ubiquitination and subsequent proteasomal degradation of target proteins. Together with CUL1, RBX1, and a F-box protein, it forms a SCF E3 ubiquitin ligase complex. The functional specificity of this complex depends on the type of F-box protein. The SCF complex serves as an adapter that links the F-box protein to CUL1 (By similarity). This protein plays a role during early flower reproductive development.
B"ALPHA	Serine/threonine protein phosphatase 2A regulatory subunit B"alpha; Regulatory subunit of type 2A protein phosphatase. It is not involved in HMGR regulation in seedlings grown in standard medium but negatively regulates root growth in response to salt.
B"BETA	Serine/threonine protein phosphatase 2A regulatory subunit B"beta; Regulatory subunit of type 2A protein phosphatase. Involved in post-transcriptional regulation of HMGR but not root growth regulation in response to salt.
B"DELTA	Probable serine/threonine protein phosphatase 2A regulatory subunit B"delta; Probable regulatory subunit of type 2A protein phosphatase.
BIG	Auxin transport protein BIG. This transporter is required for auxin efflux and polar auxin transport (PAT) influencing auxin-mediated developmental responses (e.g., cell elongation, apical dominance, lateral root production, inflorescence architecture, general growth, and development). Controls the elongation of the pedicels and stem internodes through auxin action. Involved in the expression modulation of light-regulated genes.

	Represses CAB1 and CAB3 gene expression in etiolated seedlings.
CRT1	Calreticulin-1: Molecular calcium-binding chaperone promoting folding, oligomeric assembly, and quality control in the ER via the calreticulin/calnexin cycle. This lectin may interact transiently with almost all monoglucosylated glycoproteins synthesized in the ER (By similarity).
CRT2	Calreticulin-2: Molecular calcium-binding chaperone promoting folding, oligomeric assembly, and quality control in the ER via the calreticulin/calnexin cycle. This lectin may interact transiently with almost all monoglucosylated glycoproteins synthesized in the ER (By similarity).
DRMH1	Dormancy-associated protein homolog 1.
GRF1-2	14-3-3-like protein GF14 chi is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element in plant genes that regulates nutrient metabolism.
GRF11	14-3-3-like protein GF14 omicron is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element found in plant genes.
GRF12	14-3-3-like protein GF14 iota is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element in plant genes.
GRF2	14-3-3-like protein GF14 omega is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element found in plant genes. This protein belongs to the 14-3-3 family.
GRF3	14-3-3-like protein GF14 psi is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element found in plant genes that regulates nutrient metabolism—reciprocal negative transcription regulation of miR396. Negative regulator of constitutive freezing tolerance and cold acclimation by controlling cold-induced gene expression partially through ethylene (ET)-dependent pathway; prevents ethylene (ET) biosynthesis, probably by binding 1- aminocyclopropane-1-carboxylate synthases (ACS) to reduce their stability.
GRF4	14-3-3-like protein GF14 phi is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element in plant genes.
GRF5	14-3-3-like protein GF14 upsilon is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element found in plant genes. It may be involved in cell cycle regulation by binding to soluble EDE1 and sequestering it in an inactive form during the early stages of mitosis.
GRF7	14-3-3-like protein GF14 nu is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element found in plant genes.
GRF8	14-3-3-like protein GF14 kappa is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element in plant genes that regulates nutrient metabolism. Negative regulator of freezing tolerance that modulates cold-responsive C-repeat-binding factors (CBF) DREB1A AND DREB1B proteins stability by facilitating their ubiquitin-mediated degradation.
GRF9-2	14-3-3-like protein GF14 mu is associated with a DNA binding complex that binds to the G box, a well-characterized cis-acting DNA regulatory element found in plant genes. This protein belongs to the 14-3-3 family.
ILR1	IAA-amino acid hydrolase ILR1 hydrolyzes specific amino acid conjugates of the plant growth regulator indole-3-acetic acid (IAA), including IAA-Phe, IAA-Leu, and IAA-Tyr. We can also use IAA-Ala, IAA-Gly, IAA-Met, and IAA-Glu as substrates with low efficiency.

This enzyme does not show activity with IAA-Ile, IAA-1-O-beta-D-glucose, or IAA-myo-inositol. It is the most efficient enzyme of the ILL family for IAA-Leu hydrolysis. Necessary for IAA-Leu and IAA-Phe hydrolysis in roots. May act with ILL2 to provide free IAA to germinating seedlings. This hydrolase belongs to the peptidase M20 family.

LOG3 Cytokinin riboside 5'-monophosphate phosphoribohydrolase LOG3. Cytokinin-activating enzyme working in the direct activation pathway. Phosphoribohydrolase converts inactive cytokinin nucleotides to biologically active free-base forms.

NFYB1 Nuclear transcription factor Y subunit B-1. Component of the NF-Y/HAP transcription factor complex. The NF-Y complex stimulates the transcription of various genes by recognizing and binding to a CCAAT motif in promoters belonging to the NFYB/HAP3 subunit family.

NFYB3 Nuclear transcription factor Y subunit B-3. Component of the NF-Y/HAP transcription factor complex. The NF-Y complex stimulates the transcription of various genes by recognizing and binding to a CCAAT motif in promoters.

NFYC4 Nuclear transcription factor Y subunit C-4 stimulates the transcription of various genes by recognizing and binding to a CCAAT motif in promoters (By similarity). It is involved in the abscisic acid (ABA) signaling pathway.

PP2A2-2 Serine/threonine-protein phosphatase PP2A-2 catalytic subunit: Dephosphorylates and activates the actin-depolymerizing factor ADF1, which, in turn, regulates actin cytoskeleton remodeling and is involved in the blue light photoreceptor PHOT2-mediated chloroplast avoidance movements. This protein is associated with the serine/threonine-protein phosphatase PP2A regulatory subunits A and B' to positively regulate the beta-oxidation of fatty acids and protoauxins in peroxisomes by dephosphorylating peroxisomal beta-oxidation-related proteins. It acts as a negative regulator of abscisic acid (ABA) signaling.

PP2A4 Serine/threonine-protein phosphatase PP2A-4 catalytic subunit; Functions redundantly with PP2A3 and is involved in establishing auxin gradients, apical-basal axis of polarity, and root and shoot apical meristem during embryogenesis. May dephosphorylate PIN1 and regulate its subcellular distribution for polar auxin transport. The holoenzyme composed of PP2AA1, PP2A4, and B'ZETA or B'ETA acts as a negative regulator of plant innate immunity by controlling the BAK1 phosphorylation state and activating surface-localized immune receptor complexes. This phosphatase belongs to the PPP phosphatase family.

PP2AA1 Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform; The A subunit of protein phosphatase 2A serves as a scaffolding molecule to coordinate the assembly of the catalytic subunit and a variable regulatory B subunit. It seems to act as a positive regulator of PP2A catalytic activity. This protein confers resistance to phosphatase inhibitors such as okadaic acid and cantharidin. It is involved during developmental processes such as seedling and floral development, root gravitropism, and stomatal opening regulation and regulating auxin efflux.

PP2AA2 Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform; The A subunit of protein phosphatase 2A serves as a scaffolding molecule to coordinate the assembly of the catalytic subunit and a variable regulatory B subunit. It is involved during developmental processes such as seedling and floral development. It seems to act as a negative regulator of PP2A catalytic activity. This protein associates with the serine/threonine-protein phosphatase PP2A catalytic subunit C and regulatory subunit B' to positively regulate the beta-oxidation of fatty acids and protoauxins in peroxisomes.

PP2AA3 Serine/threonine-protein phosphatase 2A 65 kDa, regulatory subunit A gamma isoform. The A subunit of protein phosphatase 2A is a scaffolding molecule to coordinate the

	assembly of the catalytic subunit and a variable regulatory B subunit. It is involved in developmental processes such as seedling and floral development. It seems to act as a negative regulator of PP2A catalytic activity.
PP2AB2	Serine/threonine protein phosphatase 2A 55 kDa regulatory subunit B beta isoform; The B regulatory subunit may modulate substrate selectivity and catalytic activity and also may direct the localization of the catalytic enzyme to a particular subcellular compartment.
SERK1	Somatic embryogenesis receptor kinase 1: Dual specificity kinase acting on serine/threonine- and tyrosine-containing substrates. Phosphorylates BRI1 on 'Ser-887' and CDC48 on at least one threonine residue and on 'Ser-41'. This protein confers embryogenic competence. This receptor acts redundantly with SERK2 as a control point for sporophytic development, controlling male gametophyte production. It is involved in the brassinolide signaling pathway.
TSB2	Tryptophan synthase beta chain 2, chloroplastic. The beta subunit synthesizes L-tryptophan from indole and L-serine and belongs to the TrpB family.

Annexed 4. Reporting guidelines for mass spectrometry (MIAPE).

1. General Features

1.1 Global descriptors

- Responsible person (or institutional role if more appropriate); provide name, affiliation, and stable contact information: Eliel Ruiz May/José Miguel Elizalde Contreras. Red de Estudios Moleculares Avanzados, Clúster Científico y Tecnológico BioMimic®, Instituto de Ecología A.C. (INECOL), Carretera Antigua a Coatepec No. 351, Congregación el Haya, CP 91070, Xalapa, Veracruz, México. eliel.ruiz@inecol.mx.
- Instrument manufacturer and model: Orbitrap Fusion Tribrid (Thermo-Fisher Scientific, San Jose, CA, USA) mass spectrometer.
- Customizations (summary): Global Settings

Method Duration (min)= 120

Ion Source Type = NSI

Spray Voltage: Positive Ion (V) = 3500

Spray Voltage: Negative Ion (V) = 600

Sweep Gas (Arb) = 0

Ion Transfer Tube Temp (°C) = 280

Internal Mass Calibration= Easy-IC

Pressure Mode = Standard

Default Charge State = 2

Experiment 1

Start Time (min) = 0

End Time (min) = 120

Cycle Time (sec) = 3

Do data dependent experiment if no target species are found = False

Scan MasterScan MSn Level = 1

Use Wide Quad Isolation = True

Detector Type = Orbitrap

Orbitrap Resolution = 120K

Mass Range = Normal

Scan Range (m/z) = 350-1500

Maximum Injection Time (ms) = 50

AGC Target = 400000

Microscans = 1

S-Lens RF Level = 60

Use ETD Internal Calibration = True

DataType = Profile

Polarity = Positive
Source Fragmentation = False
Filter MIPS
Filter Type = MIPS
MIPS On = 2
Relax Restrictions = True
Filter ChargeState
Filter Type = ChargeState
Include charge state(s) = 2-8
Include undetermined charge states = False
Include charge states 25 and higher = False
Filter DynamicExclusion
Filter Type = DynamicExclusion
Exclude after n times = 1
Exclude isotopes = True
Perform dependent scan on single charge state per precursor only = False
If occurs within (s) = 30
Exclusion duration (s) = 90
Excl. Mass Width = ppm
Mass tolerance low = 10
Mass tolerance high = 10
Filter IntensityThreshold
Filter Type = IntensityThreshold
Signal Intensity = 10000
Decision
Precursor Priority = MostIntense
Scan Event 1
ChargeRange: 3-3
AND
MZRange: 300-1600
OR
ChargeRange: 4-4
AND
MZRange: 300-1600
OR

ChargeRange: 5-5
AND
MZRange: 300-1600
OR
ChargeRange: 6-8
Scan Event 2
ChargeRange: 2-2
OR
ChargeRange: 3-3
AND
MZRange: 300-1600
OR
ChargeRange: 4-4
AND
MZRange: 300-1600
OR
ChargeRange: 5-5
AND
MZRange: 300-1600
OR
ChargeRange: 6-8
Scan Event 1
Do data dependent experiment if no target species are found = False
Scan ddMSnScan MSn Level = 2
Top N= 0
Isolation Mode = Quadrupole
Isolation Window = 1.6
Scan Range Mode = Auto Normal FirstMass = 120
ActivationType = HCD
Is Stepped Collision Energy On = False
Stepped Collision Energy (%) = 5
Multistage Activation = False
Neutral Loss Mass = 50.0001
Collision Energy (%) = 28
Detector Type = Orbitrap

Orbitrap Resolution = 30K
Maximum Injection Time (ms) = 50
AGC Target = 50000
Inject ions for all available parallelizable time = True Microscans = 1
Activation Q = 0.25
Use ETD Internal Calibration = False
DataType = Centroid
Polarity = Positive
Source Fragmentation = False
Scan Event 2
Do data dependent experiment if no target species are found = False
Scan ddMSnScan MSn Level = 2
Isolation Mode = Quadrupole
Top N= 0
Isolation Window = 1.6
Use Isolation m/z Offset = False
Multi-notch Isolation = False
Scan Range Mode = Auto Normal FirstMass = 100
ActivationType = CID Collision Energy (%) = 35
Neutral Loss Mass = 50.0001
Is Stepped Collision Energy On = False
Stepped Collision Energy (%) = 5
Multistage Activation = False
Is EThcD Active = False
Detector Type = Orbitrap
Orbitrap Resolution = 30K
Maximum Injection Time (ms) = 50
AGC Target = 50000
Inject ions for all available parallelizable time = True Microscans = 1
Activation Q = 0.25
DataType = Centroid
Polarity = Positive
Source Fragmentation = False
HPLC
Run time: 121.000 [min]

Instrument: MININT-82L3M2J_1 on minint-82l3m2j Description:

initial Instrument Setup: PumpModule.LoadingPump.%A.Equate: "H2O +0.1% Formic Acid", PumpModule.LoadingPump.%B.Equate: "ACN +0.1%Formic Acid", PumpModule.LoadingPump.%C.Equate: "%C", PumpModule.NC_Pump.%A.Equate: "%A" H2O +0.1% Formic Acid, PumpModule.NC_Pump.%B.Equate: "%B" ACN +0.1% Formic Acid

-20.000 [min] Equilibration, PumpModule.LoadingPump.Flow.Nominal: 3.000 [µl/min] PumpModule.LoadingPump.%B.Value: 0.0 [%] PumpModule.LoadingPump.%C.Value: 0.0 [%] PumpModule.LoadingPump.Curve:5, PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min] PumpModule.NC_Pump.%B.Value: 7.0 [%], PumpModule.NC_Pump.Curve: 5

0.000 [min] Inject Preparation, Wait PumpModule.LoadingPump.Ready And PumpModule.NC_Pump.Ready And ColumnOven.Ready And Sampler.Ready 0.000 [min] Inject, Sampler.Inject

0.000 [min] Start Run, ColumnOven.ColumnOven_Temp.AcqOn, PumpModule.LoadingPump.LoadingPump_Pressure.AcqOn, PumpModule.NC_Pump.NC_Pump_Flow.AcqOn,PumpModule.NC_Pump.NC_Pump_Flow_LeftBlk. AcqOn, PumpModule.NC_Pump.NC_Pump_Flow_RightBlk.AcqOn, PumpModule.NC_Pump.NC_Pump_Pressure.AcqOn

0.000 [min] Run PumpModule.LoadingPump.Flow.Nominal: 3.000 [µl/min], PumpModule.LoadingPump.%B.Value: 0.0 [%] PumpModule.LoadingPump.%C.Value: 0.0 [%], PumpModule.LoadingPump.Curve: 5, PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 7.0 [%], PumpModule.NC_Pump.Curve: 5

10.000 [min] PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 7.0 [%], PumpModule.NC_Pump.Curve: 5, ColumnOven.ValveRight: 10_1

20.000 [min] PumpModule.LoadingPump.Flow.Nominal: 3.000 [µl/min], PumpModule.LoadingPump.%B.Value: 0.0 [%], PumpModule.LoadingPump.%C.Value: 0.0 [%], PumpModule.LoadingPump.Curve: 5

30.000 [min] PumpModule.LoadingPump.Flow.Nominal: 0.300 [µl/min], PumpModule.LoadingPump.%B.Value: 50.0 [%], PumpModule.LoadingPump.%C.Value: 0.0 [%], PumpModule.LoadingPump.Curve: 5

35.000 [min] PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 19.0 [%], PumpModule.NC_Pump.Curve: 5

50.000 [min] PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 20.0 [%], PumpModule.NC_Pump.Curve: 5

65.000 [min] PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 25.0 [%], PumpModule.NC_Pump.Curve: 5

80.000 [min] PumpModule.LoadingPump.Flow.Nominal: 0.300 [µl/min], PumpModule.LoadingPump.%B.Value: 25.0 [%], PumpModule.LoadingPump.%C.Value: 0.0 [%], PumpModule.LoadingPump.Curve: 5

86.000 [min] PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 95.0 [%], PumpModule.NC_Pump.Curve: 9

90.000 [min] PumpModule.LoadingPump.Flow.Nominal: 3.000 [µl/min], PumpModule. LoadingPump.%B.Value: 0.0 [%], PumpModule.LoadingPump.%C.Value: 0.0 [%], PumpModule.LoadingPump.Curve: 5

94.000 [min] PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule. NC_Pump.%B.Value: 95.0 [%], PumpModule. NC_Pump.Curve: 5

100.000 [min] PumpModule.NC_ Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 7.0 [%], PumpModule.NC_Pump.Curve: 5

120.000 [min] PumpModule.LoadingPump.Flow.Nominal: 3.000 [µl/min], PumpModule. LoadingPump.%B.Value: 0.0 [%], PumpModule.LoadingPump.%C.Value: 0.0 [%], PumpModule. LoadingPump.Curve: 5, PumpModule.NC_Pump.Flow.Nominal: 0.250 [µl/min], PumpModule.NC_Pump.%B.Value: 7.0 [%], PumpModule.NC_Pump.Curve: 5, ColumnOven.ValveRight: 1_2

121.000 [min] Stop Run

2. Ion sources

As each spectrum is acquired using only one ionization source, select the one that applies

2.1 Electrospray Ionization (ESI)

- Supply type (static or fed): Static.
- Interface manufacturer, model: EASY-Spray™ ES081
- Sprayer type, manufacturer, model: EASY spray" nano ion source (Thermo-Fisher Scientific, San Jose, CA, USA) and interfaced with an UltiMate 3000 RSLC system (Dionex, Sunnyvale, CA, USA).
- Other parameters if discriminant for the experiment: NA

2.2 MALDI

- Plate composition (or type): NA
- Matrix composition: NA
- PSD (or LID/ISD) summary, if performed: NA
- Laser type and wavelength: NA
- Other laser and source-related parameters, if discriminating for the experiment: NA

2.3 Other ionization source

- Description of the ion source and relevant parameters: NA

3. Post-source component

As an MS spectrum or chromatogram performed on one instrument cannot be acquired using all existing analysers and detectors, select the elements that apply.

3.1 Analysers

- Ion optics, 'simple' quadrupole, hexapole, Paul trap, linear trap, magnetic sector, FT- ICR, Orbitrap: name of the analysers(s): Orbitrap Fusion Tribrid
- Time-of-flight drift tube (TOF): Reflectron status: NA

3.2 Activation / dissociation

The associated acquisition parameters are covered in 4.1

- Instrument component where the activation/dissociation occurs: Ion-Routing Multipole/Dual Pressure Linear IonTrap
- Gas type (when used): Helium
- Activation/dissociation type: CID, HCD

4. Spectrum and peak list generation and annotation

4.1 Data acquisition

- Software name and version: Xcalibur 4.0.27.10

4.2 Data analysis

- Software name and version: Proteome Discoverer 2.2 (Thermo Fisher Scientific, San Jose, CA, USA)
- Parameters used in the generation of peak lists or processed spectra: mass tolerance of 10 ppm and 0.6 Da, two missed cleavages allowed, 0.01 FDR, cysteine carbamidomethylation as fixed modification, methionine oxidation and N-terminal acetylation as dynamic modification.

4.3 Resulting data

- Location of source (‚raw') and processed files: ProteomeXchange repository with ID PXD047172 and DOI 10.6019/PXD047172
- The chromatogram(s) for SRM data and other relevant cases: NA
- m/z and intensity values: 350–1500 m/z, intensity threshold of 5.0 e3.
- MS level: MS2
- Ion mode: positive ion mode
- For MS level 2 and higher, precursor m/z and charge, if known, with the full mass spectrum/peak list containing that precursor peak, where available: precursor selection mass range of 400–1200 m/z, precursor ion exclusion width of low 18 m/z and high 5 m/z, data will be available at ProteomeXchange repository with ID PXD047172 and DOI 10.6019/PXD047172