

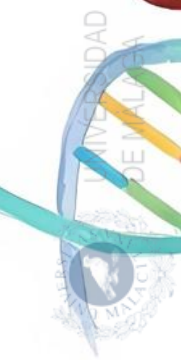
TESIS DOCTORAL

FUNCTIONAL ANALYSIS OF CANDIDATE GENES INVOLVED IN THE ORGANOLEPTIC AND NUTRITIONAL QUALITY OF STRAWBERRY FRUIT

**JOSÉ MORA
PERUJO**

Dra. Sonia Osorio Algar
Dr. José Gabriel Vallarino Castro

Programa de Doctorado en Biotecnología Avanzada,
Facultad de Ciencias, Universidad de Málaga, 2026



Functional Analysis of Candidate Genes Involved in the Organoleptic and Nutritional Quality of Strawberry Fruit

José Mora Perujo

Facultad de Ciencias. Universidad de Málaga



UNIVERSIDAD
DE MÁLAGA

AUTOR: Jose Mora Perujo.

 <https://orcid.org/0000-0002-6552-5970>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

Cualquier parte de esta obra se puede reproducir sin autorización pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): riuma.uma.es





DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña JOSÉ MORA PERUJO

Estudiante del programa de doctorado BIOTECNOLOGÍA AVANZADA de la Universidad de Málaga, autor/a de la tesis, presentada para la obtención del título de doctor por la Universidad de Málaga, titulada: FUNCTIONAL ANALYSIS OF CANDIDATE GENES INVOLVED IN THE ORGANOLEPTIC AND NUTRITIONAL QUALITY OF STRAWBERRY FRUIT

Realizada bajo la tutorización de SONIA OSORIO ALGAR y dirección de SONIA OSORIO ALGAR Y JOSÉ GABRIEL VALLARINO (si tuviera varios directores deberá hacer constar el nombre de todos)

DECLARO QUE:

La tesis presentada es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, conforme al ordenamiento jurídico vigente (Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), modificado por la Ley 2/2019, de 1 de marzo.

Igualmente asumo, ante a la Universidad de Málaga y ante cualquier otra instancia, la responsabilidad que pudiera derivarse en caso de plagio de contenidos en la tesis presentada, conforme al ordenamiento jurídico vigente.

En Málaga, a 16 de DICIEMBRE de 2025

Fdo.: JOSÉ MORA PERUJO Doctorando/a	Fdo.: SONIA OSORIO ALGAR Tutor/a
Fdo.: SONIA OSORIO ALGAR Y JOSÉ GABRIEL VALLARINO Director/es de tesis	





UNIVERSIDAD
DE MÁLAGA



Escuela de Doctorado

UNIVERSIDAD
DE MÁLAGA



EFQM  **AENOR**



Edificio Pabellón de Gobierno. Campus El Ejido.
29071
Tel.: 952 13 10 28 / 952 13 14 61 / 952 13 71 10
E-mail: doctorado@uma.es

Facultad de Ciencias

Departamento de Biología Molecular y Bioquímica

Dra. Sonia Osorio Algar, Catedrática del Departamento de Biología Molecular y Bioquímica de la Universidad de Málaga. IHSM (UMA-CSIC).

Dr. José Gabriel Vallarino Castro, Profesor Titular del Departamento de Biología Molecular y Bioquímica de la Universidad de Málaga. IHSM (UMA-CSIC).

INFORMAN:

Que D. José Mora Perujo ha realizado, bajo la dirección de ambos y supervisión de Dra. Sonia Osorio Algar como tutora, el trabajo de investigación que bajo el título “Functional analysis of candidate genes involved in the organoleptic and nutritional quality of strawberry fruit” constituye su Tesis Doctoral para aspirar a la obtención del grado de Doctor con mención internacional en Biotecnología Avanzada.

Así mismo, informan que la publicación con título “Regulation of plant tannin synthesis in crop species” en la que D. José Mora Perujo aparece como coautor y la cual avala esta tesis doctoral, no ha sido utilizada para avalar tesis doctorales anteriores y todos los coautores rechazan su uso en tesis doctorales posteriores.

Y, para que quede aquí constancia, y tenga los efectos oportunos, en cumplimiento de la legislación vigente, se extiende el presente informe

En Málaga, a 16 de diciembre de 2025

Fdo: Sonia Osorio Algar

Fdo: José Gabriel Vallarino Castro

*No podemos elegir los tiempos que nos toca
vivir, lo único que podemos hacer es decidir
qué hacer con el tiempo que se nos ha dado.*

J. R. R. Tolkien



UNIVERSIDAD
DE MÁLAGA

La ciencia no es perfecta, con frecuencia se utiliza mal, no es más que una herramienta, pero es la mejor herramienta que tenemos: se corrige a sí misma, está siempre evolucionando y se puede aplicar a todo. Con esta herramienta conquistamos lo imposible.

Carl Sagan



UNIVERSIDAD
DE MÁLAGA

I. Resumen	1
II. Aims and hypotheses	15
III. General introduction	19
1. Origins of strawberry species	21
1.1. <i>Rosaceae</i> family	21
1.2. <i>Fragaria</i> genus: characteristics and origin.....	21
2. <i>Fragaria</i> × <i>ananassa</i> genomic complexity.....	25
2.1. Octoploid strawberry genome	27
3. Strawberry fruit and quality traits	30
4. Volatile organic compounds in strawberry fruit.....	31
4.1. Functions of plant volatile organic compounds	32
4.2. Biosynthesis and function of terpenoids volatiles	33
4.3. Biosynthesis of shikimate derived volatiles	35
4.4. Fatty acid volatile-derivatives and lipoxygenase (LOX) biosynthetic pathway.....	36
5. Antioxidant compounds	37
6. Biosynthesis of polyphenols.....	39
IV. Material and methods.....	43
1. Plant material and growth conditions.....	45
1.1. Goodberry population	45
1.2. <i>Fragaria</i> × <i>ananassa</i> ‘232’ and ‘1392’ mapping population	45
1.3. <i>Fragaria</i> × <i>ananassa</i> cultivars (‘Camarosa’ and ‘Candonga’).....	46
1.4. <i>Nicotiana benthamiana</i> plants.....	46
2. Development of linkage maps and QTL analysis	46
2.1. Volatile organic compounds in Goodberry population (‘Candonga’ × ‘Senga Sengana’)	46
2.2. Ellagitannins in ‘232’ × ‘1392’ mapping population	47
3. Transcriptome analysis (RNA-seq).....	47
3.1. Goodberry population for contrasting γ-dodecalactone content	48
3.2. ‘232’ × ‘1392’ mapping population for contrasting ellagitannins content.....	48
4. RNA extraction, cDNA synthesis and gene expression analysis	48
5. Polymerase Chain Reaction (PCR) reactions.....	49
6. Competent Bacterial cells preparation	50
6.1. <i>Escherichia coli</i> TOP10 heat shock competent cells	50
6.2. <i>Agrobacterium tumefaciens</i> GV3101::pMP90 electrocompetent cells	50
6.3. <i>Agrobacterium tumefaciens</i> AGL1 heat shock competent cells.....	50
7. Searching and characterization of a candidate gene associated with γ-dodecalactone content	50

7.1. Identification of the candidate gene <i>FxaC_10g48980</i> (<i>FaLOX2</i>).....	50
7.2. Cloning of candidate gene <i>FaLOX2</i> and its tomato homolog <i>TomloxC</i>	51
7.3. Transient overexpression of ‘Candonga’ strawberry fruits	51
7.4. Generation of stable transgenic lines in <i>Solanum lycopersicum</i> cv. ‘MoneyMaker’ ...	52
7.5. Volatile organic compounds during ripening process	53
8. Searching and characterization of a candidate gene associated with ellagic acid hexose content	53
8.1. Identification of candidate gene <i>FxaC_3g11070</i> (<i>FaUDPSS</i>)	53
8.2. Cloning of <i>FxaC_3g11070</i> versions and sequencing	53
8.3. Generation of stable transgenic <i>Fragaria</i> × <i>ananassa</i> cv. ‘Candonga’ lines overexpressing <i>FaUDPSS</i>	54
8.4. TurboID proximity labelling of FaUDPSS and liquid-chromatography/mass spectrometry-mass spectrometry (LC-MS/MS)	54
9. Metabolomic analysis by mass spectrometry techniques	56
9.1. Volatile organic compounds (VOCs) extraction and analysis by Headspace-Gas Chromatography/Mass Spectrometry (HS-GC/MS)	56
9.2. Extraction, derivatization and analysis of polar primary metabolites using Gas-Chromatography/Mass Spectrometry (GC/MS).....	59
9.3. Extraction and analysis of semi-polar secondary metabolites using Ultra Performance Liquid Chromatography (UPLC)-Orbitrap Mass Spectrometry (LC/MS)	62
10. Genetic sequences retrieval and phylogenetic analysis	64
10.1. <i>In silico</i> structural data analysis of proteins	65
11. Subcellular localization of FaLOX2 and FaUDPSS by confocal microscopy imaging	65
12. Protein extraction and immunodetection assay	66
13. Genetic resources used in this work	67
14. Statistical analysis and R packages used in this study	69

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit 71

Introduction 73

1. Jasmonic acid biosynthesis and signaling pathways	73
2. Lipoxygenases: activity, occurrence and regulation	78
3. Strawberry fruit ripening process: abscisic acid (ABA) metabolism and its genetic control	80
4. Lactones: biological occurrence and importance	87

Results 89

1. Volatile content in Goodberry population.....	89
2. Quantitative Trait <i>Loci</i> (QTL) identification for volatile organic compounds.....	92
3. Identification of the candidate genes for γ -dodecalactone content in the stable mQTLs	103

4. Expression pattern of <i>FaLOX2</i> in strawberry.....	107
5. Phylogenetic analysis of <i>FaLOX2</i>	108
6. Structural analysis of <i>FaLOX2</i> protein	110
7. Transient overexpression of <i>FaLOX2</i> induces changes in strawberry volatilome	114
8. Primary and secondary metabolism changes in <i>FaLOX2</i> -overexpression.....	122
9. <i>FaLOX2</i> expression and VOCs accumulation during ripening process.....	126
10. <i>FaLOX2</i> has a cytoplasmic localization.....	130
Discussion.....	133
1. Volatile content diversity on Goodberry population.....	133
2. QTL for volatiles reveals a possible genetic hotspot in Fvb3-2.....	134
3. <i>FaLOX2</i> is responsible for the accumulation of the volatile γ -dodecalactone in strawberry fruits	135
4. <i>FaLOX2</i> is a cytoplasmic protein with involvement in jasmonate metabolism.....	137
5. Future perspectives and proposed experiments.....	138
6. Concluding remarks	139
VI. Chapter II: Characterization of a candidate gene involved in ellagitannins content in strawberry fruit	141
Introduction	143
1. Tannins and their properties: antioxidants and relationship with human health and organoleptic characteristics.....	143
2. Proanthocyanidins or condensed tannins	145
3. Hydrolysable tannins: ellagitannins and gallotannins	151
4. UDP-glycosyltransferases enzymes	155
Results	157
1. Variation of ellagitannins content in ‘232’ $\times$ ‘1392’ mapping population.....	157
2. QTLs controlling ellagitannins content in ‘232’ $\times$ ‘1392’ F ₁ population.....	159
3. Transcriptomic analysis of contrasting lines for ellagitannins content using RNA-seq	164
4. Identification of <i>FxaC_3g11070</i> as candidate for controlling ellagic acid hexose content	165
5. Discrepancies in gene sequence of <i>FxaC_3g11070</i>	173
7. Phylogenetic analysis and protein prediction of <i>FaUDPSS</i>	177
8. Subcellular localization of <i>FaUDPSS</i> protein	184
9. TurboID proximity labelling of proteins putatively interacting with <i>FaUDPSS</i>	186
Discussion.....	193
1. Ellagitannins in ‘232’ $\times$ ‘1392’ mapping population and RNA-seq analysis.....	193
2. Direct correlation between <i>FaUDPSS</i> and EAH2 content.....	193
3. <i>FxaC_3g11070</i> sequence differs from the ‘Camarosa’ reference genome sequence....	194
4. Strawberry commercial varieties and their relationship with EAH2 content.....	195

5. FaUDPSS is a very conserved cytoplasmic protein	196
6. Proteins interacting with FaUDPSS	197
7. Future perspectives and proposed experiments	198
8. Concluding remarks	199
VII. Conclusions	201
VIII. Bibliography	205
IX. Abbreviations	249
X. Annexes	253

I. Resumen



Botanical illustration of *Fragaria* × *ananassa* subsp. 'Bath' plant by Antoine Nicolas Duchesne. *Histoire Naturelle des Fraisiers*, 1759.

El reglamento de la Universidad de Málaga establece que es necesario un resumen de la tesis doctoral de al menos tres mil palabras para optar a la mención de “Doctorado internacional” de la misma, al no estar dicha tesis doctoral escrita en castellano. Por tanto, a continuación, se incluye el resumen requerido.

La presente tesis doctoral se divide en dos capítulos que recogen los resultados de dos líneas de investigación referentes a las características organolépticas del fruto de fresa. El primero de estos capítulos trata sobre la identificación de “*Quantitative Traits Loci*” (QTLs) o *loci* de caracteres cuantitativos en español, asociados a compuestos orgánicos volátiles. Este primer capítulo incluye además la caracterización de un gen candidato como principal responsable de la concentración del volátil γ -dodecalactona, importante para el aroma, en el fruto de fresa. Por otra parte, el segundo capítulo pivota entorno a la identificación de QTLs que controlan el contenido de una familia de compuestos del metabolismo secundario denominados elagitaninos o taninos solubles, reconocidos por su actividad antioxidante. De manera paralela a lo descrito para el primer capítulo, hemos incluido la caracterización de un gen candidato importante en el metabolismo de uno de los isómeros de la hexosa del ácido elágico en el fruto de fresa.

Introducción general

La fresa es una especie perteneciente al género *Fragaria*, que, a su vez, pertenece a la familia *Rosaceae*. Esta familia está compuesta por aproximadamente 100 géneros y 3.000 especies entre las cuales se pueden localizar herbáceas, ornamentales, leñosas y frutales. Muchas especies pertenecientes a esta familia han sido domesticadas a lo largo de los años para el consumo humano y entre las cuales se pueden destacar algunos frutos de interés comercial como son la fresa, frambuesa, peras, almendras, cerezas, etc.

El género *Fragaria* es conocido por el consumo de sus frutos, conocidos como fresas. Aunque, desde un punto de vista estrictamente botánico, las fresas no son frutos, sino que los verdaderos frutos son los aquenios que están embebidos en el receptáculo, la parte comestible. El receptáculo proviene del engrosamiento del receptáculo floral, siendo esta una estrategia evolutiva compartida por muchas especies dentro de este género. En este género se engloban algunas especies que comparten características tales como que son polinizadas por insectos, son plantas herbáceas perennes con hojas trifoliadas y son capaces de reproducirse asexualmente mediante estolones.

Las especies del género *Fragaria* están distribuidas principalmente en el hemisferio norte. Filogenéticamente hablando, existen dos grandes grupos con respecto al género *Fragaria*: clado vesca y clado China. En el primero de los clados se engloban especies con distintos niveles de ploidía: desde diploides (*F. vesca*, *F. bucharica* y *F. mandshurica*) hasta decaploides (*F. iturupensis*). Sin embargo, en el clado China todas las especies englobadas son diploides (*F. chinensis*) o tetraploides (*F. corymbosa*).

Históricamente se hipotetiza que la fresa cultivada (*F. × ananassa*) tiene su origen en Francia, fruto de la hibridación espontánea de dos parentales octoploides: *F. chiloensis* y *F. virginiana*. Así pues, la fresa comercial sería una especie alo-octoploide generado por una hibridación natural de las especies parentales *F. chiloensis* y *F. virginiana*. Debido a esta característica, el genoma de la fresa cultivada presenta una gran complejidad a la hora de dilucidar verdaderamente la contribución de otras especies del género *Fragaria* en la estructura del genoma de la fresa cultivada. En un principio, en 2019, se publicó una primera secuenciación del genoma de la fresa cultivada en la que se descubrió que las especies diploides de las cuales proviene dicho genoma eran *F. vesca*, *F. iinumae*, *F. nipponica* y *F. viridis* (Edger *et al.*, 2019). Sin embargo, estudios recientes han ofrecido evidencia para establecer que, coincidiendo con estudios previos, *F. vesca* y *F. iinumae* serían progenitores de dos de los subgenomas de *F. × ananassa* mientras que no existe evidencia consistente para afirmar que *F. viridis* y *F. nipponica* sean los progenitores de los otros dos subgenomas (Jin *et al.*, 2023). De hecho, actualmente se piensa que estos dos últimos subgenomas pudieron ser originados tras un evento de poliploidía entre *F. iinumae* y una especie cercana ya extinta.

El fruto de fresa se trata de un fruto muy valorado por el consumidor gracias a diversas características. Concretamente, las características organolépticas son responsables en gran parte del consumo mundial de la fresa, de entre las cuales se destacan el color, el aroma, el dulzor, la acidez, la firmeza y el contenido de compuestos antioxidantes. Respecto al aroma de este fruto, podemos decir que se trata de una combinación compleja de compuestos orgánicos volátiles (COVs). Estos compuestos presentan una baja polaridad, una masa molecular por debajo de 300 Da y una baja presión de vapor (Dudareva *et al.*, 2013). En cuanto al origen metabólico de estos COVs de fresa, podemos establecer que derivan del metabolismo secundario, pudiendo ser clasificados principalmente en cuatro grupos dependiendo de su biosíntesis: ruta de los fenilpropanoides/bencenoides, ruta de los derivados de ácidos grasos poliinsaturados, ruta de los derivados de aminoácidos y ruta de los terpenoides. La emisión de estos volátiles está ampliamente regulada y entre sus funciones biológicas en plantas destacan la atracción de insectos polinizadores y dispersores de semillas además de actuar en la defensa contra herbívoros y patógenos gracias a sus características antimicrobianas y antifúngicas.

Los compuestos volátiles derivados de la ruta de los terpenoides en el fruto de fresa son particularmente interesantes porque aportan aromas florales y cítricos, siendo atributos muy apreciados por los consumidores. Los terpenoides, también llamados isoprenoides, son una gran familia de compuestos que derivan del ensamblaje de unidades de isopreno (C_5H_8). En la fresa comercial, los compuestos más importantes para el aroma son los monoterpenos (C_{10}) y sesquiterpenos (C_{15}). Respecto a su biosíntesis en fresa, se han

descrito dos rutas principales para sintetizar los precursores de los terpenoides: ruta del mevalonato (MVA) que produce principalmente sesquiterpenos y ruta del metileritritol fosfato (PEP) que es la responsable de la mayoría de los monoterpenos. Ambas rutas producen isopentenil pirofosfato (IPP) y su isómero dimetilalil pirofosfato (DMAPP), los cuales son después ensamblados por las enzimas terpeno sintasas (TPS) para formar los compuestos volátiles (Dudareva *et al.*, 2013). Los terpenoides volátiles son conocidos por tener funciones como atractor de insectos polinizadores y diseminadores de semillas y comunicación planta-planta y planta-insecto. Además, se hipotetiza que puedan tener un posible rol en el proceso de fotosíntesis mediante la regulación del crecimiento de la propia planta (Loreto *et al.*, 2014).

Por otro lado, los compuestos volátiles derivados de ácidos grasos poliinsaturados provienen, como su nombre indica, de ácidos grasos insaturados de 18 átomos de carbono, siendo su síntesis restringida a los cloroplastos. De hecho, estos compuestos volátiles tienen como precursores, en la gran mayoría de los casos, al ácido α -linolénico o bien al ácido linoléico para producir diversos tipos de compuestos volátiles que provienen principalmente de la ruta de la lipoxigenasa (LOX). Estos volátiles también presentan varias funciones biológicas como pueden ser la defensa contra herbívoros y patógenos por sus conocidas actividades antimicrobianas y repelentes, además de actuar como señales de daño de tejidos y comunicación entre plantas mediante los denominados volátiles de estrés.

Los aminoácidos libres presentes en el fruto durante su desarrollo y maduración son transformados en compuestos volátiles por una serie de reacciones enzimáticas como transaminación, descarboxilación, reducción y esterificación. Estos compuestos son, por ejemplo, metil butanoato, etanoato 2-metilbutilo, alcohol bencílico, feniletil alcohol, eugenol, todos ellos destacables con tener aroma a dulce y floral. Estos aromas dulces y florales sirven de nuevo como una señal atrayente de insectos polinizadores y dispersores, participando en la comunicación planta-insecto; pero, también destacan por ser moléculas reguladoras del propio metabolismo secundario de la planta.

Respecto a los volátiles que derivan de la ruta de los fenilpropanoides podemos decir que es una vía metabólica secundaria que parte del aminoácido fenilalanina (sintetizado a través de la ruta del shikimato) produciendo una gran variedad de compuestos fenólicos, siendo muy importante en la defensa vegetal y atracción de polinizadores.

El fruto de fresa no es solo valorado por su sabor y aroma sino también por su alto contenido nutricional, especialmente por su riqueza en la variedad de compuestos fenólicos con actividad antioxidante, entre los cuales destacan los elagitaninos y otros tipos de taninos, siendo importantes no solo para la salud humana sino también en la defensa de la planta. Los elagitaninos son un tipo de taninos hidrolizados formados por

unidades de ácido elágico unidas a una molécula central de naturaleza monosacárido como puede ser la glucosa. Estos compuestos tienen múltiples propiedades bioactivas con beneficios para la salud humana como puede ser unas potentes actividades antioxidante, antiinflamatoria, cardioprotectora, potencialmente anticancerígena y prebiótica. Respecto a su biosíntesis, derivan de la combinación de la ruta del shikimato y la ruta de los fenilpropanoides. Las moléculas clave para la biosíntesis de los elagitaninos son tanto ácido gálico como el ácido hexahidroxidifénico (HHDP)-glucosa. En concreto, ácido gálico procedente de la ruta del shikimato se esterifica con unidades de glucosa para formar galotaninos. Posteriormente, mediante polimerización oxidativa y lactonización se forma HHDP-glucosa (hexahidroxidifenil- β -D-glucosa). Este HHDP bien se transforma en ácido elágico o bien se une en estructuras más complejas formando los elagitaninos.

Capítulo I: Caracterización de un gen candidato involucrado en el contenido del volátil γ -dodecalactona en fruto de fresa

El primer capítulo de esta Tesis Doctoral estriba sobre el estudio de los *loci* de caracteres cuantitativos o QTLs referentes a compuestos volátiles. La recolección de los datos y el análisis de QTLs para compuestos volátiles fue desarrollado en el mismo grupo de investigación previo al comienzo de esta Tesis Doctoral por la Dra. Delphine Pott.

En este estudio se trabajó con una población de *F. × ananassa* proveniente del proyecto europeo denominado Goodberry (proyecto europeo- H2020). Esta población se caracteriza por ser una F_1 (compuesta por 126 individuos) derivada del cruzamiento entre las variedades comerciales ‘Candonge’ (cultivar desarrollado en España, por tanto, estable en un ambiente mediterráneo y caracterizado por tener frutos grandes y de buena calidad) y ‘Senga Sengana’ (cultivar desarrollado en Alemania durante los años 60, estable en ambiente de entro Europa, cuyos frutos se caracterizan por un tamaño medio y de buena calidad). De este cruzamiento trabajamos con 126 líneas F_1 , las cuales se crecieron en cinco localizaciones distintas: Francia, Alemania, Italia, Polonia y España. Los frutos maduros de estas líneas F_1 fueron sujetos de un análisis metabolómico para la composición de volátiles, comprobando que el contenido de volátiles variaba dependiendo de la localización en una gran parte de los casos. Mediante un análisis de componentes principales de la variación de estos volátiles se comprobó que no existía suficiente variabilidad para poder discriminar entre las localizaciones. Sin embargo, sí se encontraba suficiente variabilidad como para poder realizar un análisis de QTLs. De hecho, con estos datos también se calculó la heredabilidad de dichos volátiles en la población de mapeo, observando que la gran parte de estos compuestos presentaban unos valores superiores a 0.5 lo que indica que su variabilidad responde en su mayoría ante un control genético y no tanto asociado al ambiente.

De este análisis de QTLs en esta población crecida en diferentes ambientes de Europa, se encontraron una serie de QTLs que denominamos como “estables” por estar presentes en al menos dos localizaciones para el mismo compuesto volátil. La gran mayoría de estos QTLs, aproximadamente un 75% fueron encontrados usando el mapa del parental femenino. Un resultado interesante fue que dentro de este 75% de QTLs encontramos aproximadamente un 67% de todos estos mapearon en el mismo grupo de ligamiento, concretamente en el Fvb3-2. Esto nos hizo pensar que el cromosoma al cual corresponde este grupo de ligamiento puede presentar un clúster de regulación para compuestos volátiles o bien que existen una serie de factores de transcripción o genes regulados de una misma forma y que contribuyen a la síntesis de una gran familia de compuestos volátiles.

Entre los compuestos volátiles analizados en esta población, destacamos una familia de compuestos llamada lactonas; que, aunque no están presentes en grandes cantidades en el fruto de fresa, tienen una importancia sensorial muy significativa principalmente en el fruto maduro. Estos compuestos aportan un aroma dulce y afrutado, a veces descrito con notas de coco y melocotón que refuerzan el aroma afrutado y tropical, y poseen umbrales de detección olfativa muy bajos. En este estudio seleccionamos dos QTLs asociados al volátil γ -dodecalactona que explicaban en torno al 20% de la variación total de este compuesto. Estos dos QTLs fueron encontrados tanto en la localización de Francia como en España. Gracias a datos de secuenciación de los individuos derivados del proyecto Goodberry, pudimos observar que existían marcadores moleculares asociados con el contenido del volátil de interés. De este modo pudimos seleccionar individuos de la población dependiendo del contenido de γ -dodecalactona en dos grupos: uno de bajo y otro de alto contenido de este volátil. Gracias al análisis transcriptómico mediante RNA-seq de estos individuos pudimos encontrar genes diferencialmente expresados en estos grupos, además de conocer cómo variaban la expresión de estos en condiciones ambientales diferentes, puesto que se analizaron las muestras procedentes de plantas crecidas en España y en Francia. Con estos datos estudiamos la expresión de los genes comprendidos en el intervalo de ambos QTLs y, además, hicimos un análisis de anotación funcional de dichos genes. El análisis de anotación funcional reveló que una gran mayoría de los genes presentaban una función “no asignada”, de modo que tuvimos que tomar un nuevo enfoque.

Siguiendo la hipótesis más probable de que la síntesis de estos compuestos (lactonas) provendría de la ruta de los derivados de ácidos grasos poliinsaturados (Dudareva *et al.*, 2013), nos centramos en los genes que presentaban cambios de expresión entre los dos grupos contratantes y que estuvieran relacionados con el metabolismo de lípidos. Los datos de expresión revelaron que tan solo el gen *FxaC_10g48980* presentaba diferencias significativas entre ambos grupos y en ambas localizaciones. La anotación funcional de

este gen reveló que la proteína para la cual codifica es una linoleato 13S-lipoxigenasa, gen que renombramos de aquí en adelante como *FaLOX2*.

Llevamos a cabo un estudio filogenético sobre dicha proteína, para el cual usamos secuencias de proteínas con función lipoxigenasa de distintas especies como fueron *Arabidopsis thaliana*, *Cucumis melo*, *F. vesca*, *Malus x domestica*, *Prunus persica* o *Solanum lycopersicum*. De este análisis pudimos inferir que la proteína de interés se agrupaba en el clado de las lipoxigenasas 13S y que se agrupaba filogenéticamente cerca de FvLOX3 y FaLOX3, además del homólogo de tomate TomLOXC. Análisis estructurales *in silico* nos indicaron que la estructura tridimensional de la proteína FaLOX2 estaba mayoritariamente compuesta por α -hélices, presentando tan solo dos regiones configuradas como β -láminas. La predicción de la estructura tridimensional realizada mediante AlphaFold3 arrojó una estructura tridimensional de la proteína mayormente globular con una superficie altamente hidrofílica y que presentaba una coordinación con un átomo de Fe, que se estabilizaba mediante la coordinación con cinco aminoácidos. Cuando se comparó dicha estructura frente a las estructuras de aquellas proteínas con las que presenta mayor cercanía filogenética se pudo comprobar que la proteína candidata carecía de dos regiones que estaban presente en muchas de las proteínas estudiadas: un posible péptido señal y una estructura de anclaje a membrana. Nuestra hipótesis es que, debido a esto, la proteína candidata presenta una localización citoplasmática mientras que aquellas que presentan las regiones anteriormente mencionadas se localizaban en cloroplastos. Adicionalmente, modelamos dicha proteína con un posible sustrato de la enzima: el ácido α -linoléico. Obtuvimos una predicción con una alta fiabilidad y donde el sustrato parecía estar cerca del átomo de Fe y del centro catalítico, permitiendo la oxidación de este compuesto y cumpliendo las predicciones de la funcionalidad catalítica de este tipo de enzimas.

Posteriormente hicimos un ensayo de localización subcelular, resultando ser una proteína localizada libre en el citoplasma. Para realizar este experimento, fusionamos dicha proteína a un fluoróforo como es GFP (*green fluorescent protein*) en el extremo carboxilo-terminal, evitando el posible centro catalítico que se encontraría en el extremo amino-terminal. De esta forma, mediante un vector de expresión, introdujimos la secuencia *FaLOX2-GFP* bajo el control de un promotor de expresión constitutiva como es 35S*CaMV* en la cepa GV3101::pMP90 de *Agrobacterium tumefaciens*. Así pues, pudimos agroinfiltrar hojas de *Nicotiana benthamiana* con esta construcción, además de emplear una construcción que tan solo expresaba GFP libre a modo de control y con una construcción basada en RFP (*red fluorescent protein*) que empleamos como control para la localización en el retículo endoplasmático. De este ensayo pudimos concluir que nuestra proteína de estudio se acumulaba en el citoplasma de las células, lo cual fue de nuevo comprobado mediante inmunodetección por *western blot*.

Una de las hipótesis que barajamos respecto a la funcionalidad de la proteína FaLOX2, es su posible papel en la ruta del ácido jasmónico ya que existe una conexión indirecta con el volátil γ -dodecalactona puesto que ambos compuestos derivan del metabolismo de los ácidos grasos poliinsaturados, especialmente ácido linoleico y ácido linolénico. De esta forma que ambas rutas comparten pasos iniciales comunes. Para explorar esta hipótesis, generamos una construcción de sobreexpresión del gen *FaLOX2* bajo el control del promotor 35SCaMV basado en el plásmido pK7GW2D. Esta construcción la agroinfiltramos en frutos *F. × ananassa* variedad ‘Candonga’, usando el propio pK7GW2D vacío como control negativo. Tras la maduración de los frutos pudimos comprobar por RT-qPCR que la expresión de *FaLOX2* era significativamente superior en los frutos en los que hicimos sobreexpresión transitoria en comparación con los frutos control. El siguiente paso fue hacer un análisis del contenido de volátiles tanto de los frutos que sobreexpresaban *FaLOX2* como de los frutos controles. Mediante este análisis pudimos observar que los frutos que sobreexpresaban *FaLOX2* poseían un mayor contenido de volátiles en comparación con los frutos control. De entre estos volátiles con mayor acumulación podían destacarse mesifurano, algunos terpenoides, pero sobre todo metil jasmonato y γ -dodecalactona. Sin embargo, otro grupo de COVs presentaban una relación inversamente proporcional con la expresión del gen *FaLOX2*, de entre los cuales se podía destacar heptanal, penten-3-ona, pentanal, *trans*-2-pentenal, etc.

Con estas mismas muestras estudiamos también tanto el metabolismo primario como el secundario. Respecto al metabolismo primario pudimos observar que los frutos con mayor expresión de *FaLOX2* acumulaban mayor cantidad de azúcares como glucosa, fructosa, sacarosa o rafinosa mientras que tenían menor contenido en compuestos como rhamnosa o putrescina. Por su parte, respecto al metabolismo secundario, la expresión de *FaLOX2* conllevaba una mayor concentración de compuestos tales como antocianinas que confieren color al fruto de fresa. Principalmente pensamos que los cambios producidos en el metabolismo primario y secundario se debían al efecto del aumento de metil jasmonato y a la señalización por esta molécula, aunque necesitaríamos de otras evidencias experimentales para poder concluirlo.

Con el objetivo de comprobar como cambian la concentración de los volátiles que veíamos alterados en los frutos con la expresión alterada del gen *FaLOX2* a lo largo de la maduración de fresa, llevamos a cabo el análisis de volátiles en estadios de maduración (verde, blanco y rojo) de fruto de fresa variedad ‘Camarosa’. Con este ensayo, pudimos comprobar que muchos compuestos volátiles aparecían durante el estadio blanco y llegaban a su máximo de acumulación en el estadio rojo, de entre ellos algunos que veían incrementada su concentración con la sobreexpresión de *FaLOX2*. De esta forma pudimos comprobar que, posiblemente, alguno de los efectos observados tras la sobreexpresión del candidato podría estar relacionados con un proceso de maduración acelerada

provocado por el aumento de metil jasmonato. Sin embargo, otras evidencias experimentales serían necesarias para poder comprobar esta hipótesis.

Capítulo II: Caracterización de un gen candidato involucrado en la síntesis de elagitaninos en el fruto de fresa

Este segundo capítulo trata sobre el estudio de QTLs para una familia de compuestos llamados elagitaninos o taninos solubles, además de la caracterización de un gen candidato para el contenido de una hexosa del ácido elágico. Para este cometido se usó una población de mapeo F_1 derivada del cruce entre dos variedades de selección de *F. × ananassa*: ‘232’ y ‘1392’. Este trabajo, al igual que el anterior análisis de QTLs, fue desarrollado en el grupo de investigación al que pertenezco con anterioridad al comienzo de esta Tesis Doctoral aunque los resultados derivados no fueron presentados hasta el momento.

Para dicho análisis se estudió el contenido en elagitaninos en fruto maduro de la población de fresa constituida por 95 líneas F_1 , durante dos años consecutivos. Se realizó un análisis de QTLs, el cual resultó con la identificación de 76 QTLs detectados para elagitaninos. De entre estos seleccionamos aquellos QTLs que explicaban en torno al 60% de la variación total en un isómero concreto la hexosa del ácido elágico. En paralelo y, al igual que hicimos en el **Capítulo I**, seleccionamos diez líneas contrastantes en el contenido de este isómero de la hexosa del ácido elágico para los QTLs qEAhex1 and qEAhex2 que explican una variación total del 70% en los niveles de este compuesto. Posteriormente, se realizó un análisis transcriptómico mediante RNA-seq en los frutos maduros de estas líneas contrastantes en el contenido de la hexosa del ácido elágico. De este análisis, se identificaron 27 genes diferencialmente expresados entre ambos grupos contrastantes localizados dentro de la región del QTL. En concreto, observamos que los genes *FvH4_1g14013* y *FvH4_1g13990* presentaban una alta tasa de cambio al comparar las líneas contrastantes en el contenido de la hexosa del ácido elágico. Curiosamente cuando hicimos la anotación funcional de estos genes candidatos, observamos que las proteínas para las que codificaban ambos genes candidatos se identificaban con una función de antocianidina-3-*O*-glucosiltransferasa. En este análisis, los genes estaban mapeados con el genoma de *F. vesca*, de modo que posteriormente hicimos BLAST frente al genoma de *F. × ananassa* cultivar ‘Camarosa’ revelando varios posibles homólogos para cada uno de los genes candidatos. En concreto, los genes identificados en el genoma de la fresa cultivada fueron *FxaC_3g15640*, *FxaC_1g15670* y *FxaC_3g11070*, los cuales presentaban diferencias significativas de expresión entre las líneas contrastantes en el contenido de la hexosa del ácido elágico. Finalmente, seleccionamos el gen *FxaC_3g11070* como candidato para posteriores análisis ya que presentaba un comportamiento más extremo en cuanto a los niveles de expresión al comparar las líneas

contrastantes, observando que cuando la concentración de la hexosa del ácido elágico era baja, la expresión de este gen era nula.

Tratando de clonar la secuencia codificante del gen *FxaC_3g11070* (renombrado como *FaUDP*) descubrimos que dicho gen no presentaba expresión en el cultivar ‘Camarosa’. Además, intentamos clonar el gen desde tejido de la población F_1 usada para el mapeo de los QTLs, encontrando una divergencia de secuencias entre distintas líneas. Entre la diversidad de secuencias pudimos encontrar una secuencia que presentaba una inserción idéntica a una secuencia intrónica presente en los parentales ‘232’ y ‘1392’ que provocaba un codón STOP prematuro; también encontramos, además, variantes del gen que presentaban una secuencia ligeramente más grande que la anotada en el genoma de referencia. Es por todo esto que nos propusimos explorar la secuencia de este gen de otras variedades de *F. × ananassa* como podrían ser ‘Camarosa’ y ‘Senga Sengana’ y la población derivada del cruce entre ambas (proyecto europeo Goodberry). De esta forma comprobamos que entre los individuos F_1 la secuencia de este gen codificaba una proteína ligeramente más larga que la anotada en la base de datos de GDR para dicho gen. De este modo, nuestra hipótesis fue que la secuencia de dicho gen estaba mal anotada en el genoma de referencia.

Además, aprovechando datos de expresión en fruto maduro de diferentes cultivares de fresa, comprobamos que este gen no presentaba expresión en las variedades de ‘Camarosa’, ‘Candongá’, ‘Clery’ y ‘Frida’, mientras que otras variedades sí presentaban expresión de dicho gen como ‘Florence’, ‘Gariguette’, ‘Senga Sengana’ y ‘Sonata’. Para posteriores análisis, tomamos como secuencia de referencia del gen *FaUDP*, la secuencia de la variedad ‘Senga Sengana’ y renombrando finalmente el gen candidato como *FaUDPSS*. Con esta información estudiamos la expresión este gen en la población F_1 derivada del cruce entre ‘Candongá’ y ‘Senga Sengana’, pudiendo clasificar a los individuos en dos grupos: uno con baja expresión del gen candidato y otro con alta expresión. Seleccionamos cinco individuos de cada una de las condiciones y analizamos el metabolismo secundario del fruto maduro. De todo el set de metabolitos secundarios estudiados, tan solo el isómero 2 de la hexosa del ácido elágico presentaba cambios significativos cuando comparábamos los dos grupos de individuos con expresión contrastante del gen *FaUDPSS*.

Hicimos un estudio filogenético de la proteína FaUDPSS junto con muchas otras proteínas anotadas como glucosiltransferasas de distintas especies vegetales (*Arabidopsis thaliana*, *Prunus dulcis*, *Fragaria vesca*, *Diospyros kaki*, etc.) y descubrimos una amplia familia filogenética con una compleja historia evolutiva en la que muchas proteínas han ido evolucionando desde ancestros comunes. En el caso de nuestra proteína FaUDPSS, observamos que se agrupaba filogenéticamente cercana a otras proteínas como FaUGT1,

FvUGT2 e incluso AtUGT72B1. Con esta información hicimos un modelado de la proteína usando AlphaFold3, observando que FUDPSS tenía una estructura tridimensional mayormente compuesta por α -hélices, pero con algunas zonas dispuestas en forma de β -láminas que están descritas como zonas de interacción con nucleótidos. También pudimos comprobar que esta proteína presentaba una superficie polar con alta tendencia hidrofílica, lo cual concordaba con los datos de predicción de localización subcelular, puesto que esta debía encontrarse en el citosol. Además, vimos que las proteínas relacionadas filogenéticamente presentaban una estructura muy similar, casi idéntica en la mayoría de los casos, por lo que hipotetizamos que las diferencias estructurales entre estas proteínas podrían deberse a diferente especificidad de catálisis al emplear diferentes sustratos. Para esta proteína también hicimos un modelado utilizando unos putativos sustratos como son el ácido elágico y la UDP-glucosa como donador de monosacáridos para la glucosilación. De la visualización de estos resultados concluimos un buen acoplamiento de ambos sustratos en la enzima y con una alta fiabilidad, constituyendo un esbozo de la posible reacción catalítica para formar la hexosa del ácido elágico.

Para estudiar la localización subcelular de la proteína FaUDPSS generamos una construcción en la que fusionamos esta proteína con GFP en el extremo amino-terminal bajo el control de un promotor *35SCaMV*, evitando de esta forma el centro catalítico que parecía encontrarse en el extremo carboxilo terminal. De este análisis pudimos concluir que esta proteína se encontraba localizada en el citosol, tal y como las predicciones bioinformáticas habían señalado.

Finalmente, quisimos estudiar el contexto de interacción de la proteína FaUDPSS para lo cual seguimos una aproximación de etiquetado por proximidad de TurboID en el cual se detectan proteínas biotiniladas que interaccionen con la proteína de interés en un radio de 10 nm. Para ello generamos la construcción TurboID-FaUDPSS y GFP-TurboID-FaUDPSS, esta última como control para comprobar que localización subcelular de la proteína no cambiaba al fusionarla a TurboID. Mediante expresión transitoria en fruto de fresa variedad ‘Camarosa’ de ambas construcciones descubrimos que la proteína de interés fusionada a TurboID seguía acumulándose en el citoplasma.

Ambas construcciones usadas para el ensayo de TurboID presentaban epítomos de hemaglutinina (HA) para su detección tras la agroinfiltración en fruto. Por tanto, comprobamos la presencia en los frutos mediante inmunodetección frente al epítipo de HA. Por su parte, la detección de biotina tanto libre como unida a otras proteínas se llevaba a cabo mediante detección con streptavidina fusionada con peroxidasa de rábano (HRP). Con esta comprobación descubrimos frutos que presentaban la construcción deseada pero también revelamos la presencia intrínseca de proteínas biotiniladas en el

fruto de fresa. Debido a esto, no podíamos asegurar que nuestra construcción fuese capaz de añadir biotina a las proteínas circundantes, de modo que infiltramos hojas de *N. benthamiana* con la construcción TurboID-FaUDPSS y añadimos biotina. Esto dio como resultado que, en presencia de biotina, la construcción era capaz de marcar proteínas circundantes. Con este resultado, analizamos las posibles proteínas marcadas en el fruto de fresa para revelar que una gran mayoría de las proteínas putativamente interactoras de FaUDPSS pertenecían al procesamiento de RNA o de proteínas, pero algunas de ellas tenían función en el metabolismo primario o secundario tales como acetil-CoA carboxilasa, BAHD-acetiltransferasa o una flavoproteína del complejo de succinato deshidrogenasa.

II. Aims and hypotheses



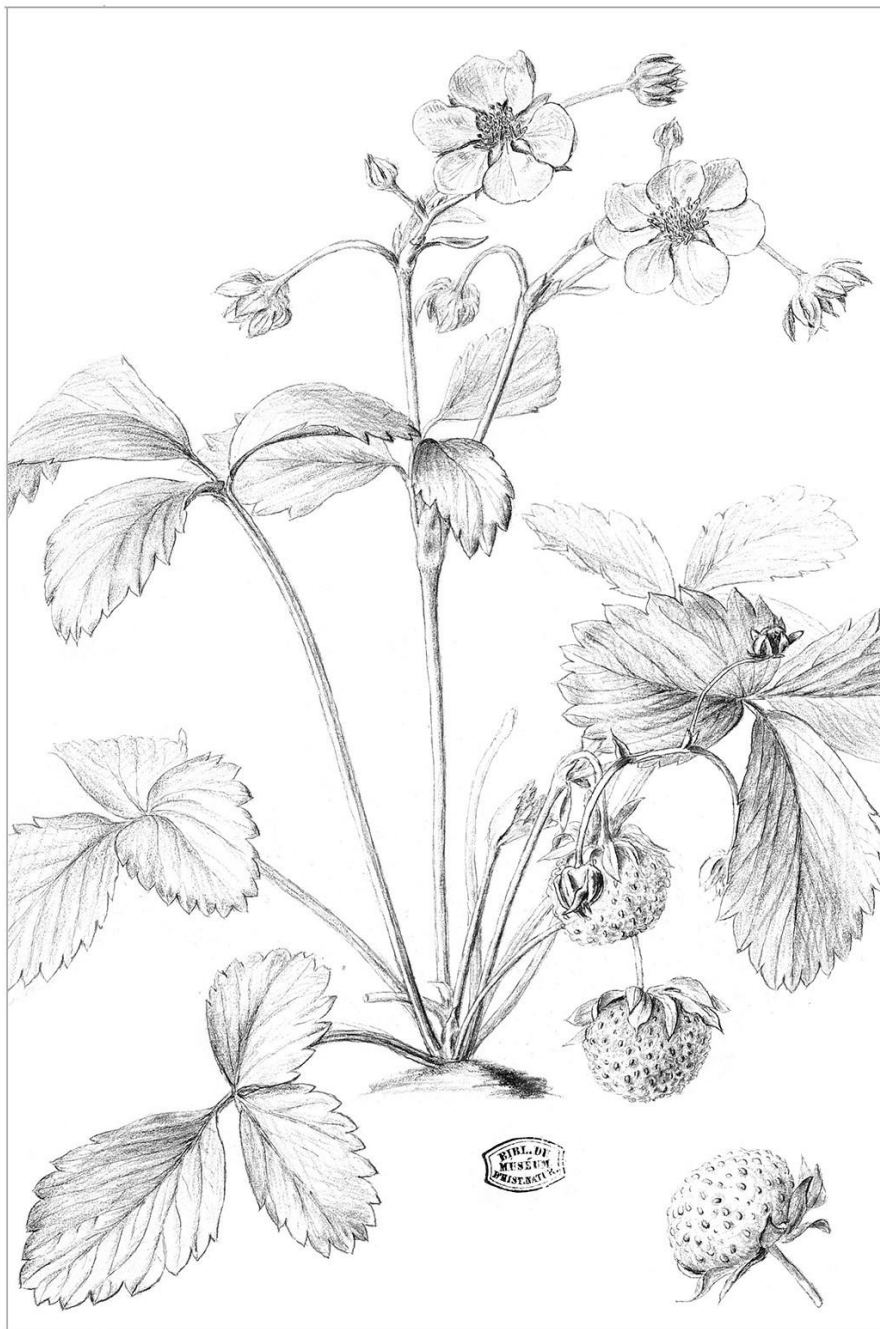
The primary aim of this doctoral thesis is to investigate the genetic determinants underlying key organoleptic traits of commercial strawberry (*Fragaria × ananassa*), with particular emphasis on aroma and tissue-specific secondary metabolism. The research is structured around the following specific aims:

- I. Identify quantitative trait loci (QTL) associated with volatile organic compounds (VOCs) production in mature strawberry fruit using an F₁ mapping population derived from the cross between ‘Candonga’ and ‘Senga Sengana’, two octoploid commercial cultivars.
- II. Identify QTLs regulating ellagitannin content in mature fruit from an octoploid strawberry F₁ mapping population generated from the cross between lines ‘232’ and ‘1392’.
- III. Functionally characterize gene *FxaC_10g48980* to elucidate its role in γ -dodecalactone biosynthesis and its potential interaction with jasmonic acid pathway in strawberry fruit.
- IV. Functionally characterize gene *FxaC_3g11070* to determine its implication on ellagitannins metabolism, specifically in the biosynthesis pathway of ellagic acid hexose derivatives in strawberry fruit.

By achieving these objectives, the study seeks to validate the central hypotheses of this PhD thesis:

- I. *FxaC_10g48980* is a candidate gene for γ -dodecalactone biosynthesis and may also contribute to jasmonic acid biosynthesis in strawberry fruit.
- II. *FxaC_3g11070* is implicated in the biosynthesis of a specific isomer of ellagic acid hexose in fruits and its expression varies among strawberry genotypes.

III. General introduction



Botanical illustration of *Fragaria* × *ananassa* subsp. 'Caroline' plant by Antoine Nicolas Duchesne. *Histoire Naturelle des Fraisiers*, 1768.

Portions of this section, as well as parts of the introduction of **Chapter II**, are partially reproduced from the following review, which has been published as part of this thesis:

Manuscript title: Regulation of plant tannin synthesis in crop species

Authors: José Mora^{†, 1}, Delphine M. Pott^{†, 2}, Sonia Osorio¹, José G. Vallarino^{1, *}

[†]: Co-authority for this publication. Both authors contributed equally to this work.

^{*}: Corresponding author.

¹: Instituto de Hortofruticultura Subtropical y Mediterránea “La Mayora” – Consejo Superior de Investigaciones Científicas – Universidad de Málaga – (IHSM-CSIC-UMA), Málaga, Spain.

²: Current address: Department of Plant Biochemistry, Centre for Plant Molecular Biology (ZMBP), Eberhard Karls University, Tübingen, Germany.

Journal: Frontier in Genetics. Section: Evolutionary and Population Genetics

Publication date: 02/05/2022

Volume: 13

DOI: [10.3389/fgene.2022.870976](https://doi.org/10.3389/fgene.2022.870976)

Contributions to this work: All the authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

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

Funding: This work was supported by grants RTI 2018-099797-B-100 (Ministerio de Ciencia, Innovación y Universidades, Spain) and UMA18-FEDERJA-179 (FEDER-Junta Andalucía). In addition, we acknowledge partial funding by the European Union’s H2020 Programme (BreedingValue; grant number 101000747), and PY20_00408 (PAIDI 2020-Junta de Andalucía). JM has received a predoctoral grant from Ministerio de Ciencia e Innovación (grant PRE 2019-091188).

1. Origins of strawberry species

1.1. *Rosaceae* family

The common word “rose” according to Greek mythology gain its significance when the Goddess of beauty, Aphrodite, gave this particular flower its name in honor to her son by rearranging his name, Eros (God of love). Later, the name of the genus type *Rosa* would give name to an entire family: *Rosaceae*.

Rosaceae family is composed of 100 genera and 3,000 species. It is the third most economically important family (Dirlewanger *et al.*, 2002) which involved fruit, nuts, ornamental, herb and woody plants. Many species of this family have been domesticated for human consumption, strawberries, almonds, pear, peach, sweet cherry and raspberry, some of the edible crops. Also, this family possesses nonedible species with high ornamental value, like roses.

Rosaceae consist of very distinct fruit types like drupe, pome, drupelet and achene. Depending on these fruit types, the family *Rosaceae* has been divided into four different subfamilies: *Rosoideae*, *Amygdaloideae/Prunoideae*, *Spiraeaoideae* and *Maloideae* (Potter *et al.*, 2002). The phylogeny of *Rosaceae* has changed and the family is divided into three basal groups depending on the nuclear and chloroplast *loci* (Potter *et al.*, 2007). For this reason, new incorporations have joined to the subfamilies, being *Amygdaloideae* (cherry, apricot, almond, etc.), *Dryadoideae* (includes *Purshia* or *Dryas*, for example) and *Rosoideae* (includes *Fragaria*, *Rubus*, or *Rosa*).

The cosmopolitan distribution of *Rosaceae*, mostly in temperate zones, together with the edible fruit are some of the reasons for their economic importance (Janick, 2010). The genetic diversity of *Rosaceae* family is very wide but the Genome Database for Rosaceae (GDR, www.rosaceae.org) was created as a repository of genetic and genomics for the *Rosaceae* family (Jung *et al.*, 2008).

1.2. *Fragaria* genus: characteristics and origin

Fragaria genus is globally one of the most important food plant groups, mostly known as strawberries, which are more than double global production compared with all other berries combined (Liston *et al.*, 2014).

Morphologically speaking, strawberries are not fruits in the botanical sense. The real fruits are dry achenes that are embedded in the receptacle. This receptacle is the thickened part of the stem from which there is a growing of the female floral organ (ovary). Therefore, the edible part of *Fragaria* species is considered as aggregate fruit with a poly-achene distribution around the receptacle. This morphological trait is a case of parallel evolving in *Rosaceae* tribe *Potentilleae* (Eriksson *et al.*, 1998).

Basically, the same life cycle is shared by almost all *Fragaria* species, where they are insect-pollinated, they present a low-growing rate with herbaceous perennials and trifolioliated leaves, they are capable of clonal growth by stolons and they have animal-dispersed accessory fruits (Johnson *et al.*, 2014). The flowers are actinomorphic due to their radial symmetry; they are white in color and usually have five petals. There are some exceptions to these characteristics: for example, *Fragaria iinumae* presents six to eight-petalled flowers and it's deciduous while usually the *Fragaria* species are evergreen (Liston *et al.*, 2014). On the other hand, five leaflets are typical of some species from the Chinese region and in some cases 4-5 leaflets are rarely observed in *Fragaria virginiana* and *Fragaria cascadiensis* (Liston *et al.*, 2014).

There are several traits used to differentiate species within the *Fragaria* genus: mature receptacles are diverse among the species (Liston *et al.*, 2014) but some of their characteristics cannot be preserved (Liston *et al.*, 2014). This is the reason why identification keys are focused on stolons morphology (Staudt, 2009) and some morphological traits of leaves as dentation, venation or texture. Moreover, the different ploidy level of *Fragaria* species tends to affect the size of the achenes, being possible to distinguish between diploid and tetraploid plants of *Fragaria mandshurica* and between diploids plants of *Fragaria vesca* and the octoploids plants (Staudt, 1962).

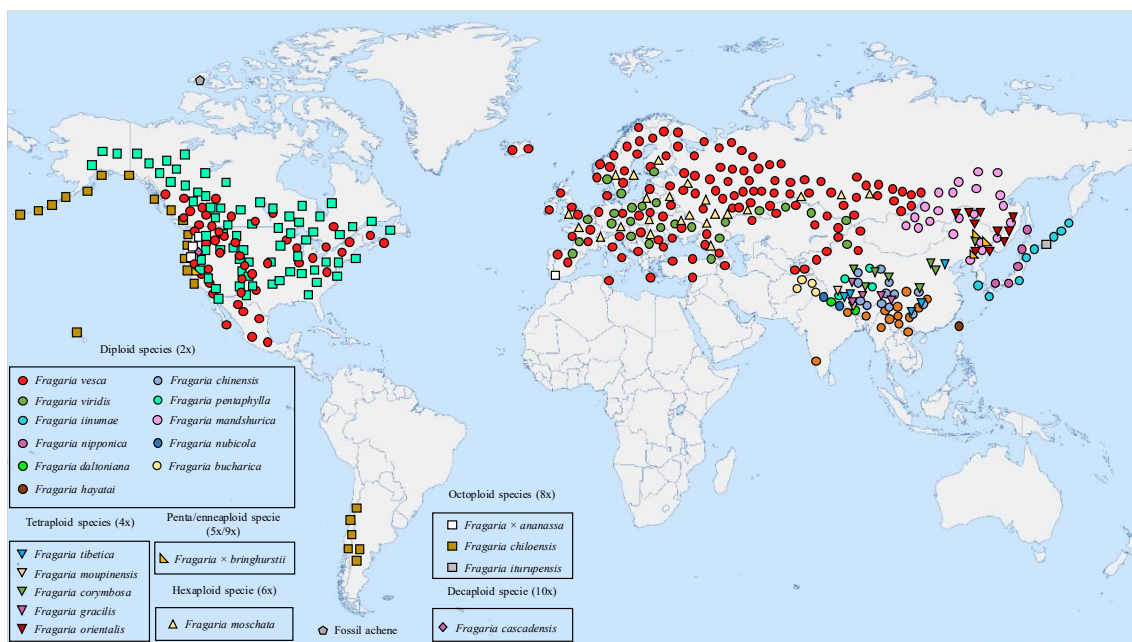


Figure 1. Approximate geographic natural distribution of *Fragaria* species. Species are grouped by ploidy level criterion. Modified from Liston *et al.*, (2014) with data from Lei *et al.*, (2017) and expowo (R package) from Zuanny *et al.*, (2024).

Natural wild occurring *Fragaria* species are spread mainly across the northern hemisphere (**Fig. 1**). Nevertheless, *Fragaria* species have a history of thousands of years

since archaeobotanical evidence showed that *Fragaria* achenes were present in pre-Columbian sites in North America (Gremillion and Sobolik, 1996; Pauketat *et al.*, 2002). Also, in their research Finn *et al.*, (2013) showed that the domestication of *Fragaria chiloensis* were carried out over a millennium ago by the indigenous Mapuche and Picunche people of Chile.

Table 1. Summary of *Fragaria* species and ploidy level grouped by clades according to phylogenetic studies. Two main phylogenetic clades were found: vesca (related to wild strawberry *Fragaria vesca*) and China clades.

Specie	Phylogenetic clade	Ploidy level	Chromosome count
<i>Fragaria</i> × <i>ananassa</i>	Vesca	Octoploid	2n= 8x= 56
<i>Fragaria bucharica</i>	Vesca	Diploid	2n= 2x= 14
<i>Fragaria cascadiensis</i>	Vesca	Decaploid	2n= 10x= 70
<i>Fragaria chiloensis</i>	Vesca	Octoploid	2n= 8x= 56
<i>Fragaria iturupensis</i>	Vesca	Octoploid/ Decaploid	2n= 8x= 56 2n= 10x= 70
<i>Fragaria mandshurica</i>	Vesca	Diploid	2n= 2x= 14
<i>Fragaria moschata</i>	Vesca	Hexaploid	2n= 6x= 42
<i>Fragaria orientalis</i>	Vesca	Tetraploid	2n= 4x= 28
<i>Fragaria vesca</i>	Vesca	Diploid	2n= 2x= 14
<i>Fragaria virginiana</i>	Vesca	Octoploid	2n= 8x= 56
<i>Fragaria chinensis</i>	China	Diploid	2n= 2x= 14
<i>Fragaria corymbosa</i>	China	Tetraploid	2n= 4x= 28
<i>Fragaria daltoniana</i>	China	Diploid	2n= 2x= 14
<i>Fragaria gracilis</i>	China	Tetraploid	2n= 4x= 28
<i>Fragaria moupinensis</i>	China	Tetraploid	2n= 4x= 28
<i>Fragaria nipponica</i>	China	Diploid	2n= 2x= 14
<i>Fragaria nubicola</i>	China	Diploid	2n= 2x= 14
<i>Fragaria pentaphylla</i>	China	Diploid	2n= 2x= 14
<i>Fragaria tibetica</i>	China	Tetraploid	2n= 4x= 28
<i>Fragaria hayatai</i>	Unresolved	Diploid	2n= 2x= 14
<i>Fragaria iinumae</i>	Unresolved	Diploid	2n= 2x= 14
<i>Fragaria nilgerrensis</i>	Unresolved	Diploid	2n= 2x= 14
<i>Fragaria viridis</i>	Unresolved (vesca)	Diploid	2n= 2x= 14

In terms of phylogeny, there are two main clades regarding the two nuclear genes *GBSSI-2* and *DHAR* (Rousseau-Gueutin *et al.*, 2009) and chloroplast genomes (Njuguna *et al.*, 2013) (**Table 1**). Both studies agreed in the differentiation of the “China clade” and “vesca clade”, comprising almost all known species of *Fragaria* except for *F. iinumae*, *F. hayatai*, *F. nilgerrensis* and *F. viridis*. However, *F. viridis* shows an 80% sister

relationship with vesca clade, as was previously described but the position of *F. nilgerrensis* remains unclear (Potter *et al.*, 2000).

Taking a closer look at the vesca clade, the authors saw that using the genes *GBSSI-2* (Rousseau-Gueutin *et al.*, 2009) and *nrITS* (Potter *et al.*, 2000) they were able to differentiate European *F. vesca* subsp. *vesca* from the American *F. vesca* subsp. *californica* and subsp. *americana*, but their position remains unclear respect to the octoploid and decaploid species (Njuguna *et al.*, 2013). According to these same authors, the possible origin of the chloroplast donor of the clade that contains the octoploid and decaploid species is North/South America. Furthermore, the authors suggested that the relationship of *F. vesca* subsp. *bracteata* and *californica* with decaploid and octoploid species were restricted to Beringian and northwestern North America locations, being a possible location for the origin of the octoploid species.

The China clade species are distributed in the homonymous country and adjacent Himalayan country, being only *F. nipponica* endemic to Japan. This clade comprises 9 species ranging from diploid to tetraploid where most of them are classified with an uncertain taxonomic status. In fact, only *F. tibetica* were classified taxonomically in 2001 (Liston *et al.*, 2014). From a phylogenetic perspective, the ancestor of this clade remains elusive and even some hypotheses support the possibility of multiple origin of this ancestor (Rousseau-Gueutin *et al.*, 2009; Njuguna *et al.*, 2013; DiMeglio *et al.*, 2014; Liston *et al.*, 2014).

On the other hand, the vesca clade comprise more ploidy levels, ranging from diploid to decaploid. Within this clade, both *F. chiloensis* and *F. virginiana* are included and they both are the octoploid ancestors of the cultivated strawberry *F. × ananassa*. Of course, this clade includes the representative specie *F. vesca* that together with *F. iinumae* have been proposed as progenitors of these species but also as progenitors of *F. iturupensis*. The vesca clade contains the uncommon decaploid species like the mentioned *F. iturupensis* (endemic from Iturup Island) and *F. cascadiensis* (detected on the west Coast). The case of *F. iturupensis* is quite exotic because both octoploid and decaploid individuals of the same taxon have been described (Hummer *et al.*, 2009).

In the study of Sun *et al.*, (2021) by analyzing the complete chloroplast genome of wild *Fragaria* species, they found the same clades found by Njuguna *et al.*, (2013) and Rousseau-Gueutin *et al.*, (2009). But they offered new information about the phylogeny of the genus. They found that *F. iinumae* seems to be the eldest of the extant species, in fact it seems to be the closest species to the common ancestor of the *Fragaria* genus, followed by *F. viridis*. These results are consistent with the present of singular characteristics in the *F. iinumae* specie, like the presence of usual seven-petalled flowers,

spinuliferous pollen grains and sunken achenes that are not so common among the *Fragaria* genus (Staudt, 2009).

Historically, the cultivated strawberry has its origin in the French region of Brest, in Brittany. Namely, in the 1760s, the French botanist Antoine Nicolas Duchesne published *Histoire Naturelle des Fraisiers* a work where he elucidated the parentage of the cultivated strawberry by crossing the octoploid species *F. chiloensis* and *F. virginiana* (Liston, 2023). The work of Duchesne was not restricted to cultivated strawberries (*F. × ananassa*) because he also tried to cross the octoploid parental *F. chiloensis* with the diploid *F. vesca* and with the hexaploid *F. moschata* but giving unsuccessful results explained by sterile offspring (Liston *et al.*, 2014).

2. *Fragaria* × *ananassa* genomic complexity

The generation of an allopolyploid from two or more diploid species is a complex process. In this formation, individuals usually contribute to the overall genome with one subgenome, which carries a big deal in terms of genomic organization. This major challenge is organizing the coexistence of several subgenomes in the same nucleus and coordinating the expression of homologues. To gain genomic stability, polyploid genomes tend to obtain a more diploid-like state by gene loss or fractionation of chromosome number reduction (Lysak *et al.*, 2006; Wendel, 2015).

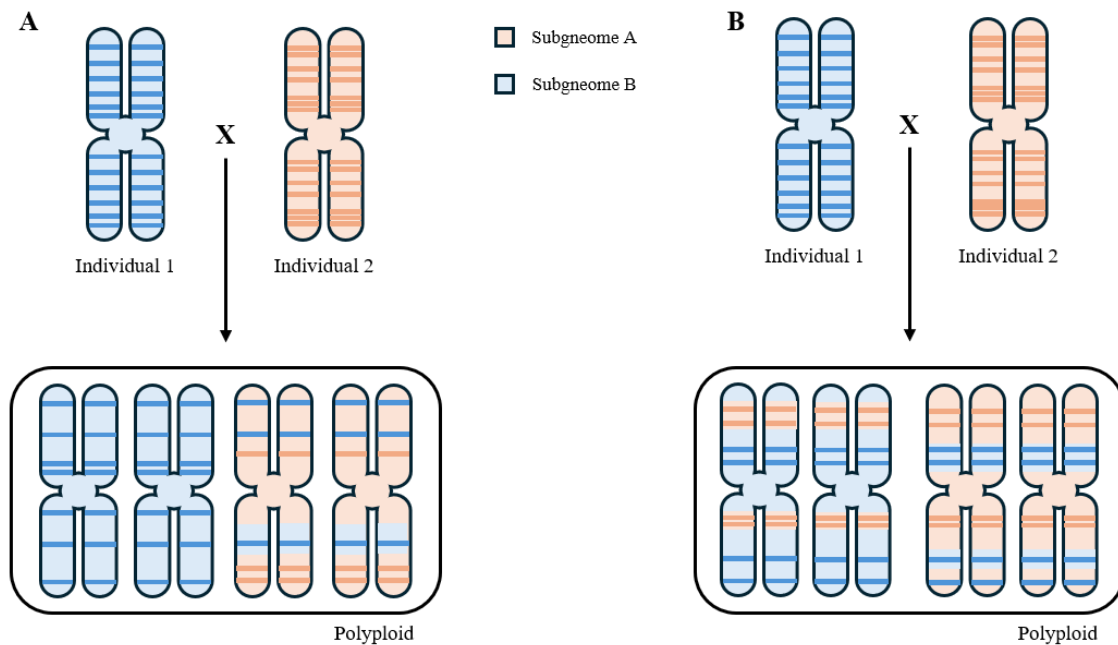


Figure 2. Schema of polyploid chromosomal organization. (A) Subgenome dominance. More gene retention, replacement of region by homologous recombination from submissive subgenome. (B) Subgenome structure evolution. Characterized by homoeologous exchanges, high synteny level, high protein identities, etc.

This deal becomes particularly challenging in the case of high-order polyploidy species, like cultivated strawberry, due to the genome complexity. The coordination of the subgenomes can be carried out through two different mechanisms: dominant subgenome evolution or parallel subgenome structure evolution (**Fig. 2**).

The subgenome dominance refers to the situation where one genome of an allopolyploid retains significantly more genes than the other submissive genome (Thomas *et al.*, 2006; Bird *et al.*, 2018). This dominant subgenome expressed more of their genes compared to the submissive one. This dominance concurs with a significant retention of tandemly duplicated genes than the submissive subgenome (Edger *et al.*, 2019; VanBuren *et al.*, 2020) and with the replacement of regions from the submissive subgenome via homoeologous exchange (Edger *et al.*, 2019). But the dominance of one subgenome is not complete, because the submissive subgenome retains genes that are significantly more expressed than the homeologs from the dominant subgenome (Edger *et al.*, 2017). In fact, in tetraploid ($2n=4x=48$) blueberry (*Vaccinium corymbosum*), one subgenome was highly expressed during fruit development while another subgenome presented higher expression in almost all developmental stages (Alger and Edger, 2020). In summary, subgenome dominance is characterized by retention of genes that are highly expressed against others and by the reduced repetitive sequences and stronger selective pressure (Schnable *et al.*, 2011; Eckardt, 2014; Parkin *et al.*, 2014; Bird *et al.*, 2018; Song *et al.*, 2024).

The second mechanism is the parallel subgenome structure evolution that is characterized by equivalent chromosome components, higher protein identities, high synteny and symmetric purifying selection of all subgenomes. Examples of dominant subgenome evolution include species like wheat (*Triticum aestivum*) (Pont *et al.*, 2013) while representants of parallel subgenome structure evolution are extended to other eukaryotic species like carp and goldfish (Li *et al.*, 2021).

The occurrence of subgenome dominance in many allopolyploids has been linked to the presence of low transposable elements (TEs) (Hollister and Gaut, 2009; Schnable and Freeling, 2011; Schnable *et al.*, 2011; Cheng *et al.*, 2016; Edger *et al.*, 2017, 2019; Wendel *et al.*, 2018). These TEs contribute highly to genome size variation, and they can alter plants gene expression (McClintock, 1984; Feschotte *et al.*, 2002; Hollister and Gaut, 2009). Due to the high proliferation and insertion rate of TEs, they are epigenetically silenced through methylation, which affects surrounding genes lowering their expression (Hollister and Gaut, 2009). This decrease in gene expression may be the reason for the association of TEs with subgenome dominance (**Fig. 3**) (Hollister and Gaut, 2009).

In subgenome dominance, the submissive subgenome exhibits lower gene expression but also has significantly higher TE density compared to the dominant one. Edger *et al.*, 2020 demonstrated in the octoploid *F. × ananassa* that the subgenome with the lowest TE density was dominant over the other three subgenomes (Edger *et al.*, 2019). The relationship between lower TE density and subgenome dominance has been observed in recent allopolyploids like wheat and cotton (Guo *et al.*, 2014; Li *et al.*, 2014; Colle *et al.*, 2019).

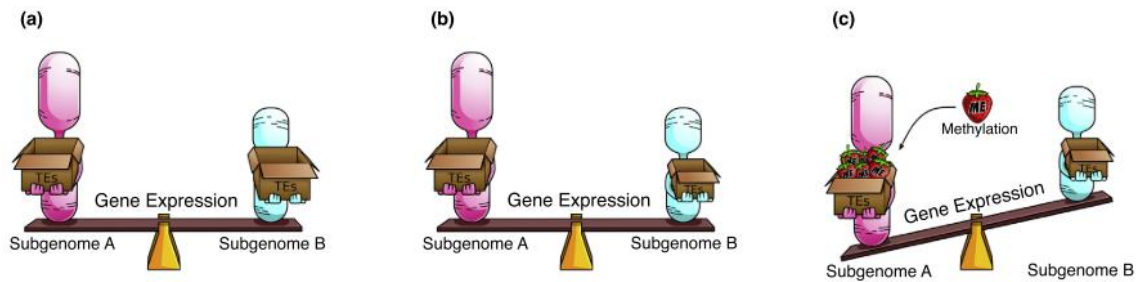


Figure 3. Subgenome domination and its relationship with transposable elements (TEs).

(a) Both subgenomes have equal amounts of TEs, resulting in an equal gene expression. (b) Subgenome A has more TEs than subgenome B, but as there isn't methylation of any subgenome the result is equal gene expression of both subgenomes (c) Subgenome A has more TEs than subgenome B and they are methylated, resulting in lower gene expression for subgenome A. Taken from Alger and Edger, (2020).

2.1. Octoploid strawberry genome

Given the fact the *Fragaria* species can be found in very different habitats (Johnson *et al.*, 2014), these species show broad facility to adapt to several environmental conditions and adaptability to pest or diseases to overcome selective pressure. Although natural variation is considered as a source of opportunities to gain resistance or adaptability to certain environments, the reality is that there is a stronger driving force for evolution. In this context, polyploidy and hybridization have played vital roles in land plan speciation (Jiao *et al.*, 2011; Ye *et al.*, 2020; Jin *et al.*, 2023).

The genus *Fragaria* presents a high interfertility between and within ploidy levels, a characteristic that induces the natural occurrence of polyploid species (Milne and Abbott, 2008). Polyploid events are common in the plant kingdom, and they consist of whole-genome duplications, that can lead to an enhancement of genetic breeding by the arise of novel and varied phenotypes (Soltis, 2005; Freeling and Thomas, 2006; Doyle *et al.*, 2008; Van De Peer *et al.*, 2017). Ploidy can be presented in two different ways: autopolyploidy and allopolyploidy, depending on a single or multiple diploid progenitor species, respectively (Stebbins, 1947; Comai, 2005). Allopolyploidy is widely present in

crops species offering desirable agronomical traits like aroma al flavor in strawberries (Ulrich and Ulbricht, 2013).

The natural hybridization of *F. virginiana* and *F. chiloensis* originated the allo-octoploid specie *F. × ananassa*, the cultivated strawberry. The first chromosome-scale genome assembly for the octoploid strawberry was published by Edger *et al.*, (2019). Previous studies selected *F. vesca* as one diploid progenitor of *F. × ananassa* based on meiotic chromosome pairing Edger *et al.*, (2019). The hypothesis was that *F. iinumae* was also a diploid progenitor, together with two unknown species (Tennessen *et al.*, 2014). The phylogenetic analysis conducted in this study placed *F. iinumae*, *F. nipponica* and *F. viridis* as three diploid progenitors of *F. × ananassa*, hypothesizing that the tetraploid and hexaploidy species may be evolutionary intermediates before the appearance of the wild octoploid specie. They also selected the subgenome of *F. vesca* progenitor as the dominant one for retaining 20% of protein-coding genes, 14% of long non-coding RNAs (lncRNA) genes and 19% fewer TEs than the chromosomes from other progenitors.

Two years later, Liu *et al.*, (2021) refined the ‘Camarosa’ genome assembly previously published (Edger *et al.*, 2019), stablishing the *F. × ananassa* genome v1.0.a2. In the next years, the proliferation of genome assemblies for several varieties of cultivated strawberry: ‘Benihoppe’ (Song *et al.*, 2024), ‘FL15.89-25’ (Fan *et al.*, 2022) and ‘Yanli’ (Mao *et al.*, 2023).

In the work of Song *et al.*, (2024), apart from publishing cv. ‘Benihoppe’ genome assembly, they studied the contribution to each of the four subgenomes (A, B, C and D) to the overall *F. × ananassa* genome. Their results suggested that subgenomes B, C and D were generated from the same ancestor, probably an unsampled or extinct population of *F. iinumae*. Their phylogenetic analysis showed that ‘Benihoppe’ assembly against the last version of *F. vesca* genome, v6.0 (Zhou *et al.*, 2023) shared a significant high degree of identity to *F. vesca* genome, and further stablishing that this species is the progenitor of subgenome A in the octoploid specie. This phylogenetic analysis demonstrates that subgenomes C and D originated from closely related ancestors, but the identities of these ancestors remain unresolved. They concluded that *F. iinumae* is the progenitor of the subgenome B and that the subgenomes C and D were phylogenetically related to this same species. These results suggested that *F. vesca*, *F. iinumae* and two *F. iinumae*-like ancestors contributed to the evolution of the genome evolution of *F. × ananassa*. Finally, this study supports the previous hypothesis of *F. vesca*-subgenome dominance, through an ABBxBx subgenome structure in where subgenome A is dominant and the others exhibit parallel evolution patterns.

Nevertheless, Jin *et al.*, (2023) showed some results that not matched with the results of Edger *et al.*, (2019) in terms of two of the diploids ancestors of the octoploid strawberry.

They showed evidence enough to assign *F. vesca* and *F. iinumae* as progenitors of subgenomes A and B, respectively (**Fig. 4**). In fact, they claimed that *F. vesca* and *F. mandshurica* (tetraploid species) are the closest living relatives to subgenome A and *F. iinumae* as the closest relative to subgenomes B, C and D. But, in contrast, *F. nipponica* and *F. viridis* were not considered as close relatives to subgenomes C and D, respectively, as thought before. The final hypothesis in this work is regarding the possibility that subgenomes B, C and D were generated after polyploidization event between *F. iinumae* and an extinct close relative.

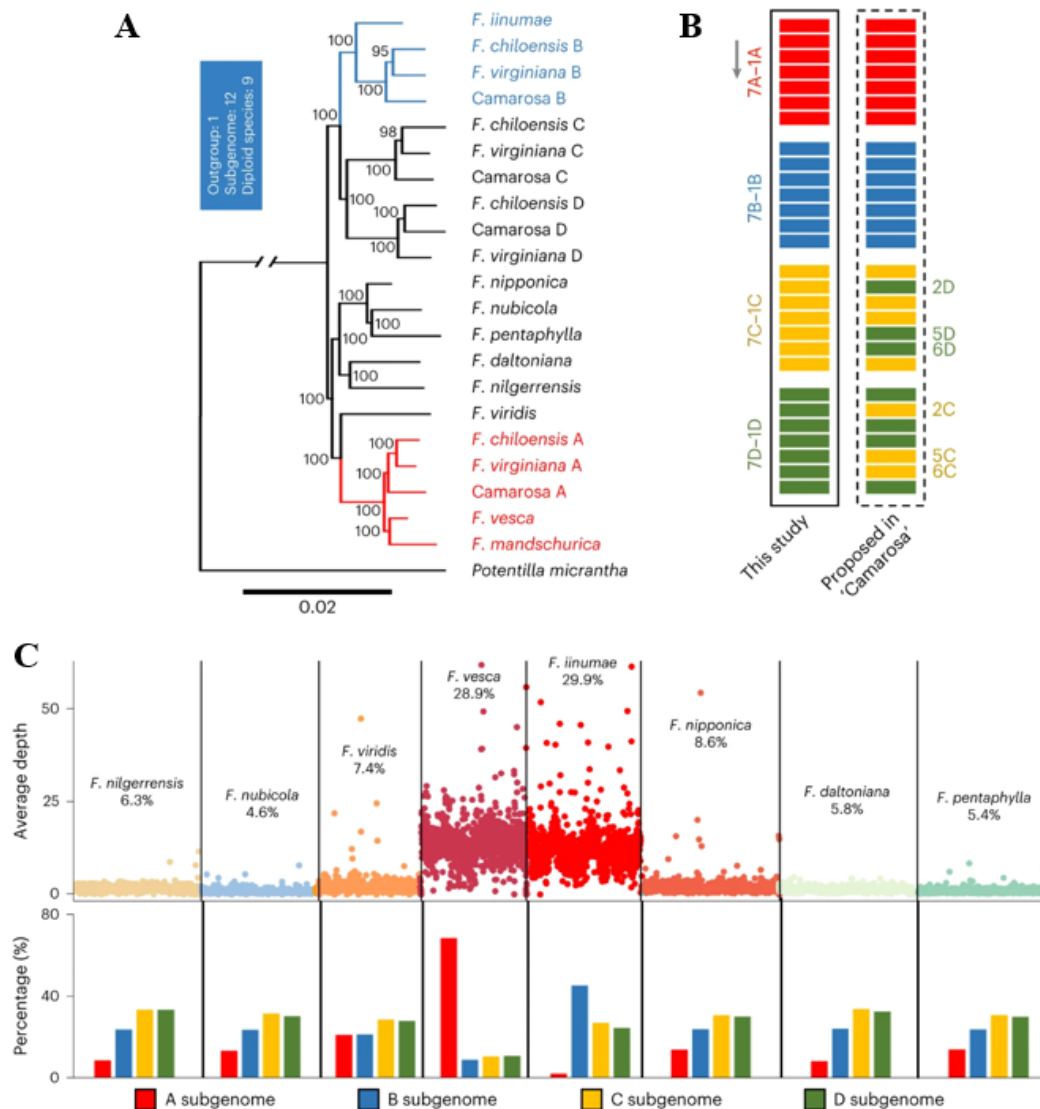


Figure 3. Subgenome assignments of the octoploid strawberry. (A) Phylogenetic analysis of three octoploids and nine diploid *Fragaria* species. Red and blue lines represent clades where *F. vesca* and *F. iinumae* are located. (B) Subgenomes assignment comparison. (C) Depth of reads of a cultivated strawberry accession mapped to the diploid strawberry genomes and distribution of reads (%) contributed by diploids to each subgenome. Modified from Jin *et al.*, (2023).

3. Strawberry fruit and quality traits

As explained before, strawberries are accessory fruits that are widely consumed by their organoleptic characteristics. The organoleptic components are related to the perception of the olfactory and taste systems, vision and tactile sensation. The perception of these components is flavor (aroma and taste), but also color and firmness. In fact, some studies about sensory perception indicate that the most influencing characteristics for consumers preference are sweetness, flavor intensity and sugar, which are produced by sugar concentration, volatiles compounds and fruit firmness, respectively (Schwieterman *et al.*, 2014; Fan *et al.*, 2021) (**Fig. 4**).

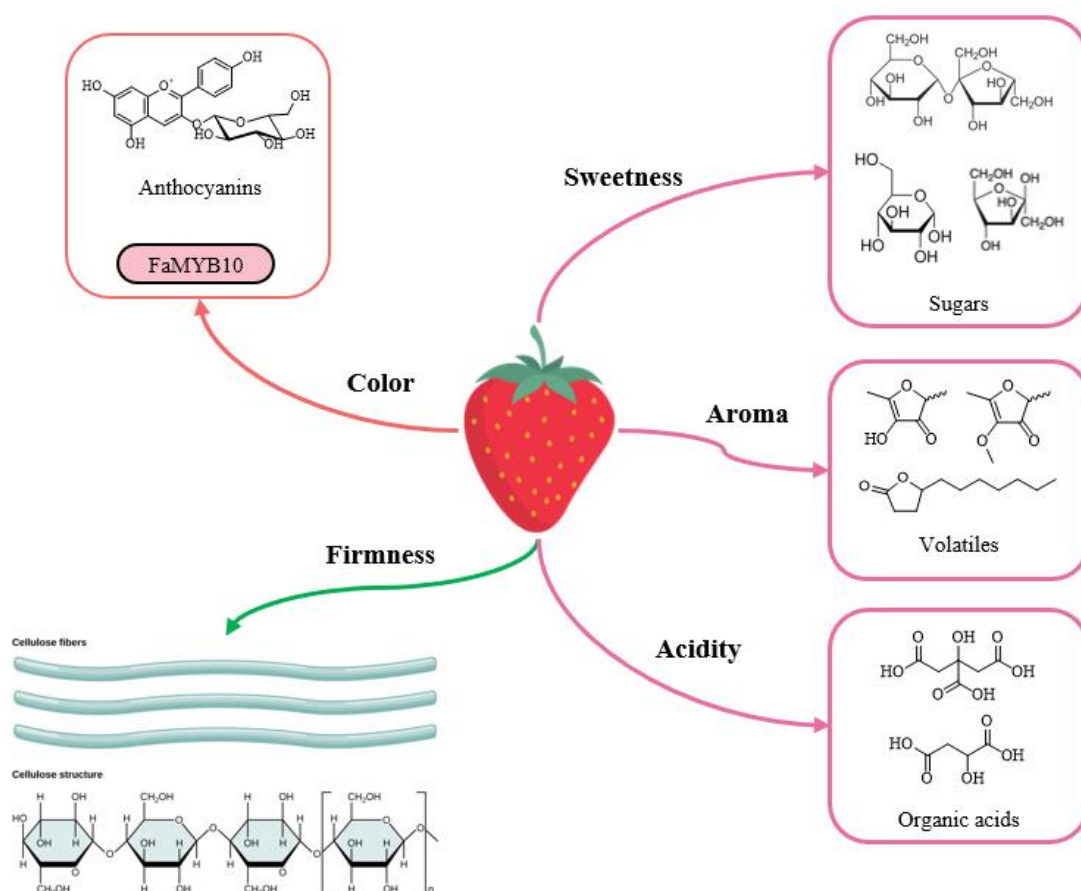


Figure 4. Main organoleptic characteristics of strawberry fruit. Flavor is the overall sense of taste (balance between sweetness and acidity) and smell. Anthocyanins offer red color to the fruit, and the firmness is due to the presence of cellulose in the fruit.

Fruit flavor is the overall perception of sugars, acids, and volatile compounds. Regarding volatiles, around 360 volatiles have been detected in strawberries but only 10-20 have been described as determinants for strawberry aroma (Schwieterman *et al.*, 2014; Urrutia *et al.*, 2017). Schwieterman *et al.*, (2014) identified 31 volatiles described to be important for strawberry flavor intensity and other 6 fatty acid-derived volatiles capable of

enhancing sweetness intensity independently of sugar content. Furthermore, in the study of (Fan *et al.*, 2021) 6 more volatiles were found to enhance sweetness perception in absence of sugar content: 1-penten-3-one, γ -dodecalactone, pentyl butanoate, hexyl butanoate, hexyl acetate and 1-methylbutyl butanoate.

Nonetheless, sensory studies select sweetness intensity as the most determinant of consumer preference (Park *et al.*, 2006; Schwieterman *et al.*, 2014; Ulrich and Olbricht, 2016). Sugar is one of the main carbohydrates in sweetness perception for it's the one mainly transported from photosynthetic to receptacle fruit sink tissues. Once arrived at the fruit cells, sucrose can be converted into starch for storage, it can be accumulated in the vacuole or simply being hydrolyzed by invertases to free glucose and fructose. Therefore, sucrose, glucose and fructose are the main soluble sugars in ripe strawberries (Liu *et al.*, 2023). Another parameter in flavor perception is the acidity of the fruit, whose main influencing molecules are organic acids. Among these organic acids, the most relevant in strawberry fruit are citrate and malate. The levels of citrate during ripening remain constant but the levels of malate increase with the advance of the maturation (Fait *et al.*, 2008). It is accepted that a high ratio between sugar and organic acids is the major determinant for consumers' acceptance (Zhang *et al.*, 2011).

Fruit characteristic color is modulated by the quantity and variability of flavonoids, and more specifically anthocyanins. The red color is mainly produced by the anthocyanin pelargonidin-3-glucoside and this compound also provides health benefits to consumers (Liu *et al.*, 2023).

Strawberry fruits soften during maturation, acquiring a melting texture, shortening their postharvest life (Brecht *et al.*, 2016). Fruit firmness is determined by the cell wall structures (primary cell wall and the middle lamella). Finally, strawberries are also appreciated for their nutritional quality due to the content in several vitamins (A, B6, E and K), thiamin, riboflavin, carotenoids and folic acid (Tulipani *et al.*, 2008, 2009; Giampieri *et al.*, 2015).

4. Volatile organic compounds in strawberry fruit

The flavor and aroma of strawberries arise from a combination of diverse types of molecules like sugars, acids and volatile organic compounds (VOCs). More than 360 VOCs have been detected in strawberry, including esters, aldehydes, ketones, alcohols, terpenes, lactones and other compounds (Zabetakis and Holden, 1997; Ménager *et al.*, 2004; Jeti *et al.*, 2007; Zorrilla-Fontanesi *et al.*, 2012).

Volatile organic compounds are a special kind of plant compound that present low polarity, therefore being lipophilic molecules, with high vapor pressure at ambient temperature and a molecular mass below 300 Da. VOCs are a chemically diverse family,

and their biosynthesis depends directly on the primary metabolism, demonstrating the high degree of connectivity between primary and secondary metabolism. The building blocks of all the diversity of VOCs branches off from only several primary metabolites: acetyl-CoA, pyruvate, phosphoenolpyruvate, and *D*-erythrose-4-phosphate. Based on the biosynthetic origin, VOCs are divided into four different classes: terpenoids, phenylpropanoids/benzenoids, polyunsaturated fatty acids (PUFA) derivatives and amino acid derivatives (**Fig. 5**).

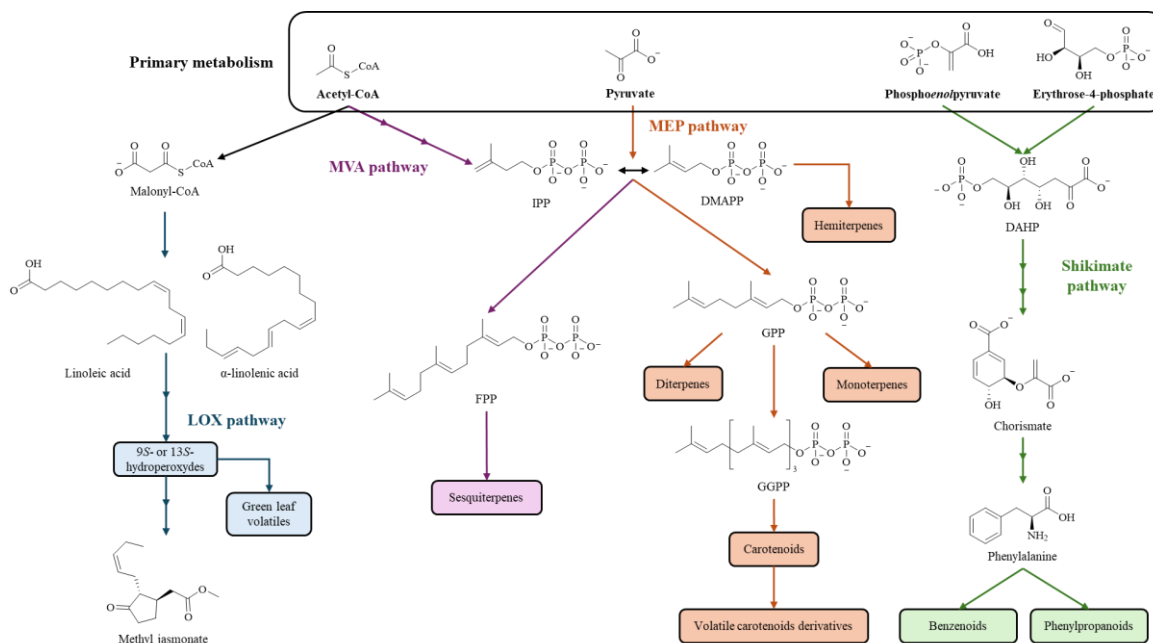


Figure 5. Overview of biosynthetic pathways leading to formation of plant volatile organic compounds (VOCs). Primary metabolite offers the precursors of all VOCs and there are four major biosynthetic pathways, namely shikimate, mevalonic acid (MVA), methylerythritol phosphate (MEP) and lipoxygenase (LOX) pathways. Abbreviations: IPP, isopentenyl diphosphate; DMPP, dimethylallyl diphosphate; FPP, farnesyl diphosphate; GPP, geranyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; DAHP, 3-deoxy-*D*-arabinoheptulosonate-7-phosphate. Modified from Dudareva *et al.*, (2013).

4.1. Functions of plant volatile organic compounds

Plant VOCs shape the chemical landscape of numerous ecosystems, mediating both intra- and interspecific interactions. They play a key role in attracting pollinators (Raguso, 2008) and seed dispersers, as well as in above- and below-ground defense against herbivores (Degenhardt *et al.*, 2009; Unsicker *et al.*, 2009; Ali *et al.*, 2012; Hiltbold and Turlings, 2012). Additionally, they contribute to protection against pathogens (Huang *et al.*, 2012) and facilitate plant-to-plant signaling (Baldwin *et al.*, 2006).

The emission of plant VOCs is spatially regulated. In angiosperms, most VOCs are released from flowers, while vegetative tissues emit small quantities of VOCs and often

in response to mechanical damage or pathogen infection (Arimura *et al.*, 2004; Knudsen *et al.*, 2006). Nevertheless, in all plant organs, VOCs synthesis and emission exhibit a changing behavior during plant development. In fact, VOCs concentration and emission increase during initial stages of development and then concentrations either fall or remain constant (Vassão *et al.*, 2006).

One of the main roles of VOCs in plant physiology is related to plant reproduction. The highest amounts of VOCs in angiosperms are emitted by the flowers, as said before. This fact answered the need to attract pollinators to ensure plant reproductive success. Indeed, floral volatile profiles are species-specific and often correlate with the type of pollinator (Jiang *et al.*, 2023). For instance, moth-pollinated plant species predominantly emit high levels of benzenoids, along with smaller amounts of terpenoids and nitrogen-containing compounds (Jiang *et al.*, 2023). In contrast, bat-pollinated flowers primarily release sulfur-containing volatiles (Von Helversen *et al.*, 2000). Also, pollinators can distinguish between complex volatile profiles (Wright *et al.*, 2005), establishing preferences towards specific profiles (Ashman, 2000).

Apart from this, floral emitted VOCs have roles beyond pollinator attraction. Some VOCs have been reported to have antimicrobial or antifungal activities (Hammer *et al.*, 2003; Huang *et al.*, 2012) but also some roles to protect the reproductive plant organs have been described (Junker *et al.*, 2011).

4.2. Biosynthesis and function of terpenoids volatiles

Terpenoids constitute the most diverse group of VOCs, with more than 80,000 compounds detected to date (Christianson, 2017). All terpenes shared the two same five-carbon precursors, isopentenyl diphosphate (IPP) and its allylic isomer, dimethylallyl diphosphate (DMAPP) (McGarvey and Croteau, 1995), both also known as isoprenoids. In plants, there are two main pathways leading to the formation of terpenoids: the methylerythritol-4-phosphate (MEP) exclusively located in plastids (Hsieh *et al.*, 2008) and the mevalonic acid pathway, located in cytoplasm but with possible contributions of endoplasmic reticulum and peroxisomes (Simkin *et al.*, 2011; Pulido *et al.*, 2012). The MEP pathway yields monoterpenes (linalool, myrcene and limonene for example) and diterpenes, while MVA pathway leads to the formation of sesquiterpenes (C₁₅) like (*E*)- β -farnesene, caryophyllene or nerolidol (Li *et al.*, 2023).

The MVA pathway consists of six enzymatic reactions starting with a condensation of three molecules of acetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA that is reduced to mevalonate. This molecule is phosphorylated twice and suffers one decarboxylation to produce IPP (Lange *et al.*, 2000).

The MEP pathway comprises seven enzymatic steps, where the first one is a condensation between *D*-glyceraldehyde 3-phosphate (GAP) and pyruvate to yield 1-deoxy-*D*-xylulose 5-phosphate. This molecule through isomerization and reduction then forms MEP. This molecule can be converted into IPP and DMAPP after 5 different enzymatic reactions. These two molecules are substrates of prenyltransferases for condensation reactions to produce prenyl diphosphates precursors: geranyl diphosphate (GPP), farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP) (Dudareva *et al.*, 2013). The MEP pathway is a source of precursors for monoterpenes (C₁₀), diterpenes (C₂₀), carotenoids and other products derived from their breakdown, like gibberelins, chlorophyll or tocopherols (Tholl, 2015). Indeed, two molecules of FPP or GGPP can be condensed to form sterol precursor squalene (C₃₀) and the carotenoid precursor octahydrolycopene (C₄₀) (Vranová *et al.*, 2013).

MEP pathway can produce both IPP and DMAPP with a molar ratio of 6:1 (Rohdich *et al.*, 2003), while IPP can be obtained also from MVA pathway. These compounds, as explained before, are isomers and they can be interconverted each other thanks to the isopentenyl diphosphate isomerase (IDI), establishing an equilibrium between them (Nakamura *et al.*, 2001). Because of this isomerization, both IPP and DMAPP serve as metabolic crosstalk between cellular compartments (Wu *et al.*, 2006; Orlova *et al.*, 2009). Transport of terpenes is thought to be under the control of ABC transporters (Shibata *et al.*, 2016; Fu *et al.*, 2017).

The biosynthesis of prenyl diphosphate intermediates is compartment-specific and depends on subcellular localization of the enzymes. In the case of the cytosol, head-to-tail condensation of two IPP molecules with one molecule of DMAPP produces FPP, which is the precursor of all sesquiterpenes. On the other hand, in the plastids head-to-tail condensation of one molecule of DMAPP with three molecules of IPP occurs to generate GPP and GGPP, precursors of mono and diterpenes, respectively (Dudareva *et al.*, 2013).

Terpenoids are known for playing a role in photosynthesis by regulating plant growth and development, pollination, and resistance to different stresses (Loreto *et al.*, 2014). Some terpenes have been found as aromatic compounds of plants to attract pollinating insects (Kaiser, 1994; Azuma *et al.*, 2001). Nevertheless, terpenes were identified as interesting molecules to attract natural enemies of some pests (Liu *et al.*, 2017).

In the case of cultivated strawberry, the terpenoid profile is prominently represented by linalool and nerolidol. In the case of wild strawberry, other terpenoids have been identified like α/β -pinene, β -phellandrene or myrtenyl acetate (Aharoni *et al.*, 2004).

4.3. Biosynthesis of shikimate derived volatiles

The amino acid phenylalanine has a pivotal role in the synthesis of VOCs, being its derivatives the second largest group of plant VOCs (Knudsen *et al.*, 2006). The precursors of shikimate pathway, phosphoenolpyruvate and *D*-erythrose 4-phosphate, have their origin in primary metabolite because they come directly from the glycolysis and the pentose phosphate pathway, respectively. These precursors through shikimate and arogenate pathways can be converted into phenylalanine establishing a connection between central carbon metabolism with this amino acid (Tzin and Galili, 2010; Maeda and Dudareva, 2012).

The biosynthesis of phenylalanine occurs within plastids (Maeda and Dudareva, 2012) but the biosynthesis of most derived VOCs takes place outside this compartment. The first step for the biosynthesis of this kind of VOCs consists of a deamination of the phenylalanine carried by the enzyme *L*-phenylalanine ammonia lyase (PAL), generating *trans*-cinnamic acid. Compounds classified as benzenoids (C₆-C₁) are produced with cinnamic acid as a precursor and it can be produced via β -oxidative pathway, non- β -oxidative pathway, or a combination of both (Boatright *et al.*, 2004; Orlova *et al.*, 2006).

The phenylpropenes (C₆-C₃) are a subfamily of phenylpropanoids where we can find some molecules such as eugenol, their derivatives, and chavicol, for example. Their biosynthesis comprises several steps shared with the one from lignin. From monolignol, two enzymatic reactions are required to remove the oxygen atom from the C9 position (Koeduka *et al.*, 2006; Dexter *et al.*, 2007). The resultant coniferyl alcohol is converted to coniferyl acetate (Dexter *et al.*, 2007) and this molecule is then substrate of eugenol synthase or isoeugenol synthase to form eugenol or isoeugenol, respectively (Koeduka *et al.*, 2008). Cinnamic acid can also be substrate of the 4-cinnamic acid hydrolase to form *p*-coumaric acid and the coumaric acetate can be used to form chavicol (Vassão *et al.*, 2006). In addition, eugenol, isoeugenol and chavicol can be modified via methylation by action of *O*-methyltransferases to form methyl eugenol, methyl isoeugenol and methyl chavicol (Gang *et al.*, 2002).

Some other VOCs with chemical structure of C₆-C₂, like phenylacetaldehyde and 2-phenylethanol, are biosynthesized from phenylalanine (Kaminaga *et al.*, 2006; Tieman *et al.*, 2007). In the case of tomato, phenylalanine is converted by action of an aromatic amino acid decarboxylase together with other enzymes to form phenylacetaldehyde (Tieman *et al.*, 2006; Dudareva *et al.*, 2013).

This pathway leads to the production of a plethora of volatile compounds that participate in organoleptic characteristics of strawberry fruits. In this case, eugenol biosynthesis is conducted by two different eugenol synthases (*FaEGS1* and *FaEGS2*) and one MYB transcription factor (*FaEOB1*) (Aragüez *et al.*, 2013; Medina-Puche *et al.*, 2014). Methyl



The first step in the LOX pathway, linoleic and linolenic acid are substrate of a regiospecific dioxygenation of the position 9 or 13 to form 9- or 13-hydroperoxy intermediates (Feussner and Wasternack, 2002). 9- and 13-hydroperoxydes are highly reactive molecules and they can be converted into a series of signaling and defense molecules called phytooxylipins by several enzymes (Griffiths, 2015).

The hydroperoxides are metabolized by three enzymes grouped into the CYP74 subfamily of cytochrome P450 enzymes: allene oxide synthase (AOS), divinyl ether synthase (DES) and hydroperoxide lyase (HPL). The bifurcation of AOS is required for jasmonate biosynthesis and will be discussed later and the HPL is the main responsible of green leaf volatile synthesis (GLVs). HPL produces the cleavage of the C-C bond of the hydroperoxides to form C₆ or C₉ aldehydes but also it can also produce 9-oxononanoic acid or 12-oxo-(9Z)-dodecenoic acid, a precursor of traumatin (Grechkin *et al.*, 2003, 2006). The C₆ or C₉ resulting aldehydes can be also reduced to alcohols by NAD-dependent alcohol dehydrogenases (ADH) (Matsui *et al.*, 2012; Mwenda and Matsui, 2014; Vincenti *et al.*, 2019) and these alcohols can be converted into their respective esters (D'Auria *et al.*, 2007).

The PUFA hydroperoxides are suitable for transformation into other compounds through different enzymes. Through a reductase, these molecules can be converted into more stable molecules, called fatty acid hydroxides (Göbel *et al.*, 2001). The action of peroxygenase (POX) or an epoxy alcohol synthase (EAS) produces the transformation of PUFA hydroperoxides into epoxy-hydroxy fatty acids (Hamberg, 1999). Finally, in low oxygen conditions lipoxygenases catalyze the breakdown of the *O-O* bond to form alkoxy radicals that can be rearranged themselves into ketodienes (Vincenti *et al.*, 2019).

The HPL branch of the LOX pathway is the main responsible for GLVs biosynthesis. This enzyme is a member of the cytochrome P450 family that has been found in both photosynthetic and non-photosynthetic tissues (Fauconnier *et al.*, 1997; Salas and Sánchez, 1999; Gargouri *et al.*, 2004; Padilla *et al.*, 2010; Jacopini *et al.*, 2016). The activity of HPL is very specific and the substrates of this enzyme must fit some structural requirements: substrates must be long-chain fatty acids (over 12 carbons) with a terminal carboxyl group, a hydroperoxyl group on C9 or C13, and a conjugated diene system (*Z*, *E*) adjacent to the hydroperoxide-bearing carbon (Hatanaka *et al.*, 1987, 1992; Fauconnier *et al.*, 1997; Noordermeer *et al.*, 2001). Furthermore, the regioselectivity of the HPL is so high that only 13(*S*)-hydroperoxides can be converted by this enzyme (Kajiwara *et al.*, 1982; Matoba *et al.*, 1985).

5. Antioxidant compounds

Antioxidant compounds are chemical species whose function is to inhibit or delay the process of oxidation (Mittler, 2017). Oxidation is the chemical process of losing

electrons, but biologically speaking, it is the production of free radical (reactive) species, which can attack electron-rich molecules such as lipids, proteins or even nucleic acids, damaging cells and tissues and thereby causing the alteration of homeostasis (Engwa, 2018). To fight against these reactive species, plants have a myriad of chemical compounds known as bio-active molecules whose activity are based on their antioxidant properties. These compounds have the ability to protect biological structures acting as scavengers of free radicals such as reactive oxygen species (ROS) and reactive nitrogen species (RNS) (Karadag *et al.*, 2009; Wang *et al.*, 2009).

One of the main chemical classes of plant antioxidants are polyphenol (phenolic) compounds. These metabolites are mainly composed of aromatic rings with hydroxyl groups capable of reacting with peroxide radicals to block degradation reactions (Hyun *et al.*, 2010). They are major and ubiquitous secondary metabolites derived from phenylalanine and shikimate pathways (Quideau *et al.*, 2011). Polyphenols range from simple molecules such as phenolic acids to more complex structures, which are highly polymerized metabolites, including tannins, flavonoids, lignans or stilbenes (Pott *et al.*, 2019).

The study of plant polyphenols started thanks to the work of four scientists: Theodore White, E. C. Bate-Smith, Tony Swain and Edwin Haslam. According to their definition of plant polyphenols, only monosubstituted phenols or compounds having di- and/or tri-hydroxyphenyl moieties can fit with the definition of “true plant polyphenol” (Haslam and Cai, 1994). Giving that, only proanthocyanidins and hydrolysable tannins, which excludes lignin polymers among others, are considered true plant polyphenols (Haslam and Cai, 1994). According to Quideau *et al.*, (2011), a polyphenol is a plant secondary metabolite derived from the phenylpropanoid-shikimate and polyketide pathways that has more than one phenolic unit and is deprived of any nitrogen-based functional group. Nowadays it is widely accepted that a polyphenol is a natural occurring compound composed by several aromatic rings substituted with hydroxyl groups. This definition comprises four principal groups of polyphenols: flavonoids, phenolic acids, lignans and stilbenes (Zhou *et al.*, 2016).

An important class of polyphenols are tannins, which have been used for ages in the treatment of animal skins for leather manufacturing, to prevent putrefaction. The reason is the ability of tannins to interact with proteins, stabilizing them and turning the skin tanned in order to be transformed into leather (Baxter *et al.*, 1997; Falcão and Araújo, 2018). The term tannin includes three families of compounds: condensed tannins (proanthocyanidins), hydrolysable tannins (gallotannins and ellagitannins) and phlorotannins (only found in marine red-brown algae). According to Zhou *et al.*, (2016) classification, condensed tannins belong to flavonoid-related compounds, while

ellagitannins belong to phenolic acid-related compounds. Tannins can be found naturally in plants and their main function is to function as a natural barrier against pathogens or herbivores (Haslam and Cai, 1994). The consumption of fruits enriched in polyphenols, such as berries, pomegranate (*Punica granatum*), persimmon (*Diospyros kaki*) and nuts are related with health benefits (Basu *et al.*, 2014; Cory *et al.*, 2018). Of course, tannins, as polyphenols, take part in crop health-promoting benefits.

6. Biosynthesis of polyphenols

Many aromatic compounds, and by extension, many antioxidant compounds, are produced by de novo synthesis via the shikimate pathway in plants (Bentley and Haslam, 1990). These aromatic compounds are produced in plastids, where this pathway is functional. Chorismate is synthesized through seven enzymatic reactions starting from phosphoenolpyruvate and erythrose-4-phosphate (**Fig. 7**) and is finally transformed into three aromatic amino acids, L-tyrosine, L-tryptophan, and L-phenylalanine (Mir *et al.*, 2015). L-phenylalanine is the precursor of antioxidant compounds like flavonoids, isoflavonoids, phenylpropanes, aurones, coumarins, and phenylpropanoid esters (Stefanachi *et al.*, 2018). Another interesting intermediate of the shikimate pathway is the 3-dehydroshikimate, which starts the biosynthesis of hydrolysable tannins (Vogt, 2010; Bontpart *et al.*, 2016).

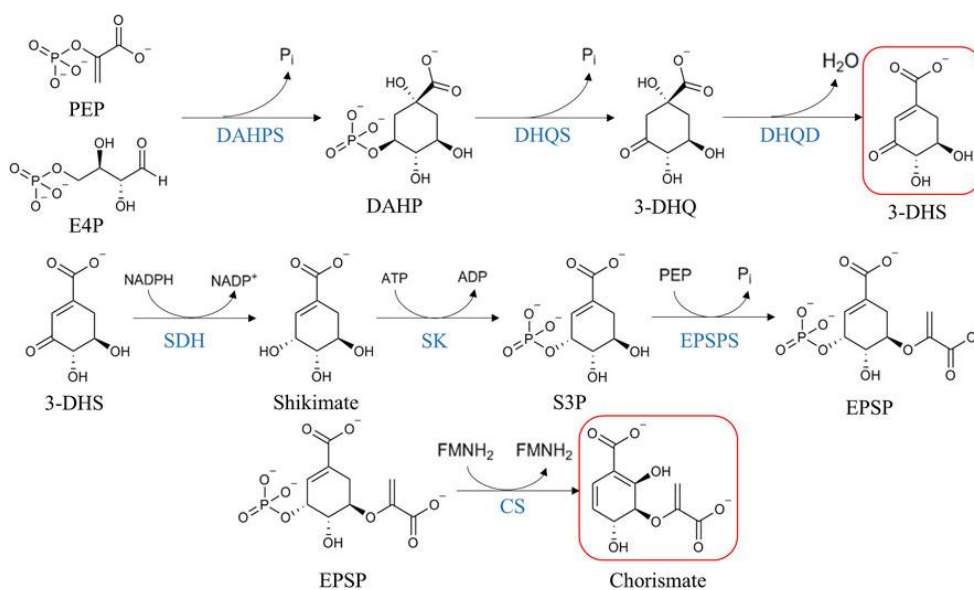


Figure 7. Schematic of shikimate pathway. Molecules highlighted in red are 3-dehydroshikimate and chorismate, precursors of hydrolysable and condensed tannins, respectively. CS: Chorismate synthase; DAHP: 3-deoxy-*D*-arabino-heptulosonate-7-phosphate; DAHPS: 3-deoxy-*D*-arabino-heptulosonate-7-phosphate synthase; DHQD: Dehydroquinate dehydratase; DHQS: Dehydroquinate synthase; 3-DHQ: 3-dehydroquinate; 3-DHS: 3-dehydroshikimate; E4P: Erythrose 4-phosphate; EPSP: 5-enolpyruvylshikimate-3-phosphate; EPSPS: 5-enolpyruvylshikimate-3-phosphate synthase; P_i: inorganic phosphate; S3P: Shikimate-3-phosphate; SK: Shikimate kinase. Taken from Mora *et al.*, (2022).

The first step of the shikimate pathway is an aldol condensation catalyzed by 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase, taking phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P) as substrates, resulting in the seven-carbon keto acid DAHP and inorganic phosphate (Mir *et al.*, 2015). In plants, DAHP synthase is regulated transcriptionally by certain environmental stimuli like mechanical wounding or biotic stresses like insects' attacks. Under these stress conditions, there is an accumulation of mRNA encoding DAHP synthase (Dyer *et al.*, 1989), strengthening the idea of the role of aromatic compounds in plant defense. In addition, DAHP synthase is controlled by feedback inhibition loops, limiting the carbon flux from primary metabolism to polyphenol synthesis (Pott *et al.*, 2019).

The next step is the transformation of DAHP into 3-dehydroquinate, catalyzed by dehydroquinate synthase (DHQS). This enzyme is a metalloenzyme nicotinamide adenine di-nucleotide (NAD^+)-dependent that acts as an enzymatic complex in the five-step transformation of DAHP. In this multistep mechanism, several reactions occur spontaneously, but only a single site catalyzes five sequential reactions. NAD^+ is required for the oxidation of C5 of DAHP and the NADH produced is then used in the reduction of the C5 carbonyl intermediate in DAHP (not shown in **Fig. 7**) (Mir *et al.*, 2015).

Then, 3-dehydroquinate is converted to 3-dehydroshikimate due to the presence of 3-dehydroquinate dehydratase (DHQase). This enzyme exists in two types: type I and type II, but with no sequential or structural homology and with no similar reaction mechanisms between them. Type I DHQase (susceptible to thermal denaturation) is found in plants, fungi and some bacteria (Chaudhuri and Coggins, 1985) and this type only participates in biosynthetic reactions, whereas type II DHQase (resistant to thermal denaturation) can participate either in biosynthetic or catabolic reactions (White *et al.*, 1990). Type I enzymes catalyze *cis*-dehydration of 3-dehydroquinate, involving the formation of a Schiff-base intermediate (Shneier *et al.*, 1991), where the conserved amino acids lysine and histidine of the active site play a critical catalytic dyad in generating the carbanion intermediate essential for the course of the reaction (Mir *et al.*, 2015).

The fourth enzyme of the shikimate pathway is the shikimate dehydrogenase (SDH) that acts in the reduction of 3-dehydroshikimate to shikimate by means of the transformation of NADPH to NADP^+ . This enzyme exists in a bifunctional complex in plants (Chaudhuri and Coggins, 1985; Coggins *et al.*, 1987). The catalysis consists of an acid-base mechanism in which a lysine and an aspartate residue form a catalytic dyad. The lysine function is to deprotonate the 3-hydroxyl group of shikimate, its amino group acting as a base (Mir *et al.*, 2015).

Shikimate kinase (SK) is the enzyme responsible for the fifth step of the shikimate pathway. This step is the phosphorylation of shikimate to shikimate-3-phosphate by

converting ATP to ADP. The SK is a simple protein (26 kDa in *A. thaliana*) that is composed of three domains: the core domain, the lid domain, and the shikimate binding domain. Within the shikimate binding domain, there are several highly conserved amino acids involved in the interaction with the substrate (Mir *et al.*, 2015).

The penultimate step of the pathway is catalyzed by 5-*enol*pyruvylshikimate 3-phosphate (EPSPS). This is a monomer protein responsible for the reversible transfer of an *enol*pyruvoyl moiety from phosphoenolpyruvate (PEP) to shikimate-3-phosphate, giving EPSP and phosphate as reversible products. It is remarkable that in this reaction, which proceeds with the cleavage of the C-O bond of PEP, not with the cleavage of the high energy P-O bond (Mir *et al.*, 2015).

Finally, the seventh and last reaction of the shikimate pathway is the one catalyzed by chorismate synthase. In this reaction, EPSPS is converted into chorismate. Although during the course of the reaction there is no redox change (Mir *et al.*, 2015), this enzyme requires the presence of FMNH₂. The presence of flavin mononucleotide may have two explanations: 1) the flavin has merely a structural role or 2) flavin restores the reduced form of sulfhydryl group structurally or catalytically important for the enzyme (Hasan and Nester, 1978).

Regulation of this pathway occurs mainly at gene expression. The cDNA of higher plants results in proteins with an amino-terminal signal sequence for plastid import. The expression of some plant genes encoding enzymes that participate in the shikimate pathway is regulated during development (Benfey and Chua, 1989; Maeda *et al.*, 2010) and by certain stimuli like wounding (Dyer *et al.*, 1989; Keith *et al.*, 1991), elicitors or pathogens (Keith *et al.*, 1991; Görlach *et al.*, 1995). At a protein level, in higher plants the enzymes possess an amino-terminal sequence required for plastid import, suggesting that the synthesis of chorismate only takes place inside the plastids (Herrmann and Weaver, 1999).



1. Plant material and growth conditions

1.1. Goodberry population

The Goodberry population (GBP) is a pseudo full-sibling F_1 population of 126 individuals derived from the crossing of two commercial strawberry varieties *F. × ananassa* cv ‘Candonge’ and cv. ‘Senga Sengana’, adapted to contrasting Southern and Northern European zones, respectively. The same population was grown in five different locations from across Europe: National Institute of Horticultural Research, (INHORT), Skierniewice, Poland (51°91’N 20°05’E); Hansabred GmbH & Co. KG, Dresden, Germany (51°09’N 13°47’E); Università Politecnica delle Marche, Ancona, Italy (43°32’N 13°22’E); INVENIO, Douville, France (45°02’N 0°61’E); and ‘El Cebollar’, Instituto Andaluz de Investigación y Formación Agraria y Pesquera (IFAPA), Huelva, Spain (37°14’N 06°48’W). Cultivation was carried out following conventional practices at each location. For each F_1 individual, 10 plants were grown in open field or under plastic tunnels, except for France where they grow in coconut fiber in a soil-less tunnel. A subset of 35 F_1 lines or the whole population was grown during 2016-2017 and 2017-2018 seasons. Ripe fruits were collected in two different harvesting times during the peak of the season. The two harvesting time were named A and B at each location in three biological replicates of at least six fruits each. The collected fruits were immediately frozen in liquid nitrogen and stored at -80°C. Samples were grounded until fine powder was obtained.

1.2. *Fragaria × ananassa* ‘232’ and ‘1392’ mapping population

The octoploid strawberry mapping population used for ellagitannins content analysis was a full-sibling family of 95 F_1 individuals derived from intraspecific cross between the breeding lines ‘232’ and ‘1392’ previously characterized for agronomic and fruit quality traits (Zorrilla-Fontanesi *et al.*, 2011). Six plants of each F_1 and parental lines were vegetative propagated and grown under commercial conditions in Moguer (Huelva, Spain) during two consecutive years.

The mapping population was previously characterized due to variation on agronomical traits (Zorrilla-Fontanesi *et al.*, 2011) as well as variation in volatile (Zorrilla-Fontanesi *et al.*, 2012), primary metabolites (Vallarino *et al.*, 2019) and secondary semi-polar metabolites (Pott *et al.*, 2020) in ripe fruits. About 25 fully mature fruits from all F_1 individuals and parental lines were harvested the same morning in one day at the pick of the harvest of the two consecutive seasons. Per line, all selected fruits were phenotypically similar for shape, size, firmness and color. Fruit from each line was pooled into three biological replicates, immediately frozen, ground in liquid nitrogen, and stored at -80 °C until analysis.

1.3. *Fragaria* × *ananassa* cultivars ('Camarosa' and 'Candonga')

The *F. × ananassa* cv. 'Camarosa' or 'Candonga' fruits were harvested in four different developmental stages corresponded to: green, G; white, W; turning (at least 25% surface red), T and red, R. Plants were grown in greenhouse conditions in IFAPA, Churriana (Spain). Pools of fruits from a single plant have been considered single biological replicates (80 plants were used in total). All fruits were frozen immediately in liquid nitrogen and achenes were removed using a scalpel on frozen fruits.

1.4. *Nicotiana benthamiana* plants

Plants were grown in chambers under long-day conditions (16 h light – 8 h dark) and a temperature of 22°C.

2. Development of linkage maps and QTL analysis

2.1. Volatile organic compounds in Goodberry population ('Candonga' × 'Senga Sengana')

This work was previously done before the start of this PhD thesis in the context of the EU initiatives -H2020 Goodberry project (<https://goodberry-eu.eu/>).

The linkage map of *F. × ananassa* (Can × SS) was constructed using single-dose markers for single chromosome out of the seven chromosomes of the parent genomes using both the strawberry DArTseq platform and the 50k Fana Axiom array (Hardigan *et al.*, 2020). JoinMap 5.1 software (Van Ooijen, 2011; Prohaska *et al.*, 2024) was used to construct the genetic map with single-dose markers from the single nucleotide polymorphism (SNP) marker Affimetrix array (Hardigan *et al.*, 2020) that segregated 1:1 (Rousseau-Gueutin *et al.*, 2008). Linkage groups were chosen from a logarithm of odds (LOD) higher than 10 and genetic maps were constructed following maximum likelihood algorithm (parameters described in Labadie *et al.*, 2022). DArTseq and Axiom SNP marker sequences were blasted to the published 'Camarosa' v1.0 genome (Edger *et al.*, 2019) and assigned to the best-matching chromosome. LGs within homologous groups were then named from LGI to LGVII, followed by a number in reference to species-derived subgenomes.

Both linkage parental maps (male and female) based on the F₁ segregating population were used separately for QTL analysis. Detection of mQTL (metabolic QTL) was performed by simple interval mapping using the R/QTL package (Broman *et al.*, 2003). Only significant QTLs were considered when their correspondent LOD scores were higher than LOD threshold at p-value ≤ 0.05. The proportion of phenotypic variance explained by a single QTL was calculated as the square of the partial correlation coefficient (r²).

2.2. Ellagitannins in ‘232’ × ‘1392’ mapping population

This work was previously done before the start of this PhD thesis, summarized and explained detailed in Delphine Pott’s PhD thesis (<https://hdl.handle.net/10630/18672>) and Vallarino *et al.*, (2019).

The map used for QTL detection was previously developed for the ‘232’ x ‘1392’ mapping population (Sánchez-Sevilla *et al.*, 2015). This map comprises 2,089 SNP and SSR markers, covering a total length of 2,489 cM. QTL analysis was conducted using MapQTL 5, and the raw data were initially analyzed with the nonparametric Kruskal-Wallis rank-sum test. Marker-QTL linkages were identified using a significance threshold of $p\text{-value} \leq 0.005$. For traits with non-normal distributions, mQTLs were identified and localized using interval mapping, with a step size of 1 cM and up to five neighboring markers. LOD significance thresholds were determined via 1,000-permutation tests for each trait. The most significant markers were subsequently employed as co-factors for restricted multiple QTL mapping (rMQM) analysis. mQTLs with LOD scores exceeding the genome-wide threshold at $p\text{-value} \leq 0.005$ were considered significant.

3. Transcriptome analysis (RNA-seq)

Paired-end libraries with an approximate insert size of 300 bp were prepared for each sample and sequenced on HiSeq2000 lanes using 2×100 bp reads. On average, 33.97 million reads were obtained per sample, ranging from 30.83 to 36.75 million reads. Raw RNA-seq reads were processed to remove low-quality nucleotides and then aligned using HISAT2 v2.1.0 (Kim *et al.*, 2015). For the contrasting ellagitannin lines, the reads were mapped to the *Fragaria vesca* v4.0.a1 genome assembly (Edger *et al.*, 2018), while for the GBP, they were aligned to the *Fragaria* × *ananassa* ‘Camarosa’ Genome Assembly v1.0 (Edger *et al.*, 2019). Default settings of HISAT2 were applied, permitting up to 40 multiple alignments per read and allowing a maximum of two mismatches during mapping.

Aligned read files were further processed with Cufflinks v2.2, following established protocols (Sánchez-Sevilla *et al.*, 2014). Reads were assembled into transcripts, which were categorized as either known (corresponding either to annotated genes in the *F. vesca* v4.0.a1 genome annotation or to annotated genes in the *F. × ananassa* ‘Camarosa’ Genome Assembly v1.0) or novel. Transcript abundance was quantified, and differential expression analyses were conducted between the samples. RNA-seq fragment counts were normalized and used to estimate the relative transcript abundances, expressed as fragments per kilobase of exon per million fragments mapped (FPKM) and as counts per million (CPM). All bioinformatics analyses were carried out at the Supercomputing and Bioinnovation Center (SCBI) in Málaga, Spain.

We conducted two independent RNA-seq analyses: one for lines with contrasting γ -dodecalactone content and another for lines with contrasting ellagitannins content.

3.1. Goodberry population for contrasting γ -dodecalactone content

We used F₁ population derived from the crossing between ‘Candonge’ and ‘Senga Sengana’ varieties (Goodberry population) for their organoleptic characteristics. Among them, the volatile organic content was selected for further experiments. Specifically, we selected lines for RNAseq analysis for their contrasting levels in γ -dodecalactone content.

Among the 109 F₁ lines, we selected 33 lines for transcriptome analysis based on their major heterozygosity. These lines were: H015, H016, H025, H032, H036, H037, H040, H048, H050, H051, H053, H054, H058, H061, H067, H069, H071, H073, H076, H083, H087, H094, H097, H098, H103, H110, H112, H114, H120, H122, H124, H125 and H126. Among these F₁ lines, we selected pool of ripe fruits contrasting in mean γ -dodecalactone content in the five countries. Total RNA was extracted from 5 lines with high γ -dodecalactone content (H016, H036, H054, H071 and H126) and 5 lines with low γ -dodecalactone content (H040, H050, H051, H098 and H112), all grown in Spain and France. Total RNA extraction was carried out following the protocol described below in this section.

3.2. ‘232’ × ‘1392’ mapping population for contrasting ellagitannins content

Semi-polar metabolites content of the ‘232’ × ‘1392’ F₁ mapping population were measured as previously described in Pott *et al.*, 2020 for the two harvesting seasons (2013 and 2014). After analyzing QTL regions associated to ellagitannins content as previously described (**section 2.2**), Following this analysis, we selected pool of ripe fruits with lines showing contrasting ellagitannins levels.

Total RNA was extracted from 5 lines with low ellagitannins content (L13-13, L62-13, L13-14, L41-14, L45-14) and 5 lines with high content (L19-13, L51-13, L19-14, L27-14, L51-14) of these compounds. RNA was extracted as explained in the next section.

4. RNA extraction, cDNA synthesis and gene expression analysis

Total RNA was isolated from strawberry fruit pools according to the method described by Gambino *et al.*, (2008) with minor modifications. Around 300 mg of ground fruit tissue was treated with 900 μ L of RNA extraction buffer composed by 100 mM Tris-HCl pH 8.0, 2 M NaCl, 25 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 2% cetyltrimethylammonium bromide (CTAB), 2.5% (m/v) polyvinylpyrrolidone (PVP-40), 2% (v/v) 2-mercaptoethanol and incubated at 65°C for 10 min. Two washing steps were performed consecutively by adding equal volumes of chloroform:isoamyl alcohol (24:1 v/v), vortexing and centrifuging at 11,000g for 10 min at 4°C. First RNA precipitation is

carried out by adding LiCl at a final concentration of 3 M, followed by a 30 min incubation at 4°C. RNA was pelleted by centrifugation at maximum speed for 20 min at 4°C and resuspended in 500 µL SSE buffer composed of 10 mM Tris-HCl pH 8.0, 1 M NaCl, 1 mM EDTA and 0.5 % (m/v) of sodium dodecyl sulphate (SDS). A final extraction step with chloroform:isoamyl alcohol was conducted, and the supernatant was precipitated with 0.7 volumes of isopropanol. After a centrifugation step at maximum speed for 15 min at 4°C, pellets were washed with 70% (v/v) ethanol and resuspended in 30 µL of H₂O. Then, DNase treatment was performed using TURBO™ DNase (Thermo Fischer) following manufacturer's instructions. Total RNA was quantified spectrophotometrically with micro-spectrophotometer (NanoDrop Technologies) and integrity was checked by loading a small quantity (around 100 ng) of total RNA to an electrophoresis agarose 1% (m/v) gel under denaturing conditions. Absence of DNA in the samples was checked by using RNA extractions as template for PCR with primers for housekeeping genes.

Retrotranscription was performed from 500 ng of RNA using reverse transcriptase reaction (RT) from iScript™ cDNA synthesis kit (Bio-Rad) following manufacturer's instructions. Integrity of cDNA was assayed by using cDNA as template for PCR with primers for housekeeping genes.

Transcript quantification of selected genes was analyzed by quantitative real-time PCR (qRT-PCR). This analysis was carried out in CFX96 or CFX384 detection systems (Bio-Rad) from 5 ng of cDNA using SsoFast™ EvaGreen supermix (Bio-Rad) according to manufacturer's instructions. Three technical replicates were used for each biological replicate. The C_q values were calculated using LinRegPCR software from relative fluorescence units (RFU) raw data. Expression data were normalized to the reference genes *FaDBP* and *FaGAPDH2* (Amil-Ruiz *et al.*, 2013; Landi *et al.*, 2014). Finally, relative gene expression was calculated using the Livak or 2^{-ΔΔC_q} method (Livak and Schmittgen, 2001). Specific primers are listed on **Table 6**.

5. Polymerase Chain Reaction (PCR) reactions

For the different purposes of this work, two polymerases were used. For confirmation of plasmid constructs, bacterial transformation and transgenic plant genotyping and for checking cDNA integrity, GoTaq® DNA Polymerase (Promega) was used following manufacturer's recommendations. On the other hand, for coding sequences (CDS) amplifications for further plasmid cloning, the proofreading iProof™ High-Fidelity DNA Polymerase (Bio-Rad) was employed following manufacturer's instructions.

6. Competent Bacterial cells preparation

6.1. *Escherichia coli* TOP10 heat shock competent cells

E. coli TOP10 strain was refreshed on liquid pre-culture of Luria-Bertani (LB) medium and grown overnight at 37°C and 180 rpm of orbital shaking. This pre-culture was used to initiate another liquid culture of LB with the same conditions and when the culture reached O.D₆₀₀= 0.25, cells were harvested by centrifugation at 3,000g for 10 min at 4°C. Cells were resuspended in cold sterile 0.1 M MgCl₂ solution.

After 30 min incubation on ice, cells were centrifuged at 3,000g for 10 min at 4°C. Harvested cells were resuspended in cold sterile TG medium (75 mM CaCl₂, 6 mM MgCl₂, 15% (v/v) glycerol). Cell suspension was incubated in cold overnight and then aliquoted in 1.5 mL tubes.

6.2. *Agrobacterium tumefaciens* GV3101::pMP90 electrocompetent cells

The *A. tumefaciens* GV3101::pMP90 strain was refreshed on plate with 50 µg/ml rifampicin and 25 µg/ml gentamycin and grown overnight at 28°C and 180 rpm of orbital shaking. Next day, a liquid culture of LB without antibiotics was inoculated with biomass from the culture plate and grown overnight at the same conditions as before. When the liquid culture reached an optical density at 600 nm (O.D₆₀₀) of 0.8-1, cells are collected by centrifugation at 5,500g for 10 min at 4°C. Cells are resuspended twice in cold sterile H₂O decreasing the total volume by half and centrifuged twice as before. Finally, cells are resuspended in sucrose-glycerol buffer (272 mM sucrose, 15% (v/v) glycerol) and aliquoted in 1.5 ml tubes. Competent cells remained viable for use up to 6 months at -80°C.

6.3. *Agrobacterium tumefaciens* AGL1 heat shock competent cells

The *A. tumefaciens* AGL1 strain was refreshed on plate with 10 µg/ml rifampicin and grown overnight at 28°C and 180 rpm of orbital shaking. Then, a liquid culture of LB with the same antibiotic and the same concentration was inoculated with biomass from the culture plate and grown until culture reached O.D₆₀₀ of 0.3-0.4. Cell suspension was centrifuged at 3,000g for 5 min at 4°C. Cells were resuspended in cold sterile 20 mM CaCl₂ solution. Cell suspension was finally aliquoted in 1.5 mL sterile tubes. Competent cells remained viable for use up to a month at -80°C.

7. Searching and characterization of a candidate gene associated with γ -dodecalactone content

7.1. Identification of the candidate gene *FxaC_10g48980 (FaLOX2)*

Molecular markers used to detect the QTLs were associated with the levels of γ -dodecalactone in the mapping population. This relationship was used to select lines with

contrasting γ -dodecalactone content and further analyses gene expression between these two contrasting groups.

Functional annotation was performed for genes in the selected QTL intervals using Mercator4 tool (Bolger *et al.*, 2021). Then, to reduce the number of candidates, only genes with putative function related to lipids or fatty acid metabolism were selected to further analyze their expression in these contrasting groups. Therefore, candidate genes were selected following gene expression and functional annotation criteria.

7.2. Cloning of candidate gene *FaLOX2* and its tomato homolog *TomloxC*

The full coding sequence of the candidate gene (*FxaC_10g48980*) was PCR-amplified from strawberry fruit (*F. × ananassa* cv. ‘Camarosa’) cDNA using specific primers (Table 7). The amplified sequence was introduced in pK7WG2D (Table 7 and Annex 1) (Karimi *et al.*, 2002) plasmid with GatewayTM cloning technology (Invitrogen) for transient and stable overexpression of the candidate *FaLOX2*. This construct overexpressed the candidate *FaLOX2* gene under the control of a 35S cauliflower mosaic virus promoter (35SCaMV) including also spectinomycin resistance as bacterial selection agent and kanamycin resistance as plant selection agent.

In the case of the homolog *TomloxC*, a construct based on pK7GWIWG2D(II) vector (Table 7 and Annex 1) was used (Karimi *et al.*, 2002). For this purpose, a fragment of 300 bp from conserved region of *Solyc01g006540* (*TomloxC*) was amplified via PCR using specific primers (Table 6). This plasmid was composed of two inverted repeats for GatewayTM cloning under the control of a 35SCaMV, leading to the formation of an RNA hairpin capable of silencing the *TomloxC* gene via RNAi. This plasmid also has a spectinomycin resistance gene as bacterial selection agent and a kanamycin resistance gene for selection *in planta*.

7.3. Transient overexpression of ‘Candonga’ strawberry fruits

The construct pK7GW2D-*FaLOX2* was introduced in *A. tumefaciens* GV3101::pMP90 as well as the empty vector pK7GW2D. Agroinfiltration with either the overexpressing construct or the empty vector as control were carried out mainly during the month of June. Briefly, *A. tumefaciens* carrying the correspondent plasmid were grown overnight in LB medium supplemented with 50 µg/ml rifampicin, 25 µg/ml gentamycin and 50 µg/ml spectinomycin until the culture reached O.D₆₀₀ = 0.8. Cultures were then centrifuged at 5,000g for 5 min at room temperature and cells were resuspended in agroinfiltration solution composed of 10 mM MgCl₂, 10 mM (2-(*N*-morpholino)ethanesulfonic acid (MES) pH 5.6 and 1 mM acetosyringone. Also, *p19* was used as suppressor of plant silencing machinery to avoid plant defense against the high number of transcripts produced by the 35SCaMV of both plasmids. For the co-infiltrations, pK7GW2D-

FaLOX2 or pK7GW2D were used at a final O.D₆₀₀= 0.6 while *p19* was used at a final O.D₆₀₀= 0.3.

Transient fruit overexpression was carried out on octoploid strawberry cv. ‘Candonga’ as previously described in Hoffmann *et al.*, (2006). Fruits were injected in green to early-white ripening stage with 100-500 µL until fruits were fully infiltrated. Fruits were harvested when fully ripe (red stage) and directly frozen in liquid N₂. Frozen fruits were pulverized independently with the help of a TissueLyser II (Qiagen). After that, RNA extraction, qRT-PCR and metabolites extractions were performed as described in this section.

7.4. Generation of stable transgenic lines in *Solanum lycopersicum* cv. ‘MoneyMaker’

Here, two different transgenic lines were generated in tomato (*Solanum lycopersicum*): (I) stable overexpression of *FaLOX2*, and (II) silencing using RNAi construct for the homologous *TomloxC*.

For that purpose, sterile tomato seeds (cv. ‘MoneyMaker’) were germinated on germination medium, composed by 4.4 g/L Murashige-Skoog including Gamborg B5 vitamins (MS), 15g/L sucrose and pH 5.9. Once cotyledons are fully expanded, they were excised at petiole and cut in half. Those explants were placed in 2Z- medium (4.4 g/L MS, 100 mg/L *myo*-inositol, 20 g/L sucrose, 2 µg/ml *trans*-zeatin, 0.2 mM acetosyringone, pH 5.8) during 24h at 25°C.

A. tumefaciens GV3101::pMP90 containing the proper construct (pGW7GWIWG2D(II)-*TomloxC*) (**Annex 1** and **Table 6**) were grown in LB medium supplemented with 50 µg/ml spectinomycin for bacterial selection. When liquid culture O.D₆₀₀ reached 0.6, the cells are collected by centrifugation at 4,500g for 20 min at 22°C. Cells were resuspended in MS-0.2% (4.4 g/L MS, 20 g/L sucrose, 100 mg/L *myo*-inositol, 0.4 µg/ml thiamine, 0.2 mM acetosyringone, pH 5.8). Tomato explants were incubated with cellular suspension of *A. tumefaciens* containing the correspondent plasmid during 5 min and gentle shaking. Then, the explants were placed in 2Z- plates with adaxial side down during in dark during 48 h at 25°C.

After 48 hours, the explants were placed in 2Z+K medium (4.4 g/L MS, 20 g/L sucrose, 100 mg/L *myo*-inositol, 2 µg/ml *trans*-zeatin, 350 mg/L timentin, 0.1 mg/L indolacetic acid, 150 mg/L kanamycin, pH 5.8) for plant selection until callus and shoots formation.

In both cases, some explants were incubated with sterile water and placed in medium without kanamycin as selective agent for regeneration positive controls and some of them were placed in medium with kanamycin for regeneration negative control.

7.5. Volatile organic compounds during ripening process

Plants of *F. × ananassa* cv. ‘Camarosa’ grown in greenhouse conditions (IFAPA, Churriana) were selected in three different maturation stages.

Five independent fruits were collected in triplicate for each developmental stage: green, white and red. These fruits were frozen in liquid nitrogen, and their achenes were removed mechanically with a scalpel blade.

Finally, 3 different pools of 5 independent fruits were established for each developmental stage. Volatile organic compounds of these samples were extracted and analyzed as described later.

8. Searching and characterization of a candidate gene associated with ellagic acid hexose content

8.1. Identification of candidate gene *FxaC_3g11070* (*FaUDPSS*)

The physical region of QTLs for ellagitannins was studied by mapping the limits in the *F. vesca* genome v4.0.a2 (Li *et al.*, 2019). Functional annotation using Mercator4 (Bolger *et al.*, 2021) and RNA-seq data generated for this purpose were used to select some genes for further study.

Protein sequences of selected genes were used to make a BLAST against ‘Camarosa’ genome v1.0.a1 (Edger *et al.*, 2019). Gene expression of *F. vesca* and *F. × ananassa* genes were compared in individuals with different ellagitannins content. Also, the expression of selected genes was also studied with contrasting lines in flavan-3-ols (monomers of proanthocyanidins or condensed tannins).

Finally, the functional annotation of genes was considered. We searched for genes with glucosyltransferase functions, making it possible to discover some genes annotated as glucosyltransferases of anthocyanidins, for example.

8.2. Cloning of *FxaC_3g11070* versions and sequencing

Several plant tissue, mainly red fruit, were collected for amplifying *FxaC_3g11070* gene. PCR amplification was accomplished by using *UTR-FaUDP* primers (Table 6) due to the conservation of the UTR zones in the case of this gene.

In all cases, gene sequences were cloned into pGEM®-T Easy vector (Promega), using manufacturer’s instructions. Gene sequencing was performed by Sanger sequencing by STAB vida services with *M13* primers (Table 6). Nucleotide sequences were then aligned using MAFFT algorithm.

8.3. Generation of stable transgenic *Fragaria* × *ananassa* cv. ‘Candonga’ lines overexpressing *FaUDPSS*

In vitro strawberry cv. ‘Candonga’ plants were used for this transformation. 4x4 mm explants were excised from young leaves and abaxial-side-down disposed on plates with N₃₀ supplemented with phytohormones (BI medium) as follows: 0.1 mg/l kinetine, 0.5 mg/l indolbutyric acid (IBA) and 2 mg/l benzyl adenine (BA). These explants were cultivated at room temperature and in absence of light for 48 hours.

A. tumefaciens AGL1 containing the proper construct (pK7GW2D-*FaUDPSS*) were grown in LB medium supplemented with 50 µg/ml spectinomycin for bacterial selection. When liquid culture O.D₆₀₀ reached 0.5 units, the cells are collected by centrifugation at 4,500g for 20 min at 22°C. Cells were resuspended in sterile liquid N₃₀ medium without hormones. ‘Candonga’ explants were incubated with the cellular suspension of *A. tumefaciens* AGL1 containing the *FaUDPSS*-overexpressing construct during 15 min at room temperature and gentle shaking. Explants were then dried and cultivated on BI medium at room temperature and in dark conditions.

After 48 hours, the explants were placed in BI medium supplemented with 50 mg/l kanamycin and 250 mg/l timentin for strain selection. After 72 hours, explants on selection medium are cultivated in light and the medium was refreshed every 4 weeks. When shoots emerged from callus, they are subcultivated individually until fully development was achieved.

In both cases, some explants were incubated with sterile water and placed in medium without kanamycin as selective agent for regeneration positive controls and some of them were placed in medium with kanamycin for regeneration negative control.

8.4. TurboID proximity labelling of *FaUDPSS* and liquid-chromatography/mass spectrometry-mass spectrometry (LC-MS/MS)

The protocol used was a modified version of Tan *et al.*, (2024). Briefly, construct based on the pGWB_{TID1} vector (**Table 7**) was performed for *FaUDPSS*. This construct expressed the cassette *3xHA-TID-FaUDPSS* under the control of a *35SCaMV* promoter. Both empty pGWB_{TID1} and pGWB_{TID1}-*FaUDPSS* were transformed into *A. tumefaciens* GV3101::pMP90 and then used to transient fruit transformation *F. × ananassa* cv. ‘Camarosa’ as described in this section. Fruits were harvested in red stage and immediately frozen in liquid nitrogen. Ground fruit tissue was stored for TurboID proximity labelling experiments.

Desalting columns (PD-10 columns, Cytiva 17-0851-01) were previously equilibrated using equilibration solution (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% 1(v/v)

glycerol, 5 mM EDTA and 2 mM DTT) according to the manual. The equilibrated columns were then conserved at room temperature until use.

About 1 g of fruit tissue is transferred to a 5 mL tube and 2 mL of protein extraction buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1 % (v/v) NP-40). The tubes were vortexed thoroughly and incubated on a rotator at 4°C for 30 min. The tubes were centrifuged at 16,000g and 4°C for 15 min and the supernatant was transferred to a new 2 mL tube. The protein extraction was added to the center of the equilibrated desalting column, and these columns were put in a 50 mL tube. The tubes were centrifuged at 1,000g and 4°C for 2 min when the protein extraction entered in the resin. The flow through was transferred to a new 2 mL tube and put on ice temporarily. About 100 µL of desalted protein extract was taken to serve as an input for western blot.

The streptavidin-C1-conjugated magnetic beads (Dynabeads™ MyOne™ Streptavidin C1, Invitrogen) were washed twice at room temperature in the protein extraction buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, 5 mM EDTA, 5 mM DTT, 1 mM PMSF, 1% (v/v) NP-40). For each sample, 80 µL of beads were used. After each washing step, the beads are collected using a magnetic rack at room temperature for at least 2 min and the solution was gently removed by pipetting. The desalted protein extracts were transferred to the pre-washed beads and incubated at 4°C for 3-4 hours under rotary agitation. The beads were captured using a magnetic rack for 4 min at room temperature and the beads were washed as follows:

- 1.7 mL of wash buffer I (0.5 % SDS) was added to the tube and keep it on a rotator at room temperature for 8 min. The beads were absorbed as before.
- 1.7 mL of wash buffer II (50 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% (v/v) Triton-X-100, 0.5% (m/v) Na-deoxycholate) was added to the tube and keep it on a rotator at room temperature for 8 min. The beads were absorbed as before.
- 1.7 mL of wash buffer III (10 mM Tris-HCl pH = 7.5, 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.5 % Na-deoxycholate) is added to the tube and keep it on a rotator at RT for 8 min. The beads are absorbed as before.
- 1.7 mL of 50 mM Tris-HCl (pH = 7.5) was added to the tube to remove the detergent. Repeat once. The beads were absorbed as before.
- The beads were finally washed three times in 1X PBS solution at room temperature for 5 min. The beads were absorbed as before.
- 1 mL of 1X PBS was added to the tube to resuspend the beads. About 100 µL of beads were taken for western blot to confirm the enrichment of biotinylated proteins.

- The rest of the beads were flash-frozen in liquid N₂ and stored at -80°C or sent immediately for LC-MS/MS analysis.

Samples were analyzed using LC-MS/MS VanquishNeo HPLC coupled to Exploris489 (both Thermo Fisher Scientific). Data processing was performed using MaxQuant software (Version 2.2.0.0 with integrated Andromeda Peptid search engine), with *Fragaria* × *ananassa* genome downloaded from GDR database (Fragaria_ananassa_v1.0.a2.proteins). Sequences for bait proteins were added to the search, iBAQ values were enabled, label-free quantification (LFQ) settings selected, match between runs enabled only within triplicates.

Statistical analysis was performed with Perseus (vs. 1.6.15.0) with log₂ LFQ transformed data and Student's *t*-test based false discovery ratio (FDR) 0.05.

8.4.1. Biotinylation assay of TID-FaUDPSS in *N. benthamiana* leaves

A. tumefaciens GV3101::pMP90 containing the constructs *FaUDPSS-TID1* and empty pGWBTID1 were grown in LB medium with the appropriate antibiotics. Cells were collected via centrifugation at 4,000g for 5 min and they were re-suspended in agroinfiltration solution (previously described) to a final O.D₆₀₀ = 0.5.

Fully expanded leaves from *N. benthamiana* were agroinfiltrated with 1 mL needleless syringe. For each replicate, three plants were selected and two leaves per plant were infiltrated. After 42 hours post-infiltration, the biotin solution (50 µM biotin, 2.5 mM MgCl₂) was infiltrated into the leaves previously agroinfiltrated with the constructs of interest.

After six hours of biotin treatment, the leaves were collected. The leaf middle vein was removed, and the rest of the material was flash-frozen in liquid nitrogen. Then, the material was grounded to a fine powder until protein extraction.

9. Metabolomic analysis by mass spectrometry techniques

9.1. Volatile organic compounds (VOCs) extraction and analysis by Headspace-Gas Chromatography/Mass Spectrometry (HS-GC/MS)

For volatile organic compound analysis, a total of 1 g of frozen fruit tissue was treated with 1 mL 20% (m/v) NaCl and incubated 5 min at room temperature. The homogenate was then centrifuged at 5,000g for 5 min at room temperature and the supernatant is added to the screw-cap headspace vials with 500 mg NaCl previously weighted. Finally, 1 µL of 1 ppm D₃₂-pentadecane (Cambridge Isotope Laboratories) as internal standard.

The volatiles were sampled by headspace-solid phase microextraction (HS-SPME) with a 65 µm polydimethylsiloxane/divinyl-benzene fiber (Supelco). Initially, vials were tempered at 50°C for 10 min. Then the volatiles were extracted by exposing the fiber to

the vial headspace for 30 min under continuous agitation and heating at 50°C. The extracted volatiles were desorbed in the GC injection port (Trace 1500, Thermo Fischer Scientific) for 2 min at 250°C in splitless mode. Incubation of the vials, extraction and desorption of the volatiles were performed automatically by a Triplus RSH autosampler (Thermo Fischer Scientific). Chromatography was performed on a DB-5ms (60mx 0.25mm X1 µm) column with helium as a carrier gas at constant flow of 1 ml/min. GC interface and MS source temperatures were 260°C and 230°C respectively. Oven temperature conditions were 40°C for 3 min, 8°C/min ramp until 250°C and then held at 250°C for 5 min. Mass spectra (ITQ900, Thermo Fischer Scientific) were recorded in scan mode in the 35-220 at an ionization energy of 70 eV and with a scanning speed of 7 scans/s.

Chromatograms derived from the HS-GC/MS were analyzed with Xcalibur™ (Thermo Fischer Scientific) software by using a database of volatile organic compounds (**Table 2**). Peak areas were normalized by comparing with the peak area of the internal standard (D₃₂-pentadecane) to correct variations in detector sensitivity and fiber aging.

Table 2. Volatile organic compounds (VOCs) analyzed. The table shows the retention times (RT) in minutes where the compound was detected, and the mass-charge ratio (m/z) used for quantification of the compound

Compound	Chemical formula	Chemical class	RT (min)	m/z
Acetone	C ₃ H ₆ O	Ketone	6.30	43.00
Ethyl acetate	C ₄ H ₈ O ₂	Ester	9.06	43.00
2-methoxy butanal	C ₅ H ₁₀ O ₂	Aldehyde	9.26	39.00
1-penten-3-ol	C ₅ H ₁₀ O	Alcohol	9.55	57.00
Isopropyl acetate	C ₅ H ₁₀ O ₂	Ester	10.36	61.00
2-pentanone	C ₅ H ₁₀ O	Ketone	11.10	86.00
Pentanal	C ₅ H ₁₀ O	Aldehyde	11.50	58.00
Isobutyl acetate	C ₆ H ₁₂ O ₂	Ester	12.00	43.00
Methyl butyrate	C ₅ H ₁₀ O ₂	Ester	12.09	74.00
3-hexenal	C ₆ H ₁₀ O	Aldehyde	14.20	69.00
Ethyl butanoate	C ₆ H ₁₂ O ₂	Ester	14.20	71.00
Hexanal	C ₆ H ₁₂ O	Aldehyde	14.34	83.00
Butyl acetate	C ₆ H ₁₂ O ₂	Ester	14.55	43.00
Butanoic acid	C ₄ H ₈ O ₂	Ester	15.20	71.00
Methyl butanoate	C ₅ H ₁₀ O ₂	Ester	15.20	71.00
Methyl pentanoate	C ₆ H ₁₂ O ₂	Ester	15.20	74.00
1-penten-3-one	C ₅ H ₈ O	Ketone	15.70	55.00
<i>trans</i> -2-hexen-1-al	C ₆ H ₁₀ O	Aldehyde	15.79	83.00
(<i>E</i>)-2-heptenal	C ₇ H ₁₂ O	Aldehyde	15.80	83.00
<i>trans</i> -2-pentenal	C ₅ H ₈ O	Aldehyde	15.80	55.00

Compound	Chemical formula	Chemical class	RT (min)	m/z
<i>trans</i> -2-hexen-1-ol	C ₆ H ₁₂ O	Alcohol	15.90	83.00
1-hexanol	C ₆ H ₁₄ O	Alcohol	16.03	56.00
Isopentyl acetate	C ₇ H ₁₄ O ₂	Ester	16.20	70.00
2-heptanone	C ₇ H ₁₄ O	Ketone	16.60	58.00
2-heptanol	C ₇ H ₁₆ O	Alcohol	16.8	45.00
Heptanal	C ₇ H ₁₄ O	Aldehyde	16.97	81.00
Pentyl acetate	C ₇ H ₁₄ O ₂	Ester	17.09	61.00
3-methyl-2-butenyl acetate	C ₇ H ₁₂ O ₂	Ester	17.29	67.00
Methyl hexanoate	C ₇ H ₁₄ O ₂	Ester	17.41	74.00
Methyl 2-hexenoate	C ₇ H ₁₂ O ₂	Ester	18.46	97.00
Acetophenone	C ₈ H ₈ O	Ketone	18.85	105.00
Benzaldehyde	C ₇ H ₆ O	Aldehyde	18.86	105.00
Myrcene	C ₁₀ H ₁₆	Terpene	18.90	105.00
β-pinene	C ₁₀ H ₁₆	Terpene	18.90	93.00
6-methyl-5-hepten-2-one	C ₈ H ₁₄ O	Ketone	18.95	108.00
Pentyl butanoate	C ₉ H ₁₈ O ₂	Ester	19.07	89.00
Hexyl butanoate	C ₁₀ H ₂₀ O ₂	Ester	19.08	89.00
Butyl butanoate	C ₈ H ₁₆ O ₂	Ester	19.10	89.00
Ethyl hexanoate	C ₈ H ₁₆ O ₂	Ester	19.14	99.00
2-pentylfuran	C ₉ H ₁₄ O	Furan	19.19	138.00
Decane	C ₁₀ H ₂₂	Alkane	19.28	57.00
(<i>Z</i>)-3-hexenyl acetate	C ₈ H ₁₄ O	Ester	19.32	67.00
Nonyl acetate	C ₁₁ H ₂₂ O ₂	Ester	19.40	56.00
Hexyl acetate	C ₈ H ₁₆ O ₂	Ester	19.41	61.00
2-hexen-1-ol acetate	C ₈ H ₁₄ O	Ester	19.46	82.00
(<i>E</i> , <i>Z</i>)-2,4-heptadienal	C ₇ H ₁₀ O	Aldehyde	19.70	81.00
α-pinene	C ₁₀ H ₁₆	Terpene	20.10	92.00
Camphene	C ₁₀ H ₁₆	Terpene	20.30	121.00
Limonene	C ₁₀ H ₁₆	Terpene	20.37	68.00
Mesifurane	C ₇ H ₁₀ O ₃	Ketone	20.60	142.00
(<i>E</i>)-2-octenal	C ₈ H ₁₄ O	Aldehyde	20.73	83.00
1-octanol	C ₈ H ₁₈ O	Alcohol	20.82	69.00
2-nonanone	C ₉ H ₁₈ O	Ketone	21.30	58.00
Nonanal	C ₉ H ₁₈ O	Aldehyde	21.60	56.00
Linalool	C ₁₀ H ₁₈ O	Terpene	21.62	93.00
Methyl benzoate	C ₈ H ₈ O ₂	Ester	21.80	105.00
Methyl octanoate	C ₉ H ₁₈ O ₂	Ester	21.90	87.00
<i>trans</i> -2-decenal	C ₁₀ H ₁₈ O	Aldehyde	22.30	70.00
(<i>E</i>)-2-nonenal	C ₉ H ₁₆ O	Aldehyde	22.80	81.00
Benzyl acetate	C ₉ H ₁₀ O ₂	Ester	23.03	108.00

Butyl hexanoate	C ₁₀ H ₂₀ O ₂	Ester	23.20	117.00
Ethyl octanoate	C ₁₀ H ₂₀ O ₂	Ester	23.34	88.00
Octyl acetate	C ₁₀ H ₂₀ O ₂	Ester	23.65	56.00
Decanal	C ₁₀ H ₂₀ O	Aldehyde	23.80	67.00
Myrtenol	C ₁₀ H ₁₆ O	Alcohol	23.90	79.00
Terpineol	C ₁₀ H ₁₈ O	Terpene	23.90	93.00
γ-terpinene	C ₁₀ H ₁₆	Terpene	23.90	121.00
Methyl salicylate	C ₈ H ₈ O ₃	Ester	23.97	120.00
Nerol	C ₁₀ H ₁₈ O	Terpene	24.20	93.00
3,4-dimethyl benzaldehyde	C ₉ H ₁₀ O	Aldehyde	24.50	133.00
Geraniol	C ₁₀ H ₁₈ O	Terpene	24.60	123.00
2-decanol	C ₁₀ H ₂₂ O	Alcohol	24.90	83.00
Methyl decanoate	C ₁₂ H ₂₄ O	Ester	25.95	74.00
Myrtenyl acetate	C ₁₂ H ₁₈ O ₂	Ester	26.45	91.00
Eugenol	C ₁₀ H ₁₂ O ₂	Alcohol	26.80	164.00
Methyl anthranilate	C ₈ H ₉ NO ₂	Ester	26.84	151.00
Octyl butanoate	C ₁₂ H ₂₄ O ₂	Ester	26.95	89.00
Hexyl hexanoate	C ₁₂ H ₂₄ O ₂	Ester	27.00	99.00
Ethyl decanoate	C ₁₂ H ₂₄ O ₂	Ester	27.10	81.00
Geranyl acetone	C ₁₃ H ₂₂ O	Ketone	28.15	107.00
Cinnamyl acetate	C ₁₁ H ₁₂ O ₂	Ester	28.30	115.00
D ₃₂ -pentadecane	C ₁₅ D ₃₂	Alkane	28.50	66.00
γ-decalactone	C ₁₀ H ₁₈ O ₂	Lactone	28.75	85.00
Methyl jasmonate	C ₁₃ H ₂₀ O ₃	Ester	29.14	151.00
Methyl dodecanoate	C ₁₃ H ₂₆ O ₂	Ester	29.70	87.00
Nerolidol	C ₁₅ H ₂₆ O	Terpene	30.09	93.00
Ethyl laurate	C ₁₄ H ₂₈ O ₂	Ester	30.35	88.00
γ-dodecalactone	C ₁₂ H ₂₂ O ₂	Lactone	32.50	85.00

9.2. Extraction, derivatization and analysis of polar primary metabolites using Gas-Chromatography/Mass Spectrometry (GC/MS)

About 120 mg of frozen strawberry fruit tissue was incubated with 1,200 µL of a solution of 8 µg/ml ribitol in methanol, during 15 min at 75°C and 800 rpm. After that, 600 µL of ultrapure water was added and followed by a centrifugation at 3,500g for 15 min. Then, about 60 µL of the upper polar phase was collected and transferred to a new 1.5 mL tube. The supernatant was then dried in a speed-vac (Eppendorf Concentrator plus), filled with Ar and kept at -80°C until use. For derivatization process through silylation, 60 µL of methoxyamine hydrochloride (30 mg/mL in pyridine) was added to the previous dried samples and incubated at 37°C and 200 rpm for 2 hours. After that, 120 µL of a mixture of 3% (v/v) of fatty acid methyl esters (FAMES) mix (**Table 3**) in *N*-methyl-*N*-

(trimethylsilyl)trifluoroacetamide (MSTFA) was added and incubated for 30 min at 37°C and 200 rpm. Finally, the mixture was transferred to a sample vial.

From each sample, 1 µl in splitless or split mode was injected with the helium carrier gas flow rate set to 2 mL/min by using the autosampler. The flow rate was kept constant with the electronic pressure control enabled. The injection temperature was set to 230°C. Chromatographic separation was performed using a 30-m MDN-35 capillary column. The temperature program was isothermal for 2 min at 80°C, followed by a 15°C/min ramp to 330°C, and holding this temperature for 6 min. Transfer line temperature was set to 250°C and match ion source conditions. The ion source was set to maximum instrument specifications and 250°C. The recorded mass range was $m/z = 70$ to $m/z = 600$ at 20 scan/s. Manual mass defect was set to 0, filament bias current was -70 V, and detector voltage was 1,700-1,850 V.

Table 3. Composition of fatty acid methyl esters (FAMES) mix. Dissolution of all listed compounds were done in chloroform.

Compound	Chemical formula	Concentration (mg/mL)
Methyl caprylate	C ₉ H ₁₈ O ₂	17.74
Methyl pelargonate	C ₁₀ H ₂₀ O ₂	17.68
Methyl caprate	C ₁₁ H ₂₂ O ₂	17.62
Methyl laurate	C ₁₃ H ₂₆ O ₂	17.60
Methyl myristate	C ₁₅ H ₃₀ O ₂	17.32
Methyl palmitate	C ₁₇ H ₃₄ O ₂	20.00
Methyl stearate	C ₁₉ H ₃₈ O ₂	10.00
Methyl eicosanoate	C ₂₁ H ₄₂ O ₂	10.00
Methyl docosanoate	C ₂₃ H ₄₆ O ₂	10.00
Methyl tetracosanoate	C ₂₅ H ₅₀ O ₂	10.00
Methyl hexacosanoate	C ₂₇ H ₅₄ O ₂	10.00
Methyl octacosanoate	C ₂₉ H ₅₈ O ₂	10.00
Methyl triacontanoate	C ₃₁ H ₆₂ O ₂	10.00
D ₆ -cholesterole	C ₂₇ D ₆ H ₄₀ O	10.00

Authentic standards were used for identification, which were listed in **Table 4**.

Table 4. Primary metabolites detected by GC/MS. m/z : mass used for intensity quantification.

Compound	Chemical formula	Chemical class	Time index	m/z
1- <i>O</i> -methyl- α -glucopyranoside	C ₇ H ₁₄ O ₆	Sugar	612,750	204.00
2-oxo-glutaric acid	C ₅ H ₆ O ₅	Organic acid	523,847	288.00
4-amino butyric acid	C ₄ H ₉ NO ₂	Organic acid	342,530	102.00
Alanine	C ₃ H ₇ NO ₂	Amino acid	354,060	116.00
Arginine	C ₆ H ₁₄ N ₄ O ₂	Amino acid	605,940	157.00

Asparagine	C ₄ H ₈ N ₂ O ₃	Amino acid	479,050	116.00
Aspartic acid	C ₄ H ₇ NO ₄	Amino acid	457,283	232.00
Citric acid	C ₆ H ₈ O ₇	Organic acid	592,883	273.00
Quinic acid	C ₇ H ₁₂ O ₆	Organic acid	577,930	255.00
Dehydroascorbic acid dimer	C ₁₂ H ₁₂ O ₁₂	Organic acid	624,990	316.00
Fructose	C ₆ H ₁₂ O ₆	Sugar	579,380	307.00
Glucose	C ₆ H ₁₂ O ₆	Sugar	590,400	160.00
Glucuronic acid	C ₆ H ₁₀ O ₇	Sugar	619,180	333.00
Fucose	C ₆ H ₁₂ O ₅	Sugar	528,270	160.00
Glyceric acid	C ₃ H ₆ O ₄	Organic acid	345,050	292.00
Pyroglutamic acid	C ₅ H ₇ NO ₃	Organic acid	506,483	156.00
Rhamnose	C ₆ H ₁₂ O ₅	Sugar	515,530	117.00
Sucrose	C ₁₂ H ₂₂ O ₁₁	Sugar	840,783	361.00
Xylose	C ₅ H ₁₀ O ₅	Sugar	493,370	307.00
α,α' -trehalose	C ₁₂ H ₂₂ O	Sugar	876,240	191.00
Erythritol	C ₄ H ₁₀ O ₄	Sugar	408,455	217.00
Fumaric acid	C ₄ H ₄ O ₄	Organic acid	371,255	245.00
Glucose-6-phosphate	C ₆ H ₁₃ O ₉ P	Sugar	776,505	160.00
Glutamine	C ₅ H ₁₀ N ₂ O ₃	Amino acid	492,650	227.00
Glycine	C ₂ H ₅ NO ₂	Amino acid	325,180	174.00
Isoleucine	C ₆ H ₁₃ NO ₂	Amino acid	319,193	158.00
Leucine	C ₆ H ₁₃ NO ₂	Amino acid	301,020	158.00
Lysine	C ₆ H ₁₄ N ₂ O ₂	Amino acid	615,467	174.00
Malic acid	C ₄ H ₆ O ₅	Organic acid	440,995	233.00
Maltotriose	C ₁₈ H ₃₂ O ₁₆	Sugar	1,972,100	204.00
Methionine	C ₅ H ₁₁ NO ₂ S	Amino acid	474,325	176.00
<i>myo</i> -inositol	C ₆ H ₁₂ O ₆	Sugar	653,910	305.00
Phenylalanine	C ₉ H ₁₁ NO ₂	Amino acid	531,145	192.00
Phosphoric acid	H ₃ PO ₄	Inorganic acid	333,520	299.00
Proline	C ₅ H ₉ NO ₂	Amino acid	338,693	142.00
Putrescine	C ₄ H ₁₂ N ₂	Amine	466,925	174.00
Pyruvic acid	C ₃ H ₄ O ₃	Organic acid	222,650	174.00
Raffinose	C ₁₈ H ₃₂ O ₁₆	Sugar	1,033,367	437.00
Ribitol	C ₅ H ₁₂ O ₅	Sugar	501,603	319.00
Serine	C ₃ H ₇ NO ₃	Amino acid	357,523	204.00
Spermidine	C ₇ H ₁₉ N ₃	Amine	728,760	174.00
Succinic acid	C ₄ H ₆ O ₄	Organic acid	365,427	247.00
Threonine	C ₄ H ₉ NO ₃	Amino acid	368,365	291.00
Tryptophan	C ₁₁ H ₁₂ N ₂ O ₂	Amino acid	790,560	202.00
Tyrosine	C ₉ H ₁₁ NO ₃	Amino acid	658,337	218.00
Urea	CH ₄ ON ₂	Amine	340,357	189.00
Valine	C ₅ H ₁₁ NO ₂	Amino acid	271,580	144.00
β -alanine	C ₃ H ₇ NO ₂	Amino acid	394,250	248.00

9.3. Extraction and analysis of semi-polar secondary metabolites using Ultra Performance Liquid Chromatography (UPLC)-Orbitrap Mass Spectrometry (LC/MS)

Metabolite extraction and UPLC-MS were performed as described previously (Pott *et al.*, 2020). In brief, 50 mg of fruit tissue was weighed in a 2 mL tube and 1 mL of pre-cooled (-20°C) solution of 1:3 (v/v) methanol:methyl *tert*-butyl ether (MTBE) was added and vortexed until fully re-suspended. The samples were then incubated at room temperature with gentle shaking followed by an incubation in an ultrasonication bath for 10 min. Then, 650 µL of 3:1 H₂O:methanol was added and mixed well. The samples were centrifuged at 15,000g for 5 min at room temperature. Then, 500 µL of the top phase were transferred to a new 1.5 mL tube and the aliquots were concentrated in a speed-vac (Eppendorf Concentrator plus), filled with Ar and kept at -80°C until use. Concentrated samples were re-suspended in 200 µL of a mixture 8:2 (v/v) H₂O:methanol.

Chromatographic separation, mass spectrometric measurements, and data analysis were performed using a Vanish UPLC system using an HSS T3 C18 reverse phase column (10032.1 mm i.d., 1.8-µm particle size) and a Orbitrap-MS/MS (Thermo-Fisher, Bremen, Germany), as described by Pott *et al.* (2020). Mass peaks from these spectra were picked and aligned (SIEVE; Thermo-Fisher) leading to peak lists containing accurate mass, retention time, and intensity of each detected peak from each sample. The accurate masses of each detected peak were searched against a compiled version of three biological databases [KNAPSAcK (<http://kanaya.naist.jp/KNAPSAcK/>), Metabolome.Jp (<http://www.metabolomics.jp/>), and KEGG (<http://www.genome.jp/kegg/>)] leading to the assignment of an elemental composition and a compound name to the masses that found a match in the database. Integration of the peak area of the corresponding molecular ion was used to quantify the metabolites.

Table 5. Summary of secondary metabolites found with the UPLC-Orbitrap-LC/MS. RT: retention time, m/z: ion used for intensity quantification. Partially modified from Hanhineva *et al.*, (2008).

Compound	Chemical class	RT (min)	Molecular formula	mode (+/-)	m/z
Caffeic acid hexose	Hydroxycinnamic acid derivative	0.70	C ₁₅ H ₁₈ O ₉	-	341.11
Galloyl hexose 1	Galloyl derivate	1.02	C ₁₃ H ₁₆ O ₁₀	-	331.03
Galloylquinic acid	Galloyl derivate	1.38	C ₁₄ H ₁₆ O ₁₀	-	343.07
Galloyl hexose 2	Galloyl derivate	2.92	C ₁₃ H ₁₆ O ₁₀	-	331.07
Di-hydroxybenzoic acid hexose	Benzoic acid derivative	3.87	C ₁₃ H ₁₆ O ₈	+	315.07
Digalloylquinic acid	Galloyl derivate	3.95	C ₂₁ H ₂₀ O ₁₄	-	495.08
Bis-HHDP glucose	Ellagitannin	4.03	C ₃₄ H ₂₄ O ₂₂	-	783.09

Di-hydroxy methyl benzoic acid hexose	Benzoic acid derivative	4.29	C ₁₄ H ₁₈ O ₈	+	329.09
Procyanidin dimer 1	Procyanidin	4.30	C ₃₀ H ₂₆ O ₁₂	-	577.14
Procyanidin trimer 1	Procyanidin	4.46	C ₄₅ H ₃₈ O ₁₈	-	865.20
Unknown ellagitannin 1	Ellagitannin	4.48	C ₈₂ H ₅₄ O ₅₂	+	934.07
Epicatechin	Flavan-3-ol	4.52	C ₁₅ H ₁₄ O ₆	-	289.07
Procyanidin trimer 2	Procyanidin	4.64	C ₄₅ H ₃₈ O ₁₈	-	865.20
Unknown ellagitannin 2	Ellagitannin	4.69	C ₈₂ H ₅₄ O ₅₂	+	934.07
Galloyl-HHDP-glucose	Ellagitannin	4.75	C ₂₇ H ₂₂ O ₁₈	-	633.07
Coumaric acid hexose 1	Hydroxycinnamic acid derivative	4.85	C ₁₅ H ₁₈ O ₈	-	325.09
Propelargonidin dimer 1	Propelargonidin	4.88	C ₃₀ H ₂₆ O ₁₁	-	561.14
Cyanidin hexose 1	Anthocyanin	4.90	C ₂₁ H ₂₁ O ₁₁	-	449.11
Propelargonidin dimer 2	Propelargonidin	5.01	C ₃₀ H ₂₆ O ₁₁	-	561.14
Coumaric acid hexose 2	Hydroxycinnamic acid derivative	5.09	C ₁₅ H ₁₈ O ₈	-	325.09
Propelargonidin dimer 4	Propelargonidin	5.17	C ₄₅ H ₃₈ O ₁₇	-	849.20
Pelargonidin hexose	Anthocyanin	5.24	C ₂₁ H ₂₂ O ₁₀	-	433.11
Procyanidin dimer 2	Procyanidin	5.32	C ₃₀ H ₂₆ O ₁₂	-	577.14
Pelargonidin rutinose 1	Anthocyanin	5.32	C ₂₇ H ₃₁ O ₁₄	-	579.15
Ellagic acid hexose 1	Ellagitannin	5.76	C ₂₀ H ₁₆ O ₁₃	-	463.05
Ellagic acid hexose 2	Ellagitannin	5.87	C ₂₀ H ₁₆ O ₁₃	-	463.05
Pelargonidin rutinose 2	Anthocyanin	5.89	C ₂₇ H ₃₁ O ₁₄	-	579.15
Procyanidin dimer 3	Procyanidin	5.90	C ₃₀ H ₂₆ O ₁₂	-	577.14
Ellagic acid pentose 1	Ellagitannin	5.96	C ₁₉ H ₁₄ O ₁₂	-	433.04
Ellagic acid hexose 3	Ellagitannin	6.01	C ₂₀ H ₁₆ O ₁₃	-	463.05
Procyanidin dimer 4	Procyanidin	6.10	C ₃₀ H ₂₆ O ₁₂	-	577.14
Perlargonidin malonyl hexose	Anthocyanin	6.17	C ₂₄ H ₂₃ O ₁₃	-	519.11
Ellagic acid pentose 2	Ellagitannin	6.20	C ₁₉ H ₁₄ O ₁₂	-	433.04
Quercetin pentose glucuronide	Flavonol	6.24	C ₂₀ H ₂₀ O ₁₂	+	611.07
Ellagic acid deoxyhexose 1	Ellagitannin	6.27	C ₂₀ H ₁₆ O ₂₂	-	447.06
Unknown ellagitannin 3	Ellagitannin	6.28	C ₈₂ H ₅₄ O ₅₂	+	934.07
Ellagic acid deoxyhexose 2	Ellagitannin	6.40	C ₂₀ H ₁₆ O ₂₂	-	447.06
Ellagic acid	Ellagitannin	6.40	C ₁₄ H ₆ O ₈	-	301.00
Propelargonidin dimer 3	Propelargonidin	6.54	C ₃₀ H ₂₆ O ₁₁	-	561.14
Galloyl-bis-HHDP-glucose	Ellagitannin	6.62	C ₄₁ H ₃₀ O ₂₆	-	935.08
Agrimoniin	Ellagitannin	6.62	C ₈₂ H ₅₄ O ₅₂	-	934.07
Quercetin glucuronide	Flavonol	6.77	C ₂₁ H ₁₈ O ₁₃	+	477.07
Quercetin hexose	Flavonol	6.83	C ₂₁ H ₂₀ O ₁₂	+	463.09
Kaempferol coumaroyl hexose 2	Flavonol	6.85	C ₃₀ H ₂₆ O ₁₃	+	593.12
Hydroxybenzoic acid hexose	Benzoic acid derivative	6.94	C ₁₃ H ₁₆ O ₈	-	301.13
Ferulic acid hexose	Hydroxycinnamic acid derivative	7.20	C ₁₆ H ₂₀ O ₉	-	355.10
Kaempferol coumaroyl hexose 3	Flavonol	7.20	C ₃₀ H ₂₆ O ₁₃	+	593.11
Epiafzelechin or phloretin hexose 1	Flavan-3-ol	7.30	C ₂₁ H ₂₄ O ₁₀	-	435.13

Compound	Chemical class	RT (min)	Molecular formula	mode (+/-)	m/z
Digalloylglucose	Galloyl derivate	7.38	C ₂₀ H ₂₀ O ₁₄	-	483.05
Kaempferol glucuronide	Flavonol	7.38	C ₂₁ H ₁₈ O ₁₂	+	461.07
Cyanidin hexose 2	Anthocyanin	7.40	C ₂₁ H ₂₁ O ₁₁	-	449.11
Ellagic acid deoxyhexose 3	Ellagitannin	7.59	C ₂₀ H ₁₆ O ₂₂	-	447.09
Kaempferol malonylhexose	Flavonol	7.80	C ₂₄ H ₂₂ O ₁₄	-	535.11
Epiafzelechin or phloretin hexose 2	Flavan-3-ol	8.03	C ₂₁ H ₂₄ O ₁₀	-	435.13
Epiafzelechin or phloretin hexose 3	Flavan-3-ol	8.60	C ₂₁ H ₂₄ O ₁₀	-	435.13
Kaempferol coumaroyl hexose 1	Flavonol	9.20	C ₃₀ H ₂₆ O ₁₃	+	593.13
Epiafzelechin or phloretin hexose 4	Flavan-3-ol	9.76	C ₂₁ H ₂₄ O ₁₀	-	435.13
Sinapic acid hexose derivative 1	Hydroxycinnamic acid derivative	9.76	C ₁₈ H ₂₆ O ₉	-	385.15
Sinapic acid hexose derivative 2	Hydroxycinnamic acid derivative	11.24	C ₁₈ H ₂₆ O ₉	-	385.15

10. Genetic sequences retrieval and phylogenetic analysis

All amino acid sequences were retrieved from several databases like Genome Database for Rosaceae (Jung *et al.*, 2019), National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>), Sol Genomics Network (Fernandez-Pozo *et al.*, 2015). Criterium for selecting sequences for phylogenetic analysis was the presence of InterPro (Blum *et al.*, 2025) motif PF00305, which was a hidden Markov model corresponding to a lipoxygenase domain, like plant lipoxygenases. In the case of glucosyltransferases, the motif PF00201 corresponding to UDP-glucuronosyl/UDP-glucosyltransferase hidden Markov model was searched.

Protein pairwise multiple sequence alignment was performed with multiple alignment using fast Fourier transform or MAFFT (Kato and Standley, 2013) and using the full length of the lipoxygenases' protein sequences. Phylogenetic tree was constructed with IQ-TREE (Minh *et al.*, 2013; Nguyen *et al.*, 2015) using the Maximum Likelihood method with 1,000 bootstrap replicates (Tamura *et al.*, 2021; Villavicencio *et al.*, 2024). The evolutionary history of lipoxygenases was inferred using the Maximum Likelihood method and the JTT matrix-based model (Jones *et al.*, 1992). The fitted model also uses probabilities for conserved sites along evolution and 4-categories gamma distribution for modelling variability all along the sequences (JTT+I+G4). The phylogenetic analysis included the sequences of several *Rosaceae* members together with *Arabidopsis thaliana*, *S. lycopersicum* and other species, involving a total of 135 different sequences.

Presence of signal peptides and transit peptides was assessed with the online tool TargetP2.0 (Armenteros *et al.*, 2019) and subcellular localizations were predicted with the DeepLoc2.0 predictor (Thummuluri *et al.*, 2022).

Finally, the phylogenetic tree was represented and modified using the online tool iTOL (Letunic and Bork, 2024).

10.1. *In silico* structural data analysis of proteins

Domain prediction was conducted using InterPro online tool (Blum *et al.*, 2025). The putative protein structures were predicted and modeled using artificial intelligence algorithms used by AlphaFold3 server (Abramson *et al.*, 2024).

Molecular docking for candidate proteins with selected ligands was carried out using AlphaFold3 algorithm provided by Protenix Server (<https://protenix-server.com/login>). Protein structures, structural alignment and figures were obtained using ChimeraX software (Meng *et al.*, 2023).

11. Subcellular localization of FaLOX2 and FaUDPSS by confocal microscopy imaging

The subcellular localization of FaLOX2 and FaUDPSS proteins was assessed by confocal microscopy by fusing a fluorophore to the candidate protein. For this purpose, constructs based on the binary plasmids pGWB5 and pGWB6 (Nakagawa *et al.*, 2007), respectively were created. A PCR amplified version of *FaLOX2* without STOP codon and *FaUDPSS* (see **Table 6** for specific primers) was introduced by Gateway™ cloning system into pGWB5 and pGWB6 backbones. The fusion proteins resulted in a FaLOX2-GFP of 99.36 kDa product under the control of a *35SCaMV* promoter present in the pGWB5 as well as 80.45 kDa for GFP-FaUDPSS under the same promoter in the pGWB6 vector.

For subcellular localization experiments *p19* was used to prevent gene silencing associated with overexpression. The fusion protein FaFAH1-mCherry (Sánchez-Sevilla *et al.*, 2014; Ruiz-Lopez *et al.*, 2021) was used in co-infiltration assays as an endoplasmic reticulum marker.

The constructs pGWB5-FaLOX2-GFP, empty pGWB5, pGWB6-FaUDPSS and empty pGWB6 (**Table 7**) were transformed into *A. tumefaciens* GV3101::pMP90 through electroporation. Those strains were then grown overnight on LB medium supplemented with 50 µg/mL rifampicin, 25 µg/mL and 50 µg/mL kanamycin. The culture was centrifuged at 3,000g for 15 min at room temperature and the pellets were resuspended in agroinfiltration solution (previously described) and incubate in the dark for 2 hours at room temperature. In the case of single-gene expression, the resuspended *Agrobacterium* cells were mixed to reach an O.D₆₀₀ of 0.6 for the construct and 0.25 for the p19. In co-

infiltration assays with FaFAH1-mCherry, the final O.D₆₀₀ of 0.4 was established for pGWB5-FaLOX2-GFP, pGWB6-GFP-FaUDPSS and FaFAH1-mCherry while the final O.D₆₀₀ of 0.15 was used for p19.

Two leaves from 3-week-old *N. benthamiana* plant, preferably the 3rd and 4th leaves from the apex, were infiltrated abaxial side using a 1 mL needleless-syringe. The plants were kept under growth conditions for 2 days after agroinfiltration.

Fluorescence of GFP and mCherry was detected using a Zeiss LSM800 confocal microscope. Excitation of GFP was achieved using a 488 nm Ar laser, while m-Cherry excitation was achieved using a wavelength of 561 nm of the same laser. Co-localization assays were assessed using sequential line scanning mode to separate signals. Microscopy objectives used were Plan-Apochromat 20X (water) and 40X (water) with up to a 2X digital zoom. All cortical plane images from *N. benthamiana* represent a maximum Z-axis projection of several planes from the cell surface to the interior. Images derived from confocal microscopy was processed using the FIJI program (Schneider *et al.*, 2012).

For imaging, 2 days post-infiltration *N. benthamiana* leaves were used to obtain leaf disks used immediately for visualization of lower epidermal cells.

Plasmolysis of cells were performed using 1 M NaCl solution where leaf disks were incubated for 10 min at room temperature and then visualized in the confocal microscope.

12. Protein extraction and immunodetection assay

N. benthamiana tissue from infiltrated leaves (as described before) was ground to fine powder using liquid N₂ with the help of a TissueLyser II (Qiagen). Proteins were extracted by incubating around 100 mg of sample with 20 µL Laemmli buffer 6X (375 mM Tris-HCl pH 6.8, 12% (m/v) SDS, 60% (v/v) glycerol and 0.03% (m/v) bromophenol blue and supplemented with 2% (v/v) 2-mercaptoethanol) at 65°C for 30 min. Samples were centrifuged at 20,000g for 1 min at room temperature and the supernatant is collected for protein electrophoresis.

Proteins are separated by size using a polyacrylamide gel electrophoresis with denaturant conditions (SDS-PAGE). For immunodetection, proteins were electrotransferred to a polyvinylidene fluoride (PVDF) membrane (Immobilon-P, Millipore, 0.45 µm pore size, IPVH00010). The membrane was blocked with 5% (m/v) skim milk in TTBS buffer (20 mM Tris, 150 mM NaCl, 0.1% (v/v) Tween 20) for 2 hours at room temperature and then incubated overnight with 1% (m/v) skim milk in TTBS supplemented with the primary antibody anti-GFP 1:600 (Santa Cruz Biotechnology, sc-9996) or anti-HA 1:3,000 (Sigma Aldrich, H3663-100UL). The membrane was then washed three times with TTBS at room temperature for 5 min. After that, the membrane was incubated with 1% (m/v) skim milk in TTBS with the secondary antibody conjugated with horseradish peroxidase

(HRP), in this case anti-mouse 1:80,000 (Sigma Aldrich, A9044) for 1 hour at room temperature. Epitome-tagged proteins were revealed on a ChemiDoc XRS+ imaging system (BioRad) with Clarity Western ECL reagent (BioRad) or SuperSignal West Atto (Thermo Scientific), depending on the sensitivity required. Following immunodetection, the membranes were stained with either Coomassie staining solution (0.2% (m/v) Coomassie R-250, 10% (v/v) glacial acetic acid and 40% (v/v) ethanol) or Ponceau staining solution (0.1% (m/v) Ponceau S and 5% (v/v) glacial acetic acid).

Immunodetection of biotin was performed with the conjugate streptavidin-HRP (Sigma Aldrich, RABHRP3-600UL). PVDF membrane blocking was performed in this case with 3% (m/v) of bovine serum albumin (BSA) for 2 hours at room temperature. Washing steps were as described above. After that, streptavidin-HRP was used in a dilution 1:15,300 in 3% (m/v) BSA. Revelation was performed as described above.

13. Genetic resources used in this work

The full set of primers, along with details on their application and target genes, is provided in **Table 6**.

Table 6. Oligonucleotides used as primers in this study. All primer sequences are represented in the 5'-3' sense. Forward and reverse primers are represented as F and R, respectively. Adapters (*attB* sites) for Gateway cloning (Invitrogen) in pDONR-Zeo are: GGGGACAAGTTTGTACAAAAAAGCAGGCTTA (GW-F) and GGGGACCACTTTGTACAAGAAAGCTGGGTA (GW-R) and are show red-highlighted. CDS: coding sequence, RNAi: RNA interference.

Gene	Sequences (5'-3')	Use
<i>FaSCL8</i>	F: ACGAAACCGTCTCTCTTCCA R: CTCGAATTGTCCGAACCTGT	PCR <i>F. × ananassa</i>
<i>FaGAPDH2</i>	F: TCCATCACTGCCACCCAGAAGACTG R: AGCAGGCAGAACCTTTCCGACAG	qRT-PCR <i>F. × ananassa</i>
<i>FaDBP</i>	F: TTGGCAGCGGGACTTTACC R: CGGTTGTGTGACGCTGTCAT	qRT-PCR <i>F. × ananassa</i>
<i>FaLOX2</i>	F: CCTTTTGCCAATGATGGACT R: TTCTTGTCCCCGTGACCTAC	qRT-PCR <i>F. × ananassa</i>
<i>FaLOX2</i>	F: GW-F-ATGAGTGAAGTAGACAAATAT R: GW-R-TTAAATAGAGATGCTATATGG	CDS cloning in pDONR-Zeo
<i>FaLOX2</i>	F: GW-F-ATGAGTGAAGTAGACAAATAT R: GW-R-CCCAATAGAGATGCTATATGG	CDS cloning without STOP codon in pDONR-Zeo
<i>TomloxC</i>	F: GW-F-ATGTTGAAGCCTCAATTTCAA R: GW-R-TAAAATTAGTGATTTACCAAG	RNAi cloning in pDONR-Zeo
<i>NPTII</i>	F: GAGGCTATTCGGCTATGACTG R: ATCGGGAGCGGCGATACCGTA	Kanamycin resistance gene for genotyping

Gene	Sequences (5'-3')	Use
<i>UTR-FaUDP</i>	F: TCTAGACACACTATTCTTGAGCT R: ACAAGGAGCAAATTAATTCACA	UTR+CDS cloning in pGEM-T Easy
<i>FaUDP</i>	F: GW-F -ATGGGTTCAACTCAACCGTA R: GW-R -CTACAAATCTTGCATTGCTTT	CDS cloning in pDONR-Zeo
<i>M13</i>	F: ATTTAGGTGACACTATAG R: TAATACGACTCACTATAGGG	Gene sequencing in pGEM-T Easy

Information about constructs generated during this work, along with description of their use and fusion proteins description were listed in **Table 7**.

Table 7. Vector and constructs used in this study and their use. Description detailed the promoter used in the construct (either *35SCaMV* or *UBQ10*), the type of gene sequence used (CDS, coding sequence; RNAi, RNA interference) and bacterial resistance/plant resistance (Zeo: zeocin, Spec: spectinomycin, Kan: kanamycin, Hyg: hygromycin). GOI: gene of interest, IOG: reverse sequence gene of interest; TID: TurboID, CDS-NS: coding sequence without STOP codon.

A. List of vectors used in this study.

Name	Description	Use
pDONR-Zeo	<i>35SCaMV::GOI</i> (Zeo ^R)	Gateway entry vector
pGEM-T Easy	<i>35SCaMV::GOI</i> (Spec ^R)	Ligation entry vector
pK7GW2D	<i>35SCaMV::GOI</i> (Spec ^R /Kan ^R)	Plant overexpression
pK7GWIWG2D(II)	<i>35SCaMV::GOI-IOG</i> (Spec ^R /Kan ^R)	Plant RNAi silencing
pGWB5	<i>35SCaMV::GOI-GFP</i> (Kan ^R /Hyg ^R)	Subcellular localization
pGWB6	<i>35SCaMV::GFP-GOI</i> (Kan ^R /Hyg ^R)	Subcellular localization
pGWB505	<i>35SCaMV::GOI-GFP</i> (Spec ^R /Kan ^R)	Subcellular localization
pH7m34GW	<i>UBQ10::GOI-mCherry</i> (Spec ^R /Hyg ^R)	Subcellular localization
pGWB TID1	<i>35SCaMV::3xHA-TID-GOI</i> (Spec ^R /Hyg ^R)	TurboID proximity labelling
pGWB TID2	<i>35SCaMV::GFP-3xHA-TID-GOI</i> (Spec ^R /Hyg ^R)	TurboID proximity labelling
pGWB502	<i>35SCaMV::GOI</i> (Spec ^R /Kan ^R)	TurboID proximity labelling

B. List of constructs generated and used in this study.

Gene	Vector	Sequence	Size (bp)	M.W. (kDa)
<i>FaLOX2</i>	pDONR-Zeo	CDS	1,860	70.63
<i>FaLOX2-NS</i>	pDONR-Zeo	CDS-NS	1,860	70.74
<i>FaLOX2</i>	pK7GW2D	CDS	1,860	70.63
<i>FaLOX2-NS</i>	pGWB5	CDS-NS	2,622	99.35
<i>TomLOXC</i>	pK7GWIWG2D(II)	RNAi	300	-
<i>FaUDPSS</i>	pDONR-Zeo	CDS	1,434	52.60
<i>FaUDPSS-TID1</i>	pGWTID1	CDS	2,514	92.06

<i>FaUDPSS-TID2</i>	pGWTID2	CDS	3,228	118.91
<i>TurboID</i>	pGWB502	CDS	1,053	38.49
<i>AtBAM1</i>	pGWB505	CDS	3,009	138.00
<i>FaFAH1</i>	pH7m34GW	CDS	1,509	57.54
<i>FaUDP-232</i>	pGEM-T Easy	CDS	1,711	-
<i>FaUDP-1392</i>	pGEM-T Easy	CDS	1,711	-
<i>FaUDP-H037</i>	pGEM-T Easy	CDS	1,443	53.02
<i>FaUDP-H051</i>	pGEM-T Easy	CDS	1,446	53.27
<i>FaUDP-L27-14</i>	pGEM-T Easy	CDS	1,590	-
<i>FaUDP-L51-14</i>	pGEM-T Easy	CDS	1,431	52.67

14. Statistical analysis and R packages used in this study

All statistical tests were performed in R with the appropriate package. For RNA-seq data analysis, the limma package from Bioconductor was used (Ritchie *et al.*, 2015) to fit data with an empirical Bayes method. Metabolite correlations were calculated with corrplot (<https://cran.r-project.org/web/packages/Correlplot/index.html>) and metan (<https://cran.r-project.org/web/packages/metan/index.html>) R packages.

Shapiro-Wilks test was applied to test normality of data distribution and Levene's test was used to test equality of variances in data distribution. Student's *t*-test or one-way ANOVA was used to check significant changes between normal data distributions. For data that did not fit a normal distribution, two different nonparametric tests were used depending on the suitability of the data: Mann-Whitney *U* test or Wilcoxon signed-rank test.

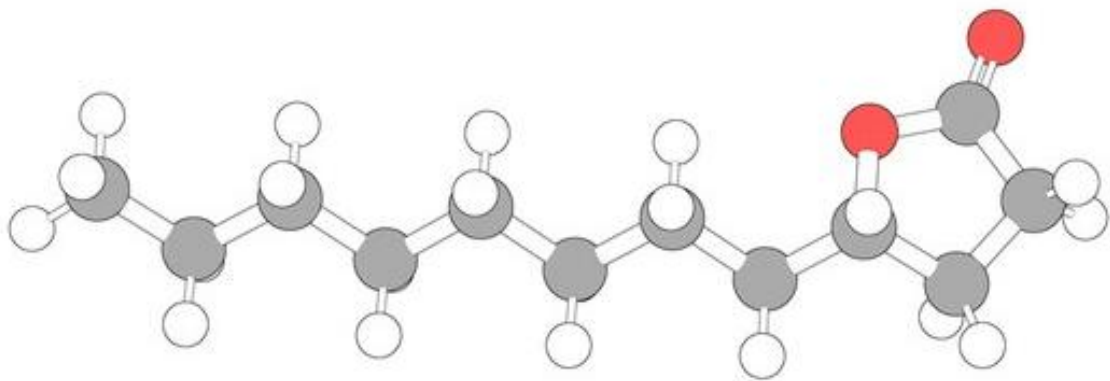
Most analysis and representations were performed using RStudio with several packages listed in **Table 8**.

Table 8. List of R packages used for this study. Some of these packages are uploaded on their correspondent GitHub page.

R package	Use	Reference
circize	Graphical representation	Gu <i>et al.</i> , 2014
ComplexHeatmap	Heatmap representation	Gu, 2022
corrplot	Correlation analysis	Wei and Simko, 2024
dplyr	Data manipulation	Wickham <i>et al.</i> , 2014
expowo	Distribution data of plants	Zuanny <i>et al.</i> , 2024
factoextra	PCA analysis	Kassambara and Mundt, 2020
FactoMineR	PCA analysis	Husson <i>et al.</i> , 2006
ggbreak	Graphical representation	Yu and Xu, 2021
ggplot2	Graphical representation	Wickham, 2016
ggsignif	Statistics graphical representation	Ahlmann-Eltze and Patil, 2022
limma	RNA-seq analysis	Ritchie <i>et al.</i> , 2015

R package	Use	Reference
Metabolomics	Primary metabolites analysis	Pérez-Martín, 2024
metan	Correlation analysis	Olivoto and Lúcio, 2020
missMDA	PCA analysis	Husson and Josse, 2010
R/QTL	QTL analysis	Broman <i>et al.</i> , 2003

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit



Introduction

1. Jasmonic acid biosynthesis and signaling pathways

Jasmonic acid (JA) biosynthetic pathway is based on unsaturated fatty acids of membrane lipids. Specifically, its biosynthesis is produced thanks to the releasement of α -linolenic acid (α -LeA, C₁₈; *cis,cis,cis*- $\Delta^9, \Delta^{12}, \Delta^{15}$) from galactolipids of chloroplast membranes by the enzyme phospholipase A1 (PLA₁). Glycerides and phospholipids on plant chloroplast membranes contain diene unsaturated fatty acids (18:2) and these compounds are substrates of fatty acid desaturase (FAD), and they yield triene unsaturated fatty acids (18:3) (Shi *et al.*, 2023).

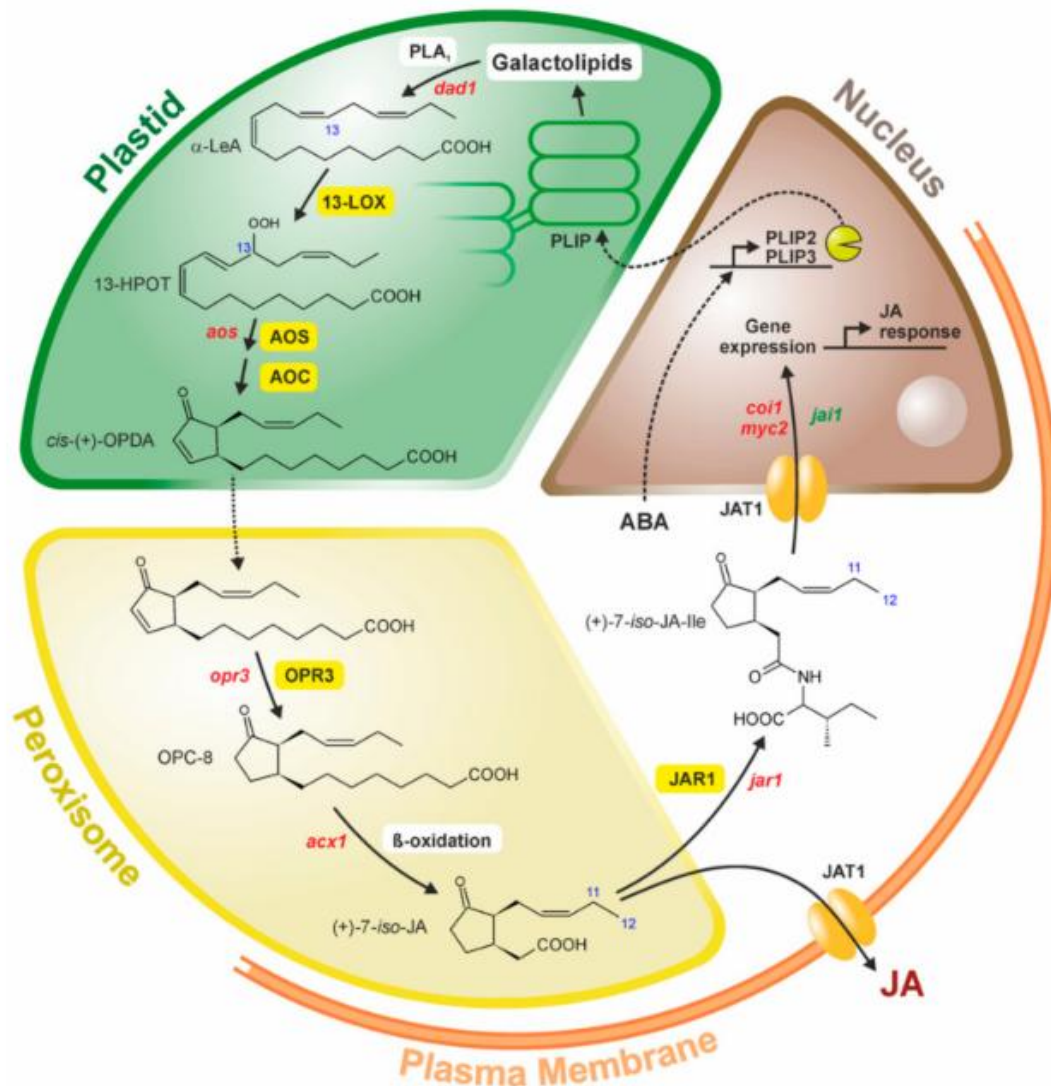


Figure 8. Biosynthesis of jasmonic acid. The biosynthesis and modification of jasmonic acid to jasmonic acid isoleucine conjugate (JA/JA-Ile) is produced by the coordination of four different compartments of a plant cell. PLIP: plastid lipase; PLA₁: phospholipase A1 of flowers; *cis*-(+)-OPDA: *cis*-(+)-12-oxophytodienoic acid; AOS: allene oxide synthase; AOC: allene oxide cyclase; OPR3: OPDA reductase 3; α -LeA: α -linolenic acid. Schema taken from Wasternack and Strnad, (2018).

Jasmonates (JAs) are primarily synthesized through the octadecanoid pathway, originating from α -LeA, or the hexadecanoid pathway, starting from hexadecatrienoic acid (HTA, C₁₆; *cis,cis,cis*- Δ^7 , Δ^{10} , Δ^{13}), both derived from chloroplast galactolipids (Wasternack and Hause, 2013). Among galactolipids, monogalactosyl diacylglycerol (MGDG) and digalactosyl diacylglycerol (DGDG) account for up to 50% and 30% of chloroplast membrane lipids, respectively. DGDG is synthesized from MGDG via the enzyme encoded by the *digalactosyl diacylglycerol synthase1 (DGD1)* gene. While both galactolipids can act as precursors for JA synthesis, MGDG is the primary one, with the MGDG:DGDG ratio significantly influencing the JA biosynthetic pathway (Li and Yu, 2018; Yu *et al.*, 2020b). Additionally, the synthesis of PUFAs from MGDG and DGDG is catalyzed by FADs, ultimately leading to the production of α -LeA and HTA, which serve as precursors for the octadecanoid and hexadecanoid pathways, respectively (He *et al.*, 2020).

The link between jasmonate and DGDG have been described (Lin *et al.*, 2016). In fact, the mutant *DGDG synthase 1, dgd1*, presents a phenotype related to changes in photosynthesis (decrease of photosynthesis ratio and altered chloroplast morphology), reduced levels of DGDG and an up-regulation of 13-LOX and phospholipase A-Iy3, an enzyme activated by DGDG. This mutant presents elevated levels of JA, JA-Ile, OPDA and arabidopsides (Lin *et al.*, 2016). Of course, several lipases, preferentially galactolipases, are involved in the levels of jasmonate. For example, PLASTID LIPASE (PLIP) contributes to jasmonate biosynthesis by exporting acyl groups from plastids for seed oil biosynthesis (Wang *et al.*, 2017). Two homologs, PLIP2 and PLIP3, possess glycerolipase A1 activity that produces the release of MGDG and phosphatidylglycerol, respectively. The regulating effect of abscisic acid (ABA) in JA biosynthesis and their crosstalk have been widely studied (O'Donnell *et al.*, 1996; Wasternack, 2015). In this case, the expression of PLIP2 and PLIP3 is ABA- and COI1-dependent (Wasternack and Strnad, 2018).

The jasmonate biosynthesis shares some enzymes with the LOX pathway because both pathways share α -linolenic acid as initial substrate. The 13(*S*)-hydroperoxide is dehydrated by AOS to unstable allene oxides, that stabilized through cyclization by allene oxide cyclase (AOC), or it can be hydrolyzed into α - or γ -ketols (Hughes *et al.*, 2009; Mosblech *et al.*, 2009). The cyclized product of α -linolenic acid, called 12-oxo-(10*Z*,15*Z*)-phytodienoic acid (OPDA), is reduced and then goes three cycles of β -oxidation catalyzed by the OPDA reductase 3 (OPR3) to yield jasmonic acid (JA). This product can be transformed into a volatile form, methyl jasmonate, by jasmonate carboxymethyl transferase (JMT) (Min *et al.*, 2005). On the other hand, the production of jasmonate through hexadecanoid pathway is considered an alternative route, beginning with HTA, undergoing the same enzymatic reaction as α -LA but yielding dinor-*cis*-12-oxo-

phytodienoic acid (dinor-OPDA). The dinor-OPDA have been classified recently as ancestral jasmonate phytohormones detected in *Marchantia polymorpha* (Wang *et al.*, 2021a). This dinor-OPDA is converted to JA by different OPRs (Soriano *et al.*, 2022).

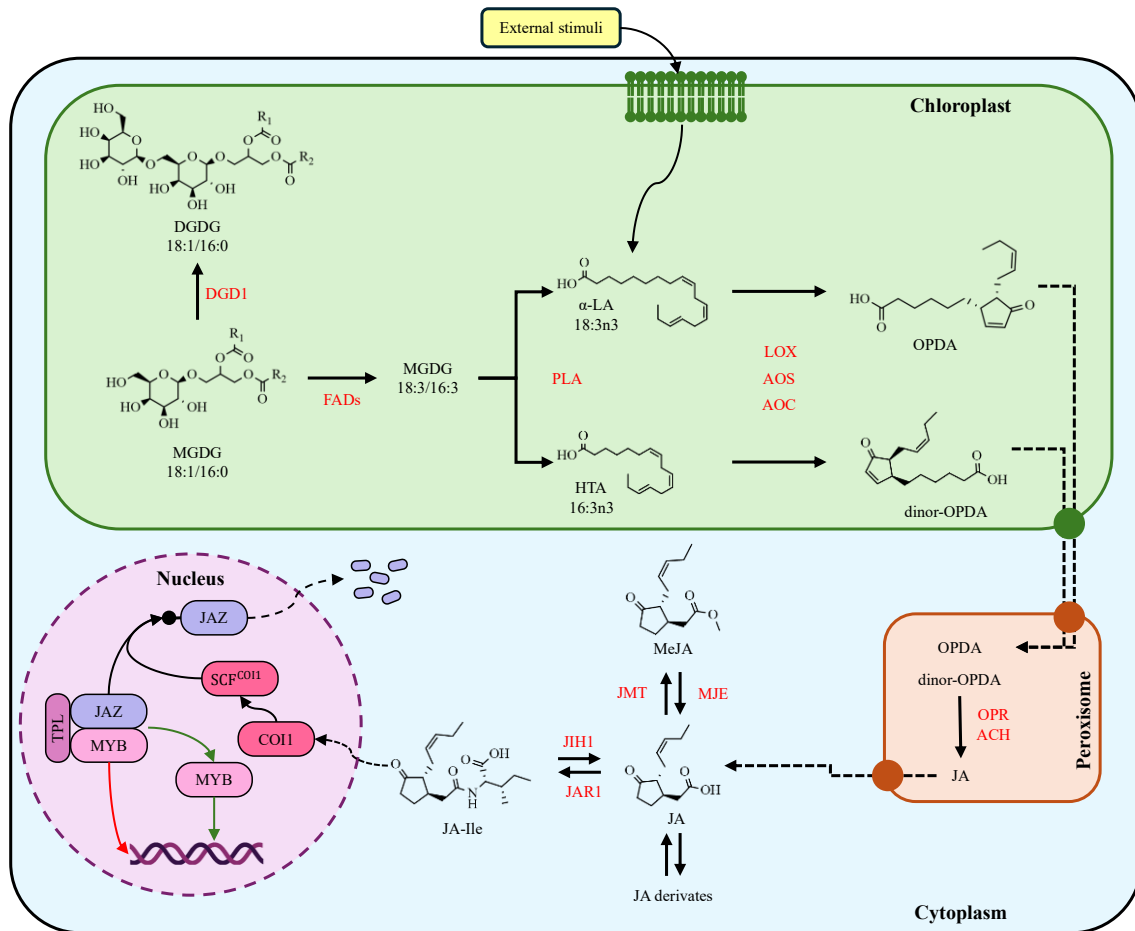


Figure 9. Jasmonic acid biosynthesis and signaling pathway. Biosynthesis of jasmonate comprises 3 different cellular compartments. The JA-Ile is recruited by nuclear COI1 for degradation of JAZ and transcriptional activation by MYB transcription factors. Dotted lines represent a transport from one organelle to another. Solid red lines represent an inhibition while the green ones represent an activation. ACH: acyl-CoA thioesterase II; AOC: allene oxide cyclase; AOS: allene oxide synthase; α -LA: α -linolenic acid; COI1: coronatine insensitive 1; DGDG: digalactosyl diacylglycerol; FAD: fatty acid desaturase; HTA: hexadecatrienoic acid; JA: jasmonic acid; MeJA: methyl jasmonate; JA-Ile: jasmonate isoleucine; JAZ: jasmonate-zim domain; JMT: jasmonate methyl transferase; JIH1: jasmonoyl-*L*-isoleucine hydrolase 1; JAR1: jasmonoyl-*L*-amino acid synthetase; LOX: lipoxygenase; MJE: methyl jasmonate esterase 1; PLA: phospholipase A1; OPDA: 12-oxo-(10Z,15Z)-phytodienoic acid; dinor-OPDA: dinor-*cis*-12-oxo-phytodienoic acid; TPL: TOPLESS. R_1 and R_2 represents fatty acids.

The canonical pathway of JA biosynthesis consists in PUFAs oxygenation, cyclization and reduction steps. In this pathway, the role of OPR3-catalyzed step is crucial in the

biosynthesis of JA. Although, there is a bypath of JA biosynthesis via dinor-OPDA and 4,5-didehydro-jasmonate (4,5-ddh-JA) (Chini *et al.*, 2018). Mutants with no OPR3 activity, *opr3-3*, accumulated 4,5-ddh-JA and JA-Ile under wounding stress but no the conjugated form 4,5-ddh-JA-Ile, which is explained by the β -oxidation of the carboxylic acid side chain of OPDA to dinor-OPDA. The alternative pathway in JA biosynthesis involves either the conversion of OPDA to dinor-OPDA within the peroxisome or the transport of dinor-OPDA from the chloroplast to the peroxisome. This dinor-OPDA is transformed into 4,5-ddh-JA and transported to cytosol where OPR2 yields JA (Wasternack and Strnad, 2018).

The regulation of JA biosynthesis courses through several pathways (**Fig. 9**). Primarily, external stimuli cause the release of α -LA from chloroplast membranes, promoting the generation of JA biosynthesis in their first steps. The expression of JA-related genes is positively regulated by JA and tissue specificity plays a role in JA biosynthesis (Schaller and Stintzi, 2009; Wasternack and Hause, 2013; Wasternack and Song, 2017). The core jasmonate signaling pathway is composed of interconnected functional modules that control the transcriptional regulation of JA-responsive genes. The JA-Ile is recruited by the nuclear receptor CORONATINE INSENSITIVE 1 (COI1). As a typical F-box protein, COI1 initially binds to SKP1-LIKE proteins (ASK1 or ASK2) and subsequently assembles the Skp1/Cul1/F-box complex (SCF^{COI1}) with *Arabidopsis* Cullin1, facilitating the recruitment of JASMONATE-ZIM DOMAIN (JAZ) proteins for ubiquitination and degradation (Devoto *et al.*, 2002; Xu *et al.*, 2002; Sheard *et al.*, 2010). JAZ proteins serve as key repressors that connect early responsive MYC transcription factors (MYC2/3/4/5) with downstream genes (Thines *et al.*, 2007; Wasternack and Song, 2017). In *A. thaliana*, 1 COI, 13 JAZ, and 4 MYC proteins have been identified so far (Xie *et al.*, 1998; Fernández-Calvo *et al.*, 2011; Howe *et al.*, 2018). Under normal conditions, when biotic effectors are absent, JA-Ile levels remain low, and JAZ family proteins accumulate, physically binding to and inhibiting MYCs (Chini *et al.*, 2007; Melotto *et al.*, 2008) through two distinct mechanisms (Chini *et al.*, 2007; Thines *et al.*, 2007; Yan *et al.*, 2007). First, MYC-bound JAZ proteins recruit the co-repressor TOPLESS (TPL) either directly via ETHYLENE-RESPONSE FACTOR-associated amphiphilic repression (EAR) motifs located at the N-terminus of specific JAZ proteins (JAZ5/6/7/8/13) or indirectly through the NOVEL INTERACTOR OF JAZ (NINJA), a protein containing EAR motifs (Pauwels *et al.*, 2010; Shyu *et al.*, 2012; Ke *et al.*, 2015). This constitutive repression prevents MYCs from binding to the G-box sequences of downstream gene promoters, effectively blocking the JA signaling pathway (Liu *et al.*, 2019). In response to various stresses, JA-Ile is rapidly activated, triggering the formation of the JAZ-JA-Ile-SCF^{COI1} E3 ubiquitin ligase complex. This complex facilitates the ubiquitination and subsequent degradation of JAZ proteins (Chini *et al.*, 2007). The degradation of JAZ proteins releases

TPL, freeing MYCs to activate the transcription of JA-responsive genes, thereby initiating the JA signaling pathway (Kazan and Manners, 2013).

On the other hand, the regulation of JA can be driven by metabolism. JA can be converted into several JA derivatives through twelve metabolites pathways known so far. These transformations are conjugation with amino acids, carboxylation, decarboxylation, methylation, esterification, sulfation, *O*-glycosylation and lactone formation. The immediate transformations of the JA are conjugation with amino acids, methylation, hydroxylation and *O*-glycosylation. When conjugated with Ile, JA-Ile can be methylated or *O*-glycosylated in the carboxyl group of isoleucine. The JA-Ile can also be hydroxylated in the C12 and then form several derivatives like 12-carboxy-Ja-Ile, 12-carboxy-JA, 12-*O*-glucosyl-JA-Ile, lactones of 12-hydroxy-JA-Ile, 12-hydroxyjasmonic acid sulfate (12-HSO₄-JA) or 12-*O*-glucosyl-JA (Koo and Howe, 2012; Wasternack and Hause, 2013; Heitz *et al.*, 2016; Widemann *et al.*, 2016; Wasternack and Strnad, 2018).

1.1. Biological role of jasmonate in strawberries

It is widely known that JA signal auto-regulates its own synthesis and gene expression. In fact, in the context of strawberries, JA was found to positively regulate JA biosynthetic pathway genes like *AOS*, *OPR*, *COII* and *JMT*. Experiments with RNAi with the grape gene *VvAOS* showed un-coloring phenotypes in strawberry fruits and, in contrast, the overexpression of the JA receptor *VvCOII* accelerated the strawberry fruit ripening process (Jia *et al.*, 2016).

Fruit ripening is the process through which fruits undergo physiological and biochemical changes at the final stage of growth to attain optimal edible quality. Ethylene has a pivotal role in the ripening process of climacteric fruits but since strawberry is classified as a non-climacteric fruit its ripening process is driven by ABA/auxins ratio (Moya-León *et al.*, 2019). Nevertheless, methyl jasmonate (MeJA) treatment can enhance the development and ripening of strawberry fruit, contributing to pigment metabolism, sugar metabolism, fruit softening, and hormone regulation. The overexpression of *FaAOC* and *FaAOS* produce an acceleration of the ripening process of strawberry fruit. The endogenous JA content in strawberry fruit rises sharply from the small fruit stage to the white fruit stage but decreases as the fruit matures, reaching its lowest level at full ripeness (Han *et al.*, 2019).

The combined effect of aroma components can objectively represent the flavor characteristics and ripeness of agricultural crops, including various fruits. These aroma components include esters (strawberries, apples, grapes), aldehydes and alcohols (peaches, watermelons, tomatoes), lactones (peaches, coconut, tomatoes), terpenes (nectarines, grapes, oranges), phenols and ethers (bananas, citrus, strawberries), as well as sulfur-containing compounds. MeJA treatment can enhance the content of flavor

compounds in strawberries, including 1-butanol, isoamyl acetate, furfural, 2-methyl butanoic acid, hexanoic acid, octyl butanoate, decyl acetate, jasmolactone, *R*- and *S*-furanol, and dodecyl acetate (de la Peña Moreno *et al.*, 2010).

Secondary metabolism is crucial for the growth and development of agricultural crops as well as their stress resistance. JAs can stimulate the synthesis of secondary metabolites, including terpenoids, alkaloids, phenylpropanoids, and flavonoids. The relationship between JA and flavonoid composition has been widely studied (Bursac Kovačević *et al.*, 2015; Fernández-Lara *et al.*, 2015). Color is a highly appreciated trait in strawberry fruit quality (Bursac Kovačević *et al.*, 2015; Fernández-Lara *et al.*, 2015). In fact, the exogenous treatment with MeJA promotes the acquisition of color due to anthocyanin accumulation (Wang, 2000; Concha *et al.*, 2013; Garrido-Bigotes *et al.*, 2018) but also produces an early upregulation of the proanthocyanidin regulation *FabHLH3* (Delgado *et al.*, 2018). Also, the MeJA treatment produced an early downregulation of the transcription factor *FaMYC2* that promotes flavonoid biosynthesis (Delgado *et al.*, 2018).

The JA signaling pathway is the primary mechanism through which plants defend against insect pests. Presumably, the treatment of plants with exogenous JA activates defensive responses of plants. In terms of resistance, either JA treatment or JA signal pathway provide the fruit with resistance to gray mold (*Botrytis cinerea*). MeJA mediates terpene synthesis in strawberry fruit to combat *B. cinerea* infection. In strawberries, MeJA induces the expression of *FaTPS1* through *FaMYC2*, leading to a rapid increase in sesquiterpenoid content, such as daulene D, which enhances the fruit's resistance to *B. cinerea* infection (Zhang *et al.*, 2022).

2. Lipxygenases: activity, occurrence and regulation

Lipxygenases (EC 1.13.11.12) are monomers of nonheme iron-containing dioxygenases with a molecular weight of 90-110 kDa widely distributed in plants and animals (Siedow, 1991; Brash, 1999). Two different domains can be distinguished in these enzymes: a β -barrel domain of 25-30 kDa in the N-terminus and a α -helix domain of 55-65 kDa in the C-terminus. Typically, the N-terminal domain contains around 160 amino acids structurally configured into a β -barrel with antiparallel plates. The C-terminal domain is composed of several α -helices, two antiparallel β -plates and an active Fe^{3+} site. The active state of the iron atom in the lipxygenases is a high-spin oxidized Fe^{3+} while the inactive state is a high-spin reduced Fe^{2+} . The iron atom is coordinated by three histidine, one isoleucine and one asparagine residues and a molecule of H_2O which stabilize the coordination complex (Singh *et al.*, 2022). The water molecule is replaced in normal conditions by the substrate of the reaction (Minor *et al.*, 1993, 1996; Singh *et al.*, 2022). In fact, native LOX contains $\text{Fe}^{2+}(\text{H}_2\text{O})$ that need first to be activated by oxidation through its own hydroperoxide product, leading to LOX $\text{Fe}^{3+}(\text{OH}^-)$ (Vincenti *et al.*, 2019).

LOX can be localized in chloroplasts, lipid bodies, vacuole or cytosol (Feussner *et al.*, 1995; Feussner and Wasternack, 2002; Porta and Rocha-Sosa, 2002). These enzymes catalyze the regio- and stereo-specific hydroperoxidation of PUFAs with (1Z,4Z)-pentadiene configuration to produce (Z, E)-hydroperoxides of the corresponding fatty acids (Hatanaka, 1993; Ivanov *et al.*, 2010). In plants, α -linolenic acid (α -LA) and linoleic acid (LA) are the common substrates of lipoxygenases (Siedow, 1991). Nevertheless, LOXs showed activity towards methyl linoleate, linolenic acid and arachidonic acid, but no activity was detected with oleic acid (Fauconnier and Marlier, 1996; Brash, 1999; Howe and Schillmiller, 2002). The degradation of hydroperoxides produces metabolites, volatile compounds, conjugated dienoic acids, MeJA, and JA. These compounds are actively involved in plant defense-related processes, including responses to biotic and abiotic stresses (Leone *et al.*, 2001).

The mechanism of action of plant-extracted LOX involves several steps: (1) hydrogen atoms are removed from the substrate while simultaneously reducing iron ions in LOX; (2) molecular oxygen reacts with substrate free radicals, the formation of peroxide free radicals; (3) oxygen is subsequently converted oxygen-free radicals; (4) finally, at the culmination of the process, Fe^{2+} in LOX reduces peroxide free radicals to form H_2O_2 , and the subsequent oxidation of Fe^{3+} facilitates the activation of LOX (Murphy, 2008). On the other hand, two models exist to explain the reaction mechanism of positional specificity of LOXs. One hypothesis suggests that fatty acid substrates firstly enter the LOX active site with their methyl terminus first. The depth and steric hindrance within the substrate-binding pocket determine how the pentadiene system aligns with the catalytic iron atom, thus influencing the positional specificity of oxygen insertion (Porta and Rocha-Sosa, 2002; Greco *et al.*, 2012). Another hypothesis proposes that regiospecificity depends on the direction of substrate entry, via the carboxylate for 9-LOXs or the methyl end for 13-LOXs, facilitating radical rearrangement (Kolomiets *et al.*, 2001; Perla *et al.*, 2014). For some plant LOXs, a combination of these models may apply, as substrate orientation may depend on the space within the binding pocket (Perla *et al.*, 2014). The carbon carrying the hydroperoxide function is a chiral carbon, leading to the existence of two stereoisomers (S, R) of the PUFA hydroperoxide formed. In plants, LOXs predominantly synthesize the (S)-diastereoisomer at high levels (95%), in contrast to spontaneous auto-oxidation, which is non-stereospecific and produces a racemic (R, S) mixture of isomers (Nanda and Yadav, 2003; Coffa *et al.*, 2005b,a). However, LOXs capable of producing the (R)-epimer have also been reported (Brash, 1999; Palmieri-Thiers *et al.*, 2009; Schiller *et al.*, 2015).

The 13-LOXs have been extensively studied in several plant species. In tomato, TomloxC and TomloxD are 13-LOXs located in chloroplast. TomloxC is responsible for the synthesis of five-carbon flavor volatiles, but its role in defense was not clear (Shen *et al.*,

2014). The mutant *spr8* is generated by a point mutation in the catalytic domain of TomloxD and this mutant shows alterations in jasmonate biosynthesis and defense against herbivores (Yan *et al.*, 2013). The role of TomloxD in JA synthesis has been reported in several studies (Yoeun *et al.*, 2013; Chen *et al.*, 2017; Oliw and Hamberg, 2017; Toporkova *et al.*, 2018). In *A. thaliana*, four different 13-LOXs (AtLOX2, AtLOX3, AtLOX4 and AtLOX6) and two 9-LOXs (AtLOX1 and AtLOX5) have been found so far. The 13-LOX are reported to play a role in JA biosynthesis, but with specific roles in other physiological processes such as wounding, natural and dark-induced senescence, fertility, and flower development (Wasternack and Strnad, 2018). The 9-LOXs are described to play a role in cell death through the synthesis of death acids (Christensen *et al.*, 2015).

LOX pathway signals are amplified by autocatalytic cycles involving Ca^{2+} and calmodulin ions. Hydroperoxide is generated in the plasma membrane from linoleate and linolenate, which facilitates the transport of Ca^{2+} ions from the extracellular to the intracellular side (Babenko *et al.*, 2017). An increase in cytoplasmic Ca^{2+} levels triggers the activation of phospholipase A, leading to the release of PUFAs from phospholipids (Liavonchanka and Feussner, 2006). Under autocatalytic cycling conditions, MeJA induces the expression of desaturase-related genes that catalyze the conversion of linoleic acid to linolenic acid (Nishiuchi *et al.*, 1997).

The role of LOXs during pathogen's attack has also been described. During pathogen attack, plants release metabolites with protective and toxic effects, while others act as signaling molecules to activate defense-related responses. For biting and chewing herbivores, the LOX pathway plays a critical role in defense mechanisms by producing key oxylipins such as 10-OPDA (a positional isomer of 12-OPDA) and 12-OPDA through the action of 13-LOXs, and 10-oxo-11-phytoenoic acid (10-OPEA) through the action of 9-LOXs (Pokotylo *et al.*, 2015).

3. Strawberry fruit ripening process: abscisic acid (ABA) metabolism and its genetic control

Strawberries are classified as non-climacteric fruits, thus relaying the ripening process to abscisic acid (ABA). The formation of the receptacle, the edible part, depends on auxins produced in the achenes, the true seeds, while ripening relies on ABA, which is synthesized in receptacle cells (Devoto *et al.*, 2002; Xu *et al.*, 2002). The ripening process of strawberry can be divided into four distinct stages defined by the color of the receptacle: green, white, turning and red. This process is subjected controlled by the main phytohormones control: auxins, gibberelins, brassinosteroids, ethylene and ABA (**Fig. 10**).

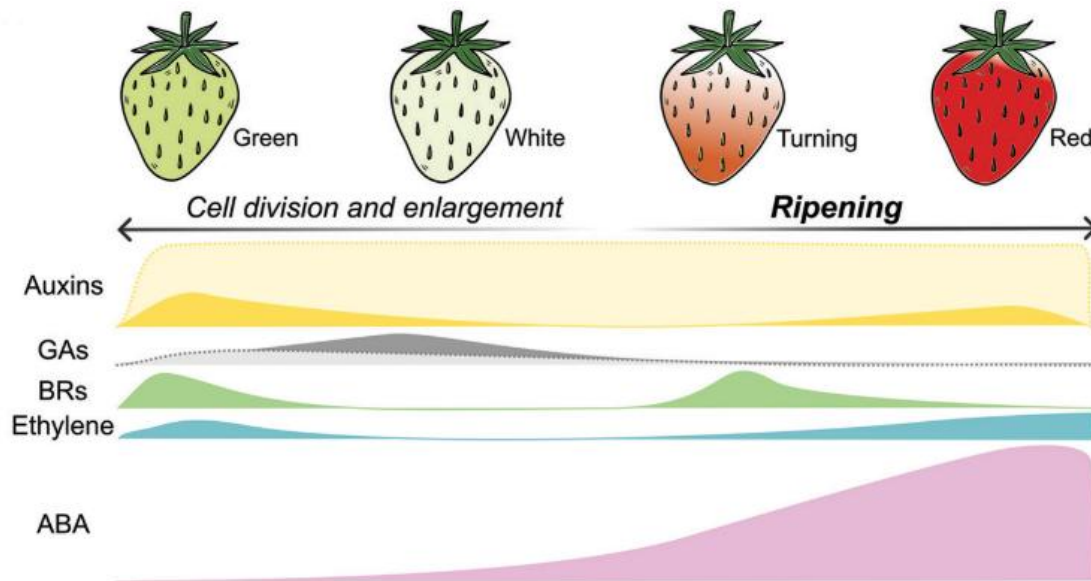


Figure 10. Scheme of the strawberry fruit ripening and the content of phytohormones. Dotted areas represent hormone levels in strawberry achene tissue. GAs: gibberelins, BRs: brassinosteroids. Schema taken from Perotti *et al.*, (2023).

The effects of auxins and gibberellic acid include strawberry fruit growth (Kang *et al.*, 2013; Liao *et al.*, 2018). In the case of strawberry, auxins are mainly synthesized in the achenes and play an important role in early stages of development where cell division and expansion are the main processes (Symons *et al.*, 2012). Gibberellic acid (GA) is thought to promote fruit elongation but not expansion. GA concentration peaks at the large green stage, which can be explained by the promoting effect of auxin on GA biosynthesis (Serrani *et al.*, 2008; Dorcey *et al.*, 2009). Brassinosteroids are steroidal phytohormones essential for plant growth, reproductive development and abiotic tolerance (Nolan *et al.*, 2020) but their function in strawberry is poorly understood. The bioactive castasterone accumulates drastically in flowers, followed by a dramatic decline during fruit ripening, suggesting near-zero actuation during fruit development and ripening (Symons *et al.*, 2012). Ethylene concentrations are higher in the green and red stages of strawberry ripening process, overlapping with an increase in respiration rate, as occurs in climacteric fruits (Perotti *et al.*, 2023).

3.1. ABA biosynthesis and catabolism

The level of ABA content in fruits is primarily regulated by a balance between its biosynthesis and catabolism. ABA is mainly synthesized via the indirect C₄₀ carotenoid pathway, a process well-characterized in *A. thaliana*. Biosynthesis begins in the plastids, where the C₄₀ carotenoid precursor zeaxanthin undergoes a series of isomerization and cleavage reactions, ultimately producing the C₁₅ molecule xanthotin. Zeaxanthin is converted to all-*trans*-violaxanthin through the action of zeaxanthin epoxidase (ZEP).

All-*trans*-violaxanthin can either isomerize into 9'-*cis*-violaxanthin via an unidentified isomerase or it can be transformed into all-*trans*-neoxanthin, which in turn may be converted into 9'-*cis*-neoxanthin by another unknown isomerase. Both 9'-*cis*-violaxanthin and 9'-*cis*-neoxanthin serve as substrates for 9-*cis*-cyclocarotenoid dioxygenase (NCED), the enzyme responsible for producing xanthoxin, being this step the one considered as the rate-limiting in ABA biosynthesis being considered as the rate-limiting in ABA synthesis (Wu *et al.*, 2023). Subsequently, in the cytosol, xanthoxin is primarily converted into abscisic acid aldehyde by the enzyme short-chain alcohol dehydrogenase (ABA2) and then into ABA by abscisic aldehyde oxidase (AAO/ABA3) (Nambara and Marion-Poll, 2005; Li *et al.*, 2022).

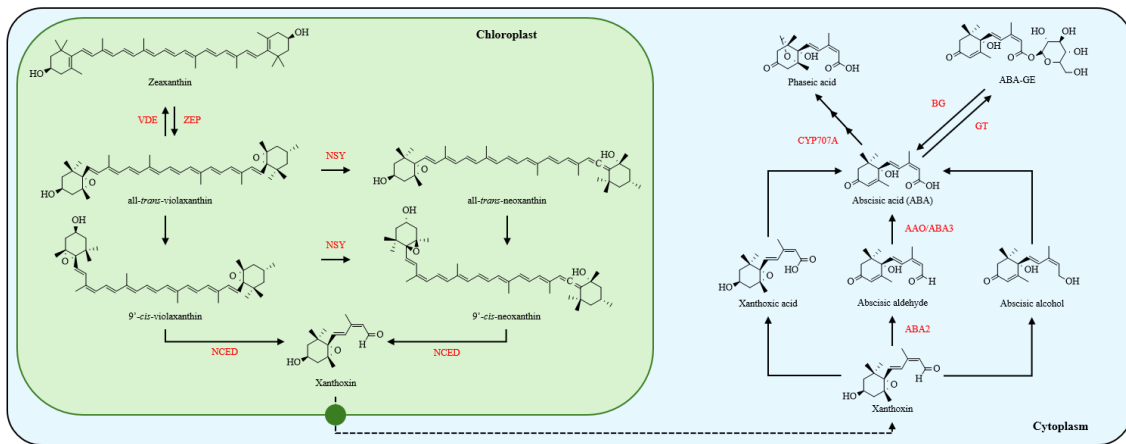


Figure 11. Absciscic acid (ABA) biosynthetic pathway. This schema shows the primary pathway for ABA synthesis in higher plants. ABA2: short-chain alcohol dehydrogenase; ABA-GE: ABA-glucose ester; AAO/ABA3: abscisic aldehyde oxidase; BG: β -glucosidase; CYP707A: cytochrome P450 monooxygenase; GT: ABA glucosyltransferase; NCED: 9'-*cis*-cyclocarotenoid dioxygenase; NSY: neoxanthin synthase; VDE: violaxanthin de-epoxidase; ZEP: zeaxanthin epoxidase. Modified from Wu *et al.*, (2023).

Two distinct pathways are involved in ABA catabolism, both of which play crucial roles in regulating fruit ripening: oxidative degradation and conjugation. The oxidative degradation of ABA is primarily catalyzed by cytochrome P450 monooxygenase of the CYP707A family. The pathway involves 8'-hydroxylation of ABA, resulting in the formation of an unstable intermediate, 8'-hydroxy-ABA, which is subsequently isomerized into phaseic acid and further reduced to dihydrophaseic acid (Rattanakon *et al.*, 2016; Figueroa *et al.*, 2021). Additionally, oxidation can occur at the C9' position, producing neophaseic acid (neoPA) through isomerization (Zhou *et al.*, 2004). The second catabolic route is ABA conjugation, mediated by ABA-glucosyltransferase (GT), which converts ABA into its primary conjugated form, ABA-glucose ester (ABA-GE). ABA-GE serves as inactive storage and transport form of ABA within the plant (Figueroa

et al., 2021). This process is reversible, as ABA-GE can be hydrolyzed back into biologically active ABA molecules by β -glucosidases (BGs).

3.2. ABA signaling cascade

ABA perception and signaling involve the coordinated action of pyrabactin resistance (PYR)/PYR-like (PYL) proteins, collectively known as regulatory components of ABA receptors (RCARs), type 2C protein phosphatases (PP2Cs), and subfamily 2 SNF1-related kinases (SnRK2s). When ABA PYR/PYL/RCARs, it facilitates their interaction with PP2Cs, effectively lifting the inhibition that PP2Cs exert on SnRK2 activity (Fujii *et al.*, 2009; Melcher *et al.*, 2009). Once activated, SnRK2s phosphorylates downstream targets, initiating ABA-mediated physiological responses. In the absence of ABA, however, PP2Cs remain active and bind to SnRK2s, keeping them in an inactive state and thereby blocking the signaling cascade (Nakashima *et al.*, 2009; Umezawa *et al.*, 2009; Shinozawa *et al.*, 2019).

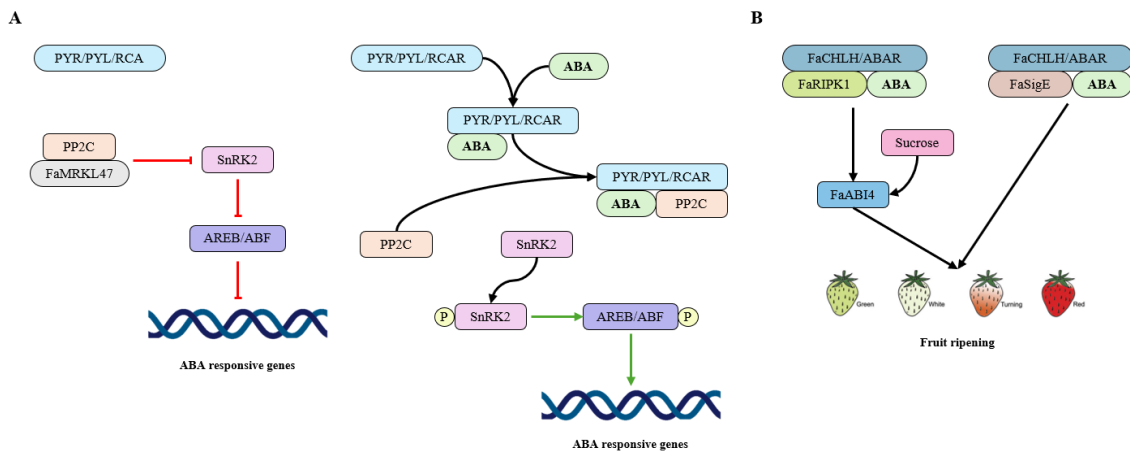


Figure 12. ABA signaling in strawberry fruits. (A) ABA perception in absence and presence of ABA. (B) Novel signaling pathway described in strawberry. Inhibition and activation are represented as red and green lines, respectively. AREB/ABF: ABA-responsive element-binding protein; FaCHLH/ABAR: ABA receptor magnesium chelatase H subunit; FaABI4: ABSCISIC ACID-INSENSITIVE 4; FaMRKL47: FERONIA/FER-like receptor kinase; FaRIPK1: Red-Initial Protein Kinase 1; FaSigE: Sigma factor E; PPC2: type 2C protein phosphatases; PYR/PYL/RCAR: Pyrabactin Resistance (PYR), PYR-like proteins (PYL), Regulatory Components of the ABA Receptors (RCAR); SnRK2: Sucrose nonfermenting 1-related protein kinase 2. Schema modified from Perotti *et al.*, (2023).

The first ABA signaling pathway involves three key components: PYR/PYL/RCAR receptors, clade A protein phosphatase 2C (PP2C, negative regulator), and sucrose-non-fermenting kinase 1-related protein kinase 2 (SnRK2, positive regulator). In the presence of ABA, the cytosolic PYR/PYL/RCAR receptors undergo a conformational change, allowing them to bind to PP2C and consequently releasing SnRK2. Activated SnRK2s

then phosphorylate downstream targets, including transcription factors of the ABA responsive element-binding factor subgroup (AREB/ABF). These phosphorylated transcriptions factors mediate the expression of ABA-responsive genes (Nakashima *et al.*, 2009; Umezawa *et al.*, 2009; Wang *et al.*, 2013; Shinozawa *et al.*, 2019). The role of various ABA signaling components has been explored through transient experiments. For instance, silencing the ABA receptor gene *FaPYR* led to a loss of color, a phenotype that could not be rescued by exogenous ABA application (Chai *et al.*, 2011). The ABSCISIC ACID INSENSITIVE (*FaABI1*), a PP2C protein, was identified as a negative regulator of fruit ripening, whose silencing caused the acceleration of the ripening process (Jia *et al.*, 2013). Overexpression of *FERONIA/FER-like receptor kinase (FaMRKL47)* delayed fruit ripening, indicating a role of this process; moreover, its interaction with *FaABI1* has been demonstrated (Jia *et al.*, 2017). Finally, RNA interference (RNAi) silencing of *FaSNRK2.6* significantly promoted ripening, while its overexpression had the opposite effect, suggesting its involvement in ABA-mediated regulation of fruit ripening (Han *et al.*, 2015).

The second ABA signaling pathway involves a putative ABA receptor gene, *FaCHLH/ABAR* also known as *FaABAR* (Hou *et al.*, 2018). In this model (**Fig. 12**), *FaABAR* and complex with Red-Initial Protein Kinase 1 (*FaRIPK1*), enabling the perception of ABA and promoting fruit ripening through the modulation of *FaABI4* expression (Hou *et al.*, 2018). Silencing of *FaABAR* led to the development of colorless fruits despite elevated ABA levels, indicating that the interruption of this signaling pathway prevents the activation of ABA responses (Jia *et al.*, 2011). Additionally, *FaABAR*-downregulated lines exhibited reduced sugar content, likely due to the upregulation of *Sigma factor E (FaSigE)*, a regulator of sugar metabolism (Jia *et al.*, 2011). Interestingly, silencing *FaSigE* resulted in decreased expression of *FaABAR* and lower ABA levels, supporting the hypothesis that *FaSigE* acts as a positive regulator of both *FaABAR* expression and fruit ripening (Zhang *et al.*, 2017).

3.3. Genetic control of strawberry fruit ripening

The ripening process of strawberry fruits involves significant modifications to primary cell wall metabolism, resulting in fruit softening, which is considered an important contribution to fruit mouthfeel (Morales-Quintana and Ramos, 2019). Transient RNAi silencing of the pectinase gene *rhamnogalacturonate lyase (FaRGLyase1)* in ‘Camarosa’ fruits led to increased pectin accumulation in the cell walls, enhancing cell wall integrity and reducing the number of intercellular spaces (Molina-Hidalgo *et al.*, 2013). In fact, the expression of *FaRGLyase1* is thought to be positively regulated by ABA, as treatment with nordihydroguaiaretic acid (NDGA), an inhibitor of NCED enzyme in the ABA synthesis, resulted in reduced expression of this gene (Han *et al.*, 2004). Similarly, silencing of β -galactosidase 4 (*Fa β Gal4*) in the ‘Camarosa’ cultivar of *F. $\times$ ananassa* led

to higher galactose levels in the cell wall and reduced pectin solubilization. The expression of *Fa β Gal4* was also positively regulated by ABA (Paniagua et al., 2016). In addition, exogenous ABA application has been shown to enhance the expression of genes such as *F. $\times$ ananassa expansin 2* (*FaEXP2*), *β -xylosidase 1* (*FaXyl1*), and *xyloglucan endotransglycosylase/hydrolase 1* (*FaXTH1*) but also produced accelerated fruit softening (Bustamante et al., 2009; Opazo et al., 2013; Nardi et al., 2014, 2016). Several transcription factors involved in fruit ripening and softening also respond to ABA treatment. For instance, SHATTER-PROOF-like (*FaSHP*), a C-type MADS-box transcription factor, acts as a positive regulator of fruit softening, coloration, and some other ripening traits (Daminato et al., 2013). *FaSHP* expression is significantly upregulated by ABA and, in turn, triggers the activation of genes such as *polygalacturonase 1* (*FaPG1*), *pectate lyase* (*FaPL*), *endo- β -1,4-glucanase 1* (*FaEG1*) as well as genes related to color and aroma (Lu et al., 2018).

FaRIF has also been identified as a central regulator of cell wall composition. Silencing *FaRIF* resulted in firmer fruit, while its overexpression led to softer receptacles. Transcriptomic analysis of these transgenic lines, the expression of arabinogalactan-proteins, responsible of cell-wall modifications, like *FaEXP1/2/3*, *FaPME38/39* and several *FaAGPs* were altered (Martín-Pizarro et al., 2021). Another key gene, *FaPRE1*, was found to influence cell wall metabolism. Its downregulation led to decreased expression of *FaPG1* and *FaRGlyase1* and produced an upregulation of cell wall genes that are expressed in early stages of strawberry fruits (Medina-Puche et al., 2019).

The most study transcriptional regulator of strawberry fruit color is *FaMYB10*. Overexpression of *FaMYB10* courses with reddening in leaves, roots and flowers and with anthocyanins accumulation in fruit (Lin-Wang et al., 2010); while its downregulation caused a decrease in anthocyanin content (Medina-Puche et al., 2014). The role of *FaMYB10* in color is so crucial that an eight-nucleotides insertion (ACTTATAC) in the coding sequence produces the white-colored fruits in the *F. $\times$ ananassa* cv. ‘Snow Princess’ (Wang et al., 2020).

Besides the regulating role on proanthocyanidins content, the paralogs *FaMYB9* and *FaMYB11* have been studied for regulating the content of C6-volatiles and esters, respectively (Lu et al., 2020, 2021a). *FaMYB9* regulates the content in hexanal and (*E*)-2-hexenal, contributing to the grassy flavor of unripe fruits (Larsen and Poll, 1990; Du et al., 2011). The protein *FaMYB9* physically interacts with *FaLOX5* and *FaMYB11* induces *FaLOX5* expression by directly binding to its promoter. Furthermore, *FaMYB11* positively regulates the expression of alcohol dehydrogenase (*FaADH*) and alcohol acyl transferase (*FaAAT*), both enzymes from fatty acid metabolism, regulating the levels of aldehydes and esters (Lu et al., 2021a).

The content of furanones in strawberry fruits is a desirable trait due to the caramel-like aroma produced by furaneol and mesifurane (Larsen and Poll, 1990). The transcription factor FaMYB98 has been identified as a direct regulator of quinone oxidoreductase (*FaQR*), responsible for transforming 4-hydroxy-5-methyl-2-methylene-3(2*H*)-furanone into furaneol. Other components participate in *FaQR* regulation, for example transcript levels of *FaMADS9*, *FaSHP* and *FaPRE1* correlate with *FaQR* downregulation (Seymour *et al.*, 2011; Daminato *et al.*, 2013; Medina-Puche *et al.*, 2019; Xie *et al.*, 2020).

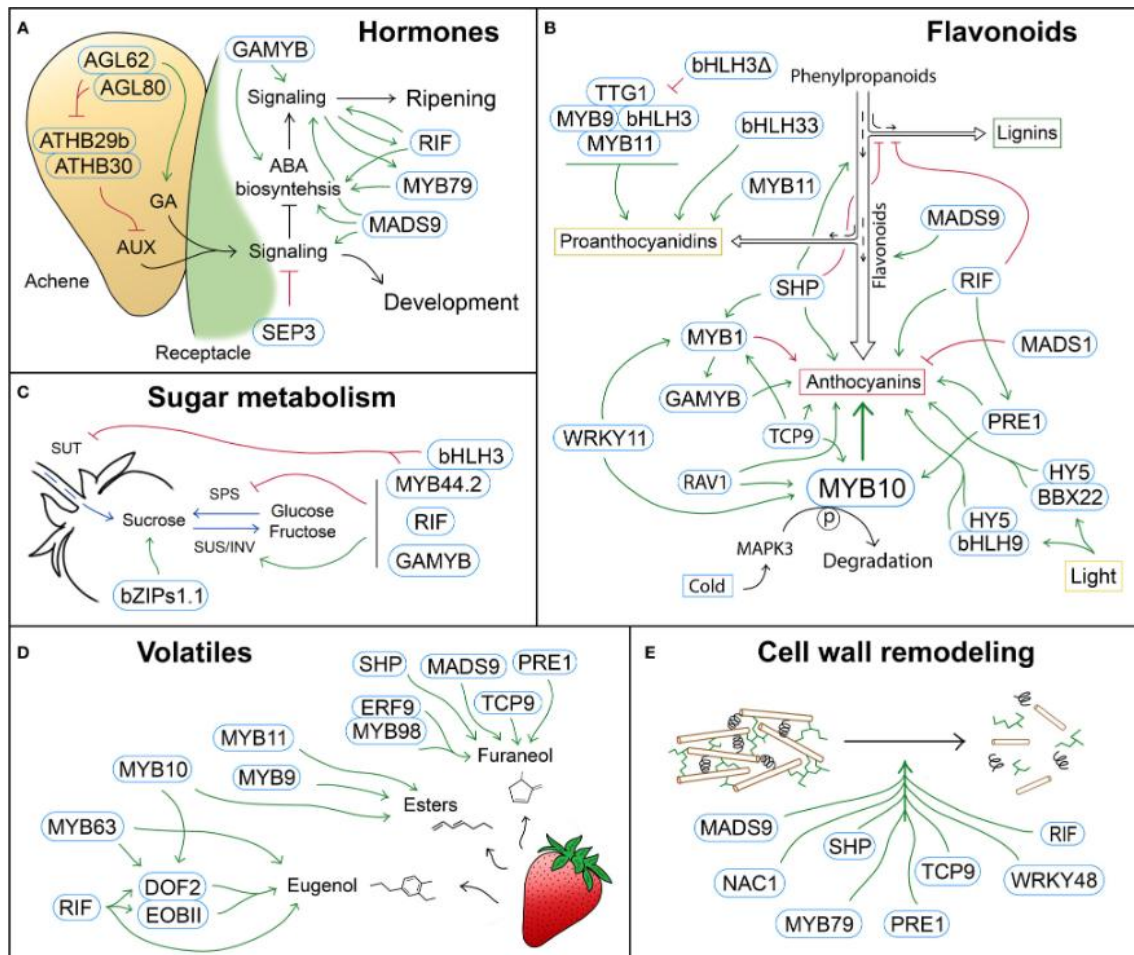


Figure 13. Genetic control by transcription factors of strawberry ripening-related processes. (A) Hormone biosynthesis regulated by auxins (AUX) and gibberellic acid (GA). (B) Regulation of phenylpropanoids biosynthesis. (C) Transcription factors involved in sugar metabolism. (D) Volatile organic compounds regulation. (E) Regulation by transcription factors of the cell wall architecture. Schema taken from Sánchez-Gómez *et al.*, (2022).

Some regulators have been described to participate in eugenol biosynthesis. For example, FaEOBII positively regulates eugenol biosynthesis and with silencing of the *FaEOBII* low levels of transcript for cinnamyl alcohol dehydrogenase (*FaCAD1*) and *FaEGS2*. Another MYB-like transcription factor, *FaMYB63* regulate eugenol biosynthetic genes like *FaCAD1*, *FaEGS1* and *FaEGS2*, but also indirectly regulate positively *FaMYB10*

and *FaEOBII* (Wang *et al.*, 2022). *FaDOF2*, a DNA binding one zing finger family member, positively regulates *FaEOBII* and *FaEGS2* gene expression (Molina-Hidalgo *et al.*, 2017).

Finally, *FaRIF* has been proposed to regulate terpenoids VOCs content, mainly linalool and nerolidol. This was discovered when *FaRIF*-silenced fruits presented a downregulation of NEROLIDOL SYNTHASE 1 (*FaNES1*), the responsible gene of linalool and nerolidol biosynthesis (Martín-Pizarro *et al.*, 2021).

4. Lactones: biological occurrence and importance

Lactones are chemically defined as cyclic esters derived from hydroxy carboxylic acids. They are named based on the size of the lactone ring, ranging from α - to ω -lactones (Fig. 14). Owing to their chemical stability, γ - and δ -lactones are the most abundant forms found in nature (Syed *et al.*, 2022). Lactones with ring sizes containing 8 to 11 carbon atoms are categorized according to their biosynthetic origin into monocyclic oxylipins, monocyclic polyketides, aliphatic bicycles, and aromatic lactones (Sartori *et al.*, 2021). Additionally, lactones with rings comprising 12 or more atoms are classified as macrolactones (Syed *et al.*, 2022). Lactones biosynthesis is closely linked to lipid metabolism, as they are naturally formed from hydroxy fatty acids.

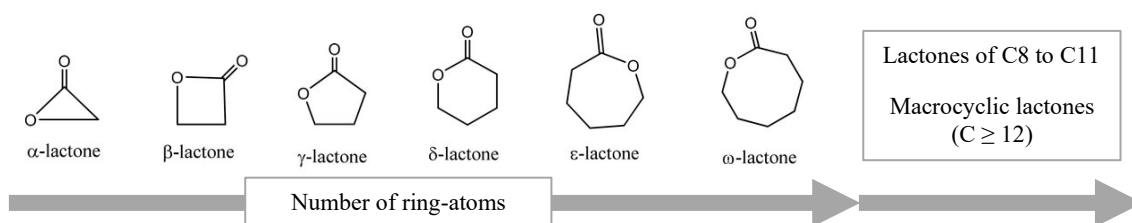


Figure 14. Chemical classification of lactones. Lactones are named according to the number of atoms participating in their ring. Modified from Syed *et al.*, 2022.

Lactones are known as flavoring agents, imparting fruity or creamy aromas and flavors reminiscent of fruits such as apricots, coconut, peach, or strawberries. Because of these sensory properties, lactones are commercially valuable compounds in the food industry. Common examples include γ -undecalactone, γ -decalactone, γ -nonalactone, and δ -decalactone (Felipe *et al.*, 2017). These compounds are commonly present in ripe fruits and contribute significantly to their complex aroma profiles (Sánchez-Sevilla *et al.*, 2014).

In strawberries, several different lactones have been identified and quantified in several varieties (Jeti *et al.*, 2007). Nevertheless, two major lactones have been identified in strawberries: γ -decalactone and γ -dodecalactone. Both contribute peach-like notes to the fruit's aroma, and they have been described as sweetness-enhancers (Fan *et al.*, 2021).

Consequently, the concentration of these lactones has emerged as a desirable trait for strawberry breeding. Genetic studies suggest that while γ -decalactone content may be controlled by a single gene, γ -dodecalactone content appears to be regulated by multiple genes (Fan *et al.*, 2021).

Results

1. Volatile content in Goodberry population

With the objective of evaluate the variability in volatiles (VOCs) content in strawberry ripe fruits in the ‘Candonga’ $\times$ ‘Senga Sengana’ strawberry bi-parental population cultivated in five different locations across Europe (France, Germany, Italy, Poland and Spain). A total of 77 VOCs were analyzed in 126 F₁ lines and the two parentals by using head space SPME gas chromatography coupled with mass spectrometry (HSPME-GC-MS). Volatile from the main classes impacting strawberry aroma were identifying including 5 alcohols, 15 aldehydes, 1 alkane, 39 esters, 8 ketones, 2 lactones and 13 terpenes.

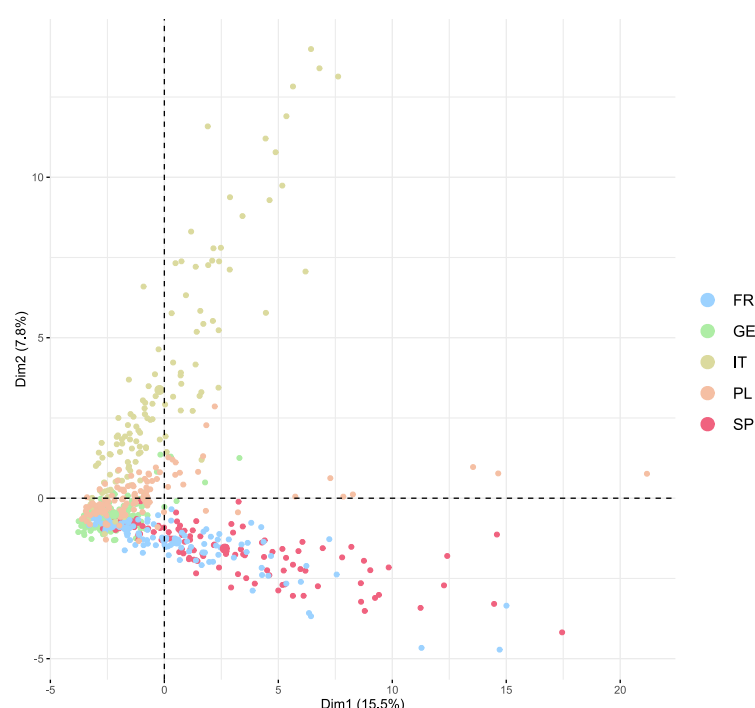


Figure 15. Principal Component Analysis (PCA) of volatile organic compounds in a total of 126 F₁ lines (Goodberry population) grown in five locations.

Dimensions 1 and 2 are showed in the graph. Five different countries are represented with colors as follow: France (FR), Germany (GE), Italy (IT), Poland (PL) and Spain (SP).

To visualize similarities between lines based on volatile profile, principal component analysis (PCA) was conducted with the full dataset (**Fig. 15**). We found that the first principal component (Dim1) accounted for 15.5% of the total variation, while PC2 (Dim 2) 7% of variation. **Annex 2** summarized the metabolites that contributed the most to the first and second dimensions, respectively. The plot showed a separation of individuals grown in Italy mainly in the dimension 2 while individuals grown in Spain and France were segregated along the dimension 1. In contrast, the individuals grown in the two locations Germany and Poland were aggregated, and only discrete separation was observed in the second dimension. Overall, the separation driven by the two dimensions was not enough to discriminate between locations.

The total of 77 VOCs were chemically diverse and grouped into seven different chemical classes (alcohols, aldehydes, alkanes, esters, ketones, lactones and terpenes) including

compounds ranging from 3 to 15 carbon atoms. Quantification of different VOCs in Goodberry population in the five above-mentioned countries was shown on **Fig. 16**.

Next, hierarchical clustering analysis (HCA) using the results of volatile profile was further investigating the relationship between metabolites (**Fig. 16**).

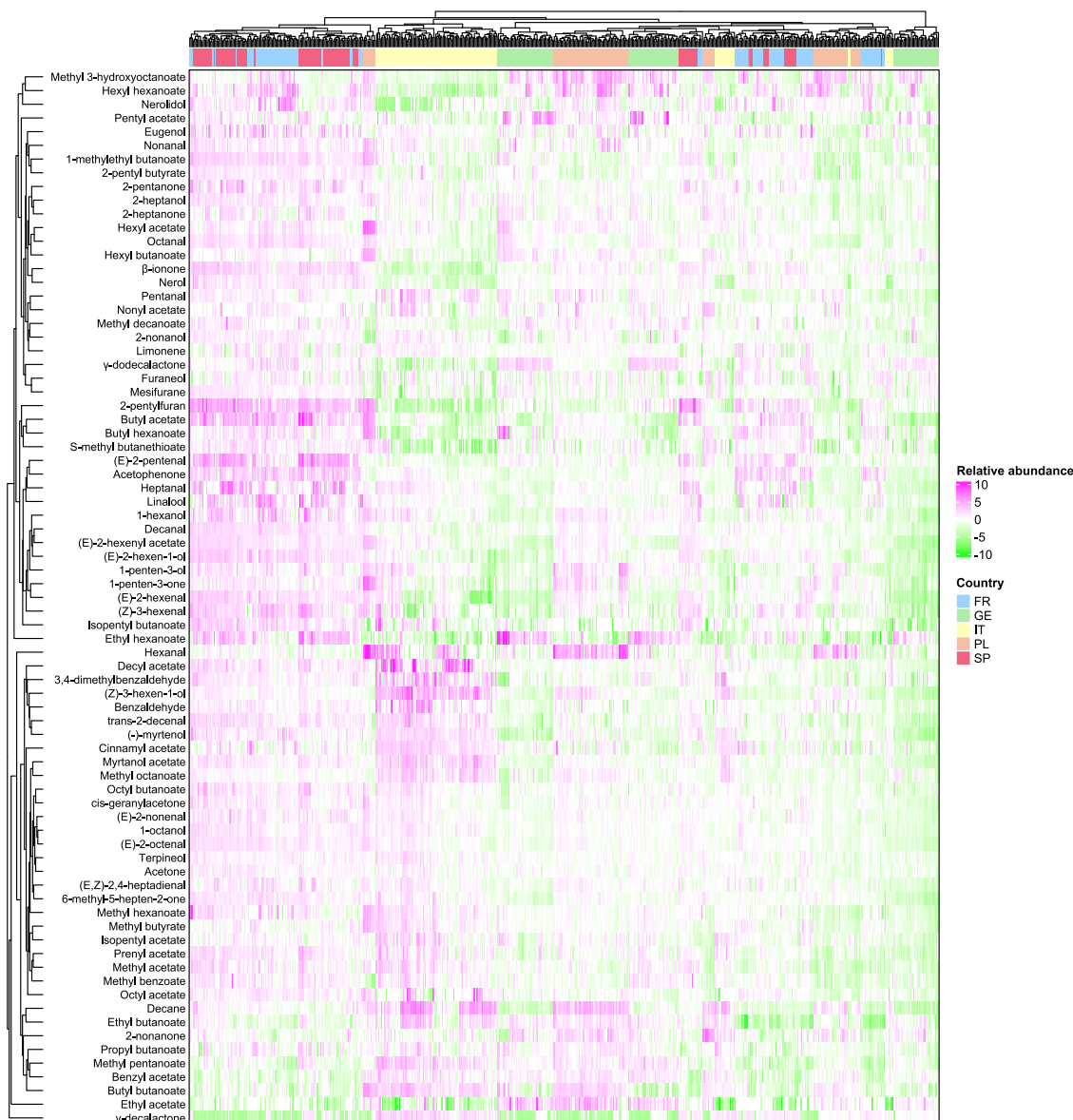


Figure 16. Volatile organic compounds profiling in Goodberry population. Data are presented as relative abundance respect to the m/z intensity of the internal standard and control sample (pool of lines). FR: France, GE: Germany, IT: Italy, PL: Poland and SP: Spain.

Metabolite clustering showed that volatiles were grouped together in several main clusters, pointing out the interconnections in fruit metabolism. Firstly, we observed that some individual grown in France and Spain clustered closely in terms of relative abundance. In general, the levels of almost all measured volatiles were the highest in comparison to the rest of the locations. In contrast, individuals grown in Poland and

Germany were also clustered together but showing the lower levels of volatile. Individuals grown in Italy showed a mixed behavior because they were grouped together and presented low or high abundance values depending on the volatile.

In addition, the heritability for each volatile was calculated by applying the variance components derived from ANOVA ($h^2 = V_G/V_P$, where V_G is the total genetic variance and V_P represents the total phenotypic variance), as previously is described by Perpiñá *et al.*, (2016). The heritability of the different volatiles is shown on **Fig. 17** and the complete information on the heritability of the five different locations is listed in **Annex 3**. A trait with a heritability above 0.5 is highly influenced by genetic control, while below 0.5 is influenced by environmental conditions (Pott *et al.*, 2020). Some volatiles presented high values of heritability, for instance (-)-myrtenol, (*E*)-2-octenal, acetone, butyl hexanoate or heptanal showing a value above 0.8. The opposite situation was observed for *cis*-geranylacetone, linalool, methyl benzoate, nonyl acetate, *trans*-2-decenal and β -ionone, whose heritability is below 0.4 and, in consequence, considered to be mainly environmentally influenced.

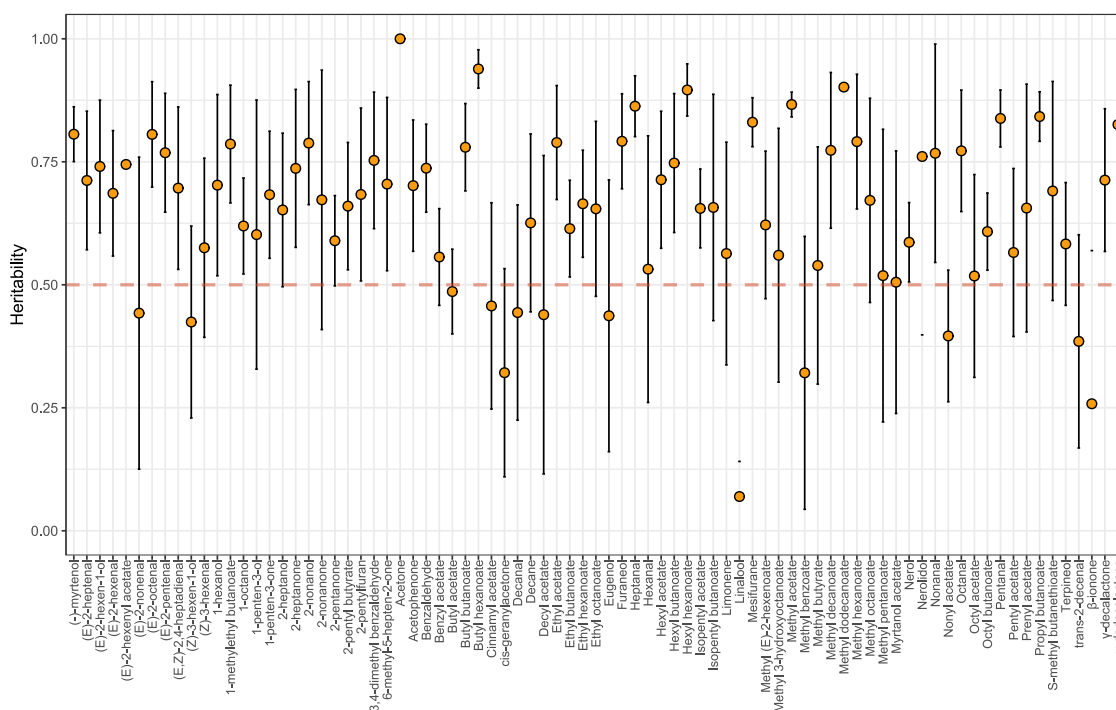


Figure 17. Heritability of volatile organic compounds in Goodberry population. Data shown corresponds to the mean of the five different locations (Germany, France, Italy, Poland and Spain) $\pm$ standard deviation. Red-dashed line represented the threshold of a heritability of 50%.

Important volatile contributors to strawberry aroma (Barbey *et al.*, 2021), like mesifurane, γ -decalactone and γ -dodecalactone displayed a heritability of 0.8, 0.7 and 0.83, respectively. Interestingly, γ -dodecalactone showed the lowest deviation when comparing

the five locations. This result pointed out that the metabolic control of γ -dodecalactone content is mainly driven by the genome. Extreme situations from linalool and acetone also were observed because they showed values of 0.07 and 1, indicating total environmental and genetic control, respectively.

2. Quantitative Trait *Loci* (QTL) identification for volatile organic compounds

The ‘Candongga’ $\times$ ‘Senga Sengana’ linkage map was constructed using single nucleotide polymorphism (SNP) markers genotype in parentals and F₁ lines. Mapping population size was reduced to 113 individuals by evaluating SNP segregation ratios, leading to the exclusion of 13 F₁ lines. From 22,960 SNP markers used initially for mapping, only 6,140 SNP markers were selected for the final maps due to the elimination of redundant markers producing conflictive mapping (**Annex 4**). A total of 28 linkage groups (LGs) were generated in both male and female maps and the markers positions in the LGs were compared with the physical position in the *F. $\times$ ananassa* ‘Camarosa’ v1.0 genome (Edger *et al.*, 2019). These LGs were grouped into seven homologous groups (HG) named 1 to 7, matching the corresponding LGs in the diploid *F. vesca* v4 reference map (Edger *et al.*, 2018).

The complete list of statistically significant Quantitative Trait *Loci* (QTLs) identified for the five different locations was summarized in **Table 9**. Alternatively, a representation of a variety of the detected QTLs was performed on **Fig. 18**. This analysis revealed a significant higher number of QTLs detected using the female map in comparison to the male map. Around the 76% of the total QTL detected corresponded to the female map where the male map only contributed with the 24%. The visualization of QTLs in the LGs allowed us to observe that approximately 67% of female QTLs were located in the chromosome Fvb3-2, which represented approximately 51% of the total QTLs detected in this study. Almost all of QTLs mapped into female Fvb3-2 shared the same genetic interval (comprised by ci-low and ci-up) indicating a possible region controlling a broad range of VOCs variation.

Focusing on lactones, QTLs explaining the variation of γ -decalactone levels can be found only in Fvb3-2 from female map in four out of the five different locations. In contrast, QTLs responsible for the variation of γ -dodecalactone levels can be found in both male and female maps. Checking the male map, the QTL qmDodeLac1 presented a narrower genetic distance than qmDodeLac2, apart from being located on different LG. Female map QTLs for γ -dodecalactone were concentrated within the LG Fvb3-2.

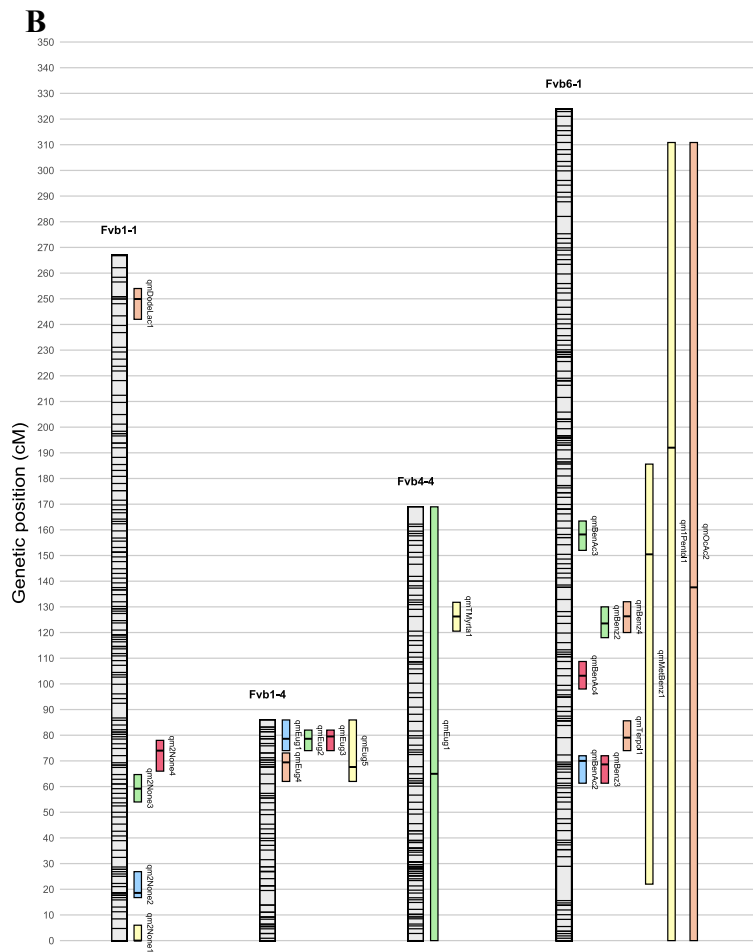


Table 9. Metabolic QTLs (mQTLs) detected for five different locations using Composite Interval Mapping (CIM). Spain (SP), Italy (IT), France (FR), Poland (PL) and Germany (GE); Male (M), Female (F); LG, *F. × ananassa* ‘Camarosa’ linkage groups; Pos, genetic position in cM; ci-low and ci-up, lower and upper genetic position of the Bayesian interval in cM, respectively; LOD, logarithm of the odds ratio; a, mean phenotypic difference between the two homozygous *loci* of the QTL; r^2 , percentage of phenotypic variance explained by the QTL. Consult **Annex 6** for further information.

QTL name	Compound	Map	LG	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r^2	p-value
qf1Hex1	1-hexanol	F	Fvb3-2	257187	SP	12.97	7.50	12.97	4.31	1.15	14.48	0.00
qf1Hex2	1-hexanol	F	Fvb3-2	10212	PL	4.72	2.00	6.00	3.24	-1.84	12.17	0.00
qf1Oct1	1-octanol	F	Fvb2-3	AX-184950666_Fa2-31	PL	22.75	10.25	50.66	2.55	0.92	7.41	0.01
qf1Oct2	1-octanol	F	Fvb3-2	cF3-2-b,loc12	FR	12.00	2.00	12.97	3.58	0.72	15.97	0.01
qf1Oct3	1-octanol	F	Fvb3-2	AX-166512951_Fa3-22	GE	9.32	2.00	10.00	3.55	0.12	12.64	0.00
qf1Oct4	1-octanol	F	Fvb3-2	102124	SP	4.72	4.00	6.00	6.54	0.58	21.87	0.00
qf1Oct5	1-octanol	F	Fvb3-2	102125	PL	4.72	2.00	6.00	3.78	-0.01	14.06	0.00
qf1Pentol1	1-penten-3-ol	F	Fvb1-3	AX-184103238_Fa1-3	GE	5.50	4.28	14.66	5.10	3.38	17.65	0.00
qf1Pentol2	1-penten-3-ol	F	Fvb3-2	102124	FR	4.72	2.00	6.00	2.93	0.19	10.70	0.01
qf2Hept1	2-heptanone	F	Fvb2-1	AX-184466077_Fa2-1	GE	12.87	8.00	33.32	3.05	-0.62	10.35	0.04
qf2Hept2	2-heptanone	F	Fvb3-2	102121	PL	4.72	2.00	6.00	2.76	-0.82	10.46	0.02
qf2Nonol1	2-nonanol	F	Fvb3-2	2571819	PL	12.97	7.50	12.97	3.57	0.51	13.33	0.00
qf2None1	2-nonanone	F	Fvb3-2	AX-184394669_Fa3-2	GE	7.50	2.00	10.00	4.82	0.04	16.77	0.00
qf34Benz1	3,4-dimethyl benzaldehyde	F	Fvb3-2	102121	FR	4.72	2.00	6.00	3.75	0.40	13.97	0.00
qf34Benz2	3,4-dimethyl benzaldehyde	F	Fvb3-2	2571826	PL	12.97	4.72	12.97	3.49	0.00	13.05	0.00
qf6M5Hep1	6-methyl-5-hepten-2-one	F	Fvb1-4	AX-166502512_Fa1-4	FR	84.81	78.36	90.00	3.37	3.61	9.93	0.01
qf6M5Hep2	6-methyl-5-hepten-2-one	F	Fvb3-2	257183	FR	12.97	4.00	12.97	4.46	0.58	14.96	0.00
qf6M5Hep3	6-methyl-5-hepten-2-one	F	Fvb3-2	10212	GE	4.72	2.00	6.00	4.09	0.02	14.41	0.00
qf6M5Hep4	6-methyl-5-hepten-2-one	F	Fvb3-2	cF3-2-b,loc12	SP	12.00	4.00	12.97	7.19	0.30	23.20	0.00

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit

qf6M5Hep5	6-methyl-5-hepten-2-one	F	Fvb3-2	2571812	PL	12.97	7.50	12.97	2.51	-0.08	9.55	0.01
qfAcphne1	Acetophenone	F	Fvb1-4	AX-184660758_Fa1-41	SP	74.69	69.18	80.00	3.44	2.41	10.43	0.00
qfAcphne2	Acetophenone	F	Fvb3-2	10212	FR	4.72	2.00	6.00	3.95	0.17	14.06	0.00
qfAcphne3	Acetophenone	F	Fvb3-2	cF3-2-b,loc81	GE	8.00	4.00	12.97	5.35	-0.07	18.44	0.00
qfAcphne4	Acetophenone	F	Fvb3-2	102125	SP	4.72	2.00	6.00	3.63	0.14	12.26	0.00
qfAcphne5	Acetophenone	F	Fvb3-2	AX-166512951_Fa3-22	PL	9.32	4.00	10.00	6.82	0.32	23.90	0.00
qfAcphne6	Acetophenone	F	Fvb3-2	AX-166512951_Fa3-21	IT	9.32	4.00	12.97	2.57	-0.07	9.25	0.03
qfAcphne7	Acetophenone	F	Fvb6-4	AX-123525449_Fa6-4	SP	80.75	0.00	81.66	2.95	0.40	0.25	0.02
qfActone1	Acetone	F	Fvb1-3	AX-184085467_Fa1-3	GE	0.00	0.00	7.33	8.25	3.96	26.48	0.00
qfActone2	Acetone	F	Fvb3-2	257181	SP	12.97	7.50	12.97	5.16	0.75	17.27	0.00
qfActone3	Acetone	F	Fvb3-2	257181	PL	12.97	7.50	12.97	8.67	-0.07	25.23	0.00
qfActone4	Acetone	F	Fvb3-3	AX-166512245_Fa3-3	SP	28.57	26.00	37.78	3.95	5.59	13.99	0.00
qfActone5	Acetone	F	Fvb3-3	cF3-3-c,loc20	PL	20.00	10.00	20.24	2.95	1.38	6.02	0.03
qfBenAc1	Benzyl acetate	F	Fvb3-2	AX-166512951_Fa3-2	FR	9.32	2.00	10.00	2.81	0.12	10.22	0.01
qfBenz1	Benzaldehyde	F	Fvb3-2	AX-166512951_Fa3-21	FR	9.32	2.00	10.00	3.60	-0.27	12.89	0.00
qfBenz2	Benzaldehyde	F	Fvb3-2	102121	SP	4.72	2.00	6.00	3.72	0.33	13.11	0.00
qfBenz3	Benzaldehyde	F	Fvb3-2	102122	PL	4.72	2.00	6.00	4.62	0.08	16.88	0.00
qfBione1	β -ionone	F	Fvb3-2	257182	FR	12.97	4.00	12.97	3.42	-0.11	12.81	0.00
qfDecAc1	Decyl acetate	F	Fvb1-3	AX-184019288_Fa1-3	FR	26.63	0.00	32.12	1.73	0.00	1.36	0.02
qfDecAc2	Decyl acetate	F	Fvb3-2	AX-166512951_Fa3-23	FR	9.32	2.00	12.97	2.16	0.06	6.72	0.00
qfDecAc3	Decyl acetate	F	Fvb3-2	2571832	PL	12.97	7.50	12.97	2.27	-0.09	9.83	0.00
qfDecAc4	Decyl acetate	F	Fvb3-4	cF3-4-a,loc98	GE	98.00	45.18	116.00	2.65	4.40	0.08	0.03
qfDecal1	Decanal	F	Fvb3-2	102124	GE	4.72	2.00	6.00	2.86	-0.01	10.32	0.03
qfDecal2	Decanal	F	Fvb3-2	AX-166512951_Fa3-24	PL	9.32	4.00	12.97	6.05	0.12	21.53	0.00
qfDecane1	Decane	F	Fvb3-2	AX-166512951_Fa3-2	FR	9.32	2.00	10.00	3.53	1.09	11.95	0.01
qfDecLac1	γ -decalactone	F	Fvb3-2	102122	IT	4.72	4.00	6.00	24.22	0.46	59.92	0.00
qfDecLac2	γ -decalactone	F	Fvb3-2	102123	FR	4.72	4.00	6.00	22.97	0.79	58.58	0.00

QTL name	Compound	Map	LG	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qfDecLac3	γ -decalactone	F	Fvb3-2	102123	SP	4.72	4.00	6.00	26.59	0.87	63.35	0.00
qfDecLac4	γ -decalactone	F	Fvb3-2	1021210	PL	4.72	4.00	6.00	26.52	0.42	65.43	0.00
qfDodeLac1	γ -dodecalactone	F	Fvb1-3	cF1-3-y.loc22	FR	22.00	4.28	27.54	4.73	0.27	7.10	0.00
qfDodeLac2	γ -dodecalactone	F	Fvb3-2	AX-166512951_Fa3-25	IT	9.32	2.00	10.00	4.27	-0.01	15.66	0.00
qfDodeLac3	γ -dodecalactone	F	Fvb3-2	AX-166512951_Fa3-21	FR	9.32	4.00	10.00	7.47	0.00	24.91	0.00
qfDodeLac4	γ -dodecalactone	F	Fvb3-2	10212	SP	4.72	2.00	6.00	4.76	0.06	16.45	0.00
qfDodeLac5	γ -dodecalactone	F	Fvb3-2	2571833	PL	12.97	7.50	12.97	11.13	-0.01	35.63	0.00
qfDodeLac6	γ -dodecalactone	F	Fvb7-4	cF7-4-y.loc56	GE	56.00	52.00	62.00	3.13	2.90	11.22	0.02
qfE2Hex1	(<i>E</i>)-2-hexenal	F	Fvb1-4	AX-184660758_Fa1-4	SP	74.69	54.00	107.93	3.15	4.04	10.78	0.02
qfE2Hex2	(<i>E</i>)-2-hexenal	F	Fvb2-1	AX-184359213_Fa2-1	GE	9.26	4.00	133.46	1.72	344.64	9.66	0.04
qfE2Hex3	(<i>E</i>)-2-hexenal	F	Fvb3-2	cF3-2-b.loc121	FR	12.00	4.00	12.97	2.76	-0.12	10.05	0.01
qfE2Hex4	(<i>E</i>)-2-hexenal	F	Fvb3-2	257183	SP	12.97	7.50	12.97	5.49	-0.30	16.99	0.00
qfE2Hex5	(<i>E</i>)-2-hexenal	F	Fvb3-2	257183	PL	12.97	2.00	12.97	2.95	-0.03	10.00	0.02
qfE2Hexol1	(<i>E</i>)-2-hexen-1-ol	F	Fvb1-4	2233	SP	94.03	89.42	100.00	2.61	7.36	7.39	0.04
qfE2Hexol2	(<i>E</i>)-2-hexen-1-ol	F	Fvb3-2	AX-166512951_Fa3-21	FR	9.32	2.00	10.00	2.59	-1.76	9.46	0.00
qfE2Hexol3	(<i>E</i>)-2-hexen-1-ol	F	Fvb3-2	257185	SP	12.97	4.00	12.97	4.40	0.82	14.06	0.00
qfE2Hexol4	(<i>E</i>)-2-hexen-1-ol	F	Fvb3-2	257185	PL	12.97	7.50	12.97	3.56	-0.38	15.03	0.00
qfE2Hexol5	(<i>E</i>)-2-hexen-1-ol	F	Fvb4-2	AX-184306385_Fa4-2	PL	49.41	18.00	88.96	2.77	3.60	1.01	0.03
qfE2Non1	(<i>E</i>)-2-nonenal	F	Fvb1-4	AX-184945096_Fa1-4	SP	71.02	66.00	76.52	3.39	2.75	9.69	0.01
qfE2Non2	(<i>E</i>)-2-nonenal	F	Fvb3-2	AX-166512951_Fa3-24	FR	9.32	2.00	10.00	3.32	0.51	11.96	0.01
qfE2Non3	(<i>E</i>)-2-nonenal	F	Fvb3-2	102123	GE	4.72	2.00	6.00	3.90	0.05	13.80	0.00
qfE2Non4	(<i>E</i>)-2-nonenal	F	Fvb3-2	10212	SP	4.72	2.00	6.00	4.76	0.57	14.81	0.00
qfE2Non5	(<i>E</i>)-2-nonenal	F	Fvb3-2	102126	PL	4.72	2.00	6.00	3.18	-0.19	11.95	0.00
qfE2Oc1	(<i>E</i>)-2-octenal	F	Fvb3-2	10212	GE	4.72	2.00	6.00	2.59	-0.16	9.40	0.03
qfE2Oc2	(<i>E</i>)-2-octenal	F	Fvb3-2	257188	SP	12.97	4.00	12.97	6.46	0.48	21.72	0.00
qfEug1	Eugenol	F	Fvb3-2	cF3-2-b.loc61	SP	6.00	2.00	10.00	2.66	-0.54	9.56	0.01

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit

qfEug2	Eugenol	F	Fvb3-2	102129	PL	4.72	2.00	6.00	3.05	-0.30	11.50	0.00
qfEug3	Eugenol	F	Fvb3-2	cF3-2-b,loc61	IT	6.00	0.00	10.00	1.82	-0.16	6.63	0.02
qfEug4	Eugenol	F	Fvb3-4	89201	GE	53.43	48.00	156.00	3.75	6.99	13.30	0.00
qfEug5	Eugenol	F	Fvb3-4	AX-184134806_Fa3-4	GE	98.13	68.00	115.68	2.59	16.12	0.61	0.00
qfEZHep1	(E,Z)-2,4-heptadienal	F	Fvb3-2	257184	FR	12.97	4.00	12.97	3.08	0.55	11.14	0.02
qfEZHep2	(E,Z)-2,4-heptadienal	F	Fvb3-2	102123	GE	4.72	2.00	10.00	4.10	0.02	11.72	0.00
qfEZHep3	(E,Z)-2,4-heptadienal	F	Fvb3-2	102122	SP	4.72	2.00	6.00	5.66	-0.01	19.25	0.00
qfEZHep4	(E,Z)-2,4-heptadienal	F	Fvb3-2	2571814	PL	12.97	7.50	12.97	5.22	0.18	18.85	0.00
qfFura1	Furaneol	F	Fvb1-3	AX-184117023_Fa1-3	FR	5.50	0.00	11.00	5.39	7.65	20.23	0.00
qfFura2	Furaneol	F	Fvb1-3	AX-184117023_Fa1-3	SP	5.50	0.00	10.00	6.66	3.57	22.40	0.00
qfFura3	Furaneol	F	Fvb1-3	AX-184512234_Fa1-32	PL	5.50	0.00	10.00	6.14	4.62	21.81	0.00
qfFura4	Furaneol	F	Fvb1-3	AX-184545710_Fa1-3	IT	0.91	0.00	11.00	7.58	5.17	24.00	0.00
qfFura5	Furaneol	F	Fvb3-2	AX-166512951_Fa3-23	GE	9.32	4.00	10.00	7.54	0.14	21.20	0.00
qfFura6	Furaneol	F	Fvb3-2	102127	GE	4.72	2.00	6.00	5.36	0.08	18.46	0.00
qfFura7	Furaneol	F	Fvb3-2	102126	GE	4.72	2.00	6.00	4.20	0.13	15.76	0.00
qfGeAc1	cis-geranilacetone	F	Fvb2-3	AX-184856604_Fa2-3	PL	31.13	0.00	50.66	1.46	3.43	4.28	0.01
qfGeAc2	cis-geranilacetone	F	Fvb3-2	10212	FR	4.72	2.00	6.00	2.41	-0.19	8.83	0.03
qfHexHex1	Hexyl hexanoate	F	Fvb3-3	AX-184271551_Fa3-3	FR	7.32	0.00	8.24	2.44	139.82	4.73	0.02
qfHexHex2	Hexyl hexanoate	F	Fvb6-2	176511	FR	0.00	0.00	5.52	2.31	-176.44	7.59	0.03
qfIsoAc1	Isoamyl acetate	F	Fvb2-1	AX-184018368_Fa2-1	GE	2.78	0.00	62.00	2.19	0.92	7.92	0.04
qfIsoAc2	Isoamyl acetate	F	Fvb2-3	AX-184104182_Fa2-3	PL	22.75	14.00	26.00	2.85	8.62	11.27	0.02
qfLimo1	Limonene	F	Fvb3-2	-	FR	9.32	4.00	10.00	9.48	-0.08	26.60	0.00
qfLimo2	Limonene	F	Fvb3-2	102125	GE	4.72	2.00	6.00	3.41	0.09	12.19	0.01
qfLimo3	Limonene	F	Fvb3-2	AX-166504873_Fa3-2	GE	8.41	2.00	10.00	3.09	0.02	11.64	0.04
qfLimo4	Limonene	F	Fvb3-2	102123	SP	4.72	4.00	6.00	9.08	-0.03	26.72	0.00
qfLimo5	Limonene	F	Fvb3-2	102123	PL	4.72	4.00	6.00	8.18	0.23	27.93	0.00
qfLimo6	Limonene	F	Fvb3-2	AX-184361324_Fa3-21	IT	8.41	2.00	12.97	4.14	-0.15	14.46	0.00
qfLimo7	Limonene	F	Fvb4-2	cF4-2-b,loc8	IT	8.00	0.00	14.00	3.07	0.92	8.13	0.02

QTL name	Compound	Map	LG	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qfLimo8	Limonene	F	Fvb5-1	14591	SP	54.43	44.29	56.00	2.89	1.43	7.09	0.04
qfLino11	Linalool	F	Fvb1-3	AX-184103238_Fa1-3	PL	5.50	0.00	32.12	2.96	1.16	5.76	0.01
qfLino12	Linalool	F	Fvb3-2	102121	FR	4.72	4.00	6.00	8.29	0.02	22.73	0.00
qfLino13	Linalool	F	Fvb3-2	AX-166512951_Fa3-21	SP	9.32	4.00	12.97	9.95	-0.04	30.72	0.00
qfLino14	Linalool	F	Fvb3-2	AX-166512951_Fa3-23	PL	9.32	2.00	10.00	4.71	-0.01	17.18	0.00
qfLino15	Linalool	F	Fvb3-2	102121	IT	4.72	2.00	10.00	4.15	-0.15	14.51	0.00
qfMButy1	Methyl butyrate	F	Fvb1-3	AX-184117023_Fa1-3	GE	5.50	0.00	32.12	1.66	1.07	6.21	0.00
qfMesif1	Mesifurane	F	Fvb1-3	AX-184512234_Fa1-3	IT	5.50	0.00	11.00	3.12	3.36	12.06	0.01
qfMesif2	Mesifurane	F	Fvb1-3	AX-184484824_Fa1-3	FR	17.44	12.00	23.89	5.98	5.65	20.83	0.00
qfMesif3	Mesifurane	F	Fvb1-3	AX-184484824_Fa1-31	SP	17.44	12.00	23.89	4.84	2.70	16.70	0.00
qfMesif4	Mesifurane	F	Fvb1-3	AX-184117023_Fa1-3	PL	5.50	0.00	11.00	6.80	3.16	23.04	0.00
qfMesif5	Mesifurane	F	Fvb3-2	10212	FR	4.72	2.00	6.00	3.01	0.12	13.94	0.04
qfMesif6	Mesifurane	F	Fvb3-2	AX-166512951_Fa3-21	GE	9.32	2.00	10.00	3.48	0.04	12.41	0.01
qfMesif7	Mesifurane	F	Fvb3-2	102124	PL	4.72	2.00	6.00	2.87	0.13	10.65	0.02
qfMetBenz1	Methyl benzonate	F	Fvb3-2	102122	GE	4.72	2.00	10.00	3.14	0.00	11.28	0.02
qfMetBenz2	Methyl benzonate	F	Fvb3-3	AX-184438095_Fa3-4	FR	7.32	0.00	8.24	2.37	2.88	4.11	0.03
qfMetBenz3	Methyl benzonate	F	Fvb5-4	cF5-4,loc52	FR	52.00	0.00	61.96	2.58	2.39	3.53	0.01
qfMetDec1	Methyl decanoate	F	Fvb3-2	2571828	PL	12.97	7.50	12.97	2.90	-0.56	11.32	0.00
qfMetDec2	Methyl decanoate	F	Fvb3-3	AX-184732992_Fa3-3	GE	13.41	4.00	15.69	3.63	2.78	12.92	0.01
qfMOHOc1	Methyl 3-hydroxyoctanoate	F	Fvb1-4	AX-184165951_Fa1-4	SP	63.61	46.02	66.00	3.35	2.35	9.59	0.01
qfMOHOc2	Methyl 3-hydroxyoctanoate	F	Fvb3-2	AX-166512951_Fa3-25	SP	9.32	4.00	12.97	5.97	0.15	17.22	0.00
qfMOHOc3	Methyl 3-hydroxyoctanoate	F	Fvb3-2	cF3-2-b,loc101	PL	10.00	2.00	12.97	3.59	-27.14	13.39	0.00
qfNerdol1	Nerolidol	F	Fvb1-3	AX-166503026_Fa1-3	SP	25.72	4.28	32.12	3.02	0.36	6.63	0.02
qfNerdol2	Nerolidol	F	Fvb3-2	9138	IT	36.20	0.00	43.63	2.17	0.27	4.30	0.02
qfNerdol3	Nerolidol	F	Fvb3-2	AX-166512951_Fa3-29	PL	9.32	2.00	12.97	4.14	-0.22	15.28	0.00

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit

qfNerdol4	Nerolidol	F	Fvb3-4	AX-184857015_Fa3-41	IT	157.26	154.00	164.00	2.21	0.37	8.02	0.02
qfNerol1	Nerol	F	Fvb3-2	9138	FR	36.20	26.96	38.00	3.43	2.11	11.05	0.01
qfNerol2	Nerol	F	Fvb3-2	84912	SP	43.63	0.00	43.63	1.58	5.10	6.27	0.05
qfNerol3	Nerol	F	Fvb3-2	102127	PL	4.72	4.00	6.00	6.53	0.09	23.02	0.00
qfNerol4	Nerol	F	Fvb3-2	AX-166504873_Fa3-2	IT	8.41	2.00	10.00	4.24	-0.03	14.79	0.00
qfNerol5	Nerol	F	Fvb3-3	cF3-3-a,loc24	GE	24.00	18.00	30.00	3.19	2.77	11.96	0.02
qfBione1	β -ionone	F	Fvb3-2	AX-166512951_Fa3-26	FR	9.32	7.50	12.97	4.52	0.25	18.64	0.00
qfBione2	β -ionone	F	Fvb2-3	60241	PL	40.32	0.00	50.66	1.77	2.06	6.84	0.02
qfBione3	β -ionone	F	Fvb3-2	10212	SP	4.72	2.00	6.00	2.51	0.26	10.71	0.03
qfNonAc1	Nonyl acetate	F	Fvb3-2	66512951_Fa	GE	9.32	0.00	10.00	2.25	-0.30	8.19	0.01
qfOcAc1	Octyl butanoate	F	Fvb3-2	AX-166512951_Fa3-21	FR	9.32	2.00	12.97	3.96	-0.20	14.22	0.00
qfOcAc2	Octyl butanoate	F	Fvb3-2	2571830	PL	12.97	7.50	12.97	2.38	0.15	9.10	0.01
qfOcBut1	Pentyl acetate	F	Fvb3-1	cF3-1,loc76	GE	76.00	2.00	98.74	1.93	4.05	4.57	0.03
qfOcBut2	Pentyl acetate	F	Fvb3-2	9138	GE	36.20	0.00	43.63	2.70	6.05	8.69	0.05
qfPent1	Pentanal	F	Fvb6-1	cF6-1-b,loc12	SP	12.00	0.00	26.58	2.47	29.16	8.12	0.03
qfPent2	Prenyl acetate	F	Fvb3-2	84911	GE	43.63	38.00	43.63	2.90	2.62	10.64	0.04
qfPrenAc1	Prenyl acetate	F	Fvb5-1	AX-184088356_Fa5-1	IT	17.60	0.00	179.94	2.39	8.87	1.40	0.00
qfPrenAc2	Prenyl acetate	F	Fvb5-4	289841	IT	4.64	0.00	10.00	1.86	13.08	6.50	0.04
qfSthioate1	<i>S</i> -methyl butanethioate	F	Fvb3-2	cF3-2-b,loc122	FR	12.00	2.00	12.97	2.59	1.69	9.46	0.01
qfTerpol1	Terpineol	F	Fvb3-2	102122	FR	4.72	4.00	6.00	7.22	-0.14	20.54	0.00
qfTerpol2	Terpineol	F	Fvb3-2	AX-166512951_Fa3-24	GE	9.32	4.00	10.00	10.75	0.51	33.57	0.00
qfTerpol3	Terpineol	F	Fvb3-2	cF3-2-b,loc6	SP	6.00	2.00	10.00	7.02	0.04	19.72	0.00
qfTerpol4	Terpineol	F	Fvb3-2	AX-166512951_Fa3-25	PL	9.32	4.00	10.00	6.05	0.09	21.53	0.00
qfTerpol5	Terpineol	F	Fvb3-2	AX-184361324_Fa3-23	IT	8.41	2.00	12.97	3.79	0.00	13.35	0.01
qfTerpol6	Terpineol	F	Fvb5-1	AX-184965167_Fa5-11	SP	61.84	56.26	63.66	5.78	1.57	14.41	0.00
qfTMyrta1	<i>trans</i> -myrtanyl acetate	F	Fvb1-3	AX-184484824_Fa1-3	FR	17.44	4.28	32.12	3.36	0.85	5.84	0.01
qfTMyrta2	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	AX-166512951_Fa3-23	IT	9.32	4.00	12.97	2.63	-0.08	9.46	0.05
qfTMyrta3	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	AX-166512951_Fa3-27	FR	9.32	2.00	10.00	4.75	0.10	16.80	0.00

QTL name	Compound	Map	LG	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qfTMyrta4	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	102122	SP	4.72	2.00	6.00	2.44	-0.24	8.80	0.03
qfTMyrta5	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	102128	PL	4.72	2.00	6.00	3.66	0.28	12.43	0.01
qfTMyrta6	<i>trans</i> -myrtanyl acetate	F	Fvb5-1	13959	PL	7.50	0.00	8.00	3.15	1.04	7.79	0.02
qfZHexAc1	(Z)-3-hexen-1-ol acetate	F	Fvb3-2	257182	FR	12.97	7.50	12.97	4.14	0.12	14.80	0.00
qfZHexAc2	(Z)-3-hexen-1-ol acetate	F	Fvb3-2	102121	GE	4.72	2.00	6.00	4.22	0.24	14.84	0.00
qfZHexAc3	(Z)-3-hexen-1-ol acetate	F	Fvb3-2	257185	SP	12.97	4.00	12.97	4.07	0.24	14.25	0.00
qfMButy1	Methyl butyrate	F	Fvb6-2	AX-166518165_Fa6-21	IT	6.52	0.00	6.52	2.79	-4.08	9.99	0.03
qm1Hex1	1-hexanol	M	Fvb7-1	229991	GE	60.24	54.00	65.79	2.96	3.94	10.66	0.04
qm1Pentol1	1-penten-3-ol	M	Fvb6-1	cM6-1,loc192	IT	192.00	0.00	310.88	1.63	16.44	2.13	0.00
qm1Pentol2	1-penten-3-ol	M	Fvb6-3	195581	PL	147.01	136.00	147.92	2.33	42.97	8.91	0.02
qm2Hept1	2-heptanone	M	Fvb3-4	AX-184250195_Fa3-41	GE	130.33	128.00	140.00	2.98	8.25	10.72	0.01
qm2Hept2	2-heptanone	M	Fvb5-2	cM5-2-b.loc50	PL	50.00	48.00	81.02	2.75	6.82	10.41	0.04
qm2Nonol1	2-nonanol	M	Fvb3-2	dAX-184641647_Fa3-21	FR	171.40	0.00	279.14	1.19	7.98	4.48	0.04
qm2Nonol2	2-nonanol	M	Fvb4-2	12890	GE	20.33	0.00	24.96	0.99	17.15	3.78	0.05
qm2None1	2-nonanone	M	Fvb1-1	291721	IT	0.00	0.00	6.00	4.87	16.70	16.93	0.00
qm2None2	2-nonanone	M	Fvb1-1	AX-184906875_Fa1-1	FR	18.58	16.76	26.86	6.17	29.33	21.08	0.00
qm2None3	2-nonanone	M	Fvb1-1	26001	GE	59.18	54.00	64.68	3.13	26.71	11.23	0.00
qm2None4	2-nonanone	M	Fvb1-1	cM1-1,loc74	SP	74.00	66.00	78.00	3.21	22.38	11.50	0.00
qm2Penton1	2-pentanone	M	Fvb3-4	AX-184842185_Fa3-4	GE	142.40	141.49	152.00	4.09	2.34	14.40	0.01
qm2Penton2	2-pentanone	M	Fvb5-2	158381	PL	18.21	10.00	21.88	3.02	3.13	11.39	0.03
qm34Benz1	Benzyl acetate	M	Fvb6-1	cM6-1.loc70	FR	70.00	61.30	72.00	3.15	2.65	11.40	0.01
qmActone1	Benzyl acetate	M	Fvb6-1	AX-184730940_Fa1-3	GE	158.20	152.00	163.44	2.59	1.52	9.39	0.02
qmBenAc4	Benzyl acetate	M	Fvb6-1	19802	SP	103.17	98.00	108.72	3.41	5.51	11.50	0.00
qmBenz1	Benzaldehyde	M	Fvb5-1	AX-184100394_Fa5-1	GE	19.33	2.78	69.26	2.96	0.65	7.66	0.05
qmBenz2	Benzaldehyde	M	Fvb6-1	19536	GE	123.54	118.00	130.00	6.37	1.14	21.52	0.00
qmBenz3	Benzaldehyde	M	Fvb6-1	AX-184420266_Fa6-1	SP	68.63	61.30	72.00	5.77	4.66	19.59	0.00

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit

qmBenz4	Benzaldehyde	M	Fvb6-1	19562	PL	126.32	120.00	132.00	3.34	1.59	12.53	0.02
qfPenAc1	β -ionone	M	Fvb1-2	AX-184132143_Fa1-41	PL	86.78	76.00	87.68	2.86	0.76	10.84	0.04
qfPenAc2	β -ionone	M	Fvb4-2	dAX-184105987_Fa4-2	FR	7.43	6.55	17.55	3.15	1.22	11.19	0.02
qfPrenAc3	β -ionone	M	Fvb5-3	281801	GE	3.67	0.00	36.00	3.93	0.24	1.58	0.00
qmDodeLac1	γ -dodecalactone	M	Fvb1-1	AX-184502996_Fa1-11	PL	249.91	242.00	254.00	3.78	0.25	12.03	0.01
qmDodeLac2	γ -dodecalactone	M	Fvb4-4	26181	GE	64.97	0.00	168.97	3.05	2.21	8.08	0.03
qmEug1	Eugenol	M	Fvb1-4	AX-184246531_Fa1-4	FR	78.61	74.00	85.95	2.76	30.11	10.33	0.00
qmEug2	Eugenol	M	Fvb1-4	AX-184246531_Fa1-4	GE	78.61	74.00	82.00	3.45	6.04	12.88	0.00
qmEug3	Eugenol	M	Fvb1-4	1526	SP	79.52	74.00	82.00	4.65	17.80	16.14	0.00
qmEug4	Eugenol	M	Fvb1-4	359941	PL	69.43	62.00	73.09	4.56	4.88	16.95	0.00
qmEug5	Eugenol	M	Fvb1-4	AX-184902981_Fa1-4	IT	67.62	62.00	85.95	2.72	11.60	10.28	0.00
qmEug6	Eugenol	M	Fvb2-2	AX-184887573_Fa2-2	FR	197.24	192.15	204.00	2.05	30.57	8.35	0.03
qmEug7	Eugenol	M	Fvb2-2	cM2-2,loc196	SP	196.00	186.00	197.24	2.50	15.46	9.77	0.02
qmIsoAc1	Isoamyl acetate	M	Fvb3-3	91221	IT	286.93	281.37	292.00	2.65	38.77	9.59	0.02
qmIsoAc2	Isoamyl acetate	M	Fvb3-3	7511	GE	94.20	10.00	108.00	2.39	17.72	8.98	0.02
qmMetAc1	Methyl acetate	M	Fvb2-1	291641	IT	234.78	228.38	252.00	3.03	9.09	10.80	0.02
qmMetBenz1	Methyl benzoate	M	Fvb6-1	AX-166518204_Fa6-1	IT	150.46	22.00	185.60	2.49	12.25	6.70	0.03
qmMetDec1	Methyl decanoate	M	Fvb3-3	cM3-3,loc34	PL	34.00	28.00	39.37	2.41	6.33	9.52	0.02
qmMOHOct1	Methyl 3-hydroxyoctonoate	M	Fvb3-2	cM3-2,loc172	GE	172.00	164.05	176.00	2.19	45.74	5.56	0.02
qmMOHOct2	Methyl 3-hydroxyoctonoate	M	Fvb3-3	cM3-3,loc208	GE	208.00	198.00	210.00	2.40	62.60	7.50	0.00
qmNerdol1	Nerolidol	M	Fvb3-3	AX-184748009_Fa3-31	GE	11.46	6.00	18.00	2.97	0.48	10.01	0.03
qmNonAc1	Nonyl acetate	M	Fvb5-2	ddAX-184088628_Fa5-2	GE	142.71	137.20	143.62	3.03	2.08	11.72	0.03
qmNonAc2	Nonyl acetate	M	Fvb7-4	cM7-4-a,loc72	IT	72.00	66.68	77.40	3.54	2.75	10.96	0.00
qmOcAc1	Octyl acetate	M	Fvb5-3	28180	GE	3.67	0.00	10.00	2.30	1.28	8.03	0.00
qmOcAc2	Octyl acetate	M	Fvb6-1	AX-166526928_Fa6-1	PL	137.60	0.00	310.88	1.79	2.12	6.45	0.03
qmPent1	Pentanal	M	Fvb2-2	61401	SP	33.36	28.00	40.00	2.58	33.92	11.01	0.03
qmPent2	Pentanal	M	Fvb5-1	AX-123361773_Fa5-1	GE	3.69	0.00	5.52	2.90	3.69	10.54	0.04

QTL name	Compound	Map	LG	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qmPrenAc1	Prenyl acetate	M	Fvb2-2	61401	IT	33.36	28.00	40.00	1.90	15.48	7.03	0.05
qmSthioate1	S-methyl butanethioate	M	Fvb5-3	169101	PL	27.68	22.16	221.97	2.49	10.62	9.56	0.03
qmTerpol1	Terpineol	M	Fvb6-1	201921	PL	79.06	74.00	85.62	3.19	1.03	11.98	0.03
qmTMyrta1	<i>trans</i> -myrtanyl acetate	M	Fvb4-4	122402	IT	126.25	120.53	131.78	2.99	1.10	10.66	0.04

During this work we focused our attention on the study of qfDodeLac3 and qfDodeLac4 which were detected in two locations, France and Spain, respectively. The main reason for choosing this QTLs was the overlapping interval and similar phenotypic variation in γ -dodecalactone content. These QTLs represented around 25% and 16% of the total variation of γ -dodecalactone across the Goodberry Population (GBP).

3. Identification of the candidate genes for γ -dodecalactone content in the stable mQTLs

With the objective of identifying candidate genes responsible for the variation of γ -dodecalactone levels, the intervals called qfDodeLac3 and qfDodeLac4 were further studied. We first investigated the relationship between molecular markers for each QTL and the γ -dodecalactone content in the whole population. For that purpose, we used γ -dodecalactone quantification data together with the SNP markers used for the linkage maps to observe the tendency of each marker associated with the content in this volatile in the mapping population. These results were presented in **Fig. 19A**.

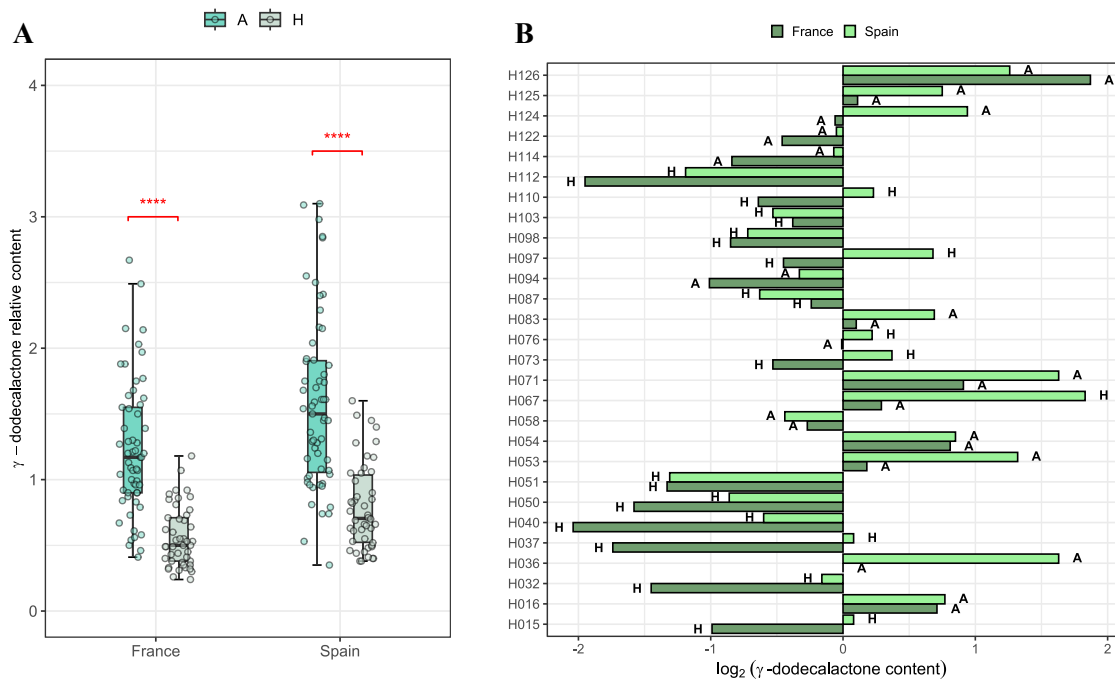


Figure 19. Genetic markers for QTLs controlling γ -dodecalactone content. (A) relationship between genetic marker and γ -dodecalactone content; (B) Goodberry individuals with contrasting γ -dodecalactone content and their genetic markers. Genetic markers: 10212 (A) and AX-166512942_Fa3-2 (H). Significance is show as “****” for p-value < 0.0001 for a Mann-Whittney *U* test.

Clearly, marker A was associated with a high content of γ -dodecalactone content while marker H seemed to be linked to lower levels of this volatile. With this information, the whole GBP was segregated into two groups depending on the molecular marker detected

for each individual. These two groups were either associated with higher content of γ -dodecalactone and another with lower content of this volatile. This segregation was presented in **Fig. 19B**. In consequence, genotypes H040, H050, H051, H098 and H112 were selected as representatives of low γ -dodecalactone content in the population, while H016, H036, H054, H071 and H126 showed high γ -dodecalactone levels to be considered as integrands of the high content group.

Next, transcriptome analysis (RNA-seq) was performed for these genotypes belonging to the two groups, genotypes showing high and low γ -dodecalactone levels. The transcriptome analysis was done in ripe fruits from genotypes grown in two locations, France and Spain. These RNA-seq experiments were analyzed separately using limma R-package and the results were represented in two independent Volcano plots. For these plots, we compared the expression between genotypes displaying high and low content of γ -dodecalactone in both locations to obtain the differentially expressed genes (DEGs) (**Fig. 20**).

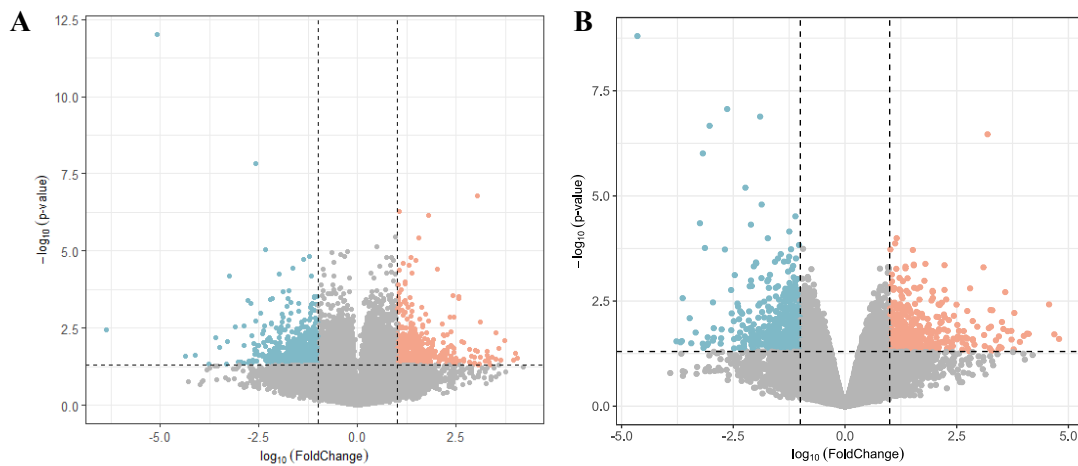


Figure 20. Volcano plots for RNA-seq selected lines from two different locations. Same individuals grown in France (**A**) and Spain (**B**) were used for the RNA-seq analysis. Lines represent the significance criteria used to select the genes: p -value < 0.05 and $\log_{10}FC > 1$ or $\log_{10}FC < -1$. Blue dots represent down-regulated genes while orange dots represent up-regulated genes.

These RNA-seq analysis revealed a total of 56,974 transcripts expressed in the fruits from Spain while fruits from France offered 107,054 transcription products, being this last location more prominent in the number of possible candidate genes. The threshold established to determine significant expression level was p -value < 0.05 while the criteria to categorize transcripts as down- and up-regulated genes were fold-change < 0.1 ($\log_{10}(FC) < -1$) for down-regulated genes and fold-change > 10 ($\log_{10}(FC) > 1$) for up-regulated genes. Following these criteria, we found a total of 521 down and 397 up-

regulated genes in samples from Spain while samples from France yielded 608 down-regulated genes and 492 up-regulated genes.

When comparing the size of the two mQTL selected for γ -dodecalactone, qfDodeLac3 and qfDodeLac4, we observed that they possessed different genetic distance. While qfDodeLac3 are localized in 6 cM including a total of 255 genes, the mQTL for qfDodeLac4 was 4 cM, comprising 146 genes.

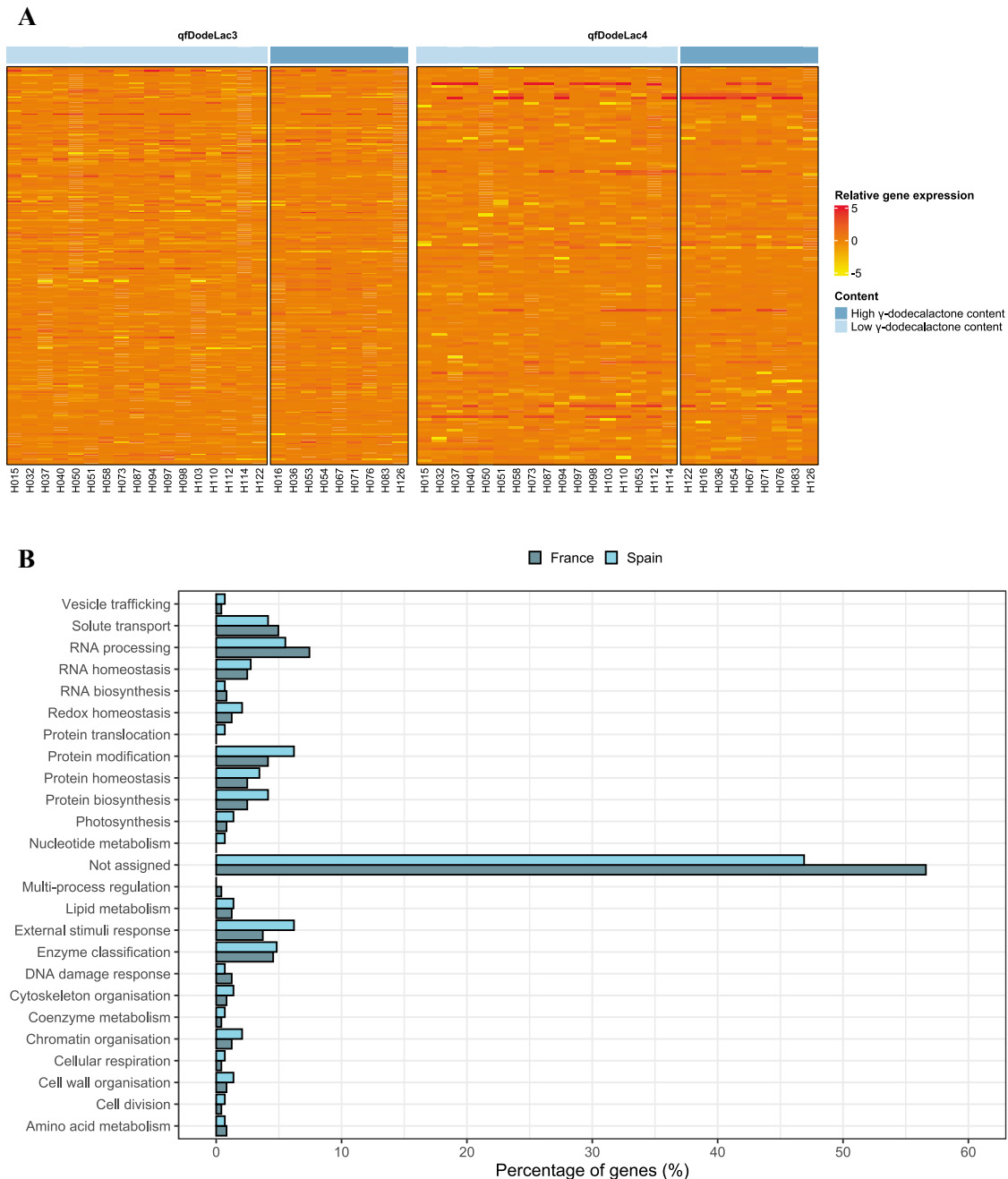


Figure 21. Gene expression and gene functional annotation of qfDodeLac3 y qfDodeLac4 for contrasting γ -dodecalactone genotypes. (A) Relative expression in $\log_2(\text{CPM})$ of genes contained within both QTLs; (B) Functional annotation of genes. Gene were annotated using the online tool of Mercator4.

Next, we studied the expression and functional annotation of the genes belonging to each QTL interval. As it is shown in **Fig. 21**, the majority of these genes were classified as “not assigned” representing almost 50 to 60% for Spain and France, respectively. The next most represented categories were those related to RNA processing and homeostasis and protein biosynthesis, modification and homeostasis, external stimuli response, and enzyme classification.

Some genes related with external stimuli were annotated as Toll-interleukin resistance (TOR) domain family proteins, tobacco mosaic virus (TMV) resistance protein or NB-ARC domain-containing disease resistance protein. About enzyme classification category, some represented genes were G-type lectin S-type receptor-like serine/threonine protein kinases, β -glucosidases or UDP-glucosyltransferases. In **Annex 7** more detailed information about gene functional annotations was summarized.

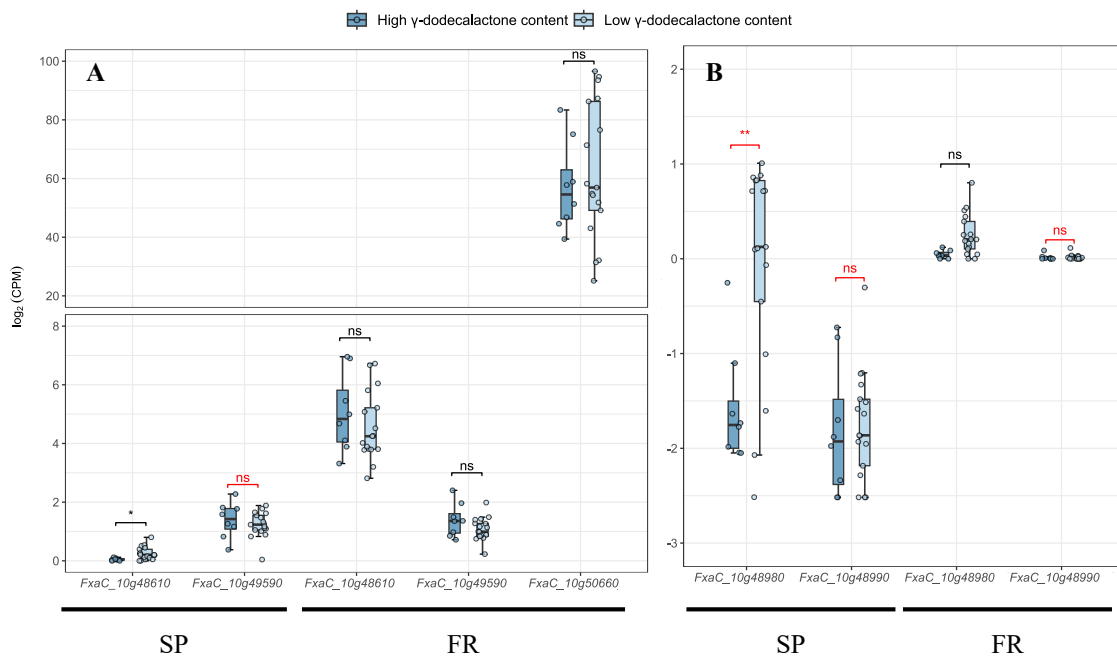


Figure 22. Lipid-related and lipoxygenase genes expression between contrasting γ -dodecalactone lines. (A) Lipid metabolism related genes; (B) Lipoxygenase genes. Statistical significance is shown as “*” for a p-value < 0.05 and “” for a p-value < 0.005. Two statistical tests are shown in this graph: Student’s t-test (in black) and Mann-Whitney *U* test (in red). SP: Spain; FR: France.**

Taking a closer look at genes related to lipids, the expression of some genes when comparing the contrasting γ -dodecalactone groups were observed that only two genes, *FxaC_10g48610* and *FxaC_10g48980*, from Spain data showed statistically significant differences in between the contrasting groups. This suggested that these two genes could be selected as possible candidates for qfDodeLac3 and qfDodeLac4 and, therefore, they can be pointed as candidate genes for controlling the content of γ -dodecalactone. The

expression of both genes is lower when the concentration of γ -dodecalactone is higher, which indicates a possible negative relationship between the expression of these two candidate genes and γ -dodecalactone levels (**Fig. 22**).

On **Table 10** the functional annotation and the expression data of both pre-selected genes were summarized. For further study, we selected *FxaC_10g48980* based on its expression level in both location, France and Spain. In particular, it was the only gene that presented a p-value under 0.05 in both independent RNAseq analyses.

FxaC_10g48980 was categorized as a linoleate 13S-lipoxygenase. Lipoxygenases (LOXs) are commonly classified into two families: 9-LOX and 13-LOX, based on the position at which they oxygenate their respective fatty acid substrates. In *F. × ananassa*, several LOXs genes have been identified and are typically named using the prefix *FaLOX* followed by a number. In accordance with this established nomenclature, we have renamed *FxaC_10g48980* as *FaLOX2*.

Table 10. Summary of RNAseq data of selected genes. logFC refers to the change between contrasting γ -dodecalactone levels. FC: Fold-change; FDR: False discovery ratio

Gene	Functional annotation	RNAseq Spain			RNAseq France		
		logFC	p-value	FDR	logFC	p-value	FDR
<i>FxaC_10g48610</i>	MSF1-like family protein	0.16	0.43	0.93	0.11	0.64	1.00
<i>FxaC_10g48980</i>	Linoleate 13S-lipoxygenase	-1.99	$4.03 \cdot 10^{-4}$	0.67	1.00	$1.85 \cdot 10^{-2}$	0.95

4. Expression pattern of *FaLOX2* in strawberry

We investigated the expression levels of *FaLOX2* across various tissues and developmental stages in *F. × ananassa*. For this purpose, we utilized tissue-specific RNAseq data from Sánchez-Sevilla *et al.* (2017), which includes gene expression profiles in achenes, receptacle, leaves, and roots, as well as four ripening stages for both achenes and receptacle. As references, we compared the expression of *FaLOX2* with the expression of commonly used qRT-PCR reference genes in strawberry, including *FaDBP*, *FaGAPDH2*, and *FaACT1* (Landi *et al.*, 2014; Ye *et al.*, 2021). The expression profiles of *FaLOX2*, alongside these reference genes were shown on **Fig. 23**.

Overall *FaLOX2* expression was lower compared to housekeeping genes mentioned above. In the ‘Camarosa’ RNAseq dataset, *FaLOX2* expression peaked in leaf tissue and progressively decreased during fruit ripening. The highest expression was observed in green achene (GA) and green receptacle (GR) and white receptacle (WR) stages. In leaves, *FaLOX2* expression reached approximately 60 FPKM, which was comparable to the lowest expression levels observed for *FaACT1* and *FaDBP*. Specially, *FaACT1* and

FaDBP expression ranged from 50 to 200 FPKM while *FaGAPDH2* showed markedly higher expression, exceeding 2,500 FPKM.

As shown in **Fig. 23**, the highest expression of *FaLOX2* was observed in leaf tissue. Within strawberry fruit, *FaLOX2* expression was the highest in green achenes as well as during the green and white stages of the receptacle. The highest expression of *FaLOX2* was observed in leaf tissue. However, within strawberry fruit, *FaLOX2* showed the highest expression in green achenes as well as in the green and white stages of the receptacle.

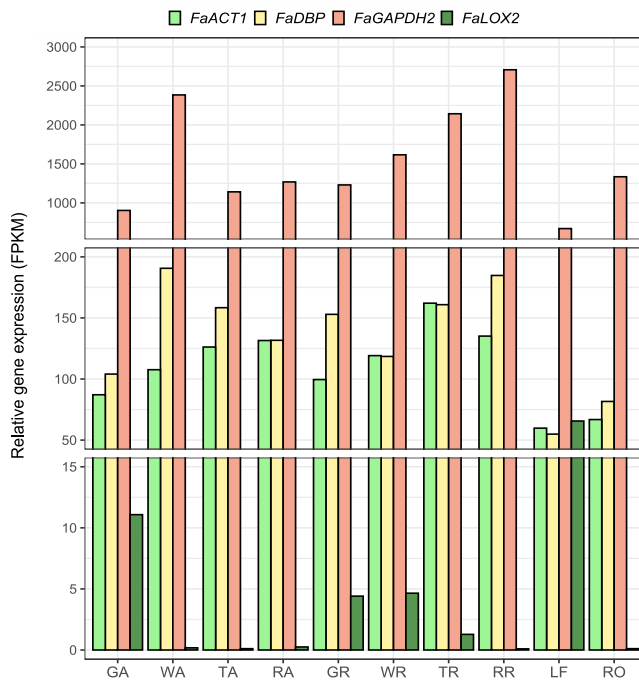


Figure 23. *FaLOX2* expression pattern in strawberry. Data from ‘Camarosa’ RNAseq (Sánchez-Sevilla *et al.*, 2017) in different developmental stages (G, green; W, white; T, turning; R, red) and different strawberry tissues (A, achene; R, receptacle; LF, leaf; RO, roots); The expression levels of reference genes *FaACT1*, *FaDBP* and *FaGAPDH2* were included for comparisons to assess *FaLOX2* expression relative to commonly used housekeeping genes.

Given the low endogenous expression levels of *FaLOX2*, the silencing of the gene was deemed unsuitable. Instead, subsequent experiments will focus on scenarios involving *FaLOX2* overexpression.

5. Phylogenetic analysis of *FaLOX2*

Next, in order to better characterize the potential function of *FaLOX2* gene, we next performed a phylogenetic analysis including *FaLOX2* and several other lipoxygenases from different plant species: *A. thaliana*, *Cucumis melo*, *Cucumis sativum*, *F. × ananassa*, *F. vesca*, *Malus x domestica*, *Olea europaea*, *Prunus dulcis*, *Prunus persica*, *Pisum sativa*, *Solanum lycopersicum*, *Solanum tuberosum* and *Vitis vinifera*. Most of the *F. × ananassa* LOXs were identified using the PF00305 hidden Markov model, which corresponds to the lipoxygenase domain. Using this hidden Markov model, 49 distinct LOX genes were retrieved from the GDR database. Additional LOX sequences were included in the analysis based on literature reports and manual searches for the PF00305 motif.

Subcellular localization ■ Cytoplasm ■ Chloroplast ■ Nucleus ■ Cell membrane

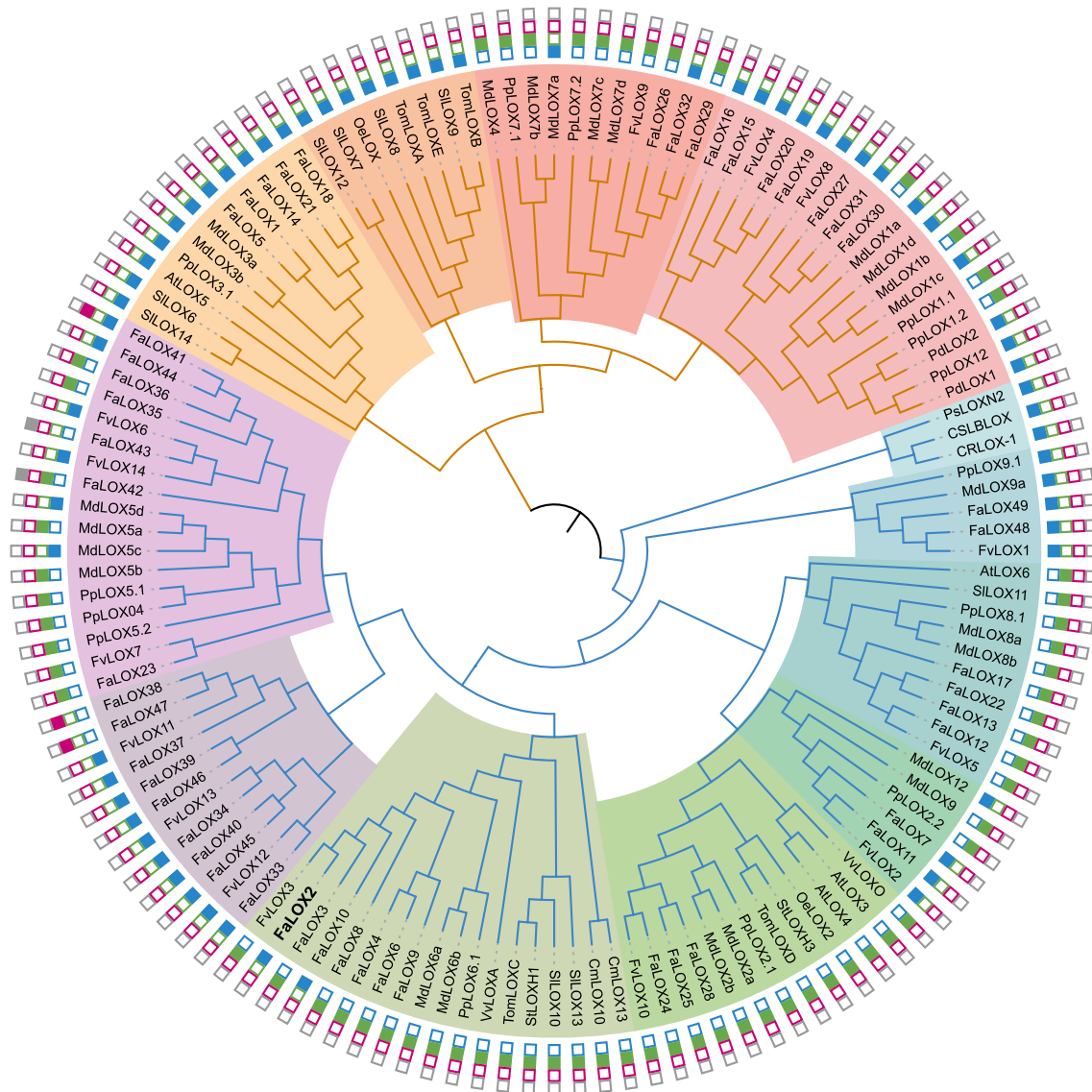


Figure 24. Phylogenetic tree of lipoxxygenases. The present tree was based on amino acid sequences from several members of *Rosaceae* family (apple, strawberry, almond and peach), but also other species like tomato or *Arabidopsis thaliana*. Brown-color lines represent the 9-LOX subfamily of lipoxxygenases while blue-color lines represent the 13-LOX subfamily. The species used in this study are *Arabidopsis thaliana* (At), *Cucumis melo* (Cm), *Cucumis sativum* (CR and CS), *Fragaria × ananassa* (Fa), *Fragaria vesca* (Fv), *Malus x domestica* (Md), *Olea europaea* (Oe), *Prunus dulcis* (Pd), *Prunus persica* (Pp), *Pisum sativa* (Ps), *Solanum lycopersicum* (Sl or TomLOX), *Solanum tuberosum* (St) and *Vitis vinifera* (Vv). For each sequence the colored squares represent the subcellular localization: blue (cytoplasm), green (chloroplast), magenta (nucleus) and grey (cell membrane).

All selected LOX sequences were analyzed phylogenetically, and results were summarized on **Fig. 24**. Additional details, including protein accession numbers and

phylogenetic group classifications were displayed on **Annex 8**. As observed in the phylogenetic tree, the 9-LOX family appeared less diverse than the 13-LOX family in terms of evolution history. The 13-LOX clade was more closely related to the common ancestor of lipoxygenases and encompasses a wide range of proteins with varied subcellular localizations. In contrast, the 9-LOX family appeared to be more specialized with its members grouped into four different clades: 9-A, 9-B, 9-C and 9-D, respectively.

A total of 135 LOX proteins were included in the phylogenetic analysis. Among them, 85 clustered within the 13-LOX family, representing more than 60% of the total dataset. In contrast, the 9-LOX family was represented by 50 different proteins and exhibited fewer diversification events compared to the 13-LOX clade. Most proteins within the 13-LOX family were predicted to localize in the chloroplast, consistent with the presence of an N-terminus chloroplast transit peptide. However, a subset of 13-LOX proteins was predicted to localize in other compartments, including the nucleus or cell membrane. On the other hand, 9-LOX proteins were predicted to be localized either in the cytoplasm or chloroplast, with subcellular localization largely determined by the presence or absence of the chloroplast transit peptide determinant in the compartmentalization of the proteins (see **Annex 8**).

FaLOX2 was positioned within 13-F clade, which includes LOX proteins from several species, most prominently *F. vesca* or *F. × ananassa*. Its closest phylogenetic homologs were identified as FvLOX3 and FaLOX3 and FaLOX10. Notably, FaLOX2, along with FaLOX10, were an exception in terms of predicted subcellular localization within the 13-F clade. Other members of the same clade possessed chloroplast transit peptides and were likely targeted to the chloroplast. Despite the difference in predicted localization, FaLOX2 shared close phylogenetic relationships with FvLOX3 and FaLOX3. Additionally, the presence of TomLOXC from *S. lycopersicum* (Shen *et al.*, 2014) within the same clade suggested a possible homolog of FaLOX2.

6. Structural analysis of FaLOX2 protein

LOX are non-heme iron dioxygenases whose catalytic Fe^{3+} center is coordinated octahedrally by three histidines, the terminal carboxylate of an isoleucine, one asparagine and a water molecule. This ionic form is essential for O_2 insertion into poly-unsaturated fatty acid (Murphy, 2008).

AlphaFold3 prediction of FaLOX2 (see **Fig. 25** and **Annex 9**) yielded a compact, predominantly α -helical monomer (approx. 82%) and only 3 regions with β -sheet disposition. Focusing on the metal site (**Fig. 25B**), AlphaFold correctly places Fe^{3+} in an octahedral pocket ligated by His277, His282, His469, Asn473 and the C-terminal Ile619, matching the generic plant LOX motif and the residues proposed by Singh *et al.*, (2022). A bound water occupies the sixth coordination position, completing the sphere.

The FaLOX2 protein had a dominant negative Coulombic surface potential, which indicated a high presence of negatively charged amino acid exposed to the surface in a neutral pH medium. Furthermore, the protein presented a polar surface (**Fig. 25D**), which produces hydrophilic characteristics to the protein. The fact that FaLOX2 had charged and, therefore, hydrophilic surface would make it suitable of interacting with a polar medium, like the cellular cytoplasmic environment.

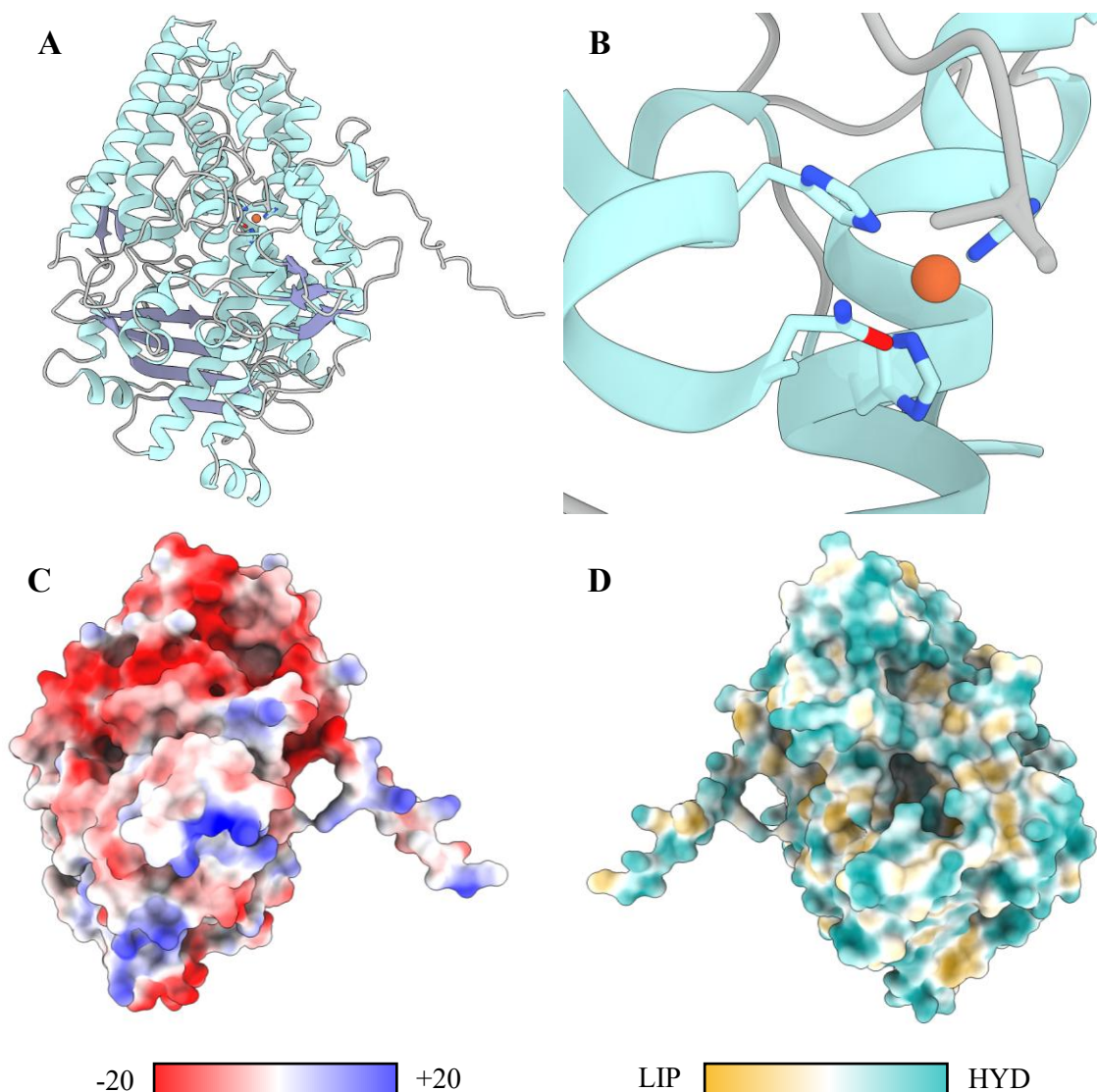


Figure 25. FaLOX2 protein structural analysis. (A) Protein structure predicted by AlphaFold3; (B) Protein iron-binding (Fe^{3+}) motif coordinated with 5 amino acid residues; (C) Coulombic surface potential; (D) Hydrophilic surface coloring (HYD: Hydrophilic, LIP: Lipophilic or hydrophobic). Structural visualizations were performed with ChimeraX.

To explore substrate binding, the FaLOX2 protein was co-modelled with α -linolenic acid using the AlphaFold3 algorithm. The resulting complex showed high confidence across both global and interface metrics, with predicted template modelling (pTM) and interface

predicted template modelling (ipTM) exceeding 0.9, indicating excellent alignment quality and interface accuracy.

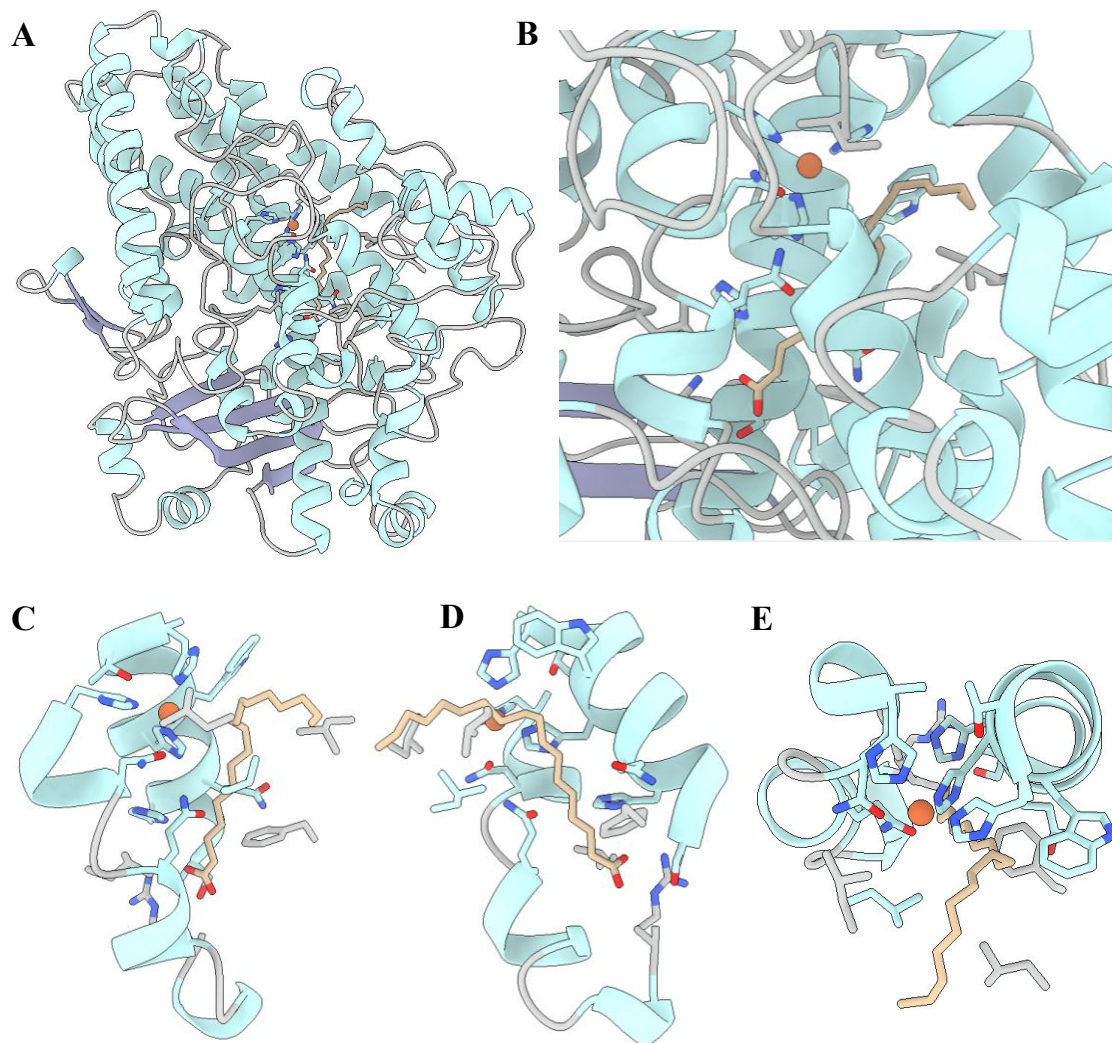


Figure 26. FaLOX2 protein interaction with α -linolenic acid using AlphaFold3. pTM= 0.97; ipTM= 0.95. (A) Global view of predicted model; (B) Closer view of the Fe^{3+} motif; (C-E) Different views of amino acid residues interacting with both Fe^{3+} atom (orange) and α -linolenic acid. Structural visualizations were performed with ChimeraX.

These scores support the reliability of the predicted FaLOX2- α -linolenic acid interaction geometry. The substrate was positioned in a hydrophobic channel oriented toward the Fe^{3+} catalytic site, consistent with known LOX substrate-binding modes (Schneider *et al.*, 2007; Singh *et al.*, 2022). This configuration facilitates dioxygen insertion at the central (C-13) position of α -linolenic acid, supporting the functional assignment of FaLOX2 as a 13S-lipoxygenase. For better visualization we only displayed α -linolenic acid, Fe^{3+} atom and the amino acids involved in the interaction with both substrate and metal atom (**Fig. 26C-E**).

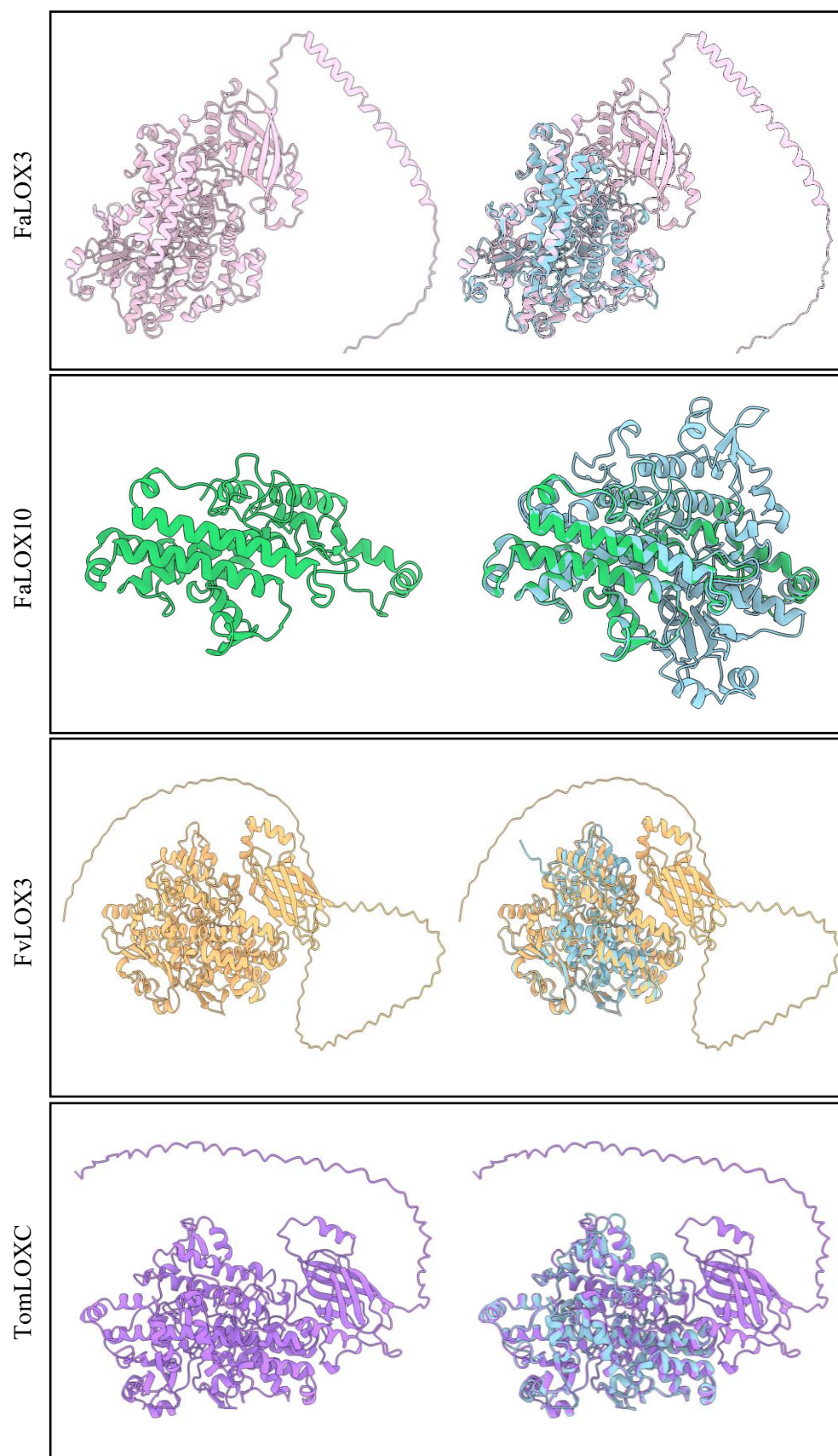


Figure 27. Lipoxigenases protein structural comparison. For each protein, on the left their respective protein structure prediction by AlphaFold3 and on the right the superposition with the FaLOX2 protein (cyan).

The prediction highlighted the five amino acids used for coordinating with the Fe^{3+} atom: His277, His282, His469, Asn473 and Ile619. Nevertheless, new amino acid residues were highlighted for maybe being involved in the interaction with the α -linolenic acid: Arg486, Gln476, Gln273, His272, Leu131, Leu324, Leu535, Phe335, Ser269 and Trp278.

In addition to FaLOX2, several phylogenetically related LOX proteins were modelled using AlphaFold3, including FvLOX3, FaLOX10 and TomLOXC (**Fig. 27**). FvLOX3, FaLOX10 were selected due to their high sequence similarity to FaLOX2 with *Fragaria* species, while TomLOXC (from *S. lycopersicum*) was included for its known function as a 13-LOX involved in volatile organic compounds (Shen *et al.*, 2014). Prediction results from AlphaFold3 were listed with quality parameters on **Annex 9**.

Among them, FaLOX10 presented the simplest structure, comprising only 306 amino acids, arranged predominantly in α -helices, with a minor β -sheet region. Despite partial structure alignment with FaLOX2, FaLOX10 lacked a complete Fe^{3+} coordination sphere, with only three conserved residues (His156, Asn160 and Ile306). The absence of the full triad of histidines necessary for metal binding likely precludes catalytic activity, as reflected in AlphaFold's low-confidence prediction for metal coordination in this protein.

In contrast, FaLOX3, FvLOX3 and TomLOXC each adopted a canonical LOX fold, including a complete Fe^{3+} binding site with five coordinating residues (three His, one Asn and one Ile). Although these residues are found at non-identical positions due to sequence length differences, their spatial arrangement in the 3D structures was consistent with functional LOX enzymes. A particular conserved feature across these three proteins was the presence of an N-terminal disordered region followed by a well-structured jelly roll fold. This fold, composed of two four-stranded antiparallel β -sheets, formed a hydrophobic cavity lined with apolar residues, consistent with previously reported membrane interaction or signal peptide domains. In all three proteins, this structural motif was confined to the first 200-300 amino acids, suggesting a shared architecture in their N-terminal domains.

FaLOX3 and FvLOX3 displayed strong structural homology with FaLOX2, especially in the catalytic region, whereas TomLOXC, while still aligning well, showed greater sequence divergence. Sequence alignment (**Annex 10**) indicated that TomLOXC shared approximately 50% identity and 60% similarity with FaLOX2, supporting its use as a comparative model for functional inference.

7. Transient overexpression of *FaLOX2* induces changes in strawberry volatilome

Here, we observed that the expression of *FaLOX2* gene was significantly different between control fruits and 35S-*FaLOX2* fruits considering either *FaDBP* or *FaGAPDH2*

as housekeeping genes, being an overexpression between 2.5 to 10.0 times. The endogenous expression levels of *FaLOX2* in strawberry fruit were relatively low compared to constitutively expressed housekeeping genes (**Fig. 23**). To gain insight into the function role of *FaLOX2* during strawberry fruit development, we performed a transient overexpression assay using the constitutive *CaMV35S* promoter in *F. × ananassa* cv. ‘Candonga’ under greenhouse conditions. Overexpression was conducted via agroinfiltration using the pK7GW2D-*FaLOX2* construct, with empty vector (pK7GW2D) as a negative control (**Table 7** and **Annex 1**). Expression levels were quantified by qRT-PCR, employing *FaDBP* and *FaGAPDH2* as housekeeping genes as previously described by Landi *et al.*, (2014) and Ye *et al.*, (2021). A total of ten *FaLOX2*-overexpressing fruits (F) and seven control fruits (P) were analyzed (**Fig. 28**). The results showed that *FaLOX2* was significantly upregulated in infiltrated fruits, with expression levels ranging from 2.5 to 10.0-fold higher than controls, regardless of the housekeeping gene used for normalization. These data confirm the effectiveness of the transient expression system and establish a suitable platform for downstream functional analyses.

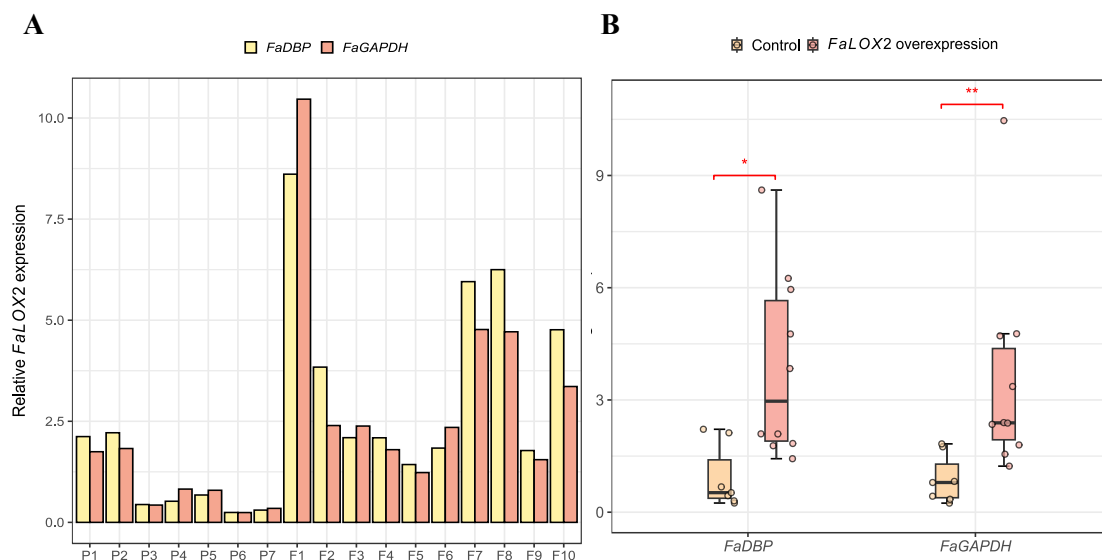


Figure 28. Relative gene expression of *FaLOX2* in agroinfiltrated ‘Candonga’ fruits. (A) Individual data from each agroinfiltrated fruit (P, control; F, overexpression). **(B)** *FaLOX2* expression boxplot. Cq values from *FaLOX2* gene were relativized with *FaDBP* and *FaGAPDH2* and with the control values. Relative gene expression is represented as $2^{-\Delta\Delta Cq}$ data. Significance is show as “*” for p-value < 0.05 and “**” for p-value < 0.005 for a Mann-Whittney *U* test.

From this gene expression experiment, two experimental groups of samples were established: fruits agroinfiltrated with the empty vector pK7GW2D (control, designed as P1-P7) and fruits transiently overexpression *FaLOX2* via the pK7GW2D-*FaLOX2* construct (named F1-F10). Seven independent fruits were selected as control fruits (from

P1 to P7) and ten independent fruits were selected as *35SCaMV-FaLOX2* overexpressed fruits (from F1 to F10). Each fruit represented an independent biological replicate, given that individual agroinfiltration events were performed on separate ‘Candonga’ fruits. Despite the inherent variability in transformation efficiency, each fruit was treated as an autonomous unit for both gene expression and metabolite profiling.

To investigate metabolic changes associated with *FaLOX2* overexpression, we conducted a comprehensive analysis of volatile organic compounds (VOCs), as well as primary and secondary metabolites, using each sample independently. A principal component analysis (PCA) was conducted to explore global variation in volatile profiles between control and *FaLOX2*-overexpression fruits (**Fig. 29**). The first two principal components accounted for approximately 60% of the total variation, with the first dimension (PC1) contributing 47% while the second dimension (PC2) contributing nearly 13%. Out of 16 total components, only PC1 and PC2 explained more than 10% of the total variance individually. The highest contribution of volatiles to PC1 was approximately 2.5%, attributed to linalool, while the highest contribution to PC2 exceeded 6%, corresponding to heptanal.

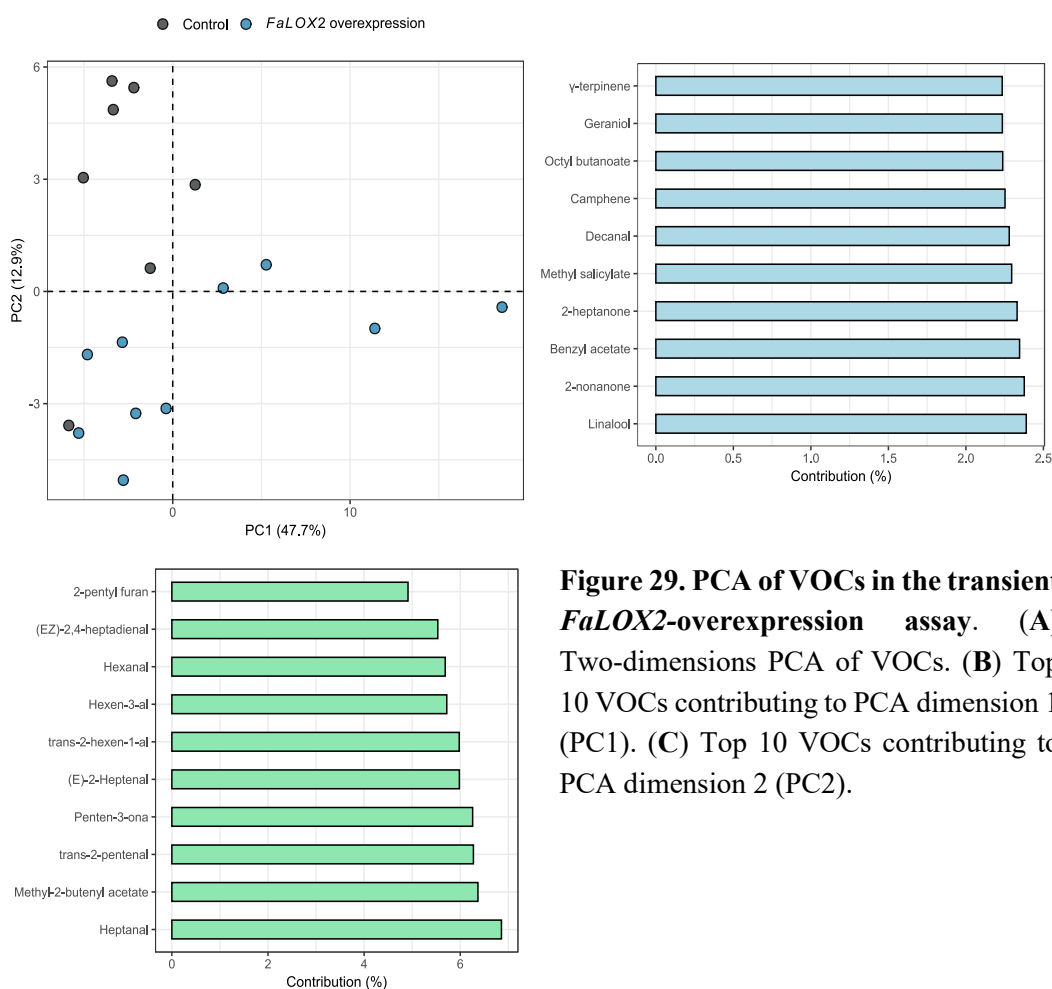


Figure 29. PCA of VOCs in the transient *FaLOX2*-overexpression assay. (A) Two-dimensions PCA of VOCs. (B) Top 10 VOCs contributing to PCA dimension 1 (PC1). (C) Top 10 VOCs contributing to PCA dimension 2 (PC2).

Overall, the independent fruit samples were clearly separated into two main groups based on their treatment: control (P) and *FaLOX2*-overexpressing (F) fruits. The PCA results revealed that PC1 predominantly separated *FaLOX2*-overexpressing samples, while PC2 contributed to further segregation between control and overexpressing fruits. Notably, one control fruit clustered with the *FaLOX2*-overexpression group, which may be explained by its elevated levels of several VOCs contributing strongly to PC2, particularly those among the top 10 most influential volatiles. Each individual fruit was used as a biological replicate for volatile compound extraction and analysis.

The results of the volatile metabolite profiling following transient *FaLOX2* overexpression are summarized in **Fig. 30**. A total of 84 different volatile compounds were identified and quantified, as presented in **Table 4**. These compounds were classified according to their chemical structure into seven different categories: alcohol, aldehyde, alkane, ester, ketone, lactone and terpene. Additionally, the volatiles were grouped by their carbon chain length, resulting in 13 carbon-based classes, ranging from 3 to 15 carbon atoms. This classification provides insight into the structural diversity and biosynthetic origins of the volatiles affected by *FaLOX2* activity.

To further assess the variation in volatile profiles between *FaLOX2*-overexpressing and control fruits, a hierarchical clustering analysis (HCA) was performed using Pearson correlation as the distance metric. Interestingly, we observed higher levels of ester compounds (39 esters from a total of 84 VOCs) when comparing *FaLOX2*-overexpressing to control fruits. Aldehydes, the second most abundant class (15 compounds), generally decreased in concentration in the *FaLOX2*-overexpression group, in some cases by as much as two-fold compared to controls. When VOCs were classified by carbon chain length, compounds containing 10 carbon atoms were the most prevalent (19 chemicals), comprising primarily terpenes and a subset of esters. Eight-carbon compounds formed the second largest group, largely dominated by esters. The terpene class displayed a strong clustering tendency, likely due to the presence of numerous structural isomers within the dataset. Similarly, esters also clustered tightly, consistent with their increased accumulation in response to *FaLOX2* overexpression. These clustering patterns suggest a coordinated modulation of specific metabolic pathways, particularly those involved in ester and terpene biosynthesis. Among the detected volatile compounds, only two lactones were detected: γ -decalactone and γ -dodecalactone. Both compounds increased their respective concentrations when comparing 35S-*FaLOX2* to control fruits. Remarkably, γ -decalactone levels were up to 100-times higher in some *FaLOX2*-overexpressing fruits, while γ -dodecalactone levels increased approximately three-fold. This upregulation mirrored the pattern observed for most ester compounds, whereas classes such as aldehydes and ketones generally showed a reduction in abundance in response to *FaLOX2* overexpression.

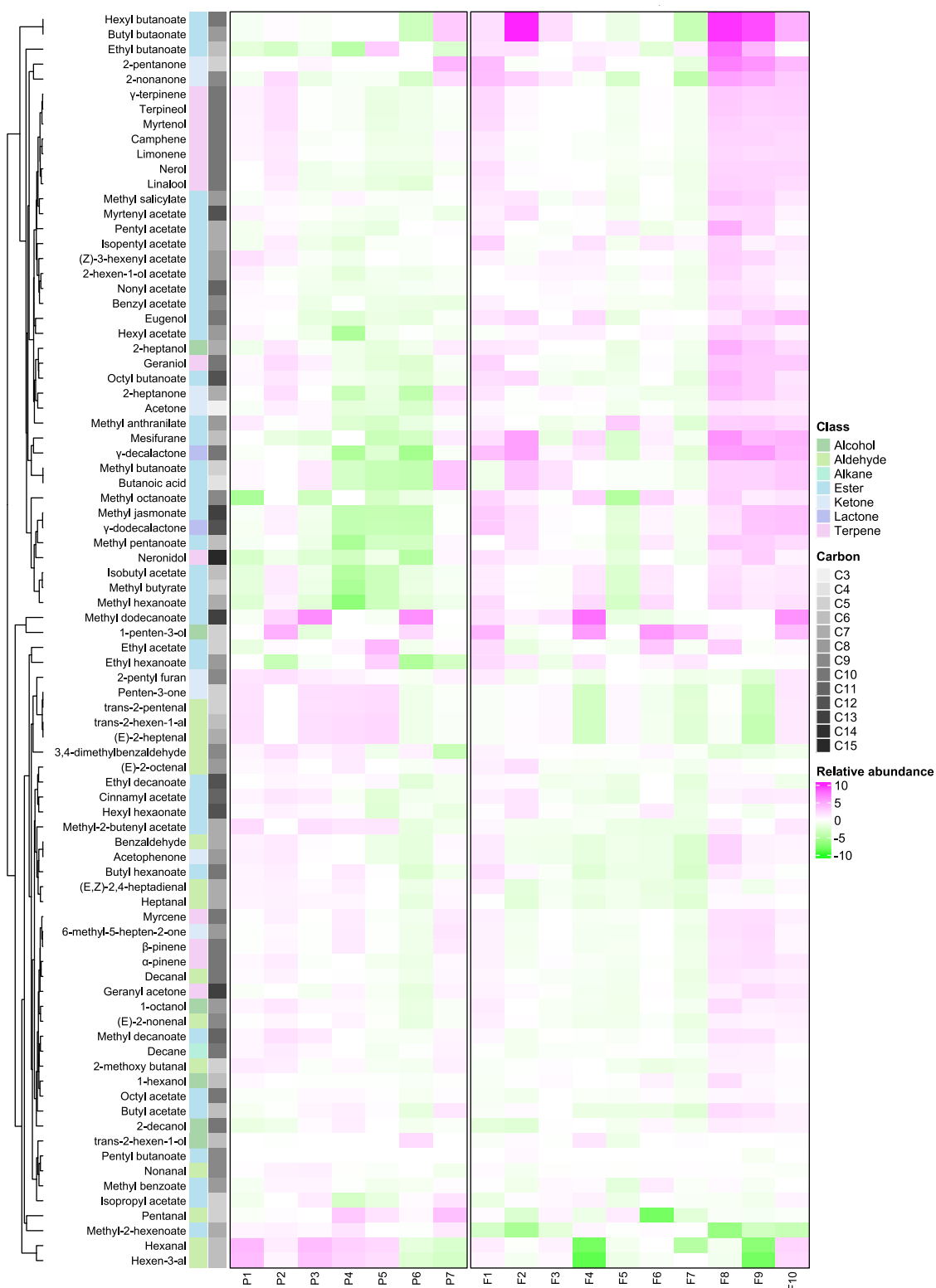


Figure 30. Volatile organic compounds profile of *FaLOX2* transient overexpression. The graph displays the accumulation of volatiles relativized to D₃₂-pentadecane intensity for each sample. Volatile organic compounds are classified into 7 chemical classes and according to the number of carbon atoms.

V. Chapter I: Characterization of candidate gene for γ -dodecalactone content in strawberry fruit

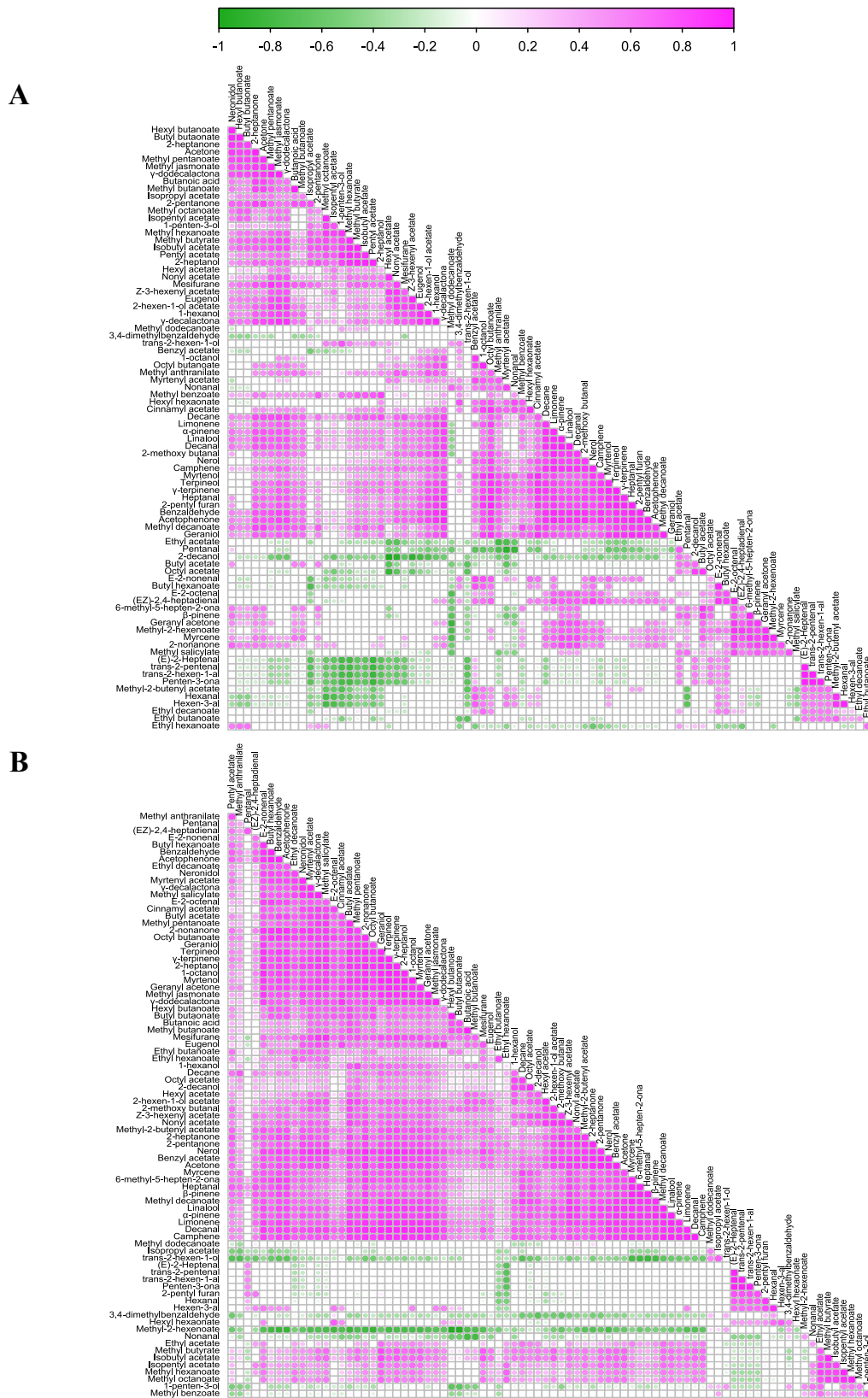


Figure 31. Spearman correlation of VOCs in *FaLOX2* transient overexpression. (A) VOCs Spearman correlation of control fruits (agroinfiltrated with empty vector). **(B)** VOCs Spearman correlation of 35S-*FaLOX2* overexpressed fruits. Only significant (p -value < 0.01) Spearman's rank correlation were displayed in **A** and **B**. Both correlograms were displayed with hierarchical clustering (HCA).

Additionally, several volatiles with a high loading in the PCA (**Fig. 29**) were significantly altered by *FaLOX2* overexpression. Key contributors to PC2 included aldehydes and ketones, such as heptanal, trans-2-pentenal, and penten-3-one, each contributing over 6% to PC2 variance. Notably, these compounds all showed a consistent decrease in concentration in *FaLOX2*-overexpressing fruits, reinforcing the observed metabolic shift toward esters and lactones, and away from aldehyde- and ketone-related pathways.

Spearman correlations analysis of volatile compounds was conducted independently for control (P) and *FaLOX2*-overexpressing (F) fruit samples to examine metabolic coordination under differing gene expression contexts. Using the dataset of 84 volatile organic compounds (VOCs) per group, approximately 3,500 pairwise correlations were calculated for each condition (**Fig. 31**). The *FaLOX2*-overexpressing fruits exhibited a greater number of positive correlations, suggesting enhanced co-accumulation of VOCs under elevated *FaLOX2* expression. In contrast, control fruits showed a larger proportion of negative correlations, indicating more divergent volatile behavior in the absence of overexpression.

A particularly strong positive correlation ($p > 0.9$) was observed between γ -dodecalactone and methyl jasmonate in both experimental conditions. This correlation remained stable across treatments and was further emphasized by their co-clustering in the *FaLOX2*-overexpressing group. This relationship suggests a possible metabolic linkage between jasmonate biosynthesis, the lipoxygenase (LOX) pathway, and the formation of specific lactones, likely due to their shared fatty acid precursors (see **Figs. 5, 6, and 9**). Moreover, volatile accumulation patterns confirmed that *FaLOX2*-overexpressing fruits, particularly samples F8, F9, and F10, showed a higher overall abundance of VOCs, notably esters and terpenes, compared to controls. Conversely, certain aldehydes exhibited an inverse pattern, increasing in the overexpression condition, despite the overall reduction of this class (**Fig. 30**).

In total, 16 VOCs were significantly over-accumulated, and 8 VOCs were significantly down-accumulated in *FaLOX2*-overexpressing fruits. Interestingly, although γ -decalactone and γ -dodecalactone are structurally similar, their correlation patterns diverged. γ -decalactone clustered more closely with mesifurane, a key aromatic compound in strawberry known for its desirable sensory profile (Barbey *et al.*, 2021), whereas its association with γ -dodecalactone was relatively weaker. This distinction may reflect differential regulation or metabolic routing within the lactone biosynthetic network.

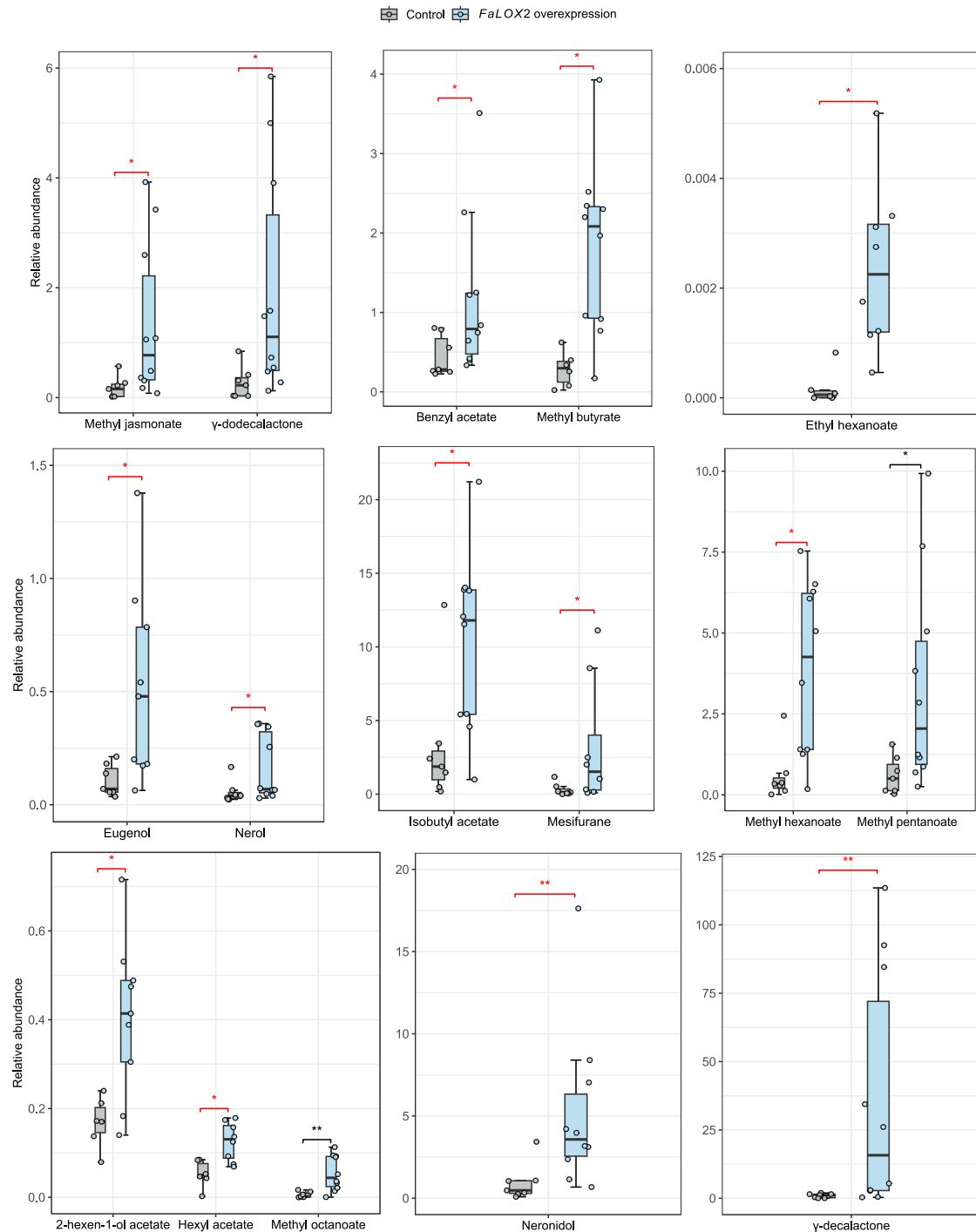


Figure 32. Differentially up-accumulated VOCs in *FaLOX2* transient overexpressing fruits. Statistical significance is shown as “*” for a p -value < 0.05 and “**” for a p -value < 0.005 . Two statistical tests are shown in this graph: Student’s t -test (in black) and Wilcoxon signed-rank test (in red).

Taking a deeper look at the VOCs content in between control and *FaLOX2*-overexpressing ‘Candonga’ fruits reveals distinct differences in metabolite accumulation (Fig. 32). Among the VOCs significantly up-accumulated in the overexpression

condition, γ -dodecalactone and methyl jasmonate stand out prominently. This finding supports the selection of *FaLOX2* as a candidate gene linked to QTLs controlling variation in these volatiles. The observed increase in methyl jasmonate is likely attributable to its shared biosynthetic pathway with γ -dodecalactone, suggesting coordinated regulation mediated by *FaLOX2*. Additionally, γ -decalactone is also elevated in *FaLOX2*-overexpressing fruits, possibly reflecting the biosynthesis of multiple related lactones (e.g., γ -deca- and γ -dodecalactone) derived from common fatty acid precursors.

Other VOCs with relatively low abundance in ‘Candonga’ fruits but influenced by *FaLOX2* overexpression include esters such as ethyl hexanoate, hexyl acetate, and methyl octanoate. Among terpenes, the levels of nerol and nerolidol appear to be positively affected by *FaLOX2* overexpression. In contrast, several aldehydes containing 5 to 7 carbon atoms exhibit significant down-accumulation in *FaLOX2*-overexpressing fruits (**Fig. 33**). Similarly, volatiles such as penten-3-one, 2-pentyl furan, and methyl-2-hexenoate show decreasing trends consistent with those aldehydes. Only one ester, methyl-2-hexanoate, displayed a marked reduction, approximately five-fold, upon *FaLOX2* overexpression. The ketone penten-3-one also demonstrated a notable decrease in concentration under the same conditions.

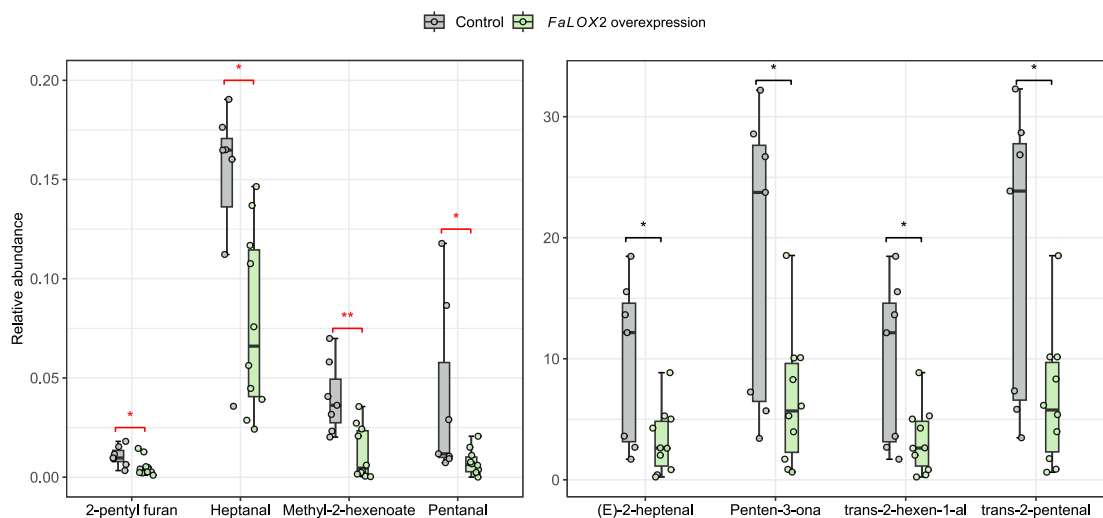
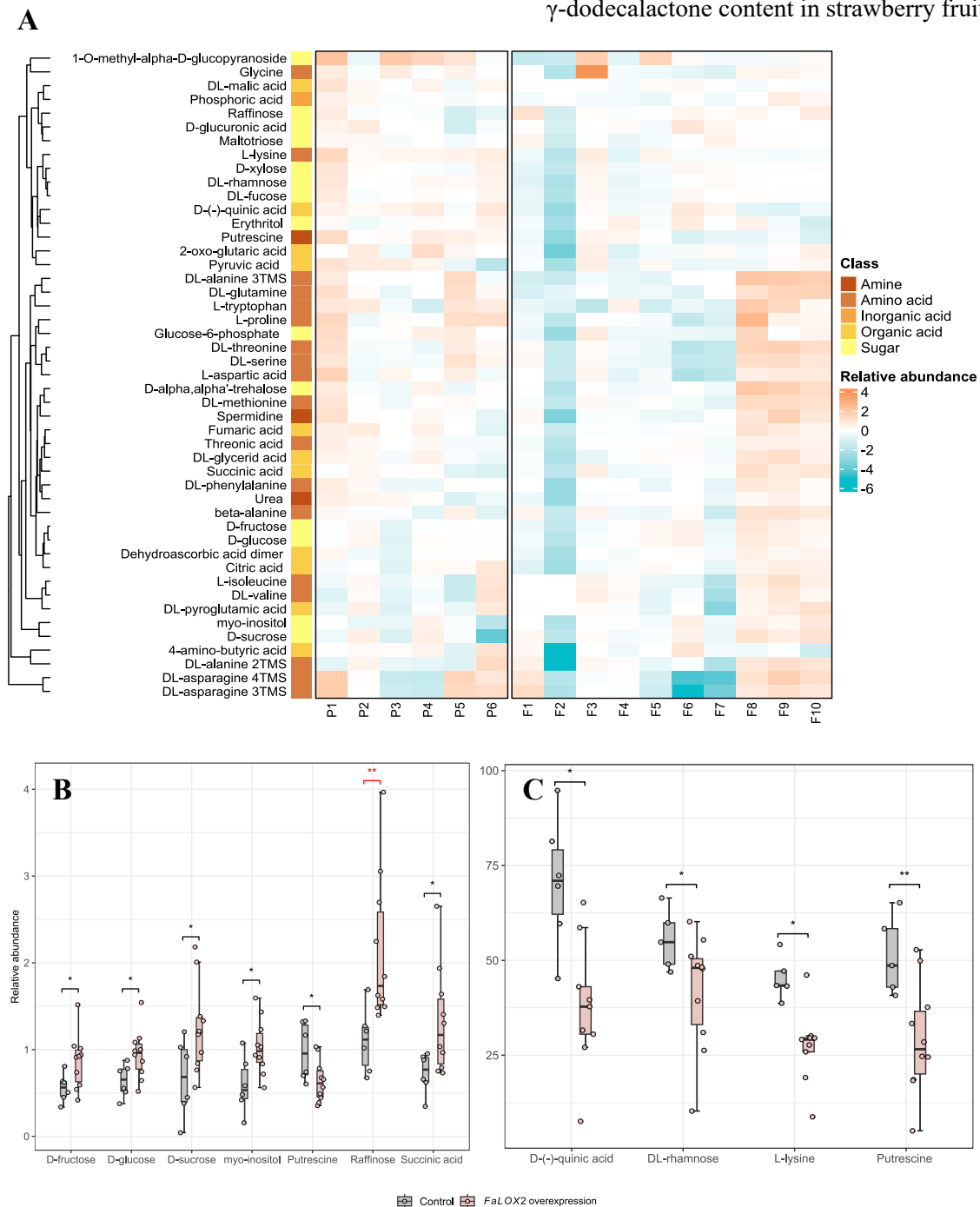


Figure 33. Differentially down-accumulated VOCs in *FaLOX2* transient overexpressing fruits. Statistical significance is shown as “*” for a p-value < 0.05 and “**” for a p-value < 0.005. Two statistical tests are shown in this graph: Student’s t-test (in black) and Wilcoxon signed-rank test (in red).

8. Primary and secondary metabolism changes in *FaLOX2*-overexpression

Next, we explore primary and secondary metabolism in the transient overexpression experiment. For the analysis of primary metabolism, only polar compounds were extracted and analyzed using Gas Chromatography-Mass Spectrometry. A summary of compounds detected for primary metabolism were presented on **Fig. 34A**.



Among the primary metabolites identified, sugars and amino acids were the most abundant categories, followed by organic acids, particularly those from the Krebs cycle. The heatmap (**Figure 34A**) revealed several trends, including an overall accumulation of

sugars in most *FaLOX2*-overexpressing fruits (F samples). An exception was sample F2, which exhibited lower sugar levels compared to other F samples, a pattern not observed in the VOC data. Conversely, samples F8, F9, and F10 displayed high levels of most primary metabolites. The primary metabolites most affected by *FaLOX2* overexpression included sugars, nitrogen-containing compounds (such as amines and amino acids), and, to a lesser extent, organic acids from the Krebs cycle. Notably, the main sugars found in strawberry fruit, glucose, fructose, and sucrose were significantly accumulated in the overexpressing samples (**Fig. 34B**). The increase in the levels of these important sugars for strawberry fruits correlates with a higher Brix grade in *FaLOX2*-overexpressing fruits and, therefore, increases the overall sugar perception of these fruits (**Annex 10**). Other important sugars, such as raffinose and *myo*-inositol, showed similar trends. Among the organic acids, succinic acid was the only compound with a marked change.

When metabolite levels were normalized to the dry weight of the samples (**Fig. 34C**), changes independent of water content were observed. For instance, putrescine showed reduced levels in the F samples, consistent with results based on fresh weight. Similarly, the sugar rhamnose and amino acid, lysine also decreased. A significant reduction was also noted in quinic acid, an organic acid involved in the shikimate pathway, which links primary and secondary metabolism. Heatmap of primary metabolites relativized to dry fruit weight could be consulted on **Annex 10**.

Concerning secondary metabolites, as shown in **Fig. 35**, we also observed significant alterations on *FaLOX2*-overexpressing fruits. The heatmap in **Fig. 35A** demonstrated a high diversity of secondary metabolites, classified into ten distinct groups, primarily including phenylpropanoids, phytohormones (indolacetic acid), tannins, and flavonoids. There were clear differences between control and *FaLOX2*-overexpressing samples, particularly in samples P5–P7, which had lower levels of most metabolites under control conditions, but accumulated these compounds upon *FaLOX2* overexpression. All metabolites showing significant changes exhibited a common pattern: increased accumulation in the *FaLOX2*-overexpressing samples compared to controls. The most affected class was hydroxycinnamic acid derivatives, with notable compounds including coumaric acid hexose 1 and 2, and sinapic acid hexose derivatives 1 and 2. The second most affected class was anthocyanins, represented by cyanidin hexose 1, pelargonidin malonyl hexose, and pelargonidin hexose. Additional changes were observed in benzoic acid derivatives (e.g., di-hydroxy methyl benzoic acid hexose, hydroxybenzoic acid hexose), flavan-3-ols (e.g., (epi)-afzelechin, phloretin hexoses 1 and 2), and flavonols (e.g., kaempferol coumaroyl hexose 3).

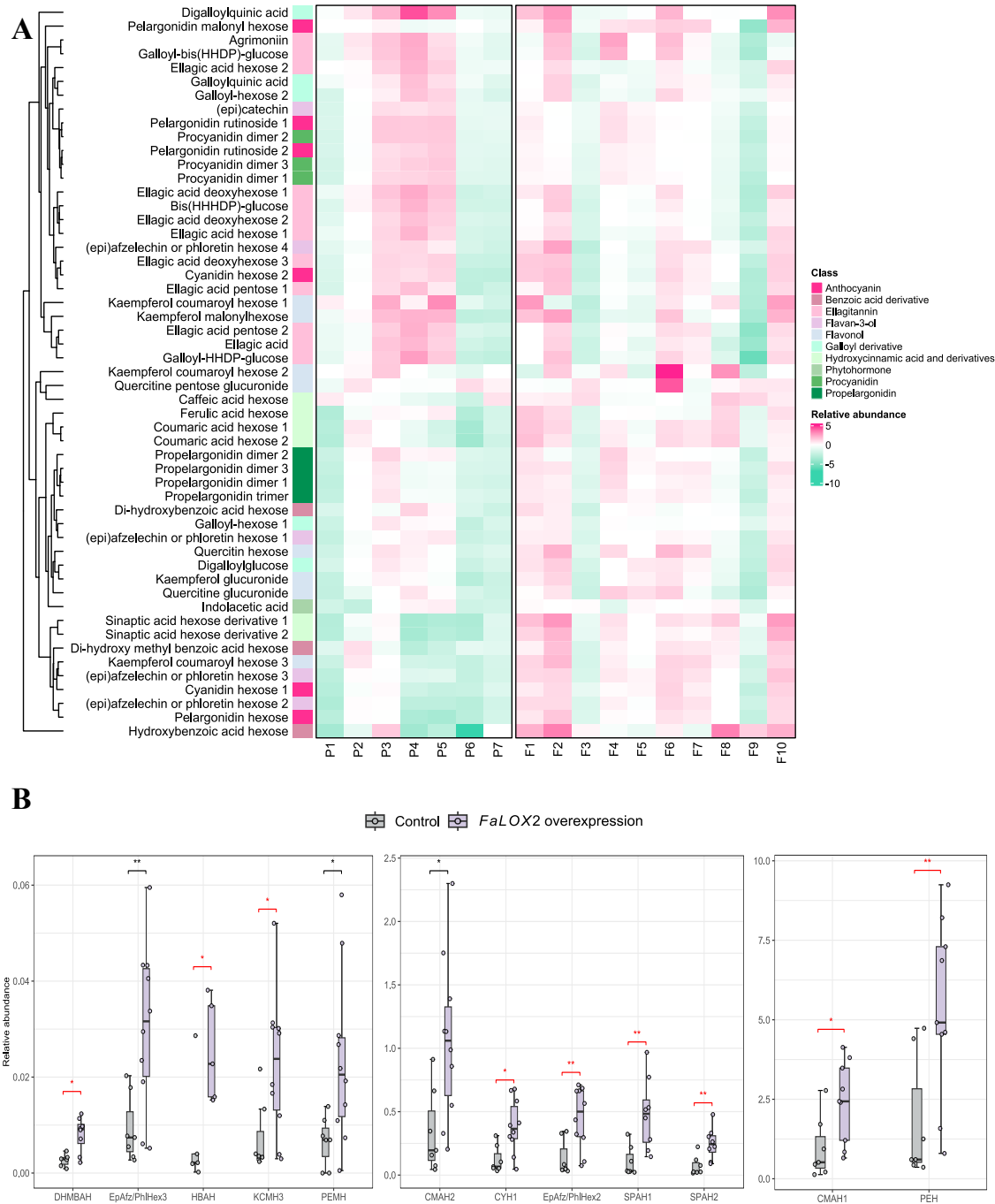


Figure 35. Changes in secondary metabolism in transient *FaLOX2* overexpression. (A) Heatmap of secondary metabolism; **(B)** Statistically significant changes in primary metabolites. Statistical significance is shown as “*” for a p-value < 0.05 and “***” for a p-value < 0.005. Two statistical tests are shown in this graph: Student’s t-test (in black) and Mann-Whitney *U* test (in red). EpAfz/PhlHex2: epiafzelechin or phloretin hexose 2; EpAfz/PhlHex3: epiafzelechin or phloretin hexose 3; CMAH1: Coumaric acid hexose 1; CMAH2: Coumaric acid hexose 2; SPAH1: Sinaptic acid hexose derivative 1; SPAH2: Sinaptic acid hexose derivative 2; KCMH3: Kaempferol coumaroyl hexose 3; PEH: Pelargonidin hexose; PEMH: Pelargonidin malonyl hexose; CYH1: Cyanidin hexose 1; HBAH: Hydroxybenzoic acid hexose; DHMBAH: Di-hydroxy methyl benzoic acid hexose.

9. *FaLOX2* expression and VOCs accumulation during ripening process

Strawberry fruits of the cultivar ‘Camarosa’ were collected at three different developmental stages: green, white and red. For each stage, three independent biological replicates were harvest, with each replicate consisting of a pool of five fruits. Prior to analysis, achenes were removed to specifically study the volatiles produced by the edible receptacle tissue.

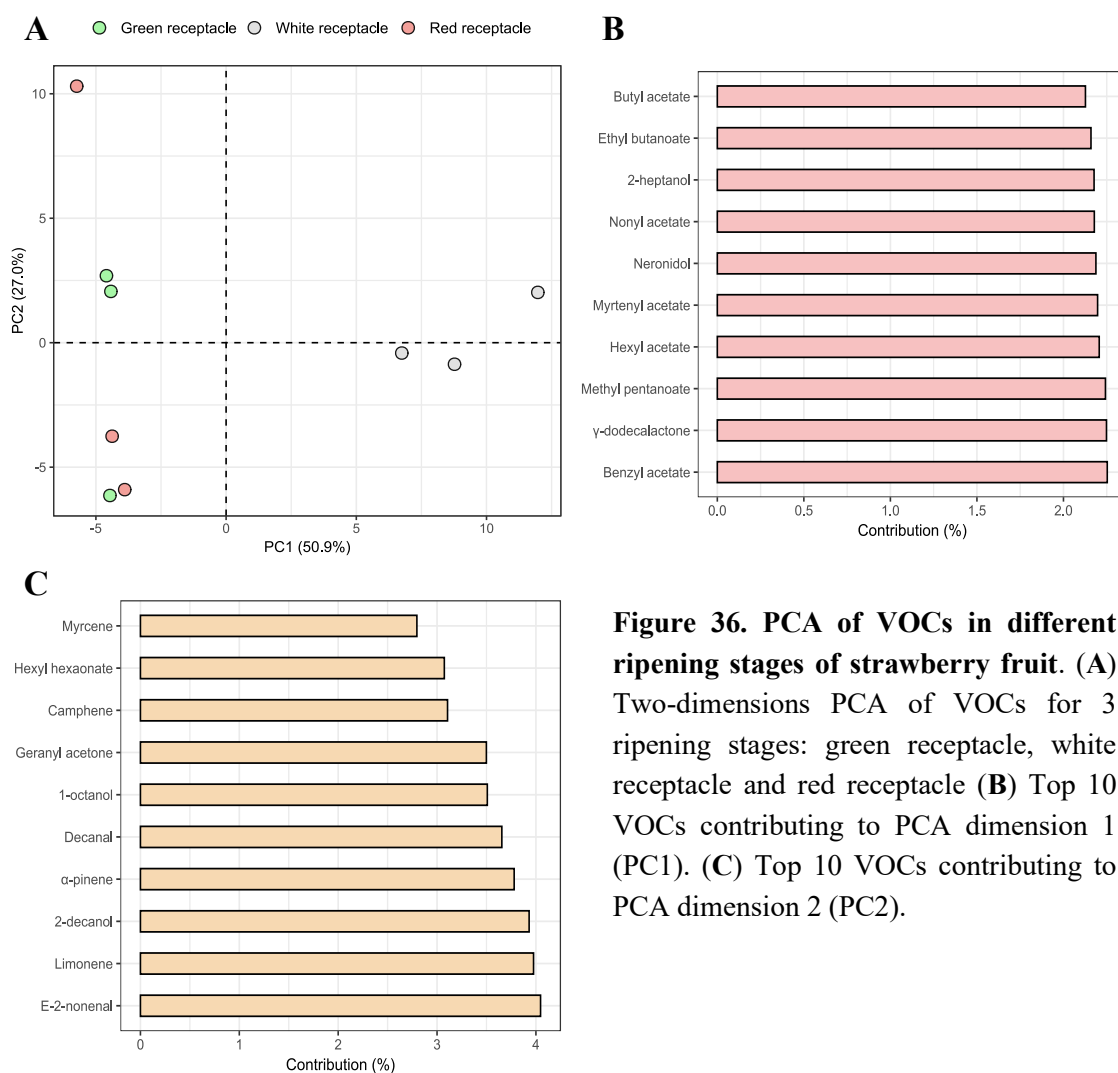


Figure 36. PCA of VOCs in different ripening stages of strawberry fruit. (A) Two-dimensions PCA of VOCs for 3 ripening stages: green receptacle, white receptacle and red receptacle (B) Top 10 VOCs contributing to PCA dimension 1 (PC1). (C) Top 10 VOCs contributing to PCA dimension 2 (PC2).

The variability in VOCs across the different developmental stages was assessed using PCA based on a dataset comprising 81 detected volatiles. The results of the PCA are presented in **Fig. 36**. This multivariate analysis explained nearly 80% of the total variability within the dataset. The first component (PC1) contributed approximately 51%, while the second component (PC2) accounted for 27% of the total variance. PCA revealed that white stage samples formed a distinct cluster, clearly separated from the green and red stages along PC1. In contrast, green and red stages were discriminated along PC2, although no complete separation was observed between these two stages, suggesting some overlap in their VOC profiles.

Based on the PCA loading scores, benzyl acetate and γ -dodecalactone were identified as the major contributors to PC1, each explaining approximately 2.5% of the total variance. In PC2, the compounds with the highest contributions were *E*-2-nonenal and limonene, both showing loading values twice those observed for PC1.

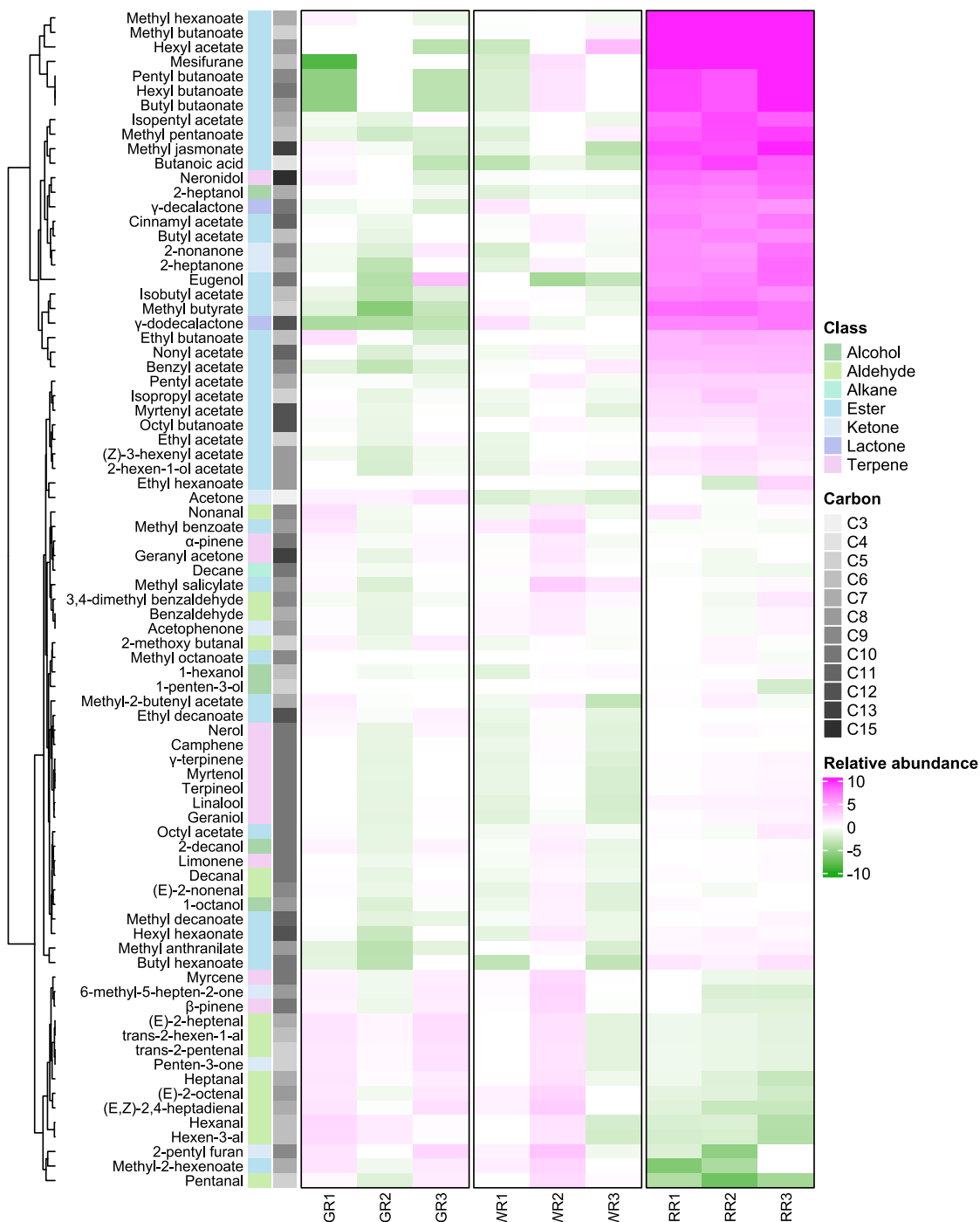


Figure 37. Changes in volatile organic compounds during strawberry fruit ripening. Three developmental stages of fruit ripening process were selected: GR: green receptacle; WR: white receptacle and RR: red receptacle. Relative intensity was calculated using D₃₂-pentadecane peaks intensity as reference. Volatile organic compounds are classified into 7 chemical classes and according to the number of carbon atoms.

The entire volatile dataset was represented as HCA analysis as shown in **Fig. 37**. Two major trends emerged from this analysis: a group of VOCs that increased in concentration during ripening and another group that decreased as the fruit matured. Nearly 70% of the esters detected in the heatmap showed a clear increase in concentration as ripening progressed. These esters clustered together, likely due to their coordinated accumulation pattern. In addition to esters, several aldehydes, terpenes, lactones and ketones exhibited significantly higher concentrations in the ripe receptacle (red) compared to the green or white developmental stages.

Conversely, a notable depletion of VOCs was observed in certain chemical classes. Approximately 60% of the aldehydes showed significantly reduced levels during ripening. Similarly, some ketones, such as penten-3-one and terpenes, including myrcene and β -pinene, also declined in concentration. While many other VOCs displayed ripening-associated changes, their shifts were less pronounced than those observed in esters or aldehydes. Notably, most terpenes increased moderately at the red receptacle stage. In general, VOCs concentrations in the green and white stages were close to the detection limits of the analytical equipment, suggesting that the volatile production is minimal in the immature stages. Some examples of changes in volatiles concentrations with different ripening stages are shown in **Fig. 38**.

A total of 81 volatile compounds were detected in samples across the three developmental stages. From this initial number, only 36 compounds exhibited statistically significant changes based on one-way ANOVA. The ester class, which dominates strawberry aroma, was the most represented, with 27 out of 36 significant VOCs belonging to this group. All esters exhibited an increase in their respective concentrations from green to red stages. Only three terpenoids showed significant increases during ripening: geraniol, linalool, and nerolidol, with linalool displaying the greatest accumulation. Although 15 aldehydes were detected, none showed statistically significant differences across developmental stages. However, most displayed a general trend of decline, as seen with hexanal and pentanal in **Fig. 38**. Among ketones, only acetone and 2-nonanone significantly increased during ripening.

Key aroma-contributing volatile organic compounds, as identified by Fan *et al.*, (2021), also underwent substantial changes. Mesifurane showed a dramatic 90-fold increase from the white to the red stage, remaining nearly undetectable in green and white stages. Similarly, two major lactones, γ -decalactone and γ -dodecalactone, exhibited similar patterns, peaking in red stage. Among these, γ -decalactone reached a higher relative level (6 units), while γ -dodecalactone was approximately 10 times lower, at around 0.20 units. Finally, methyl anthranilate, another key compound, also showed a significant increase (nearly 4-fold) from green to red stage.

The plant hormone methyl jasmonate followed a similar trend, with minimal levels in green and white stages, and a sharp increase at the red stage, reaching nearly 3 units. This suggests a burst in methyl jasmonate biosynthesis during ripening and supports its regulatory role in strawberry maturation, as proposed by Han *et al.* (2019).

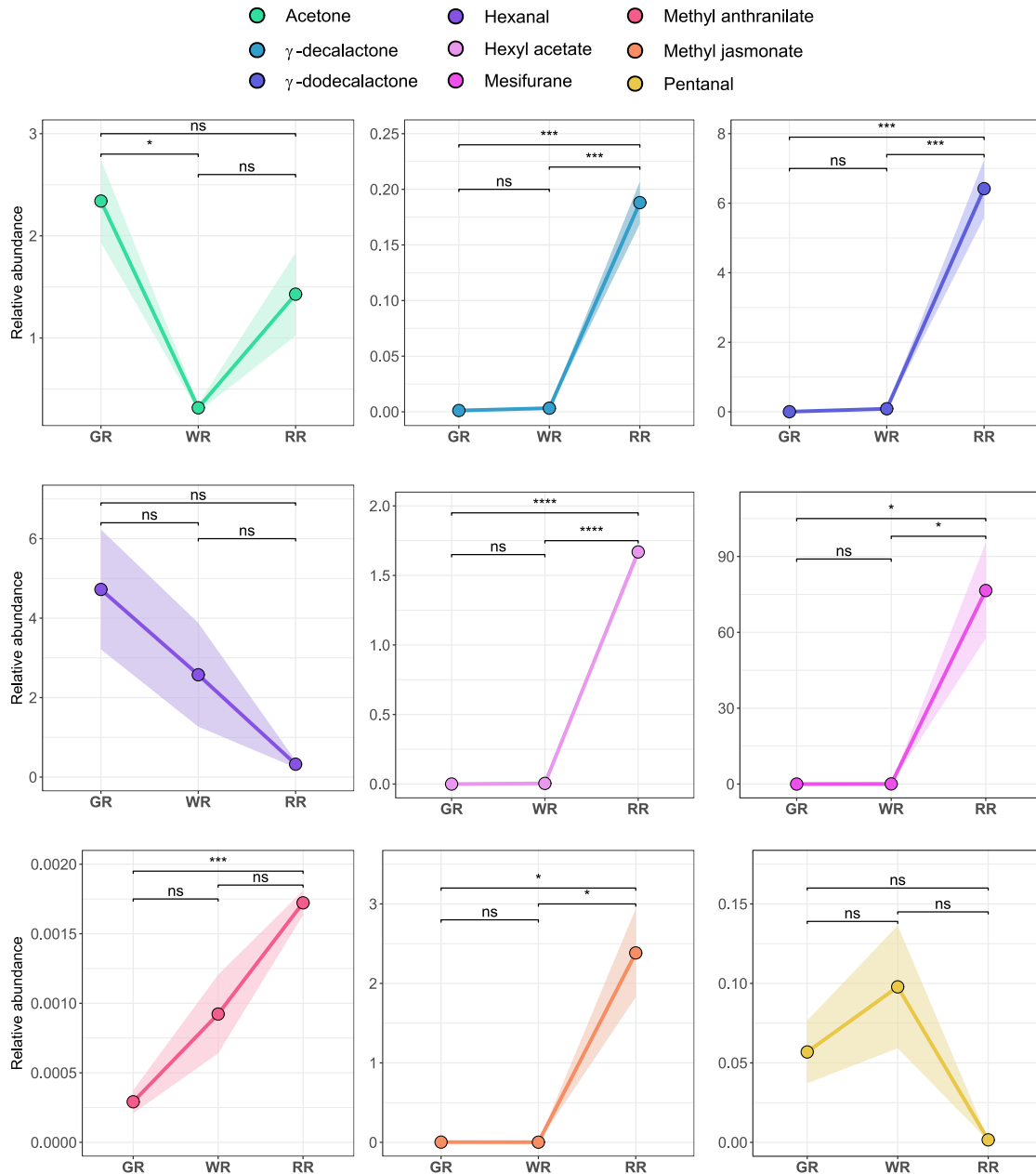


Figure 38. Volatiles changes during strawberry fruit ripening. Data are shown as mean $\pm$ standard error. GR: green receptacle; WR: white receptacle; TR: turning receptacle; RR: red receptacle. Statistical significance is represented as “ns” for non-significant, “*” for p-value < 0.05 and “****” for a p-value < 0.0005, for a one-way ANOVA.

Regarding to gene expression, RNA-seq data from Sánchez-Sevilla *et al.* (2017) indicated that *FaLOX2* expression peaked in leaf tissue (**Fig. 23**). In contrast, *FaLOX2* expression

declined during ripening in both achene and receptacle tissues, with higher expression observed at the green and white stages. Nevertheless, the expression levels of *FaLOX2* remained relatively low compared to the housekeeping gene *FvH4_1g3980* (also known as *FvActin*), which showed FPKM values of approximately 100–150 units in achene and receptacle tissues.

10. FaLOX2 has a cytoplasmic localization

Subcellular localization of FaLOX2 protein was assessed by cloning full-length coding sequence of *FaLOX2* gene in frame to the green fluorescent protein (GFP) reporter under the control of the *CaMV35S* promoter. The fusion protein resulted in the attachment of GFP tag at the C-terminal domain of FaLOX2 using the vector pGWB5 (further information on **Table 7** and **Annexes 9** and **12**).

The resulting construct (*35S::FaLOX2-GFP*) was transiently expressed in *N. benthamiana* leaf epidermal cells via *A. tumefaciens*-mediated infiltration. Fluorescence signals were observed using confocal laser scanning microscopy at 48–72 hours post-infiltration. The FaLOX2-GFP fusion protein exhibited a clear cytosolic localization, with no fluorescence signal detected in other organelles such as endoplasmic reticulum, chloroplasts or mitochondria. As a control, leaves infiltrated with *35S::GFP* alone displayed ubiquitous GFP fluorescence throughout the cytoplasm and nucleus (**Fig. 39**). To localize more accurately the fusion protein FaLOX2-GFP we co-infiltrate this construct together with an endoplasmic reticulum marker, such as FaFAH1 (Ruiz-Lopez *et al.*, 2021; Sánchez-Sevilla *et al.*, 2014). As **Fig. 39C** suggested, there was no co-localization of FaFAH1 and FaLOX2, implicating a different subcellular localization from endoplasmic reticulum.

FaLOX2. In this case, we agroinfiltrated *N. benthamiana* leaves independently using FaLOX2-GFP, GFP and AtBAM1-GFP (cell membrane marker, kindly borrowed from Araceli Castillo's lab). Leaves discs were harvested 48 h post-infiltration and were incubated 10 min in 1 M NaCl to achieve the shrinking of cells' protoplasm. After this treatment, protein strands derived from the retracement of cell membrane could be seen, such as the case of AtBAM1 on **Fig. 39E1-E3**. On the contrary, neither free GFP (**Fig. 39D1-D3**) nor FaLOX2-GFP (**Fig. 39F1-F3**) presented this protein strands, therefore demonstrating that both are not related to cell membranes and are widespread across the cytoplasm.

These observations suggest that FaLOX2 localizes to the cytoplasm, consistent with its proposed role in lipid-derived signaling and volatile biosynthesis, processes known to occur primarily in the cytosolic compartment of plant cells.

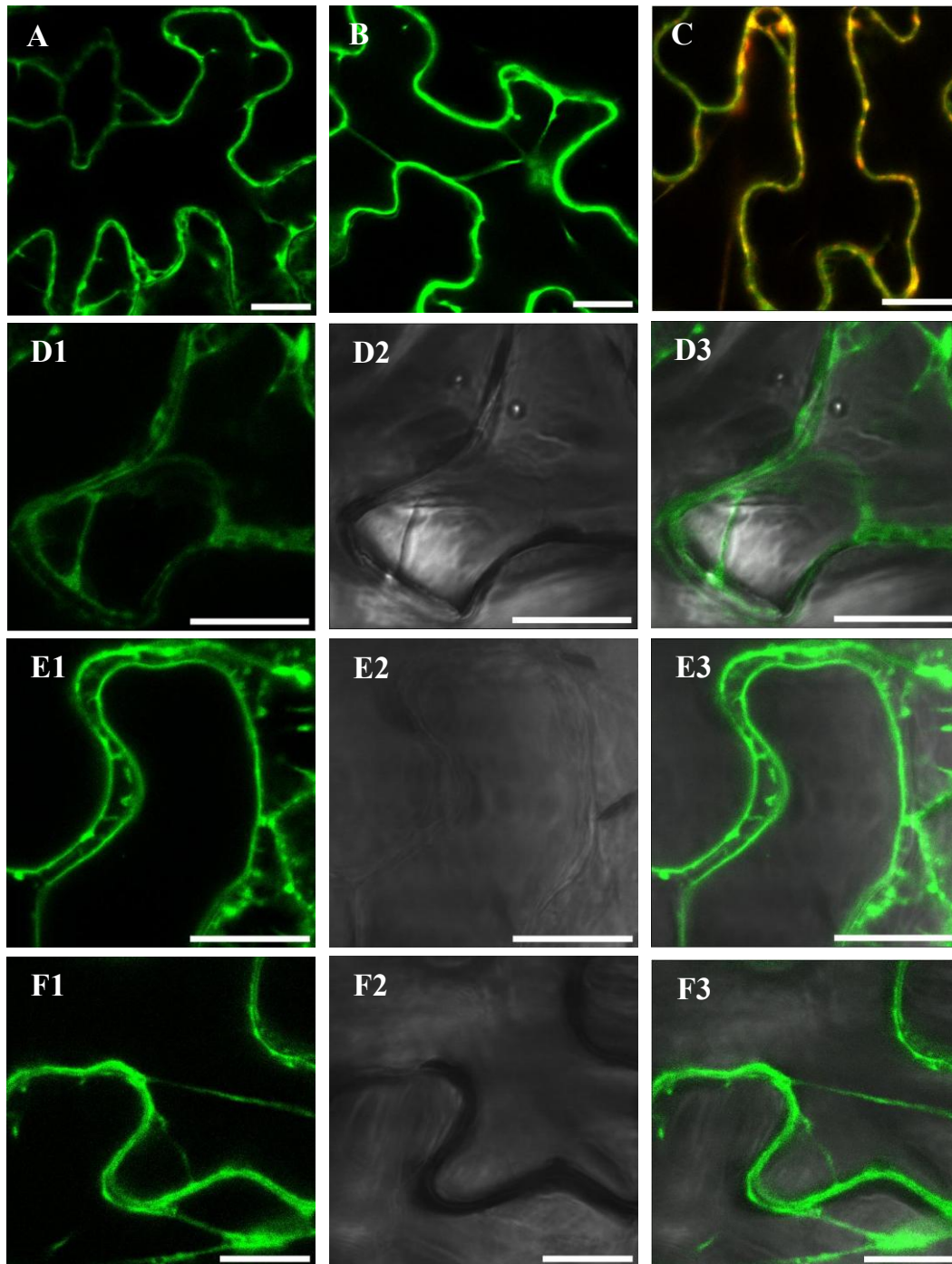


Figure 39. Subcellular localization of FaLOX2 protein. (A) free GFP; (B) FaLOX2-GFP; (C) Co-infiltration of FaLOX2-GFP with FaFAH1-RFP. **D1-F3** comprise a plasmolysis assay with 1 M NaCl during 10 min, where (**D1-D3**) free GFP dark field, bright field and merge; (**E1-E3**) AtBAM1 dark field, bright field and merge and (**F1-F3**) FaLOX2-GFP dark field, bright field and merge. White scale bar represents 20 μ m.

Moreover, plasmolysis assay was performed to discard a cell membrane localization for *N. benthamiana* leaf discs were collected after confocal microscopy and total protein extraction was performed. **Fig. 40** showed the results of the immunodetection for the

samples after confocal microscopy. In this experiment, GFP was detected as a band of 26.77 kDa, FaLOX2 with GFP tag had a molecular mass of 99.36 kDa and AtBAM1 had 138 kDa. Nevertheless, AtBAM1 couldn't be detected by western blot. There was no free GFP besides from sample A, so there was no degradation of the constructs with GFP used as tag.

H

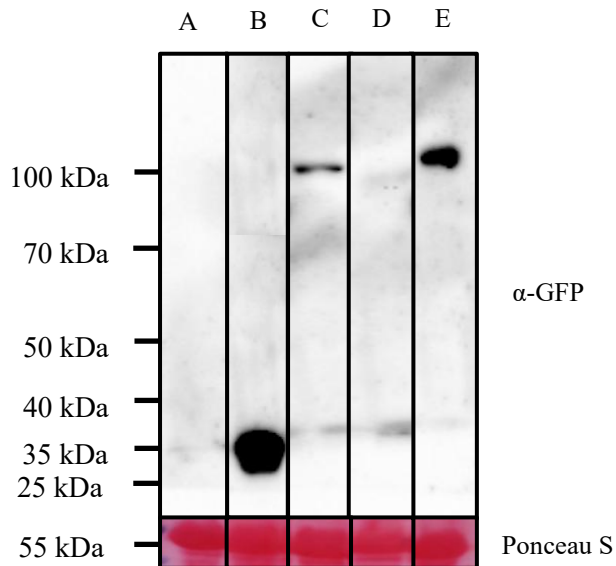


Figure 40. Western blot of GFP for *N. benthamiana* leaves used in confocal microscopy. A: naïve (non-infiltrated sample), B: free GFP, C: FaLOX2-GFP, D: AtBAM1 and E: co-infiltration of FaLOX2 and FaFAH1. Ponceau S: Ponceau staining.

Discussion

1. Volatile content diversity on Goodberry population

The analysis of volatile organic compounds (VOCs) in the ‘Candongra’ $\times$ ‘Senga Sengana’ bi-parental population across five diverse European environments revealed substantial chemical and biological diversity in strawberry volatile profiles. Despite the detection of 77 VOCs from multiple classes including esters, aldehydes, terpenes, and lactones, the principal component analysis (PCA) explained only a modest portion of the variance (PC1: 15.5%, PC2: 7%), suggesting that both genotype and environment interact in a complex manner to shape the overall volatile profile. While some geographical trends were observed such as the partial separation of lines grown in Italy in PC2 and the close grouping of lines grown in France and Spain in PC1, these patterns were not strong enough to differentiate between locations. This implies a high level of plasticity in VOC composition in response to local environmental conditions.

Hierarchical clustering further supported this variability, with individuals grown in France and Spain generally showing higher VOC levels, potentially due to similar climatic conditions, while those grown in Poland and Germany exhibited lower overall volatile content. Lines grown in Italy showed more inconsistent patterns, likely reflecting intermediate or fluctuating environmental effects. Importantly, heritability estimates revealed that certain VOCs, including key aroma contributors such as mesifurane, γ -decalactone, and γ -dodecalactone, are strongly genetically controlled ($h^2 > 0.7$), indicating that these compounds are promising targets for marker-assisted selection in breeding programs aimed at improving fruit aroma. In contrast, volatiles like linalool, which showed very low heritability ($h^2 = 0.07$), appear to be highly influenced by the environment, highlighting the need for environmentally stable conditions or controlled growth settings when targeting such traits. Notably, γ -dodecalactone displayed both high heritability (0.83) and low variability across locations, reinforcing its potential as a robust biomarker for breeding strawberries with consistent sensory quality. Conversely, the dichotomy between acetone ($h^2 = 1.0$) and linalool ($h^2 = 0.07$) emphasizes the range of genetic determinism among VOCs and underlines the importance of dissecting the genetic architecture behind individual aroma compounds.

The genetic control of VOCs with high heritability should be driven by one or multiple genes, being these genes affecting their biosynthesis or catabolism. In fact, some compounds with high heritability have been demonstrated to have a genetic explanation, like *FaQR* for furaneol biosynthesis (Medina-Puche *et al.*, 2019; Xie *et al.*, 2020), *FaNES1* for nerolidol biosynthesis (Medina-Puche *et al.*, 2019) or *FaFAD1* for the biosynthesis of γ -decalactone (Sánchez-Sevilla *et al.*, 2014). The work of Fan *et al.*, (2021) suggested that γ -decalactone followed a bimodal distribution, thus supporting the

hypothesis that the presence of this volatile was controlled by a single gene (Chambers *et al.*, 2014). Among the two more prominent lactones, γ -dodecalactone displayed higher heritability value, indicating a stronger genetic control over this volatile. In fact, the presence of γ -dodecalactone follows a normal distribution, what suggest polygenic control (Fan *et al.*, 2021).

These findings collectively demonstrate that while the environment plays a crucial role in shaping the strawberry volatilome, specific VOCs are under strong genetic regulation, offering clear avenues for genetic improvement of fruit aroma in cultivated strawberries.

2. QTL for volatiles reveals a possible genetic hotspot in Fvb3-2

The QTL analysis for volatile compounds revealed that over 50% of the total QTL identified in this work were located on linkage group Fvb2-3 of the female genetic map. Remarkably, this represented more than 60% of all QTL detected in the female map, indicating a strong enrichment in this specific region.

The diversity of compounds mapping to Fvb3-2 was substantial and included compounds such as acetone, acetophenone, benzaldehyde, decanal, 2-nonanol, γ -decalactone and γ -dodecalactone. These compounds are structurally diverse and originate from distinct biosynthetic pathways. While some like aldehydes and alcohols may share steps in the lipoxygenase (LOX) pathway, such as eugenol (phenylpropanoids pathway), mesifurane (as previously described by Wein *et al.*, 2002; Zorrilla-Fontanesi *et al.*, 2012), and terpineol (a terpenoid), are not biochemically related. Despite this biochemical diversity, QTLs for these volatiles are all mapped to the same genomic region.

Interestingly, the peak LOD scores for most of these QTLs coincided with the position of specific molecular markers: 12929910|F|0-9:A>C-9:A>C (12.97 cM), AX-166512951 (9.32 cM) and 1805207|F|0-15:C>A-15:C>A (4.72 cM). This suggests that the interval between 12.97 and 4.72 cM may represent a QTL hotspot associated with VOC biosynthesis.

One possible explanation for this is the presence of a metabolic gene cluster in this genomic region. Such gene clusters, comprised of non-homologous but functionally related genes involved in secondary metabolism, have been described in several plant species (Nützmann *et al.*, 2016; Bharadwaj *et al.*, 2021). These gene clusters often include genes encoding enzymes for rate-limiting or pathway-committed steps, as well as enzymes responsible for modification of pathway end-products. Notable examples have been identified in *Arabidopsis thaliana* (Field and Osbourn, 2008; Field *et al.*, 2011; Chen *et al.*, 2019), *Oryza sativa* (Shimura *et al.*, 2007; Swaminathan *et al.*, 2009; Wang *et al.*, 2011; Wu *et al.*, 2011), *Solanum lycopersicum* (Itkin *et al.*, 2013; Matsuba *et al.*, 2013, 2015; Zhou and Pichersky, 2020) and *Zea mays* (Frey *et al.*, 1997). To date, however, no

such metabolic cluster has been reported in the *Fragaria* genus. The presence of a gene cluster in this region of *F. × ananassa* may be supported by subgenome dominance and local gene duplications, both of which are known mechanisms for cluster formation (Schlöpfer *et al.*, 2017). This grouping of QTLs for biochemically related volatiles, such as various terpenes or structurally related γ -decalactone and γ -dodecalactone support this possibility. In *S. lycopersicum*, a terpene synthase cluster is thought to have evolved through tandem duplications of terpenoid synthase and *cis*-prenyltransferases genes, even though some genes in the cluster may be non-functional (Matsuba *et al.*, 2013).

An alternative hypothesis is the existence of a master regulator, possibly a transcription factor, in the Fvb3-2 region that modulates VOC biosynthesis or catabolism across multiple pathways. Transcription factors such as *FaRIF* (Martín-Pizarro *et al.*, 2021) and *FaMYB10* (Lin-Wang *et al.*, 2010) have been shown to regulate multiple aspects of fruit metabolism, making them plausible candidates for such regulation. To determine whether this genomic region harbors a metabolic gene cluster or a shared regulatory mechanism, further functional analysis is required. Candidate genes within the Fvb3-2 interval should be subjected to in planta overexpression or silencing experiments to assess their impact on VOC profiles. Additionally, RNA-seq of lines with contrasting volatile profiles could identify differentially expressed genes (DEGs) within this region, shedding light on its functional role in VOC biosynthesis.

More experiments are needed to elucidate whether there is a genetic cluster or if there are common genes responsible for the regulation of some VOCs. Functional studies of genes restricted to Fvb3-2 can be conducted to reveal the possible changes in VOCs with classic overexpression and silencing assays *in planta*. Another approach would be to make RNA-seq of lines with contrasting lines in different volatiles and observe the DEGs within this zone.

3. *FaLOX2* is responsible for the accumulation of the volatile γ -dodecalactone in strawberry fruits

The lines used for RNA-seq were initially selected based on their heterozygosity across multiple traits. However, it is possible that other lines, excluded from this analysis, exhibited greater differences in γ -dodecalactone content. This may explain why the genes listed in Table 10 did not show statistically significant false discovery ratio (FDR) values. Consequently, candidate gene selection was based on functional annotation, as the genes with FDR < 0.05 were associated only with protein or RNA homeostasis.

QTLs for γ -dodecalactone are predominantly located on linkage group Fvb3-2. In this study, we identified one candidate gene for γ -dodecalactone content within this linkage group. Interestingly, a gene previously linked to γ -decalactone (a compound structurally similar to γ -dodecalactone) was also located on the corresponding chromosome candidate

gene (Sánchez-Sevilla *et al.*, 2014). Specifically, this gene, *FaFAD1*, encodes a fatty acid desaturase involved in the synthesis of linoleic acid (18:2) in the endoplasmic reticulum. *FaFAD1* is believed to act on membrane fatty acids, generating intermediates that feed into other metabolic routes, including the LOX pathway. Our hypothesis is that our candidate gene, annotated as lipoxygenase (*FaLOX2*), contributes to the LOX pathway by catalyzing the oxidation of α -linolenic acid. However, *FaLOX2* is unlikely to directly catalyze the transformation into γ -dodecalactone, as this lactone forms via the self-esterification of CoA-activated 4-hydroxydodecanoid acid in filamentous fungi (Silva *et al.*, 2019). Rather, *FaLOX2* may play a more upstream role, funneling substrates into VOC biosynthesis pathway that leads to γ -dodecalactone. This hypothesis was reinforced by the observation that *FaLOX2*-overexpressing fruits accumulated significant high levels γ -decalactone, γ -dodecalactone and methyl jasmonate. These results suggest that *FaLOX2* functions in the initial steps of the LOX pathway, which overlaps with the jasmonate biosynthesis pathway. Additional LOX-related volatiles, such as ethyl hexanoate, were also elevated in *FaLOX2*-overexpressing fruits. Several other VOCs were affected by *FaLOX2* overexpression, which may be an indirect result of increased methyl jasmonate levels. Exogenous application of methyl jasmonate has been shown to accelerate ripening in strawberries (Han *et al.*, 2019). This effect could be mediated through crosstalk between jasmonate and abscisic acid (ABA) signaling pathways, leading to altered VOCs profiles as a consequence of advance of fruits maturations.

In fact, some VOCs showed the same tendency during fruit maturation process and in response to *FaLOX2* overexpression. For example, mesifurane and eugenol, which are synthesized via pathways unrelated to LOX, increased in the later stages of fruit development and similarly increased in the *FaLOX2*-overexpressing fruits. A comparable pattern was observed for 2-hexen-1-ol acetate, benzyl acetate, methyl butyrate, nerolidol, and others. Conversely, VOCs that typically decline during ripening, such as 2-pentyl furan, heptanal, methyl-2-hexenoate or penten-3-on, were also found at lower concentration in the *FaLOX2*-overexpressing fruits compared to the control.

The analysis of primary metabolites further supports this model. During strawberry ripening, the most prominent metabolic changes include increased of glucose, fructose and sucrose (Fait *et al.*, 2008; Zhang *et al.*, 2011; Liu *et al.*, 2020; Enomoto, 2021). These three sugars were present at higher concentrations in the *FaLOX2*-overexpressing fruits, consistent with a significantly increased Brix value (**Annex 10**). These changes may be hormonally regulated by jasmonate and ABA during ripening (Sánchez-Gómez *et al.*, 2022). The observed increase in raffinose may be linked to its role in stress response and wound signaling (Couée *et al.*, 2006; Mi *et al.*, 2008; Sengupta *et al.*, 2015), aligning with jasmonate's known function in stress responses. Meanwhile, the decrease in rhamnose

could reflect reduced demand for pectin and cell wall glycoproteins biosynthesis as fruit tissues soften during ripening (Jiang *et al.*, 2021).

The accumulation of anthocyanins, a class of secondary metabolites, was also enhanced in *FaLOX2*-overexpressing fruits, likely in response to increased methyl jasmonate levels as previously described (Delgado *et al.*, 2018). This effect may be mediated through the modulation of the FaMYB10-FabHLH3 complex, FaMYC2 or other transcription factors involved in anthocyanin biosynthesis (Delgado *et al.*, 2018).

Expression profile of *FaLOX2* across different tissues and three stages of maturation revealed that this gene has the highest expression in leaf tissue. In receptacle, expression peaked in green and white stages and then decreased until almost no expression in red stage. One possible explanation for this behavior is that during the ripening process of strawberry, some signal triggers the expression of the *FaLOX2* and this enzyme through its action over α -linolenic acid contributing to the accumulation of methyl jasmonate. Once jasmonate levels reach a functional threshold, negative feedback mechanisms may suppress *FaLOX2* expression to maintain hormonal balance.

Further experiments are needed to fully elucidate the role of *FaLOX2* in strawberry VOC and hormone metabolism. Current efforts are focused on the analysis of stable transgenic lines generated in ‘Chandler’ genetic background. We have *in vitro* strawberry plants that putatively overexpress *FaLOX2* under the control of a constitutive promoter. Our future objectives will be to study level of overexpression by RT-qPCR, select plants with optimal levels of *FaLOX2* expression and then study the metabolic, proteomic and transcriptomic landscape of plants grown in greenhouse conditions. Additionally, to investigate the role of *FaLOX2* in jasmonate-mediated defense responses, we plan to perform pathogen colonization assays using *Botrytis cinerea* and other necrotrophic fungi on *FaLOX2*-overexpressing fruits.

4. *FaLOX2* is a cytoplasmic protein with involvement in jasmonate metabolism

Plant lipoxygenases were classified in two groups by phylogenetic analysis. This analysis places *FaLOX2* within 13S family of lipoxygenase, in accordance with our results from functional analysis. *FaLOX2* was located within 13S-F clade and, together with *FaLOX10*, were the only ones with a predicted subcellular localization different from chloroplast. Interestingly, *FaLOX2* was phylogenetically grouped close to *FvLOX3*, a protein that has been described as upregulated in response to methyl jasmonate (Li *et al.*, 2022). In this work the authors explore the regulation of different *F. vesca* lipoxygenases, and some interesting results can be applied to our candidate. *FvLOX3* expression pattern matches partially with data expression from *FaLOX2* because both are highly expressed in leaves, but *FvLOX3* also expresses in flowers. Also, *FvLOX3* was described to be

expressed under salt stress and salicylic acid treatment and downregulated under ABA treatment. Finally, they analyzed the promoter of several *F. vesca* LOX, and they found that *FvLOX3* promoter possesses elements involved in circadian rhythms as well as elements involved in mechanical injury response (WUN motif), methyl jasmonate response elements (CGTCA- and TGACG-motifs) and ABA response elements (ABRE).

Structurally speaking, FaLOX2 and FvLOX3 are very similar. Indeed, both structures overlap with each other almost identically and the residues for coordinating the Fe atom were the same and in the same disposition. Only a short distance difference can be seen regarding two α -helixes zones. What is remarkable is the presence of the jelly roll fold region that creates a hydrophobic cavity. Our hypothesis is that these antiparallel β -sheets structures are the ones involved in anchoring to biological membranes. The presence of this structural domain together with the presence of a signal peptide in the protein N-terminus would be responsible for the chloroplast subcellular localization of FvLOX3. On the other hand, TomLOXC is a well-studied LOX in *S. lycopersicum* that has been described to be relevant for the synthesis of C5 and C6 aldehydes and alcohols, compounds associated to flavor in tomato fruit (Chen *et al.*, 2004; Shen *et al.*, 2014). According to our work, FaLOX2 protein shared the same tridimensional structure than TomLOXC, despite the putative protein regions important for the chloroplastidic localization of TomLOXC (Chen *et al.*, 2004).

FaLOX2 was predicted to accumulate in the cytoplasm, so we performed an agroinfiltration of *N. benthamiana* with the fusion protein FaLOX2-GFP. The fusion in the C-terminal domain of the FaLOX2 protein was based on an InterPro analysis (Blum *et al.*, 2025) in where the lipoxygenase domain was more proximal to the N-terminal domain. Nevertheless, future experiments contemplate the possibility of generating the GFP-FaLOX2 construct and checking the subcellular localization again. With the current data from confocal microscopy, the FaLOX2 accumulates in the cytoplasm of *N. benthamiana* leaves, discarding the possibility of reticular localization due to the co-infiltration of FaFAH1. Furthermore, with the plasmolysis assay we also discarded the possibility of FaLOX2 to be associated to cell membranes.

5. Future perspectives and proposed experiments

With our current data, we are designing new experiments to characterize more deeply FaLOX2 protein. Firstly, to address the subcellular localization we would like to design chimeric proteins that carry the full coding sequence of FaLOX2 but with an N-terminal part from either FvLOX3 or TomLOXC. One chimeric construction would be the jelly roll fold (from FvLOX3 or TomLOXC) fused to the FaLOX2 CDS and another would be the signal peptide (again either from FvLOX3 or TomLOXC) fused to the FaLOX2 CDS. These two chimeric constructs would be used to check the subcellular localization and

test if the presence of either the jelly roll fold or the signal peptide is enough to change the subcellular localization from cytoplasm to chloroplast.

We currently have transgenic *S. lycopersicum* cv. ‘MoneyMaker’ lines. Initially we created the *35SCaMV::FaLOX2*-overexpressing lines and *TomLOXC*-RNAi silencing lines and we are focusing our efforts on obtaining homozygous lines. The volatile γ -dodecalactone has not been described to be present in *S. lycopersicum*, so one goal would be to check if the presence of *FaLOX2* is enough to produce γ -dodecalactone in a heterologous system. On the other hand, with *TomLOXC*-silenced lines we want to check if the transient expression of *FaLOX2* in tomato fruit will complement the phenotype of these lines. With this information we would also discover the possible actuation of *FaLOX2* protein in jasmonate signaling in a climacteric-model fruit by volatilome analysis, for example.

For the biochemical characterization of the *FaLOX2*, one possibility would be to purify fusion protein *FaLOX2*-GFP from agroinfiltrated *N. benthamiana* leaves and conduct enzymatic assays. In this case we could test several substrates, like α -linolenic acid, with different pH and temperature conditions and analyze the resulting products by HS-GC/MS.

Finally, regarding the involvement of *FaLOX2* in abiotic and biotic stresses we performed an assay of wounding based on the publication from Ibanez et al., 2019. We performed circular wounds of 5 mm diameter in strawberry ‘Camarosa’ and ‘Candongia’ plants, collect fruits tissue. With these fruits we will measure the expression of our candidate gene *FaLOX2* together with *FaJMT* as positive control by RT-qPCR. The jasmonate methyl transferase (*FaJMT*) gene is upregulated with the mechanical wounding (Ibanez et al., 2019) and we would expect similar results with the *FaLOX2* expression.

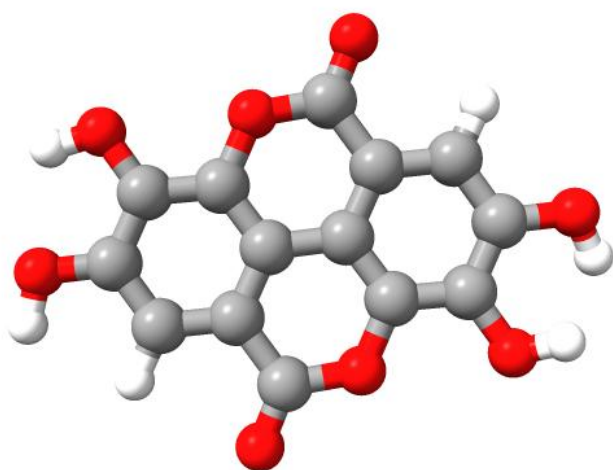
6. Concluding remarks

This chapter was about the volatiles genetic control in a mapping population derived from the crossing of ‘Camarosa’ and ‘Senga Sengana’. The previous work in our group revealed a high number of QTLs controlling volatile compounds in this mapping population. We selected a QTL that explained around 25% of the total variation of γ -dodecalactone that is important for sweetness perception. We selected a linoleate 13S-lipoxygenase as the candidate gene of this QTL and the transient overexpression of this gene resulted in an overaccumulation of γ -dodecalactone and methyl jasmonate, as well as sugars important for strawberry fruit. The effect of anthocyanins accumulation may occur through jasmonate signaling as a side effect of the overexpression of the candidate. The overaccumulation of some other VOCs would be an effect of an accelerated ripening process produced by the jasmonate-related signaling produced by the rise of methyl jasmonate. We also discovered that the protein codified by the candidate gene is located

in the cytoplasm in opposition to other lipoxygenases that are present in chloroplast. The cytoplasmic localization of our candidate can be explained by the absence of the signal peptide and the jelly roll fold for anchoring to biological membranes.

Finally, the proposed experiments would offer us more confident information about the characterization of the candidate protein. The establishment of the stable overexpression in 'Chandler' background would offer us new scenarios for exploring the functionality of this gene.

VI. Chapter II: Characterization of a candidate gene involved in ellagitannins content in strawberry fruit



Introduction

1. Tannins and their properties: antioxidants and relationship with human health and organoleptic characteristics

The antioxidant properties of tannins have been thoroughly examined from natural sources or from food grade preparations. For example, Shin *et al.*, (2014) studied the antioxidant activities of both hydrolysable and condensed tannins in three cultivars of persimmon. The radical scavenging activity (RSA) is often used to determine the antioxidant capacities and particularly, the 2,2'-azinobis-3-ethylbenzthiazolin-6-sulfonic acid (ABTS) RSA assay allows to measure the antioxidant activities of both hydrophilic and hydrophobic antioxidant chemicals (Floegel *et al.*, 2011). Interestingly, 1,000 µg/ml of soluble tannins from three persimmon cultivars showed higher ABTS RSA than 100 µg/ml of L-ascorbic acid (Shin *et al.*, 2014). The 2,2-diphenyl-1-picrylhydrazyl (DPPH) RSA assay (Blois, 1958) was also conducted in the same study to provide evidence that the DPPH RSA of unripe persimmons was significantly higher than those of ripe persimmons, mirroring the trend of tannin accumulation, which decreases during fruit ripening. Additionally, Ricci *et al.*, (2016) studied the antioxidant activity of commercial food grade tannins (extracted from grape and *Quercus* tissues) in wine samples. Samples with higher tannin content (measured as (+)-catechin equivalents) showed, in most cases, the highest values of DPPH radical scavenging activity. The reducing power of the samples, estimated with the ferric-reducing antioxidant power (FRAP) assay (Benzie and Strain, 1999), ranged between 0.513 and 0.662 mM of FeSO₄·7H₂O, providing effective reducing properties in some cases. Finally, and at least for wine, it seems that based on ABTS, FRAP, and DPPH assays, hydrolysable tannins show higher antioxidant activity than condensed tannins (Pascual *et al.*, 2017; Vignault *et al.*, 2018).

The use of tannins in food and beverages provides certain advantages, like protection from oxidation processes. In the industry, tannins are used as flavoring agents, and, for example, commercial tannins are added for the clarification of grape musts and wines (Ricci *et al.*, 2016). This helps improve color stability and avoid oxidation (Versari *et al.*, 2013), offering a natural alternative to synthetic antioxidants. The oxidation of certain compounds plays a crucial role in the flavor of some beverages like beer, especially during aging processes (Vanderhaegen *et al.*, 2006). To further confirm the effect of tannins on beverages, De Francesco *et al.*, (2020) added various tannin-rich extracts to beer. This addition led to an improvement in turbidity, color formation, foam quality, citrus and spicy notes, and an increase in the body of the beer.

As previously mentioned, the antioxidant capacities of tannins are of high importance for human health, providing neuroprotective, cardioprotective, and even antitumoral activities. In a recent review, Maugeri *et al.*, (2022) summarized the *in vitro* and *in vivo*

effects of tannins in several experimental models and provided evidence of anti-bacterial and anti-inflammatory effects. The anti-inflammatory effects of tannins are related to palliating oxidative stress, reducing ROS and NOS, avoiding DNA damage, and taking part in some signaling routes (Maugeri *et al.*, 2022). In addition, the health benefits of tannins have been thoroughly reviewed elsewhere (Basu *et al.*, 2014; Cory *et al.*, 2018; Lu *et al.*, 2021b).

Tannins are present in foods and beverages, providing them with the perception of astringency, which is felt as a sensation of dryness in the mouth and tongue. This event is thought to be due to the precipitation of oral proteins and mucopolysaccharides when they interact with tannins (Baxter *et al.*, 1997). The mechanism of astringency perception has been widely discussed, and some hypotheses have been proposed: 1) interaction between tannins and epithelial receptors (Ma *et al.*, 2014), 2) physical movement of mouth musculature (Pires *et al.*, 2020) or 3) increasing of the oral friction caused by protein-tannins aggregates (Rossetti *et al.*, 2009). As tannins are widely distributed among plants, astringency is also present in many food-sources, such as rice (Shao *et al.*, 2018), tea (Xu *et al.*, 2018), grape (Sun *et al.*, 2020), cocoa (Misnawi *et al.*, 2005), and also in some cereals like sorghum and millet (Ahmad *et al.*, 2018). Nevertheless, not all plant sources have the same quantity of tannins. In the case of proanthocyanidins, the amount of (+)-catechin in some sorghum and millet varieties ranges from 172 to 179 mg/100 g for sorghum and from 172 to 174 mg/100 g (Ahmad *et al.*, 2018), respectively. On the other hand, 11–54 mg/g of (+)-catechin were found in eight different cultivars of persimmon (Akagi *et al.*, 2010b). Concerning ellagitannins, the content in strawberry fruits was around 2–18 mg/g (Karlińska *et al.*, 2021).

Interestingly, astringency is often accompanied by bitterness, which is a source of confusion between both taste properties (Arnold *et al.*, 1980; Green, 1993; Carpenter *et al.*, 2019). However, the perception of astringency seems to be a secondary attribute when both astringency and bitterness are simultaneously perceived in the same product (Arnold *et al.*, 1980; Green, 1993). Interestingly, the astringency/bitterness properties of tannins depend on their degree of polymerization. Molecules with low molecular weight tend to be bitterer, whereas tannins with a high degree of polymerization appear to be more astringent (Robichaud and Noble, 1990; Peleg *et al.*, 1999; Hufnagel and Hofmann, 2008). There is a real challenge in enhancing the tannin content of foods and beverages because of their modulation in organoleptic characteristics, which can lead to non-pleasing foodstuffs. It is needed to better understand the food matrix effects on the perception of astringency and bitterness and to discover new ways to modulate tannins' contributions to taste perception while maintaining or even increasing these health-promoting compounds.

Some research has already been carried out to solve this particular conflict. For example, the addition of some proteins leads to the removal of highly reactive tannins and the modification of astringency and bitterness (Siebert *et al.*, 1996; Siebert, 1999). Proteins like β -casein, β -lactoglobulin, Ca-caseinate, Na-caseinate, and different types of gelatins have been assayed for their potential bitter-taste-masking (Bohin *et al.*, 2012). For example, the addition of gelatin to red wines causes a decrease in astringency due to the specific removal of high molecular weight galloylated proanthocyanidins (Sarni-Manchado *et al.*, 1999). The contribution of polysaccharides to astringency is also important. Soluble polysaccharides, derived, for example, from pectins during the ripening process, can compete with salivary proteins for polyphenolic substrates like tannins, which lead to a modification in astringency perception (Ozawa *et al.*, 1987).

2. Proanthocyanidins or condensed tannins

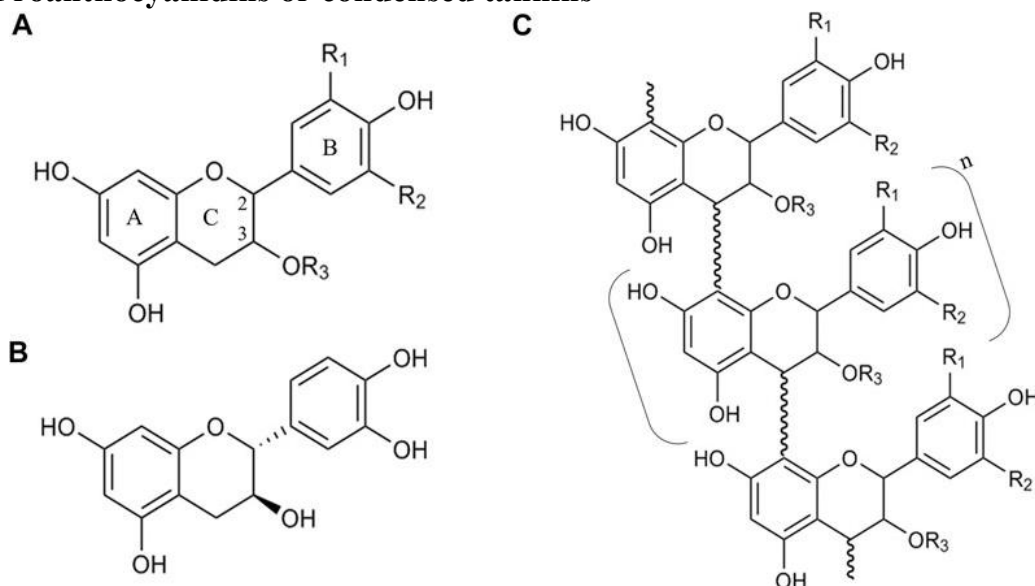


Figure 41. Schematic of the structure of catechins and proanthocyanidins. (A) General structure of a catechin, where (A–C) indicate the distinct rings of the carbon backbone C6–C3–C6, numbers 2 and 3 indicate the chiral carbons (stereochemistry not shown), and R_1 , R_2 and R_3 are the possible substituent groups. (B) An example of a catechin (+)-catechin. (C) General structure of a proanthocyanidin, where n indicates the number of repetitions of the monomer in brackets. The bond between the monomers is called an interflavan linkage. Schema taken from Mora *et al.*, (2022).

Condensed tannins, or proanthocyanidins, are a highly diverse class of polymers and/or oligomers of a flavonoid called flavan-3-ol. Structurally speaking, flavanols are composed of a carbon backbone of C6–C3–C6, where the C6 compounds correspond to aromatic rings (A- and B-ring) and the C3 is a dihydropyran ring (C-ring) (Aron and Kennedy, 2008). The A-ring is similar to a resorcinol moiety and the B-ring is similar to a catechol moiety (**Fig. 41A**). The monomers of flavan-3-ols are called catechins, and

they differ among them in the stereochemistry of the C2 and C3, the presence or absence of galloyl groups, and the hydroxylation of the B-ring (Watrelet and Norton, 2020). When the substituents are $R_1 = R_3 = H$ and $R_2 = OH$, then the molecule is (+)-catechin (Figure 2B). As flavan-3-ols have two chiral carbons, they possess four diastereoisomers, so one isomer of (+)-catechin (2*R*, 3*S* configuration) is (+)-epicatechin (2*S*, 3*S* configuration). In nature, the configuration 2*R* is more common than the 2*S* configuration (Aron and Kennedy, 2008).

Aron and Kennedy, (2008) described thirteen types of proanthocyanidins based upon the hydroxylation of A-ring and B-ring carbons and the C3 from the C-ring. Hydroxylation of C8 (A-ring) is the most uncommon modification of proanthocyanidins, present only in the cases of proteracacinidin and promelacacinidin. By contrast, in the case of C3 (C-ring), only in 5 types a proton is present instead of a hydroxyl group (Aron and Kennedy, 2008). In this classification, it is not contemplated that C3 substituents despite hydroxyl groups.

2.1. Proanthocyanidins biosynthesis, transport and accumulation

Proanthocyanidins are synthesized from L-phenylalanine through the phenylpropanoid and flavonoid pathways. First, L-phenylalanine is transformed into cinnamic acid or cinnamate by the phenylalanine ammonia lyase (PAL), and the cinnamate is hydroxylated to p-coumaric acid by means of cinnamate 4-hydroxylase (C4H). The next step is the addition of coenzyme A (CoA) to the p-coumaric acid by the catalysis of 4-coumaroyl-CoA ligase (4CL), yielding p-coumaroyl-CoA (Vogt, 2010; Deng and Lu, 2017). 4-coumaroyl-CoA and 3 molecules of malonyl-CoA are the substrates of chalcone synthase (CHS), which catalyzes the first step of the flavonoid pathway, yielding naringenin chalcone, and chalcone isomerase (CHI) then converts this molecule to the flavanone naringenin. Naringenin serves as a substrate for flavonoid 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H) and flavonoid 3'5'-hydroxylase (F3'5'H). If naringenin is the substrate of F3'5'H and F3H, the product is a dihydroflavonol called dihydromyricetin, while the combined action of F3'H and F3H yields the production of dihydroquercetin. Dihydroflavonol 4-reductase (DFR) converts dihydromyricetin to leucodelphinidin ($R_1 = R_2 = OH$ in **Fig. 41A**) and dihydroquercetin to leucocyanidin ($R_1 = OH$ and $R_2 = H$ in **Fig. 41A**). Then, the anthocyanidin synthase (ANS), also called leucoanthocyanidin dioxygenase (LDOX), produced delphinidin and cyanidin from leucodelphinidin and leucocyanidin, respectively. Additionally, dihydrokaempferol is derived from naringenin and can be converted into leucopelargonidin ($R_1 = R_2 = H$ in **Fig. 41A**) and pelargonidin by DFR and ANS catalysis, respectively. Cyanidin, delphinidin, and pelargonidin can be transformed into 2,3-*cis*-flavan-3-ols epicatechin, epigallocatechin, and epiafzelechin, respectively, by the action of anthocyanidin reductase (ANR). On the other hand, 2,3-*trans*-flavan-3-ols are synthesized from

leucocyanidin, leucodelphinidin, and leucopelargonidin by the action of leucoanthocyanidin reductase (LAR) (**Fig. 42**) (Akagi *et al.*, 2011; Yu *et al.*, 2019; Wei *et al.*, 2021).

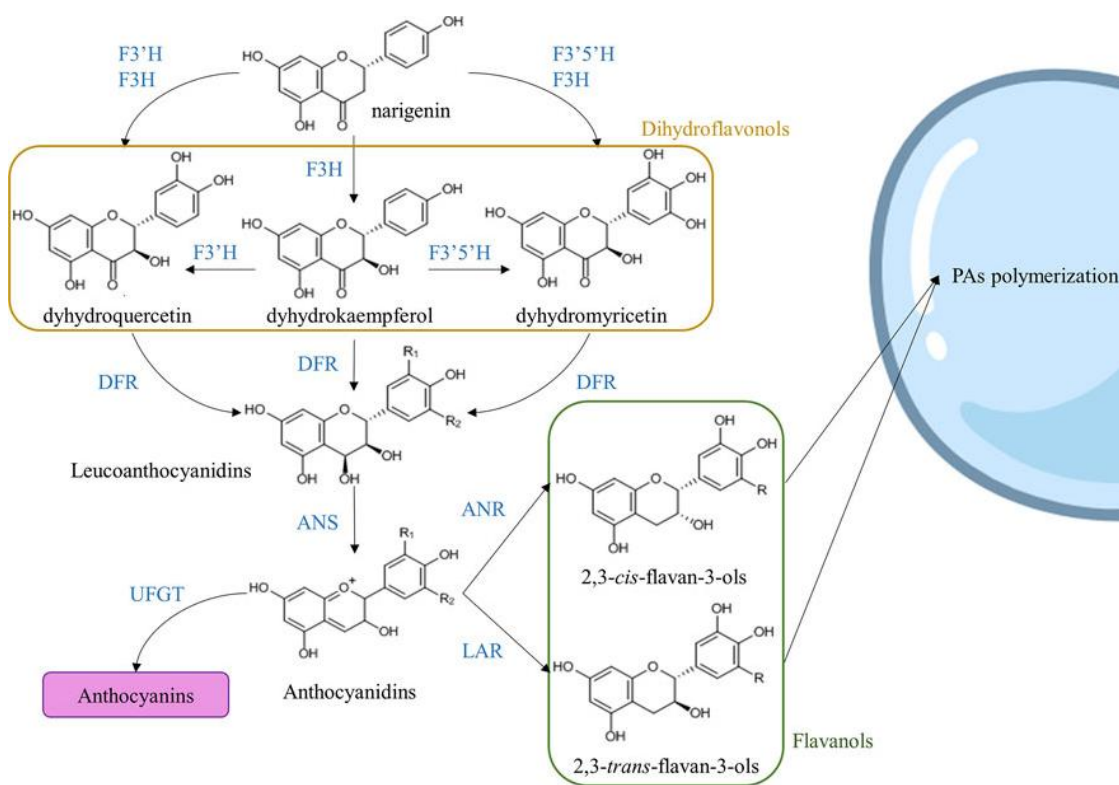


Figure 42. Simplified flavonoid pathway. One branch of this pathway yields flavanols that can be transported to the vacuole by several proposed mechanisms. Abbreviations: flavonoid 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), flavonoid 3'5'-hydroxylase (F3'5'H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), UDP-glucose flavonoid 3-O-glucosyltransferase (UFGT), anthocyanidin reductase (ANR), and leucoanthocyanidin reductase (LAR). Taken from Mora *et al.*, 2022.

PAs are found to accumulate exclusively in the vacuoles, forming glycoside conjugates (Grotewold, 2004; Kitamura *et al.*, 2010; Fontes *et al.*, 2011). In particular, some vacuoles specialized in PA accumulation are known as “tannin vacuoles” (Terrier *et al.*, 2009). The knowledge about biophysical and biochemical characteristics of vacuoles' membranes is very elusive, converting tannin transport into these organelles in a poorly known mechanism. One mechanism proposed for the transport of PAs from the cytosol to the vacuole is membrane transporters. There are three kinds of putative membrane transporters for PAs: ATP binding cassette (ABC) transporters, mammalian bilitranslocase (BLT) transporters, and multidrug detoxification and extrusion (MATE) transporters (Rousserie *et al.*, 2019). Only ABC transporters act as primary active

transporters, using ATP for transport, while BLT and MATE transporters function as secondary active transporters, using an electrochemical gradient to move PAs.

The name “condensed tannins” for PAs is because the acid hydrolysis of these compounds releases anthocyanidins due to the breakage of interflavan linkages (Deng and Lu, 2017). As mentioned before, PAs are built of flavan-3-ol monomers linked by carbon-carbon bonds (Dixon *et al.*, 2013). Actual knowledge indicates that (+)-catechins and (-)-epicatechins are thought to be starter units for PA polymerization (Xie *et al.*, 2003; Kitamura *et al.*, 2004; Dixon *et al.*, 2005). Due to the interflavan linkages, two types of PAs can be distinguished. The A-type consists of subunits linked by C4-C8 and/or C4-C6 (A-ring carbons) and C2-O-C7 or C2-O-C5 bonds. On the other hand, B-type PAs presents only the two possible linkages C4-C8 and/or C4-C6 (Yu *et al.*, 2020). The distribution of A-type and B-type PAs is different in nature. A-type PAs are mainly found in berries like bilberry (*Vaccinium myrtillus*) and cranberry (*Vaccinium oxycoccus*), cinnamon (*Cinnamomum spp.*) and peanuts (*Arachis hypogaea*), for example (Gabetta *et al.*, 2000; Gu *et al.*, 2002, 2003), while B-types are widely distributed (Gabetta *et al.*, 2000; Gu *et al.*, 2003; Suvanto *et al.*, 2020). Condensed tannins are responsible for the bitterness and astringency of many fruits. For example, in apples, astringency is correlated to the mean degree of polymerization (mDP) and the galloylation degree of these compounds, but it is negatively correlated with the hydroxylation level of the B-ring of flavan-3-ols (Saxe *et al.*, 2021).

2.2. Proanthocyanidin genetic regulation

Regulation of proanthocyanidin biosynthesis occurs mainly at the transcriptional level, especially for the genes encoding biosynthetic enzymes (Koes *et al.*, 2005; Xu *et al.*, 2015). PAs synthesis is regulated during fruit growth, as their accumulation starts with the development of the fruit and stops when it starts ripening (Wei *et al.*, 2021). The regulation of PAs is well-known in *A. thaliana* where there are 6 families of transcription factors thought to participate in PAs synthesis regulation: MYB, basic helix-loop-helix (bHLH, also known as MYC), tryptophan-aspartic acid repeat (WDR)-containing protein 40 (WD40), WRKY, and MADS-box proteins and zinc fingers (Terrier *et al.*, 2009; Rousserie *et al.*, 2019). The MYB-bHLH-WD40 (MBW) complexes are essential regulator elements for PAs accumulation and gene expression (**Fig. 43**) (Koes *et al.*, 2005; Ramsay and Glover, 2005; Wei *et al.*, 2021). Due to the ubiquity of bHLH and WD40 in transcriptional regulation, it is the C-terminal part of the MYB protein in the MBW-complex that is responsible for the specificity in activating target genes. Meanwhile, bHLH is responsible for binding to the nucleotide sequence (normally E-box or G-box *cis*-elements) (Xu *et al.*, 2014), and WD40 is crucial for the stabilization of the complex (Huang *et al.*, 2014; Yu *et al.*, 2020a). Both transcription factors, MYB and bHLH, can specifically interact with nucleotide sequences within the flavonoid promoter.

MYBCORE and AC-rich elements are the two main binding sites for MYBs in the regulation of flavonoids and PAs, although the exact motif for MYB binding remains unknown in most PAs structural genes (Constabel, 2018). In *A. thaliana*, the MBW-complex is composed of *AtMYB123* (or Transparent Testa, TT2), *AtbHLH42* (or TT8), and Transparent Testa Glabra31 (or TTG3), which is the WD40 protein. This complex is responsible for the activation of *ANR* (Rousserie *et al.*, 2019), *LDOX* and *DFR* genes, leading to the accumulation of PAs in the seed coat (Nesi *et al.*, 2001; Baudry *et al.*, 2004). In particular, TTG2-like proteins are thought to be crucial in PA transportation (Debeaujon *et al.*, 2001; Gonzalez *et al.*, 2016). Homologs of *TT2*, *TT8*, and *TTG1* in other fruits and plants present a role in flavonoid and PA regulation (Yu *et al.*, 2020a).

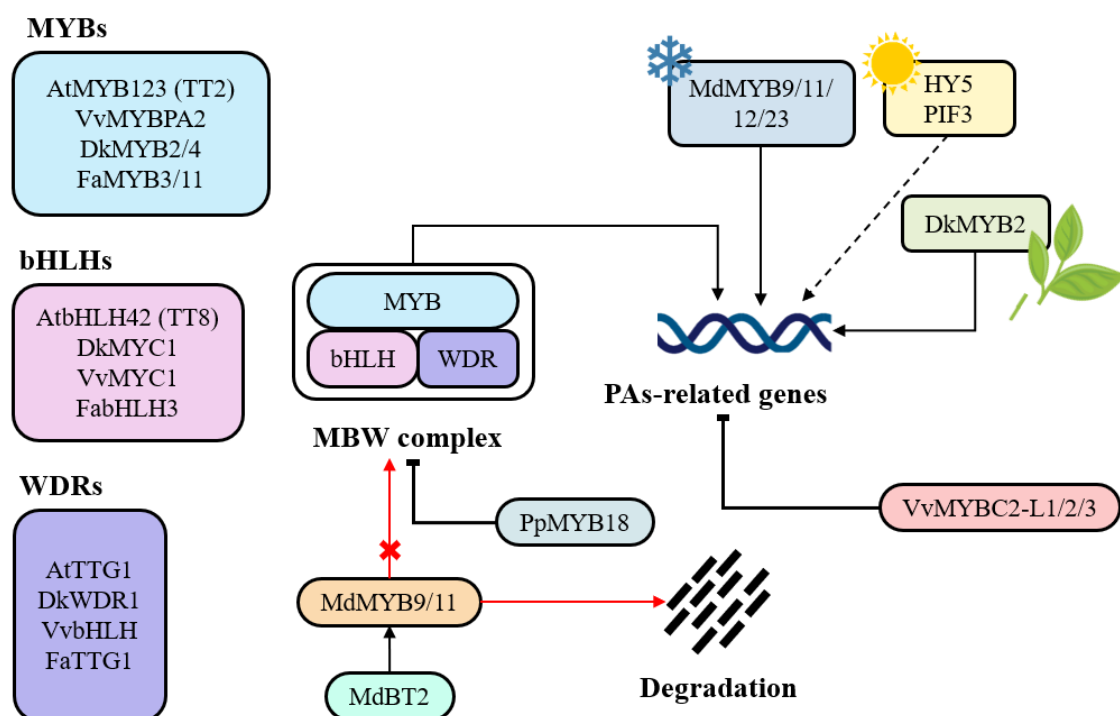


Figure 43. General regulation scheme of proanthocyanidin synthesis. The scheme shows regulation via MBW-complex formation and cold-, light- and wounding-induced regulation. The degradation of MdMYB9/11 through BT2 and the proteasome is also shown with the red arrows, resulting in the impossibility of these two transcription factors to form the MBW-complex. Arrows indicate positive regulation, while flat-ended lines mean negative regulation (repression). Dotted lines indicate a process not fully elucidated. At: *Arabidopsis thaliana*, Dk: *Diospyros kaki*, Fa: *Fragaria F. × ananassa*, Md: *Malus domestica*, Pp: *Prunus persica*, Vv: *Vitis vinifera*. Schema modified from Mora *et al.*, (2022).

Besides regulation at the gene level, the biosynthesis of PAs is affected by biotic and abiotic stresses and plant hormones. Environmental or abiotic stresses, such as low temperature, drought, or wounding, among others, may lead to the production of ROS and the consequent oxidative stress responses (Suzuki *et al.*, 2012). As described

previously, flavonoids and other antioxidant compounds can palliate the action of ROS, and the genes responsible for their synthesis can be induced under these circumstances (Winkel-Shirley, 2002). Water quantity has been proven to influence the phenolic content of grapes (Castellarin *et al.*, 2007). Light also plays a part in the regulation of PAs, and it has been suggested that light-signaling proteins, such as HY5 and PIF3, may interact with some promoters of PAs-related genes, like *CHS*, *CHI*, and *DFR* (Shin *et al.*, 2007). The action of plant hormones courses through their interaction with MBW-complexes, thus regulating flavonoid genes at the transcriptional or post-transcriptional level (Jaakola, 2013). In addition to abiotic stresses, plants are also exposed to the influence of fungi, bacteria, viruses, and herbivores. In this context, accumulation of PAs is a plant defense strategy against these biotic stresses (Dixon *et al.*, 2005; Ullah *et al.*, 2017). The two main hormones related to biotic stresses are salicylic and jasmonic acids, suggesting a possible role for both molecules in PA regulation.

2.3. Proanthocyanidin regulation in strawberries

Other crops have been the object of study in what pertains to PAs regulation. In reference to strawberries (*F. × ananassa*), the major and common compounds related to PAs are catechin, proanthocyanidin B1, proanthocyanidin trimer, and proanthocyanidin B3 (Giampieri *et al.*, 2012; Fierascu *et al.*, 2020). In strawberry, epicatechins and catechins levels, as well as procyanidin levels, decrease with ripening (Aaby *et al.*, 2012), specifically in the final maturation stage, this tendency has been observed in other species like bilberry (Suvanto *et al.*, 2020). The functional orthologues of *AtTT2*, *AtTT8*, and *AtTTG1* have been elucidated in strawberry, being *FaMYB9/FaMYB11*, *FabHLH3*, and *FaTTG1*, respectively (Schaart *et al.*, 2013). In a yeast-two hybrid assay, *FaMYC1* and *FabHLH33* were also discovered to be related to PA synthesis, but their implication remains elusive. It is speculated that *FaMYC1* and *FabHLH33* could dimerize in the same way as other bHLHs (Feller *et al.*, 2006). Derived from the same work, Schaart *et al.*, (2013) hypothesized that *FaMYB5* and *FabHLH3Δ*, a truncated version of *FabHLH3* identified in strawberries, act as negative regulators of flavonoid biosynthesis. *FaMYB5* may combine with *FaMYB1*, a confirmed repressor of PAs biosynthesis with high sequence homology to *VvMYBC2-L1* (Aharoni *et al.*, 2001; Paolucci *et al.*, 2011), and they together interfere with the activity of MBW-complexes. On its part, *FabHLH3Δ* is capable of interacting with *FaMYB9* or *FaMYB11* and causing the blockage of their binding sites for bHLHs, avoiding the MBW-complex formation (Schaart *et al.*, 2013). Interestingly, it was found that *FaTT12-1* is a putative orthologue of *AtTT12*, which is light-sensitive. In fact, red light can significantly activate its expression (Chen *et al.*, 2018), suggesting a possible role in PA biosynthesis regulation. Delgado *et al.*, (2018) demonstrated that treatment with methyl jasmonate caused upregulation of *FaANS*, *FaLAR*, and *FaUFGT* genes and a downregulation of the *FaANR* gene compared to the

control, leading to increased anthocyanin levels and decreased PAs content. On the other hand, treatment with methyl jasmonate and jarin-1, an inhibitor of the bioactive form of jasmonate, jasmonoyl-isoleucine, caused the opposite effect. The increase of PAs in fruits treated with methyl jasmonate and jarin-1 could be explained by the upregulation of *FaMYB9*, *FaMYB11*, and *FabHLH33* together with *FaANR* (Delgado *et al.*, 2018). Together, these data suggest a connection between the phytohormone jasmonic acid and PA biosynthesis during strawberry fruit ripening. Inoculation experiments with *Colletotrichum acutatum* and *B. cinerea* in strawberries at different maturity stages showed that fungal infection was much higher in the case of riper fruits, but almost absent in the case of unripe strawberries, where the expression levels of *LAR* and *ANR* were found to be higher (Guidarelli *et al.*, 2011; Haile *et al.*, 2019). This matches with the fact that flavan-3-ols and PA concentrations decrease as ripening progresses (Carbone *et al.*, 2009; Nagpala *et al.*, 2016), reinforcing the role of PAs in biotic stresses.

3. Hydrolysable tannins: ellagitannins and gallotannins

The presence of hydrolysable tannins in food and beverages is more uncommon than the presence of condensed tannins. They can be found in black currant (*Ribes nigrum*), blackberries (*Rubus fruticosus*), strawberries, raspberries (*Rubus occidentalis*), and pomegranate (*Punica granatum*), guava (*Pisidium spp.*), mango (*Mangifera indica*) and in nuts like almonds (*Prunus dulcis*), pecans (*Carya illinoensis*), and walnuts (*Juglans regia*), with the total absence in species like legumes and cereal grains (Landete, 2011; Smeriglio *et al.*, 2017).

3.1. Biosynthesis of hydrolysable tannins

On the opposite of proanthocyanidins, much less is known regarding hydrolysable tannins synthesis and regulation. Structurally, hydrolysable tannins are based on a central sugar core, mainly β -D-glucose, in which phenolic compounds are esterified through its hydroxyl groups (Grundhöfer *et al.*, 2001; Quideau *et al.*, 2011). There are two major classes of these compounds depending on the phenolic groups: gallotannins (gallic acid) and ellagitannins (hexahydroxydiphenic acid, HHDP) (Bar-Ya'akov *et al.*, 2019). Acid hydrolysis of gallotannins yields sugar and gallic acid, whereas the hydrolysis of ellagitannins provides sugar, gallic acid, and ellagic acid (**Fig. 44**) (Smeriglio *et al.*, 2017). Gallotannins are the simplest hydrolysable tannins and are rarely found in nature. They are present in some woody species like mango and almonds; on the other hand, ellagitannins are more common in fruits, nuts and seeds (Landete, 2011; Smeriglio *et al.*, 2017). In fact, a myriad of ellagitannin compounds have been identified to date, principally due to their ability to oligomerize (Okuda and Ito, 2011).

The key molecule for hydrolysable tannin biosynthesis is gallic acid, which is derived from 3-dehydroshikimate via a shikimate dehydrogenase (Niemetz and Gross, 2005; Muir

et al., 2011; Bontpart *et al.*, 2016). Furthermore, gallic acid is relevant for the formation of epigallocatechin and epicatechin gallate (Liu *et al.*, 2012), which are proanthocyanidin monomers. Downregulation of the shikimate dehydrogenase, *DkSDH*, in persimmon resulted in a decrease in epigallocatechin content in PCNA-type mutants (Akagi *et al.*, 2009), which matches the implication of gallic acid in both types of tannins. Studies in *Camellia sinensis* and *V. vinifera* indicate the role of shikimate dehydrogenase in generating gallic acid (Bontpart *et al.*, 2016; Huang *et al.*, 2019). Gallic acid is then esterified with UDP-glucose (Gross, 1983), to form 1-*O*-galloyl- β -*D*-glucopyranose (also known as β -glucogallin), which is the simplest gallotannin. Ono *et al.*, (2016) found two UDP-glycosyltransferases (UGTs) in *P. granatum*, *PgUGT84A23* and *PgUGHT84A24*, related to β -glucogallin synthesis. According to studies in oak (*Quercus rubur* and *Quercus rubra*) and sumac (*Rhus typhina*), the next steps of galloylation occur with a remarkable specificity following the sequence: β -glucogallin, 1,6-digalloylglucose, 1,2,3-trigalloylglucose, 1,2,3,6-tetragalloylglucose and finally, 1,2,3,4,6-pentagalloylglucose (Schmid *et al.*, 1987; Gross and Denzel, 1991). These subsequent additions of galloyl moieties (Haslam and Cai, 1994) result in the formation of 1,2,3,4,6-pentagalloyl-*D*-glucose, the common precursor of both gallotannins and ellagitannins (**Fig. 44**). Pentagalloylglucose can be oxidized to generate 3,4,5,3',4',5'-hexahydroxyphenoyl (HHDP) residues, a molecule that once liberated from the sugar core in the form of hexahydroxydiphenic acid, spontaneously lactonizes into ellagic acid (Niemetz and Gross, 2005; Smeriglio *et al.*, 2017). The oxidation of pentagalloylglucose occurs via oxygen-dependent laccase-type enzyme catalyzation in *Tellima grandiflora* (Grundhöfer *et al.*, 2001; Niemetz and Gross, 2003a,b; Niemetz *et al.*, 2003).

Besides β -*D*-glucose, which is by far the most common sugar core for gallotannins and ellagitannins, other exotic polyols may appear, such as xylose, sorbitol, fructose, or even not-sugar compounds like shikimic, and quinic acids. However, their abundance in nature is extraordinarily low and they are only found in maple (*Acer negundo*), chestnut (*Castanea sativa*) and witch-hazel (*Hamamelis sp.*) (Smeriglio *et al.*, 2017). Galloylation reactions with β -glucogallin as substrate can yield a wide range of molecules from di- to octagalloylglucoses, although the most common is 1,2,3,4,6-pentagalloyl-*D*-glucose (Smeriglio *et al.*, 2017). The enzyme β -glucogallin-dependent galloyltransferases (Niemetz and Gross, 2005) is thought to be involved in the modification of gallotannins, one example is the transformation of hexa-/heptagalloylglucoses into 3-*O*-trigalloyl-1,2,4,6-*O*-tetragalloyl- β -*D*-glucopyranose, and other gallotannins with high degree of galloylation (Niemetz and Gross, 2005). Two galloyl residues can form what is called a meta- or para-depside bond, which is the result of the esterification between the carbonyl group of gallic acid and the aromatic meta- or para-hydroxyl group and comprises one of the most common modifications of gallotannins (Smeriglio *et al.*, 2017).

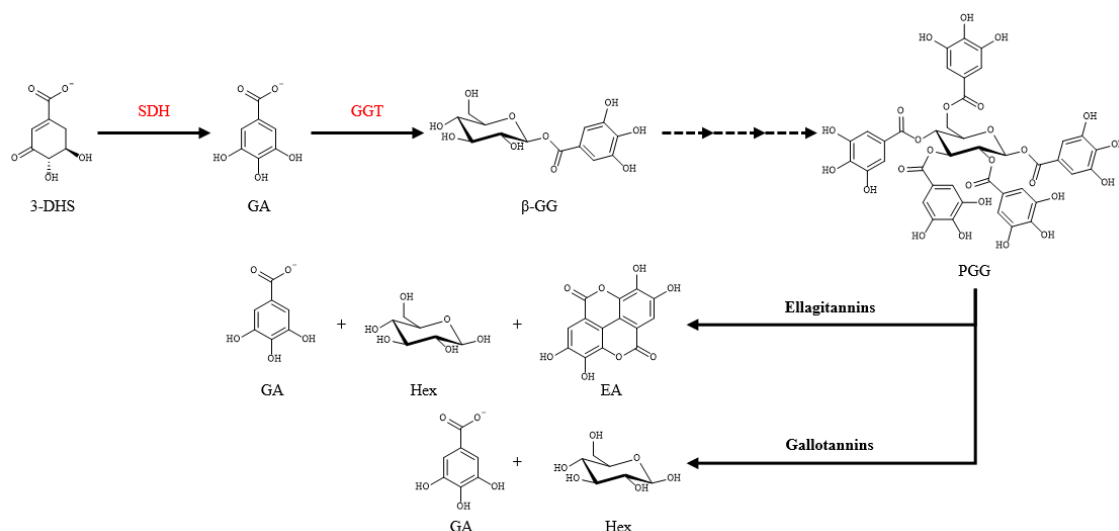


Figure 44. Biosynthetic pathway of hydrolysable tannins. The hydrolysis of gallotannins and ellagitannins yields gallic acid and sugar (glucose represented) and gallic acid, ellagic acid, and sugar, respectively. Dashed lines represent multistep reactions. β-GG: β-glucogallin; 3-DHS: 3-dehydroshikimate; EA: ellagic acid; GA: gallic acid; GGT: β-glucogallin-dependent galloyltransferases; Hex: hexose; PPG: pentagalloylglucose; SDH: shikimate dehydrogenase. Modified from Mora *et al.*, (2022).

3.2. Occurrence and regulation of hydrolysable tannins

Ellagitannins can be present in monomeric forms (e.g. casuarticin, corilagin, castalagin, eugenin, geraniin, corilagin, galloyl-HDDP-glucose, and potentillin), dimeric forms (e.g. sanguin H-2, sanguin H-6, agrimoniin, and lambertianin A), trimeric forms (e.g. nupharin C, nupharin E) or even oligomeric forms (e.g. hirtellin A). Ellagic acid is one of the most important hydrolysable tannins in *F. × ananassa* fruits (Ariza *et al.*, 2018). It has been recently found that agrimoniin, which is a dimeric ellagitannin composed of two units of potentillin linked by a dehydrogalloyl group (Karlińska *et al.*, 2021), is highly present in ‘Camino Real’, ‘San Andreas’ and ‘Festival’ strawberry cultivars (Villamil-Galindo *et al.*, 2021). In fact, agrimoniin has been proposed as a taxonomic marker for the *Rosaceae* family, to which strawberries belong (Grochowski *et al.*, 2017). Other common ellagitannins in the *Fragaria* genus are sanguin H-6, galloyl-bis-HHDP-glucose, and lambertianin C (Giampieri *et al.*, 2012).

Regarding ellagitannin biosynthesis, formation of HHDP through the C4-C6 bonding of two respective gallic acids leads to the synthesis of tellimagrandin II as the first product of the pathway in a reaction catalyzed by a laccase-type polyphenol oxidase, as described in strawberry (Karlińska *et al.*, 2021). Esterification of two more gallic acids at C2 and C3 and a possible de-esterification in C1 of tellimagrandin II result in several monomers: pedunculagin, potentillin and casuarticin (Niemetz and Gross, 2005; Karlińska *et al.*, 2021). Further steps implicate dimerization or condensation reactions, forming 3 types of

C-O-C in the case of the *Roseaceae* family: GOD, DOG, and GOG, depending on the donor and the acceptor of the reaction. The GOG-type bond implies two galloyl groups, the DOG-type bond (most frequently found in oligomers) requires a hydroxyl group from an HHDP as donor and a galloyl group as acceptor, and finally, the GOD-type bond is created from a galloyl hydroxyl donor to form an ether linkage with a HHDP as acceptor (Quideau, 2009).

Pomegranate fruit is another example of a reservoir of hydrolysable tannins, with more than 30 chemical species identified in the fruit peel, juice, and seeds (Fischer *et al.*, 2011; Mena *et al.*, 2012; Ito *et al.*, 2014; Ambigaipalan *et al.*, 2017; Wu and Tian, 2017; Liu and Seeram, 2018). The most representative ellagitannins in pomegranate fruit peel are α - and β -punicalagin (Seeram *et al.*, 2005), and gallic acid, punicalagin, and ellagic acid glycosides count as hydrolysable tannins present in both juice and fruit peel. Only 3,3'-di-*O*-methylellagic acid and 3,3',4'-tri-*O*-methylellagic acid are present exclusively in the seeds (Bar-Ya'akov *et al.*, 2019). Once again, the accumulation of tannins can be modified by regulating the expression of genes encoding key enzymes in their biosynthesis. Under osmotic stress, two shikimate dehydrogenase transcripts, *PgSDH3s* and *PgSDH4*, were accumulated, which is thought to increase hydrolysable tannin concentration (Habashi *et al.*, 2019). Moreover, Habashi *et al.*, (2019) reported that sucrose and red light stress could affect the accumulation of gallic acid and hydrolysable tannins with the change of SDH expression in pomegranates.

Tannases are a family of enzymes also known as tannin acyl-hydrolases (EC 3.1.1.20) capable of breaking carboxylic ester bonds present in gallotannins and galloylated flavanols (Ramírez *et al.*, 2008; Rodríguez *et al.*, 2009). These enzymes have been extensively studied in bacteria and fungi (Banerjee *et al.*, 2012), but there is still much to learn in plants. The work of Dai *et al.*, (2020) revealed that transient overexpression and RNAi of a strawberry tannase (*FaTA*) provoke an alteration in ellagic acid content; in addition, several enzymes belonging to the class I carboxylesterase clade were identified in *V. vinifera*, *J. regia*, *C. clementine*, *D. kaki*, and *C. sinensis*, indicating the hypothetical existence of more tannase genes in plants. A recent work in Juglandaceae (Wang *et al.*, 2021b) reveals regulatory cis-elements found in tannase promoters: E-box and ARR1AT, involved in brassinolide and cytokine responsiveness, respectively, and W-box and WUN-motifs implicated in wound abiotic stress responses. Also, motifs related to flavonoid biosynthesis (MYB-related motifs) and MYC motifs, related to cell growth, were found in tannases from Chinese hickory (*Carya cathayensis*) and pecan (*Carya illinoensis*). These last results suggest the possibility of studying the regulation of tannases belonging to the same clad described by Dai *et al.*, (2020) in other species.

4. UDP-glycosyltransferases enzymes

Glycosylation is a crucial metabolic process that supports the detoxification, storage, and excretion of xenobiotics and endogenous substrates. These reactions are driven primarily by glycosyltransferases (GTs), that are enzymes capable of transferring a monosaccharide from an activated sugar donor to an acceptor substrate (Lairson *et al.*, 2008). The most studied type of glycosylation is the one catalyzed by uridine diphosphate (UDP)-dependent glycosyltransferases (UGTs), a type of GTs belonging to GT family 1. Nucleotide-activated sugars are used as glycosyl by UGTs to form *O*-, *N*-, *S*-, and *C*-glycosides (Lairson *et al.*, 2008; Tegl and Nidetzky, 2020). UGTs utilize a variety of sugar donors, including UDP-glucose, UDP-glucuronic acid, UDP-galactose, and UDP-xylose. Among these, UDP-glucuronic acid is the most extensively studied in animal UGTs, while UDP-glucose is predominantly associated with plant UGTs (Meech *et al.*, 2012; Tiwari *et al.*, 2016). The addition of glucuronide is a process in phase II metabolism that increases the solubility of xenobiotics to facilitate their elimination, which is a crucial process in detoxification (Lethe *et al.*, 2024).

UGT families are classified based on phylogeny, with families 1 to 2 being specific to humans and families 71 to 100 corresponding to plant species. These UGT families share a minimum of 40% sequence homology and are further divided into subfamilies, represented by a letter that exhibits at least 60% homology (Ross *et al.*, 2001). In mammals and humans, UGTs are membrane-bound enzymes located in the lumen of the endoplasmic reticulum, containing transmembrane-spanning regions or motifs. In contrast, plant UGTs are freely soluble in the cytoplasm (Ross *et al.*, 2001).

UGTs are characterized by the presence of a GT-B fold, comprising two Rossmann domains separated by a catalytic cleft (Lairson *et al.*, 2008). But the defining protein motif of UGTs is a highly conserved consensus sequence of 40-50 amino acid length classified as PS00375 in PROSITE (Sigrist *et al.*, 2013). This conserved protein motif is usually located in the C-terminal domain of the protein and is thought to be involved in the binding of the protein to the nucleotide moiety of the UDP-activated sugar (Ross *et al.*, 2001). This C-terminal domain is also named plant secondary product glycosyltransferase (PSPG) motif (Hughes and Hughes, 1994; Osmani *et al.*, 2008). On the other hand, the N-terminal domain is the responsible for interacting with the acceptor substrate (Lairson *et al.*, 2008; Rahimi *et al.*, 2019).

4.1. Substrate recognition and reaction mechanism of glycosyltransferases

As an example of plant GTs, UGT71G1 presents a consensus amino acid region that participates in the recognition of the sugar moiety of UDP-glucose: Asn-361, Glu-381, Gln-382 and Trp-360 (Shao *et al.*, 2005). Thr-143 has been also reported as important for binding of the glucose moiety through the 6'OH (Maharjan *et al.*, 2020). The selection of

UDP-glucose over another UDP-sugar is partially determined by the last amino acid of its consensus sequence, being Gln-382. In fact, in homologs capable of binding UDP-galactose this last residue is replaced with histidine, but the discrimination of UDP-glucose and UDP-galactose appears to be more complex than only this substitution. The discrimination of UDP-glucose over UDP-glucuronic acid is due to the presence of Thr-143 able to form a hydrogen bond with the 6'-hydroxyl group of UDP-glucose (Lethe *et al.*, 2024). Humans GTs mostly use UDP-glucuronic acid as sugar donor molecule, instead of UDP-glucose. On the other hand, plant GTs use UDP-glucose as the donor molecule for glycosylation, leading to the formation of a substrate-3-*O*- β -*D*-glucoside with the releasing of UDP as a side product (Lethe *et al.*, 2024).

4.2. Glycosylation of plant secondary metabolites and occurrence in strawberry

UGTs, along with acyltransferases, methyltransferases, and cytochrome P450s, play a key role in the diversification of secondary metabolites, modifying their physicochemical properties (Gachon *et al.*, 2005).

The glycosylation at 3-OH position of anthocyanins has been described as a crucial modification for the chemical stability of these compounds (Gachon *et al.*, 2005; Bowles *et al.*, 2006). Strawberry accumulates high levels of antioxidants such as anthocyanins (Pombo *et al.*, 2011) and flavonoids. Some interesting compounds from strawberries have been described as substrates for glycosylation, for example quercetin and kaempferol are the two most important flavonols in this plant, and their occurrence is as glucosides (Häkkinen *et al.*, 1999). Specifically, quercetin 3-*O*/7-*O*/4-*O*-glucosides, kaempferol 3-*O*/7-*O*-glucosides as well as rhamnetin 7-*O*-glucoside have been detected in strawberry fruit (Griesser *et al.*, 2008). The VOC 4-hydroxy-2,5-dimethyl-3(2*H*)-furanone (furanol) is susceptible to glycosylation to form furaneol-glucoside (Song *et al.*, 2015). In the study of Song *et al.*, (2015), they demonstrate that *UGT71K3*-silenced strawberry fruits decreased furaneol-glucoside levels compared with the control (Song *et al.*, 2016). A functional characterization of several strawberry enzymes belonging to the family UGT71 was performed (Song *et al.*, 2015). In this study, the downregulation of *UGT71W2* resulted in lower levels of kaempferol glucoside, 3-hydroxycoumarin glucoside, quercetin glucoside, kaempferol-3-(6'-coumaroyl) glucoside and ABA-glucoside. Also, to investigate the regioselectivity of the members of UGT71, they used quercetin as substrate, due to five hydroxyl groups susceptible to form an *O*-glycosidic bond. They found out that *UGT71A33* formed mainly 7- and 3'-glucosides while *UGT71A34* and *UGT71A35* were responsible for the accumulation of 7- and 3-glucosides, respectively. Interestingly, only *UGT71W2* was responsible for the accumulation of quercetin 3-glucoside where diglucosides of quercetin only appeared by the action of *UGT71A33* and *UGT71A34*.

Results

1. Variation of ellagitannins content in ‘232’ × ‘1392’ mapping population

Red fruits from F₁ mapping population (‘232’ × ‘1392’) were collected from two consecutive years: 2013 and 2014. These fruits were grounded to powder and then used for secondary metabolites analysis as explained in the correspondent section (**Material and Methods**). Ellagitannins were evaluated in the two parental lines and their F₁ progeny by Ultra Performance-Liquid Chromatography coupled to Tandem Mass Spectrometry (UPLC-Orbitrap-MS/MS). These results were previously collected and presented in Delphine Pott’s PhD thesis (<https://hdl.handle.net/10630/18672>), and partially published in Pott *et al.*, (2020).

Ellagitannins is a highly diverse family of secondary metabolism with multiple isomeric forms. The ellagitannin content in parental and F₁ lines was collected from Delphine Pott’s PhD thesis (<https://hdl.handle.net/10630/18672>) and summarized in **Table 11**.

Table 11. Relative concentration of ellagitannins identified in the UPLC-Orbitrap-MS/MS. Retention time (RT), molecular ion (m/z). Mean values of the parental ‘232’ and the range (min, max) of the F₁ population in both years (2013 and 2014) are shown. Bold values indicate a significant difference between ‘232’ and ‘1392’ parental lines (p-value < 0.05). Data are relativized to the mean value of the ‘1392’ parental line.

Metabolite	RT (min)	m/z	mode (+/-)	232 parental		2013		2014	
				2013	2014	min	max	min	max
HHDP glucose 1	1.58	481.06	-	1.28	0.79	0.57	4.65	0.36	3.51
HHDP glucose 2	2.06	481.06	-	1.05	0.65	0.49	2.57	0.21	2.76
Bis(HHDP) glucose 1	3.66	783.07	-	1.16	1.00	0.35	3.81	0.27	3.74
Bis(HHDP) glucose 2	3.80	783.07	-	0.86	0.61	0.55	1.98	0.37	2.96
Bis(HHDP) glucose 3	4.52	783.07	-	0.83	0.56	0.51	1.99	0.38	2.76
Galloyl-HHDP-glucose	5.09	633.07	-	1.06	0.76	0.47	2.61	0.20	3.49
Di-galloyl-HHDP glucose	5.95	785.08	-	0.51	0.44	0.27	1.88	0.00	2.86
Ellagic acid deoxyhexose 1	6.67	447.06	-	1.00	0.73	0.60	2.02	0.41	2.49
Ellagic acid deoxyhexose 2	6.79	447.06	-	1.22	0.98	0.65	4.40	0.41	3.89
Ellagic acid	7.00	301.00	-	1.05	0.69	0.30	2.60	0.22	6.42
Galloyl-bis(HHDP)-glucose 1	6.16	935.08	-	0.50	0.40	0.23	2.05	0.10	7.58
Galloyl-bis(HHDP)-glucose 2	6.29	935.08	-	0.66	0.67	0.26	2.33	0.14	5.52
Galloyl-bis(HHDP)-glucose 3	6.76	935.07	-	1.19	0.90	0.29	3.26	0.15	12.78
Galloyl-bis(HHDP)-glucose 4	6.88	935.08	-	3.21	2.18	0.57	10.91	0.25	6.72
Tri-galloyl-HHDP glucose 1	6.16	937.08	-	0.45	0.45	0.15	2.05	0.17	6.83

Metabolite	RT (min)	m/z	mode (+/-)	232 parental		2013		2014	
				2013	2014	min	max	min	max
Tri-galloyl-HHDP glucose 2	6.29	937.08	-	0.64	0.58	0.25	2.45	0.14	5.55
Unknown ellagitannin 1	4.48	935.08	+	0.86	0.38	0.47	2.85	0.06	3.26
Unknown ellagitannin 2	4.69	935.08	+	1.29	0.88	0.63	3.63	0.12	14.73
Unknown ellagitannin 3	6.28	935.08	+	0.67	0.25	0.00	2.84	0.00	4.06
Ellagic acid hexose 1	5.76	463.05	-	0.94	0.62	0.38	2.23	0.00	3.18
Ellagic acid hexose 2	5.87	463.05	-	1.55	0.80	0.00	4.75	0.00	3.13
Ellagic acid hexose 3	6.01	463.05	-	1.02	0.71	0.53	1.96	0.31	2.69
Tris-galloyl-glucose 1	5.09	635.08	-	1.05	0.75	0.47	2.60	0.17	3.73
Tris-galloyl-glucose 2	5.69	635.09	-	0.29	0.31	0.00	1.29	0.00	2.98
Rhoipteleatin H 1	3.95	965.05	-	0.55	0.73	0.00	3.72	0.00	18.60
Castalagin 1	5.62	933.06	-	0.63	0.61	0.33	2.46	0.23	3.81
Castalagin 2	5.94	933.06	-	1.24	0.63	0.37	3.05	0.17	6.23
Castalagin 3	6.28	933.06	-	0.79	0.70	0.31	2.69	0.20	4.29
Castalagin 4	6.37	933.06	-	1.50	0.85	0.42	4.22	0.00	9.78
Castalagin 5	6.87	933.06	-	0.72	0.80	0.33	3.62	0.23	6.41
Castalagin 6	7.46	933.06	-	0.77	0.85	0.37	3.77	0.18	7.44
Lagerstannin A 1	4.79	799.06	-	0.42	0.75	0.14	1.81	0.00	5.24
Lagerstannin A 2	5.77	799.06	-	0.45	0.75	0.18	1.90	0.14	3.72
Lagerstannin B 1	4.82	949.06	-	0.96	0.81	0.38	2.28	0.07	6.31
Lagerstannin B 2	7.30	949.06	-	0.47	0.65	0.14	2.23	0.10	2.77
Lagerstannin B 3	7.42	949.06	-	0.81	0.53	0.00	2.25	0.00	3.28
Lagerstannin B 4	8.25	949.06	-	0.52	0.65	0.20	2.92	0.15	2.57

A total of 37 distinct ellagitannins compounds were detected over two consecutive years. A high number of isomers were identified for each ellagitannin, which resulted in grouping them into only 14 main categories; the remaining compounds were differentiated based solely on their retention times. In general, all detected ellagitannins exhibited minimum relative concentrations close to 0, reflecting low accumulation levels in some F₁ individuals. Nonetheless, sharp increases in the relative concentrations were observed for certain compounds, such as galloyl-bis-(HDDP)-glucose 3 (2014), galloyl-bis-(HDDP)-glucose 3 (2013) or rhoipteleatin H1 (2014).

Significant differences in ellagitannin concentrations were detected for F₁ individuals derived from the crossing of ‘232’ and ‘1392’ parentals. Specifically, 21 from 37 detected ellagitannins compounds showed statistically significant variation between the parental lines. The ellagitannins with significant changes in the F₁ population were grouped into 11 distinct categories: HDDP glucose, bis-(HDDP)-glucose, di-galloyl-HDDP-glucose,

galloyl-bis(HDDP)-glucose, tri-galloyl-HDDP-glucose, unknown ellagitannin, ellagic acid hexose, tris-galloyl-glucose, castalagin, lagerstannin A and lagerstannin B.

2. QTLs controlling ellagitannins content in ‘232’ × ‘1392’ F₁ population

All statistically significant QTLs detected for ellagitannins content were listed on **Table 12**. Additionally, **Fig. 45** showed a representation of some QTLs in their respective linkage groups in the two consecutive years of this study.

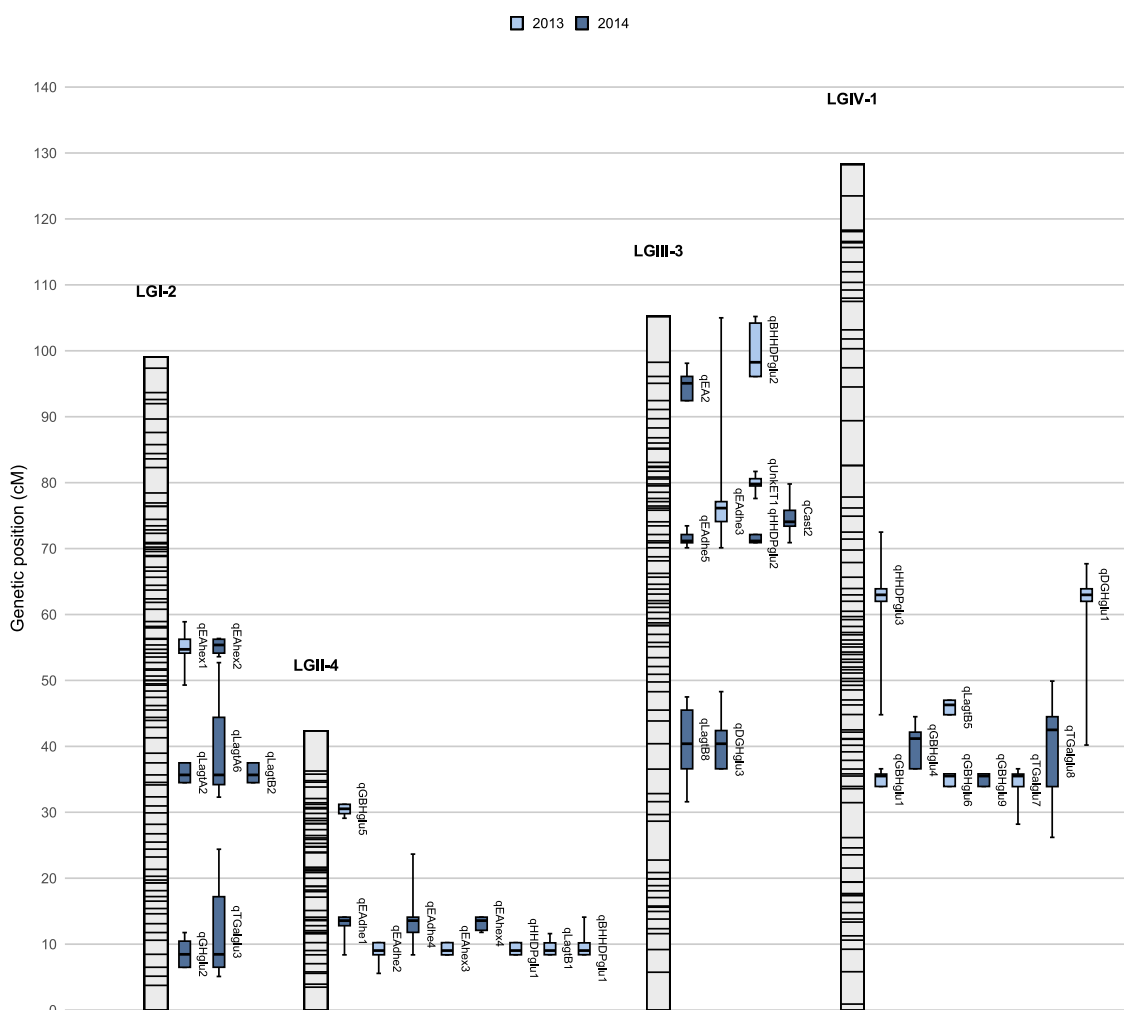


Figure 45. Representation of some QTLs controlling ellagitannins content in ‘232’ × ‘1392’ F₁ population based on *F. vesca* genome. Four different linkage groups (LG) are represented depending on the genetic distance (cM). QTLs are represented as 1-LOD interval for the colored rectangles, whiskers for 2-LOD interval and black lines for the genetic position where LOD score is the highest in the interval. Black lines crossing linkage groups represents genetic markers used to generate the linkage map. Further information is provided on **Annex 13** and **Annex 14**.

Ellagitannins content seemed to be controlled widespread in the genome due to QTLs presence in every linkage group (LGI – LGVII) and more specifically, in our study we found QTLs in 19 from 33 possible linkage groups (see **Annex 13**). The total number of QTLs identified ascended to 76 where 43 of them were detected with fruits harvested in 2013 from mapping population and the rest were detected using 2014 season.

In this study, 15 different compounds were studied for ellagitannins content comprising from chemically simple ellagitannins like ellagic acid or HHDP-glucose to more complex chemical species like lagerstanin A and B or rhoipteleatin H. In this whole set of ellagitannins only one chemical structure remains unclear, being called unknown ellagitannin and identified through m/z pattern and retention time. From this QTL analysis, we identify LGII-4, LGIII-3, LGIV-1 as the linkage groups where more QTLs were detected, with 9 QTLs per linkage group in any case.

Almost all identified QTLs were co-localized within specific regions across different linkage groups. For ellagic acid hexose, four QTLs were identified, corresponding to two isomeric forms of the same compound, distinguishable only by their retention times and classified as ellagic acid hexose 2 (RT = 5.87 min) and ellagic acid hexose 3 (RT = 6.01 min) (see **Table 11**). In addition, QTLs with overlapping LOD positions were observed, such as those associated with ellagic acid deoxyhexose. Specifically, qEAdhe1, qEAdh2, and qEAdh4 were all mapped to nearly identical physical positions within LGII-4. This pattern was frequently observed for multiple ellagitannins, suggesting the presence of potential metabolic hotspots.

Specifically for this work, we focused our attention on the QTLs controlling around 50-70% of the total variation of EAH2: qEAhex1 and qEAhex2. These two QTLs shared the same physical position in the linkage maps, where only qEAhex1 had a wider 2-LOD interval.

Table 12. Metabolic QTLs (mQTLs) of ellagitannins detected for two years using restricted multiple QTL mapping (rMQM) analysis. Metabolites are shown with their retention time (RT) in minutes (min); Group, linkage group based on *F. vesca* reference genome; Pos, genetic position in cM; 2LODl and 2LODu, lower and upper genetic position of the 2-LOD Bayesian interval in cM, respectively; LOD, logarithm of the odds ratio; a, mean phenotypic difference between the two homozygous *loci* of the QTL; r^2 , percentage of phenotypic variance explained by the QTL; KW, significance level of the Kruskal-Wallis test.

QTL name	Metabolite – RT (min)	Year	Group	Marker	Pos (cM)	2LODl (cM)	2LODu (cM)	LOD	a	r^2	KW
qBHHPglu1	Bis-HHDP-glucose - 3.80	2013	LGII-4	25187-5:T>C-	9.03	8.40	14.10	6.45	0.03	26.40	*
qBHHPglu2	Bis-HHDP-glucose - 3.80	2013	LGIII-3	F13596-62:A>G-LG3	98.26	96.10	105.20	4.81	0.03	18.80	*
qBHHPglu3	Bis-HHDP-glucose - 4.52	2013	LGVII-3	27124-23:A>G-	42.28	40.70	54.80	6.48	0.01	21.40	***
qCast1	Castalagin - 5.62	2014	LGVII-4	F11347-65:C>A-LG7**	13.05	6.40	21.70	5.53	0.15	25.70	*
qCast10	Castalagin - 7.46	2013	LGVI-3	M27561-10:A>T-	26.74	4.20	33.10	5.72	0.09	23.20	***
qCast11	Castalagin - 7.46	2013	LGVI-5	37297-9:A>T-	38.36	30.20	46.20	6.92	0.09	30.70	*
qCast2	Castalagin - 5.94	2014	LGIII-3	34601--	74.06	70.90	79.80	5.67	0.11	19.90	*
qCast3	Castalagin - 6.28	2013	LGIV-4	ChFaM023-151h	107.85	104.00	124.20	4.92	0.04	13.00	*
qCast4	Castalagin - 6.28	2013	LGVI-2	M14618-24:C>T-LG6	24.48	2.00	33.30	6.22	0.04	15.20	*
qCast5	Castalagin - 6.28	2013	LGVI-3	10796-19:G>C-LG6	24.20	21.90	24.60	5.31	0.04	12.40	*
qCast6	Castalagin - 6.28	2014	LGVII-4	F03318-60:G>C-LG7	8.73	6.40	21.70	4.93	0.18	21.20	*
qCast7	Castalagin - 6.37	2014	LGVII-3	F34133-39:T>C-LG7	21.02	19.10	22.30	6.25	0.24	19.70	*
qCast8	Castalagin - 6.87	2013	LGII-2	M28484-36:G>T-	65.96	63.50	81.80	4.96	0.03	10.30	*
qCast9	Castalagin - 6.87	2013	LGIV-4	ChFaM148-156/158	114.77	111.60	124.20	9.14	0.03	22.90	*
qDGHglu1	Di-galloyl-HHDP-glucose - 5.95	2013	LGIV-1	F33514-56:C>A-LG4	62.98	40.20	67.70	5.09	0.07	16.80	***
qDGHglu2	Di-galloyl-HHDP-glucose - 5.95	2013	LGVI-5	M33723-47:A>T-LG6	40.61	28.20	48.20	5.05	0.07	17.60	*
qDGHglu3	Di-galloyl-HHDP-glucose - 5.95	2014	LGIII-3	16712-44:C>A-	40.40	36.60	48.30	6.57	0.04	34.30	*
qEA1	Ellagic acid - 7.00	2014	LGVII-2	SAAT-S2	44.82	39.20	53.10	5.61	0.16	20.60	***
qEA2	Ellagic acid - 7.00	2014	LGIII-3	F11907-10:C>T-LG3	95.07	92.44	98.11	5.53	0.16	18.10	**
qEAdhe1	Ellagic acid deoxyhexose 1 - 6.67	2014	LGII-4	11775-27:A>G-LG2	13.57	8.38	14.10	7.18	0.06	25.60	***
qEAdhe2	Ellagic acid deoxyhexose 2 - 6.79	2013	LGII-4	25187-5:T>C-	9.03	5.57	10.23	6.57	0.01	28.00	****
qEAdhe3	Ellagic acid deoxyhexose 2 - 6.79	2013	LGIII-3	M41100-38:G>A-LG3	76.14	70.12	105.00	4.19	0.01	16.20	*

QTL name	Metabolite – RT (min)	Year	Group	Marker	Pos (cM)	2LODl (cM)	2LODu (cM)	LOD	a	r ²	KW
qEAdhe4	Ellagic acid deoxyhexose 2 - 6.79	2014	LGII-4	11775-27:A>G-LG2	13.57	8.38	23.66	3.44	0.02	9.60	**
qEAdhe5	Ellagic acid deoxyhexose 2 - 6.79	2014	LGIII-3	F23153-22:C>T-LG3	71.16	70.12	73.45	4.91	0.02	14.30	*
qEAhex1	Ellagic acid hexose 2 - 5.87	2013	LGI-2	M10828-43:T>G-	54.72	49.30	58.90	23.34	0.08	70.70	****
qEAhex2	Ellagic acid hexose 2 - 5.87	2014	LGI-2	30785-64:G>T-LG1	55.38	53.60	56.36	21.31	0.05	51.40	****
qEAhex3	Ellagic acid hexose 3 - 6.01	2013	LGII-4	25187-5:T>C-	9.03	8.38	10.23	7.27	0.05	29.50	****
qEAhex4	Ellagic acid hexose 3 - 6.01	2014	LGII-4	11775-27:A>G-LG2	13.57	11.79	14.10	4.51	0.12	20.40	****
qGBHglu1	Galloyl-bis-HHDP-glucose - 6.16	2013	LGIV-1	ChFaM017-136	35.50	33.92	36.60	13.99	0.07	31.50	****
qGBHglu2	Galloyl-bis-HHDP-glucose - 6.16	2013	LGVI-2	M31478-44:T>A-LG6	16.03	13.74	29.70	5.55	0.07	10.30	*
qGBHglu3	Galloyl-bis-HHDP-glucose - 6.16	2013	LGVII-3	F39028-48:A>C-LG7	47.58	42.28	55.98	5.51	0.07	10.70	*
qGBHglu4	Galloyl-bis-HHDP-glucose - 6.16	2014	LGIV-1	F20883-65:G>T-LG4	41.18	36.60	44.50	8.16	0.20	28.60	***
qGBHglu5	Galloyl-bis-HHDP-glucose - 6.88	2013	LGII-4	ChFaM004-147	30.53	29.10	31.21	8.65	0.04	12.70	*
qGBHglu6	Galloyl-bis-HHDP-glucose - 6.88	2013	LGIV-1	ChFaM017-136	35.50	33.92	35.83	27.17	0.04	42.40	****
qGBHglu7	Galloyl-bis-HHDP-glucose - 6.88	2013	LGIV-4	21004-37:C>T-LG4	19.31	1.40	20.82	7.66	0.04	7.70	***
qGBHglu8	Galloyl-bis-HHDP-glucose - 6.88	2013	LGVI-2	M18650-56:C>T-	13.74	4.00	24.50	9.39	0.04	11.60	*
qGBHglu9	Galloyl-bis-HHDP-glucose - 6.88	2014	LGIV-1	ChFaM017-136	35.50	33.92	35.83	13.63	0.26	49.70	****
qGHglu1	Galloyl-HHDP-glucose - 5.09	2013	LGVI-2	M25388-24:A>G-	18.80	10.95	29.73	7.89	0.06	34.30	*
qGHglu2	Galloyl-HHDP-glucose - 5.09	2014	LGI-2	F12694-18:A>T-LG1	8.46	6.49	11.76	7.67	0.08	22.00	*
qGHglu3	Galloyl-HHDP-glucose - 5.09	2014	LGIII-1	25437-34:C>T-	22.68	19.67	23.10	5.08	0.08	15.70	*
qGHglu4	Galloyl-HHDP-glucose - 5.09	2014	LGVII-2	21145-30:G>A-LG7	46.80	44.82	50.35	7.59	0.08	23.50	***
qHHDPglu1	HHDP-glucose - 1.58	2013	LGII-4	25187-5:T>C-	9.03	8.39	10.23	7.13	0.01	32.90	*
qHHDPglu2	HHDP-glucose - 1.58	2014	LGIII-3	F23153-22:C>T-LG3	71.16	70.90	72.13	5.92	0.02	20.90	***
qHHDPglu3	HHDP-glucose - 2.06	2013	LGIV-1	F33514-56:C>A-LG4	62.98	44.80	72.50	5.59	0.05	18.80	****
qHHDPglu4	HHDP-glucose - 2.06	2013	LGIV-3	LOX-S2	38.51	32.20	53.20	4.22	0.05	14.30	*
qLagtA1	Lagerstannin A - 4.79	2013	LGII-5	M43853-58:C>A-LG2	56.42	55.70	58.10	9.20	0.10	42.20	***
qLagtA2	Lagerstannin A - 4.79	2014	LGI-2	F32156-17:A>C-LG1	35.67	34.50	37.50	5.48	0.07	29.90	*
qLagtA3	Lagerstannin A - 4.79	2014	LGVI-1	M17235-40:A>T-LG6	9.96	0.00	38.10	5.32	0.07	18.80	*

VI. Chapter II: Candidate gene for ellagitannins
content in strawberry fruit

qLagtA4	Lagerstannin A - 5.77	2013	LGII-5	15703-58:G>A-LG2	56.93	55.70	58.10	8.72	0.05	31.00	***
qLagtA5	Lagerstannin A - 5.77	2013	LGIV-3	F21099-40:A>G-	44.78	32.20	53.20	6.87	0.05	23.80	*
qLagtA6	Lagerstannin A - 5.77	2014	LGI-2	F32156-17:A>C-LG1	35.67	32.30	52.70	4.49	0.20	30.20	*
qLagtA7	Lagerstannin A - 5.77	2014	LGV-1	F34200-22:T>A-LG5	26.06	13.80	31.30	5.62	0.20	24.70	***
qLagtB1	Lagerstannin B - 4.82	2013	LGII-4	25187-5:T>C-	9.03	8.40	11.60	5.33	0.09	26.70	*
qLagtB2	Lagerstannin B - 4.82	2014	LGI-2	F32156-17:A>C-LG1	35.67	34.50	37.50	6.52	0.20	30.70	**
qLagtB3	Lagerstannin B - 4.82	2014	LGVII-1	F37091-34:C>A-LG6	35.86	34.60	38.10	7.35	0.21	20.90	*
qLagtB4	Lagerstannin B - 7.30	2013	LGII-5	M43853-58:C>A-LG2	56.42	55.70	58.10	8.46	0.10	36.60	*
qLagtB5	Lagerstannin B - 7.42	2013	LGIV-1	F20746-41:A>T-LG4	46.29	44.80	47.00	6.37	0.08	16.50	*
qLagtB6	Lagerstannin B - 7.42	2013	LGIV-3	F23458-34:A>C-LG4	43.37	42.20	46.70	8.17	0.08	21.90	*
qLagtB7	Lagerstannin B - 7.42	2013	LGV-4	M16560-9:T>A-	42.71	41.10	46.80	6.96	0.08	17.20	**
qLagtB8	Lagerstannin B - 7.42	2014	LGIII-3	16712-44:C>A-	40.40	31.60	47.50	4.83	0.14	23.20	*
qRhoH1	Rhoipteleatin H - 3.95	2013	LGV-4	17622-5:A>T-	33.61	19.20	42.70	5.02	0.53	20.70	***
qRhoH2	Rhoipteleatin H - 3.95	2014	LGVII-1	F37091-34:C>A-LG6	35.86	25.10	37.40	5.72	0.38	18.00	***
qRhoH3	Rhoipteleatin H - 3.95	2014	LGVII-2	20385-29:A>C-	56.27	31.40	62.80	4.66	0.38	14.90	*
qTGalglu1	Tris-galloyl-glucose - 5.09	2013	LGVII-2	M25388-24:A>G-	18.80	13.70	27.70	8.96	0.04	28.20	*
qTGalglu2	Tris-galloyl-glucose - 5.09	2013	LGVII-3	14529-36:G>A-LG7	47.80	40.70	62.20	3.79	0.04	9.20	*
qTGalglu3	Tris-galloyl-glucose - 5.09	2014	LGI-2	F12694-18:A>T-LG1	8.46	5.10	24.40	5.64	0.10	15.70	*
qTGalglu4	Tris-galloyl-glucose - 5.09	2014	LGVII-2	21145-30:G>A-LG7	46.80	37.20	53.60	4.41	0.10	12.60	**
qTGalglu5	Tris-galloyl-glucose - 5.69	2013	LGIII-1	F18418-49:A>C-LG3	46.19	43.00	47.80	6.96	0.03	19.60	****
qTGalglu6	Tris-galloyl-glucose - 5.69	2013	LGV-3	M12353-41:C>T-	47.95	25.60	54.10	4.25	0.03	11.50	**
qTGalglu7	Tri-galloyl-HHDP-glucose - 6.16	2013	LGIV-1	ChFaM017-136	35.50	28.20	36.60	10.01	0.14	31.10	****
qTGalglu8	Tri-galloyl-HHDP-glucose - 6.16	2014	LGIV-1	11205-28:A>G-LG4	42.50	26.20	49.90	5.85	0.09	21.60	****
qTGalglu9	Tri-galloyl-HHDP-glucose - 6.16	2014	LGV-1	M11115-55:G>A-LG5	13.77	4.90	16.10	4.74	0.09	19.50	*
qUnkET1	Unknown ellagitannin - 4.48	2013	LGIII-3	F40029-28:G>A-LG3	79.80	77.60	81.70	6.22	0.04	15.10	*
qUnkET2	Unknown ellagitannin - 4.48	2014	LGII-5	11687-58:T>A-LG2	37.00	34.70	37.80	6.19	0.05	22.90	*
qUnkET3	Unknown ellagitannin - 4.69	2013	LGVII-1	F10992-38:T>A-LG7	23.34	9.70	32.80	4.80	0.06	10.50	**
qUnkET4	Unknown ellagitannin - 4.69	2014	LGVII-1	F14673-22:C>G-LG7	45.87	40.80	49.80	5.01	0.45	17.00	*

3. Transcriptomic analysis of contrasting lines for ellagitannins content using RNA-seq

Focusing on ellagic acid derivatives, only ellagic acid hexose 1 and 2 presented statistical differences in 2014 and 2013, respectively (**Table 12**). These isomers are distinguished based on their retention times showing that ellagic acid hexose 2 (EAH2) was more retained in the chromatographic column than the isomer 1. Nevertheless, details about chemical structure about EAH2 are unknown by now.

Based on EAH2 content some F₁ individuals were classified into low and high EAH2 relative content. The mean levels of EAH2 in the F₁ population during 2013 and 2014 and EAH2 relative concentration of selected F₁ lines were represented on **Fig. 46**.

Lines selected for RNA-seq analysis (**Fig. 46B**) presented either absent levels (L13-14, L41-14, L45-14, L13-13 and L62-13) or relative high levels of ellagitannins (L19-14, L51-14, L27-14, L19-13 and L51-13), making suitable to study the expression profile of both circumstances. Indeed, in both years the results were reproducible for the same levels were obtained in the high or low EAH situations.

The selection of differentially expressed genes (DEGs) was based on fold-change (FC) and p-value. More specifically, we use similar criteria to that in **Chapter I** for selecting significant transcripts: p-value < 0.05 and log(FC) < -1 for down-regulated genes and log(FC) > 1 for up-regulated genes. With these criteria, 586 down-regulated genes arose from the RNA-seq data, and 429 genes were classified as up-regulated.

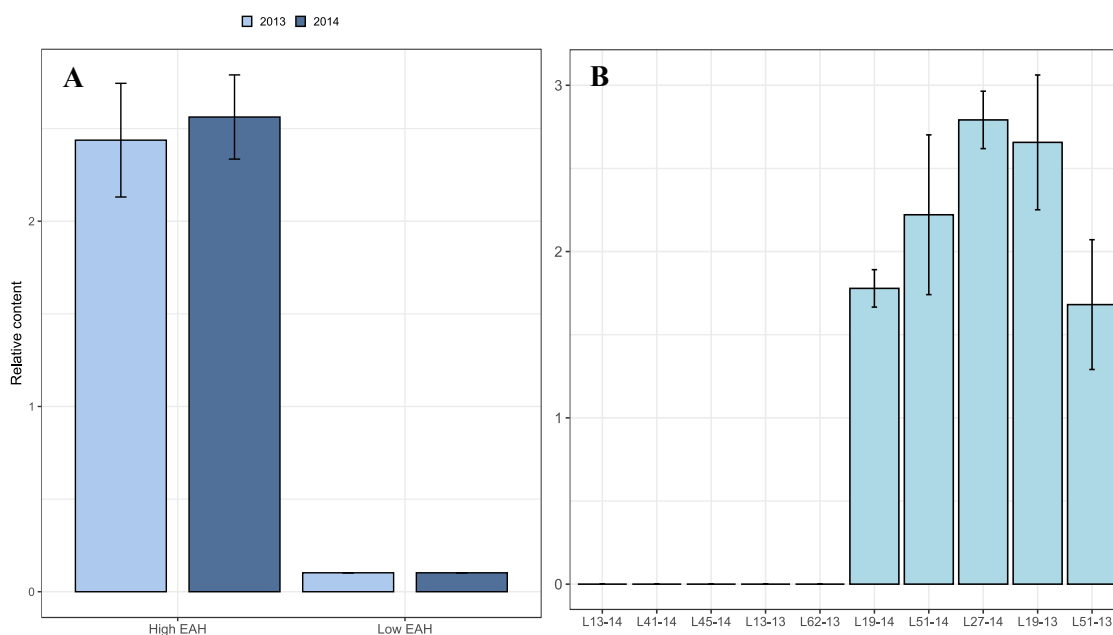


Figure 46. Ellagic acid hexose 2 (EAH2) content in '232' × '1392' F₁ population. (A) EAH2 relative content in two consecutive years; (B) EAH2 relative content in selected lines for RNA-seq.

Changing the significance statistic criterium for the false discovery ratio (FDR) and applying a filter of $FDR < 0.05$, then a total of 27 DEGs were found, where 8 of them were down-regulated genes and 19 were up-regulated genes. A volcano plot was built to observe statistically significant DEGs where orange-colored and blue-colored points represented up and down-regulated genes, respectively. The volcano plot was represented on **Fig. 47**.

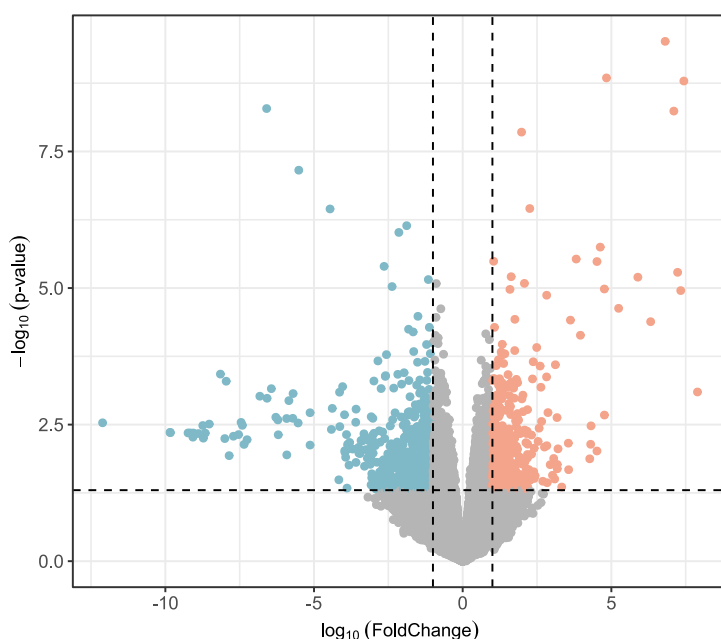


Figure 47. Volcano plot of differentially expressed genes (DEGs) for ellagitannins content. The criteria for DEGs selection were a $\log_{10}(FC) < -1$ or $\log_{10}(FC) > 1$ and a $p\text{-value} < 0.05$. Blue and orange dots represented down and up-regulated genes, respectively.

4. Identification of *FxaC_3g11070* as candidate for controlling ellagic acid hexose content

The physical regions corresponding to qEAhex1 and qEAhex2 spanned 9.60 and 2.76 cM, respectively. The overlapping region between both QTLs was defined by the boundaries of qEAhex2 and encompassed approximately 500 genes. By integrating RNA-seq data, only differentially expressed genes (DEGs) within these QTL intervals were considered, narrowing the candidate gene list to 172.

Gene sequences delimited within qEAhex1 and qEAhex2 were collected from GDR database. A molecular function was assigned to each of the gene sequences using Mercator4 tool (Bolger *et al.*, 2021). A summary of the functional annotations of these DEGs were collected in **Fig. 48**.

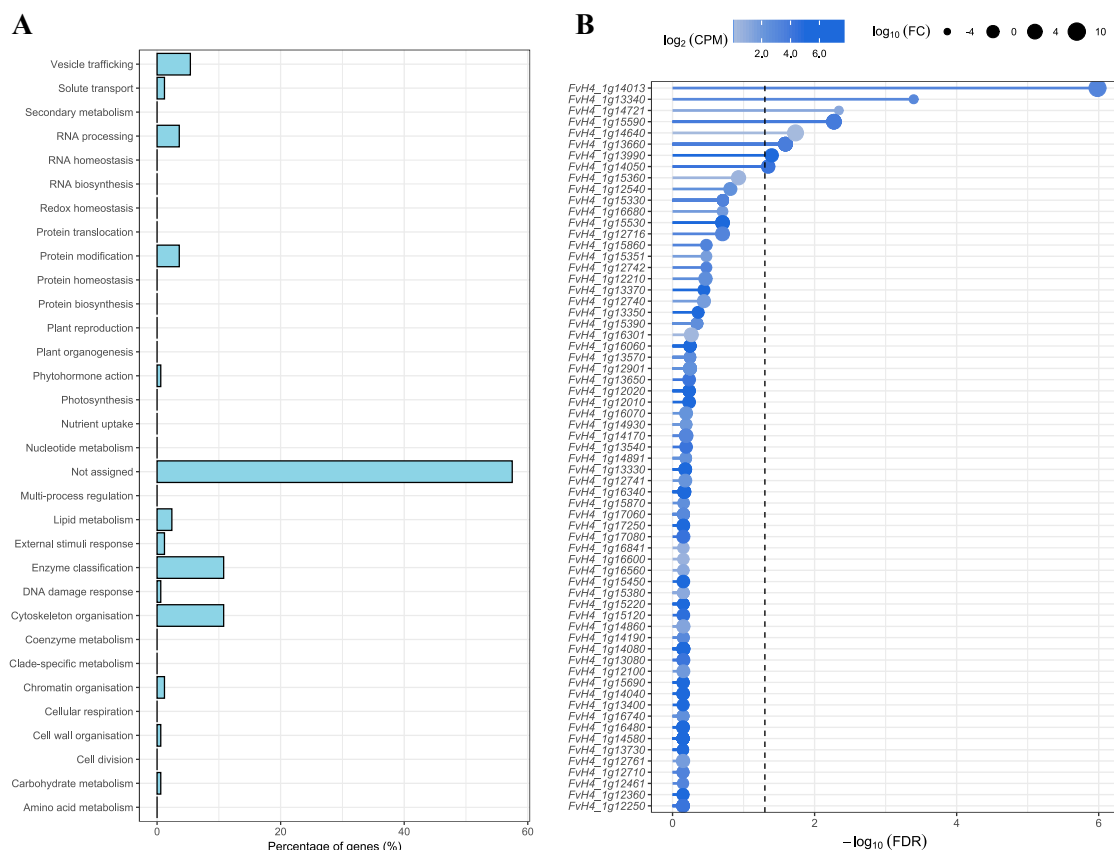


Figure 48. Genes for QTLs controlling ellagic acid hexose content. (A) Functional annotation of selected genes within the QTL (172 genes); (B) RNA-seq data representation of genes within the QTLs (172 genes). Genes were annotated with Mercator4 (Bolger *et al.*, 2021). Genes are annotated with *F. vesca* genome v4 (Edger *et al.*, 2017). Dotted line in B represents a threshold for $\text{FDR} = 0.05$.

The majority of genes located within the QTL regions were not assigned to any functional category, with approximately 60% lacking annotated molecular function. Among the annotated genes, the most represented functional categories were enzyme classification and cytoskeleton organization, followed by vesicle trafficking. The “not assigned” category encompassed a broad range of molecular functions, including genes involved in secondary metabolism to other cellular processes not adequately represented in the predefined categories.

Fig. 48B showed the DEGs within the QTL interval, along with their expression changes ($\log_{10}(\text{FC})$) and associated false discovery ratio (FDR). Based on the statistical threshold of $\text{FDR} < 0.05$, only eight genes were selected for further analysis. The integration of RNA-seq expression data and functional annotation enabled the identification of candidate genes potentially involved in ellagic acid hexose accumulation, as summarized in **Table 13**.

Table 13. Summary of selected DEGs in the QTLs interval. Information from RNA-seq analysis was used together with functional annotation. Gene expression is presented as counts per million (CPM). “-” indicates missing molecular function based on Mercator4 software. Genes are shown based on *F. vesca* reference genome v4 (Edger *et al.*, 2017).

Gene	logFC	log ₂ CPM	FDR	Molecular function
<i>FvH4_1g14013</i>	9.06	3.19	$1.04 \cdot 10^{-6}$	Anthocyanidin 3- <i>O</i> -glucosyltransferase
<i>FvH4_1g13340</i>	-3.28	2.93	$4.03 \cdot 10^{-5}$	-
<i>FvH4_1g14721</i>	-3.46	1.12	$4.55 \cdot 10^{-3}$	Transcription factor (BEL)
<i>FvH4_1g15590</i>	3.64	3.79	$5.34 \cdot 10^{-3}$	Phospholipase A2 (pPLA2-II)
<i>FvH4_1g14640</i>	6.69	0.11	$1.86 \cdot 10^{-2}$	-
<i>FvH4_1g13660</i>	1.83	3.63	$2.57 \cdot 10^{-2}$	Integrin-like protein (AT14a)
<i>FvH4_1g13990</i>	1.34	5.69	$4.02 \cdot 10^{-2}$	Anthocyanidin 3- <i>O</i> -glucosyltransferase
<i>FvH4_1g14050</i>	1.14	4.26	$4.50 \cdot 10^{-2}$	MFS-type solute transporter

Genes lacking annotated molecular functions were excluded from further studies. Among the remaining candidates, *FvH4_1g14013* emerged as the top candidate due to the highest fold-change and the lowest FDR among the selected DEGs. This gene encodes an anthocyanidin-3-*O*-glucosyltransferase, an enzyme potentially involved in the biosynthesis or modification of metabolites associated with the identified QTLs. Given that this molecular function appeared twice among the selected genes, *FvH4_1g13990*, which encoded a similar enzyme, was also prioritized as a second candidate.

To further characterize these genes, homologous sequences in *F. × ananassa* ('Camarosa' genome v1.0.a1 from Edger *et al.*, 2019) were identified using BLAST searches against their predicted protein sequences. Additionally, each sequence retrieved from the 'Camarosa' genome v1.0.a1 (Edger *et al.*, 2019) was subsequently BLASTed against the updated version of the 'Camarosa' genome (Liu *et al.*, 2021), and the corresponding identifiers from the new version were retrieved.

The identified homologs of the candidate gene *FvH4_1g13990* were:

- snap_masked-Fvb1-2-processed-gene-90.25-mRNA-1 (*FxaC_2g19460*).
- augustus_masked-Fvb1-3-processed-gene-79.10-mRNA-1 (*FxaC_3g15640*).
- augustus_masked-Fvb1-3-processed-gene-79.11-mRNA-1 (*FxaC_3g15650*).

Similarly, the second candidate gene, *FvH4_1g14013* exhibited three homologous sequences:

- maker-Fvb1-4-snap-gene-68.38-mRNA-1 (*FxaC_1g15670*).
- maker-Fvb1-3-augustus-gene-52.35-mRNA-1 (*FxaC_3g11070*).
- maker-Fvb1-2-snap-gene-90.66-mRNA-1 (*FxaC_2g19510*).

Homologous amino acid sequences for each *F. vesca* candidate gene were retrieved and aligned pairwise with their corresponding *F. vesca* counterparts. The results of these sequence comparisons were presented in **Table 14**.

Every *F. × ananassa* protein sequence was annotated as anthocyanidin 3-*O*-glucosyltransferase, as well as the *F. vesca* candidate genes. All homologs of *FvH4_1g13990* exhibited a substantial reduction in protein sequence length, resulting in low sequence identity and similarity (below 30% in all cases). In contrast, a completely different situation was observed with homologs of *FvH4_1g14013*, as the protein encoded by *FxaC_3g11070* showed an almost identical amino acid length, sequence identity, and sequence similarity. Nevertheless, protein codified by *FxaC_1g15670* and *FxaC_2g19510* exhibited larger amino acids sequences with the consequent reduction of sequence identity and similarity below 50% in any case.

Table 14. Protein sequence comparison of *Fragaria × ananassa* homologs of candidate genes associated with EAH2 content. Sequences were retrieved from GDR database and pairwise alignment were performed to compare between sequences. Functional annotations were assigned using Mercator4 (Bolger *et al.*, 2021). Aa: number of amino acids; ID: sequence identity; SM: sequence similarity.

Protein	Aa	% ID	% SM	Functional annotation
Homologs to candidate gene <i>FvH4_1g13990</i>				
<i>FxaC_2g19460</i>	137	26.70	27.30	Anthocyanidin 3- <i>O</i> -glucosyltransferase
<i>FxaC_3g15640</i>	119	11.20	14.80	Anthocyanidin 3- <i>O</i> -glucosyltransferase
<i>FxaC_3g15650</i>	144	14.80	20.10	Anthocyanidin 3- <i>O</i> -glucosyltransferase
Homologs to candidate gene <i>FvH4_1g14013</i>				
<i>FxaC_1g15670</i>	952	49.40	49.70	Anthocyanidin 3- <i>O</i> -glucosyltransferase
<i>FxaC_2g19510</i>	1,033	43.70	54.70	Anthocyanidin 3- <i>O</i> -glucosyltransferase
<i>FxaC_3g11070</i>	477	94.60	96.10	Anthocyanidin 3- <i>O</i> -glucosyltransferase

All *F. × ananassa* homologs presented the same functional annotation, therefore developing biological roles as anthocyanidin-3-*O*-glucosyltransferase in the strawberry plant context.

From these homologs, only 2 homologs to each of the *F. vesca* genes showed expression in the RNA-seq data. For the case of *FvH4_1g13990* only *FxaC_3g15640* and *FxaC_3g15650* were detected in the RNA-seq analysis. On the other hand, only *FxaC_1g15670* and *FxaC_3g11070* homologs to *FvH4_1g14013* gene were detected in the same set of RNA-seq data.

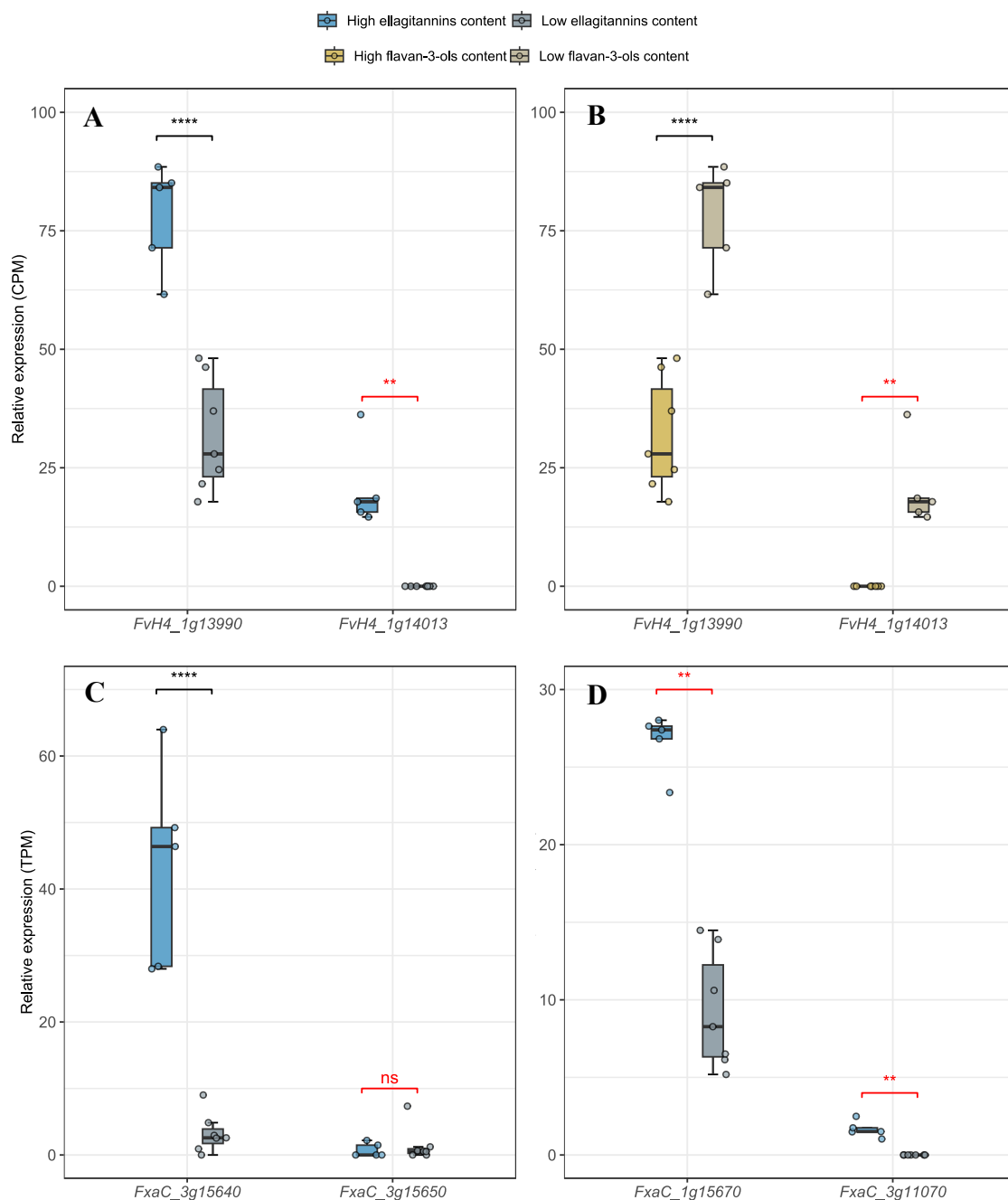


Figure 49. Gene expression of QTLs candidates and their *F. x ananassa* homologs. (A) Candidate genes expression related to ellagitannins content; (B) Candidate genes expression related to proanthocyanidins content; (C) Gene expression of *F. x ananassa* homologs to *FvH4_1g13990*; (D) Gene expression of *F. x ananassa* homologs to *FvH4_1g14013*. Two statistical tests are shown in this graph: Student's *t*-test (in black) and Mann-Whitney *U* test (in red).

Since ellagitannins and proanthocyanidins derived from different branches of the shikimate pathway, we tried to associate *F. x ananassa* candidates with concentrations of

ellagitannins and flavan-3-ols (monomers of proanthocyanidins). Gene expressions of the candidates and their respective homologs were summarized in **Fig. 49**.

When we observed the gene expression of lines with high content in ellagitannins, both *F. vesca* genes had relatively high expression while lower to null expression was detected in lines with low content in ellagitannins. Expression of these genes had an inverted behavior when we checked the content in flavan-3-ols. Specifically, the gene *FvH4_1g14013* had no expression in lines with low ellagitannins content and high flavan-3-ols, showing more drastic changes in between the two groups.

In the case of *F. × ananassa* homologs from *FvH4_1g13990*, only *FxaC_3g15640* presented statistical changes between high and low content of ellagitannins. On the other hand, homologs from *FvH4_1g14013* correlated their expression with the content of ellagitannins in both cases. Particularly, *FxaC_3g11070* shared the same behavior as *FvH4_1g14013*: they had no expression when the content of ellagitannins is low and when the content of ellagitannins is higher, the expression arose.

Finally, *FxaC_3g11070* was selected as the main candidate gene for EAH2 content, as it exhibited the same expression trend as *FvH4_1g14013* with respect to ellagitannin levels, along with high sequence identity and similarity at the protein level.

4. Gene expression of ellagitannins candidate in Goodberry population

To assess the potential involvement of these candidate genes in ellagitannins biosynthesis, expression profiles were analyzed within the GoodBerry population (GBP, **Chapter I**) using RNA-seq data. To test the consistency of our hypothesis regarding the involvement of *FxaC_3g11070* in EAH2 content, we analyzed the expression of this candidate gene in additional strawberry varieties ('Candonga' and 'Senga Sengana'). The aim was to determine whether a consistent relationship between gene expression and EAH2 concentration could be observed across different genetic backgrounds.

In order to do that, we checked the expression of the pre-selected candidates in the GBP, and the results were exhibited on **Fig. 50**. We first tried to associate one genetic marker with the content of EAH2 but there was no correlation between any genetic marker and the content in that secondary metabolite.

Both *FxaC_3g11070* and *FxaC_1g15670* were studied for 'Candonga' and 'Senga Sengana' varieties grown in two different locations: France and Spain. Values from different locations were almost similar, presenting similar tendency even considering the two genes. Nevertheless, *FxaC_1g15670* had more overall expression than *FxaC_3g11070* considering strawberry variety or the different location. In both cases, 'Candonga' variety presented lower levels of expression for both candidates without distinction for each location. In fact, the levels of *FxaC_3g11070* were absent in

‘Candonga’ variety, while *FxaC_1g15670* has significantly low expression in this same cultivar than in ‘Senga Sengana’.

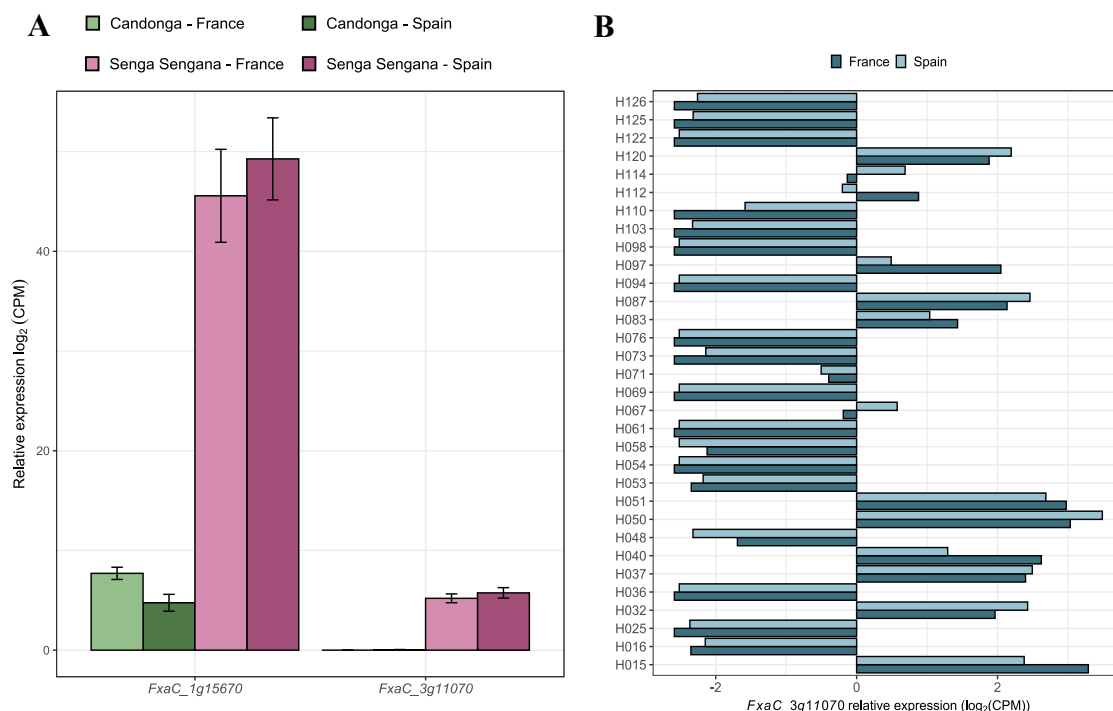


Figure 50. Ellagitannins candidates in Goodberry population. (A) *FvH4_1g14013* homologs expression of the GBP parents in two locations; (B) *FxaC_3g11070* expression in GBP F₁ population in two locations.

Then, the expression of the *FxaC_3g11070* was studied in the whole GBP F₁ population and we were able to distinguish individuals with low and high expression of the gene. From this population, 19 individuals were classified in the group of low content in EAH2 while 10 individuals presented high levels of EAH2. In this case, we selected H015, H032, H037, H050, H051 and H087 as representatives of the group of the high expression of the gene; while H016, H053, H054, H058, H098 and H122 were individuals with low *FxaC_3g11070*.

Secondary metabolites analysis with UPLC-Orbitrap-LC/MS were extracted and analyzed. A total of 27 secondary metabolites were detected belonging to three different chemical classifications: ellagitannins, flavan-3-ols and procyanidin. A total of 17 chemical compounds were assigned to the ellagitannin's category. The distinguishing of different isomers was done depending on their corresponding retention time (Table 5). The results of this secondary metabolite analysis were represented in Fig. 51.

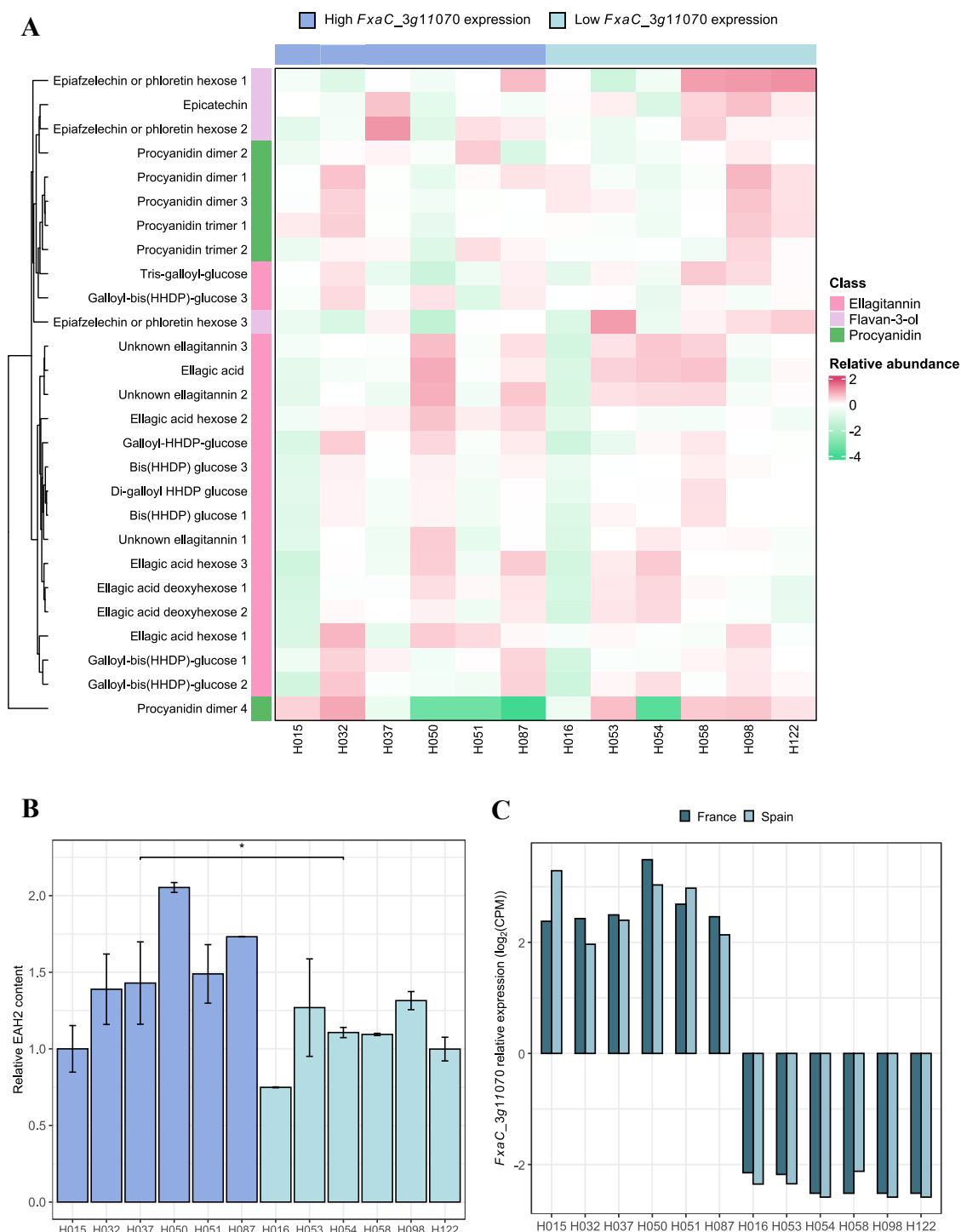


Figure 51. Secondary metabolites in Goodberry population with *FxaC_3g11070* (UDP) expression. (A) Heatmap of the secondary metabolite studied in GBP F₁ population; **(B)** Ellagic acid hexose 2 content in the selected F₁ lines; **(C)** *FxaC_3g11070* expression in the selected lines in two locations (France and Spain). Statistical significance was tested with Student's *t*-test (p-value < 0.05).

This analysis led us to detect several compounds related to ellagic acid metabolism, such as the proper ellagic acid, three different isomers of ellagic acid hexose and two isomers

from ellagic acid deoxyhexose. Only the isomer EAH2 was found statistically significant when comparing groups with high and low candidate gene expression. Nevertheless, no clear lineal correlation was found between gene expression and EAH2 accumulation (Annex 15).

5. Discrepancies in gene sequence of *FxaC_3g11070*

The *FxaC_3g11070* coding sequence was amplified from red fruit tissue in all cases using UDP primers (Table 6).

We discovered that *FxaC_3g11070* didn't have any expression in *F. × ananassa* cv. 'Camarosa', as shown in Fig. 52. In contrast, *F. vesca* homolog *FvH4_1g14013* had the highest expression in leaf tissue. In our case, the candidate could not be cloned using any tissue from 'Camarosa' variety.

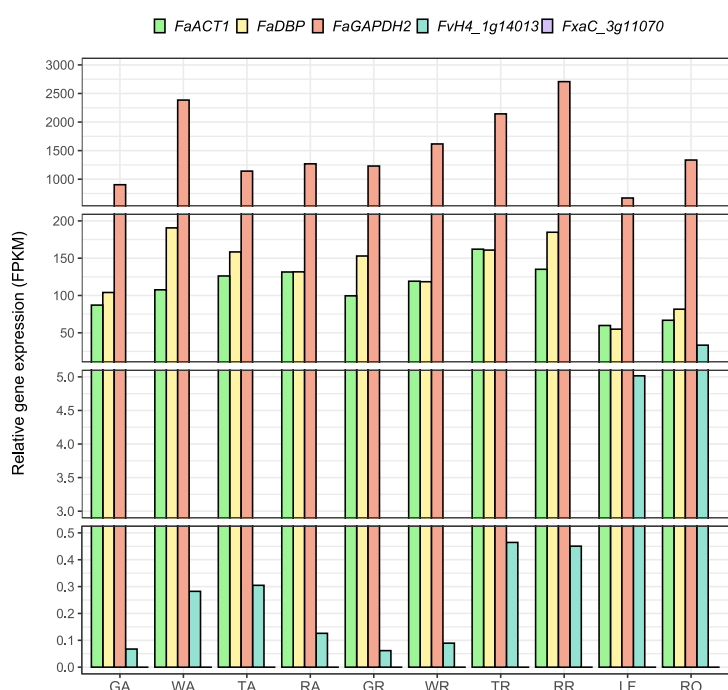


Figure 52. Gene expression of *F. × ananassa* cv. 'Camarosa' (*FxaC_3g11070*) and *F. vesca* homolog (*FvH4_1g14013*).

Housekeeping genes (*FaACT1*, *FaDBP* and *FaGAPDH2*) were used for comparison. GA: green achene; WA: white achene; TA: turning achene; RA: red achene; GR: green receptacle; WR: white receptacle; TR: turning receptacle; RR: red receptacle; LF: leaf; RO: roots.

Firstly, genetic material from F₁ individuals derived from the crossing of '232' and '1392' was used to obtain gene sequence. Individuals with high ellagitannin content were selected to clone the sequence, assuming the positive correlation between gene expression and EAH2 content. The two chosen individuals to obtain genetic sequences were L27-14 and L51-14.

DNA sequencing of cloned region resulted in a set of single nucleotide polymorphism (SNP) for the sequence of L51-14. One characteristic SNP was the change of a thymine for a cytosin in the STOP codon TGA, provoking the appearance of an arginine at the C-terminus of the protein. Consequently, this protein had 5 extra amino acids than the gene sequence present in GDR database. On the other hand, individual L27-14 presented an

insertion of 141 bp that produces a premature STOP codon, resulting in a defective protein.

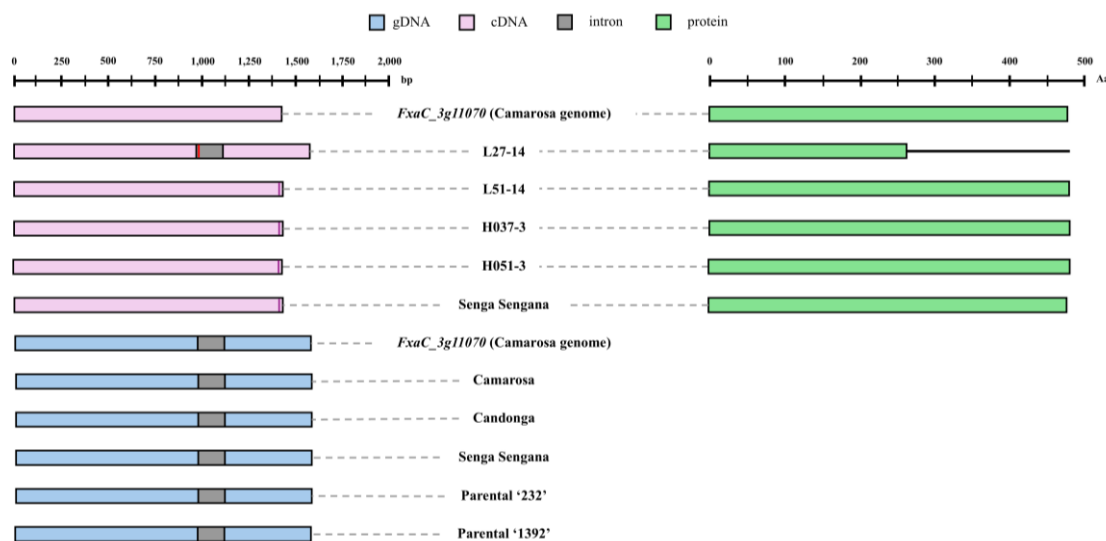


Figure 53. Summary of DNA and protein alignment of different sequences of *FxaC_3g11070*. DNA alignments are shown in base-pair (bp) and protein alignments are shown in number of amino acids (Aa). Magenta line represents a change in the STOP codon for an arginine while red line represents a premature STOP codon.

We moved then to the GBP, the crossing between ‘Candonga’ and ‘Senga Sengana’ varieties. In this case, we cloned the gene from two different individuals: H037-3 and H051-3. The sequences from these two individuals differed in several SNPs but the length was the same and the SNP that changed the STOP codon for an arginine was conserved in both sequences. Then, ‘Candonga’ and ‘Senga Sengana’ fruit tissue was used to clone *FxaC_3g11070*. No amplification was accomplished with ‘Candonga’ variety but the sequence from ‘Senga Sengana’ was like the ones from H037-3 and H051-3, conservating the same length and some of the SNPs compared to the reference sequence from the ‘Camarosa’ genome available on GDR database.

Lastly, cloned sequences differed from the one from ‘Camarosa’ reference genome and shared common changes among L51-14, H037-3, H051-3 and ‘Senga Sengana’ individuals. All the information about alignments was summarized on **Fig. 53** and extra information was available on **Annexes 16** and **17**.

Finally, for further studies we selected ‘Senga Sengana’ sequence as representative of *FxaC_3g11070*. Then, *FxaC_3g11070* was named *FaUDPSS*.

6. Gene expression of candidate gene in strawberry varieties

Previously, we demonstrated that *FxaC_3g11070* expression was positively correlated with EAH2 concentration in the F1 mapping population (‘Candonga’ × ‘Senga Sengana’)

and showed that gene expression patterns varied among the ‘Camarosa’, ‘Candonga’, and ‘Senga Sengana’ varieties. In this context, we hypothesized that *FxaC_3g11070* could represent a desirable trait for strawberry breeding. To investigate this further, we examined the expression profile of the candidate gene across a broader set of commercial strawberry cultivars. For this purpose, we analyzed gene expression data from six commercial varieties (‘Candonga’, ‘Clery’, ‘Florence’, ‘Frida’, ‘Gariguette’, and ‘Sonata’) cultivated in four distinct locations across Europe. Expression profiles of *FxaC_1g15671* and *FxaC_3g11070* are presented in **Fig. 54**.

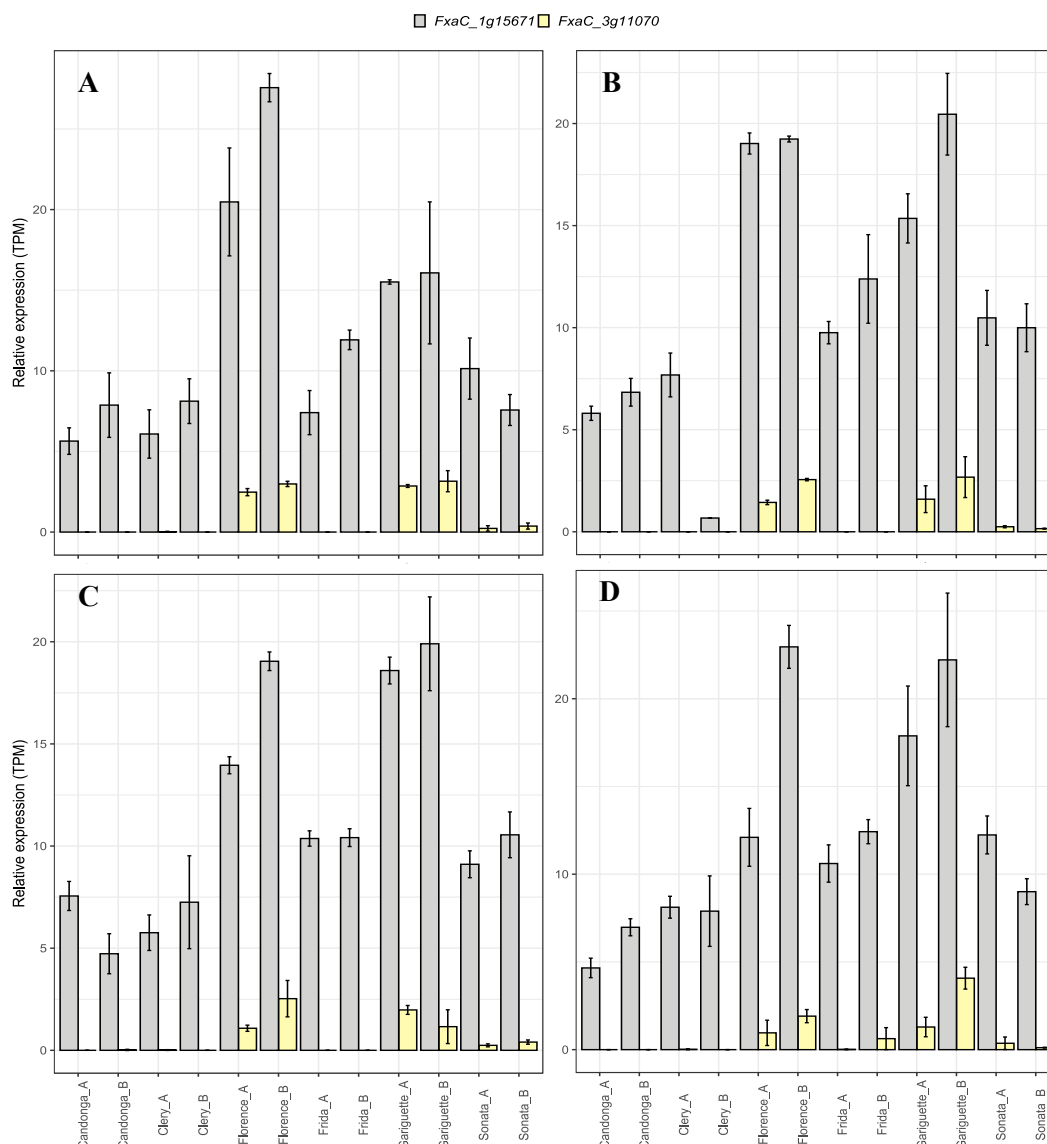


Figure 54. Candidate genes expression in four different locations. (A) France; **(B)** Germany; **(C)** Poland and **(D)** Norway. Data shown as mean $\pm$ standard deviation. A and B following the name of the variety corresponds to two different harvesting times.

FxaC_1g15671 showed higher expression level than its homolog *FxaC_3g11070* and the expression levels are consistent in the four locations but having differences among the

varieties. On the other hand, *FxaC_3g11070* only had expression in ‘Florence’, ‘Frida’ and ‘Sonata’ varieties but with low levels. The rest of the varieties didn’t show expression for this gene, which matches the data from **Fig. 50** about no expression in the ‘Camarosa’ parental.

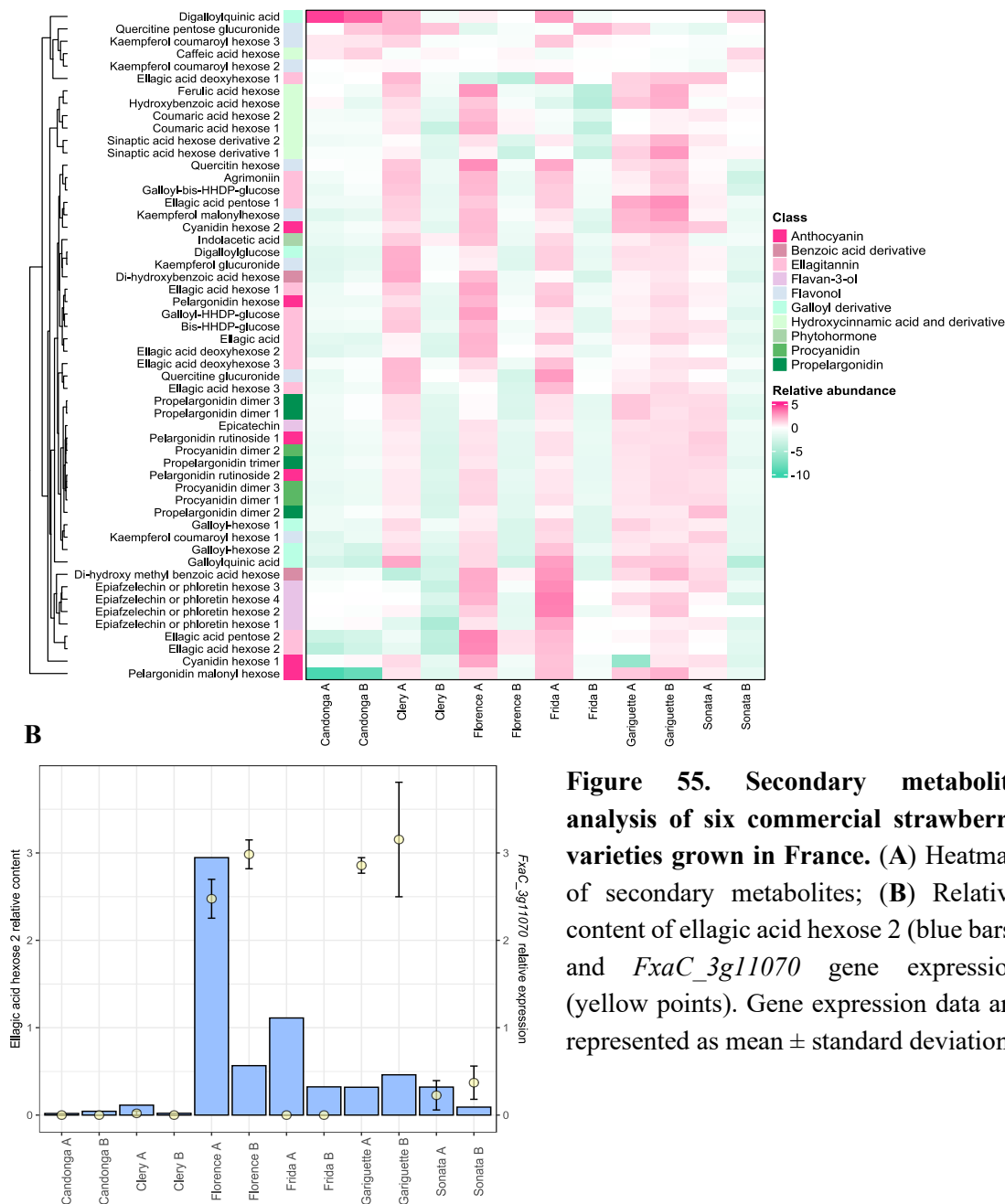


Figure 55. Secondary metabolite analysis of six commercial strawberry varieties grown in France. (A) Heatmap of secondary metabolites; **(B)** Relative content of ellagic acid hexose 2 (blue bars) and *FxaC_3g11070* gene expression (yellow points). Gene expression data are represented as mean $\pm$ standard deviation.

We also performed a secondary metabolite analysis of the different commercial varieties with different harvesting times (A and B). We were able to detect a set of 54 different secondary metabolites comprising ten chemical classes. Again, different isomers are identified by their retention time (**Table 5**). These results were shown on **Fig. 55**.

Several samples followed the same tendency: A-harvesting time resulted in higher values in all varieties except for ‘Candonga’ variety. There was a positive tendency between

EAH2 accumulation and *FxC_3g11070* gene expression. In the samples with no expression of the candidate gene, the amount of the secondary metabolite was low but when the samples had expression of the gene, the concentration increased. The only exception was 'Frida' variety because they showed no expression of the candidate gene but the concentration of the EAH2 was higher than in 'Candongra' and 'Clery' varieties.

7. Phylogenetic analysis and protein prediction of *FaUDPSS*

Protein sequences were retrieved mainly from two databases: GDR and NCBI. Sequence selection was performed by searching for PF00201 hidden Markov model corresponding to UDP-glucuronosyl/UDP-glycosyltransferase motifs. A total of 120 different sequences were recruited from seven species: *A. thaliana*, *Camellia sinensis*, *Diospyros kaki*, *F. × ananassa*, *F. vesca*, *S. lycopersicum* and *V. vinifera*. Of these sequences, 69 corresponded to commercial strawberry (*F. × ananassa*) and 7 were selected from *F. vesca*, with over 60% of the sequences belonging to the *Fragaria* genus.

Phylogeny of UDPSS and several other glucosyltransferases were studied, and the results are summarized in **Fig. 56. Annex 18** presented extra information about glucosyltransferase proteins studied in the phylogenetic tree.

Clade A contains the highest number of proteins from *A. thaliana*. This tree suggested that the glucosyltransferases selected in this study derived from a common ancestor and consecutive clades evolved from previous clades. In this case, A-clade is phylogenetically elder than the other clades and L-clade is the newest.

The candidate *FaUDPSS* was clustered in the E clade of the tree, this clade being the one with more members. All the proteins of this clade presented a sequence comprising 460 to 515 amino acids, except for *FaUGT11* that was 205 amino acid long. This clade included four proteins from *A. thaliana* (*AtUGT71B1*, *AtUGT71C5*, *AtUGT72B1* and *AtUGT88A1*) and the rest of the proteins belonged to *F. vesca* and *F. × ananassa*. The closest proteins to *FaUDPSS* were *FaUGT1* and *FvUGT2*, which are the proteins codified by *FxC_1g15671* and *FvH4_1g14013*, respectively. *FaUDP* was the 'Camarosa' reference sequence present in the GDR database, and it is related closely with *FaUGT4*. All these proteins were clustered together in a sub-clade where *AtUGT72B1* (Rehman et al., 2018) was phylogenetically close to *FaUGT57*.

Regarding subcellular localization, almost 95% of all proteins presented cytoplasmic localization, followed by an extracellular localization. Only *FaUGT49* presented a chloroplastidic localization and *FaUGT44* was predicted to be on the endoplasmic reticulum. Regarding the candidate *FaUDPSS*, this protein was predicted to be present in the cytoplasm, as well as *FaUDP* protein.

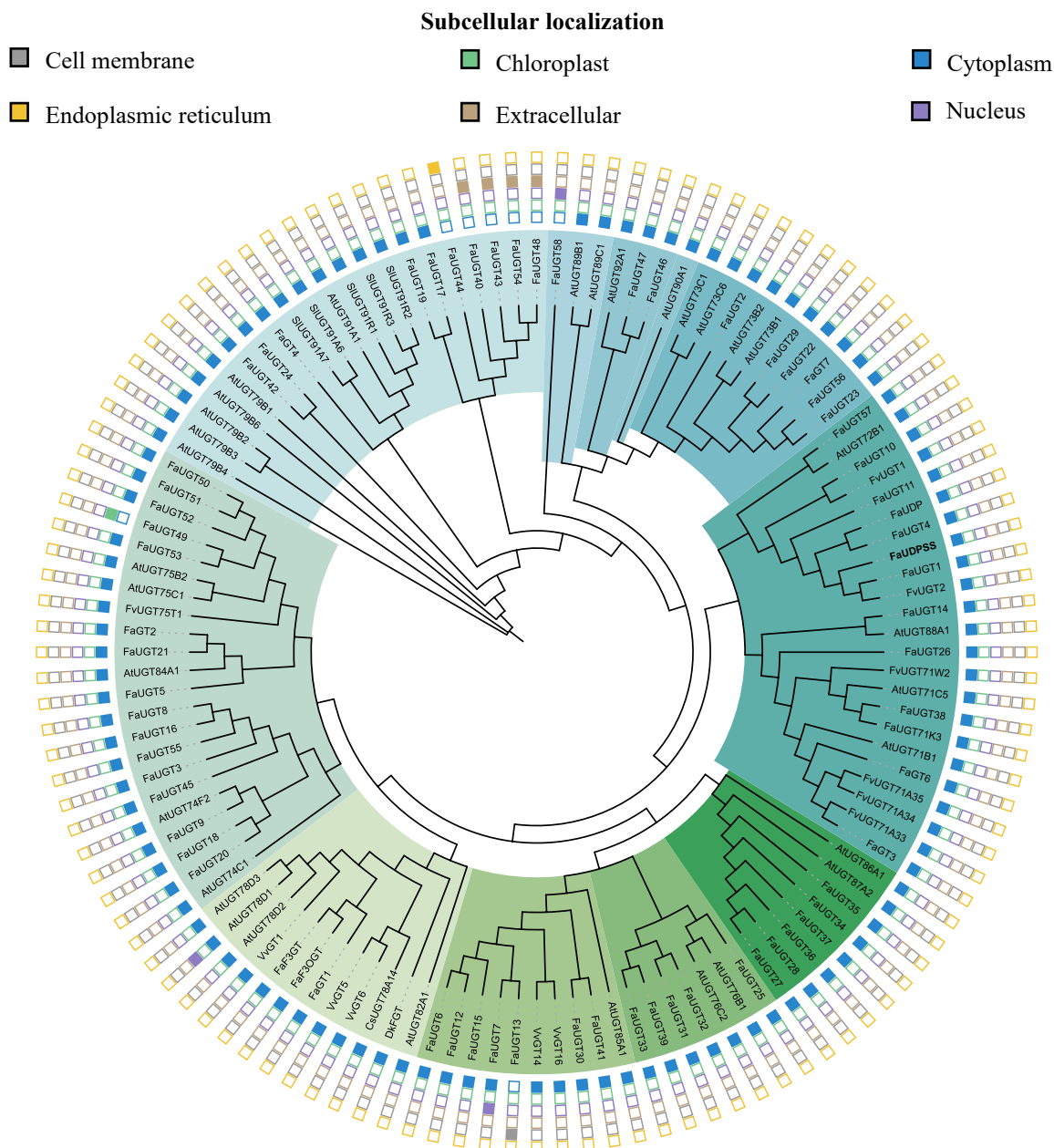


Figure 56. Phylogenetic tree of glucosyltransferases. The present tree was based on amino acid sequences of several species. Colors represent different groups based on phylogenetic relationships. The species used in this study are *Arabidopsis thaliana* (At), *Camellia sinensis* (Cs), *Diospyros kaki* (Dk), *Fragaria × ananassa* (Fa), *Fragaria vesca* (Fv), *Solanum lycopersicum* (Sl) and *Vitis vinifera* (Vv). For each sequence the colored squares represent the subcellular localization: blue (cytoplasm), green (chloroplast), magenta (nucleus) and grey (cell membrane).

Using AlphaFold3 we predicted the tridimensional structure of FaUDPSS and this information was recruited on **Fig. 57**. The protein conformation was mainly globular with amino acids disposed predominantly in α -helixes, with only two β -sheets regions.

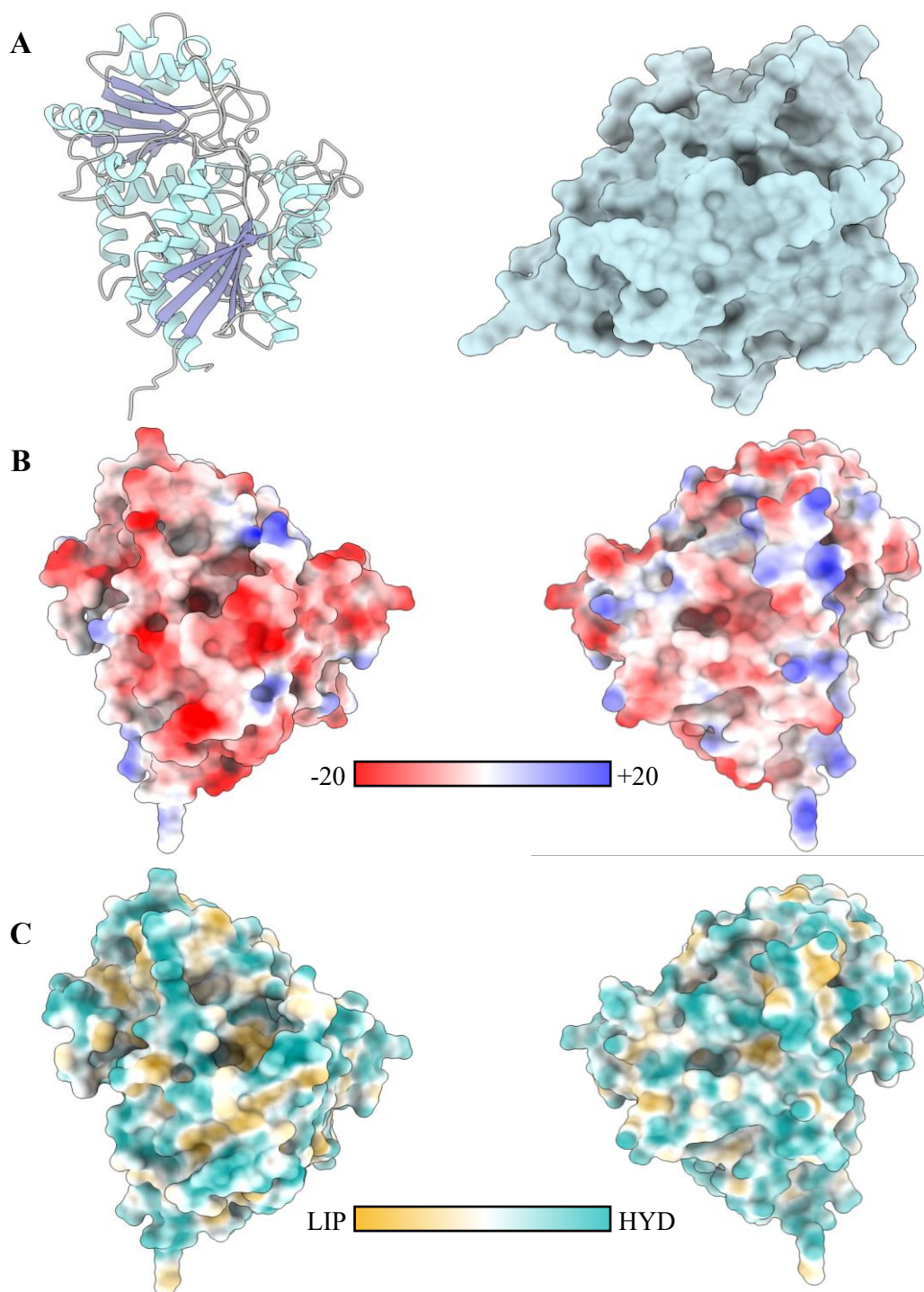


Figure 57. Protein modelling and visualization of FaUDPSS. (A) Representation of tridimensional structure of FaUDPSS colored by secondary structure and representation of protein surface; (B) Coulombic surface potential; (C) Hydrophilic surface coloring (HYD: Hydrophilic, LIP: Lipophilic or hydrophobic). Structural modelling and visualizations were performed with AlphaFold3 and ChimeraX, respectively.

Regarding Coulombic surface potential, the protein showed high electronegativity in almost all the surface, while electropositive zones were much more uncommon. For this

reason, most of the protein with polar amino acids were exposed to the medium. **Fig. 57C** showed the tendency of each amino acid to interact with polar molecules like water (hydrophilic) or with nonpolar molecules (hydrophobic). Almost all the surface of FaUDPSS was analyzed as hydrophilic, where hydrophobic residues were very restricted to certain inner zones of the protein external surface.

We proposed ellagic acid as substrate of glycosylation by FaUDPSS and using UDP-glucose as a sugar-donor molecule. For this purpose, we modelled FaUDPSS with ellagic acid and UDP-glucose and the results were displayed on **Fig. 58**. Modeling values were higher to 0.90 (ipTM= 0.99 and pTM= 0.98) indicating high confidence for the prediction with both molecules.

Interaction between FaUDPSS and the two substrates was highlighted on **Fig. 58B**. Some amino acids seemed to be involved in the interaction with some substrates. Ellagic acid was physically proximal to Ala391, Asp183, Asp315, His19, Phe117, Phe318, Pro184 and Val85 and therefore these amino acids could be involved in molecular docking. On the other hand, UDP-glucose was a larger molecule and presented more interactions than ellagic acid. In fact, the amino acids highlighted were: Ala351, Ala391, Asn372, Cys139, Gln353, Gln393, Glu376, Glu392, His368, Ile21, Ser280, Ser373, Pro138, Trp350 and Trp371.

The visualization of the FaUDPSS protein surface (**Fig. 58A**) allowed us to identify three protein pockets, designated as S1, S2, and P. We hypothesized that these regions are involved in the passage of substrates through the pockets, necessary for the enzymatic reaction. For each of the pockets, we estimated the corresponding area by assuming a circular surface. The diameter was measured as the distance between amino acids within each pocket, which was then used to calculate the radius.

A closer view of S1 revealed the presence of the uracil motif of UDP-glucose, which could indicate the entrance of this molecule to the catalytic domain of the protein. Electrostatic surface coloring showed a positive charge of amino acids involved in the formation of the pocket, corresponding to positively charged amino acids at biological conditions: Arg248, Arg251, Trp350 and Pro352. The presence of positively charged amino acid residues conferred a hydrophilic character to the S1 pocket, thereby creating a favorable environment for the accommodation of a polar ligand such as the UDP-sugar. The area of this pocket was estimated to be $0.40 \pm 0.09 \text{ nm}^2$ (**Annexes 19A and 19B**), which appears sufficient to accommodate UDP-glucose, whose topological polar surface area is reported as 0.29 nm^2 according to [PubChem](#).

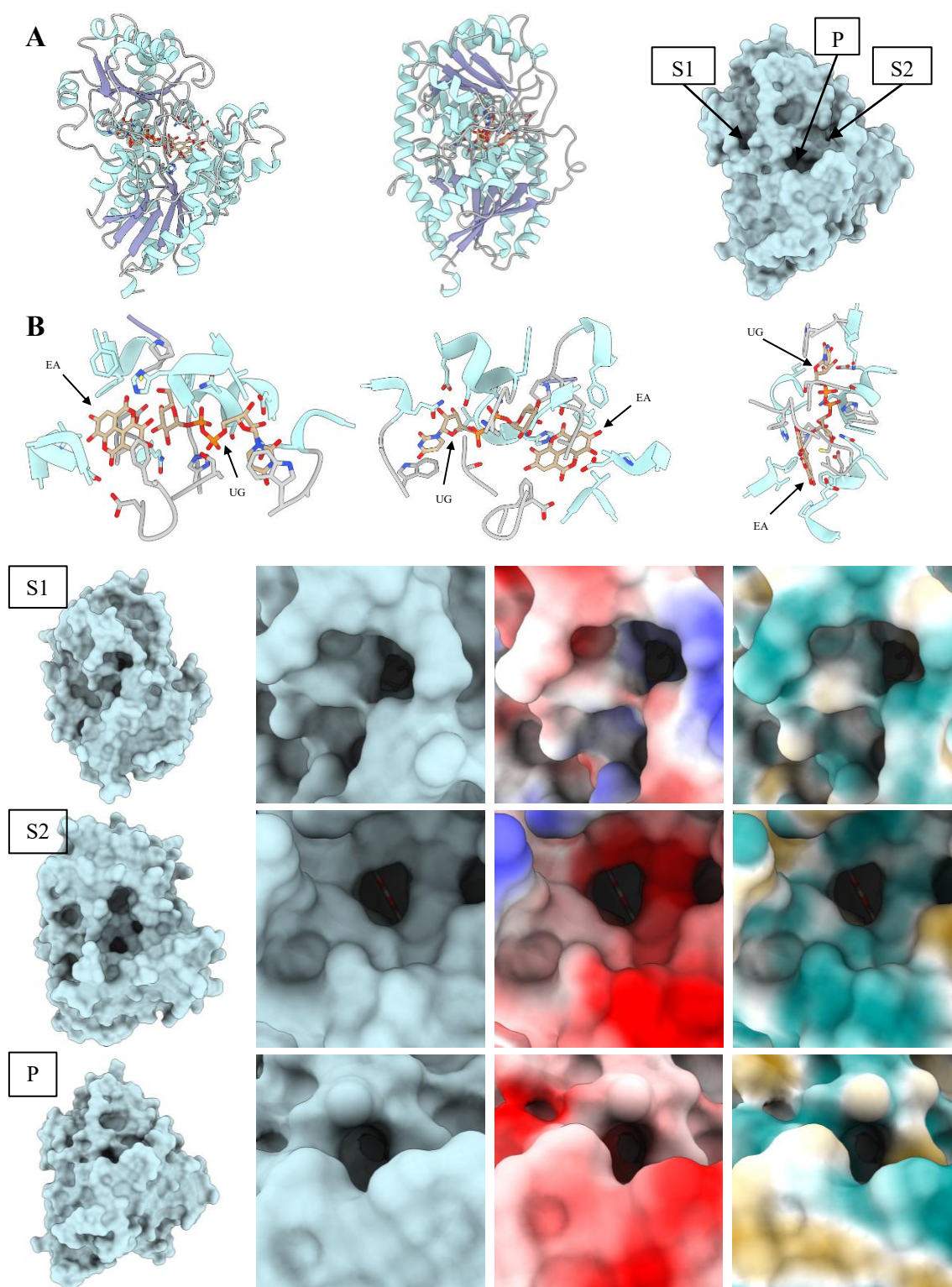


Figure 58. FaUDPSS protein prediction with ellagic acid and UDP-glucose as putative substrates. (A) General view of FaUDPSS colored by secondary structure and protein external surface. In general surface imaging three different pockets are highlighted: S1, S2 and P. (B) Closer views to ellagic acid (EA) and UDP-glucose (UG) and interacting amino acids of FaUDPSS. Each protein pocket is shown in a general and a closer view and electrostatic and hydrophilic surfaces are displayed (with the same scale showed on Fig. 57). Proteins were modelled with AlphaFold3 and align and visualize with ChimeraX.

S2 pocket seemed to be responsible for the entrance of ellagic acid through the catalytic domain. This region was delimited by both polar and non-polar amino acids: Asp183, Asp315, Tyr390, Val84 and Val85. These amino acids defined an area of approximately $0.43 \pm 0.05 \text{ nm}^2$, while the topological polar surface area of ellagic acid was 0.13 nm^2 , suggesting that this cavity is sufficient to accommodate the molecule. The S2 pocket exhibited a predominantly negative Coulombic potential, attributable to the presence of deprotonated carboxylate groups from aspartate residues and the hydroxyl moiety of tyrosine. The presence of polar amino acids conferred hydrophilic properties to the S2 pocket, despite the inclusion of non-polar residues such as valine.

Finally, no molecule was detected within the P pocket, which exhibited a cavity with an estimated area of approximately $0.39 \pm 0.1 \text{ nm}^2$. This pocket is delineated by five amino acids: Gly16, Glu51, Phe318, Phe319, and Ser48. Coulombic surface coloring revealed an electronegative potential, likely attributable to the negatively charged side chain of glutamate and the hydroxyl group of serine. The presence of this electronegative potential suggested a hydrophilic environment near the P pocket, despite the non-polar nature of phenylalanine and glycine, which did not contribute to the pocket's hydrophilicity. Although we did not find any evidence of a molecule related to this P pocket, some molecules must be involved in the catalytic mechanism of UDP-glycosylation.

UGTs related to FaUDPSS were selected based on phylogenetic proximity (**Fig. 56**) to investigate the three-dimensional architecture of these proteins. For this purpose, we selected FaUDP, FaUGT1, FvUGT2, and the *A. thaliana* representative AtUGT72B1. We modeled them with AlphaFold3 and compared them structurally and the results were summarized on **Fig. 59**.

Tridimensional structures for selected proteins revealed a common structure as described before: globular structure with two β -sheets regions and mostly amino acids disposed of α -helixes. This indicates a possible conserved structure for UGTs implicated in the reaction mechanism of the enzyme. A detailed amino acid sequences alignment was built on **Annex 20**.

The FaUDP protein exhibited nearly 100% identity with FaUDPSS, with the only difference being a change in the stop codon and the addition of few more amino acids to the *FxaC_3g11070* sequence. Almost identical tridimensional structures were observed in the case of FaUGT1 and FvUGT2, which corresponded to *FxaC_1g15371* and *FvH4_1g14013*, respectively. The representative of *A. thaliana* from E-clade was AtUGT72B1, which matched almost perfectly with the structure of FaUDPSS. These comparisons reinforce the idea of a common shared tridimensional structure appropriate for glycosylation reactions with high substrate promiscuity for the enzymes.

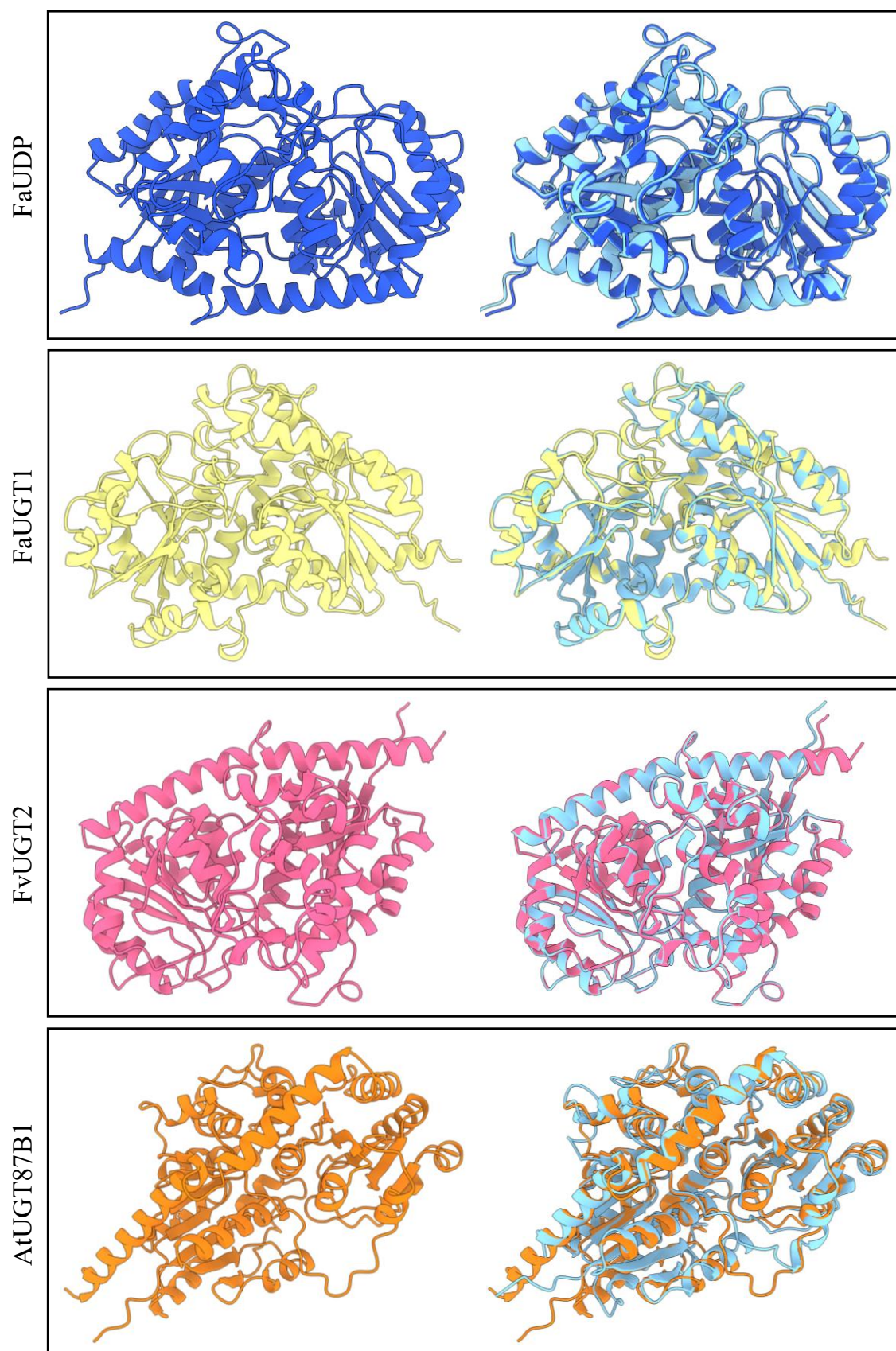


Figure 59. Structural comparisons of UGTs proteins. For each protein, on the left their respective protein structure prediction by AlphaFold3 and on the right the superposition with the FaUDPSS protein (cyan). FaUDP refers to the protein codified by the gene *FxaC_3g11070* obtained from 'Camarosa' reference genome (Egder *et al.*, 2019)

8. Subcellular localization of FaUDPSS protein

Confocal microscopy was applied to discover the subcellular localization of the FaUDPSS protein using *N. benthamiana* leaves and using GFP-FaUDPSS fusion protein and free GFP as a control. Tridimensional structure of GFP-FaUDPSS can be consulted on **Annexes 9** and **12**.

Confocal images obtained from FaUDPSS were presented on **Fig. 60**. Also, plasmolysis assay with 1 M NaCl for 10 min was performed to discard membrane localization of the candidate protein.

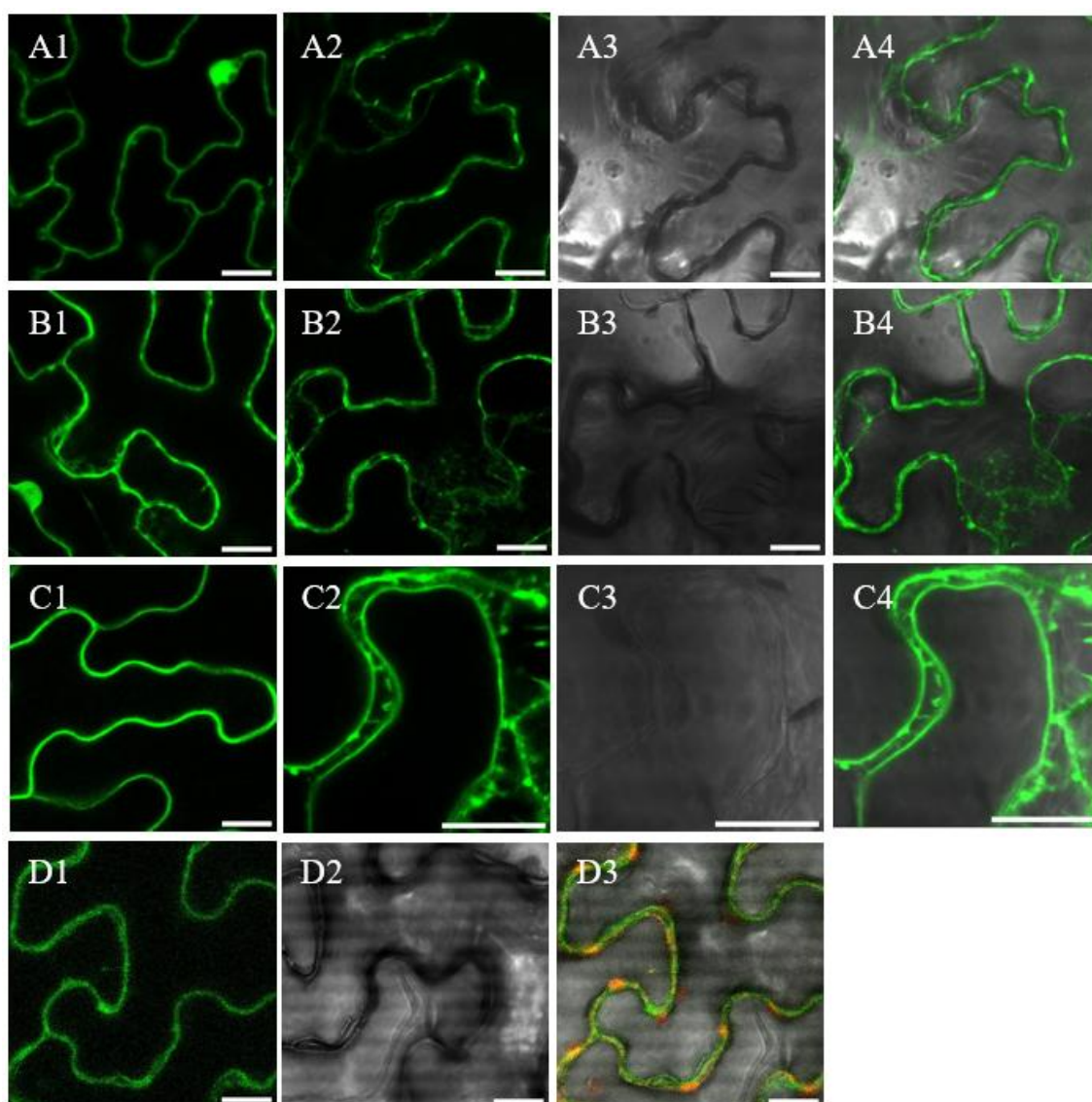


Figure 60. Subcellular localization and plasmolysis assay of FaUDPSS protein. (A) free GFP; (B) GFP-FaUDPSS; (C) AtBAM1; (D) Co-infiltration of FaUDPSS-GFP with FaFAH1-RFP. A/B/C/1 comprise dark field photos in normal conditions while A/B/C/2-4 show dark field, bright field and merge, respectively of each protein with 1 M NaCl for 10 min. **D1-3** show dark field, bright field and merge of the co-infiltration FaUDPSS and FaFAH1. White scale bar represents 20 μm .

The protein FaUDPSS was fused to GFP by the amino-terminal domain of the protein in the vector pGWB6 (**Table 2**). As we could observe, free GFP was spread across the cytoplasm, but it also accumulated in nuclei as could be seen in **Fig. 60A1**. With NaCl treatment, no change in subcellular localization was produced and no protein strands were observed. In the case of FaUDPSS, a similar localization to free GFP was observed and no proteins strands were detected, discarding the membrane localization for the candidate protein. AtBAM1 was used as positive control for plasmolysis assay due to the presence of protein strands with the retraction of the membrane observed in **Fig. 60C2-4**. The co-infiltration of GFP-FaUDPSS and FaFAH1-RFP allowed to see the protein fused to GFP in the whole cytoplasm when the RFP was associated to the endoplasmic reticulum.

A western blot for GFP was performed to confirm that no signal from the confocal was due to the degradation of fusion proteins. These results were presented on **Fig. 61**.

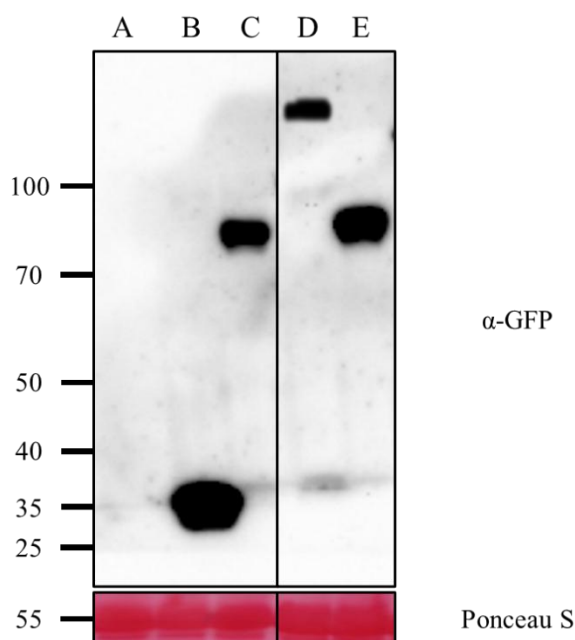


Figure 61. Western blot of GFP for *N. benthamiana* leaves used in confocal microscopy of FaUDPSS protein. A: naïve (untreated sample), B: free GFP, C: FaUDPSS-GFP, D: AtBAM1; E: co-infiltration of FaUDPSS and FaFAH1; Ponceau S: Ponceau staining. Protein sizes are represented in kDa.

Free GFP was detected at approximately 26.77 kDa while FaUDPSS possessed 80.45 kDa, which can be observed in C and F. The plasma membrane protein, AtBAM1, was detected at around 138.00 kDa, corresponding to the larger protein detected in the western blot.

The next step was to generate the TurboID (TID) fusion protein for future experiments. Also, we generated a construct of this fusion protein with GFP tag. With this protein TID-GFP-FaUDPSS we used confocal microscopy to check if this fusion protein presented the same subcellular localization as GFP-FaUDPSS.

9. TurboID proximity labelling of proteins putatively interacting with FaUDPSS

Firstly, the subcellular localization of the fusion protein TID-GFP-FaUDPSS was studied to check if the presence of TID changed the accumulation of the candidate protein. Confocal images of TID-GFP-FaUDPSS were presented on **Fig. 62**.

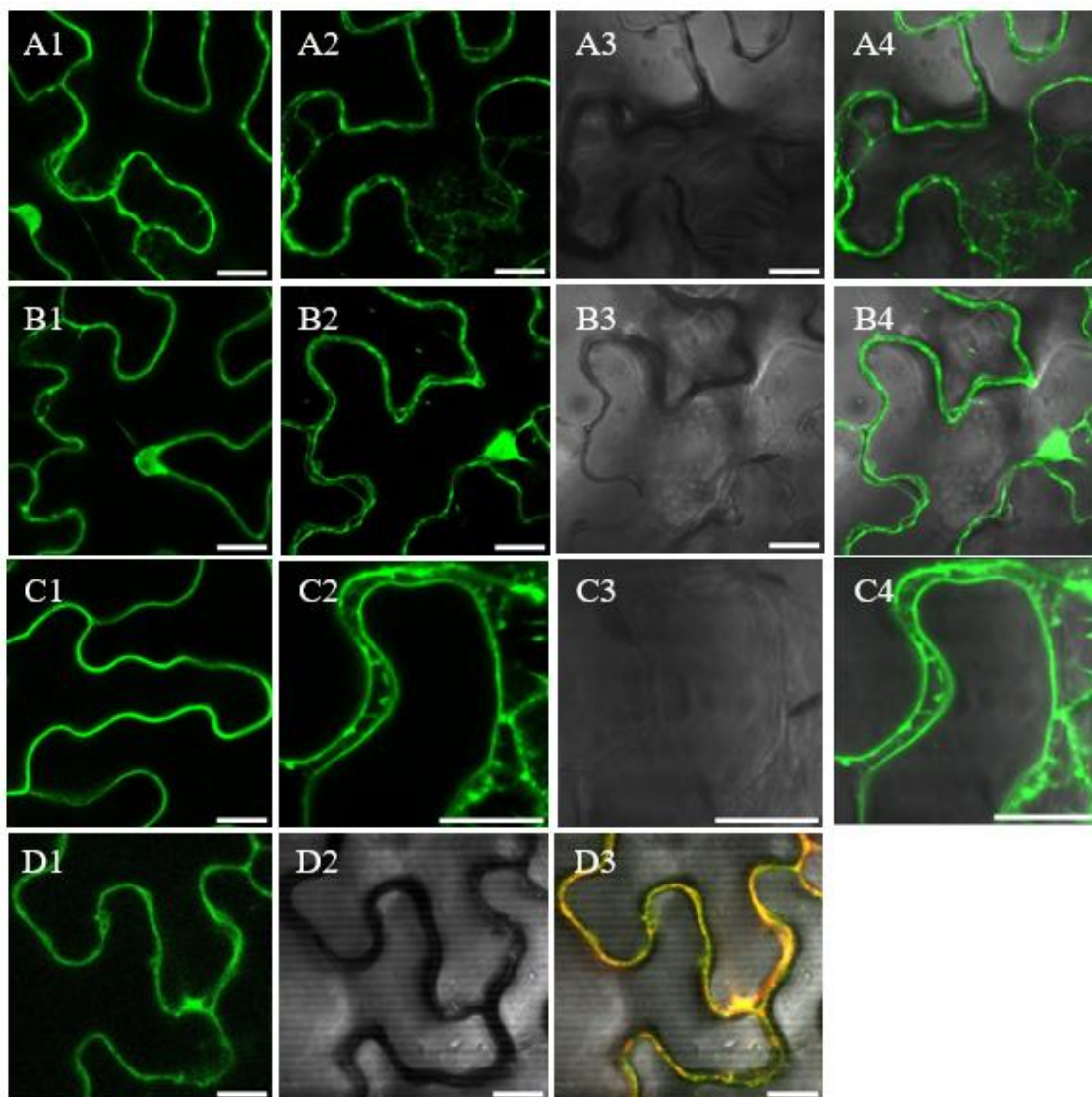


Figure 62. Subcellular localization and plasmolysis assay of TurboID-FaUDPSS fusion protein. (A) GFP-FaUDPSS; (B) TID-GFP-FaUDPSS; (C) AtBAM1; (D) Co-infiltration of TID-GFP-FaUDPSS with FaFAH1-RFP. **A/B/C/1** comprise dark field photos in normal conditions while **A/B/C/2-4** show dark field, bright field and merge, respectively of each protein with 1 M NaCl for 10 min. **D1-3** show dark field, bright field and merge of the co-infiltration TID-GFP-FaUDPSS and FaFAH1-RFP. White scale bar represents 20 μ m.

Interestingly, TID-GFP-FaUDPSS was located in the cytoplasm but also it accumulated in the nuclei. The co-infiltration with FaFAH1-RFP showed the endoplasmic reticulum

in red and the fusion protein spread across the cytoplasm. Confocal images of GFP-FaUDPSS and TID-GFP-FaUDPSS were similar even in signal intensity, suggesting the same nucleus-cytoplasm subcellular localization.

AtBAM1 was again used as control of plasmolysis assay to check the protein strands produced when cellular membranes were retracted. In the case of TID-fusion, the protein of interest didn't show any protein strand when treated with 1 M NaCl for 10 min, indicating no localization in plasma membrane.

An immunodetection using an α -GFP antibody was performed to discard the possibility of protein degradation in confocal assay from **Fig. 62**. The results of this western blot analysis was presented on **Fig. 63**.

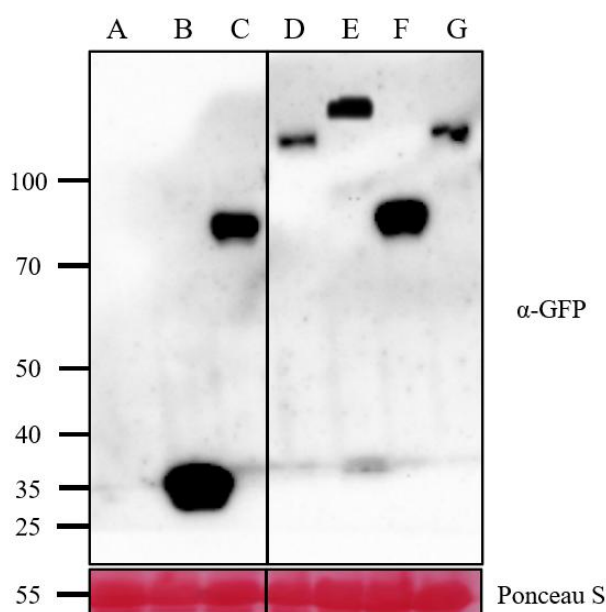


Figure 63. Western blot of GFP for *N. benthamiana* leaves used in confocal microscopy of FaUDPSS protein. A: naïve (untreated sample), B: free GFP, C: FaUDPSS-GFP, D: TurboID-GFP-FaUDPSS; E: AtBAM1; F: co-infiltration of FaUDPSS and FaFAH1; G: co-infiltration of TurboID-GFP-FaUDPSS with FaFAH1. Ponceau S: Ponceau staining. Protein sizes are represented in kDa.

From the western blot analysis, we did not detect any degradation of the fusion proteins used in the confocal microscopy analysis (**Fig. 62**). In fact, we detected the same protein sizes for free GFP, GFP-FaUDPSS and AtBAM1 (**Fig. 61**).

The new fusion protein TID-GFP-FaUDPSS was detected at around 120 kDa, corresponding to the sum of the three individual proteins TID (38.49 kDa), GFP (26.77 kDa) and FaUDPSS (52.60 kDa). The same consistent molecular mass was detected when leaves were co-infiltrated with TID-GFP-FaUDPSS and FaFAH1.

To test if the fusion protein TID-FaUDPSS had biotin ligase activity, we performed an assay in *N. benthamiana* leaves using biotin using constructs FaUDPSS-TID1 and empty vector pGWB-TID1 as control.

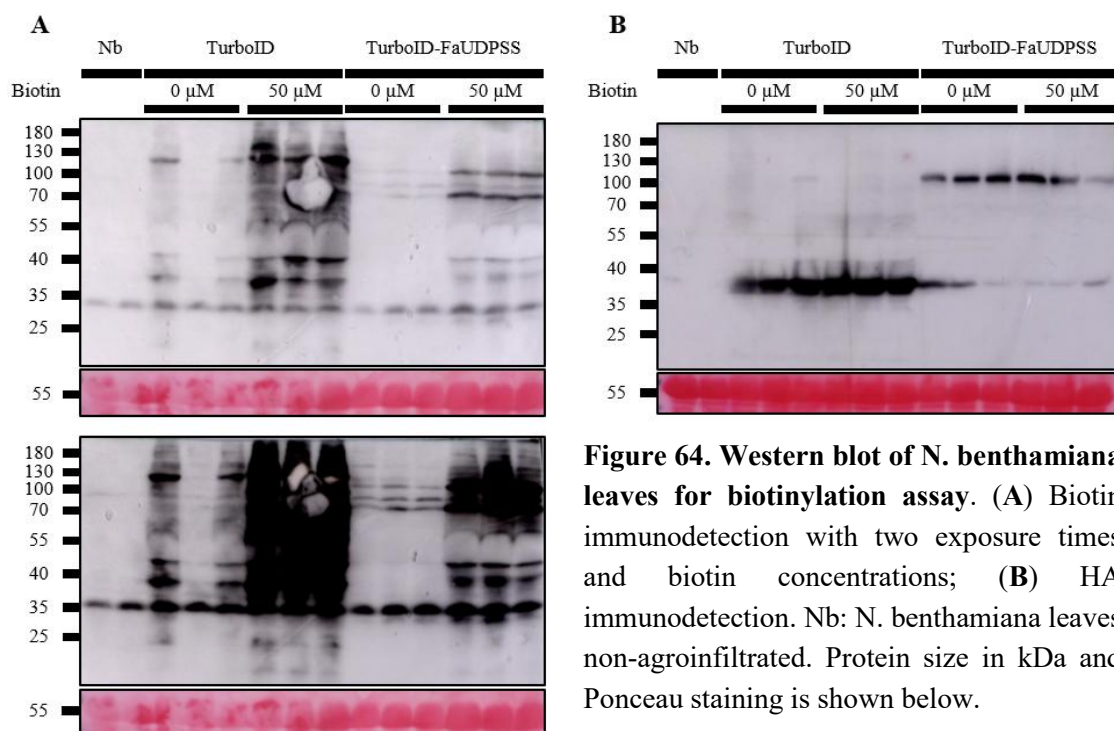


Figure 64. Western blot of *N. benthamiana* leaves for biotinylation assay. (A) Biotin immunodetection with two exposure times and biotin concentrations; (B) HA immunodetection. Nb: *N. benthamiana* leaves non-agroinfiltrated. Protein size in kDa and Ponceau staining is shown below.

Non-agroinfiltrated *N. benthamiana* leaves did not show any signal either in biotin or in HA immunodetection. **Fig. 64B** showed signal for free TID protein (38.49 kDa) and for the fusion protein TID-FaUDPSS (92.06 kDa) indicating the presence of both proteins in the samples, with no difference between presence and absence of biotin. **Fig. 64A** showed that in the presence of 50 μM of biotin both free TID and TID-FaUDPSS were able to biotinylate proteins in their proximity. This became more prominent when we increased the exposure time with the revealing reagent.

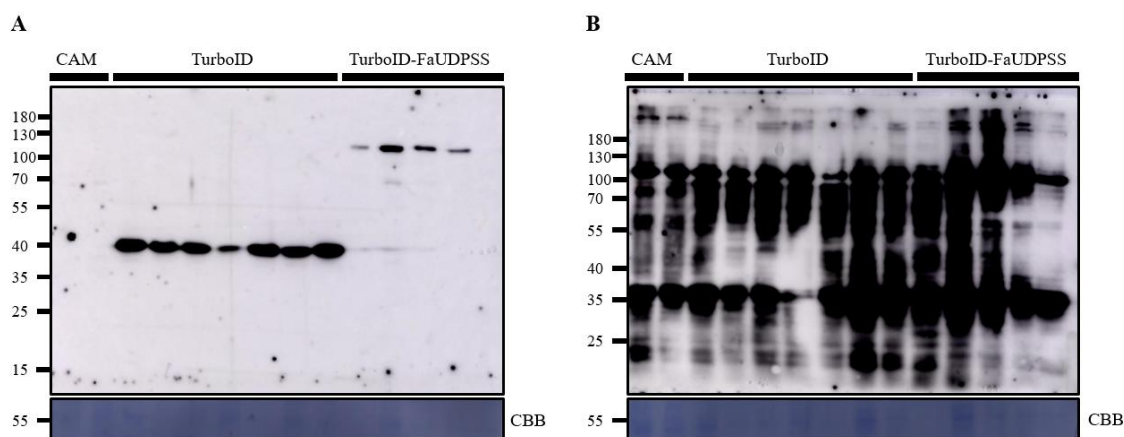


Figure 65. Western blot of 'Camarosa' fruits. (A) Immunodetection with α -HA; (B) Biotin immunodetection with streptavidin-HRP. CAM: non-agroinfiltrated 'Camarosa' fruits; CCB: Coomassie Brilliant Blue staining. Protein sizes were represented as kDa. Protein concentrations were calculated in **Annex 21**.

For discovering the interactors of FaUDPSS protein, 14 different ‘Camarosa’ fruits were used in three experimental conditions: negative control (non-agroinfiltrated fruits), positive control (agroinfiltration with free TID in pGWB502 vector) and FaUDPSS samples (fruits agroinfiltrated with TID-FaUDPSS based on pGWB-TID1 vector). Total proteins were extracted and used for a western blot with α -HA and streptavidin-HRP to check the presence of the TID and the biotinylation of proteins, respectively. Results of these immunodetections were presented on **Fig. 65**.

Most of the agroinfiltrated fruits were detected with α -HA, indicating the presence of the TID protein, alone at 38.48 kDa or fused to FaUDPSS at 92.06 kDa. CAM samples didn’t present a signal for HA, indicating no contamination of the samples. On the other hand, in **Fig. 65B** there was a pattern of biotinylated proteins in the CAM samples, which indicated a high concentration of biotin in ‘Camarosa’ fruits. Also, a different pattern of biotinylated proteins could be seen between free TID and fusion protein. From these experiments, we selected 3 samples with free TID and 3 with TID-FaUDPSS to make biotinylated proteins enrichment.

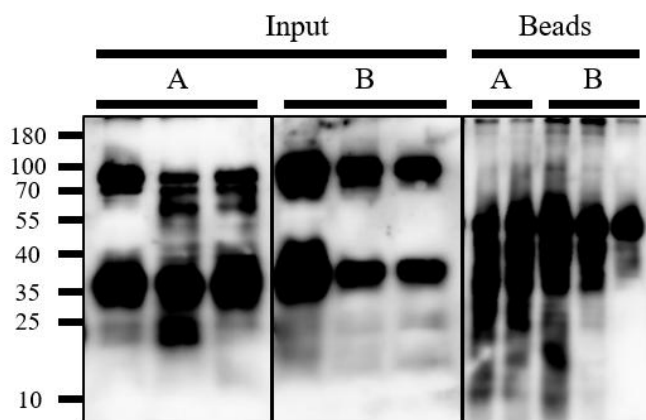


Figure 66. Purification of biotinylated proteins. A: control samples (free TID); B: fruits agroinfiltrated with TID-FaUDPSS fusion protein. Protein sizes were represented as kDa.

The purification of biotinylated proteins was accomplished by using magnetic beads coupled with streptavidin. The western blot for biotin detection before and after purification was presented on **Fig. 66**. The input used in this figure was the protein extraction after the desalting column and there were indeed differences after the purification with the magnetic beads, making an enrichment of biotinylated proteins.

The proteomics analysis revealed some putative interactors of the candidate protein FaUDPSS. The proteins listed in **Table 15** were selected based on the following criteria: they were detected in at least two out of three samples and exhibited a twofold difference between the two experimental conditions.

From 441 detected proteins in mass spectrometry, only the 31 ones listed below were selected as putative interactors, where a vast majority of them were components of ribosomal subunits, being involved in RNA processing. Besides these, some proteins

putatively involved in metabolism were acetyl-CoA carboxylase, oxalyl-CoA synthetase or ATP synthetase, for example.

Table 15. Putative protein interactors of FaUDPSS found by TurboID proximity labelling. Proteins were identified by their protein name with ‘Camarosa’ reference genome (Tianjia *et al.*, 2021). Aa: number of amino acids. Functional annotation was annotated using Mercator4 and Swiss-Prot databases.

Gene name	Aa	Functional annotation
<i>FxaC_1g31961.tl</i>	156	Component eS31 of small ribosomal-subunit (SSU) proteome
<i>FxaC_2g12301.tl</i>	367	Flavoprotein component of succinate dehydrogenase complex (SDH1)
<i>FxaC_4g14710.tl</i>	88	Component eL13 of large ribosomal-subunit (LSU) proteome
<i>FxaC_5g10091.tl</i>	255	Molecular co-chaperone (P23)
<i>FxaC_8g33400.tl</i>	568	β -subunit of methylcrotonoyl-CoA carboxylase complex
<i>FxaC_9g13710.tl</i>	123	Component uL29 of large ribosomal-subunit (LSU) proteome
<i>FxaC_11g04950.tl</i>	181	Component uL5 of large ribosomal-subunit (LSU) proteome
<i>FxaC_12g39160.tl</i>	258	Component uL2 of large ribosomal-subunit (LSU) proteome
<i>FxaC_13g55130.tl</i>	259	Component eS1 of small ribosomal-subunit (SSU) proteome
<i>FxaC_14g23891.tl</i>	204	Component uS9 of small ribosomal-subunit (SSU) proteome
<i>FxaC_17g54370.tl</i>	2,268	Acetyl-CoA carboxylase
<i>FxaC_18g02500.tl</i>	198	Acetyl-CoA carboxylase
<i>FxaC_18g02530.tl</i>	2,268	Acetyl-CoA carboxylase
<i>FxaC_18g44250.tl</i>	524	Oxalyl-CoA synthetase
<i>FxaC_19g15150.tl</i>	551	β -subunit of ATP synthase peripheral MF1 subcomplex
<i>FxaC_19g19470.tl</i>	92	Component eL43 of large ribosomal-subunit (LSU) proteome
<i>FxaC_19g22310.tl</i>	273	RRM-domain containing protein
<i>FxaC_21g27720.tl</i>	149	Component eL28 of large ribosomal-subunit (LSU) proteome
<i>FxaC_21g35330.tl</i>	773	Seed biotin containing protein (SBP65)
<i>FxaC_22g17940.tl</i>	62	Component eS30 of small ribosomal-subunit (SSU) proteome
<i>FxaC_23g14330.tl</i>	162	Component eL24 of large ribosomal-subunit (LSU) proteome
<i>FxaC_23g28590.tl</i>	647	Molecular chaperone (Hsp70-1/2/3/4/5)
<i>FxaC_23g54311.tl</i>	263	14-3-3 phosphoprotein-binding protein (GRF)
<i>FxaC_23g61220.tl</i>	545	RopGEF guanine nucleotide exchange factor (SWAP70)
<i>FxaC_25g44710.tl</i>	572	Peptide N4 N-acetylglucosaminyl asparagine amidase
<i>FxaC_26g16040.tl</i>	1,018	E1 component of 2-oxoglutarate dehydrogenase complex
<i>FxaC_26g18861.tl</i>	331	BAHD-acyltransferase
<i>FxaC_26g28290.tl</i>	148	Nucleoside diphosphate kinase 1
<i>FxaC_27g16432.tl</i>	307	C2-domain containing protein
<i>FxaC_28g04511.tl</i>	117	Component eL36 of large ribosomal-subunit (LSU) proteome

Finally, we checked the number of peptides detected for each protein in the two different experimental conditions. For this purpose, we calculate the peptide ratio (R) which was the difference between peptides in the sample (P_s) and in the control (P_c) divided by the total of peptides (P_T) found for each interactor.

$$R = \frac{P_S - P_C}{P_T}$$

If $R < 0$, more peptides were found in control conditions than in fruits with the overexpression of FaUDPSS and if $R > 0$, then more proteins were found in the FaUDPSS overexpression. Results of this analysis were represented in **Fig. 67**.

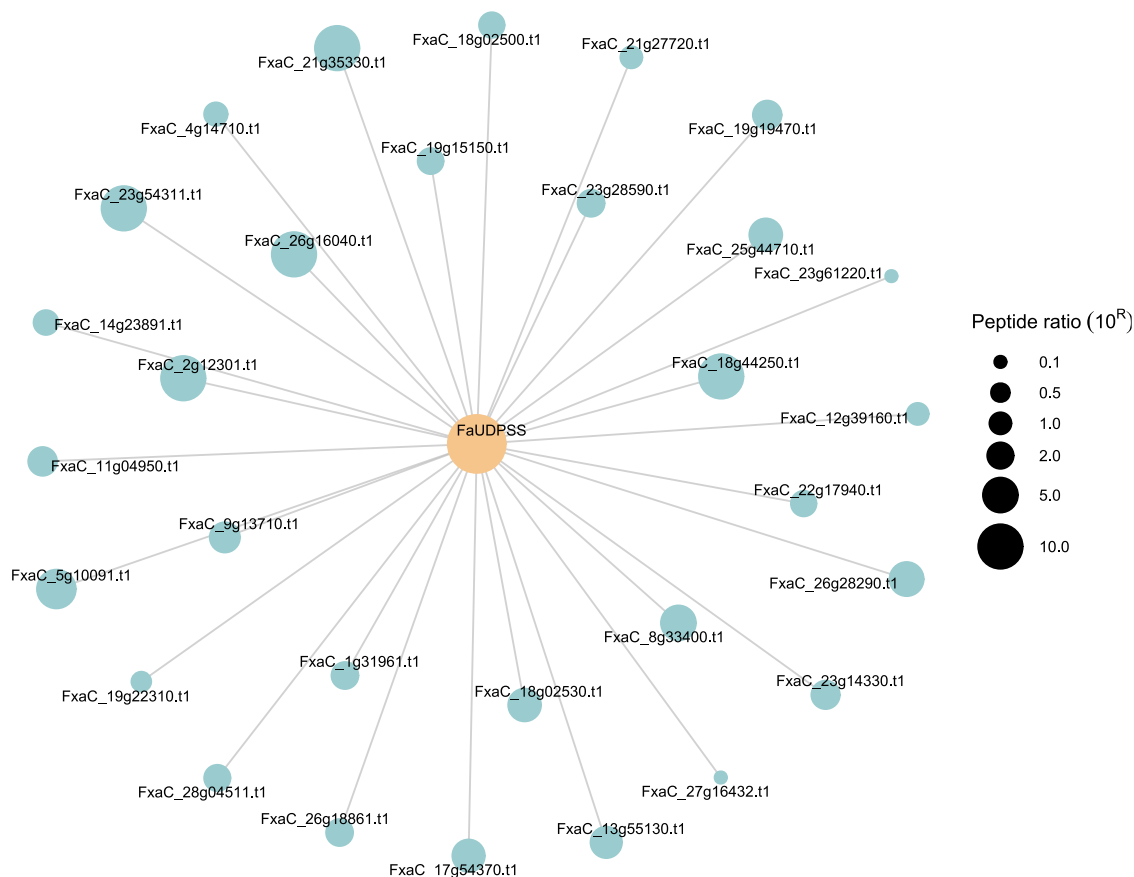


Figure 67. Putative protein interactors of FaUDPSS. Points sizes represent the peptide ratio calculated for each interactor in a scale from 0.1 (PS= 0) to 10 (PC= 0).

From this representation, we could see the representativity of each protein in the proteomics analysis. Some examples of genes that codified proteins with less peptides found in the control were *FxaC_2g12301.t1*, *FxaC_5g10091.t1*, *FxaC_18g44250.t1*, *FxaC_21g35330.t1* or *FxaC_26g16040.t1*. The opposite situation held genes that codified proteins like *FxaC_1g31961.t1*, *FxaC19g22310.t1* or *FxaC_27g16432.t1*.

The expression of the genes putatively interacting with FaUDPSS was presented in **Fig. 68**. Firstly, no significant differences can be observed among either strawberry varieties or countries. Almost all genes presented low levels of expression below 200 TPM, except for *FxaC_1g31961.t1* which possessed the highest expression in the whole set studied. In addition, *FxaC_23g28590.t1* showed relatively high expression values compared to the

rest of the genes selected. Consistently, in the case of the GBP the highest expression was shown by *FxaC_1g31961.t1* and *FxaC_23g28590.t1*, respectively.

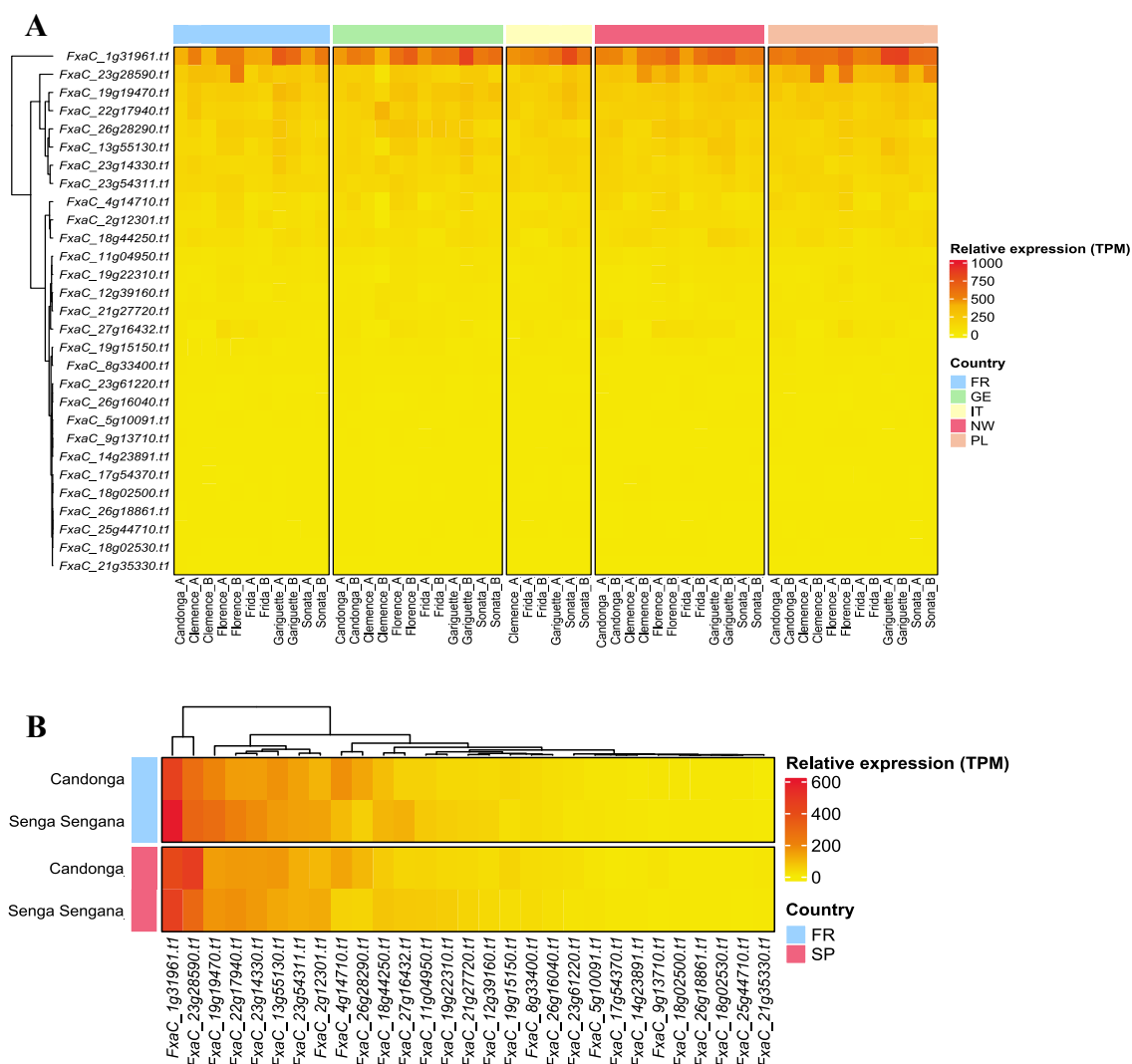


Figure 68. Relative expression in TPM of the putative interactors of FaUDPSS. (A) Candidates' expression in different commercial strawberries and different locations (FR: France, GE: Germany, IT: Italy, NW: Norway, PL: Poland). In this varieties, A and B represents two different harvesting times. **(B)** Expression in parentals from GBP: 'Candonga' and 'Senga Sengana' varieties.

Expression of some selected genes from **Fig. 68A** can be consulted in **Annex 22** for further information.

Discussion

1. Ellagitannins in ‘232’ x ‘1392’ mapping population and RNA-seq analysis

In this study, the levels of ellagitannins associated QTLs were analyzed to investigate the genetic basis underlying their accumulation. Among the simplest ellagitannins, only ellagic acid hexose 1 (EAH1) and ellagic acid hexose 2 (EAH2) showed statistically significant differences between the parental lines. Based on these findings, EAH2 was selected for further transcriptomic analysis using RNA-seq on lines with contrasting levels of this metabolite.

From the QTL mapping, only the locus qEAhex1 was considered for in-depth analysis, as it was consistently detected in the 2013 dataset, a year when significant differences in EAH2 levels observed between the parental lines. Although qEAhex1 and qEAhex2 overlap within the same genomic region, suggesting a broader QTL relevant to ellagitannin content, only qEAhex1 was prioritized. Candidate genes selection within this region was initially based on statistical association; however, functional annotation was also taken into account during the final decision. Genes encoding putative anthocyanidin 3-*O*-glucosyltransferase were prioritized due to their potential involvement in the ellagitannins biosynthesis pathway. One of the strongest candidates, *FvH4_1g14013*, was found to be unexpressed in lines with low ellagitannins and high flavan-3-ols, possible indicating a redirection of metabolic flux (Mora *et al.*, 2022). A similar expression pattern was observed for *FxaC_3g11070*, leading to the hypothesis that this gene's activity or lack thereof, might influence the accumulation of EAH2, especially considering the relatively low abundance of this metabolite in our samples.

Although *FxaC_1g15670* was not initially selected, its elevated expression levels suggest it could still play a role in ellagitannin biosynthesis. In this context, *FvH4_1g13990* was excluded due to its lower fold change compared to *FvH4_1g14013*, although we cannot entirely rule out the possibility that its *F. × ananassa* homologs may contribute to increased EAH2 concentration.

2. Direct correlation between *FaUDPSS* and EAH2 content

We investigated the candidate expression in another population different from the mapping population. We used the Goodberry population (used in **Chapter I**) to observe that in ‘Camarosa’ parental there was no expression of *FxaC_3g11070* while ‘Senga Sengana’ was chosen as a possible source of the gene sequence. These results were consistent in both locations studied in this case: Spain and France.

Due to the contrasting expression of both parentals (‘Camarosa’ and ‘Senga Sengana’), the offspring presented a high diversity in *FxaC_3g11070* gene expression, allowing us to choose individuals with high and low gene expression. Secondary metabolite extraction

showed a positive correlation between *FxaC_3g11070* expression and EAH2 accumulation. The group with low gene expression presented overall significant low EAH2 concentration and high gene expression group resulted in a greater EAH2 concentration. Nevertheless, we didn't observe any change in flavan-3-ols when our hypothesis was that an increase in ellagitannins would probably reflect a depletion of proanthocyanidins or their monomers.

We performed a transient overexpression assay through agroinfiltration of 'Candonga' fruits. In this experiment, fruits were infiltrated with the construct *35SCaMV::FaUDPSS* and an empty vector to check if the overexpression of the candidate gene was able to raise the concentration of EAH2. 'Candonga' variety was selected as genetic background for not having *FxaC_3g11070* expression and low amount of EAH2, making possible to check if the overexpression of this candidate gene produces an accumulation of EAH2.

In this same scenario, we performed stable overexpression in 'Candonga' genetic background of *FxaC_3g11070* under the control of a *35SCaMV* promoter. With this stable transformation we would be able to see the contribution of the candidate gene to EAH2 concentration in a variety where the concentration of this ellagitannin is very low and the expression of the gene is absent.

3. *FxaC_3g11070* sequence differs from the 'Camarosa' reference genome sequence

According to our data, *FxaC_3g11070* doesn't express in 'Camarosa' variety of octoploid strawberry. This was checked using cDNA from leaves and from red fruit of 'Camarosa', 'Candonga' and 'Senga Sengana', where only cDNA from 'Senga Sengana' was able to generate a PCR-amplification (data not shown). Nevertheless, *F. vesca* homolog *FvH4_1g14013* has higher expression in root tissue. Given the fact that ellagitannins functions are poorly understood, the expression of a gene responsible for the concentration of some ellagitannins in roots remains unknown. But one hypothesis could be the use of this compounds for chelating Fe ions in the soil. Ellagitannins with HDDP unities, like agrimoniin can efficiently chelates Fe atoms (Moilanen *et al.*, 2016). Another hypothesis is that these compounds participate in a crosstalk with pathogen or beneficial microorganisms from the rhizosphere to either establish protection against harmful microorganisms or to enhance the fitness of the roots.

The amplification of the *FxaC_3g11070* coding sequence revealed discrepancies with the established 'Camarosa' reference genome (Edger *et al.*, 2019). The F₁ population derived from the crossing of '232' and '1392' showed diverse genetic sequences, like an insertion of a 141 bp cataloged as intron in the parentals' gDNA. Among the F₁ population only one individual studied showed a gene sequence similar to the one annotated in the reference genome but with a few more amino acids long. This sequence was selected as

the one representative for the gene giving the fact that it was similar in F₁ individuals from Goodberry population, which gives us consistent evidence that even in two different populations, the gene sequence was conserved. Nevertheless, in the case of the GBP, the offspring should have inherited the sequence from the ‘Senga Sengana’ parental because ‘Camarosa’ didn’t show any expression while the F₁ individuals did. In the case of the mapping population for ellagitannins, the *FxaC_3g11070*-expressing individuals have a broad range of different gene sequences, derived from the breeding of the parentals. Specifically, with our data, ‘Camarosa’ has no expression of this gene which can influence the expression of the ‘1392’ parental considering the breeding history of both parentals: ‘232’ (Sel. 4-43 × Vilanova) and ‘1392’ (Gaviota × ‘Camarosa’) (Zorrilla-Fontanesi *et al.*, 2011). *FxaC_3g11070* gene expression of Vilanova and Gaviota “grand parentals” remains unknown, and this could be the reason why some F₁ individuals hasn’t any expression or the gene sequence has high diversity even in the same offspring.

4. Strawberry commercial varieties and their relationship with EAH2 content

Once we observe that *FxaC_3g11070* was expressing or not depending on the strawberry variety studied, we focus on studying its expression in six commercial varieties bred in four different locations (France, Germany, Norway and Poland). Data showed that only Florence, Gariguet and Sonata have *FxaC_3g11070* expression and TPM values were similar in the four locations. Despite the different geographical and climatic conditions of each location, the expression of the candidate gene was similar, suggesting a low environmental influence. It was interesting that the expression of the gene didn’t change too much with the different harvesting times (A and B) and maybe this supposition is supported in the fact that probably these two harvesting times were selected during red ripe stage. We also checked the expression of the close homolog *FxaC_1g15671* whose TPM values were higher. In fact, all varieties expressed this gene in contrast to what happens with our candidate *FxaC_3g11070*.

Then we proceeded with the secondary metabolite extraction and the focus on EAH2. In general, varieties with no gene expression showed low levels of EAH2 except for ‘Frida’ variety whose levels were high for a non-*FxaC_3g11070*-expressing individual. The case of Gariguet is also intriguing due to its relatively high expression of the gene and low amount of EAH2. These results should be taken as preliminary results given the fact that only varieties from one location were studied. If our hypothesis is verified and the *FxaC_3g11070* is “variety-specific” we can use the expression of this gene together with the EAH2 concentration as a breeding marker for future projects. This makes sense as ellagic acid is very important in strawberry fruit (Ariza *et al.*, 2018) but maybe EAH2 is more specific within the same species, providing a way of distinguishing varieties regarding the ellagitannins content and antioxidative properties of these strawberry varieties.

5. FaUDPSS is a very conserved cytoplasmic protein

Phylogenetic analysis of UGTs revealed a case of continuous phylogenetic evolution from the same ancestor. We considered clade A as the phylogenetically more ancient and clade L the most recent, respectively, of all clades.

The protein of interest cloned from ‘Senga Sengana’ background (FaUDPSS) is comprised within clade E. Reference ‘Camarosa’ gene sequence of *FxaC_3g11070* (FaUDP) is closely related to FaUGT4 but also close to FaUDPSS. In addition, FaUDPSS is very close to FaUGT1 and FvUGT2, which was reinforced with the structural superposition of each protein structure. According to our data, structural comparisons of several UGTs resulted in a very conserved protein structure, with minor displacements between proteins. Despite the broad range of substrates subtle for glycosylation, all the proteins shared the same structural domains used for the recognition of substrates (Gharabli *et al.*, 2023). Our hypothesis is that the catalytic domains are very conserved for recognizing the UDP-activated hexoses and that the presence in different positions of several key amino acids are determinant for discriminating against different types of hexoses, as has been proposed in bibliography (Lethe *et al.*, 2024). Nevertheless, the binding pocket would have responsibility for the recognition of a plethora of possible substrates, but the amino acid implicated in this recognition should be able to interact with hydroxyl groups, which are a very common functional group found in flavonoids and tannins (Mora *et al.*, 2022).

The subcellular localization of most UGT was predicted to be cytoplasmic which matches the structural visualization of the protein. The protein FaUDPSS presents a high Coulombic potential, and a tendency of hydrophilic amino acids exposed to the medium. Similar properties should be displayed by the rest of the UGTs and together with the absence of neither signal peptide nor transmembrane domains, making UGTs suitable for being dispersed in the cytoplasm.

The subcellular localization was checked experimentally by the fusion protein GFP-FaUDPSS. The GFP tag was fused in the N-terminal domain of FaUDPSS due to the presence of the catalytic domain in the C-terminus, avoiding the dysfunctionality of the protein by the tridimensional alteration of the protein of interest. FaUDPSS is widespread in cytoplasm according to our assays, but some of the protein is condensed in the nucleus. Our hypothesis is that maybe this protein is found in the nucleus due to permeability across nucleus pores thanks to the reduced size of the fusion protein (80.45 kDa). Also, the plasmolysis revealed that FaUDPSS is not associated with membranes, which is consistent with the fact that this protein didn’t possess any transmembrane domains and no larger regions with hydrophobic amino acids exposed to the external medium. In addition, the synthesis of most flavonoids and tannins takes place in the cytoplasm (Mora

et al., 2022), so if the candidate participates in any of these routes should be present in the same cellular compartment.

6. Proteins interacting with FaUDPSS

Following the same strategy, we fused the TID protein to the N-terminus of the protein FaUDPSS. Firstly, we locate the GFP signal in the cytoplasm which indicates that the fusion with the TID didn't change the subcellular localization.

We performed the agroinfiltration of 'Camarosa' fruits for TID experiments following a similar approach than the one used recently in the same species (Martín-Pizarro *et al.*, 2025). Immunodetection of the fusion protein TID-FaUDPSS was performed against HA epitome, resulting in several fruits overexpressing either TID or TID-FaUDPSS. With the biotin detection we discovered that 'Camarosa' fruits presented high number of biotinylated proteins by detection with streptavidin-HRP. This high number of biotinylated proteins may be caused by the high concentration of biotin in strawberry fruits that reach about 15 ng/g in fresh fruit (Staggs *et al.*, 2004). Even considering this fact, the biotinylating pattern was different between free TID and the fusion protein in fruits. Nevertheless, we needed to prove that the fusion protein TID-FaUDPSS was able to biotinylate the proteins of its surroundings, and this was performed using *N. benthamiana* leaves. Indeed, we saw that the fusion protein was able to add this modification to certain proteins generating a band pattern with 50 μ M biotin. During the purification of biotinylated proteins, one aliquot from the free TID samples after beads were lost, so we made the detection of biotin with five samples instead of six after the incubation with the bead conjugated with streptavidin.

Although no significant signals were detected in between experimental groups, we thought that this could be produced by the low concentration of proteins in the samples (see **Annex 19**). However, we checked the number of peptides found in the control and the FaUDPSS conditions to look for proteins with a FC of twice whose peptides were detected in at least two of the three replicates, as described previously. According to Mercator4, some putative interactors of FaUDPSS belong to ribosome organization, which can be understood under the fact that the transcription and translation machinery of the cells are activated during the transient transformation. Interestingly, we found some interactors that participate in primary metabolism like acetyl-CoA carboxylase or succinate dehydrogenase. Interestingly, we found among the list a BAHD-acyltransferase, the same enzyme detected in our group for being responsible for the accumulation of another ellagitannins. However, the possible interactions between these candidates with FaUDPSS remain unknown and even we modelled some interactions between proteins with AlphaFold3, but with poor results. This approximation would not be the most accurate one because the biosynthesis of ellagitannins is poorly understood

and we don't know the stoichiometry of this protein and the reaction mechanism below the synthesis of these tannins. We also checked the expression of each putative interactor in different varieties and different populations, and we saw that peptides related to ribosome and RNA process presented the highest expression in the whole set, but some putative interactors related to primary metabolism also presented high expression levels.

7. Future perspectives and proposed experiments

With our current data, we are designing new experiments to characterize more deeply biochemically and functionally the *FxaC_3g11070* gene and the protein codified by this very gene.

Our current efforts are directed to analyze the transient overexpression in 'Candonga' fruits. Firstly, we will check the expression of *FxaC_3g11070* by RT-qPCR with specific primers and then we will extract primary and secondary metabolites. We would expect an increase in the EAH2 produced by the overexpression of *FxaC_3g11070*, but we cannot discard the possibility of the increase of other isomers, as well as a decrease in the concentration of flavan-3-ols and proanthocyanidins. We could also see if this overexpression affects the primary metabolite regarding shikimate pathway intermediates (Mora *et al.*, 2022). Following this same strategy, we will also check the *FxaC_3g11070* expression in our stable overexpression lines and prove that the results provided by the transient overexpression were consistent.

Furthermore, we can drive phenotype analysis regarding fruit production, sugar content, Brix grade, color and more. Also, one experiment we would like to conduct is to isolate EAH2 using LC/MS and fraction collectors. We would select the m/z peak corresponding to EAH2 by retention time and then we could send it to nuclear magnetic resonance (¹H-RMN). With this experiment we could elucidate the chemical identity of EAH2.

Furthermore, with transient and stable overexpression of the gene we want to test the antioxidant capacity of the fruits following Trolox equivalent antioxidant capacity (TEAC) assay. We could see the contribution of EAH2 to the total antioxidant capacity of the strawberry fruit.

Regarding the *FxaC_3g11070* gene expression in strawberry varieties, we want to check the EAH2 content in all the four locations of the study and try to relate this content to the gene expression. Our main goal is to test if the *FxaC_3g11070* expression follows a tendency with the EAH2 content in all varieties and check if it's consistent among several more strawberry varieties. However, nowadays we are focusing our attention on cloning the promoter region of *FxaC_3g11070* in 'Camarosa', 'Candonga' and 'Senga Sengana' varieties to check if these promoter regions differ from one variety to another. Our hypothesis is that maybe 'Candonga' and 'Camarosa' promoter sequences contain

changes compared to ‘Senga Sengana’ promoter (insertions, deletions, transposons, etc.), resulting in a non-expression of the candidate gene.

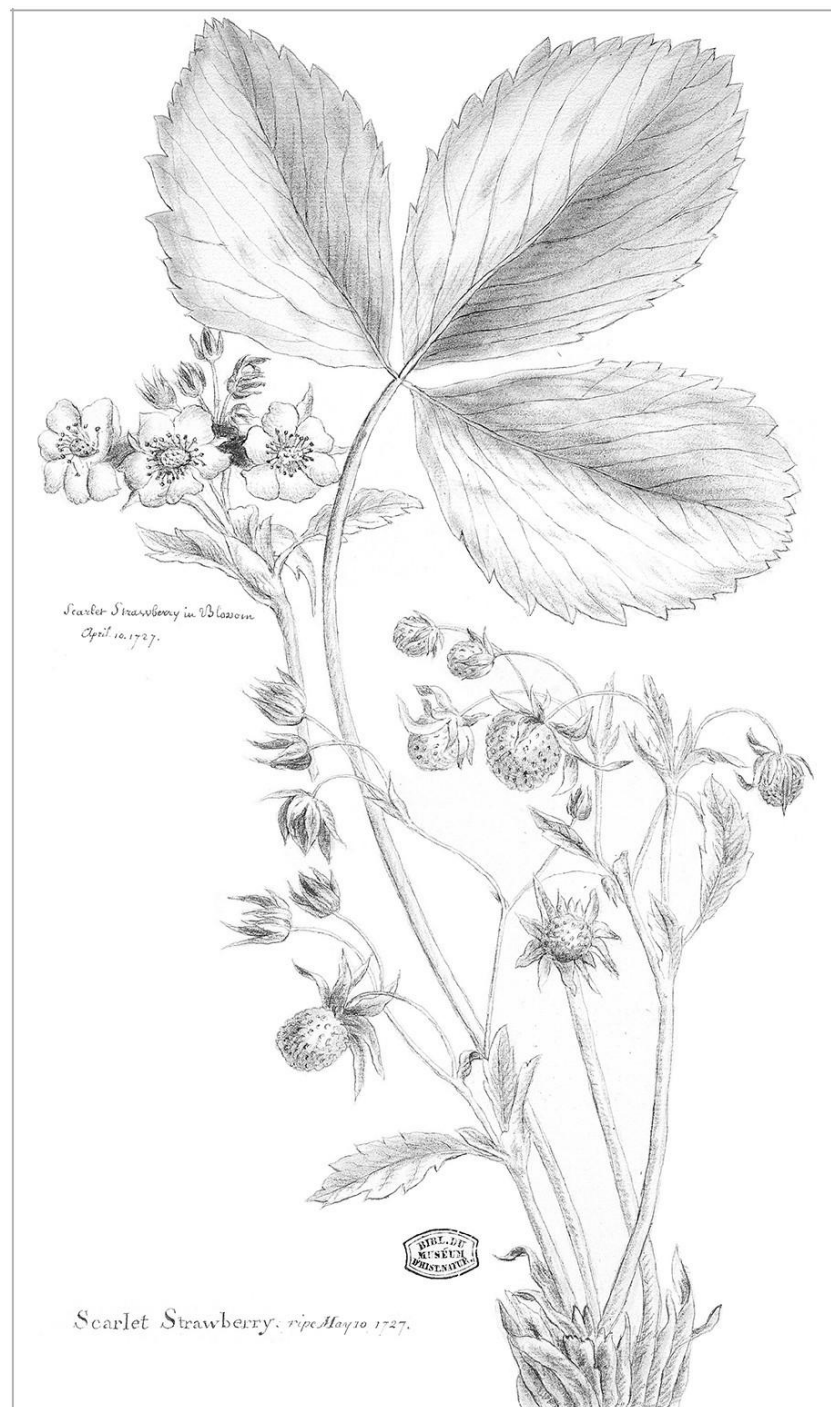
The putative interactors of FaUDPSS should be checked and study the *in vitro* interactions between the candidate and other proteins. For this purpose, some experiments can be driven like the yeast-to-hybrid approach with the putative interactors. Also, we could investigate the subcellular localization of the interactors because if the subcellular localization is the same, the probability of the interactions would be higher. Additionally, co-immunoprecipitation of FaUDPSS and some interactors in *N. benthamiana* leaves would inform us if there’s any physical interaction between them.

Finally, the same approach that the one proposed in **Chapter I** can be driven for this candidate: purification of FaUDPSS using GFP-tag. With this purified FaUDPSS we could test enzymatically some substrates like ellagic acid, and some other molecules better narrow the function of the enzyme. This information together with the possible information from the ¹H-RMN would serve for the characterization of the enzyme.

8. Concluding remarks

During this chapter we have explored the implication of a candidate gene in the content of ellagitannins. Specifically, previously to the start of this PhD thesis, a population with contrasting lines in ellagitannins were selected. Using this population one QTL was identified to control around 50-70% of the total content of EAH2. Within this QTL, we found one candidate gene annotated as anthocyanidin 3-*O*-glucosyltransferase that may play a role in EAH2 biosynthesis. The candidate gene showed expression depending on the variety studied: ‘Camarosa’, ‘Candongá’, Clery and Frida showed no expression while Florence, Gariguette, ‘Senga Sengana’ and Sonata presented gene expression. In this context, we discovered that the genome reference sequence of ‘Camarosa’ differs from the real gene sequence that we cloned from ‘Senga Sengana’ variety. We also discovered a tendency between candidate expression and EAH2 accumulation even in different populations and varieties. Finally, we putatively discovered protein interactors of FaUDPSS with different functions ranging from RNA processing to primary metabolism related proteins. Further proposed experiments are required for the complete characterization of the gene and for elucidating the biosynthesis of some ellagitannins.

VII. Conclusions

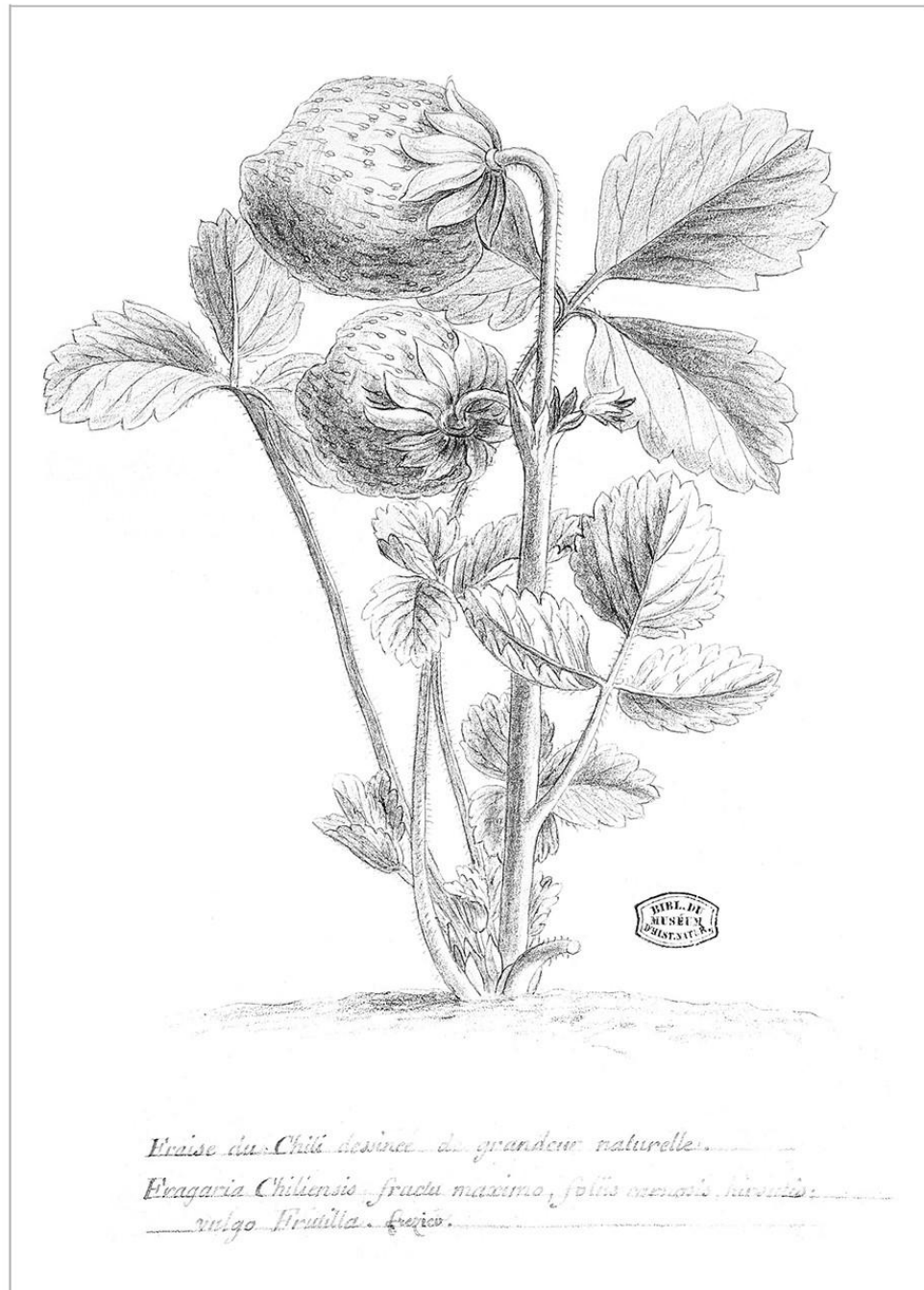


Botanical illustration of *Fragaria virginiana* plant by Antoine Nicolas Duchesne. *Histoire Naturelle des Fraisiers*, 1766.

The main conclusions derived from both chapters of this PhD thesis are listed below:

1. QTL analysis of VOCs in Goodberry population revealed stable QTLs found in more than two different locations, thus establishing new future perspectives for studying VOCs regulation.
2. Female map revealed high number of QTLs within linkage group Fvb3-2 probably revealing a hotspot for VOCs biosynthesis and/or regulation.
3. The transient overexpression of *FaLOX2* drives changes in VOCs content as well as an accumulation of sugars and anthocyanins and other flavonoids. Specifically, γ -dodecalactone and methyl jasmonate were overaccumulated in the transient *FaLOX2*-overexpressing fruits.
4. *FaLOX2* seems to be related to jasmonate biosynthesis due to the accumulation of methyl jasmonate, probably contributing to the ripening process and the metabolism side-effects derived from fruit's maturation.
5. Ellagitannins related QTLs were found as stable QTLs in the mapping population '232' $\times$ '1392' within two consecutive years.
6. There is a positive correlation between *FaUDP* and EAH2 concentration in Goodberry population.
7. *FaUDP* coding sequence is different from the one annotated in 'Camarosa' reference genome.
8. *FaUDP* expression pattern is varietal-specific, showing no expression in 'Camarosa', 'Candongá', Clery and Frida while Florence, Gariguet, Sonata and 'Senga Sengana' showed expression of this gene. Also, the concentration of EAH2 correlates in most cases with *FaUDP* expression pattern.
9. TurboID proximity labeling of *FaUDPSS* reveals putative interactors with proteins related to RNA and protein processing and primary and secondary metabolism.

VIII. Bibliography



Botanical illustration of *Fragaria chilensis* plant by Antoine Nicolas Duchesne. *Histoire Naturelle des Fraisiers*, 1765.

- Aaby K, Mazur S, Nes A, Skrede G.** 2012. Phenolic compounds in strawberry (*Fragaria x ananassa* Duch.) fruits: Composition in 27 cultivars and changes during ripening. *Food Chemistry* **132**, 86–97. doi: [10.1016/j.foodchem.2011.10.037](https://doi.org/10.1016/j.foodchem.2011.10.037).
- Abramson J, Adler J, Dunger J, et al.** 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500. doi: [10.1038/s41586-024-07487-w](https://doi.org/10.1038/s41586-024-07487-w).
- Aharoni A.** 2008. Non-targeted analysis of spatial metabolite composition in strawberry (*Fragaria* × *ananassa*) flowers. *Phytochemistry* **69**, 2463–2481. doi: [10.1016/j.phytochem.2008.07.009](https://doi.org/10.1016/j.phytochem.2008.07.009).
- Aharoni A, Giri AP, Verstappen FWA, Berteaux CM, Sevenier R, Sun Z, Jongsma MA, Schwab W, Bouwmeester HJ.** 2004. Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *Plant Cell* **16**, 3110–3131. doi: [10.1105/tpc.104.023895](https://doi.org/10.1105/tpc.104.023895).
- Aharoni A, De Vos CHR, Wein M, Sun Z, Greco R, Kroon A, Mol JNM, O'Connell AP.** 2001. The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant Journal* **28**, 319–332. doi: [10.1046/j.1365-3113X.2001.01154.x](https://doi.org/10.1046/j.1365-3113X.2001.01154.x).
- Ahmad F, Pasha I, Saeed M, Asgher M.** 2018. Biochemical profiling of Pakistani sorghum and millet varieties with special reference to anthocyanins and condensed tannins. *International Journal of Food Properties* **21**, 1586–1597. doi: [10.1080/10942912.2018.1502198](https://doi.org/10.1080/10942912.2018.1502198).
- Akagi T, Ikegami A, Suzuki Y, Yoshida J, Yamada M, Sato A, Yonemori K.** 2009. Expression balances of structural genes in shikimate and flavonoid biosynthesis cause a difference in proanthocyanidin accumulation in persimmon (*Diospyros kaki* Thunb.) fruit. *Planta* **230**, 899–915. doi: [10.1007/s00425-009-0991-6](https://doi.org/10.1007/s00425-009-0991-6).
- Akagi T, Katayama-Ikegami A, Yonemori K.** 2011. Proanthocyanidin biosynthesis of persimmon (*Diospyros kaki* Thunb.) fruit. *Scientia Horticulturae* **130**, 373–380. doi: [10.1016/j.scienta.2011.07.021](https://doi.org/10.1016/j.scienta.2011.07.021).
- Alger EI, Edger PP.** 2020. One subgenome to rule them all: underlying mechanisms of subgenome dominance. *Current Opinion in Plant Biology* **54**, 108–113. doi: [10.1016/j.pbi.2020.03.004](https://doi.org/10.1016/j.pbi.2020.03.004).
- Ali JG, Alborn HT, Campos-Herrera R, Kaplan F, Duncan LW, Rodriguez-Saona C, Koppenhöfer AM, Stelinski LL.** 2012. Subterranean, herbivore-induced plant volatile increases biological control activity of multiple beneficial nematode species in distinct habitats. *PLOS ONE* **7**, e38146. doi: [10.1371/journal.pone.0038146](https://doi.org/10.1371/journal.pone.0038146).
- Ambigaipalan P, de Camargo AC, Shahidi F.** 2017. Identification of phenolic antioxidants and bioactives of pomegranate seeds following juice extraction using HPLC-DAD-ESI-MSn. *Food Chemistry* **221**, 1883–1894. doi: [10.1016/j.foodchem.2016.10.058](https://doi.org/10.1016/j.foodchem.2016.10.058).
- Amil-Ruiz F, Garrido-Gala J, Blanco-Portales R, Foltá KM, Muñoz-Blanco J, Caballero JL.** 2013. Identification and validation of reference genes for transcript normalization in strawberry (*Fragaria* × *ananassa*) defense responses. *PLOS ONE* **8**, e70603. doi: [10.1371/journal.pone.0070603](https://doi.org/10.1371/journal.pone.0070603).

- Aragüez I, Osorio S, Hoffmann T, Rambla JL, Medina-Escobar N, Granell A, Botella MÁ, Schwab W, Valpuesta V.** 2013. Eugenol production in achenes and receptacles of strawberry fruits is catalyzed by synthases exhibiting distinct kinetics. *Plant Physiology* **163**, 946–958. doi: [10.1104/pp.113.224352](https://doi.org/10.1104/pp.113.224352).
- Arimura GI, Ozawa R, Kugimiya S, Takabayashi J, Bohlmann J.** 2004. Herbivore-induced defense response in a model legume. Two-spotted spider mites induce emission of (*E*)- β -ocimene and transcript accumulation of (*E*)- β -ocimene synthase in *Lotus japonicus*. *Plant Physiology* **135**, 1976–1983. doi: [10.1104/pp.104.042929](https://doi.org/10.1104/pp.104.042929).
- Ariza MT, Reboredo-Rodríguez P, Cervantes L, Soria C, Martínez-Ferri E, González-Barreiro C, Cancho-Grande B, Battino M, Simal-Gándara J.** 2018. Bioaccessibility and potential bioavailability of phenolic compounds from achenes as a new target for strawberry breeding programs. *Food Chemistry* **248**, 155–165. doi: [10.1016/j.foodchem.2017.11.105](https://doi.org/10.1016/j.foodchem.2017.11.105).
- Armenteros JJA, Salvatore M, Emanuelsson O, Winther O, Von Heijne G, Elofsson A, Nielsen H.** 2019. Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance* **2**. doi: [10.26508/lsa.201900429](https://doi.org/10.26508/lsa.201900429).
- Arnold RA, Noble AC, Singleton VL.** 1980. Bitterness and astringency of phenolic fractions in wine. *Journal of Agricultural and Food Chemistry* **28**, 675–678. doi: [10.1021/jf60229a026](https://doi.org/10.1021/jf60229a026).
- Aron PM, Kennedy JA.** 2008. Flavan-3-ols: Nature, occurrence and biological activity. *Molecular Nutrition and Food Research* **52**, 79–104. doi: [10.1002/mnfr.200700137](https://doi.org/10.1002/mnfr.200700137).
- Ashman TL.** 2000. Pollinator selectivity and its implications for the evolution of dioecy and sexual dimorphism. *Ecology* **81**, 2577–2591. doi: [10.1890/0012-9658\(2000\)081\[2577:PSAIIF\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[2577:PSAIIF]2.0.CO;2).
- Azuma H, Toyota M, Asakawa Y.** 2001. Intraspecific variation of floral scent chemistry in *Magnolia kobus* DC. (*Magnoliaceae*). *Journal of Plant Research* **114**, 411–422. doi: [10.1007/pl00014006](https://doi.org/10.1007/pl00014006).
- Babenko LM, Shcherbatiuk MM, Skaterna TD, Kosakivska I V.** 2017. Lipxygenases and their metabolites in formation of plant stress tolerance. *Ukrainian Biochemical Journal* **89**, 5–21. doi: [10.15407/ubj89.01.005](https://doi.org/10.15407/ubj89.01.005).
- Baldwin IT, Halitschke R, Paschold A, Von Dahl CC, Preston CA.** 2006. Volatile signaling in plant-plant interactions: ‘Talking trees’ in the genomics era. *Science* **311**, 812–815. doi: [10.1126/science.1118446](https://doi.org/10.1126/science.1118446).
- Banerjee A, Jana A, Pati BR, Mondal KC, Das Mohapatra PK.** 2012. Characterization of tannase protein sequences of bacteria and fungi: An *in silico* study. *Protein Journal* **31**, 306–327. doi: [10.1007/s10930-012-9405-x](https://doi.org/10.1007/s10930-012-9405-x).
- Bar-Ya’akov I, Tian L, Amir R, Holland D.** 2019. Primary metabolites, anthocyanins, and hydrolyzable tannins in the pomegranate fruit. *Frontiers in Plant Science* **10**, 620. doi: [10.3389/fpls.2019.00620](https://doi.org/10.3389/fpls.2019.00620).
- Basu A, Nguyen A, Betts NM, Lyons TJ.** 2014. Strawberry as a functional food: an evidence-based review. *Critical reviews in food science and nutrition* **54**, 790–806. doi: [10.1080/10408398.2011.608174](https://doi.org/10.1080/10408398.2011.608174).
- Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L.** 2004. TT2,

TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant Journal* **39**, 366–380. doi: [10.1111/j.1365-3113X.2004.02138.x](https://doi.org/10.1111/j.1365-3113X.2004.02138.x).

Baxter NJ, Lilley TH, Haslam E, Williamson MP. 1997. Multiple interactions between polyphenols and a salivary proline-rich protein repeat result in complexation and precipitation. *Biochemistry* **36**, 5566–5577. doi: [10.1021/bi9700328](https://doi.org/10.1021/bi9700328).

Benfey PN, Chua NH. 1989. Regulated genes in transgenic plants. *Science* **244**, 174–181. doi: [10.1126/science.244.4901.174](https://doi.org/10.1126/science.244.4901.174).

Bentley R, Haslam E. 1990. The shikimate pathway - A metabolic tree with many branches. *Critical Reviews in Biochemistry and Molecular Biology* **25**, 307–384. doi: [10.3109/10409239009090615](https://doi.org/10.3109/10409239009090615).

Benzie IFF, Strain JJ. 1999. Ferric reducing/antioxidant power assay: Direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. *Methods in Enzymology*. *Methods Enzymol*, 15–27. doi: [10.1016/S0076-6879\(99\)99005-5](https://doi.org/10.1016/S0076-6879(99)99005-5).

Bharadwaj R, Kumar SR, Sharma A, Sathishkumar R. 2021. Plant metabolic gene clusters: evolution, organization, and their applications in synthetic biology. *Frontiers in Plant Science* **12**, 697318. doi: [10.3389/fpls.2021.697318](https://doi.org/10.3389/fpls.2021.697318).

Bird KA, VanBuren R, Puzey JR, Edger PP. 2018. The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytologist* **220**, 87–93. doi: [10.1111/nph.15256](https://doi.org/10.1111/nph.15256).

Blois MS. 1958. Antioxidant determinations by the use of a stable free radical. *Nature* **181**, 1199–1200. doi: [10.1038/1811199a0](https://doi.org/10.1038/1811199a0).

Blum M, Andreeva A, Florentino LC, et al. 2025. InterPro: the protein sequence classification resource in 2025. *Nucleic acids research* **53**, D444–D456. doi: [10.1093/nar/gkaf1082](https://doi.org/10.1093/nar/gkaf1082).

Boatright J, Negre F, Chen X, Kish CM, Wood B, Peel G, Orlova I, Gang D, Rhodes D, Dudareva N. 2004. Understanding *in vivo* benzenoid metabolism in petunia petal tissue. *Plant Physiology* **135**, 1993–2011. doi: [10.1104/pp.104.045468](https://doi.org/10.1104/pp.104.045468).

Bohin MC, Vincken J-P, van der Hijden HTWM, Gruppen H. 2012. Efficacy of food proteins as carriers for flavonoids. *Journal of Agricultural and Food Chemistry* **60**, 4136–4143. doi: [10.1021/jf205292r](https://doi.org/10.1021/jf205292r).

Bolger M, Schwacke R, Usadel B. 2021. MapMan visualization of RNA-Seq data using Mercator4 functional annotations. *Methods in Molecular Biology* **2354**, 195–212. doi: [10.1007/978-1-0716-1609-3_9](https://doi.org/10.1007/978-1-0716-1609-3_9).

Bontpart T, Marlin T, Vialet S, Guiraud JL, Pinasseau L, Meudec E, Sommerer N, Cheynier V, Terrier N. 2016. Two shikimate dehydrogenases, *VvSDH3* and *VvSDH4*, are involved in gallic acid biosynthesis in grapevine. *Journal of Experimental Botany* **67**, 3537–3550. doi: [10.1093/jxb/erw184](https://doi.org/10.1093/jxb/erw184).

Bowles D, Lim EK, Poppenberger B, Vaistij FE. 2006. Glycosyltransferases of lipophilic small molecules. *Annual Review of Plant Biology* **57**, 567–597. doi: [10.1146/annurev.arplant.57.032905.105429](https://doi.org/10.1146/annurev.arplant.57.032905.105429).

- Brash AR.** 1999. Lipoxygenases: Occurrence, functions, catalysis, and acquisition of substrate. *Journal of Biological Chemistry* **274**, 23679–23682. doi: [10.1074/jbc.274.34.23679](https://doi.org/10.1074/jbc.274.34.23679).
- Brecht JK, Loaza FE, Nunes MCN, Emond JP, Uysal I, Badia F, Wells J, Saenz J.** 2016. Reducing strawberry waste and losses in the postharvest supply chain via intelligent distribution management. *Acta Horticulturae* **1120**, 253–259. doi: [10.17660/ActaHortic.2016.1120.38](https://doi.org/10.17660/ActaHortic.2016.1120.38).
- Broman KW, Wu H, Sen S, Churchill GA.** 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890. doi: [10.1093/bioinformatics/btg112](https://doi.org/10.1093/bioinformatics/btg112).
- Bursać Kovačević D, Putnik P, Dragović-Uzelac V, Vahčić N, Babojelić MS, Levaj B.** 2015. Influences of organically and conventionally grown strawberry cultivars on anthocyanins content and color in purees and low-sugar jams. *Food Chemistry* **181**, 94–100. doi: [10.1016/j.foodchem.2015.02.063](https://doi.org/10.1016/j.foodchem.2015.02.063).
- Bustamante CA, Civello PM, Martínez GA.** 2009. Cloning of the promoter region of β -xylosidase (*FaXylI*) gene and effect of plant growth regulators on the expression of *FaXylI* in strawberry fruit. *Plant Science* **177**, 49–56. doi: [10.1016/j.plantsci.2009.03.015](https://doi.org/10.1016/j.plantsci.2009.03.015).
- Carbone F, Preuss A, De Vos RCH, D'Amico E, Perrotta G, Bovy AG, Martens S, Rosati C.** 2009. Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits. *Plant, Cell and Environment* **32**, 1117–1131. doi: [10.1111/j.1365-3040.2009.01994.x](https://doi.org/10.1111/j.1365-3040.2009.01994.x).
- Carpenter G, Cleaver L, Blakeley M, Hasbullah N, Houghton J, Gardner A.** 2019. Wine astringency reduces flavor intensity of Brussels sprouts. *Journal of Texture Studies* **50**, 71–74. doi: [10.1111/jtxs.12378](https://doi.org/10.1111/jtxs.12378).
- Castellarin SD, Matthews MA, Di Gaspero G, Gambetta GA.** 2007. Water deficits accelerate ripening and induce changes in gene expression regulating flavonoid biosynthesis in grape berries. *Planta* **227**, 101–112. doi: [10.1007/s00425-007-0598-8](https://doi.org/10.1007/s00425-007-0598-8).
- Chai YM, Jia HF, Li CL, Dong QH, Shen YY.** 2011. FaPYR1 is involved in strawberry fruit ripening. *Journal of Experimental Botany* **62**, 5079–5089. doi: [10.1093/jxb/err207](https://doi.org/10.1093/jxb/err207).
- Chambers AH, Pillet J, Plotto A, Bai J, Whitaker VM, Foltá KM.** 2014. Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC Genomics* **15**, 1–15. doi: [10.1186/1471-2164-15-217](https://doi.org/10.1186/1471-2164-15-217).
- Chaudhuri S, Coggins JR.** 1985. The purification of shikimate dehydrogenase from *Escherichia coli*. *The Biochemical journal* **226**, 217–223. doi: [10.1042/bj2260217](https://doi.org/10.1042/bj2260217).
- Chen G, Hackett R, Walker D, Taylor A, Lin Z, Grierson D.** 2004. Identification of a specific isoform of tomato lipoxygenase (*TomloxC*) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiology* **136**, 2641. doi: [10.1104/PP.104.041608](https://doi.org/10.1104/PP.104.041608).
- Chen Y, Jerneřen F, Oliw EH.** 2017. Purification and site-directed mutagenesis of linoleate 9S-dioxygenase-allene oxide synthase of *Fusarium oxysporum* confirms the oxygenation mechanism. *Archives of Biochemistry and Biophysics* **625–626**, 24–29. doi: [10.1016/j.abb.2017.05.007](https://doi.org/10.1016/j.abb.2017.05.007).
- Chen Q, Jiang T, Liu YX, et al.** 2019. Recently duplicated sesterterpene (C25) gene

clusters in *Arabidopsis thaliana* modulate root microbiota. *Science China Life Sciences* **62**, 947–958. doi: [10.1007/s11427-019-9521-2](https://doi.org/10.1007/s11427-019-9521-2).

Chen X, Xiong J, Huang S, Li X, Zhang Y, Zhang L, Wang F. 2018. Analytical profiling of proanthocyanidins from *Acacia mearnsii* bark and *in vitro* assessment of antioxidant and antidiabetic potential. *Molecules* **23**, 2891. doi: [10.3390/molecules23112891](https://doi.org/10.3390/molecules23112891).

Cheng F, Sun C, Wu J, Schnable J, Woodhouse MR, Liang J, Cai C, Freeling M, Wang X. 2016. Epigenetic regulation of subgenome dominance following whole genome triplication in *Brassica rapa*. *New Phytologist* **211**, 288–299. doi: [10.1111/nph.13884](https://doi.org/10.1111/nph.13884).

Chini A, Fonseca S, Fernández G, et al. 2007. The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* **448**, 666–671. doi: [10.1038/nature06006](https://doi.org/10.1038/nature06006).

Chini A, Monte I, Zamarreño AM, et al. 2018. An OPR3-independent pathway uses 4,5-didehydrojasmonate for jasmonate synthesis. *Nature Chemical Biology* **14**, 171–178. doi: [10.1038/nchembio.2540](https://doi.org/10.1038/nchembio.2540).

Christensen SA, Huffaker A, Kaplan F, et al. 2015. Maize death acids, 9-lipoxygenase-derived cyclopent(a)nonenes, display activity as cytotoxic phytoalexins and transcriptional mediators. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 11407–11412. doi: [10.1073/pnas.1511131112](https://doi.org/10.1073/pnas.1511131112).

Christianson DW. 2017. Structural and chemical biology of terpenoid cyclases. *Chemical Reviews* **117**, 11570–11648. doi: [10.1021/acs.chemrev.7b00287](https://doi.org/10.1021/acs.chemrev.7b00287).

Coffa G, Imber AN, Maguire BC, Laxmikanthan G, Schneider C, Gaffney BJ, Brash AR. 2005a. On the relationships of substrate orientation, hydrogen abstraction, and product stereochemistry in single and double dioxygenations by soybean lipoxygenase-1 and its Ala542Gly mutant. *Journal of Biological Chemistry* **280**, 38756–38766. doi: [10.1074/jbc.M504870200](https://doi.org/10.1074/jbc.M504870200).

Coffa G, Schneider C, Brash AR. 2005b. A comprehensive model of positional and stereo control in lipoxygenases. *Biochemical and Biophysical Research Communications* **338**, 87–92. doi: [10.1016/j.bbrc.2005.07.185](https://doi.org/10.1016/j.bbrc.2005.07.185).

Coggins JR, Boocock MR, Chaudhuri S, Lambert JM, Lumsden J, Nimmo GA, Smith DDS. 1987. The arom multifunctional enzyme from *Neurospora crassa*. *Methods in Enzymology*. *Methods Enzymol*, 325–341. doi: [10.1016/S0076-6879\(87\)42044-2](https://doi.org/10.1016/S0076-6879(87)42044-2).

Colle M, Leisner CP, Wai CM, et al. 2019. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *GigaScience* **8**, 1–15. doi: [10.1093/gigascience/giz012](https://doi.org/10.1093/gigascience/giz012).

Comai L. 2005. The advantages and disadvantages of being polyploid. doi: [10.1038/nrg1711](https://doi.org/10.1038/nrg1711).

Concha CM, Figueroa NE, Poblete LA, Oñate FA, Schwab W, Figueroa CR. 2013. Methyl jasmonate treatment induces changes in fruit ripening by modifying the expression of several ripening genes in *Fragaria chiloensis* fruit. *Plant Physiology and Biochemistry* **70**, 433–444. doi: [10.1016/j.plaphy.2013.06.008](https://doi.org/10.1016/j.plaphy.2013.06.008).

Constabel CP. 2018. Molecular Controls of Proanthocyanidin Synthesis and Structure: Prospects for Genetic Engineering in Crop Plants. *Journal of Agricultural and Food Chemistry* **66**, 9882–9888. doi: [10.1021/acs.jafc.8b02950](https://doi.org/10.1021/acs.jafc.8b02950).

- Constantin A-E.** 2019. Significance Brackets for 'ggplot2'. R package version 0.6.0. CRAN: Contributed Packages 4. doi: [10.32614/cran.package.ggsignif](https://doi.org/10.32614/cran.package.ggsignif).
- Cory H, Passarelli S, Szeto J, Tamez M, Mattei J.** 2018. The role of polyphenols in human health and food systems: a mini-review. *Frontiers in Nutrition* **5**, 87. doi: [10.3389/fnut.2018.00087](https://doi.org/10.3389/fnut.2018.00087).
- Couée I, Sulmon C, Gouesbet G, El Amrani A.** 2006. Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *Journal of Experimental Botany* **57**, 449–459. doi: [10.1093/jxb/erj027](https://doi.org/10.1093/jxb/erj027).
- D'Auria JC, Reichelt M, Luck K, Svatoš A, Gershenzon J.** 2007. Identification and characterization of the BAHD acyltransferase malonyl CoA: Anthocyanidin 5-O-glucoside-6"-O-malonyltransferase (*At5MAT*) in *Arabidopsis thaliana*. *FEBS Letters* doi: [10.1016/j.febslet.2007.01.060](https://doi.org/10.1016/j.febslet.2007.01.060).
- Dai X, Liu Y, Zhuang J, et al.** 2020. Discovery and characterization of tannase genes in plants: roles in hydrolysis of tannins. *New Phytologist* **226**, 1104–1116. doi: [10.1111/nph.16425](https://doi.org/10.1111/nph.16425).
- Daminato M, Guzzo F, Casadoro G.** 2013. A SHATTERPROOF-like gene controls ripening in non-climacteric strawberries, and auxin and abscisic acid antagonistically affect its expression. *Journal of Experimental Botany* **64**, 3775–3786. doi: [10.1093/jxb/ert21](https://doi.org/10.1093/jxb/ert21).
- Debeaujon I, Peeters AJM, Leon-Kloosterziel KM, Koornneef M.** 2001. The TRANSPARENT TESTA12 gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *The Plant Cell* **13**, 853. doi: [10.2307/3871345](https://doi.org/10.2307/3871345).
- Degenhardt J, Hiltbold I, Köllner TG, Frey M, Gierl A, Gershenzon J, Hibbard BE, Ellersieck MR, Turlings TCJ.** 2009. Restoring a maize root signal that attracts insect-killing nematodes to control a major pest. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13213–13218. doi: [10.1073/pnas.0906365106](https://doi.org/10.1073/pnas.0906365106).
- Delgado LD, Zúñiga PE, Figueroa NE, Pastene E, Escobar-Sepúlveda HF, Figueroa PM, Garrido-Bigotes A, Figueroa CR.** 2018. Application of a JA-ile biosynthesis inhibitor to methyl jasmonate-treated strawberry fruit induces upregulation of specific MBW complex-related genes and accumulation of proanthocyanidins. *Molecules* **23**, 1433. doi: [10.3390/molecules23061433](https://doi.org/10.3390/molecules23061433).
- Deng Y, Lu S.** 2017. Biosynthesis and regulation of phenylpropanoids in plants. *Critical Reviews in Plant Sciences* **36**, 257–290. doi: [10.1080/07352689.2017.1402852](https://doi.org/10.1080/07352689.2017.1402852).
- Devoto A, Nieto-Rostro M, Xie D, Ellis C, Harmston R, Patrick E, Davis J, Sherratt L, Coleman M, Turner JG.** 2002. COI1 links jasmonate signalling and fertility to the SCF ubiquitin-ligase complex in *Arabidopsis*. *Plant Journal* **32**, 457–466. doi: [10.1046/j.1365-313X.2002.01432.x](https://doi.org/10.1046/j.1365-313X.2002.01432.x).
- Dexter R, Qualley A, Kish CM, Ma CJ, Koeduka T, Nagegowda DA, Dudareva N, Pichersky E, Clark D.** 2007. Characterization of a petunia acetyltransferase involved in the biosynthesis of the floral volatile isoeugenol. *Plant Journal* **49**, 265–275. doi: [10.1111/j.1365-313X.2006.02954.x](https://doi.org/10.1111/j.1365-313X.2006.02954.x).

- DiMeglio LM, Staudt G, Yu H, Davis TM.** 2014. A phylogenetic analysis of the genus *Fragaria* (strawberry) using intron-containing sequence from the *ADH-1* gene. *PLoS ONE* **9**, e102237. doi: [10.1371/journal.pone.0102237](https://doi.org/10.1371/journal.pone.0102237).
- Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arús P, Laigret F.** 2002. Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theoretical and Applied Genetics* **105**, 127–138. doi: [10.1007/s00122-002-0867-7](https://doi.org/10.1007/s00122-002-0867-7).
- Dixon RA, Liu C, Jun JH.** 2013. Metabolic engineering of anthocyanins and condensed tannins in plants. *Current Opinion in Biotechnology* **24**, 329–335. doi: [10.1016/j.copbio.2012.07.004](https://doi.org/10.1016/j.copbio.2012.07.004).
- Dixon RA, Xie DY, Sharma SB.** 2005. Proanthocyanidins - A final frontier in flavonoid research? *New Phytologist* **165**, 9–28. doi: [10.1111/j.1469-8137.2004.01217.x](https://doi.org/10.1111/j.1469-8137.2004.01217.x).
- Dorcey E, Urbez C, Blázquez MA, Carbonell J, Perez-Amador MA.** 2009. Fertilization-dependent auxin response in ovules triggers fruit development through the modulation of gibberellin metabolism in *Arabidopsis*. *Plant Journal* **58**, 318–332. doi: [10.1111/j.1365-313X.2008.03781.x](https://doi.org/10.1111/j.1365-313X.2008.03781.x).
- Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF.** 2008. Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* **42**, 443–461. doi: [10.1146/annurev.genet.42.110807.091524](https://doi.org/10.1146/annurev.genet.42.110807.091524).
- Du X, Plotto A, Baldwin E, Rouseff R.** 2011. Evaluation of volatiles from two subtropical strawberry cultivars using GC-olfactometry, GC-MS odor activity values, and sensory analysis. *Journal of Agricultural and Food Chemistry* **59**, 12569–12577. doi: [10.1021/jf2030924](https://doi.org/10.1021/jf2030924).
- Dudareva N, Klempien A, Muhlemann JK, Kaplan I.** 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist* **198**, 16–32. doi: [10.1111/nph.12145](https://doi.org/10.1111/nph.12145).
- Dyer WE, Henstrand JM, Handa AK, Herrmann KM.** 1989. Wounding induces the first enzyme of the shikimate pathway in *Solanaceae*. *Proceedings of the National Academy of Sciences* **86**, 7370–7373. doi: [10.1073/pnas.86.19.7370](https://doi.org/10.1073/pnas.86.19.7370).
- Eckardt NA.** 2014. Genome dominance and interaction at the gene expression level in allohexaploid wheat. *Plant Cell* **26**, 1834. doi: [10.1105/tpc.114.127183](https://doi.org/10.1105/tpc.114.127183).
- Edger PP, Poorten TJ, VanBuren R, Hardigan MA, Colle M, McKain MR, Smith RD, Teresi SJ, Nelson ADL, Wai CM, Alger EI, Bird KA, Yocca AE, Pumplun N, Ou S, Ben-Zvi G, Brodt A, Baruch K, Swale T, Shiue L, Acharya CB, Cole GS, Mower JP, Childs LK, Jiang N, Lyons E, Feeling M, Puzey JP, Knapp SJ.** 2019. Origin and evolution of the octoploid strawberry genome. *Nature Genetics* **51**, 541–547. doi: [10.1038/s41588-019-0356-4](https://doi.org/10.1038/s41588-019-0356-4).
- Edger PP, Smith R, McKain MR, Cooley AM, Vallejo-Marin M, Yuan Y, Bewick AJ, Ji L, Platts AE, Bowman MJ, Childs KL, Washburn JD, Schmitz RJ, Smith GD, Pires JC, Puzey JR.** 2017. Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower. *Plant Cell* **29**, 2150–2167. doi: [10.1105/tpc.17.00010](https://doi.org/10.1105/tpc.17.00010).

- Edger PP, VanBuren R, Colle M, Colle M, Poorten TJ, Wai CM, Niederhuth CE, Alger EI, Ou S, Acharya CB, W J, Callow P, McKain MR, Shi J, Collier C, Xiong Z, Mower JP, Slovin JP, Hytönen T, Jiang N, Childs KL, Knapp S.** 2018. Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *GigaScience* 7, 2. doi: [10.1093/gigascience/gix124](https://doi.org/10.1093/gigascience/gix124).
- Engwa GA.** 2018. Free radicals and the role of plant phytochemicals as antioxidants against oxidative stress-related diseases. *Phytochemicals - Source of Antioxidants and Role in Disease Prevention*. InTech. doi: [10.5772/intechopen.76719](https://doi.org/10.5772/intechopen.76719).
- Enomoto H.** 2021. Unique distribution of ellagitannins in ripe strawberry fruit revealed by mass spectrometry imaging. *Current Research in Food Science* 4, 821–828. doi: [10.1016/j.crfs.2021.11.006](https://doi.org/10.1016/j.crfs.2021.11.006).
- Eriksson T, Donoghue MJ, Hibbs MS.** 1998. Phylogenetic analysis of *Potentilla* using DNA sequences of nuclear ribosomal internal transcribed spacers (ITS), and implications for the classification of *Rosoideae* (*Rosaceae*). *Plant Systematics and Evolution* 211, 155–179. doi: [10.1007/BF00985357](https://doi.org/10.1007/BF00985357).
- Fait A, Hanhineva K, Beleggia R, Dai N, Rogachev I, Nikiforova VJ, Fernie AR, Aharoni A.** 2008. Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiology* 148, 730–750. doi: [10.1104/PP.108.120691](https://doi.org/10.1104/PP.108.120691).
- Falcão L, Araújo MEM.** 2018. Vegetable tannins used in the manufacture of historic leathers. *Molecules* 23, 1081. doi: [10.3390/molecules23051081](https://doi.org/10.3390/molecules23051081).
- Fan Z, Hasing T, Johnson TS, Garner DM, Barbey CR, Colquhoun TA, Sims CA, Resende MFR, Whitaker VM.** 2021. Strawberry sweetness and consumer preference are enhanced by specific volatile compounds. *Horticulture Research* 8, 66. doi: [10.1038/s41438-021-00502-5](https://doi.org/10.1038/s41438-021-00502-5).
- Fan Z, Tieman DM, Knapp SJ, et al.** 2022. A multi-omics framework reveals strawberry flavor genes and their regulatory elements. *New Phytologist* 236, 1089–1107. doi: [10.1111/nph.18416](https://doi.org/10.1111/nph.18416).
- Fauconnier ML, Marlier M.** 1996. An efficient procedure for the production of fatty acid hydroperoxides from hydrolyzed flax seed oil and soybean lipoxygenase. *Biotechnology Techniques* 10, 839–844. doi: [10.1007/bf00154668](https://doi.org/10.1007/bf00154668).
- Fauconnier ML, Perez AG, Sanz C, Marlier M.** 1997. Purification and characterization of tomato leaf (*Lycopersicon esculentum* Mill.) hydroperoxide lyase. *Journal of Agricultural and Food Chemistry* 45, 4232–4236. doi: [10.1021/jf9701042](https://doi.org/10.1021/jf9701042).
- Felipe L, Oliveira A, Bicas J.** 2017. Bioaromas - perspectives for sustainable development. *Trends in Food Sciences & Technology* 62, 141–153. doi: [10.1016/j.tifs.2017.02.005](https://doi.org/10.1016/j.tifs.2017.02.005).
- Feller A, Hernandez JM, Grotewold E.** 2006. An ACT-like domain participates in the dimerization of several plant basic-helix-loop-helix transcription factors. *Journal of Biological Chemistry* 281, 28964–28974. doi: [10.1074/jbc.M603262200](https://doi.org/10.1074/jbc.M603262200).
- Fernández-Calvo P, Chini A, Fernández-Barbero G, et al.** 2011. The *Arabidopsis* bHLH transcription factors *MYC3* and *MYC4* are targets of *JAZ* repressors and act

additively with *MYC2* in the activation of jasmonate responses. *Plant Cell* **23**, 701–715. doi: [10.1105/tpc.110.080788](https://doi.org/10.1105/tpc.110.080788).

Fernández-Lara R, Gordillo B, Rodríguez-Pulido FJ, Lourdes González-Miret M, del Villar-Martínez AA, Dávila-Ortiz G, Heredia FJ. 2015. Assessment of the differences in the phenolic composition and color characteristics of new strawberry (*Fragaria x ananassa* Duch.) cultivars by HPLC-MS and imaging tristimulus colorimetry. *Food Research International* **76**, 645–653. doi: [10.1016/j.foodres.2015.07.038](https://doi.org/10.1016/j.foodres.2015.07.038).

Fernandez-Pozo N, Menda N, Edwards JD, et al. 2015. The Sol Genomics Network (SGN)-from genotype to phenotype to breeding. *Nucleic Acids Research* **43**, D1036–D1041. doi: [10.1093/nar/gku1195](https://doi.org/10.1093/nar/gku1195).

Feschotte C, Jiang N, Wessler SR. 2002. Plant transposable elements: Where genetics meets genomics. *Nature Reviews Genetics* **3**, 329–341. doi: [10.1038/nrg793](https://doi.org/10.1038/nrg793).

Feussner I, Hause B, Vörös K, Parthier B, Wasternack C. 1995. Jasmonate-induced lipoxygenase forms are localized in chloroplasts of barley leaves (*Hordeum vulgare* cv. Salome). *The Plant Journal* **7**, 949–957. doi: [10.1046/j.1365-313x.1995.07060949.x](https://doi.org/10.1046/j.1365-313x.1995.07060949.x).

Feussner I, Wasternack C. 2002. The lipoxygenase pathway. *Annual Review of Plant Biology* **53**, 275–297. doi: [10.1146/annurev.arplant.53.100301.135248](https://doi.org/10.1146/annurev.arplant.53.100301.135248).

Field B, Fiston-Lavier AS, Kemen A, Geisler K, Quesneville H, Osbourn AE. 2011. Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16116–16121. doi: [10.1073/pnas.1109273108](https://doi.org/10.1073/pnas.1109273108).

Field B, Osbourn AE. 2008. Metabolic diversification - Independent assembly of operon-like gene clusters in different plants. *Science* **320**, 543–547. doi: [10.1126/science.1154990](https://doi.org/10.1126/science.1154990).

Figuroa NE, Hoffmann T, Olbricht K, Abrams SR, Schwab W. 2021. Contrasting dynamics in abscisic acid metabolism in different *Fragaria spp.* During fruit ripening and identification of the enzymes involved. *Journal of Experimental Botany* **72**, 1245–1259. doi: [10.1093/jxb/eraa503](https://doi.org/10.1093/jxb/eraa503).

Finn CE, Retamales JB, Lobos GA, Hancock JF. 2013. The chilean strawberry (*Fragaria chiloensis*): Over 1000 years of domestication. *HortScience* **48**, 418–421. doi: [10.21273/hortsci.48.4.418](https://doi.org/10.21273/hortsci.48.4.418).

Fischer UA, Carle R, Kammerer DR. 2011. Identification and quantification of phenolic compounds from pomegranate (*Punica granatum* L.) peel, mesocarp, aril and differently produced juices by HPLC-DAD-ESI/MSn. *Food Chemistry* **127**, 807–821. doi: [10.1016/j.foodchem.2010.12.156](https://doi.org/10.1016/j.foodchem.2010.12.156).

Floegel A, Kim DO, Chung SJ, Koo SI, Chun OK. 2011. Comparison of ABTS/DPPH assays to measure antioxidant capacity in popular antioxidant-rich US foods. *Journal of Food Composition and Analysis* **24**, 1043–1048. doi: [10.1016/j.jfca.2011.01.008](https://doi.org/10.1016/j.jfca.2011.01.008).

Fontes N, Gerós H, Delrot S. 2011. Grape berry vacuole: A complex and heterogeneous membrane system specialized in the accumulation of solutes. *American Journal of Enology and Viticulture* **62**, 270–278. doi: [10.5344/ajev.2011.10125](https://doi.org/10.5344/ajev.2011.10125).

de Francesco G, Bravi E, Sanarica E, Marconi O, Cappelletti F, Perretti G. 2020.

Effect of addition of different phenolic-rich extracts on beer flavour stability. *Foods* **9**, 1638. doi: [10.3390/foods9111638](https://doi.org/10.3390/foods9111638).

Freeling M, Thomas BC. 2006. Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Research* **16**, 805–814. doi: [10.1101/gr.3681406](https://doi.org/10.1101/gr.3681406).

Frey M, Chomet P, Glawischnig E, et al. 1997. Analysis of a chemical plant defense mechanism in grasses. *Science* **277**, 696–699. doi: [10.1126/science.277.5326.696](https://doi.org/10.1126/science.277.5326.696).

Fu X, Shi P, He Q, et al. 2017. *AaPDR3*, a PDR transporter 3, is involved in sesquiterpene β -caryophyllene transport in *Artemisia annua*. *Frontiers in Plant Science* **8**, 261357. doi: [10.3389/fpls.2017.00723](https://doi.org/10.3389/fpls.2017.00723).

Fujii H, Chinnusamy V, Rodrigues A, Rubio S, Antoni R, Park SY, Cutler SR, Sheen J, Rodriguez PL, Zhu JK. 2009. In vitro reconstitution of an abscisic acid signalling pathway. *Nature* **462**, 660–664. doi: [10.1038/nature08599](https://doi.org/10.1038/nature08599).

Gabetta B, Fuzzati N, Griffini A, Lolla E, Pace R, Ruffilli T, Peterlongo F. 2000. Characterization of proanthocyanidins from grape seeds. *Fitoterapia* **71**, 162–175. doi: [10.1016/S0367-326X\(99\)00161-6](https://doi.org/10.1016/S0367-326X(99)00161-6).

Gachon CMM, Langlois-Meurinne M, Saindrenan P. 2005. Plant secondary metabolism glycosyltransferases: The emerging functional analysis. *Trends in Plant Science* **10**, 542–549. doi: [10.1016/j.tplants.2005.09.007](https://doi.org/10.1016/j.tplants.2005.09.007).

Gambino G, Perrone I, Gribaudo I. 2008. A rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochemical Analysis* **19**, 520–525. doi: [10.1002/pca.1078](https://doi.org/10.1002/pca.1078).

Gang DR, Lavid N, Zubieta C, Chen F, Beuerle T, Lewinsohn E, Noel JP, Pichersky E. 2002. Characterization of phenylpropene *O*-methyltransferases from sweet basil: Facile change of substrate specificity and convergent evolution within a plant *O*-methyltransferase family. *Plant Cell* **14**, 505–519. doi: [10.1105/tpc.010327](https://doi.org/10.1105/tpc.010327).

Gargouri M, Drouet P, Legoy MD. 2004. Hydroperoxide-lyase activity in mint leaves: Volatile C6-aldehyde production from hydroperoxy-fatty acids. *Journal of Biotechnology* **111**, 59–65. doi: [10.1016/j.jbiotec.2004.03.005](https://doi.org/10.1016/j.jbiotec.2004.03.005).

Garrido-Bigotes A, Figueroa PM, Figueroa CR. 2018. Correction to: Jasmonate metabolism and its relationship with abscisic acid during strawberry fruit development and ripening. *Journal of Plant Growth Regulation* **37**, 705–706. doi: [10.1007/s00344-017-9755-x](https://doi.org/10.1007/s00344-017-9755-x).

Gharabli H, Della Gala V, Welner DH. 2023. The function of UDP-glycosyltransferases in plants and their possible use in crop protection. *Biotechnology Advances* **67**, 108182. doi: [10.1016/j.biotechadv.2023.108182](https://doi.org/10.1016/j.biotechadv.2023.108182).

Giampieri F, Forbes-Hernandez TY, Gasparrini M, Alvarez-Suarez JM, Afrin S, Bompadre S, Quiles JL, Mezzetti B, Battino M. 2015. Strawberry as a health promoter: An evidence based review. *Food and Function*. doi: [10.1039/c5fo00147a](https://doi.org/10.1039/c5fo00147a).

Giampieri F, Tulipani S, Alvarez-Suarez JM, Quiles JL, Mezzetti B, Battino M. 2012. The strawberry: Composition, nutritional quality, and impact on human health. *Nutrition* **28**, 9–19. doi: [10.1016/j.nut.2011.08.009](https://doi.org/10.1016/j.nut.2011.08.009).

- Göbel C, Feussner I, Schmidt A, Scheel D, Sanchez-Serrano J, Hamberg M, Rosahl S.** 2001. Oxylipin profiling reveals the preferential stimulation of the 9-Lipoxygenase pathway in elicitor-treated potato cells. *Journal of Biological Chemistry* **276**, 6267–6273. doi: [10.1074/jbc.M008606200](https://doi.org/10.1074/jbc.M008606200).
- Gonzalez A, Brown M, Hatlestad G, et al.** 2016. TTG2 controls the developmental regulation of seed coat tannins in *Arabidopsis* by regulating vacuolar transport steps in the proanthocyanidin pathway. *Developmental Biology* **419**, 54–63. doi: [10.1016/j.ydbio.2016.03.031](https://doi.org/10.1016/j.ydbio.2016.03.031).
- Görlach J, Raesecke HR, Rentsch D, Regenass M, Roy P, Zala M, Keel C, Boller T, Amrhein N, Schmid J.** 1995. Temporally distinct accumulation of transcripts encoding enzymes of the prechorismate pathway in elicitor-treated, cultured tomato cells. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3166–3170. doi: [10.1073/pnas.92.8.3166](https://doi.org/10.1073/pnas.92.8.3166).
- Grechkin AN, Brühlmann F, Mukhtarova LS, Gogolev Y V., Hamberg M.** 2006. Hydroperoxide lyases (CYP74C and CYP74B) catalyze the homolytic isomerization of fatty acid hydroperoxides into hemiacetals. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1761**, 1419–1428. doi: [10.1016/j.bbalip.2006.09.002](https://doi.org/10.1016/j.bbalip.2006.09.002).
- Grechkin AN, Mukhtarova LS, Hamberg M.** 2003. Detection of an enol intermediate in the hydroperoxide lyase chain cleavage reaction. *FEBS Letters* **549**, 31–34. doi: [10.1016/S0014-5793\(03\)00758-0](https://doi.org/10.1016/S0014-5793(03)00758-0).
- Greco M, Chiappetta A, Bruno L, Bitonti MB.** 2012. In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *Journal of Experimental Botany* **63**, 695–709. doi: [10.1093/jxb/err313](https://doi.org/10.1093/jxb/err313).
- Green BG.** 1993. Oral astringency: A tactile component of flavor. *Acta Psychologica* **84**, 119–125. doi: [10.1016/0001-6918\(93\)90078-6](https://doi.org/10.1016/0001-6918(93)90078-6).
- Gremillion KJ, Sobolik KD.** 1996. Dietary variability among prehistoric forager-farmers of eastern North America. *Current Anthropology* **37**, 529–539. doi: [10.1086/204515](https://doi.org/10.1086/204515).
- Griesser M, Vitzthum F, Fink B, Bellido ML, Raasch C, Munoz-Blanco J, Schwab W.** 2008. Multi-substrate flavonol *O*-glucosyltransferases from strawberry (*Fragaria × ananassa*) achene and receptacle. *Journal of Experimental Botany* **59**, 2611–2625. doi: [10.1093/JXB/ERN117](https://doi.org/10.1093/JXB/ERN117).
- Griffiths G.** 2015. Biosynthesis and analysis of plant oxylipins. *Free Radical Research* **49**, 565–582. doi: [10.3109/10715762.2014.1000318](https://doi.org/10.3109/10715762.2014.1000318).
- Grochowski DM, Skalicka-Woźniak K, Orhan IE, Xiao J, Locatelli M, Piwowarski JP, Granica S, Tomczyk M.** 2017. A comprehensive review of agrimoniin. *Annals of the New York Academy of Sciences* **1401**, 166–180. doi: [10.1111/nyas.13421](https://doi.org/10.1111/nyas.13421).
- Gross GG.** 1983. Partial purification and properties of UDP-glucose:vanillate 1-*O*-glucosyl transferase from oak leaves. *Phytochemistry* **22**, 2179–2182. doi: [10.1016/S0031-9422\(00\)80141-7](https://doi.org/10.1016/S0031-9422(00)80141-7).
- Gross GG, Denzel K.** 1991. Biosynthesis of gallotannins. β -glucogallin-dependent galloylation of 1,6-digalloylglucose to 1,2,6-trigalloylglucose. *Zeitschrift für Naturforschung - Section C Journal of Biosciences* **46**, 389–394. doi: [217](https://doi.org/10.1515/znc-1991-</p>
</div>
<div data-bbox=)

[5-609](#).

Grotewold E. 2004. The challenges of moving chemicals within and out of cells: insights into the transport of plant natural products. *Planta* **219**, 906–909. doi: [10.1007/s00425-004-1336-0](#).

Grundhöfer P, Niemetz R, Schilling G, Gross GG. 2001. Biosynthesis and subcellular distribution of hydrolyzable tannins. *Phytochemistry* **57**, 915–927. doi: [10.1016/S0031-9422\(01\)00099-1](#).

Gu Z. 2022. Complex heatmap visualization. *iMeta* **1**, e43. doi: [10.1002/imt2.43](#).

Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812. doi: [10.1093/bioinformatics/btu393](#).

Gu L, Kelm M, Hammerstone JF, Beecher G, Cunningham D, Vannozzi S, Prior RL. 2002. Fractionation of polymeric procyanidins from lowbush blueberry and quantification of procyanidins in selected foods with an optimized normal-phase HPLC-MS fluorescent detection method. *Journal of Agricultural and Food Chemistry* **50**, 4852–4860. doi: [10.1021/jf020214v](#).

Gu L, Kelm MA, Hammerstone JF, Beecher G, Holden J, Haytowitz D, Prior RL. 2003. Screening of foods containing proanthocyanidins and their structural characterization using LC-MS/MS and thiolytic degradation. *Journal of Agricultural and Food Chemistry* **51**, 7513–7521. doi: [10.1021/jf034815d](#).

Guidarelli M, Carbone F, Mourgues F, Perrotta G, Rosati C, Bertolini P, Baraldi E. 2011. *Colletotrichum acutatum* interactions with unripe and ripe strawberry fruits and differential responses at histological and transcriptional levels. *Plant Pathology* **60**, 685–697. doi: [10.1111/j.1365-3059.2010.02423.x](#).

Guo H, Wang X, Gundlach H, Mayer KFX, Peterson DG, Scheffler BE, Chee PW, Paterson AH. 2014. Extensive and biased intergenomic nonreciprocal DNA exchanges shaped a nascent polyploid genome, *Gossypium* (Cotton). *Genetics* **197**, 1153–1163. doi: [10.1534/genetics.114.166124](#).

Habashi R, Hacham Y, Dhakarey R, Matityahu I, Holland D, Tian L, Amir R. 2019. Elucidating the role of shikimate dehydrogenase in controlling the production of anthocyanins and hydrolysable tannins in the outer peels of pomegranate. *BMC Plant Biology* **19**, 476. doi: [10.1186/s12870-019-2042-1](#).

Haile ZM, Nagpala-De Guzman EG, Moretto M, Sonogo P, Engelen K, Zoli L, Moser C, Baraldi E. 2019. Transcriptome profiles of strawberry (*Fragaria vesca*) fruit interacting with *Botrytis cinerea* at different ripening stages. *Frontiers in Plant Science* **10**, 1131. doi: [10.3389/fpls.2019.01131](#).

Häkkinen SH, Kärenlampi SO, Heinonen IM, Mykkänen HM, Törrönen AR. 1999. Content of the flavonols quercetin, myricetin, and kaempferol in 25 edible berries. *Journal of Agricultural and Food Chemistry* **47**, 2274–2279. doi: [10.1021/jf9811065](#).

Hamberg M. 1999. An epoxy alcohol synthase pathway in higher plants: Biosynthesis of antifungal trihydroxy oxylipins in leaves of potato. *Lipids* **34**, 1131–1142. doi: [10.1007/s11745-999-0464-7](#).

Hammer KA, Carson CF, Riley T V. 2003. Antifungal activity of the components of

Melaleuca alternifolia (tea tree) oil. *Journal of Applied Microbiology* **95**, 853–860. doi: [10.1046/j.1365-2672.2003.02059.x](https://doi.org/10.1046/j.1365-2672.2003.02059.x).

Han Y, Chen C, Yan Z, Li J, Wang Y. 2019. The methyl jasmonate accelerates the strawberry fruits ripening process. *Scientia Horticulturae* **249**, 250–256. doi: [10.1016/j.scienta.2019.01.061](https://doi.org/10.1016/j.scienta.2019.01.061).

Han Y, Dang R, Li J, Jiang J, Zhang N, Jia M, Wei L, Li Z, Li B, Jia W. 2015. Sucrose nonfermenting1-related protein kinase2.6, an ortholog of open stomata1, is a negative regulator of strawberry fruit development and ripening. *Plant Physiology* **167**, 915–930. doi: [10.1104/pp.114.251314](https://doi.org/10.1104/pp.114.251314).

Han SY, Kitahata N, Saito T, Kobayashi M, Shinozaki K, Yoshida S, Asami T. 2004. A new lead compound for abscisic acid biosynthesis inhibitors targeting 9-cis-epoxycarotenoid dioxygenase. *Bioorganic and Medicinal Chemistry Letters* **14**, 3033–3036. doi: [10.1016/j.bmcl.2004.04.035](https://doi.org/10.1016/j.bmcl.2004.04.035).

Hanhineva K, Rogachev I, Kokko H, Mintz-Oron S, Venger I, Kärenlampi S,

Hardigan MA, Feldmann MJ, Lorant A, Bird KA, Famula R, Acharya C, Cole G, Edger PP, Knapp SJ. 2020. Genome synteny has been conserved among the octoploid progenitors of cultivated strawberry over millions of years of evolution. *Frontiers in Plant Science* **10**, 504896. doi: [10.3389/fpls.2019.01789](https://doi.org/10.3389/fpls.2019.01789).

Hasan N, Nester EW. 1978. Purification and properties of chorismate synthase from *Bacillus subtilis*. *Journal of Biological Chemistry* **253**, 4993–4998. doi: [10.1016/s0021-9258\(17\)34646-x](https://doi.org/10.1016/s0021-9258(17)34646-x).

Haslam E, Cai Y. 1994. Plant polyphenols (vegetable tannins): Gallic acid metabolism. *Natural Product Reports* **11**, 41–66. doi: [10.1039/NP9941100041](https://doi.org/10.1039/NP9941100041).

Hatanaka A. 1993. The biogenesis of green odour by green leaves. *Phytochemistry* **34**, 1201–1218. doi: [10.1016/0031-9422\(91\)80003-J](https://doi.org/10.1016/0031-9422(91)80003-J).

Hatanaka A, Kajiwarra T, Matsui K, Toyota H. 1992. Substrate specificity of tea leaf hydroperoxide lyase. *Zeitschrift für Naturforschung - Section C Journal of Biosciences* **47**, 677–679. doi: [10.1515/znc-1992-9-1006](https://doi.org/10.1515/znc-1992-9-1006).

Hatanaka A, Kajiwarra T, Sekiya J. 1987. Biosynthetic pathway for C6-aldehydes formation from linolenic acid in green leaves. *Chemistry and Physics of Lipids* **44**, 341–361. doi: [10.1016/0009-3084\(87\)90057-0](https://doi.org/10.1016/0009-3084(87)90057-0).

He M, Qin CX, Wang X, Ding NZ. 2020. Plant unsaturated fatty acids: Biosynthesis and regulation. *Frontiers in Plant Science* **11**, 511331. doi: [10.3389/fpls.2020.00390](https://doi.org/10.3389/fpls.2020.00390).

Heitz T, Smirnova E, Widemann E, Aubert Y, Pinot F, Ménard R. 2016. The rise and fall of jasmonate biological activities. *Subcellular Biochemistry* **86**, 405–426. doi: [10.1007/978-3-319-25979-6_16](https://doi.org/10.1007/978-3-319-25979-6_16).

Von Helversen O, Winkler L, Bestmann HJ. 2000. Sulphur-containing ‘perfumes’ attract flower-visiting bats. *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology* **186**, 143–153. doi: [10.1007/s003590050014](https://doi.org/10.1007/s003590050014).

Herrmann KM, Weaver LM. 1999. The shikimate pathway. *Annual Review of Plant Biology* **50**, 473–503. doi: [10.1146/annurev.arplant.50.1.473](https://doi.org/10.1146/annurev.arplant.50.1.473).

Hiltpold I, Turlings TCJ. 2012. Manipulation of chemically mediated interactions in

agricultural soils to enhance the control of crop pests and to improve crop yield. *Journal of Chemical Ecology* **38**, 641–650. doi: [10.1007/s10886-012-0131-9](https://doi.org/10.1007/s10886-012-0131-9).

Hincha DK, Zuther E, Hundertmark M, Heyer AG. 2006. The role of compatible solutes in plant freezing tolerance: A case study on raffinose. *Cold Hardiness in Plants: Molecular Genetics, Cell Biology, and Physiology*, 203–218. doi: [10.1079/9780851990590.0203](https://doi.org/10.1079/9780851990590.0203).

Hoffmann T, Kalinowski G, Schwab W. 2006. RNAi-induced silencing of gene expression in strawberry fruit (*Fragaria* × *ananassa*) by agroinfiltration: a rapid assay for gene function analysis. *The Plant Journal* **48**, 818–826. doi: [10.1111/j.1365-313x.2006.02913.x](https://doi.org/10.1111/j.1365-313x.2006.02913.x).

Hollister JD, Gaut BS. 2009. Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Research* **19**, 1419–1428. doi: [10.1101/gr.091678.109](https://doi.org/10.1101/gr.091678.109).

Hou BZ, Xu C, Shen YY. 2018. A leu-rich repeat receptor-like protein kinase, FaRIPK1, interacts with the ABA receptor, FaABAR, to regulate fruit ripening in strawberry. *Journal of Experimental Botany* **69**, 1569–1582. doi: [10.1093/jxb/erx488](https://doi.org/10.1093/jxb/erx488).

Howe GA, Major IT, Koo AJ. 2018. Modularity in Jasmonate Signaling for Multistress Resilience. *Annual Review of Plant Biology* **69**, 387–415. doi: [10.1146/annurev-arplant-042817-040047](https://doi.org/10.1146/annurev-arplant-042817-040047).

Howe GA, Schilmiller AL. 2002. Oxylin metabolism in response to stress. *Current Opinion in Plant Biology* **5**, 230–236. doi: [10.1016/S1369-5266\(02\)00250-9](https://doi.org/10.1016/S1369-5266(02)00250-9).

Hsieh MH, Chang CY, Hsu SJ, Chen JJ. 2008. Chloroplast localization of methylerythritol 4-phosphate pathway enzymes and regulation of mitochondrial genes in *ispD* and *ispE* albino mutants in *Arabidopsis*. *Plant Molecular Biology* **66**, 663–673. doi: [10.1007/s11103-008-9297-5](https://doi.org/10.1007/s11103-008-9297-5).

Huang K, Li M, Liu Y, et al. 2019. Functional analysis of 3-dehydroquinate dehydratase/shikimate dehydrogenases involved in shikimate pathway in *Camellia sinensis*. *Frontiers in Plant Science* **10**, 1268. doi: [10.3389/fpls.2019.01268](https://doi.org/10.3389/fpls.2019.01268).

Huang M, Sanchez-Moreiras AM, Abel C, Sohrabi R, Lee S, Gershenzon J, Tholl D. 2012. The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (*E*)- β -caryophyllene, is a defense against a bacterial pathogen. *New Phytologist* **193**, 997–1008. doi: [10.1111/j.1469-8137.2011.04001.x](https://doi.org/10.1111/j.1469-8137.2011.04001.x).

Huang YF, Vialet S, Guiraud JL, Torregrosa L, Bertrand Y, Cheynier V, This P, Terrier N. 2014. A negative MYB regulator of proanthocyanidin accumulation, identified through expression quantitative locus mapping in the grape berry. *New Phytologist* **201**, 795–809. doi: [10.1111/nph.12557](https://doi.org/10.1111/nph.12557).

Hufnagel JC, Hofmann T. 2008. Quantitative reconstruction of the nonvolatile sensometabolome of a red wine. *Journal of Agricultural and Food Chemistry* **56**, 9190–9199. doi: [10.1021/jf801742w](https://doi.org/10.1021/jf801742w).

Hughes RK, De Domenico S, Santino A. 2009. Plant cytochrome CYP74 family: Biochemical features, endocellular localisation, activation mechanism in plant defence and improvements for industrial applications. *ChemBioChem* **10**, 1122–1133. doi: [10.1002/cbic.200800633](https://doi.org/10.1002/cbic.200800633).

- Hughes J, Hughes MA.** 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* crantz) cotyledons. *Mitochondrial DNA* **5**, 41–49. doi: [10.3109/10425179409039703](https://doi.org/10.3109/10425179409039703).
- Hummer KE, Nathewet P, Yanagi T.** 2009. Decaploidy in *Fragaria iturupensis* (*Rosaceae*). *American Journal of Botany* **96**, 713–716. doi: [10.3732/ajb.0800285](https://doi.org/10.3732/ajb.0800285).
- Husson F, Josse J.** 2012. missMDA: Handling missing values with/in multivariate data analysis (principal component methods). R package version **1.2**. doi: [10.32614/cran-package.missmda](https://doi.org/10.32614/cran-package.missmda).
- Husson F, Josse J, Le S, Maintainer JM.** 2020. Multivariate exploratory data analysis and data mining. *Cran* **1**, 1–130. doi: [10.32614/cran.package.factominer](https://doi.org/10.32614/cran.package.factominer).
- Hyun J, Woo Y, Hwang DS, Jo G, Eom S, Lee Y, Park JC, Lim Y.** 2010. Relationships between structures of hydroxyflavones and their antioxidative effects. *Bioorganic and Medicinal Chemistry Letters* **20**, 5510–5513. doi: [10.1016/j.bmcl.2010.07.068](https://doi.org/10.1016/j.bmcl.2010.07.068).
- Ibanez F, Bang WY, Lombardini L, Cisneros-Zevallos L.** 2019. Solving the controversy of healthier organic fruit: Leaf wounding triggers distant gene expression response of polyphenol biosynthesis in strawberry fruit (*Fragaria x ananassa*). *Scientific Reports* **9**, 1–11. doi: [10.1038/s41598-019-55033-w](https://doi.org/10.1038/s41598-019-55033-w).
- Itkin M, Heinig U, Tzfadia O, et al.** 2013. Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science* **341**, 175–179. doi: [10.1126/science.1240230](https://doi.org/10.1126/science.1240230).
- Ito H, Li P, Koreishi M, Nagatomo A, Nishida N, Yoshida T.** 2014. Ellagitannin oligomers and a neolignan from pomegranate arils and their inhibitory effects on the formation of advanced glycation end products. *Food Chemistry* **152**, 323–330. doi: [10.1016/j.foodchem.2013.11.160](https://doi.org/10.1016/j.foodchem.2013.11.160).
- Ivanov I, Heydeck D, Hofheinz K, Roffeis J, O'Donnell VB, Kuhn H, Walther M.** 2010. Molecular enzymology of lipoxygenases. *Archives of Biochemistry and Biophysics* **503**, 161–174. doi: [10.1016/j.abb.2010.08.016](https://doi.org/10.1016/j.abb.2010.08.016).
- Jaakola L.** 2013. New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends in Plant Science* **18**, 477–483. doi: [10.1016/j.tplants.2013.06.003](https://doi.org/10.1016/j.tplants.2013.06.003).
- Jacopini S, Mariani M, de Caraffa VBB, Gambotti C, Vincenti S, Desjobert JM, Muselli A, Costa J, Berti L, Maury J.** 2016. Olive recombinant hydroperoxide lyase, an efficient biocatalyst for synthesis of green leaf volatiles. *Applied Biochemistry and Biotechnology* **179**, 671–683. doi: [10.1007/s12010-016-2023-x](https://doi.org/10.1007/s12010-016-2023-x).
- Janick J.** 2010. The origins of fruits, fruit growing, and fruit breeding. *Plant Breeding Reviews* **25**, 255–321. doi: [10.1002/9780470650301.ch8](https://doi.org/10.1002/9780470650301.ch8).
- Jetti RR, Yang E, Kurnianta A, Finn C, Qian MC.** 2007. Quantification of selected aroma-active compounds in strawberries by headspace solid-phase microextraction gas chromatography and correlation with sensory descriptive analysis. *Journal of Food Science* **72**, S487–S496. doi: [10.1111/j.1750-3841.2007.00445.x](https://doi.org/10.1111/j.1750-3841.2007.00445.x).
- Jia HF, Chai YM, Li CL, Lu D, Luo JJ, Qin L, Shen YY.** 2011. Absciscic acid plays an important role in the regulation of strawberry fruit ripening. *Plant Physiology* **157**, 188–199. doi: [10.1104/pp.111.177311](https://doi.org/10.1104/pp.111.177311).

- Jia M, Ding N, Zhang Q, et al.** 2017. A FERONIA-like receptor kinase regulates strawberry (*Fragaria × ananassa*) fruit ripening and quality formation. *Frontiers in Plant Science* **8**, 266186. doi: [10.3389/fpls.2017.01099](https://doi.org/10.3389/fpls.2017.01099).
- Jia HF, Lu D, Sun JH, Li CL, Xing Y, Qin L, Shen YY.** 2013. Type 2C protein phosphatase ABI1 is a negative regulator of strawberry fruit ripening. *Journal of Experimental Botany* **64**, 1677–1687. doi: [10.1093/jxb/ert028](https://doi.org/10.1093/jxb/ert028).
- Jia H, Zhang C, Pervaiz T, Zhao P, Liu Z, Wang B, Wang C, Zhang L, Fang J, Qian J.** 2016. Jasmonic acid involves in grape fruit ripening and resistant against *Botrytis cinerea*. *Functional and Integrative Genomics* **16**, 79–94. doi: [10.1007/s10142-015-0468-6](https://doi.org/10.1007/s10142-015-0468-6).
- Jiang N, Dillon FM, Silva A, Gomez-Cano L, Grotewold E.** 2021. Rhamnose in plants - from biosynthesis to diverse functions. *Plant Science* **302**, 110687. doi: [10.1016/j.plantsci.2020.110687](https://doi.org/10.1016/j.plantsci.2020.110687).
- Jiang Y, Ma X, Zhang W, Hu Q, Zheng J, Qian R.** 2023. Diversity and intraspecific variation of floral volatile compounds underscoring the terpenoids and methyl esters among the wild *Clematis* species and their cultivars. *Ornamental Plant Research* **3**. doi: [10.48130/opr-2023-0020](https://doi.org/10.48130/opr-2023-0020).
- Jiao Y, Wickett NJ, Ayyampalayam S, et al.** 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* **473**, 97–100. doi: [10.1038/nature09916](https://doi.org/10.1038/nature09916).
- Jin X, Du H, Zhu C, Wan H, Liu F, Ruan J, Mower JP, Zhu A.** 2023. Haplotype-resolved genomes of wild octoploid progenitors illuminate genomic diversifications from wild relatives to cultivated strawberry. *Nature Plants* 2023 9:8 **9**, 1252–1266. doi: [10.1038/s41477-023-01473-2](https://doi.org/10.1038/s41477-023-01473-2).
- Johnson AL, Govindarajulu R, Ashman TL.** 2014. Bioclimatic evaluation of geographical range in *Fragaria* (*Rosaceae*): Consequences of variation in breeding system, ploidy and species age. *Botanical Journal of the Linnean Society* **176**, 99–114. doi: [10.1111/boj.12190](https://doi.org/10.1111/boj.12190).
- Jones DT, Taylor WR, Thornton JM.** 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **8**, 275–282. doi: [10.1093/bioinformatics/8.3.275](https://doi.org/10.1093/bioinformatics/8.3.275).
- Jung S, Lee T, Cheng CH, et al.** 2019. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Research* **47**, D1137–D1145. doi: [10.1093/nar/gky1000](https://doi.org/10.1093/nar/gky1000).
- Jung S, Staton M, Lee T, Blenda A, Svancara R, Abbott A, Main D.** 2008. GDR (Genome Database for Rosaceae): Integrated web-database for *Rosaceae* genomics and genetics data. *Nucleic Acids Research* **36**, D1034–D1040. doi: [10.1093/nar/gkm803](https://doi.org/10.1093/nar/gkm803).
- Junker RR, Gershenzon J, Unsicker SB.** 2011. Floral odor bouquet loses its ant repellent properties after inhibition of terpene biosynthesis. *Journal of Chemical Ecology* **37**, 1323–1331. doi: [10.1007/s10886-011-0043-0](https://doi.org/10.1007/s10886-011-0043-0).
- Kaiser R.** 1994. Trapping, investigation and reconstitution of flower scents. *Perfumes*, 213–250. doi: [10.1007/978-94-011-3826-0_7](https://doi.org/10.1007/978-94-011-3826-0_7).
- Kajiwarra T, Sekiya J, Asano M, Hatanaka A.** 1982. Enantioselectivity of enzymatic cleavage reaction of 13-hydro-peroxylinolenic acid to C6-aldehyde and C12-oxoacid in

- tea chloroplasts. *Agricultural and Biological Chemistry* **46**, 3087–3088. doi: [10.1080/00021369.1982.10865560](https://doi.org/10.1080/00021369.1982.10865560).
- Kaminaga Y, Schnepf J, Peel G, et al.** 2006. Plant phenylacetaldehyde synthase is a bifunctional homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation. *Journal of Biological Chemistry* **281**, 23357–23366. doi: [10.1074/jbc.M602708200](https://doi.org/10.1074/jbc.M602708200).
- Kang C, Darwish O, Geretz A, Shahan R, Alkharouf N, Liu Z.** 2013. Genome-scale transcriptomic insights into early-stage fruit development in woodland strawberry *Fragaria vesca*. *Plant Cell* **25**, 1960–1978. doi: [10.1105/tpc.113.111732](https://doi.org/10.1105/tpc.113.111732).
- Karadag A, Ozcelik B, Saner S.** 2009. Review of methods to determine antioxidant capacities. *Food Analytical Methods* **2**, 41–60. doi: [10.1007/s12161-008-9067-7](https://doi.org/10.1007/s12161-008-9067-7).
- Karimi M, Inzé D, Depicker A.** GATEWAYTM vectors for *Agrobacterium*-mediated plant transformation. *Trends in Plant Science* **7**, 193–195. doi: [10.1016/s1360-1385\(02\)02251-3](https://doi.org/10.1016/s1360-1385(02)02251-3).
- Karlińska E, Masny A, Cieślak M, Macierzyński J, Pecio Ł, Stochmal A, Kosmala M.** 2021. Ellagitannins in roots, leaves, and fruits of strawberry (*Fragaria × ananassa* Duch.) vary with developmental stage and cultivar. *Scientia Horticulturae* **275**, 109665. doi: [10.1016/j.scienta.2020.109665](https://doi.org/10.1016/j.scienta.2020.109665).
- Kassambara A, Mundt F.** 2020. Extract and visualize the results of multivariate data analyses. *Cran*, 84. doi: [10.32614/cran.package.factoextra](https://doi.org/10.32614/cran.package.factoextra).
- Katoh K, Standley DM.** 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772–780. doi: [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010).
- Kazan K, Manners JM.** 2013. MYC2: The master in action. *Molecular Plant* **6**, 686–703. doi: [10.1093/mp/sss128](https://doi.org/10.1093/mp/sss128).
- Ke J, Ma H, Gu X, Thelen A, Brunzelle JS, Li J, Xu HE, Melcher K.** 2015. Structural basis for recognition of diverse transcriptional repressors by the TOPLESS family of corepressors. *Science Advances* **1**. doi: [10.1126/sciadv.1500107](https://doi.org/10.1126/sciadv.1500107).
- Keith B, Dong X, Ausubel FM, Fink GR.** 1991. Differential induction of 3-deoxy-*D*-arabino-heptulosonate 7-phosphate synthase genes in *Arabidopsis thaliana* by wounding and pathogenic attack. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 8821–8825. doi: [10.1073/pnas.88.19.8821](https://doi.org/10.1073/pnas.88.19.8821).
- Kim D, Langmead B, Salzberg SL.** 2015. HISAT: A fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357–360. doi: [10.1038/nmeth.3317](https://doi.org/10.1038/nmeth.3317).
- Kitamura S, Matsuda F, Tohge T, Yonekura-Sakakibara K, Yamazaki M, Saito K, Narumi I.** 2010. Metabolic profiling and cytological analysis of proanthocyanidins in immature seeds of *Arabidopsis thaliana* flavonoid accumulation mutants. *The Plant Journal* **62**, 549–559. doi: [10.1111/j.1365-3113X.2010.04174.x](https://doi.org/10.1111/j.1365-3113X.2010.04174.x).
- Kitamura S, Shikazono N, Tanaka A.** 2004. TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant Journal* **37**, 104–114. doi: [10.1046/j.1365-3113X.2003.01943.x](https://doi.org/10.1046/j.1365-3113X.2003.01943.x).
- Knudsen JT, Eriksson R, Gershenzon J, Ståhl B.** 2006. Diversity and distribution of

floral scent. *Botanical Review* **72**, 1–120. doi: [10.1663/0006-8101\(2006\)72\[1:DADOFS\]2.0.CO;2](https://doi.org/10.1663/0006-8101(2006)72[1:DADOFS]2.0.CO;2).

Koeduka T, Fridman E, Gang DR, et al. 2006. Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10128–10133. doi: [10.1073/pnas.0603732103](https://doi.org/10.1073/pnas.0603732103).

Koeduka T, Louie G V., Orlova I, et al. 2008. The multiple phenylpropene synthases in both *Clarkia breweri* and *Petunia hybrida* represent two distinct protein lineages. *Plant Journal* **54**, 362–374. doi: [10.1111/j.1365-3113X.2008.03412.x](https://doi.org/10.1111/j.1365-3113X.2008.03412.x).

Koes R, Verweij W, Quattrocchio F. 2005. Flavonoids: A colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* **10**, 236–242. doi: [10.1016/j.tplants.2005.03.002](https://doi.org/10.1016/j.tplants.2005.03.002).

Kolomiets M V., Hannapel DJ, Chen H, Tymeson M, Gladon RJ. 2001. Lipoxygenase is involved in the control of potato tuber development. *Plant Cell* **13**, 613–626. doi: [10.1105/tpc.13.3.613](https://doi.org/10.1105/tpc.13.3.613).

Koo AJK, Howe GA. 2012. Catabolism and deactivation of the lipid-derived hormone jasmonoyl-isoleucine. *Frontiers in Plant Science* **3**, 21250. doi: [10.3389/fpls.2012.00019](https://doi.org/10.3389/fpls.2012.00019).

de la Peña Moreno F, Monagas M, Blanch GP, Bartolomé B, Ruiz del Castillo ML. 2010. Enhancement of anthocyanins and selected aroma compounds in strawberry fruits through methyl jasmonate vapor treatment. *European Food Research and Technology* **230**, 989–999. doi: [10.1007/s00217-010-1243-5](https://doi.org/10.1007/s00217-010-1243-5).

Labadie M, Vallin G, Potier A, et al. 2022. High resolution quantitative trait *locus* mapping and whole genome sequencing enable the design of an anthocyanidin reductase-specific homoeo-allelic marker for fruit colour improvement in octoploid strawberry (*Fragaria* × *ananassa*). *Frontiers in Plant Science* **13**, 869655. doi: [10.3389/fpls.2022.869655](https://doi.org/10.3389/fpls.2022.869655).

Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyl transferases: Structures, functions, and mechanisms. *Annual Review of Biochemistry* **77**, 521–555. doi: [10.1146/annurev.biochem.76.061005.092322](https://doi.org/10.1146/annurev.biochem.76.061005.092322).

Landete JM. 2011. Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health. *Food Research International* **44**, 1150–1160. doi: [10.1016/j.foodres.2011.04.027](https://doi.org/10.1016/j.foodres.2011.04.027).

Landi L, Feliziani E, Romanazzi G. 2014. Expression of defense genes in strawberry fruits treated with different resistance inducers. *Journal of Agricultural and Food Chemistry* **62**, 3047–3056. doi: [10.1021/JF404423x](https://doi.org/10.1021/JF404423x).

Lange BM, Rujan T, Martin W, Croteau R. 2000. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 13172–13177. doi: [10.1073/pnas.240454797](https://doi.org/10.1073/pnas.240454797).

Larsen M, Poll L. 1990. Odour thresholds of some important aroma compounds in raspberries. *Zeitschrift für Lebensmittel-Untersuchung und -Forschung* **191**, 129–131. doi: [10.1007/bf01202638](https://doi.org/10.1007/bf01202638).

Lei JJ, Xue L, Guo RX, Dai HP. 2017. The *Fragaria* species native to China and their

geographical distribution. *Acta Horticulturae* **1156**, 37–45. doi: [10.17660/ActaHortic.2017.1156.5](https://doi.org/10.17660/ActaHortic.2017.1156.5).

Leone A, Melillo MT, Bleve-Zacheo T. 2001. Lipoxygenase in pea roots subjected to biotic stress. *Plant Science* **161**, 703–717. doi: [10.1016/S0168-9452\(01\)00458-7](https://doi.org/10.1016/S0168-9452(01)00458-7).

Lethe MCL, Paris V, Wang X, Chan CTY. 2024. Similarities in structure and function of UDP-glycosyltransferase homologs from human and plants. *International Journal of Molecular Sciences* 2024, Vol. 25, Page 2782 **25**, 2782. doi: [10.3390/ijms25052782](https://doi.org/10.3390/ijms25052782).

Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**, W78–W82. doi: [10.1093/nar/gkae268](https://doi.org/10.1093/nar/gkae268).

Li A, Liu D, Wu J, et al. 2014. mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* **26**, 1878–1900. doi: [10.1105/tpc.114.124388](https://doi.org/10.1105/tpc.114.124388).

Li Y, Pi M, Gao Q, Liu Z, Kang C. 2019. Updated annotation of the wild strawberry *Fragaria vesca* V4 genome. *Horticulture Research*. doi: [10.1038/s41438-019-0142-6](https://doi.org/10.1038/s41438-019-0142-6).

Li JT, Wang Q, Huang Yang M Di, et al. 2021. Parallel subgenome structure and divergent expression evolution of allo-tetraploid common carp and goldfish. *Nature Genetics* **53**, 1493–1503. doi: [10.1038/s41588-021-00933-9](https://doi.org/10.1038/s41588-021-00933-9).

Li Z Qi, Xie Q, Yan J Hui, Chen J Qing, Chen Q Xi. 2022. Genome-wide identification and characterization of the abiotic-stress-responsive lipoxygenase gene family in diploid woodland strawberry (*Fragaria vesca*). *Journal of Integrative Agriculture* **21**, 1982–1996. doi: [10.1016/S2095-3119\(21\)63819-4](https://doi.org/10.1016/S2095-3119(21)63819-4).

Li HM, Yu CW. 2018. Chloroplast galactolipids: The link between photosynthesis, chloroplast shape, jasmonates, phosphate starvation and freezing tolerance. *Plant and Cell Physiology* **59**, 1128–1134. doi: [10.1093/pcp/pcy088](https://doi.org/10.1093/pcp/pcy088).

Li C, Zha W, Li W, Wang J, You A. 2023. Advances in the biosynthesis of terpenoids and their ecological functions in plant resistance. *International Journal of Molecular Sciences* **24**, 11561. doi: [10.3390/ijms241411561](https://doi.org/10.3390/ijms241411561).

Liao X, Li M, Liu B, Yan M, Yu X, Zi H, Liu R, Yamamuro C. 2018. Interlinked regulatory loops of ABA catabolism and biosynthesis coordinate fruit growth and ripening in woodland strawberry. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E11542–E11550. doi: [10.1073/pnas.1812575115](https://doi.org/10.1073/pnas.1812575115).

Liavonchanka A, Feussner I. 2006. Lipoxygenases: Occurrence, functions and catalysis. *Journal of Plant Physiology* **163**, 348–357. doi: [10.1016/j.jplph.2005.11.006](https://doi.org/10.1016/j.jplph.2005.11.006).

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie TK, Espley R V., Hellens RP, Allan AC. 2010. An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in *Rosaceae*. *BMC Plant Biology* **10**, 1–17. doi: [10.1186/1471-2229-10-50](https://doi.org/10.1186/1471-2229-10-50).

Lin YT, Chen LJ, Herrfurth C, Feussner I, Li HM. 2016. Reduced biosynthesis of digalactosyldiacylglycerol, a major chloroplast membrane lipid, leads to oxylipin overproduction and phloem cap lignification in *Arabidopsis*. *Plant Cell* **28**, 219–232. doi: [10.1105/tpc.15.01002](https://doi.org/10.1105/tpc.15.01002).

- Liston A.** 2023. On the origin of strawberries. *Nature Plants* **9**, 1176–1177. doi: [10.1038/s41477-023-01488-9](https://doi.org/10.1038/s41477-023-01488-9).
- Liston A, Cronn R, Ashman TL.** 2014. *Fragaria*: A genus with deep historical roots and ripe for evolutionary and ecological insights. *American Journal of Botany* **101**, 1686–1699. doi: [10.3732/ajb.1400140](https://doi.org/10.3732/ajb.1400140).
- Liu XJ, An XH, Liu X, Hu DG, Wang XF, You CX, Hao YJ.** 2017. MdSnRK1.1 interacts with MdJAZ18 to regulate sucrose-induced anthocyanin and proanthocyanidin accumulation in apple. *Journal of Experimental Botany* **68**, 2977–2990. doi: [10.1093/jxb/erx150](https://doi.org/10.1093/jxb/erx150).
- Liu Y, Du M, Deng L, Shen J, Fang M, Chen Q, Lu Y, Wang Q, Li C, Zhai Q.** 2019. Myc2 regulates the termination of jasmonate signaling via an autoregulatory negative feedback loop. *Plant Cell* **31**, 106–127. doi: [10.1105/tpc.18.00405](https://doi.org/10.1105/tpc.18.00405).
- Liu Y, Gao L, Liu L, Yang Q, Lu Z, Nie Z, Wang Y, Xia T.** 2012. Purification and characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant (*Camellia sinensis*). *Journal of Biological Chemistry* **287**, 44406–44417. doi: [10.1074/jbc.M112.403071](https://doi.org/10.1074/jbc.M112.403071).
- Liu HT, Ji Y, Liu Y, et al.** 2020. The sugar transporter system of strawberry: genome-wide identification and expression correlation with fruit soluble sugar-related traits in a *Fragaria* × *ananassa* germplasm collection. *Horticulture Research* 2020 7:1 **7**, 1–19. doi: [10.1038/s41438-020-00359-0](https://doi.org/10.1038/s41438-020-00359-0).
- Liu T, Li M, Liu Z, Ai X, Li Y.** 2021. Reannotation of the cultivated strawberry genome and establishment of a strawberry genome database. *Horticulture Research* **8**, 1–9. doi: [10.1038/s41438-021-00476-4](https://doi.org/10.1038/s41438-021-00476-4).
- Liu Z, Liang T, Kang C.** 2023. Molecular bases of strawberry fruit quality traits: Advances, challenges, and opportunities. *Plant Physiology* **193**, 900–914. doi: [10.1093/plphys/kiad376](https://doi.org/10.1093/plphys/kiad376).
- Liu Y, Seeram NP.** 2018. Liquid chromatography coupled with time-of-flight tandem mass spectrometry for comprehensive phenolic characterization of pomegranate fruit and flower extracts used as ingredients in botanical dietary supplements. *Journal of Separation Science* **41**, 3022–3033. doi: [10.1002/jssc.201800480](https://doi.org/10.1002/jssc.201800480).
- Livak KJ, Schmittgen TD.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **25**, 402–408. doi: [10.1006/meth.2001.1262](https://doi.org/10.1006/meth.2001.1262).
- Loreto F, Dicke M, Schnitzler JP, Turlings TCJ.** 2014. Plant volatiles and the environment. *Plant, Cell and Environment* **37**, 1905–1908. doi: [10.1111/pce.12369](https://doi.org/10.1111/pce.12369).
- Lu W, Chen J, Ren X, Yuan J, Han X, Mao L, Ying T, Luo Z.** 2018. One novel strawberry MADS-box transcription factor FaMADS1a acts as a negative regulator in fruit ripening. *Scientia Horticulturae* **227**, 124–131. doi: [10.1016/j.scienta.2017.09.042](https://doi.org/10.1016/j.scienta.2017.09.042).
- Lu H, Luo Z, Li D, Jiang Y, Li L.** 2021a. FaMYB11 promotes the accumulation of volatile esters by regulating FaLOX5 during strawberry (*Fragaria* × *ananassa*) ripening. *Postharvest Biology and Technology* **178**, 111560. doi: [10.1016/j.postharvbio.2021.111560](https://doi.org/10.1016/j.postharvbio.2021.111560).
- Lu H, Luo Z, Wang L, Liu W, Li D, Belwal T, Xu Y, Li L.** 2020. FaMYB9 is involved

in the regulation of C6 volatile biosynthesis in strawberry. *Plant Science* **293**, 110422. doi: [10.1016/j.plantsci.2020.110422](https://doi.org/10.1016/j.plantsci.2020.110422).

Lu W, Shi Y, Wang R, Su D, Tang M, Liu Y, Li Z. 2021b. Antioxidant activity and healthy benefits of natural pigments in fruits: A review. *International Journal of Molecular Sciences* **22**, 4945. doi: [10.3390/ijms22094945](https://doi.org/10.3390/ijms22094945).

Luedemann A, Von Malotky L, Erban A, Kopka J. 2012. TagFinder: Preprocessing software for the fingerprinting and the profiling of gas chromatography-mass spectrometry based metabolome analyses. *Methods in Molecular Biology* **860**, 255–286. doi: [10.1007/978-1-61779-594-7_16](https://doi.org/10.1007/978-1-61779-594-7_16).

Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I. 2006. Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5224–5229. doi: [10.1073/pnas.0510791103](https://doi.org/10.1073/pnas.0510791103).

Ma W, Guo A, Zhang Y, Wang H, Liu Y, Li H. 2014. A review on astringency and bitterness perception of tannins in wine. *Trends in Food Science and Technology* **40**, 6–19. doi: [10.1016/j.tifs.2014.08.001](https://doi.org/10.1016/j.tifs.2014.08.001).

Maeda H, Dudareva N. 2012. The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annual Review of Plant Biology* **63**, 73–105. doi: [10.1146/annurev-arplant-042811-105439](https://doi.org/10.1146/annurev-arplant-042811-105439).

Maeda H, Shasany AK, Schnepf J, Orlova I, Taguchi G, Cooper BR, Rhodes D, Pichersky E, Dudareva N. 2010. RNAi suppression of Arogenate Dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in *Petunia* petals. *Plant Cell* **22**, 832–849. doi: [10.1105/tpc.109.073247](https://doi.org/10.1105/tpc.109.073247).

Maharjan R, Fukuda Y, Shimomura N, Nakayama T, Okimoto Y, Kawakami K, Nakayama T, Hamada H, Inoue T, Ozaki SI. 2020. An ambidextrous polyphenol glycosyltransferase PaGT2 from *Phytolacca americana*. *Biochemistry* **59**, 2551–2561. doi: [10.1021/acs.biochem.0c00224](https://doi.org/10.1021/acs.biochem.0c00224).

Mao J, Wang Y, Wang B, et al. 2023. High-quality haplotype-resolved genome assembly of cultivated octoploid strawberry. *Horticulture Research* **10**. doi: [10.1093/hr/uhad002](https://doi.org/10.1093/hr/uhad002).

Martín-Pizarro C, Franco-Zorrilla JM, Perotti MF, Lozano-Durán R, Qin G, Posé D. 2025. FaRIF at the core of strawberry fruit ripening: Deciphering its targets and interaction networks. *bioRxiv*. doi: [10.1101/2025.01.13.631932](https://doi.org/10.1101/2025.01.13.631932).

Martín-Pizarro C, Vallarino JG, Osorio S, et al. 2021. The NAC transcription factor FaRIF controls fruit ripening in strawberry. *Plant Cell* **33**, 1574–1593. doi: [10.1093/plcell/koab070](https://doi.org/10.1093/plcell/koab070).

Matoba T, Hidaka H, Kitamura K, Kaizuma N, Kito M. 1985. Contribution of hydroperoxide lyase activity to n-hexanal formation in soybean. *Journal of Agricultural and Food Chemistry* **33**, 856–858. doi: [10.1021/jf00065a022](https://doi.org/10.1021/jf00065a022).

Matsuba Y, Nguyen TTH, Wiegert K, et al. 2013. Evolution of a complex locus for terpene biosynthesis in *Solanum*. *Plant Cell* **25**, 2022–2036. doi: [10.1105/tpc.113.111013](https://doi.org/10.1105/tpc.113.111013).

Matsuba Y, Zi J, Jones AD, Peters RJ, Pichersky E. 2015. Biosynthesis of the diterpenoid lycosantalol via neryleryl diphosphate in *Solanum lycopersicum*. *PLOS*

ONE **10**, e0119302. doi: [10.1371/journal.pone.0119302](https://doi.org/10.1371/journal.pone.0119302).

Matsui K, Sugimoto K, Mano J, Ozawa R, Takabayashi J. 2012. Differential metabolisms of green leaf volatiles in injured and intact parts of a wounded leaf meet distinct ecophysiological requirements. *PLOS ONE* **7**, e36433. doi: [10.1371/journal.pone.0036433](https://doi.org/10.1371/journal.pone.0036433).

Maugeri A, Lombardo GE, Cirimi S, Süntar I, Barreca D, Laganà G, Navarra M. 2022. Pharmacology and toxicology of tannins. *Archives of Toxicology* **96**, 1257–1277. doi: [10.1007/s00204-022-03250-0](https://doi.org/10.1007/s00204-022-03250-0).

McClintock B. 1984. The significance of responses of the genome to challenge. *Science* **226**, 792–801. doi: [10.1126/science.15739260](https://doi.org/10.1126/science.15739260).

McGarvey DJ, Croteau R. 1995. Terpenoid metabolism. *The Plant Cell* **7**, 1015–1026. doi: [10.1105/TPC.7.7.1015](https://doi.org/10.1105/TPC.7.7.1015).

Medema MH, Kottmann R, Yilmaz P, et al. 2015. Minimum information about a biosynthetic gene cluster. *Nature Chemical Biology* **11**, 625–631. doi: [10.1038/nchembio.1890](https://doi.org/10.1038/nchembio.1890).

Medina-Puche L, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodríguez-Franco A, Caballero JL, Schwab W, Muñoz-Blanco J, Blanco-Portales R. 2014. MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria* × *ananassa* fruits. *Journal of Experimental Botany* **65**, 401–417. doi: [10.1093/jxb/ert377](https://doi.org/10.1093/jxb/ert377).

Medina-Puche L, Martínez-Rivas FJ, Molina-Hidalgo FJ, Mercado JA, Moyano E,

Meech R, Miners JO, Lewis BC, MacKenzie PI. 2012. The glycosidation of xenobiotics and endogenous compounds: Versatility and redundancy in the UDP-glycosyltransferase superfamily. *Pharmacology and Therapeutics* **134**, 200–218. doi: [10.1016/j.pharmthera.2012.01.009](https://doi.org/10.1016/j.pharmthera.2012.01.009).

Melcher K, Ng LM, Zhou XE, et al. 2009. A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. *Nature* **462**, 602–608. doi: [10.1038/nature08613](https://doi.org/10.1038/nature08613).

Melotto M, Mecey C, Niu Y, et al. 2008. A critical role of two positively charged amino acids in the Jas motif of *Arabidopsis* JAZ proteins in mediating coronatine- and jasmonoyl isoleucine-dependent interactions with the COI1 F-box protein. *Plant Journal* **55**, 979–988. doi: [10.1111/j.1365-313X.2008.03566.x](https://doi.org/10.1111/j.1365-313X.2008.03566.x).

Mena P, Calani L, Dall'Asta C, Galaverna G, García-Viguera C, Bruni R, Crozier A, Del Rio D. 2012. Rapid and comprehensive evaluation of (poly)phenolic compounds in pomegranate (*Punica granatum* L.) Juice by UHPLC-MSn. *Molecules* **17**, 14821–14840. doi: [10.3390/molecules171214821](https://doi.org/10.3390/molecules171214821).

Ménager I, Jost M, Aubert C. 2004. Changes in physicochemical characteristics and volatile constituents of strawberry (cv. Cigaline) during maturation. *Journal of Agricultural and Food Chemistry* **52**, 1248–1254. doi: [10.1021/jf0350919](https://doi.org/10.1021/jf0350919).

Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* **32**, e4792. doi: [10.1002/pro.4792](https://doi.org/10.1002/pro.4792).

Mi SK, Song MC, Eun YK, Yang JI, Hwangbo H, Young CK, Ryu CM, Kwang YY,

- Gap CC, Baik HC.** 2008. Galactinol is a signaling component of the induced systemic resistance caused by *Pseudomonas chlororaphis* O6 root colonization. *Molecular Plant-Microbe Interactions* **21**, 1643–1653. doi: [10.1094/MPMI-21-12-1643](https://doi.org/10.1094/MPMI-21-12-1643).
- Milne RI, Abbott RJ.** 2008. Reproductive isolation among two interfertile *Rhododendron* species: Low frequency of post-F1 hybrid genotypes in alpine hybrid zones. *Molecular Ecology* **17**, 1108–1121. doi: [10.1111/j.1365-294X.2007.03643.x](https://doi.org/10.1111/j.1365-294X.2007.03643.x).
- Min SS, Dong GK, Lee SH.** 2005. Isolation and characterization of a jasmonic acid carboxyl methyltransferase gene from hot pepper (*Capsicum annuum* L.). *Journal of Plant Biology* **48**, 292–297. doi: [10.1007/bf03030525](https://doi.org/10.1007/bf03030525).
- Minh BQ, Nguyen MAT, Von Haeseler A.** 2013. Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* **30**, 1188–1195. doi: [10.1093/molbev/mst024](https://doi.org/10.1093/molbev/mst024).
- Minor W, Bolin JT, Steczko J, Axelrod B, Otwinowski Z.** 1993. Crystallographic determination of the active site iron and its ligands in soybean lipoxygenase L-1. *Biochemistry* **32**, 6320–6323. doi: [10.1021/bi00076a003](https://doi.org/10.1021/bi00076a003).
- Minor W, Steczko J, Stec B, Otwinowski Z, Bolin JT, Walter R, Axelrod B.** 1996. Crystal structure of soybean lipoxygenase L-1 at 1.4 Å resolution. *Biochemistry* **35**, 10687–10701. doi: [10.1021/bi960576u](https://doi.org/10.1021/bi960576u).
- Mir R, Jallu S, Singh TP.** 2015. The shikimate pathway: Review of amino acid sequence, function and three-dimensional structures of the enzymes. *Critical Reviews in Microbiology* **41**, 172–189. doi: [10.3109/1040841X.2013.813901](https://doi.org/10.3109/1040841X.2013.813901).
- Misnawi, Jinap S, Jamilah B, Nazamid S.** 2005. Changes in polyphenol ability to produce astringency during roasting of cocoa liquor. *Journal of the Science of Food and Agriculture* **85**, 917–924. doi: [10.1002/jsfa.1954](https://doi.org/10.1002/jsfa.1954).
- Mittler R.** 2017. ROS Are Good. *Trends in Plant Science* **22**, 11–19. doi: [10.1016/j.tplants.2016.08.002](https://doi.org/10.1016/j.tplants.2016.08.002).
- Moilanen J, Karonen M, Tähtinen P, Jacquet R, Quideau S, Salminen JP.** 2016. Biological activity of ellagitannins: Effects as anti-oxidants, pro-oxidants and metal chelators. *Phytochemistry* **125**, 65–72. doi: [10.1016/j.phytochem.2016.02.008](https://doi.org/10.1016/j.phytochem.2016.02.008).
- Molina-Hidalgo FJ, Franco AR, Villatoro C, Medina-Puche L, Mercado JA, Hidalgo MA, Monfort A, Caballero JL, Muñoz-Blanco J, Blanco-Portales R.** 2013. The strawberry (*Fragaria×ananassa*) fruit-specific rhamnogalacturonate lyase 1 (*FaRGLyase1*) gene encodes an enzyme involved in the degradation of cell-wall middle lamellae. *Journal of Experimental Botany* **64**, 1471–1483. doi: [10.1093/jxb/ers386](https://doi.org/10.1093/jxb/ers386).
- Molina-Hidalgo FJ, Medina-Puche L, Cañete-Gómez C, et al.** 2017. The fruit-specific transcription factor FaDOF2 regulates the production of eugenol in ripe fruit receptacles. *Journal of Experimental Botany* **68**, 4529–4543. doi: [10.1093/jxb/erx257](https://doi.org/10.1093/jxb/erx257).
- Mora J, Pott DM, Osorio S, Vallarino JG.** 2022. Regulation of Plant Tannin Synthesis in Crop Species. *Frontiers in Genetics* **0**, 887. doi: [10.3389/fgene.2022.870976](https://doi.org/10.3389/fgene.2022.870976).
- Morales-Quintana L, Ramos P.** 2019. Chilean strawberry (*Fragaria chiloensis*): An integrative and comprehensive review. *Food Research International* **119**, 769–776. doi: [10.1016/j.foodres.2018.10.059](https://doi.org/10.1016/j.foodres.2018.10.059).

- Mosblech A, Feussner I, Heilmann I.** 2009. Oxylipins: Structurally diverse metabolites from fatty acid oxidation. *Plant Physiology and Biochemistry* **47**, 511–517. doi: [10.1016/j.plaphy.2008.12.011](https://doi.org/10.1016/j.plaphy.2008.12.011).
- Moya-León MA, Mattus-Araya E, Herrera R.** 2019. Molecular events occurring during softening of strawberry fruit. *Frontiers in Plant Science* **10**, 452316. doi: [10.3389/fpls.2019.00615](https://doi.org/10.3389/fpls.2019.00615).
- Muir RM, Ibáñez AM, Uratsu SL, et al.** 2011. Mechanism of gallic acid biosynthesis in bacteria (*Escherichia coli*) and walnut (*Juglans regia*). *Plant Molecular Biology* **75**, 555–565. doi: [10.1007/s11103-011-9739-3](https://doi.org/10.1007/s11103-011-9739-3).
- Muñoz P, Castillejo C, Gómez JA, Miranda L, Lesemann S, Olbricht K, Petit A, Chartier P, Haugeneder A, Trinkl J, Mazzoni L, Masny A, Zurawicz E, Ziegler F, Usadel B, Schwab W, Denoyes B, Mezzeti B, Osorio S, Sánchez-Sevilla JF, Amaya I.** 2023. QTL analysis for ascorbic acid content in strawberry fruit reveals a complex genetic architecture and association with GDP-L-galactose phosphorylase. *Horticulture Research* **10**. doi: [10.1093/hr/uhad006](https://doi.org/10.1093/hr/uhad006).
- Murphy PA.** 2008. Soybean proteins. *Soybeans: Chemistry, Production, Processing, and Utilization*, 229–267. doi: [10.1016/b978-1-893997-64-6.50011-1](https://doi.org/10.1016/b978-1-893997-64-6.50011-1).
- Mwenda CM, Matsui K.** 2014. The importance of lipoxygenase control in the production of green leaf volatiles by lipase-dependent and independent pathways. *Plant Biotechnology* **31**, 445–452. doi: [10.5511/plantbiotechnology.14.0924a](https://doi.org/10.5511/plantbiotechnology.14.0924a).
- Nagpala EG, Guidarelli M, Gasperotti M, Masuero D, Bertolini P, Vrhovsek U, Baraldi E.** 2016. Polyphenols variation in fruits of the susceptible strawberry cultivar Alba during ripening and upon fungal pathogen interaction and possible involvement in unripe fruit tolerance. *Journal of Agricultural and Food Chemistry* **64**, 1869–1878. doi: [10.1021/acs.jafc.5b06005](https://doi.org/10.1021/acs.jafc.5b06005).
- Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T.** 2007. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *Journal of Bioscience and Bioengineering* **104**, 34–41. doi: [10.1263/jbb.104.34](https://doi.org/10.1263/jbb.104.34).
- Nakamura A, Shimada H, Masuda T, Ohta H, Takamiya K ichiro.** 2001. Two distinct isopentenyl diphosphate isomerases in cytosol and plastid are differentially induced by environmental stresses in tobacco. *FEBS Letters* **506**, 61–64. doi: [10.1016/S0014-5793\(01\)02870-8](https://doi.org/10.1016/S0014-5793(01)02870-8).
- Nakashima K, Fujita Y, Kanamori N, et al.** 2009. Three arabidopsis SnRK2 protein kinases, SRK2D/SnRK2.2, SRK2E/SnRK2.6/OST1 and SRK2I/SnRK2.3, involved in ABA signaling are essential for the control of seed development and dormancy. *Plant and Cell Physiology* **50**, 1345–1363. doi: [10.1093/pcp/pcp083](https://doi.org/10.1093/pcp/pcp083).
- Nambara E, Marion-Poll A.** 2005. Absciscic acid biosynthesis and catabolism. *Annual Review of Plant Biology* **56**, 165–185. doi: [10.1146/annurev.arplant.56.032604.144046](https://doi.org/10.1146/annurev.arplant.56.032604.144046).
- Nanda S, Yadav JS.** 2003. Lipoxygenase biocatalysis: A survey of asymmetric oxygenation. *Journal of Molecular Catalysis B: Enzymatic* **26**, 3–28. doi: [10.1016/S1381-1177\(03\)00146-2](https://doi.org/10.1016/S1381-1177(03)00146-2).

- Nardi CF, Villarreal NM, Dotto MC, Ariza MT, Vallarino JG, Martínez GA, Valpuesta V, Civello PM.** 2016. Influence of plant growth regulators on Expansin2 expression in strawberry fruit. Cloning and functional analysis of *FaEXP2* promoter region. *Postharvest Biology and Technology* **114**, 17–28. doi: [10.1016/j.postharvbio.2015.11.008](https://doi.org/10.1016/j.postharvbio.2015.11.008).
- Nardi CF, Villarreal NM, Opazo MC, Martínez GA, Moya-León MA, Civello PM.** 2014. Expression of *FaXTH1* and *FaXTH2* genes in strawberry fruit. Cloning of promoter regions and effect of plant growth regulators. *Scientia Horticulturae* **165**, 111–122. doi: [10.1016/j.scienta.2013.10.035](https://doi.org/10.1016/j.scienta.2013.10.035).
- Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L.** 2001. The *Arabidopsis* TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *The Plant Cell* **13**, 2099–2114. doi: [10.1105/tpc.010098](https://doi.org/10.1105/tpc.010098).
- Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ.** 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**, 268–274. doi: [10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300).
- Niemetz R, Gross GG.** 2003a. Oxidation of pentagalloylglucose to the ellagitannin, tellimagrandin II, by a phenol oxidase from *Tellima grandiflora* leaves. *Phytochemistry* **62**, 301–306. doi: [10.1016/S0031-9422\(02\)00557-5](https://doi.org/10.1016/S0031-9422(02)00557-5).
- Niemetz R, Gross GG.** 2003b. Ellagitannin biosynthesis: Laccase-catalyzed dimerization of tellimagrandin II to cornusiin E in *Tellima grandiflora*. *Phytochemistry* **64**, 1197–1201. doi: [10.1016/j.phytochem.2003.08.013](https://doi.org/10.1016/j.phytochem.2003.08.013).
- Niemetz R, Gross GG.** 2005. Enzymology of gallotannin and ellagitannin biosynthesis. *Phytochemistry* **66**, 2001–2011. doi: [10.1016/j.phytochem.2005.01.009](https://doi.org/10.1016/j.phytochem.2005.01.009).
- Niemetz R, Schilling G, Gross GG.** 2003. Biosynthesis of the dimeric ellagitannin, cornusiin E, in *Tellima grandiflora*. *Phytochemistry* **64**, 109–114. doi: [10.1016/S0031-9422\(03\)00280-2](https://doi.org/10.1016/S0031-9422(03)00280-2).
- Nishiuchi T, Hamada T, Kodama H, Iba K.** 1997. Wounding changes the spatial expression pattern of the *Arabidopsis* plastid ω -3 fatty acid desaturase gene (*FAD7*) through different signal transduction pathways. *Plant Cell* **9**, 1701–1712. doi: [10.1105/tpc.9.10.1701](https://doi.org/10.1105/tpc.9.10.1701).
- Njuguna W, Liston A, Cronn R, Ashman TL, Bassil N.** 2013. Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Molecular Phylogenetics and Evolution* **66**, 17–29. doi: [10.1016/j.ympev.2012.08.026](https://doi.org/10.1016/j.ympev.2012.08.026).
- Nolan TM, Vukasinović N, Liu D, Russinova E, Yin Y.** 2020. Brassinosteroids: Multidimensional regulators of plant growth, development, and stress responses. *Plant Cell* **32**, 298–318. doi: [10.1105/tpc.19.00335](https://doi.org/10.1105/tpc.19.00335).
- Noordermeer MA, Veldink GA, Vliegthart JFG.** 2001. Fatty acid hydroperoxide lyase: A plant cytochrome P450 enzyme involved in wound healing and pest resistance. *ChemBioChem* **2**, 494–504. doi: [10.1002/1439-7633\(20010803\)2:7/8<494::aid-cbic494>3.0.co;2-1](https://doi.org/10.1002/1439-7633(20010803)2:7/8<494::aid-cbic494>3.0.co;2-1).
- Nützmann HW, Huang A, Osbourn A.** 2016. Plant metabolic clusters – from genetics to genomics. *New Phytologist* **211**, 771–789. doi: [10.1111/nph.13981](https://doi.org/10.1111/nph.13981).

- O'Donnell PJ, Calvert C, Atzorn R, Wasternack C, Leyser HMO, Bowles DJ.** 1996. Ethylene as a signal mediating the wound response of tomato plants. *Science* **274**, 1914–1917. doi: [10.1126/science.274.5294.1914](https://doi.org/10.1126/science.274.5294.1914).
- Okuda T, Ito H.** 2011. Tannins of constant structure in medicinal and food plants—hydrolyzable tannins and polyphenols related to tannins. *Molecules* **16**, 2191–2217. doi: [10.3390/molecules16032191](https://doi.org/10.3390/molecules16032191).
- Olivoto T, Lúcio ADC.** 2020. metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution* **11**, 783–789. doi: [10.1111/2041-210X.13384](https://doi.org/10.1111/2041-210X.13384).
- Oliw EH, Hamberg M.** 2017. An allene oxide and 12-oxophytodienoic acid are key intermediates in jasmonic acid biosynthesis by *Fusarium oxysporum*. *Journal of Lipid Research* **58**, 1670–1680. doi: [10.1194/jlr.M077305](https://doi.org/10.1194/jlr.M077305).
- Ono NN, Qin X, Wilson AE, Li G, Tian L.** 2016. Two UGT84 family glycosyltransferases catalyze a critical reaction of hydrolyzable tannin biosynthesis in pomegranate (*Punica granatum*). (C Xu, Ed.). *PLoS ONE* **11**, e0156319. doi: [10.1371/journal.pone.0156319](https://doi.org/10.1371/journal.pone.0156319).
- Van Ooijen JW.** 2011. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genetics Research* **93**, 343–349. doi: [10.1017/S0016672311000279](https://doi.org/10.1017/S0016672311000279).
- Opazo MC, Lizana R, Pimentel P, Herrera R, Moya-León MA.** 2013. Changes in the mRNA abundance of *FcXTH1* and *FcXTH2* promoted by hormonal treatments of *Fragaria chiloensis* fruit. *Postharvest Biology and Technology* **77**, 28–34. doi: [10.1016/j.postharvbio.2012.11.007](https://doi.org/10.1016/j.postharvbio.2012.11.007).
- Orlova I, Marshall-Colón A, Schnepf J, et al.** 2006. Reduction of benzenoid synthesis in petunia flowers reveals multiple pathways to benzoic acid and enhancement in auxin transport. *Plant Cell* **18**, 3458–3475. doi: [10.1105/tpc.106.046227](https://doi.org/10.1105/tpc.106.046227).
- Orlova I, Nagegowda DA, Kish CM, et al.** 2009. The small subunit of snapdragon geranyl diphosphate synthase modifies the chain length specificity of tobacco geranylgeranyl diphosphate synthase in planta. *Plant Cell* **21**, 4002–4017. doi: [10.1105/tpc.109.071282](https://doi.org/10.1105/tpc.109.071282).
- Osmani SA, Bak S, Imberty A, Olsen CE, Møller BL.** 2008. Catalytic key amino acids and UDP-sugar donor specificity of a plant glucuronosyltransferase, UGT94B1: Molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. *Plant Physiology* **148**, 1295–1308. doi: [10.1104/pp.108.128256](https://doi.org/10.1104/pp.108.128256).
- Ozawa T, Lilley TH, Haslam E.** 1987. Polyphenol interactions: astringency and the loss of astringency in ripening fruit. *Phytochemistry* **26**, 2937–2942. doi: [10.1016/S0031-9422\(00\)84566-5](https://doi.org/10.1016/S0031-9422(00)84566-5).
- Padilla MN, Hernández ML, Pérez AG, Sanz C, Martínez-Rivas JM.** 2010. Isolation, expression, and characterization of a 13-hydroperoxide lyase gene from olive fruit related to the biosynthesis of the main virgin olive oil aroma compounds. *Journal of Agricultural and Food Chemistry* **58**, 5649–5657. doi: [10.1021/jf9045396](https://doi.org/10.1021/jf9045396).
- Palmieri-Thiers C, Canaan S, Brunini V, Lorenzi V, Tomi F, Desseyn JL, Garscha U, Oliw EH, Berti L, Maury J.** 2009. A lipoxygenase with dual positional specificity is expressed in olives (*Olea europaea* L.) during ripening. *Biochimica et Biophysica Acta* -

Molecular and Cell Biology of Lipids **1791**, 339–346. doi: [10.1016/j.bbalip.2009.02.012](https://doi.org/10.1016/j.bbalip.2009.02.012).

Paolocci F, Robbins MP, Passeri V, Hauck B, Morris P, Rubini A, Arcioni S, Damiani F. 2011. The strawberry transcription factor FaMYB1 inhibits the biosynthesis of proanthocyanidins in *Lotus corniculatus* leaves. *Journal of Experimental Botany* **62**, 1189–1200. doi: [10.1093/jxb/erq344](https://doi.org/10.1093/jxb/erq344).

Park J Il, Lee YK, Chung W Il, Lee IH, Choi JH, Lee WM, Ezura H, Lee SP, Kim IJ. 2006. Modification of sugar composition in strawberry fruit by antisense suppression of an ADP-glucose pyrophosphorylase. *Molecular Breeding* **17**, 269–279. doi: [10.1007/s11032-005-5682-9](https://doi.org/10.1007/s11032-005-5682-9).

Parkin IAP, Koh C, Tang H, et al. 2014. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. *Genome Biology* **15**, 1–18. doi: [10.1186/gb-2014-15-6-r77](https://doi.org/10.1186/gb-2014-15-6-r77).

Pascual O, Vignault A, Gombau J, Navarro M, Gómez-Alonso S, García-Romero E, Canals JM, Hermosín-Gutiérrez I, Teissedre PL, Zamora F. 2017. Oxygen consumption rates by different oenological tannins in a model wine solution. *Food Chemistry* **234**, 26–32. doi: [10.1016/j.foodchem.2017.04.148](https://doi.org/10.1016/j.foodchem.2017.04.148).

Pauketat TR, Kelly LS, Fritz GJ, Lopinot NH, Elias S, Hargrave E. 2002. The Residues of feasting and public ritual at early Cahokia. *American Antiquity* **67**, 257–279. doi: [10.2307/2694566](https://doi.org/10.2307/2694566).

Pauwels L, Barbero GF, Geerinck J, et al. 2010. NINJA connects the co-repressor TOPLESS to jasmonate signalling. *Nature* **464**, 788–791. doi: [10.1038/nature08854](https://doi.org/10.1038/nature08854).

Van De Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics* **18**, 411–424. doi: [10.1038/nrg.2017.26](https://doi.org/10.1038/nrg.2017.26).

Peleg H, Gacon K, Schlich P, Noble AC. 1999. Bitterness and astringency of flavan-3-ol monomers, dimers and trimers. *Journal of the Science of Food and Agriculture* **79**, 1123–1128. doi: [10.1002/\(SICI\)1097-0010\(199906\)79:8<1123::AID-jsfa336>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0010(199906)79:8<1123::AID-jsfa336>3.0.CO;2-D).

Pérez-Martín JE. 2024. GitHub - jjoseenrique/Metabolomics: Custom metabolomic tools. Accessed February 2025. <https://github.com/jjoseenrique/Metabolomics>.

Perla V, Jayanty SS, Holm DG, Davidson RD. 2014. Relationship between tuber storage proteins and tuber powdery scab resistance in potato. *American Journal of Potato Research* **91**, 233–245. doi: [10.1007/s12230-013-9343-y](https://doi.org/10.1007/s12230-013-9343-y).

Perotti MF, Posé D, Martín-Pizarro C. 2023. Non-climacteric fruit development and ripening regulation: ‘the phytohormones show’. *Journal of Experimental Botany* **74**, 6237–6253. doi: [10.1093/jxb/erad271](https://doi.org/10.1093/jxb/erad271).

Perpiñá G, Esteras C, Gibon Y, Monforte AJ, Picó B. 2016. A new genomic library of melon introgression lines in a cantaloupe genetic background for dissecting desirable agronomical traits. *BMC Plant Biology* **16**, 1–21. doi: [10.1186/s12870-016-0842-0](https://doi.org/10.1186/s12870-016-0842-0).

Pires MA, Pastrana LM, Fucinõs P, Abreu CS, Oliveira SM. 2020. Sensorial Perception of astringency: Oral mechanisms and current analysis methods. *Foods* **9**, 1124. doi: [10.3390/foods9081124](https://doi.org/10.3390/foods9081124).

Pokotylo I V., Kolesnikov YS, Derevyanchuk M V., Kharitonenko AI, Kravets VS.

2015. Lipoxygenases and plant cell metabolism regulation. *Ukrainian Biochemical Journal* **87**, 41–55. doi: [10.15407/ubj87.02.041](https://doi.org/10.15407/ubj87.02.041).
- Pombo MA, Martínez GA, Civello PM.** 2011. Cloning of *FaPAL6* gene from strawberry fruit and characterization of its expression and enzymatic activity in two cultivars with different anthocyanin accumulation. *Plant Science* **181**, 111–118. doi: [10.1016/j.plantsci.2011.04.012](https://doi.org/10.1016/j.plantsci.2011.04.012).
- Pont C, Murat F, Guizard S, et al.** 2013. Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. *Plant Journal* **76**, 1030–1044. doi: [10.1111/tpj.12366](https://doi.org/10.1111/tpj.12366).
- Porta H, Rocha-Sosa M.** 2002. Plant lipoxygenases. Physiological and molecular features. *Plant Physiology* **130**, 15–21. doi: [10.1104/pp.010787](https://doi.org/10.1104/pp.010787).
- Pott DM, Osorio S, Vallarino JG.** 2019. From central to specialized metabolism: An overview of some secondary compounds derived from the primary metabolism for their role in conferring nutritional and organoleptic characteristics to fruit. *Frontiers in Plant Science* **10**. doi: [10.3389/fpls.2019.00835](https://doi.org/10.3389/fpls.2019.00835).
- Pott DM, Vallarino JG, Cruz-Rus E, Willmitzer L, Sánchez-Sevilla JF, Amaya I, Osorio S.** 2020. Genetic analysis of phenylpropanoids and antioxidant capacity in strawberry fruit reveals mQTL hotspots and candidate genes. *Scientific Reports* **10**, 1-15. doi: [10.1038/s41598-020-76946-x](https://doi.org/10.1038/s41598-020-76946-x).
- Potter D, Eriksson T, Evans RC, et al.** 2007. Phylogeny and classification of *Rosaceae*. *Plant Systematics and Evolution* **266**, 5–43. doi: [10.1007/s00606-007-0539-9](https://doi.org/10.1007/s00606-007-0539-9).
- Potter D, Gao F, Bortiri PE, Oh SH, Baggett S.** 2002. Phylogenetic relationships in *Rosaceae* inferred from chloroplast *matK* and *trnL-trnF* nucleotide sequence data. *Plant Systematics and Evolution* **231**, 77–89. doi: [10.1007/s006060200012](https://doi.org/10.1007/s006060200012).
- Potter D, Luby JJ, Harrison RE.** 2000. Phylogenetic relationships among species of *Fragaria* (*Rosaceae*) inferred from non-coding nuclear and chloroplast DNA sequences. *Systematic Botany* **25**, 337–348. doi: [10.2307/2666646](https://doi.org/10.2307/2666646).
- Prohaska A, Petit A, Lesemann S, et al.** 2024. Strawberry phenotypic plasticity in flowering time is driven by the interaction between genetic loci and temperature. *Journal of Experimental Botany*. doi: [10.1093/jxb/erae279](https://doi.org/10.1093/jxb/erae279).
- Pulido P, Perello C, Rodriguez-Concepcion M.** 2012. New insights into plant Isoprenoid metabolism. *Molecular Plant* **5**, 964–967. doi: [10.1093/mp/sss088](https://doi.org/10.1093/mp/sss088).
- Quideau S.** 2009. Chemistry and biology of ellagitannins: An underestimated class of bioactive plant polyphenols. *WORLD SCIENTIFIC*. doi: [10.1142/6795](https://doi.org/10.1142/6795).
- Quideau S, Deffieux D, Douat-Casassus C, Pouységu L.** 2011. Plant polyphenols: Chemical properties, biological activities, and synthesis. doi: [10.1002/anie.201000044](https://doi.org/10.1002/anie.201000044).
- Raguso RA.** 2008. Wake up and smell the roses: The ecology and evolution of floral scent. *Annual Review of Ecology, Evolution, and Systematics* **39**, 549–569. doi: [10.1146/annurev.ecolsys.38.091206.095601](https://doi.org/10.1146/annurev.ecolsys.38.091206.095601).
- Rahimi S, Kim J, Mijakovic I, Jung KH, Choi G, Kim SC, Kim YJ.** 2019. Triterpenoid-biosynthetic UDP-glycosyltransferases from plants. *Biotechnology Advances* **37**, 107394. doi: [10.1016/j.biotechadv.2019.04.016](https://doi.org/10.1016/j.biotechadv.2019.04.016).

- Ramírez L, Arrizon J, Sandoval G, Cardador A, Bello-Mendoza R, Lappe P, Mateos-Díaz JC.** 2008. A new microplate screening method for the simultaneous activity quantification of feruloyl esterases, tannases, and chlorogenate esterases. *Applied Biochemistry and Biotechnology* **151**, 711–723. doi: [10.1007/s12010-008-8319-8](https://doi.org/10.1007/s12010-008-8319-8).
- Ramsay NA, Glover BJ.** 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Science* **10**, 63–70. doi: [10.1016/j.tplants.2004.12.011](https://doi.org/10.1016/j.tplants.2004.12.011).
- Rattanakon S, Ghan R, Gambetta GA, Deluc LG, Schlauch KA, Cramer GR.** 2016. Abscisic acid transcriptomic signaling varies with grapevine organ. *BMC Plant Biology* **16**, 1–14. doi: [10.1186/s12870-016-0763-y](https://doi.org/10.1186/s12870-016-0763-y).
- Ricci A, Olejar KJ, Parpinello GP, Mattioli AU, Teslić N, Kilmartin PA, Versari A.** 2016. Antioxidant activity of commercial food grade tannins exemplified in a wine model. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment* **33**, 1761–1774. doi: [10.1080/19440049.2016.1241901](https://doi.org/10.1080/19440049.2016.1241901).
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK.** 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47. doi: [10.1093/nar/gkv007](https://doi.org/10.1093/nar/gkv007).
- Robichaud JL, Noble AC.** 1990. Astringency and bitterness of selected phenolics in wine. *Journal of the Science of Food and Agriculture* **53**, 343–353. doi: [10.1002/jsfa.2740530307](https://doi.org/10.1002/jsfa.2740530307).
- Rodríguez-Franco A, Caballero JL, Muñoz-Blanco J, Blanco-Portales R.** 2019. An atypical HLH transcriptional regulator plays a novel and important role in strawberry ripened receptacle. *BMC Plant Biology* **19**, 1–16. doi: [10.1186/s12870-019-2092-4](https://doi.org/10.1186/s12870-019-2092-4).
- Rodríguez H, Curiel JA, Landete JM, de las Rivas B, de Felipe FL, Gómez-Cordovés C, Mancheño JM, Muñoz R.** 2009. Food phenolics and lactic acid bacteria. *International Journal of Food Microbiology* **132**, 79–90. doi: [10.1016/j.ijfoodmicro.2009.03.025](https://doi.org/10.1016/j.ijfoodmicro.2009.03.025).
- Rohdich F, Zepeck F, Adam P, et al.** 2003. The deoxyxylulose phosphate pathway of isoprenoid biosynthesis: Studies on the mechanisms of the reactions catalyzed by IspG and IspH protein. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1586–1591. doi: [10.1073/pnas.0337742100](https://doi.org/10.1073/pnas.0337742100).
- Ross J, Li Y, Lim EK, Bowles DJ.** 2001. Higher plant glycosyltransferases. *Genome Biology* **2**, 1–6. doi: [10.1186/gb-2001-2-2-reviews3004](https://doi.org/10.1186/gb-2001-2-2-reviews3004).
- Rossetti D, Bongaerts JHH, Wantling E, Stokes JR, Williamson AM.** 2009. Astringency of tea catechins: More than an oral lubrication tactile percept. *Food Hydrocolloids* **23**, 1984–1992. doi: [10.1016/j.foodhyd.2009.03.001](https://doi.org/10.1016/j.foodhyd.2009.03.001).
- Rousseau-Gueutin M, Gaston A, Aïnouche A, Aïnouche ML, Olbricht K, Staudt G, Richard L, Denoyes-Rothan B.** 2009. Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): New insights from phylogenetic analyses of low-copy nuclear genes. *Molecular Phylogenetics and Evolution* **51**, 515–530. doi: [10.1016/j.ympev.2008.12.024](https://doi.org/10.1016/j.ympev.2008.12.024).
- Rousseau-Gueutin M, Lerceteau-Köhler E, Barrot L, Sargent DJ, Monfort A, Simpson D, Arús P, Guérin G, Denoyes-Rothan B.** 2008. Comparative genetic mapping between octoploid and diploid *Fragaria* species reveals a high level of

colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. *Genetics* **179**, 2045–2060. doi: [10.1534/genetics.107.083840](https://doi.org/10.1534/genetics.107.083840).

Rousserie P, Rabot A, Geny-Denis L. 2019. From flavanols biosynthesis to wine tannins: What place for grape seeds? *Journal of Agricultural and Food Chemistry* **67**, 1325–1343. doi: [10.1021/acs.jafc.8b05768](https://doi.org/10.1021/acs.jafc.8b05768).

Ruiz-Lopez N, Pérez-Sancho J, del Valle AE, et al. 2021. Synaptotagmins at the endoplasmic reticulum–plasma membrane contact sites maintain diacylglycerol homeostasis during abiotic stress. *Plant Cell* **33**, 2431–2453. doi: [10.1093/plcell/koab122](https://doi.org/10.1093/plcell/koab122).

Salas JJ, Sánchez J. 1999. Hydroperoxide lyase from olive (*Olea europaea*) fruits. *Plant Science* **143**, 19–26. doi: [10.1016/S0168-9452\(99\)00027-8](https://doi.org/10.1016/S0168-9452(99)00027-8).

Sarni-Manchado P, Deleris A, Avallone S, Cheynier V, Moutounet M. 1999. Analysis and characterization of wine condensed tannins precipitated by proteins used as fining agent in enology. *American Journal of Enology and Viticulture* **1**, 81–86. doi: [10.5344/ajev.1999.50.1.81](https://doi.org/10.5344/ajev.1999.50.1.81).

Sartoni S, Diaz M, Diaz-Muñoz G. 2021. Lactones: classification, synthesis, biological activities, and industrial applications. *Tetrahedron* **84**, 132001. doi: [10.1016/j.tet.2021.132001](https://doi.org/10.1016/j.tet.2021.132001).

Sánchez-Gómez C, Posé D, Martín-Pizarro C. 2022. Insights into transcription factors controlling strawberry fruit development and ripening. *Frontiers in Plant Science* **13**, 1022369. doi: [10.3389/fpls.2022.1022369/bibtex](https://doi.org/10.3389/fpls.2022.1022369/bibtex).

Sánchez-Sevilla JF, Cruz-Rus E, Valpuesta V, Botella MA, Amaya I. 2014. Deciphering gamma-decalactone biosynthesis in strawberry fruit using a combination of genetic mapping, RNA-Seq and eQTL analyses. *BMC Genomics* **15**, 1–15. doi: [10.1186/1471-2164-15-218](https://doi.org/10.1186/1471-2164-15-218).

Sánchez-Sevilla JF, Horvath A, Botella MA, Gaston A, Folta K, Kilian A, Denoyes B, Amaya I. 2015. Diversity arrays technology (DArT) marker platforms for diversity analysis and linkage mapping in a complex crop, the octoploid cultivated strawberry (*Fragaria × ananassa*). *PLOS ONE* **10**, e0144960. doi: [10.1371/journal.pone.0144960](https://doi.org/10.1371/journal.pone.0144960).

Sánchez-Sevilla JF, Vallarino JG, Osorio S, Bombarely A, Posé D, Merchante C, Botella MA, Amaya I, Valpuesta V. 2017. Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (*Fragaria × ananassa*). *Scientific reports* **7**, 13737. doi: [10.1038/s41598-017-14239-6](https://doi.org/10.1038/s41598-017-14239-6).

Sarni-Manchado P, Deleris A, Avallone S, Cheynier V, Moutounet M. 1999. Analysis and characterization of wine condensed tannins precipitated by proteins used as fining agent in enology. *American Journal of Enology and Viticulture* **50**, 81–86. doi: [10.5344/ajev.1999.50.1.81](https://doi.org/10.5344/ajev.1999.50.1.81).

Saxe HJ, Horibe T, Balan B, Butterfield TS, Feinberg NG, Zabaneh CM, Jacobson AE, Dandekar AM. 2021. Two UGT84A family glycosyltransferases regulate phenol, flavonoid, and tannin metabolism in *Juglans regia* (English walnut). *Frontiers in Plant Science* **12**, 178. doi: [10.3389/fpls.2021.626483](https://doi.org/10.3389/fpls.2021.626483).

Schaart JG, Dubos C, Romero De La Fuente I, van Houwelingen AMML, de Vos RCH, Jonker HH, Xu W, Routaboul JM, Lepiniec L, Bovy AG. 2013. Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling

- proanthocyanidin biosynthesis in strawberry (*Fragaria* × *ananassa*) fruits. *New Phytologist* **197**, 454–467. doi: [10.1111/nph.12017](https://doi.org/10.1111/nph.12017).
- Schaller A, Stintzi A. 2009. Enzymes in jasmonate biosynthesis - Structure, function, regulation. *Phytochemistry* **70**, 1532–1538. doi: [10.1016/j.phytochem.2009.07.032](https://doi.org/10.1016/j.phytochem.2009.07.032).
- Schiller D, Contreras C, Vogt J, Dunemann F, Defilippi BG, Beaudry R, Schwab W. 2015. A dual positional specific lipoxygenase functions in the generation of flavor compounds during climacteric ripening of apple. *Horticulture Research* **2**. doi: [10.1038/hortres.2015.3](https://doi.org/10.1038/hortres.2015.3).
- Schläpfer P, Zhang P, Wang C, *et al.* 2017. Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. *Plant Physiology* **173**, 2041–2059. doi: [10.1104/pp.16.01942](https://doi.org/10.1104/pp.16.01942).
- Schmid SW, Denzel K, Gross GG, Schilling G. 1987. Enzymatic synthesis of 1,6-digalloylglucose from β-glucogallin by β-glucogallin: β-glucogallin 6-O-galloyltransferase from oak leaves. *Zeitschrift für Naturforschung - Section C Journal of Biosciences* **42**, 87–92. doi: [10.1515/znc-1987-1-215](https://doi.org/10.1515/znc-1987-1-215).
- Schnable JC, Freeling M. 2011. Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. *PLOS ONE* **6**, e17855. doi: [10.1371/journal.pone.0017855](https://doi.org/10.1371/journal.pone.0017855).
- Schnable JC, Springer NM, Freeling M. 2011. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4069–4074. doi: [10.1073/pnas.1101368108](https://doi.org/10.1073/pnas.1101368108).
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675. doi: [10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089).
- Schwieterman ML, Colquhoun TA, Jaworski EA, *et al.* 2014. Strawberry flavor: Diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLoS ONE* **9**, e88446. doi: [10.1371/journal.pone.0088446](https://doi.org/10.1371/journal.pone.0088446).
- Seeram N, Lee R, Hardy M, Heber D. 2005. Rapid large scale purification of ellagitannins from pomegranate husk, a by-product of the commercial juice industry. *Separation and Purification Technology* **41**, 49–55. doi: [10.1016/j.seppur.2004.04.003](https://doi.org/10.1016/j.seppur.2004.04.003).
- Sengupta S, Mukherjee S, Basak P, Majumder AL. 2015. Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Frontiers Research Foundation*. doi: [10.3389/fpls.2015.00656](https://doi.org/10.3389/fpls.2015.00656).
- Serrani JC, Ruiz-Rivero O, Fos M, García-Martínez JL. 2008. Auxin-induced fruit-set in tomato is mediated in part by gibberellins. *Plant Journal* **56**, 922–934. doi: [10.1111/j.1365-3113X.2008.03654.x](https://doi.org/10.1111/j.1365-3113X.2008.03654.x).
- Seymour GB, Ryder CD, Cevik V, Hammond JP, Popovich A, King GJ, Vrebalov J, Giovannoni JJ, Manning K. 2011. A SEPALLATA gene is involved in the development and ripening of strawberry (*Fragaria* × *ananassa* Duch.) fruit, a non-climacteric tissue. *Journal of Experimental Botany* **62**, 1179–1188. doi: [10.1093/jxb/erq360](https://doi.org/10.1093/jxb/erq360).
- Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *Plant Cell* **17**, 3141–3154. doi: [10.1105/tpc.105.035055](https://doi.org/10.1105/tpc.105.035055).

- Shao Y, Hu Z, Yu Y, Mou R, Zhu Z, Beta T.** 2018. Phenolic acids, anthocyanins, proanthocyanidins, antioxidant activity, minerals and their correlations in non-pigmented, red, and black rice. *Food Chemistry* **239**, 733–741. doi: [10.1016/j.foodchem.2017.07.009](https://doi.org/10.1016/j.foodchem.2017.07.009).
- Sheard LB, Tan X, Mao H, et al.** 2010. Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature* **468**, 400–407. doi: [10.1038/nature09430](https://doi.org/10.1038/nature09430).
- Shen J, Tieman D, Jones JB, Taylor MG, Schmelz E, Huffaker A, Bies D, Chen K, Klee HJ.** 2014. A 13-lipoxygenase, TomloxC, is essential for synthesis of C5 flavour volatiles in tomato. *Journal of Experimental Botany* **65**, 419–428. doi: [10.1093/jxb/ert382](https://doi.org/10.1093/jxb/ert382).
- Shi R, Yu J, Chang X, Qiao L, Liu X, Lu L.** 2023. Recent advances in research into jasmonate biosynthesis and signaling pathways in agricultural crops and products. *Processes* 2023, Vol. 11, Page 736 **11**, 736. doi: [10.3390/pr11030736](https://doi.org/10.3390/pr11030736).
- Shibata Y, Ojika M, Sugiyama A, Yazaki K, Jones DA, Kawakita K, Takemoto D.** 2016. The full-size ABCG transporters Nb-ABCG1 and Nb-ABCG2 function in pre- and postinvasion defense against *Phytophthora infestans* in *Nicotiana benthamiana*. *Plant Cell* **28**, 1163–1181. doi: [10.1105/tpc.15.00721](https://doi.org/10.1105/tpc.15.00721).
- Shimura K, Okada A, Okada K, et al.** 2007. Identification of a biosynthetic gene cluster in rice for momilactones. *Journal of Biological Chemistry* **282**, 34013–34018. doi: [10.1074/jbc.m703344200](https://doi.org/10.1074/jbc.m703344200).
- Shin J, Park E, Choi G.** 2007. PIF3 regulates anthocyanin biosynthesis in an HY5-dependent manner with both factors directly binding anthocyanin biosynthetic gene promoters in *Arabidopsis*. *Plant Journal* **49**, 981–994. doi: [10.1111/j.1365-3113.2006.03021.x](https://doi.org/10.1111/j.1365-3113.2006.03021.x).
- Shin Y-J, Shon M-S, Kim G-N, Lee S-C.** 2014. Antioxidant and anti-adipogenic activities of persimmon tannins. *Food Science and Biotechnology* **23**, 1689–1694. doi: [10.1007/s10068-014-0230-1](https://doi.org/10.1007/s10068-014-0230-1).
- Shinozawa A, Otake R, Takezawa D, et al.** 2019. SnRK2 protein kinases represent an ancient system in plants for adaptation to a terrestrial environment. *Communications Biology* **2**, 1–13. doi: [10.1038/s42003-019-0281-1](https://doi.org/10.1038/s42003-019-0281-1).
- Shneier A, Abell C, Kleanthous C, Deka R, Coggins JR.** 1991. Observation of an imine intermediate on dehydroquinase by electrospray mass spectrometry. *Journal of the American Chemical Society* **113**, 9416–9418. doi: [10.1021/ja00024a085](https://doi.org/10.1021/ja00024a085).
- Shyu C, Figueroa P, de Pew CL, Cooke TF, Sheard LB, Moreno JE, Katsir L, Zheng N, Browse J, Howe GA.** 2012. JAZ8 lacks a canonical degron and has an EAR motif that mediates transcriptional repression of jasmonate responses in *Arabidopsis*. *Plant Cell* **24**, 536–550. doi: [10.1105/tpc.111.093005](https://doi.org/10.1105/tpc.111.093005).
- Siebert KJ.** 1999. Effects of protein-polyphenol interactions on beverage haze, stabilization, and analysis. *Journal of Agricultural and Food Chemistry* **47**, 353–362. doi: [10.1021/jf980703o](https://doi.org/10.1021/jf980703o).
- Siebert KJ, Troukhanova N V., Lynn PY.** 1996. Nature of polyphenol-protein interactions. *Journal of Agricultural and Food Chemistry* **44**, 80–85. doi: [10.1021/jf9502459](https://doi.org/10.1021/jf9502459).
- Siedow JN.** 1991. Plant lipoxygenase: Structure and function. *Annual Review of Plant*

Physiology and Plant Molecular Biology **42**, 145–188. doi: [10.1146/annurev.pp.42.060191.001045](https://doi.org/10.1146/annurev.pp.42.060191.001045).

Sigrist CJA, De Castro E, Cerutti L, Cucho BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. 2013. New and continuing developments at PROSITE. Nucleic Acids Research **41**, D344–D347. doi: [10.1093/nar/gks1067](https://doi.org/10.1093/nar/gks1067).

Silva R, Aguiar TQ, Coelho E, Jiménez A, Revuelta JL, Domingues L. 2019. Metabolic engineering of *Ashbya gossypii* for deciphering the de novo biosynthesis of γ -lactones. Microbial Cell Factories **18**, 1–11. doi: [10.1186/s12934-019-1113-1](https://doi.org/10.1186/s12934-019-1113-1).

Simkin AJ, Guirimand G, Papon N, Courdavault V, Thabet I, Ginis O, Bouzid S, Giglioli-Guivarc'h N, Clastre M. 2011. Peroxisomal localisation of the final steps of the mevalonic acid pathway in *planta*. Planta **234**, 903–914. doi: [10.1007/s00425-011-1444-6](https://doi.org/10.1007/s00425-011-1444-6).

Singh P, Arif Y, Miszczuk E, Bajguz A, Hayat S. 2022. Specific roles of Lipoxygenases in development and responses to stress in plants. Plants **11**, 979. doi: [10.3390/plants11070979](https://doi.org/10.3390/plants11070979).

Smeriglio A, Barreca D, Bellocchio E, Trombetta D. 2017. Proanthocyanidins and hydrolysable tannins: occurrence, dietary intake and pharmacological effects. British Journal of Pharmacology **174**, 1244–1262. doi: [10.1111/bph.13630](https://doi.org/10.1111/bph.13630).

Soltis PS. 2005. Ancient and recent polyploidy in angiosperms. New Phytologist **166**, 5–8. doi: [10.1111/j.1469-8137.2005.01379.x](https://doi.org/10.1111/j.1469-8137.2005.01379.x).

Song C, Gu L, Liu J, Zhao S, Hong X, Schulenburg K, Schwab W. 2015. Functional characterization and substrate promiscuity of UGT71 glycosyltransferases from Strawberry (*Fragaria* $\times$ *ananassa*). Plant and Cell Physiology. doi: [10.1093/pcp/pcv151](https://doi.org/10.1093/pcp/pcv151).

Song C, Hong X, Zhao S, Liu J, Schulenburg K, Huang FC, Franz-Oberdorf K, Schwab W. 2016. Glucosylation of 4-hydroxy-2,5-dimethyl-3(2*H*)-furanone, the key strawberry flavor compound in strawberry fruit. Plant Physiology **171**, 139–151. doi: [10.1104/pp.16.00226](https://doi.org/10.1104/pp.16.00226).

Song Y, Peng Y, Liu L, Li G, Zhao X, Wang X, Cao S, Muyle A, Zhou Y, Zhou H. 2024. Phased gap-free genome assembly of octoploid cultivated strawberry illustrates the genetic and epigenetic divergence among subgenomes. Horticulture Research **11**. doi: [10.1093/hr/uhad252](https://doi.org/10.1093/hr/uhad252).

Soriano G, Kneeshaw S, Jimenez-Aleman G, Zamarreño ÁM, Franco-Zorrilla JM, Rey-Stolle MF, Barbas C, García-Mina JM, Solano R. 2022. An evolutionarily ancient fatty acid desaturase is required for the synthesis of hexadecatrienoic acid, which is the main source of the bioactive jasmonate in *Marchantia polymorpha*. New Phytologist **233**, 1401–1413. doi: [10.1111/nph.17850](https://doi.org/10.1111/nph.17850).

Staggs CG, Sealey WM, McCabe BJ, Teague AM, Mock DM. 2004. Determination of the biotin content of select foods using accurate and sensitive HPLC/avidin binding. Journal of Food Composition and Analysis **17**, 767–776. doi: [10.1016/j.jfca.2003.09.015](https://doi.org/10.1016/j.jfca.2003.09.015).

Staudt G. 1962. Taxonomic studies in the genus *Fragaria* typification of *Fragaria* species known at the time of Linnaeus. Canadian Journal of Botany. doi: [10.1139/b62-081](https://doi.org/10.1139/b62-081).

Staudt G. 2009. Strawberry biogeography, genetics and systematics. Acta Horticulturae

842, 71–84. doi: [10.17660/ActaHortic.2009.842.1](https://doi.org/10.17660/ActaHortic.2009.842.1).

Stebbins GL. 1947. Types of polyploids: Their classification and significance. *Advances in Genetics* **1**, 403–429. doi: [10.1016/S0065-2660\(08\)60490-3](https://doi.org/10.1016/S0065-2660(08)60490-3).

Stefanachi A, Leonetti F, Pisani L, Catto M, Carotti A. 2018. Coumarin: A natural, privileged and versatile scaffold for bioactive compounds. *Molecules* **23**, 250. doi: [10.3390/molecules23020250](https://doi.org/10.3390/molecules23020250).

Sun X, Cheng X, Zhang J, Ju Y, Que Z, Liao X, Lao F, Fang Y, Ma T. 2020. Letting wine polyphenols functional: Estimation of wine polyphenols bioaccessibility under different drinking amount and drinking patterns. *Food Research International* **127**, 108704. doi: [10.1016/j.foodres.2019.108704](https://doi.org/10.1016/j.foodres.2019.108704).

Sun J, Sun R, Liu H, et al. 2021. Complete chloroplast genome sequencing of ten wild *Fragaria* species in China provides evidence for phylogenetic evolution of *Fragaria*. *Genomics* **113**, 1170–1179. doi: [10.1016/j.ygeno.2021.01.027](https://doi.org/10.1016/j.ygeno.2021.01.027).

Suvanto J, Karppinen K, Riihinen K, Jaakola L, Salminen JP. 2020. Changes in the proanthocyanidin composition and related gene expression in bilberry (*Vaccinium myrtillus* L.) tissues. *Journal of Agricultural and Food Chemistry* **68**, 7378–7386. doi: [10.1021/acs.jafc.0c02158](https://doi.org/10.1021/acs.jafc.0c02158).

Suzuki N, Koussevitzky S, Mittler R, Miller G. 2012. ROS and redox signalling in the response of plants to abiotic stress. *Plant, Cell and Environment* **35**, 259–270. doi: [10.1111/j.1365-3040.2011.02336.x](https://doi.org/10.1111/j.1365-3040.2011.02336.x).

Swaminathan S, Morrone D, Wang Q, Bruce Fulton D, Peters RJ. 2009. CYP76M7 is an ent-cassadiene C11 α -hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *Plant Cell* **21**, 3315–3325. doi: [10.1105/tpc.108.063677](https://doi.org/10.1105/tpc.108.063677).

Syed N, Singh S, Chaturvedi S, Nannaware A, Khare S, Rout P. 2022. Production of lactones for flavoring and pharmacological purposes from unsaturated lipids: an industrial perspective. *Critical Reviews in Food Science and Nutrition* **63**, 10047–10078. doi: [10.1080/10408398.2022.2068124](https://doi.org/10.1080/10408398.2022.2068124).

Symons GM, Chua YJ, Ross JJ, Quittenden LJ, Davies NW, Reid JB. 2012. Hormonal changes during non-climacteric ripening in strawberry. *Journal of Experimental Botany* **63**, 4741–4750. doi: [10.1093/jxb/ers147](https://doi.org/10.1093/jxb/ers147).

Taiyun W, Simko V. 2021. R package ‘corrplot’: Visualization of a Correlation Matrix. <https://github.com/taiyun/corrplot>.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution* **38**, 3022–3027. doi: [10.1093/molbev/msab120](https://doi.org/10.1093/molbev/msab120).

Tan H, Zhou Y, Dinius E, Lozano-Durán R. 2024. The Ti-TAN plasmid toolbox for TurboID-based proximity labeling assays in *Nicotiana benthamiana*. *Journal of Integrative Plant Biology* **66**, 166–168. doi: [10.1111/jipb.13610/supinfo](https://doi.org/10.1111/jipb.13610/supinfo).

Tegl G, Nidetzky B. 2020. Leloir glycosyltransferases of natural product C-glycosylation: structure, mechanism and specificity. *Biochemical Society Transactions* **48**, 1583–1598. doi: [10.1042/BST20191140](https://doi.org/10.1042/BST20191140).

Tennessen JA, Govindarajulu R, Ashman TL, Liston A. 2014. Evolutionary origins

and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biology and Evolution*. doi: [10.1093/gbe/evu261](https://doi.org/10.1093/gbe/evu261).

Terrier N, Torregrosa L, Ageorges A, Vialet S, Verriès C, Cheynier V, Romieu C. 2009. Ectopic expression of *VvMybPA2* promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiology* **149**, 1028–1041. doi: [10.1104/PP.108.131862](https://doi.org/10.1104/PP.108.131862).

Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, Nomura K, He SY, Howe GA, Browse J. 2007. JAZ repressor proteins are targets of the SCFCOI1 complex during jasmonate signalling. *Nature* **448**, 661–665. doi: [10.1038/nature05960](https://doi.org/10.1038/nature05960).

Tholl D. 2015. Biosynthesis and biological functions of terpenoids in plants. *Advances in Biochemical Engineering/Biotechnology* **148**, 63–106. doi: [10.1007/10_2014_295](https://doi.org/10.1007/10_2014_295).

Thomas BC, Pedersen B, Freeling M. 2006. Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Genome Research* **16**, 934–946. doi: [10.1101/gr.4708406](https://doi.org/10.1101/gr.4708406).

Thumulari V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. 2022. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research* **50**, W228–W234. doi: [10.1093/nar/gkac278](https://doi.org/10.1093/nar/gkac278).

Tieman DM, Loucas HM, Kim JY, Clark DG, Klee HJ. 2007. Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol. *Phytochemistry* **68**, 2660–2669. doi: [10.1016/j.phytochem.2007.06.005](https://doi.org/10.1016/j.phytochem.2007.06.005).

Tieman D, Taylor M, Schauer N, Fernie AR, Hanson AD, Klee HJ. 2006. Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8287–8292. doi: [10.1073/pnas.0602469103](https://doi.org/10.1073/pnas.0602469103).

Tiwari P, Sangwan RS, Sangwan NS. 2016. Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes. *Biotechnology Advances* **34**, 714–739. doi: [10.1016/j.biotechadv.2016.03.006](https://doi.org/10.1016/j.biotechadv.2016.03.006).

Toporkova YY, Gorina SS, Bessolitsyna EK, Smirnova EO, Fatykhova VS, Brühlmann F, Ilyina TM, Mukhtarova LS, Grechkin AN. 2018. Double function hydroperoxide lyases/epoxyalcohol synthases (CYP74C) of higher plants: identification and conversion into allene oxide synthases by site-directed mutagenesis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1863**, 369–378. doi: [10.1016/j.bbalip.2018.01.002](https://doi.org/10.1016/j.bbalip.2018.01.002).

Tulipani S, Mezzetti B, Battino M. 2009. Impact of strawberries on human health: Insight into marginally discussed bioactive compounds for the Mediterranean diet. *Public Health Nutrition* **12**, 1656–1662. doi: [10.1017/s1368980009990516](https://doi.org/10.1017/s1368980009990516).

Tulipani S, Mezzetti B, Capocasa F, Bompadre S, Beekwilder J, De Vos CHR, Capanoglu E, Bovy A, Battino M. 2008. Antioxidants, phenolic compounds, and nutritional quality of different strawberry genotypes. *Journal of Agricultural and Food Chemistry*. doi: [10.1021/jf0719959](https://doi.org/10.1021/jf0719959).

Tzin V, Galili G. 2010. The biosynthetic pathways for shikimate and aromatic amino

acids in *Arabidopsis thaliana*. The American Society of Plant Biologists. doi: [10.1199/tab.0132](https://doi.org/10.1199/tab.0132).

Ullah C, Unsicker SB, Fellenberg C, Constabel CP, Schmidt A, Gershenzon J, Hammerbacher A. 2017. Flavan-3-ols are an effective chemical defense against rust infection. *Plant Physiology* **175**, 1560–1578. doi: [10.1104/pp.17.00842](https://doi.org/10.1104/pp.17.00842).

Ulrich D, Hoberg E, Rapp A, Kecke S. 1997. Analysis of strawberry flavour - discrimination of aroma types by quantification of volatile compounds. *European Food Research and Technology* **205**, 218–223. doi: [10.1007/s002170050154](https://doi.org/10.1007/s002170050154).

Ulrich D, Olbricht K. 2016. A search for the ideal flavor of strawberry - Comparison of consumer acceptance and metabolite patterns in *Fragaria* × *ananassa* Duch. *Journal of Applied Botany and Food Quality* **89**, 223–234. doi: [10.5073/jabfq.2016.089.029](https://doi.org/10.5073/jabfq.2016.089.029).

Ulrich D, Ulbricht K. 2013. Diversity of volatile patterns in sixteen *Fragaria vesca* L. Accessions in comparison to cultivars of *Fragaria* × *ananassa*. *Journal of Applied Botany and Food Quality* **86**, 37–46. doi: [10.5073/jabfq.2013.086.006](https://doi.org/10.5073/jabfq.2013.086.006).

Umezawa T, Sugiyama N, Mizoguchi M, Hayashi S, Myouga F, Yamaguchi-Shinozaki K, Ishihama Y, Hirayama T, Shinozaki K. 2009. Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 17588–17593. doi: [10.1073/pnas.0907095106](https://doi.org/10.1073/pnas.0907095106).

Unsicker SB, Kunert G, Gershenzon J. 2009. Protective perfumes: the role of vegetative volatiles in plant defense against herbivores. *Current Opinion in Plant Biology* **12**, 479–485. doi: [10.1016/j.pbi.2009.04.001](https://doi.org/10.1016/j.pbi.2009.04.001).

Urrutia M, Rambla JL, Alexiou KG, Granell A, Monfort A. 2017. Genetic analysis of the wild strawberry (*Fragaria vesca*) volatile composition. *Plant Physiology and Biochemistry* **121**, 99–117. doi: [10.1016/j.plaphy.2017.10.015](https://doi.org/10.1016/j.plaphy.2017.10.015).

Vallarino JG, Pott DM, Cruz-Rus E, Miranda L, Medina-Minguez JJ, Valpuesta V, Fernie AR, Sánchez-Sevilla JF, Osorio S, Amaya I. 2019. Identification of quantitative trait loci and candidate genes for primary metabolite content in strawberry fruit. *Horticulture Research* doi: [10.1038/s41438-018-0077-3](https://doi.org/10.1038/s41438-018-0077-3). doi: [10.1038/s41438-018-0077-3](https://doi.org/10.1038/s41438-018-0077-3).

VanBuren R, Man Wai C, Wang X, et al. 2020. Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nature Communications* **11**, 884. doi: [10.1038/s41467-020-14724-z](https://doi.org/10.1038/s41467-020-14724-z).

Vanderhaegen B, Neven H, Verachtert H, Derdelinckx G. 2006. The chemistry of beer aging - A critical review. *Food Chemistry* **95**, 357–381. doi: [10.1016/j.foodchem.2005.01.006](https://doi.org/10.1016/j.foodchem.2005.01.006).

Vassão DG, Gang DR, Koeduka T, Jackson B, Pichersky E, Davin LB, Lewis NG. 2006. Chavicol formation in sweet basil (*Ocimum basilicum*): Cleavage of an esterified C9 hydroxyl group with NAD(P)H-dependent reduction. *Organic and Biomolecular Chemistry* **4**, 2733–2744. doi: [10.1039/b605407b](https://doi.org/10.1039/b605407b).

Versari A, Du Toit W, Parpinello GP. 2013. Oenological tannins: A review. *Australian Journal of Grape and Wine Research* **19**, 1–10. doi: [10.1111/ajgw.12002](https://doi.org/10.1111/ajgw.12002).

Vignault A, González-Centeno MR, Pascual O, Gombau J, Jourdes M, Moine V,

- Iturmendi N, Canals JM, Zamora F, Teissedre PL.** 2018. Chemical characterization, antioxidant properties and oxygen consumption rate of 36 commercial oenological tannins in a model wine solution. *Food Chemistry* **268**, 210–219. doi: [10.1016/j.foodchem.2018.06.031](https://doi.org/10.1016/j.foodchem.2018.06.031).
- Villamil-Galindo E, Van de Velde F, Piagentini AM.** 2021. Strawberry agro-industrial by-products as a source of bioactive compounds: effect of cultivar on the phenolic profile and the antioxidant capacity. *Bioresources and Bioprocessing* **8**, 61. doi: [10.1186/s40643-021-00416-z](https://doi.org/10.1186/s40643-021-00416-z).
- Villavicencio JD, Tobar J, Zoffoli JP, O'Brien JA, Contreras C.** 2024. Identification, characterization, and expression of lipoxygenase genes in sweet cherry (*Prunus avium* L.) cv. Regina and their relationship with the development of an herbaceous off-flavor during fruit ripening. *Plant Physiology and Biochemistry* **206**, 108271. doi: [10.1016/j.plaphy.2023.108271](https://doi.org/10.1016/j.plaphy.2023.108271).
- Vincenti S, Mariani M, Alberti JC, Jacopini S, de Caraffa VBB, Berti L, Maury J.** 2019. Biocatalytic synthesis of natural green leaf volatiles using the lipoxygenase metabolic pathway. *Catalysts* **9**, 873. doi: [10.3390/catal9100873](https://doi.org/10.3390/catal9100873).
- Vogt T.** 2010. Phenylpropanoid biosynthesis. doi: [10.1093/mp/ssp106](https://doi.org/10.1093/mp/ssp106).
- Vranová E, Coman D, Gruissem W.** 2013. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annual Review of Plant Biology* **64**, 665–700. doi: [10.1146/annurev-arplant-050312-120116](https://doi.org/10.1146/annurev-arplant-050312-120116).
- Wang SY.** 2000. Effect of methyl jasmonate on water stress in strawberry. *Acta Horticulturae* **516**, 89–95. doi: [10.17660/actahortic.2000.516.10](https://doi.org/10.17660/actahortic.2000.516.10).
- Wang K, Froehlich JE, Zienkiewicz A, Hersh HL, Benning C.** 2017. A plastid phosphatidylglycerol lipase contributes to the export of acyl groups from plastids for seed oil biosynthesis. *Plant Cell* **29**, 1678–1696. doi: [10.1105/tpc.17.00397](https://doi.org/10.1105/tpc.17.00397).
- Wang Q, Hillwig ML, Peters RJ.** 2011. CYP99A3: Functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. *Plant Journal* **65**, 87–95. doi: [10.1111/j.1365-313x.2010.04408.x](https://doi.org/10.1111/j.1365-313x.2010.04408.x).
- Wang J, Sakurai H, Kato N, Kaji T, Ueda M.** 2021a. Syntheses of dinor-cis/iso-12-oxo-phytodienoic acid (dn-cis/iso-OPDAs), ancestral jasmonate phytohormones of the bryophyte *Marchantia polymorpha* L., and their catabolites. *Scientific Reports* **11**, 1–9. doi: [10.1038/s41598-021-81575-z](https://doi.org/10.1038/s41598-021-81575-z).
- Wang J, Wang K, Lyu S, et al.** 2021b. Genome-wide identification of tannase genes and their function of wound response and astringent substances accumulation in *Juglandaceae*. *Frontiers in Plant Science* **12**, 739. doi: [10.3389/fpls.2021.664470](https://doi.org/10.3389/fpls.2021.664470).
- Wang J, Wu A, Xu Y, Liu J, Qian X.** 2009. M2-A induces apoptosis and G2-M arrest via inhibiting PI3 K/Akt pathway in HL60 cells. *Cancer Letters* **283**, 193–202. doi: [10.1016/j.canlet.2009.03.039](https://doi.org/10.1016/j.canlet.2009.03.039).
- Wang P, Xue L, Batelli G, Lee S, Hou YJ, Van Oosten MJ, Zhang H, Tao WA, Zhu JK.** 2013. Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11205–11210. doi: [10.1073/pnas.1308974110](https://doi.org/10.1073/pnas.1308974110).

- Wang H, Zhang H, Yang Y, et al.** 2020. The control of red colour by a family of MYB transcription factors in octoploid strawberry (*Fragaria × ananassa*) fruits. *Plant Biotechnology Journal* **18**, 1169–1184. doi: [10.1111/pbi.13282](https://doi.org/10.1111/pbi.13282).
- Wasternack C.** 2015. How jasmonates earned their laurels: Past and present. *Journal of Plant Growth Regulation* **34**, 761–794. doi: [10.1007/s00344-015-9526-5](https://doi.org/10.1007/s00344-015-9526-5).
- Wasternack C, Hause B.** 2013. Jasmonates: Biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in *Annals of Botany*. *Annals of Botany* **111**, 1021–1058. doi: [10.1093/aob/mct067](https://doi.org/10.1093/aob/mct067).
- Wasternack C, Song S.** 2017. Jasmonates: Biosynthesis, metabolism, and signaling by proteins activating and repressing transcription. *Journal of Experimental Botany* **68**, 1303–1321. doi: [10.1093/jxb/erw443](https://doi.org/10.1093/jxb/erw443).
- Wasternack C, Strnad M.** 2018. Jasmonates: News on occurrence, biosynthesis, metabolism and action of an ancient group of signaling compounds. *International Journal of Molecular Sciences* **19**, 2539. doi: [10.3390/ijms19092539](https://doi.org/10.3390/ijms19092539).
- Watrelet AA, Norton EL.** 2020. Chemistry and reactivity of tannins in *Vitis spp.*: A review. *Molecules* **25**, 2110. doi: [10.3390/molecules25092110](https://doi.org/10.3390/molecules25092110).
- Wei X, Ju Y, Ma T, Zhang J, Fang Y, Sun X.** 2021. New perspectives on the biosynthesis, transportation, astringency perception and detection methods of grape proanthocyanidins. *Critical Reviews in Food Science and Nutrition* **61**, 2372–2398. doi: [10.1080/10408398.2020.1777527](https://doi.org/10.1080/10408398.2020.1777527).
- Wein M, Lavid N, Lunkenbein S, Lewinsohn E, Schwab W, Kaldenhoff R.** 2002. Isolation, cloning and expression of a multifunctional *O*-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2*H*)-furanone, one of the key aroma compounds in strawberry fruits. *Plant Journal* **31**, 755–765. doi: [10.1046/j.1365-3113x.2002.01396.x](https://doi.org/10.1046/j.1365-3113x.2002.01396.x).
- Wendel JF.** 2015. The wondrous cycles of polyploidy in plants. *American Journal of Botany* **102**, 1753–1756. doi: [10.3732/ajb.1500320](https://doi.org/10.3732/ajb.1500320).
- Wendel JF, Lisch D, Hu G, Mason AS.** 2018. The long and short of doubling down: polyploidy, epigenetics, and the temporal dynamics of genome fractionation. *Current Opinion in Genetics and Development* **49**, 1–7. doi: [10.1016/j.gde.2018.01.004](https://doi.org/10.1016/j.gde.2018.01.004).
- White PJ, Young J, Hunter IS, Nimmo HG, Coggins JR.** 1990. The purification and characterization of 3-dehydroquinase from *Streptomyces coelicolor*. *Biochemical Journal* **265**, 735–738. doi: [10.1042/bj2650735](https://doi.org/10.1042/bj2650735).
- Wickham H.** 2016. ggplot2. doi: [10.1007/978-3-319-24277-4](https://doi.org/10.1007/978-3-319-24277-4).
- Wickman H, François R, Henry L, Muller K.** 2021. dplyr: A grammar of data manipulation. CRAN Repository **3**, 1–2. doi: [10.32614/cran.package.dplyr](https://doi.org/10.32614/cran.package.dplyr).
- Widemann E, Smirnova E, Aubert Y, Miesch L, Heitz T.** 2016. Dynamics of jasmonate metabolism upon flowering and across leaf stress responses in *Arabidopsis thaliana*. *Plants* **5**, 539–547. doi: [10.3390/plants5010004](https://doi.org/10.3390/plants5010004).
- Winkel-Shirley B.** 2002. Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology* **5**, 218–223. doi: [10.1016/s1369-5266\(02\)00256-X](https://doi.org/10.1016/s1369-5266(02)00256-X).
- Wright GA, Lutmerding A, Dudareva N, Smith BH.** 2005. Intensity and the ratios of

compounds in the scent of snapdragon flowers affect scent discrimination by honeybees (*Apis mellifera*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* **191**, 105–114. doi: [10.1007/s00359-004-0576-6](https://doi.org/10.1007/s00359-004-0576-6).

Wu W, Cao SF, Shi LY, Chen W, Yin XR, Yang ZF. 2023. Absciscic acid biosynthesis, metabolism and signaling in ripening fruit. *Frontiers in Plant Science* **14**, 1279031. doi: [10.3389/fpls.2023.1279031](https://doi.org/10.3389/fpls.2023.1279031).

Wu Y, Hillwig ML, Wang Q, Peters RJ. 2011. Parsing a multifunctional biosynthetic gene cluster from rice: Biochemical characterization of CYP71Z6 & 7. *FEBS Letters* **585**, 3446–3451. doi: [10.1016/j.febslet.2011.09.038](https://doi.org/10.1016/j.febslet.2011.09.038).

Wu S, Schalk M, Clark A, Miles RB, Coates R, Chappell J. 2006. Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nature Biotechnology* **24**, 1441–1447. doi: [10.1038/nbt1251](https://doi.org/10.1038/nbt1251).

Wu S, Tian L. 2017. Diverse phytochemicals and bioactivities in the ancient fruit and modern functional food pomegranate (*Punica granatum*). *Molecules* **22**, 1606. doi: [10.3390/molecules22101606](https://doi.org/10.3390/molecules22101606).

Xie DX, Feys BF, James S, Nieto-Rostro M, Turner JG. 1998. COI1: An *Arabidopsis* gene required for jasmonate-regulated defense and fertility. *Science* **280**, 1091–1094. doi: [10.1126/science.280.5366.1091](https://doi.org/10.1126/science.280.5366.1091).

Xie YG, Ma YY, Bi PP, Wei W, Liu J, Hu Y, Gou YJ, Zhu D, Wen YQ, Feng JY. 2020. Transcription factor FvTCP9 promotes strawberry fruit ripening by regulating the biosynthesis of abscisic acid and anthocyanins. *Plant Physiology and Biochemistry* **146**, 374–383. doi: [10.1016/j.plaphy.2019.11.004](https://doi.org/10.1016/j.plaphy.2019.11.004).

Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. 2003. Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* **299**, 396–399. doi: [10.1126/science.1078540](https://doi.org/10.1126/science.1078540).

Xu F, Gao X, Xi Z, Zhang H, Peng X, Wang Z, Wang T, Meng Y. 2015. Application of exogenous 24-epibrassinolide enhances proanthocyanidin biosynthesis in *Vitis vinifera* ‘Cabernet Sauvignon’ berry skin. *Plant Growth Regulation* **75**, 741–750. doi: [10.1007/s10725-014-9976-y](https://doi.org/10.1007/s10725-014-9976-y).

Xu W, Grain D, Bobet S, Le Gourrierc J, Thévenin J, Kelemen Z, Lepiniec L, Dubos C. 2014. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in *Arabidopsis* seed. *New Phytologist* **202**, 132–144. doi: [10.1111/nph.12620](https://doi.org/10.1111/nph.12620).

Xu L, Liu F, Lechner E, Genschik P, Crosby WL, Ma H, Peng W, Huang D, Xie D. 2002. The SCFCO11 ubiquitin-ligase complexes are required for jasmonate response in *Arabidopsis*. *Plant Cell* **14**, 1919–1935. doi: [10.1105/tpc.003368](https://doi.org/10.1105/tpc.003368).

Xu YQ, Zhang YN, Chen JX, Wang F, Du QZ, Yin JF. 2018. Quantitative analyses of the bitterness and astringency of catechins from green tea. *Food Chemistry* **258**, 16–24. doi: [10.1016/j.foodchem.2018.03.042](https://doi.org/10.1016/j.foodchem.2018.03.042).

Yan Y, Stolz S, Chételat A, Reymond P, Pagni M, Dubugnon L, Farmer EE. 2007. A downstream mediator in the growth repression limb of the jasmonate pathway. *Plant Cell* **19**, 2470–2483. doi: [10.1105/tpc.107.050708](https://doi.org/10.1105/tpc.107.050708).

Yan L, Zhai Q, Wei J, et al. 2013. Role of tomato lipoxygenase D in wound-induced

jasmonate biosynthesis and plant immunity to insect herbivores. *PLOS Genetics* **9**, e1003964. doi: [10.1371/journal.pgen.1003964](https://doi.org/10.1371/journal.pgen.1003964).

Ye CY, Wu D, Mao L, et al. 2020. The genomes of the allohexaploid *Echinochloa crus-galli* and its progenitors provide insights into polyploidization-driven Adaptation. *Molecular Plant* **13**, 1298–1310. doi: [10.1016/j.molp.2020.07.001](https://doi.org/10.1016/j.molp.2020.07.001).

Ye Y, Lu Y, Wang G, Liu Y, Zhang Y, Tang H. 2021. Stable reference gene selection for qRT-PCR normalization in strawberry (*Fragaria x ananassa*) leaves under different stress and light-quality conditions. *Horticulturae* **7**, 11. doi: [10.3390/horticulturae7110452/s1](https://doi.org/10.3390/horticulturae7110452/s1).

Yoeun S, Rakwal R, Han O. 2013. Dual positional substrate specificity of rice allene oxide synthase-1: Insight into mechanism of inhibition by type II ligand imidazole. *BMB Reports* **46**, 151–156. doi: [10.5483/bmbrep.2013.46.3.117](https://doi.org/10.5483/bmbrep.2013.46.3.117).

Yu D, Huang T, Tian B, Zhan J. 2020a. Advances in biosynthesis and biological functions of proanthocyanidins in horticultural plants. *Foods* **9**, 1774. doi: [10.3390/foods9121774](https://doi.org/10.3390/foods9121774).

Yu K, Jun JH, Duan C, Dixon RA. 2019. VvLAR1 and VvLAR2 are bifunctional enzymes for proanthocyanidin biosynthesis in grapevine. *Plant Physiology* **180**, 1362–1374. doi: [10.1104/pp.19.00447](https://doi.org/10.1104/pp.19.00447).

Yu CW, Lin YT, Li H min. 2020b. Increased ratio of galactolipid MGDG : DGDG induces jasmonic acid overproduction and changes chloroplast shape. *New Phytologist* **228**, 1327–1335. doi: [10.1111/nph.16766](https://doi.org/10.1111/nph.16766).

Yu G, Xu S. 2021. ggbreak: Set Axis Break for ‘ggplot2’. doi: [10.32614/cran.package.ggbreak](https://doi.org/10.32614/cran.package.ggbreak).

Zabetakis I, Holden MA. 1997. Strawberry flavour: Analysis and biosynthesis. *Journal of the Science of Food and Agriculture* **74**, 421–434. doi: [10.1002/\(SICI\)1097-0010\(199708\)74:4<421::AID-JSFA817>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0010(199708)74:4<421::AID-JSFA817>3.0.CO;2-6).

Zhang S, Hou B, Chai L, Yang A, Yu X, Shen Y. 2017. Sigma factor *FaSigE* positively regulates strawberry fruit ripening by ABA. *Plant Growth Regulation* **83**, 417–427. doi: [10.1007/s10725-017-0308-x](https://doi.org/10.1007/s10725-017-0308-x).

Zhang Z, Lu S, Yu W, Ehsan S, Zhang Y, Jia H, Fang J. 2022. Jasmonate increases terpene synthase expression, leading to strawberry resistance to *Botrytis cinerea* infection. *Plant Cell Reports* **41**, 1243–1260. doi: [10.1007/s00299-022-02854-1](https://doi.org/10.1007/s00299-022-02854-1).

Zhang J, Wang X, Yu O, Tang J, Gu X, Wan X, Fang C. 2011. Metabolic profiling of strawberry (*Fragaria × ananassa* Duch.) during fruit development and maturation. *Journal of Experimental Botany* **62**, 1103–1118. doi: [10.1093/jxb/erq343](https://doi.org/10.1093/jxb/erq343).

Zhou R, Cutler AJ, Ambrose SJ, et al. 2004. A New abscisic acid catabolic pathway. *plant physiology* **134**, 361–369. doi: [10.1104/pp.103.030734](https://doi.org/10.1104/pp.103.030734).

Zhou F, Pichersky E. 2020. The complete functional characterisation of the terpene synthase family in tomato. *New Phytologist* **226**, 1341–1360. doi: [10.1111/nph.16431](https://doi.org/10.1111/nph.16431).

Zhou Y, Xiong J, Shu Z, et al. 2023. The telomere-to-telomere genome of *Fragaria vesca* reveals the genomic evolution of *Fragaria* and the origin of cultivated octoploid strawberry. *Horticulture Research* **10**. doi: [10.1093/hr/uhad027](https://doi.org/10.1093/hr/uhad027).

- Zhou Y, Zheng J, Li Y, Xu DP, Li S, Chen YM, Li H Bin.** 2016. Natural polyphenols for prevention and treatment of cancer. *Nutrients* **8**, 515. doi: [10.3390/nu8080515](https://doi.org/10.3390/nu8080515).
- Zorrilla-Fontanesi Y, Cabeza A, Domínguez P, Medina JJ, Valpuesta V, Denoyes-Rothan B, Sánchez-Sevilla JF, Amaya I.** 2011. Quantitative trait *loci* and underlying candidate genes controlling agronomical and fruit quality traits in octoploid strawberry (*Fragaria* × *ananassa*). *Theoretical and Applied Genetics* 2011 123:5 **123**, 755–778. doi: [10.1007/S00122-011-1624-6](https://doi.org/10.1007/S00122-011-1624-6).
- Zorrilla-Fontanesi Y, Rambla JL, Cabeza A, Medina JJ, Sánchez-Sevilla JF, Valpuesta V, Botella MA, Granell A, Amaya I.** 2012. Genetic analysis of strawberry fruit aroma and identification of *O*-methyltransferase FaOMT as the locus controlling natural variation in mesifurane content. *Plant Physiology* **159**, 851–870. doi: [10.1104/pp.111.188318](https://doi.org/10.1104/pp.111.188318).
- Zuanny DC, Vilela B, Moonlight PW, Särkinen TE, Cardoso D.** 2024. expowo: An R package for mining global plant diversity and distribution data. *Applications in Plant Sciences* **12**, e11609. doi: [10.1002/aps3.11609](https://doi.org/10.1002/aps3.11609).
- Zuther E, Schulz E, Childs LH, Hinch DK.** 2012. Clinal variation in the non-acclimated and cold-acclimated freezing tolerance of *Arabidopsis thaliana* accessions. *Plant, Cell and Environment* **35**, 1860–1878. doi: [10.1111/j.1365-3040.2012.02522.x](https://doi.org/10.1111/j.1365-3040.2012.02522.x).

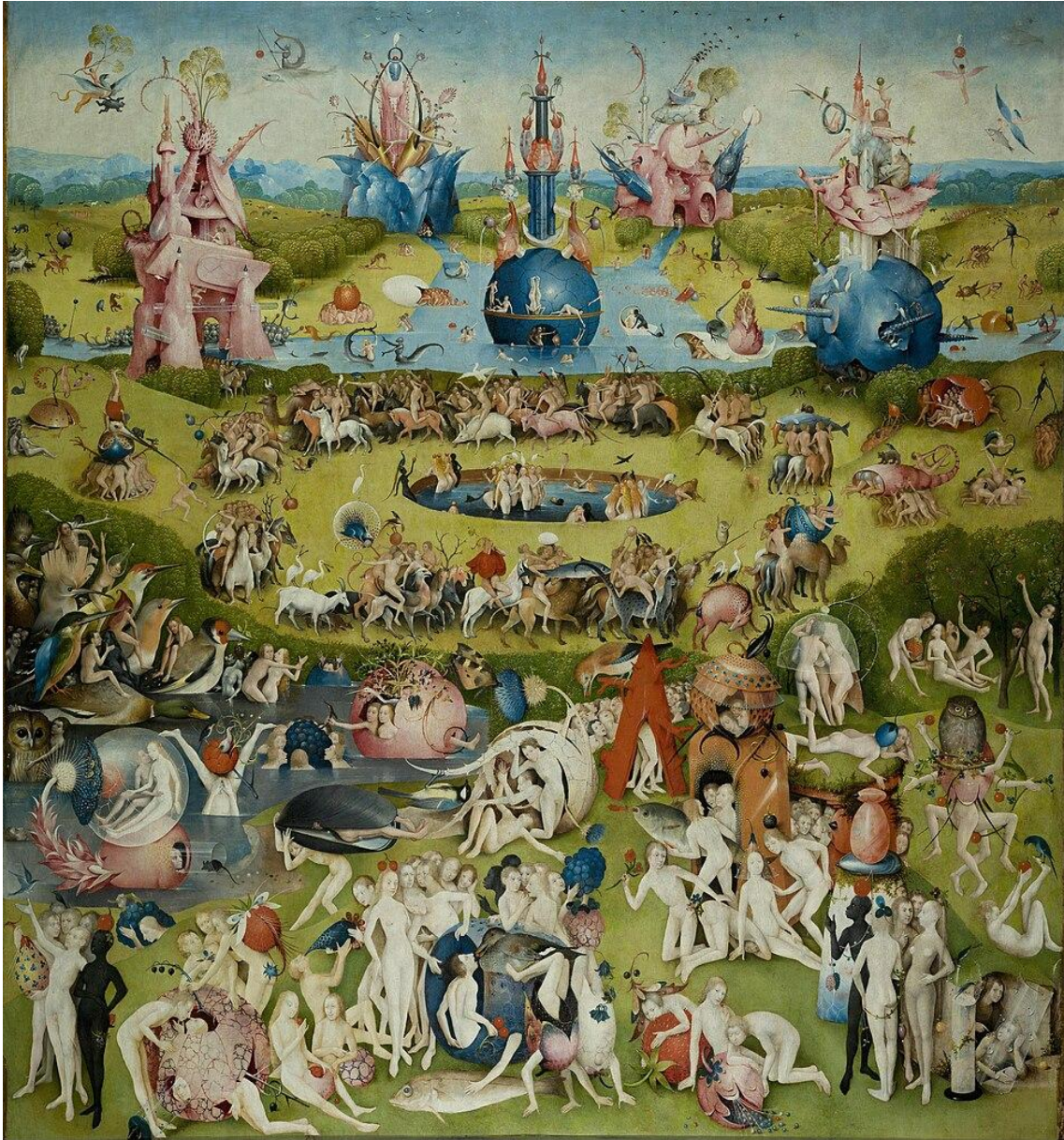
IX. Abbreviation



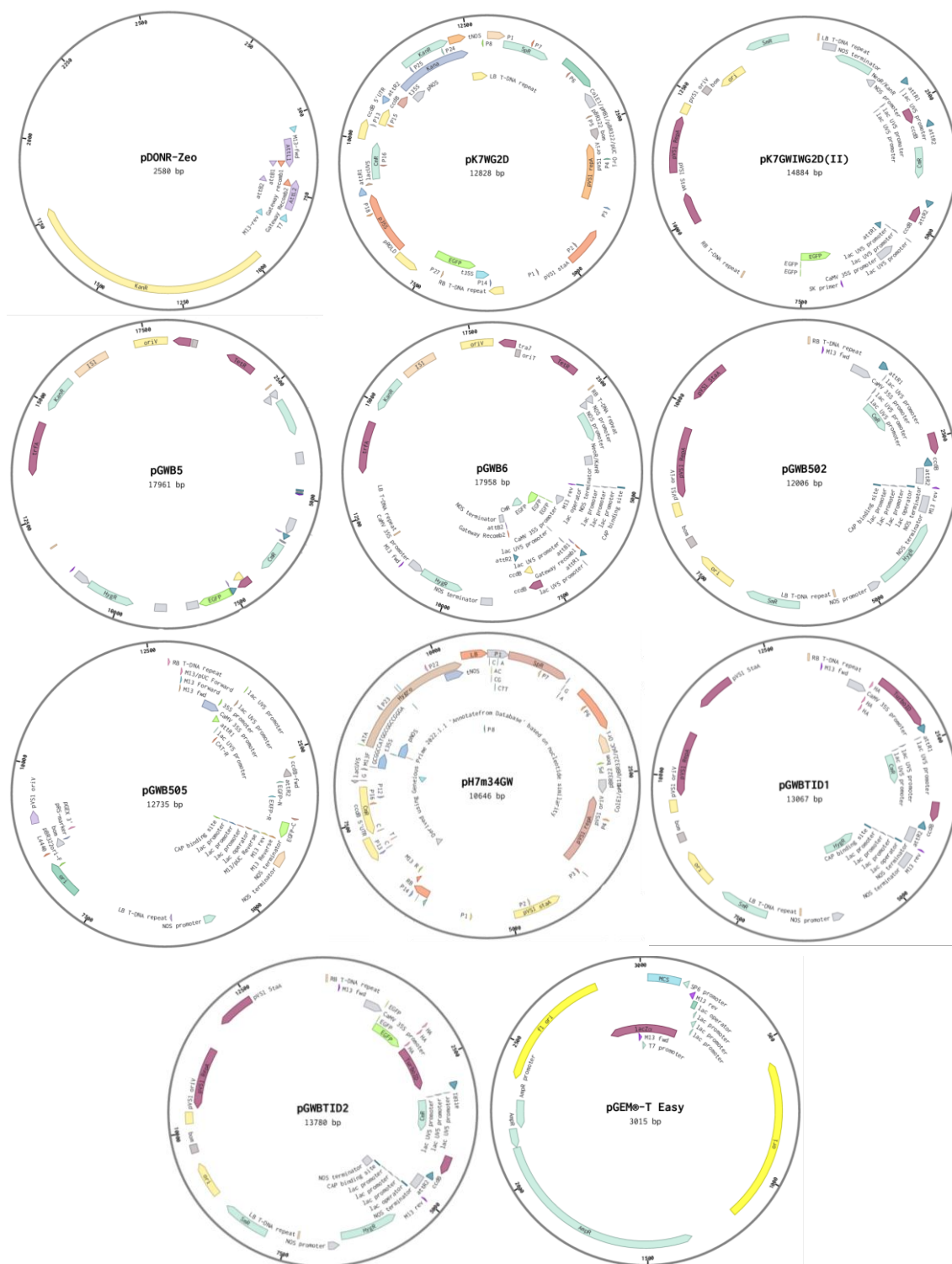
Venus and Adonis, Antonio Canova (c. XIX). In Greek mythology, strawberries are associated with Aphrodite (Venus in Roman mythology), due to their heart-shaped form and red color, which evoke attributes traditionally linked to the goddess. According to one myth, strawberries are said to have originated from her tears, shed as she mourned the death of her lover Adonis. These tears, mingled with Adonis's blood, are believed to have given rise to the first strawberry plants.

α-LeA: α -linolenic acid	LC/MS: Liquid chromatography-mass spectrometry
ABA: abscisic acid	LOD: logarithm of the odds ratio
cDNA: complementary DNA	LOX: lipoxygenase
CDS: Coding sequence	m/z: mass-to-charge ratio
Chr: chromosome	MeJA: methyl jasmonate
CIM: composite interval mapping	PA: proanthocyanidin
CPM: Counts per Million	PUFA: polyunsaturated fatty acid
EAH2: Ellagic acid hexose 2	pTM: predicted template moleling
ET: ellagitannin	QTL: Quantitative trait <i>loci</i>
FAD: Fatty acid desaturase	RFP: Red Fluorescent Protein
FaUDP: <i>FxC_3g11070</i> gene (CDS from cv. 'Camarosa' reference genome)	RNAi: RNA interference
FaUDPSS: <i>Fragaria</i> $\times$ <i>ananassa</i> UDP-glycosyltransferase (CDS from cv. 'Senga Sengana')	SIM: simple interval mapping
FaUDPSS: <i>FxC_3g11070</i> gene (CDS from cv. 'Senga Sengana')	TE: Low transposable elements
FC: Fold Change	TID: TurboID
FDR: False Discovery Ratio	TPM: Transcripts Per Kilobase Million
FPKM: Fragments Per Kilobase Million	UGT: UDP-dependent glycosyltransferase
GBP: Goodberry project	VOC: Volatile organic compound
GC/MS: Gas chromatography-mass spectrometry	
gDNA: genomic DNA	
GDR: Genome Database for Rosaceae	
GFP: Green Fluorescent Protein	
HS-GC/MS: headspace-gas chromatography/mass spectrometry	
ipTM: interface predicted template modeling	
JA: jasmonic acid/jasmonate	

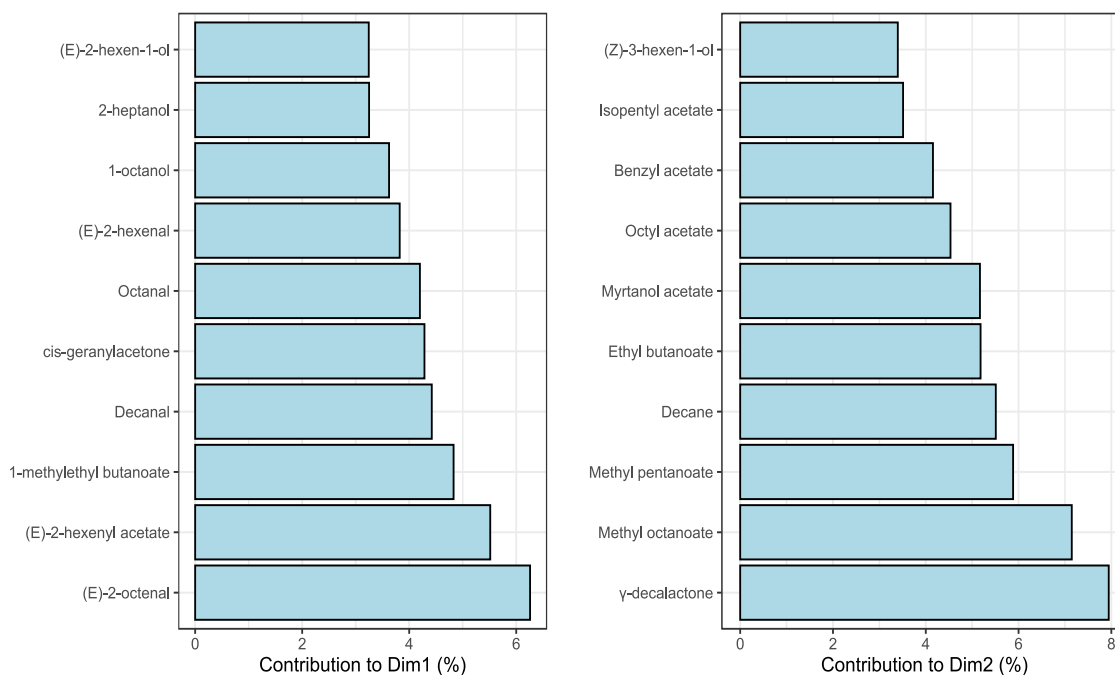
X. Annexes



The Garden of Earthly Delights, Hieronymus Bosch (c. XVI). Representations of strawberry fruits are observed, as an allusion to fleeting carnal pleasures.



Annex 1. Maps of genetic vectors used in this work. All vectors represent empty vectors (no exogenous genes cloned in the sequence) from **Table 7A**. Some annotations are shown in the figures for Gateway cloning (*attB1*, *attB2*, *attL1* and *attL2*), antibiotic resistances (*AmpR*, *CmR*, *KanR*, *NeoR*, *SmR*), promoters and terminators (*CaMV35S* and *NOS* terminator) and more.



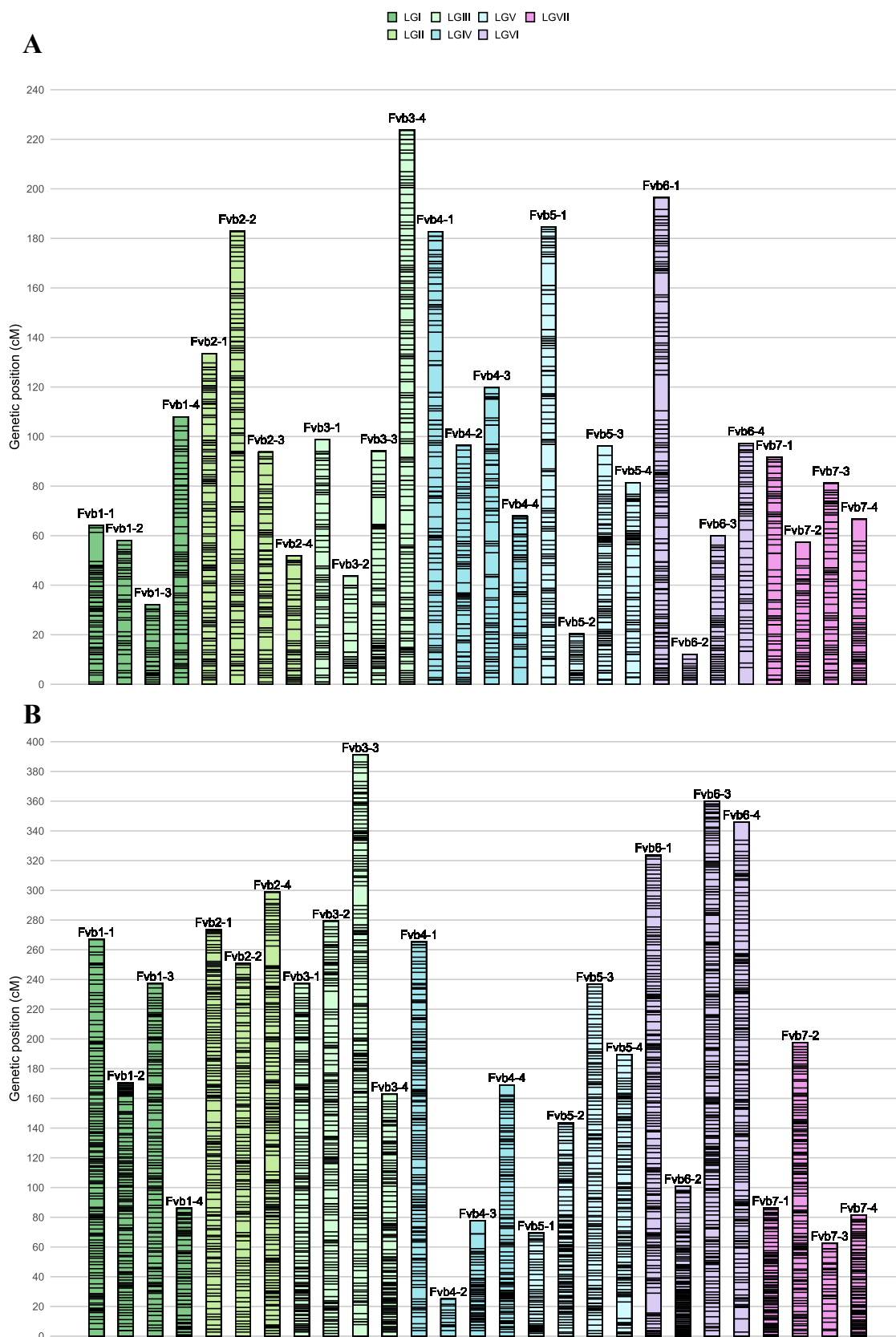
Annex 2. Metabolites contributing to PCA dimensions for Goodberry population. In this figure are present the top 10 metabolites whose contributions were the highest to dimension 1 and 2, respectively, of the PCA for VOCs in Goodberry population.

Annex 3. Heritability of volatile organic compounds in Goodberry population. Heritability for each volatile was calculated by applying the variance components derived from ANOVA ($h^2 = V_G/V_P$, where V_G is the total genetic variance and V_P represents the total phenotypic variance). GE: Germany, FR: France, SP: Spain, PL: Poland, IT: Italy.

Metabolite	GE	FR	SP	PL	IT
(-)-myrtenol	0.78	0.71	0.86	0.86	0.82
(E)-2-heptenal	0.67	0.49	0.73	0.74	0.93
(E)-2-hexen-1-ol	0.68	0.74	0.72	0.58	0.99
(E)-2-hexenal	0.56	0.56	0.68	0.74	0.90
(E)-2-hexenyl acetate	0.74	nd	nd	nd	nd
(E)-2-nonenal	0.85	0.39	-0.01	0.24	0.74
(E)-2-octenal	0.85	0.68	0.82	0.70	0.97
(E)-2-pentenal	0.63	0.66	0.73	0.87	0.95
(E,Z)-2,4-heptadienal	0.78	0.44	0.59	0.77	0.91
(Z)-3-hexen-1-ol	0.12	0.60	0.28	0.50	0.62
(Z)-3-hexenal	0.56	0.40	0.69	0.37	0.86
1-hexanol	0.57	0.89	0.72	0.43	0.91
1-methylethyl butanoate	0.58	0.79	0.78	0.83	0.94
1-octanol	0.56	0.68	0.47	0.63	0.76
1-penten-3-ol	0.25	0.67	0.79	0.33	0.96
1-penten-3-one	0.61	0.60	0.81	0.53	0.86

2-heptanol	0.69	0.54	0.45	0.68	0.90
2-heptanone	0.90	0.52	0.63	0.69	0.94
2-nonanol	0.86	0.73	0.84	0.58	0.94
2-nonanone	0.69	0.77	0.77	0.17	0.95
2-pentanone	0.47	0.50	0.61	0.66	0.71
2-pentyl butyrate	0.68	0.74	0.66	0.42	0.80
2-pentylfuran	0.79	0.37	0.61	0.79	0.86
3,4-dimethyl benzaldehyde	0.86	0.78	0.51	0.72	0.90
6-methyl-5-hepten-2-one	0.59	0.76	0.82	0.43	0.93
Acetone	1.00	1.00	1.00	1.00	1.00
Acetophenone	0.60	0.75	0.71	0.53	0.92
Benzaldehyde	0.61	0.78	0.65	0.81	0.84
Benzyl acetate	0.41	0.64	0.51	0.53	0.69
Butyl acetate	0.47	0.34	0.56	0.48	0.58
Butyl butanoate	0.75	0.80	0.62	0.85	0.87
Butyl hexanoate	0.93	0.90	0.97	0.90	0.99
Cinnamyl acetate	0.32	0.50	0.63	0.13	0.70
<i>cis</i> -geranylacetone	0.38	0.17	0.17	0.18	0.71
Decanal	0.77	0.21	0.52	0.19	0.53
Decane	0.50	0.68	0.69	0.36	0.89
Decyl acetate	0.06	0.53	0.53	0.13	0.95
Ethyl acetate	0.77	nd	nd	0.66	0.94
Ethyl butanoate	0.53	0.52	0.73	0.54	0.74
Ethyl hexanoate	0.70	0.50	0.70	0.82	0.59
Ethyl octanoate	nd	0.41	0.72	nd	0.83
Eugenol	0.15	0.62	0.16	0.38	0.87
Furaneol	0.76	0.63	0.83	0.81	0.92
Heptanal	0.76	0.85	0.87	0.90	0.94
Hexanal	0.80	0.28	0.49	0.20	0.88
Hexyl acetate	0.68	0.51	0.66	0.81	0.92
Hexyl butanoate	0.88	0.49	0.72	0.80	0.85
Hexyl hexanoate	0.87	0.83	0.88	0.92	0.98
Isopentyl acetate	0.61	0.60	0.61	0.64	0.81
Isopentyl butanoate	0.28	nd	0.82	0.67	0.86
Limonene	0.40	0.61	0.60	0.28	0.94
Linalool	-0.01	0.07	0.03	0.20	0.06
Mesifurane	0.80	0.80	0.82	0.80	0.93
Methyl (<i>E</i>)-2-hexenoate	0.66	0.37	0.68	0.78	nd
Methyl 3- hydroxyoctanoate	0.82	0.35	0.16	0.68	0.79
Methyl acetate	0.82	0.86	0.88	0.88	0.89
Methyl benzoate	0.06	0.66	0.05	0.18	0.66
Methyl butyrate	0.27	0.56	0.60	0.32	0.95
Methyl decanoate	0.67	0.93	1.00	0.66	0.61
Methyl dodecanoate	nd	0.90	nd	nd	nd
Methyl hexanoate	0.89	0.60	0.81	0.68	0.98

Metabolite	GE	FR	SP	PL	IT
Methyl octanoate	0.85	0.57	0.48	0.47	0.99
Methyl pentanoate	0.74	0.11	0.29	0.50	0.94
Myrtanol acetate	0.65	0.69	0.02	0.43	0.75
Nerol	0.63	0.55	0.63	0.44	0.67
Nerolidol	0.94	0.04	0.97	0.93	0.94
Nonanal	0.35	0.74	0.93	0.86	0.96
Nonyl acetate	0.41	0.33	0.28	0.31	0.65
Octanal	0.78	0.69	0.59	0.87	0.93
Octyl acetate	0.63	0.21	0.37	0.80	0.57
Octyl butanoate	0.59	0.61	0.51	0.59	0.75
Pentanal	0.74	0.82	0.88	0.84	0.91
Pentyl acetate	0.49	0.67	0.36	0.46	0.85
Prenyl acetate	0.73	0.20	0.73	0.66	0.96
Propyl butanoate	0.80	0.78	0.92	0.85	0.86
S-methyl butanethioate	0.83	0.32	0.84	0.55	0.92
Terpineol	0.46	0.50	0.59	0.55	0.82
<i>trans</i> -2-decenal	0.13	0.63	0.62	0.40	0.15
β -ionone	0.16	0.10	0.14	0.01	0.87
γ -decalactone	0.70	0.66	0.79	0.49	0.92
γ -dodecalactone	0.81	0.78	0.83	0.78	0.93



Annex 4. Representation of ‘Candonga’ x ‘Senga Sengana’ linkage maps. Seven different linkage groups (LGs) were found. Horizontal lines within each chromosome represent genetic markers used in the mapping analysis. (A) Female map; (B) Male map.

Annex 5. Summary of the linkage maps of the segregating population ‘Candonga’ × ‘Senga Sengana’. Each subdivision of the seven linkage groups (LG) was represented with correspondent number of genetic markers and size. Marker density was calculated as marker/cM. Linkage groups were assigned according to the nomenclature of ‘Camarosa’ reference genome (Edger *et al.*, 2019)

LG	Female Map			Male Map		
	Markers	Size (cM)	Density	Markers	Size (cM)	Density
Fvb1-1	84.00	64.07	1.31	148.00	266.84	0.55
Fvb1-2	50.00	58.05	0.86	137.00	170.51	0.80
Fvb1-3	65.00	32.12	2.02	132.00	237.12	0.56
Fvb1-4	88.00	107.93	0.82	71.00	85.95	0.83
Fvb2-1	151.00	49.94	3.02	158.00	273.56	0.58
Fvb2-2	138.00	182.76	0.76	112.00	250.76	0.45
Fvb2-3	104.00	93.96	1.11	-	-	-
Fvb2-4	53.00	51.86	1.02	179.00	298.65	0.60
Fvb3-1	78.00	98.74	0.79	162.00	237.12	0.68
Fvb3-2	38.00	43.63	0.87	154.00	279.14	0.55
Fvb3-3	114.00	94.04	1.21	207.00	391.09	0.53
Fvb3-4	182.00	223.60	0.81	137.00	162.77	0.84
Fvb4-1	121.00	182.71	0.66	157.00	265.18	0.59
Fvb4-2	92.00	96.35	0.95	17.00	24.96	0.68
Fvb4-3	84.00	119.74	0.70	63.00	77.66	0.81
Fvb4-4	60.00	67.83	0.88	112.00	168.97	0.66
Fvb5-1	152.00	184.51	0.82	61.00	69.26	0.88
Fvb5-2	38.00	20.22	1.88	141.00	143.62	0.98
Fvb5-3	130.00	96.17	1.35	131.00	236.75	0.55
Fvb5-4	73.00	81.37	0.90	131.00	189.47	0.69
Fvb6-1	151.00	196.41	0.77	191	323.82	0.59
Fvb6-2	19.00	12.07	1.57	181	100.87	1.79
Fvb6-3	67.00	59.92	1.12	237	360.00	0.66
Fvb6-4	67.00	97.23	0.69	199	345.89	0.58
Fvb7-1	76.00	91.59	0.83	120	85.99	1.40
Fvb7-2	74.00	57.34	1.29	133	197.50	0.67
Fvb7-3	96.00	81.10	1.18	40	62.40	0.64
Fvb7-4	75.00	66.59	1.13	109	81.36	1.34
Total	2,520.00	2,611.84		3,620.00	5,387.19	

Annex 6. Metabolic QTLs (mQTLs) detected using Single Interval Mapping (SIM). Spain (SP), Italy (IT), France (FR), Poland (PL) and Germany (GE); Male (M), Female (F); Chr, *F.* × *ananassa* ‘Camarosa’ chromosome; Pos, genetic position in cM; ci-low and ci-up, lower and upper genetic position of the Bayesian interval in cM, respectively; LOD, logarithm of the odds ratio; a, mean phenotypic difference between the two homozygous loci of the QTL; r^2 , percentage of phenotypic variance explained by the QTL.

QTL name	Compound	Map	Chr	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r^2	p-value
qf1Hex3	1-hexanol	F	Fvb3-2	257186	SP	12.97	2.00	12.97	4.14	1.15	14.48	0.00
qf1Hex4	1-hexanol	F	Fvb3-2	257186	PL	12.97	2.00	12.97	3.75	-1.84	13.94	0.00
qf1Oct6	1-octanol	F	Fvb3-2	263121	FR	39.89	28.79	43.63	3.17	2.84	11.47	0.02
qf1Oct7	1-octanol	F	Fvb3-2	257185	FR	12.97	4.00	12.97	4.73	0.72	16.59	0.00
qf1Oct8	1-octanol	F	Fvb3-2	cF3-2-b,loc8	GE	8.00	2.00	12.97	5.35	-0.07	18.44	0.00
qf1Oct9	1-octanol	F	Fvb3-2	257189	SP	12.97	4.00	12.97	7.98	0.58	26.00	0.00
qf1Oct10	1-octanol	F	Fvb3-2	2571817	PL	12.97	4.00	12.97	5.11	-0.01	18.50	0.00
qf1Pentol1	1-penten-3-ol	F	Fvb1-3	AX-166503026_Fa1-3	GE	25.72	0.00	32.12	1.89	1.07	7.06	0.00
qf1Pentol2	1-penten-3-ol	F	Fvb3-2	AX-166512951_Fa3-22	FR	9.32	2.00	12.97	3.20	0.19	11.63	0.00
qf2Hept3	2-heptanone	F	Fvb2-2	AX-184030860_Fa2-2	GE	23.14	0.00	182.76	2.89	0.78	10.43	0.05
qf2Hept4	2-heptanone	F	Fvb3-2	257187	PL	12.97	2.00	12.97	2.83	-0.82	10.71	0.02
qf2Nonol2	2-nonanol	F	Fvb3-2	2571818	PL	12.97	4.00	12.97	3.57	0.51	13.33	0.00
qf2None2	2-nonanone	F	Fvb3-2	102121	GE	4.72	2.00	12.97	3.14	0.00	11.28	0.02
qf34Benz3	3,4-dimethyl benzaldehyde	F	Fvb3-2	263123	FR	39.89	28.79	43.63	3.70	3.00	13.24	0.01
qf34Benz4	3,4-dimethyl benzaldehyde	F	Fvb3-2	2571810	FR	12.97	4.00	12.97	4.80	0.40	16.81	0.00
qf34Benz5	3,4-dimethyl benzaldehyde	F	Fvb3-2	2571825	PL	12.97	2.00	12.97	3.49	0.00	13.05	0.00
qf6M5Hep5	6-methyl-5-hepten-2-one	F	Fvb3-2	AX-89887538_Fa3-2	SP	39.89	26.00	43.63	3.33	2.47	11.90	0.02
qf6M5Hep6	6-methyl-5-hepten-2-one	F	Fvb3-2	257182	FR	12.97	4.00	12.97	4.22	0.58	14.96	0.00
qf6M5Hep7	6-methyl-5-hepten-2-one	F	Fvb3-2	25718	GE	12.97	2.00	12.97	4.21	0.02	14.79	0.00

QTL name	Compound	Map	Chr	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qf6M5Hep8	6-methyl-5-hepten-2-one	F	Fvb3-2	257183	SP	12.97	4.00	12.97	7.05	0.30	23.53	0.00
qf6M5Hep9	6-methyl-5-hepten-2-one	F	Fvb3-2	2571811	PL	12.97	2.00	12.97	2.51	-0.08	9.55	0.01
qfAcphne8	Acetophenone	F	Fvb1-4	AX-184962799_Fa1-4	SP	74.69	54.00	107.93	2.92	2.41	10.43	0.02
qfAcphne9	Acetophenone	F	Fvb3-2	AX-166512951_Fa3-21	FR	9.32	2.00	12.97	4.39	0.17	15.51	0.00
qfAcphne10	Acetophenone	F	Fvb3-2	AX-184647040_Fa3-2	GE	7.50	2.00	12.97	4.82	0.04	16.77	0.00
qfAcphne11	Acetophenone	F	Fvb3-2	AX-166512951_Fa3-21	SP	9.32	2.00	12.97	3.86	0.14	13.56	0.00
qfAcphne12	Acetophenone	F	Fvb3-2	AX-166512951_Fa3-21	PL	9.32	4.00	12.97	6.82	0.32	23.90	0.00
qfAcphne13	Acetophenone	F	Fvb3-2	AX-166512951_Fa3-2	IT	9.32	2.00	12.97	2.57	-0.07	9.25	0.03
qfActone3	Acetone	F	Fvb1-1	AX-123363457_Fa1-1	GE	23.14	2.00	64.07	2.62	9.71	9.50	0.05
qfActone4	Acetone	F	Fvb1-3	88	GE	0.00	0.00	23.89	8.08	3.96	26.48	0.00
qfActone5	Acetone	F	Fvb1-3	AX-184085467_Fa1-31	GE	0.00	0.00	23.89	5.17	3.38	17.86	0.00
qfActone6	Acetone	F	Fvb3-2	25718	SP	12.97	4.72	12.97	5.02	0.75	17.27	0.00
qfActone7	Acetone	F	Fvb3-2	25718	PL	12.97	4.00	12.97	7.26	-0.07	25.23	0.00
qfActone8	Acetone	F	Fvb3-3	25627	SP	25.79	0.00	37.78	4.10	5.59	14.34	0.00
qfActone9	Acetone	F	Fvb3-3	AX-184732992_Fa3-3	PL	13.41	0.00	37.78	3.36	2.57	12.57	0.01
qfBenz4	Benzaldehyde	F	Fvb3-2	AX-166512951_Fa3-2	FR	9.32	2.00	12.97	3.60	-0.27	12.89	0.00
qfBenz5	Benzaldehyde	F	Fvb3-2	257182	SP	12.97	2.00	12.97	4.44	0.33	15.42	0.00
qfBenz6	Benzaldehyde	F	Fvb3-2	2571810	PL	12.97	4.00	12.97	5.13	0.08	18.58	0.00
qfBenzAc2	Benzyl acetate	F	Fvb3-2	257181	FR	12.97	2.00	12.97	2.90	0.12	10.55	0.01
qfBione3	β -ionone	F	Fvb3-2	257181	FR	12.97	2.00	12.97	3.42	-0.11	12.81	0.00
qfDecAc5	Decyl acetate	F	Fvb3-2	AX-166512951_Fa3-22	FR	9.32	0.00	12.97	1.81	0.06	6.72	0.01
qfDecAc6	Decyl acetate	F	Fvb3-2	2571831	PL	12.97	2.00	12.97	2.58	-0.09	9.83	0.00
qfDecAc7	Decyl acetate	F	Fvb3-4	9278	GE	2.74	0.00	219.93	2.56	4.40	9.29	0.04
qfDecAc8	Decyl acetate	F	Fvb4-2	cF4-2-a,loc8	PL	8.00	0.00	96.35	1.90	1.42	7.34	0.04
qfDecal3	Decanal	F	Fvb3-2	257182	GE	12.97	4.00	12.97	10.96	0.51	34.10	0.00

qfDecal4	Decanal	F	Fvb3-2	257181	SP	12.97	2.00	12.97	3.17	0.55	11.29	0.02
qfDecal5	Decanal	F	Fvb3-2	2571822	PL	12.97	4.00	12.97	6.37	0.12	22.50	0.00
qfDecLac5	γ -decalactone	F	Fvb3-2	AX-184260226_Fa3-2	IT	38.98	34.00	43.63	5.15	0.16	17.66	0.00
qfDecLac6	γ -decalactone	F	Fvb3-2	26312	FR	39.89	36.20	43.63	6.60	0.27	22.38	0.00
qfDecLac7	γ -decalactone	F	Fvb3-2	AX-184260226_Fa3-2	SP	38.98	32.53	43.63	5.18	0.30	17.74	0.00
qfDecLac8	γ -decalactone	F	Fvb3-2	cF3-2-a,loc40	PL	40.00	34.37	43.63	5.61	0.14	20.11	0.00
qfDecLac9	γ -decalactone	F	Fvb3-2	AX-166512951_Fa3-24	IT	9.32	6.00	12.97	27.31	0.46	64.34	0.00
qfDecLac10	γ -decalactone	F	Fvb3-2	257183	FR	12.97	8.41	12.97	26.79	0.79	64.23	0.00
qfDecLac11	γ -decalactone	F	Fvb3-2	AX-166512951_Fa3-26	SP	9.32	8.41	10.00	32.42	0.87	70.59	0.00
qfDecLac12	γ -decalactone	F	Fvb3-2	AX-166512951_Fa3-27	PL	9.32	8.41	12.97	36.03	0.42	76.38	0.00
qfDecLac13	γ -decalactone	F	Fvb4-2	10997	GE	27.99	0.00	96.35	3.03	0.73	10.90	0.03
qfDecLac14	γ -decalactone	F	Fvb6-2	17651	GE	0.00	0.00	10.15	2.61	-7.58	9.46	0.04
qfDodeLac7	γ -dodecalactone	F	Fvb3-2	AX-184260226_Fa3-21	IT	38.98	24.00	43.63	3.33	0.23	11.81	0.01
qfDodeLac8	γ -dodecalactone	F	Fvb3-2	AX-89887538_Fa3-2	PL	39.89	30.00	43.63	3.03	0.25	11.41	0.05
qfDodeLac9	γ -dodecalactone	F	Fvb3-2	cF3-2-b,loc8	IT	8.00	2.00	12.97	4.52	-0.01	15.69	0.00
qfDodeLac10	γ -dodecalactone	F	Fvb3-2	257184	FR	12.97	4.00	12.97	7.47	0.00	24.93	0.00
qfDodeLac11	γ -dodecalactone	F	Fvb3-2	25718	SP	12.97	4.00	12.97	5.90	0.06	19.96	0.00
qfDodeLac12	γ -dodecalactone	F	Fvb3-2	AX-166512951_Fa3-210	PL	9.32	4.00	12.97	11.05	-0.01	35.75	0.00
qfDodeLac13	γ -dodecalactone	F	Fvb7-4	24360	GE	11.08	0.00	60.00	3.76	2.90	13.34	0.00
qfE2Hex6	(E)-2-hexenal	F	Fvb1-4	cF1-4,loc74	SP	74.00	54.00	107.93	3.03	4.04	10.80	0.04
qfE2Hex7	(E)-2-hexenal	F	Fvb2-1	6594	GE	0.00	0.00	50.00	2.13	0.92	7.99	0.05
qfE2Hex8	(E)-2-hexenal	F	Fvb3-2	cF3-2-b,loc12	FR	12.00	2.00	12.97	2.76	-0.12	10.05	0.01
qfE2Hex9	(E)-2-hexenal	F	Fvb3-2	257182	SP	12.97	2.00	12.97	4.93	-0.30	16.99	0.00
qfE2Hex10	(E)-2-hexenal	F	Fvb3-2	257182	PL	12.97	2.00	12.97	2.63	-0.03	10.00	0.05
qfE2Hexol6	(E)-2-hexen-1-ol	F	Fvb3-2	AX-166512951_Fa3-2	FR	9.32	2.00	12.97	2.59	-1.76	9.46	0.00
qfE2Hexol7	(E)-2-hexen-1-ol	F	Fvb3-2	257184	SP	12.97	2.00	12.97	4.02	0.82	14.06	0.00
qfE2Hexol8	(E)-2-hexen-1-ol	F	Fvb3-2	257184	PL	12.97	2.00	12.97	4.07	-0.38	15.03	0.00

QTL name	Compound	Map	Chr	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qfE2Non5	(E)-2-nonenal	F	Fvb3-2	AX-89887538_Fa3-2	SP	39.89	26.00	43.63	3.05	2.95	10.87	0.03
qfE2Non6	(E)-2-nonenal	F	Fvb3-2	AX-166512951_Fa3-23	FR	9.32	2.00	12.97	3.32	0.51	11.96	0.01
qfE2Non7	(E)-2-nonenal	F	Fvb3-2	cF3-2-b,loc12	GE	12.00	0.00	12.97	2.28	-0.30	8.30	0.00
qfE2Non8	(E)-2-nonenal	F	Fvb3-2	25718	SP	12.97	4.00	12.97	5.03	0.57	17.29	0.00
qfE2Non9	(E)-2-nonenal	F	Fvb3-2	2571821	PL	12.97	4.00	12.97	3.94	-0.19	14.59	0.00
qfE2Oc3	(E)-2-octenal	F	Fvb3-2	AX-89887538_Fa3-21	SP	39.89	28.79	43.63	3.08	2.45	10.99	0.03
qfE2Oc4	(E)-2-octenal	F	Fvb3-2	AX-89887538_Fa3-21	SP	39.89	28.79	43.63	3.08	2.45	10.99	0.03
qfE2Oc5	(E)-2-octenal	F	Fvb3-2	257181	GE	12.97	2.00	12.97	3.73	0.12	13.24	0.00
qfE2Oc6	(E)-2-octenal	F	Fvb3-2	25718	GE	12.97	2.00	12.97	2.74	-0.16	9.90	0.02
qfE2Oc7	(E)-2-octenal	F	Fvb3-2	257187	SP	12.97	4.00	12.97	6.49	0.48	21.72	0.00
qfE2Oc8	(E)-2-octenal	F	Fvb3-2	2571816	PL	12.97	2.00	12.97	2.50	0.07	9.54	0.03
qfEug6	Eugenol	F	Fvb3-2	2571811	FR	12.97	0.00	12.97	1.86	-1.27	6.91	0.05
qfEug7	Eugenol	F	Fvb3-2	AX-184670931_Fa3-2	SP	10.23	2.00	12.97	2.70	-0.54	9.70	0.01
qfEug8	Eugenol	F	Fvb3-2	cF3-2-b,loc12	PL	12.00	2.00	12.97	3.19	-0.30	11.99	0.00
qfEug9	Eugenol	F	Fvb3-2	cF3-2-b,loc6	IT	6.00	0.00	12.97	1.82	-0.16	6.63	0.02
qfEug10	Eugenol	F	Fvb3-4	8920	GE	53.43	0.00	89.68	3.75	6.99	13.30	0.00
qfEZHep5	(E,Z)-2,4-heptadienal	F	Fvb3-2	257183	FR	12.97	2.00	12.97	3.08	0.55	11.14	0.02
qfEZHep6	(E,Z)-2,4-heptadienal	F	Fvb3-2	102122	GE	4.72	2.00	12.97	3.27	0.02	11.72	0.01
qfEZHep7	(E,Z)-2,4-heptadienal	F	Fvb3-2	257186	SP	12.97	4.00	12.97	6.26	-0.01	21.05	0.00
qfEZHep8	(E,Z)-2,4-heptadienal	F	Fvb3-2	257186	SP	12.97	4.00	12.97	6.26	-0.01	21.05	0.00
qfEZHep9	(E,Z)-2,4-heptadienal	F	Fvb3-2	2571813	PL	12.97	4.00	12.97	5.22	0.18	18.85	0.00
qfFura8	Furaneol	F	Fvb1-3	AX-184512234_Fa1-3	FR	5.50	0.00	23.89	5.89	7.65	20.23	0.00
qfFura9	Furaneol	F	Fvb1-3	AX-184484824_Fa1-3	SP	17.44	0.00	23.89	6.78	3.57	22.75	0.00
qfFura10	Furaneol	F	Fvb1-3	AX-184512234_Fa1-31	PL	5.50	0.00	22.00	6.14	4.62	21.81	0.00

qfFura11	Furaneol	F	Fvb1-3	88	IT	0.00	0.00	20.00	7.37	5.17	24.27	0.00
qfFura12	Furaneol	F	Fvb3-2	AX-166512951_Fa3-22	GE	9.32	2.00	12.97	6.26	0.14	21.20	0.00
qfFura13	Furaneol	F	Fvb3-2	AX-184293605_Fa3-2	GE	11.14	2.00	12.97	5.78	0.08	19.76	0.00
qfFura14	Furaneol	F	Fvb3-2	AX-166512951_Fa3-24	GE	9.32	2.00	12.97	4.88	0.13	16.95	0.00
qfFura15	Furaneol	F	Fvb7-4	cF7-4-y.loc8	FR	8.00	0.00	62.00	2.78	5.54	10.12	0.04
qfFura16	Furaneol	F	Fvb7-4	AX-184765761_Fa7-2	IT	34.54	0.00	60.00	3.24	3.73	11.52	0.02
qfGeAc3	cis-geranilacetone	F	Fvb2-3	AX-184950666_Fa2-32	PL	22.75	0.00	50.66	1.30	3.43	5.06	0.04
qfGeAc4	cis-geranilacetone	F	Fvb3-2	cF3-2-b.loc8	FR	8.00	2.00	12.97	2.67	-0.19	9.74	0.01
qfHept1	Heptanal	F	Fvb2-3	30126	PL	27.47	0.00	50.66	3.54	1.93	13.22	0.01
qfHept2	Heptanal	F	Fvb3-2	25718	SP	12.97	2.00	12.97	2.74	2.57	9.83	0.03
qfHept3	Heptanal	F	Fvb3-2	257188	PL	12.97	4.00	12.97	4.67	0.16	17.07	0.00
qfHexHex3	Hexyl hexanoate	F	Fvb3-2	25718	FR	12.97	0.00	12.97	2.17	-16.40	8.01	0.04
qfIsoAc3	Isoamyl acetate	F	Fvb2-3	AX-184950666_Fa2-3	PL	22.75	0.00	50.66	2.85	8.62	11.27	0.02
qfLimo9	Limonene	F	Fvb3-2	AX-89887538_Fa3-2	FR	39.89	28.79	43.63	3.57	2.12	12.79	0.01
qfLimo10	Limonene	F	Fvb3-2	8491	SP	43.63	24.00	43.63	3.20	2.43	11.39	0.02
qfLimo11	Limonene	F	Fvb3-2	cF3-2-b.loc121	FR	12.00	4.00	12.97	8.36	-0.08	27.45	0.00
qfLimo12	Limonene	F	Fvb3-2	AX-166512951_Fa3-21	GE	9.32	2.00	12.97	3.51	0.09	12.51	0.01
qfLimo13	Limonene	F	Fvb3-2	102124	GE	4.72	2.00	12.97	3.27	0.02	11.72	0.02
qfLimo14	Limonene	F	Fvb3-2	AX-166512951_Fa3-2	SP	9.32	4.00	12.97	8.77	-0.03	28.19	0.00
qfLimo15	Limonene	F	Fvb3-2	2571815	PL	12.97	4.00	12.97	9.45	0.23	31.52	0.00
qfLimo16	Limonene	F	Fvb3-2	AX-184361324_Fa3-2	IT	8.41	2.00	12.97	4.14	-0.15	14.46	0.00
qfLino16	Linalool	F	Fvb3-2	AX-89887538_Fa3-2	FR	39.89	28.00	43.63	3.87	2.71	13.79	0.01
qfLino17	Linalool	F	Fvb3-2	8491	SP	43.63	24.00	43.63	3.18	2.01	11.31	0.02
qfLino18	Linalool	F	Fvb3-2	25718	FR	12.97	4.00	12.97	7.60	0.02	25.29	0.00
qfLino19	Linalool	F	Fvb3-2	AX-166512951_Fa3-2	SP	9.32	4.00	12.97	9.72	-0.04	30.72	0.00
qfLino110	Linalool	F	Fvb3-2	2571820	PL	12.97	4.00	12.97	5.05	-0.01	18.31	0.00
qfLino111	Linalool	F	Fvb3-2	10212	IT	4.72	2.00	12.97	4.15	-0.15	14.51	0.00
qfMesif8	Mesifurane	F	Fvb1-3	AX-184545710_Fa1-31	IT	0.91	0.00	25.72	3.43	3.36	12.15	0.00

QTL name	Compound	Map	Chr	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qfMesif9	Mesifurane	F	Fvb1-3	AX-184117023_Fa1-31	FR	5.50	0.00	22.00	6.77	5.65	22.88	0.00
qfMesif10	Mesifurane	F	Fvb1-3	AX-184512234_Fa1-3	SP	5.50	0.00	23.89	5.25	2.70	17.96	0.00
qfMesif11	Mesifurane	F	Fvb1-3	AX-184512234_Fa1-33	PL	5.50	0.00	23.89	6.54	3.16	23.04	0.00
qfMesif12	Mesifurane	F	Fvb3-2	257181	IT	12.97	4.00	12.97	2.58	0.78	9.28	0.05
qfMesif13	Mesifurane	F	Fvb3-2	AX-166504873_Fa3-2	FR	8.41	2.00	12.97	4.01	0.12	14.26	0.00
qfMesif14	Mesifurane	F	Fvb3-2	AX-166512951_Fa3-2	GE	9.32	2.00	12.97	3.48	0.04	12.41	0.01
qfMesif15	Mesifurane	F	Fvb3-2	AX-166512951_Fa3-2	PL	9.32	2.00	12.97	2.97	0.13	11.20	0.02
qfMesif16	Mesifurane	F	Fvb6-3	AX-123364217_Fa6-3	FR	7.34	0.00	25.00	3.08	-3.21	11.16	0.03
qfMetBenz4	Methyl benzoate	F	Fvb3-2	AX-166512951_Fa3-22	FR	9.32	2.00	12.97	2.48	0.06	9.09	0.02
qfMetBenz5	Methyl benzoate	F	Fvb3-2	AX-184647040_Fa3-21	GE	7.50	2.00	12.97	3.99	0.05	14.10	0.00
qfMetBenz6	Methyl benzoate	F	Fvb3-2	AX-166512951_Fa3-22	SP	9.32	2.00	12.97	2.78	-0.34	9.96	0.05
qfMetBenz7	Methyl benzoate	F	Fvb3-3	8160	SP	15.69	0.00	30.00	3.02	5.12	10.77	0.03
qfMetButy1	Methyl butyrate	F	Fvb3-2	cF3-2-b,loc12	FR	12.00	2.00	12.97	2.95	0.34	10.78	0.03
qfMetButy2	Methyl butyrate	F	Fvb6-2	AX-166518165_Fa6-2	IT	6.52	0.00	6.52	2.79	-4.08	9.99	0.03
qfMetDec3	Methyl decanoate	F	Fvb3-2	2571827	PL	12.97	2.00	12.97	2.90	-0.56	11.32	0.00
qfMetDec4	Methyl decanoate	F	Fvb3-4	AX-123366976_Fa3-4	GE	14.30	0.00	94.00	2.01	16.12	7.60	0.04
qfMetHex1	Methyl hexanoate	F	Fvb3-2	257181	FR	12.97	4.00	12.97	3.30	1.11	11.90	0.01
qfMetHex2	Methyl hexanoate	F	Fvb3-2	cF3-2-b,loc10	GE	10.00	2.00	12.97	3.05	0.49	10.97	0.03
qfMetHex3	Methyl hexanoate	F	Fvb3-2	102121	GE	4.72	2.00	12.97	2.83	0.02	10.22	0.04
qfMOHOc4	Methyl 3-hydroxyoctanoate	F	Fvb1-4	AX-184945096_Fa1-41	SP	71.02	53.33	96.00	3.19	2.35	11.33	0.01
qfMOHOc5	Methyl 3-hydroxyoctanoate	F	Fvb3-2	257182	SP	12.97	4.00	12.97	5.08	0.15	17.46	0.00
qfMOHOc6	Methyl 3-hydroxyoctanoate	F	Fvb3-2	cF3-2-b,loc10	PL	10.00	2.00	12.97	3.59	-27.14	13.39	0.00
qfNerdol10	Nerolidol	F	Fvb3-4	AX-184857015_Fa3-4	IT	157.26	0.00	223.60	2.21	0.37	8.02	0.02
qfNerdol7	Nerolidol	F	Fvb2-2	cF2-2,loc20	SP	20.00	0.00	43.35	2.82	0.70	10.09	0.03
qfNerdol8	Nerolidol	F	Fvb3-2	AX-166512951_Fa3-28	PL	9.32	2.00	12.97	4.14	-0.22	15.28	0.00

qfNerol6	Nerol	F	Fvb3-2	AX-184361324_Fa3-24	IT	8.41	2.00	12.97	4.24	-0.03	14.79	0.00
qfNerol7	Nerol	F	Fvb3-3	AX-184571583_Fa3-3	GE	14.78	0.00	37.78	3.64	2.78	12.93	0.01
qfNerol8	Nerol	F	Fvb3-2	263122	FR	39.89	26.00	43.63	3.60	2.11	12.91	0.00
qfNerol9	Nerol	F	Fvb3-2	84911	SP	43.63	0.00	43.63	1.71	5.10	6.27	0.02
qfNerol10	Nerol	F	Fvb3-2	2571824	PL	12.97	4.00	12.97	7.12	0.09	24.82	0.00
qfNerol11	Nerol	F	Fvb3-2	257189	FR	12.97	2.00	12.97	3.68	0.03	13.15	0.00
qfOcAc3	Octyl acetate	F	Fvb2-3	AX-184104182_Fa2-31	PL	22.75	0.00	50.66	1.95	2.06	7.50	0.01
qfOcAc4	Octyl acetate	F	Fvb3-2	cF3-2-b,loc121	GE	12.00	2.00	12.97	3.05	-0.01	10.97	0.02
qfOcBut3	Octyl butanoate	F	Fvb3-2	AX-166512951_Fa3-2	FR	9.32	2.00	12.97	3.96	-0.20	14.22	0.00
qfOcBut4	Octyl butanoate	F	Fvb3-2	2571829	PL	12.97	2.00	12.97	2.38	0.15	9.10	0.01
qfPent3	Pentanal	F	Fvb3-2	cF3-2-b,loc8	SP	8.00	2.00	12.97	2.55	0.26	10.86	0.02
qfPrenAc4	Prenyl acetate	F	Fvb1-4	AX-184945096_Fa1-4	IT	71.02	0.00	74.69	1.92	16.57	7.11	0.03
qfPrenAc5	Prenyl acetate	F	Fvb3-2	8491	GE	43.63	25.12	43.63	2.95	2.62	10.64	0.03
qfPrenAc6	Prenyl acetate	F	Fvb3-2	cF3-2-b,loc8	GE	8.00	2.00	12.97	4.80	-0.49	16.69	0.00
qfSthioate2	<i>S</i> -methyl butanethioate	F	Fvb3-2	25718	FR	12.97	2.00	12.97	2.69	1.69	9.81	0.01
qfTerpol10	Terpineol	F	Fvb3-2	AX-166512951_Fa3-24	SP	9.32	2.00	12.97	5.92	0.04	20.01	0.00
qfTerpol11	Terpineol	F	Fvb3-2	2571823	PL	12.97	4.00	12.97	6.17	0.09	21.88	0.00
qfTerpol12	Terpineol	F	Fvb3-2	AX-184361324_Fa3-22	IT	8.41	2.00	12.97	3.79	0.00	13.35	0.01
qfTerpol13	Terpineol	F	Fvb5-1	AX-184965167_Fa5-1	SP	61.84	0.00	69.16	4.12	1.57	14.41	0.00
qfTerpol7	Terpineol	F	Fvb3-2	AX-89887538_Fa3-21	FR	39.89	24.00	43.63	2.73	1.50	9.94	0.05
qfTerpol8	Terpineol	F	Fvb3-2	257182	FR	12.97	4.00	12.97	6.19	-0.14	21.16	0.00
qfTerpol9	Terpineol	F	Fvb3-2	257183	GE	12.97	2.00	12.97	4.84	-0.30	16.82	0.00
qfTMyrta10	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	257184	SP	12.97	2.00	12.97	2.60	-0.24	9.33	0.02
qfTMyrta11	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	AX-166512951_Fa3-26	PL	9.32	2.00	12.97	3.82	0.28	14.18	0.00
qfTMyrta8	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	AX-166512951_Fa3-22	IT	9.32	2.00	12.97	2.63	-0.08	9.46	0.05

QTL name	Compound	Map	Chr	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qfTMyrta9	<i>trans</i> -myrtanyl acetate	F	Fvb3-2	cF3-2-b,loc121	FR	12.00	2.00	12.97	4.76	0.10	16.84	0.00
qfZHexAc4	(<i>Z</i>)-3-hexen-1-ol acetate	F	Fvb3-2	257181	FR	12.97	4.00	12.97	4.14	0.12	14.80	0.00
qfZHexAc5	(<i>Z</i>)-3-hexen-1-ol acetate	F	Fvb3-2	257181	GE	12.97	2.00	12.97	4.49	0.24	15.72	0.00
qfZHexAc6	(<i>Z</i>)-3-hexen-1-ol acetate	F	Fvb3-2	cF3-2-b,loc12	GE	12.00	2.00	12.97	2.68	-0.09	9.70	0.03
qfZHexAc7	(<i>Z</i>)-3-hexen-1-ol acetate	F	Fvb3-2	257184	SP	12.97	2.00	12.97	4.07	0.24	14.25	0.00
qm1Hex2	1-hexanol	M	Fvb7-1	22999	GE	60.24	46.00	85.99	2.96	3.94	10.66	0.04
qm1Pentol3	1-penten-3-ol	M	Fvb6-1	18710	IT	180.99	0.00	323.82	1.00	16.44	3.69	0.02
qm1Pentol4	1-penten-3-ol	M	Fvb6-3	19558	PL	147.01	119.39	249.59	2.33	42.97	8.91	0.02
qm2Hept3	2-heptanone	M	Fvb3-4	AX-184250195_Fa3-4	GE	130.33	73.07	162.77	2.98	8.25	10.72	0.01
qm2Hept4	2-heptanone	M	Fvb5-2	20582	PL	47.71	0.00	78.00	2.91	6.82	11.00	0.02
qm2Nonol3	2-nonanol	M	Fvb3-2	dAX-184641647_Fa3-2	FR	171.40	0.00	279.14	1.19	7.98	4.48	0.04
qm2Nonol4	2-nonanol	M	Fvb4-2	13133	GE	0.00	0.00	24.96	1.01	17.15	3.88	0.04
qm2None5	2-nonanone	M	Fvb1-1	29172	IT	0.00	0.00	92.00	4.87	16.70	16.93	0.00
qm2None6	2-nonanone	M	Fvb1-1	2756	FR	57.34	0.00	80.32	7.33	29.33	24.53	0.00
qm2None7	2-nonanone	M	Fvb1-1	2600	GE	59.18	0.00	90.00	3.13	26.71	11.23	0.00
qm2None8	2-nonanone	M	Fvb1-1	2600	SP	59.18	0.00	92.00	3.42	22.38	12.21	0.00
qm2Penton4	2-pentanone	M	Fvb3-2	AX-184416025_Fa3-2	FR	18.38	0.00	83.54	2.88	4.38	10.45	0.04
qm2Penton5	2-pentanone	M	Fvb3-3	AX-184093094_Fa3-3	IT	334.54	312.00	391.09	3.22	2.12	11.44	0.02
qm2Penton6	2-pentanone	M	Fvb3-4	AX-184874789_Fa3-41	GE	130.33	108.27	162.77	4.40	2.34	15.41	0.00
qm2Penton7	2-pentanone	M	Fvb5-2	15838	PL	18.21	0.00	52.00	3.02	3.13	11.39	0.03
qm2Penton8	2-pentanone	M	Fvb5-3	14536	IT	177.56	138.67	236.75	3.14	1.76	11.17	0.02
qm34Benz1	3,4-dimethyl benzaldehyde	M	Fvb1-1	cM1-1,loc22	IT	22.00	0.00	218.00	1.77	12.32	6.56	0.05
qmBenAc5	Benzyl acetate	M	Fvb6-1	cM6-1.loc102	FR	102.00	14.68	172.00	3.43	2.65	12.35	0.00

qmBenAc6	Benzyl acetate	M	Fvb6-1	cM6-1,loc122	GE	122.00	50.00	176.00	3.03	1.52	10.90	0.01
qmBenAc7	Benzyl acetate	M	Fvb6-1	cM6-1,loc100	SP	100.00	22.00	158.20	3.35	5.51	11.87	0.00
qmBenz5	Benzaldehyde	M	Fvb6-1	33174	GE	141.29	98.00	166.00	6.44	1.14	21.73	0.00
qmBenz6	Benzaldehyde	M	Fvb6-1	cM6-1,loc74	SP	74.00	56.00	160.00	5.82	4.66	19.71	0.00
qmBenz7	Benzaldehyde	M	Fvb6-1	AX-184730940_Fa1-3	PL	158.20	98.00	183.76	3.46	1.59	12.94	0.01
qmBione4	β -ionone	M	Fvb1-1	AX-184412870_Fa1-1	GE	197.48	82.14	266.84	3.22	0.33	11.54	0.01
qmBione5	β -ionone	M	Fvb1-2	AX-184132143_Fa1-4	PL	86.78	0.00	119.78	2.86	0.76	10.84	0.04
qmBione6	β -ionone	M	Fvb4-2	12878	FR	3.74	0.00	24.96	3.00	1.22	11.31	0.02
qmDodeLac3	γ -dodecalactone	M	Fvb1-1	AX-184502996_Fa1-1	PL	249.91	236.00	262.00	3.20	0.25	12.03	0.04
qmDodeLac4	γ -dodecalactone	M	Fvb1-4	811	PL	38.94	8.00	66.00	3.16	0.25	11.89	0.04
qmDodeLac5	γ -dodecalactone	M	Fvb3-4	AX-184351052_Fa3-4	GE	130.33	26.00	152.00	3.51	2.77	12.51	0.01
qmEug8	Eugenol	M	Fvb1-4	35994	FR	69.43	28.00	85.95	2.92	30.11	10.59	0.00
qmEug9	Eugenol	M	Fvb1-4	1526	GE	79.52	37.10	82.00	3.62	6.04	12.88	0.00
qmEug10	Eugenol	M	Fvb1-4	35994	SP	69.43	38.94	82.00	4.73	17.80	16.36	0.00
qmEug11	Eugenol	M	Fvb1-4	35994	PL	69.43	37.10	82.00	4.64	4.88	16.95	0.00
qmEug12	Eugenol	M	Fvb1-4	AX-184235467_Fa1-4	IT	67.62	26.91	82.00	2.87	11.60	10.28	0.00
qmEug13	Eugenol	M	Fvb2-2	3266	FR	174.60	72.00	249.85	2.34	30.57	8.59	0.00
qmEug14	Eugenol	M	Fvb2-2	3308	GE	176.44	88.50	241.46	2.35	5.42	8.56	0.00
qmEug15	Eugenol	M	Fvb2-2	30222	SP	178.27	90.00	249.85	2.92	15.46	10.43	0.00
qmEug16	Eugenol	M	Fvb2-2	AX-184887573_Fa2-2	PL	197.24	78.00	248.00	2.91	4.26	10.99	0.01
qmEug17	Eugenol	M	Fvb2-2	AX-184887573_Fa2-2	IT	197.24	0.00	248.00	1.81	10.31	6.61	0.03
qmEZHep1	(<i>E,Z</i>)-2,4-heptadienal	M	Fvb4-4	cM4-4-a,loc128	IT	128.00	80.00	168.97	3.30	2.01	11.72	0.01
qmIsoAc3	Isoamyl acetate	M	Fvb3-3	9122	IT	286.93	122.00	391.09	2.65	38.77	9.59	0.02
qmIsoAc4	Isoamyl acetate	M	Fvb3-3	AX-184168298_Fa3-3	GE	179.40	2.00	258.20	2.79	17.72	10.41	0.00
qmMetAc2	Methyl acetate	M	Fvb2-1	29164	IT	234.78	122.00	246.00	3.03	9.09	10.80	0.02
qmMetBenz2	Methyl benzoate	M	Fvb7-1	23208	IT	70.41	47.37	85.99	2.51	14.83	9.04	0.03
qmMetdec2	Methyl decanoate	M	Fvb3-3	29366	PL	72.99	0.00	270.00	2.62	6.33	10.31	0.01

QTL name	Compound	Map	Chr	Marker	Country	Pos (cM)	ci-low (cM)	ci-up (cM)	LOD	a	r ²	p-value
qmMetOc1	Methyl octanoate	M	Fvb1-1	26002	GE	59.18	0.00	124.64	2.75	4.63	10.53	0.04
qmMetOc2	Methyl octanoate	M	Fvb5-2	13532	SP	7.39	0.00	143.62	0.72	79.93	2.69	0.04
qmMOct3	Methyl 3-hydroxyoctanoate	M	Fvb3-3	AX-184448181_Fa3-3	GE	183.97	72.00	391.09	2.13	62.60	8.04	0.02
qmMOct4	Methyl 3-hydroxyoctanoate	M	Fvb4-3	AX-123367232_Fa4-3	GE	68.82	44.00	77.66	2.24	-30.06	8.43	0.01
qmNerdol1	Nerolidol	M	Fvb3-3	AX-184748009_Fa3-3	GE	11.46	0.00	202.00	2.77	0.48	10.01	0.04
qmNerdol2	Nerolidol	M	Fvb3-4	AX-184842185_Fa3-4	IT	142.40	0.00	162.77	1.96	0.33	7.12	0.05
qmNonAc3	Nonyl acetate	M	Fvb5-2	ddAX-184952715_Fa5-2	GE	142.71	100.29	143.62	3.03	2.08	11.72	0.03
qmNonAc4	Nonyl acetate	M	Fvb7-4	24736	IT	72.16	33.43	81.36	3.01	2.75	10.98	0.00
qmNonAc5	Nonyl acetate	M	Fvb7-4	24728	IT	54.40	12.82	58.00	2.54	2.04	9.36	0.03
qmOcAc3	Octyl acetate	M	Fvb5-3	cM5-3,loc4	GE	4.00	0.00	236.75	2.17	1.28	8.04	0.00
qmOcAc4	Octyl acetate	M	Fvb6-1	19904	PL	99.43	0.00	292.00	2.01	2.12	7.75	0.01
qmPent3	Pentanal	M	Fvb2-2	6140	SP	33.36	0.00	50.00	2.58	33.92	11.01	0.03
qmPent4	Pentanal	M	Fvb5-1	16140	GE	11.94	0.00	49.84	3.12	3.69	11.28	0.02
qmPrenAc2	Prenyl acetate	M	Fvb2-2	6140	IT	33.36	0.00	154.00	1.90	15.48	7.03	0.05
qmTerpol2	Terpineol	M	Fvb6-1	20192	PL	79.06	38.20	87.44	3.19	1.03	11.98	0.03
qmTMyrta2	<i>trans</i> -myrtanyl acetate	M	Fvb4-4	122401	IT	126.25	90.94	153.99	2.99	1.10	10.66	0.04

Annex 7. Functional annotation of genes restricted to qfDodeLac3. All the genes delimited within the QTL for γ -dodecalactone in the female map (qfDodeLac3) selected for further studies. Functional annotation was carried out using Mercator4, with MapMan BINcodes and by using BLASTP (amino acid sequences) with TAIR and TREMBL databases.

Gene	MapMan BINcode	Description	BLASTP_TAIR	BLASTP_TREMBL
<i>FxaC_10g48280</i>	35.1	Not assigned	Cellulose synthase like E1	Putative cellulose synthase (UDP-forming)
<i>FxaC_10g48560</i>	2.4.4.2.2	Cellular respiration	Transmembrane protein G1P-related 1	Putative heme-transporting ATPase
<i>FxaC_10g48600</i>	35.1	Not assigned	A20/AN1-like zinc finger family protein	Putative transcription regulator A20-like family
<i>FxaC_10g48720</i>	35.1	Not assigned	Leucine-rich repeat (LRR) family protein	Putative histone deacetylase
<i>FxaC_10g48610</i>	5.2.3.3.1.1.1	Lipid metabolism	MSF1-like family protein	Putative PRELI/MSF1 domain-containing protein
<i>FxaC_10g48670</i>	18.4.6.5	Protein modification	AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kinase family protein	Uncharacterized protein
<i>FxaC_10g48680</i>	35.2	Not assigned	MIF4G domain-containing protein / MA3 domain-containing protein	Putative initiation factor eIF-4 gamma, MA3
<i>FxaC_10g48790</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g48840</i>	50.2.7	Enzyme classification	Receptor-like protein kinase 1	Uncharacterized protein
<i>FxaC_10g48860</i>	35.1	Not assigned	Tetratricopeptide repeat (TPR)-like superfamily protein	Putative tetratricopeptide-like helical domain-containing protein
<i>FxaC_10g48980</i>	11.7.1.2	Phytohormone action	Lipoxygenase 2	Lipoxygenase
<i>FxaC_10g49090</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative toll-like receptor, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49000</i>	35.1	Not assigned	Topoisomerase 6 subunit B	DNA topoisomerase 6 subunit B
<i>FxaC_10g49410</i>	26.8.2.1	External stimuli response	Disease resistance protein (TIR-NBS-LRR class)	Putative TIR domain, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49590</i>	5.1.4.5	Lipid metabolism	NAD(P)-binding Rossmann-fold superfamily protein	Putative enoyl-[acyl-carrier-protein] reductase (NADH)
<i>FxaC_10g49750</i>	35.1	Not assigned	Calcium-dependent lipid-binding (CaLB domain) plant phosphoribosyltransferase family protein	Putative C2 domain, phosphoribosyltransferase
<i>FxaC_10g49810</i>	19.2.2.1.5.3.5	Protein homeostasis	NA	E3 ubiquitin-protein ligase
<i>FxaC_10g49850</i>	24.2.9.2.2	Solute transport	Cation efflux family protein	Putative cation efflux protein

Gene	MapMan BINcode	Description	BLASTP_TAIR	BLASTP_TREMBL
<i>FxaC_10g49900</i>	35.1	Not assigned	Tetratricopeptide repeat (TPR)-like superfamily protein	Putative pentatricopeptide
<i>FxaC_10g49970</i>	35.2	Not assigned	Family of unknown function (DUF716)	Uncharacterized protein
<i>FxaC_10g50020</i>	18.3.4.1.4	Protein modification	Hexosyl transferase	Putative K domain, splicing factor 1, helix-hairpin domain-containing protein
<i>FxaC_10g50170</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g50350</i>	24.2.3.5.1	Solute transport	Amino acid permease family protein	Putative amino acid/polyamine transporter I
<i>FxaC_10g50380</i>	35.1	Not assigned	Receptor like protein 2	Putative non-specific serine/threonine protein kinase
<i>FxaC_10g50450</i>	9.2.1.1.2	Secondary metabolism	Galactose oxidase/kelch repeat superfamily protein	Putative F-box domain, galactose oxidase, beta-propeller
<i>FxaC_10g50510</i>	16.4.6.2	RNA processing	O-Glycosyl hydrolases family 17 protein	Putative glucan endo-1,3-beta-D-glucosidase
<i>FxaC_10g50630</i>	35.1	Not assigned	LRR family protein	Putative non-specific serine/threonine protein kinase
<i>FxaC_10g50840</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g50800</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative toll-like receptor, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g51050</i>	35.1	Not assigned	RING/U-box superfamily protein	Putative aminoacyltransferase, E1 ubiquitin-activating enzyme
<i>FxaC_10g51160</i>	35.1	Not assigned	Pentatricopeptide repeat (PPR) superfamily protein	Pentatricopeptide repeat-containing protein
<i>FxaC_10g51170</i>	35.1	Not assigned	purin-rich alpha 1	Putative purine-rich element binding protein family
<i>FxaC_10g48130</i>	35.1	Not assigned	Uncharacterised protein family (UPF0172)	Putative ER membrane protein complex subunit 8/9
<i>FxaC_10g48200</i>	18.4.6.5	Protein modification	D6 protein kinase like 2	Uncharacterized protein
<i>FxaC_10g48220</i>	35.1	Not assigned	Sec14p-like phosphatidylinositol transfer family protein	Putative CRAL-TRIO lipid binding domain, CRAL/TRIO domain-containing protein
<i>FxaC_10g48270</i>	35.1	Not assigned	cellulose synthase like E1	Uncharacterized protein
<i>FxaC_10g48340</i>	22.3.5.1	Vesicle trafficking	Emp24/gp25L/p24 family/GOLD family protein	Uncharacterized protein
<i>FxaC_10g48240</i>	35.1	Not assigned	Plant protein of unknown function (DUF827)	Putative WEB family protein
<i>FxaC_10g48390</i>		Not assigned	Pentatricopeptide repeat (PPR) superfamily protein	Uncharacterized protein

<i>FxaC_10g48420</i>	35.2	Not assigned	NOL1/NOP2/sun family protein / antitermination NusB domain-containing protein	Putative 16S rRNA (Cytosine(967)-C(5))-methyltransferase
<i>FxaC_10g48430</i>	16.8.2.1.1	RNA processing	Polynucleotidyl transferase, ribonuclease H-like superfamily protein	Putative poly(A)-specific ribonuclease
<i>FxaC_10g48530</i>	24.2.11.4	Solute transport	ZIP metal ion transporter family	Putative zinc/iron permease
<i>FxaC_10g48690</i>	35.2	Not assigned	MIF4G domain-containing protein / MA3 domain-containing protein	Putative initiation factor eIF-4 gamma, MA3
<i>FxaC_10g48700</i>	35.2	Not assigned	Putative glycosyl hydrolase of unknown function (DUF1680)	Putative alpha-L-arabinofuranosidase B, arabinose-binding domain, six-hairpin glycosidase
<i>FxaC_10g48710</i>	17.1.2.1.32	Protein biosynthesis	Ribosomal protein L31e family protein	Uncharacterized protein
<i>FxaC_10g48640</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g48750</i>	15.5.32	RNA biosynthesis	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	Putative transcription factor bHLH family
<i>FxaC_10g48770</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g48850</i>	13.1.1.1.2	Cell cycle organisation	Cyclin family protein	Putative cyclin
<i>FxaC_10g48880</i>	50.3.2	Enzyme classification	beta glucosidase 13	Putative glycosidase
<i>FxaC_10g48780</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g48970</i>	50.3.2	Enzyme classification	beta glucosidase 17	Beta-glucosidase 3
<i>FxaC_10g49070</i>	11.10.1.10.1	Phytohormone action	ETHYLENE-INSENSITIVE3-like 3	Putative transcription factor EIL family
<i>FxaC_10g49080</i>	12.3.3.9.8	Chromatin organisation	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	Putative [histone H3]-lysine-36 demethylase
<i>FxaC_10g49100</i>	26.8.2.1	External stimuli response	Toll-Interleukin-Resistance (TIR) domain family protein	Putative TIR domain-containing protein
<i>FxaC_10g48990</i>	11.7.1.2	Phytohormone action	Lipoxygenase 2	Lipoxygenase
<i>FxaC_10g49130</i>	6.2.1.1.2	Nucleotide metabolism	Carbamoyl phosphate synthetase A	Putative carbamoyl-phosphate synthase (Glutamine-hydrolyzing)
<i>FxaC_10g49190</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g49320</i>	35.2	Not assigned	NA	Uncharacterized protein

Gene	MapMan BINcode	Description	BLASTP_TAIR	BLASTP_TREMBL
<i>FxaC_10g49420</i>	17.5.3.1	Protein biosynthesis	Eukaryotic elongation factor 5A-1	Eukaryotic translation initiation factor 5A
<i>FxaC_10g49640</i>	15.5.16	RNA biosynthesis	KH domain-containing protein / zinc finger (CCCH type) family protein	Putative transcription factor C3H family
<i>FxaC_10g49500</i>	35.2	Not assigned	alpha/beta-Hydrolases superfamily protein	Putative alpha/Beta hydrolase
<i>FxaC_10g49550</i>	15.5.2.2	RNA biosynthesis	DNA-binding bromodomain-containing protein	Putative transcription factor MYB-HB-like family
<i>FxaC_10g49710</i>	17.1.2.2.2.1.2	Protein biosynthesis	Noc2p family	Putative nucleolar complex protein
<i>FxaC_10g49740</i>	17.4.1.2.3.4	Protein biosynthesis	NagB/RpiA/CoA transferase-like superfamily protein	Uncharacterized protein
<i>FxaC_10g49830</i>	35.1	Not assigned	P-loop containing nucleoside triphosphate hydrolases superfamily protein	Putative nitric-oxide synthase (NADPH)
<i>FxaC_10g49870</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative toll-like receptor, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49880</i>	12.3.6.2	Chromatin organisation	Protein kinase family protein	Uncharacterized protein
<i>FxaC_10g49980</i>	19.2.2.1.5.3.2	Protein homeostasis	Histone mono-ubiquitination 2	E3 ubiquitin protein ligase
<i>FxaC_10g50010</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g50040</i>	21.4.1.1.4	Cell wall organisation	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	Putative plant lipid transfer protein/Par allergen
<i>FxaC_10g50050</i>	35.2	Not assigned	Protein of unknown function, DUF538	Uncharacterized protein
<i>FxaC_10g50320</i>	18.4.24.1.2.5	Protein modification	protein phosphatase 2A-3	Serine/threonine-protein phosphatase
<i>FxaC_10g50330</i>	9.1.6.1.1	Secondary metabolism	Phytoene synthase	Putative 15-cis-phytoene synthase
<i>FxaC_10g50440</i>	9.2.1.1.2	Secondary metabolism	acyl-activating enzyme 18	Putative acetate--CoA ligase
<i>FxaC_10g50570</i>	35.1	Not assigned	alpha/beta-Hydrolases superfamily protein	Putative chlorophyllase
<i>FxaC_10g50580</i>	18.3.4.1.1.1	Protein modification	UDP-Glycosyltransferase superfamily protein	Putative phosphatidylinositol N-acetylglucosaminyltransferase
<i>FxaC_10g50600</i>	15.5.2.1	RNA biosynthesis	myb domain protein 5	Transcription factor MYB5

<i>FxaC_10g50670</i>	20.1.1.1	Cytoskeleton organisation	tubulin alpha-5	Tubulin alpha chain
<i>FxaC_10g50760</i>	15.3.3.4.2	RNA biosynthesis	RNA polymerase II transcription mediators	Mediator of RNA polymerase II transcription subunit 13
<i>FxaC_10g50640</i>	24.2.13	Solute transport	Mitochondrial substrate carrier family protein	Putative mitochondrial carrier domain-containing protein
<i>FxaC_10g50660</i>	24.2.13	Solute transport	Mitochondrial substrate carrier family protein	Putative mitochondrial carrier domain-containing protein
<i>FxaC_10g50750</i>	17.7.4.4	Protein biosynthesis	elongation factor P (EF-P) family protein	Putative translation elongation factor P
<i>FxaC_10g50810</i>	19.4.1.2	Protein homeostasis	NA	NA
<i>FxaC_10g50820</i>	35.2	Not assigned	Protein of Unknown Function (DUF239)	Putative neprosin
<i>FxaC_10g50830</i>	17.2.13	Protein biosynthesis	Lysyl-tRNA synthetase, class II	Lysine--tRNA ligase
<i>FxaC_10g50920</i>	16.5.2.1	RNA processing	evolutionarily conserved C-terminal region 2	Putative YTH domain-containing protein
<i>FxaC_10g50940</i>	4.2.7.2.1.2	Amino acid metabolism	branched-chain alpha-keto acid decarboxylase E1 beta subunit	Putative 3-methyl-2-oxobutanoate dehydrogenase (2-methylpropanoyl-transferring)
<i>FxaC_10g50870</i>	1.1.8.1.6.2.2	Photosynthesis	TCP-1/cpn60 chaperonin family protein	Putative chaperonin Cpn60/TCP-1 family, groEL-like apical domain, groEL-like equatorial
<i>FxaC_10g50890</i>	18.4.1.37.1	Protein modification	Protein kinase superfamily protein	Uncharacterized protein
<i>FxaC_10g50950</i>	15.4.1	RNA biosynthesis	TATA binding protein 2	Putative TATA-box binding protein
<i>FxaC_10g51030</i>	18.1.1.6.2.2	Protein modification	Calmodulin-binding protein	Putative mannose-6-phosphate receptor binding domain, glucosidase 2 subunit beta
<i>FxaC_10g51020</i>	35.2	Not assigned	Antitermination NusB domain-containing protein	Putative NusB antitermination factor
<i>FxaC_10g51090</i>	35.1	Not assigned	Uncharacterised conserved protein UCP022280	Uncharacterized protein
<i>FxaC_10g51130</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g51080</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g51220</i>	35.2	Not assigned	Oxidoreductases	Putative alpha-ketoglutarate-dependent dioxygenase AlkB
<i>FxaC_10g48110</i>	35.1	Not assigned	RNA-binding (RRM/RBD/RNP motifs) family protein	Putative nucleotide-binding alpha-beta plait domain-containing protein
<i>FxaC_10g48400</i>	35.1	Not assigned	ATP binding;nucleic acid binding;helicases	Putative tetratricopeptide-like helical domain-containing protein

Gene	MapMan BINcode	Description	BLASTP_TAIR	BLASTP_TREMBL
<i>FxaC_10g48350</i>	35.1	Not assigned	RPM1-interacting protein 4 (RIN4) family protein	Putative RIN4, pathogenic type III effector avirulence factor Avr cleavage
<i>FxaC_10g48450</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g48440</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g48470</i>	18.4.1.19	Protein modification	NA	Nicotinamide/nicotinic acid mononucleotide adenylyltransferase (Fragment)
<i>FxaC_10g48500</i>	15.5.15	RNA biosynthesis	Outer envelope protein of 80 kDa	Putative bacterial surface antigen (D15)
<i>FxaC_10g48410</i>	19.2.2.1.5.1.1	Protein homeostasis	Ubiquitin-protein ligase 1	Putative aminoacyltransferase, E1 ubiquitin-activating enzyme
<i>FxaC_10g48630</i>	35.1	Not assigned	Pentatricopeptide repeat (PPR) superfamily protein	Putative pentacotriptide-repeat region of PROPR
<i>FxaC_10g48810</i>	50.2.7	Enzyme classification	Receptor-like protein kinase 1	Uncharacterized protein
<i>FxaC_10g48870</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g48920</i>	50.3.2	Enzyme classification	beta glucosidase 12	Uncharacterized protein
<i>FxaC_10g49110</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative toll-like receptor, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49270</i>	50.2.7	Enzyme classification	F-box family protein	Putative F-box domain-containing protein
<i>FxaC_10g49330</i>	7.14.2	Coenzyme metabolism	Unknown protein	UPF0061 protein azo1574
<i>FxaC_10g49360</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative TIR domain, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49440</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative TIR domain, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49560</i>	14.3.6.2.1	DNA damage response	Nse4, component of Smc5/6 DNA repair complex	Non-structural maintenance of chromosomes element 4
<i>FxaC_10g49460</i>	26.8.2.1	External stimuli response	NB-ARC domain-containing disease resistance protein	Putative toll-like receptor, P-loop containing nucleoside triphosphate hydrolase
<i>FxaC_10g49760</i>	35.2	Not assigned	PRLI-interacting factor, putative	Uncharacterized protein
<i>FxaC_10g49770</i>	21.7.1	Cell wall organisation	Glucan synthase-like 12	Putative 1,3-beta-glucan synthase

<i>FxaC_10g49780</i>	35.2	Not assigned	Late embryogenesis abundant protein (LEA) family protein	Uncharacterized protein
<i>FxaC_10g49890</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g49910</i>	35.2	Not assigned	Unknown protein	Uncharacterized protein
<i>FxaC_10g50340</i>	35.2	Not assigned	Lipase/lipoxygenase, PLAT/LH2 family protein	Putative PLAT/LH2 domain-containing protein
<i>FxaC_10g50390</i>	24.2.3.5.1	Solute transport	Amino acid permease family protein	Putative amino acid/polyamine transporter I
<i>FxaC_10g50130</i>	35.2	Not assigned	DNAJ heat shock N-terminal domain-containing protein	Putative DnaJ domain-containing protein
<i>FxaC_10g50080</i>	35.1	Not assigned	Leucine-rich receptor-like protein kinase family protein	Putative leucine-rich repeat-containing, plant-type, leucine-rich repeat domain, L
<i>FxaC_10g50260</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g50270</i>	35.2	Not assigned	RING/U-box superfamily protein	Putative transcription factor C2H2 family
<i>FxaC_10g50300</i>	35.1	Not assigned	P-loop containing nucleoside triphosphate hydrolases superfamily protein	DEAD-box ATP-dependent RNA helicase 37-like
<i>FxaC_10g50490</i>	35.2	Not assigned	NA	Putative replication protein A, OB
<i>FxaC_10g50520</i>	15.3.5.3.7	RNA biosynthesis	TBP-associated factor 7	Putative TAFII55 protein region
<i>FxaC_10g50540</i>	35.1	Not assigned	Calmodulin-binding family protein	Uncharacterized protein
<i>FxaC_10g50560</i>	35.1	Not assigned	ACD1-like	Putative pheophorbide a oxygenase
<i>FxaC_10g50590</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g50620</i>	35.2	Not assigned	Zinc ion binding	Putative Zinc finger, SWIM-type
<i>FxaC_10g50500</i>	19.2.2.8.3.3.1	Protein homeostasis	Ubiquitin-conjugating enzyme family protein	Putative aminoacyltransferase, E1 ubiquitin-activating enzyme
<i>FxaC_10g50930</i>	16.5.2.1	RNA processing	evolutionarily conserved C-terminal region 2	Putative YTH domain-containing protein
<i>FxaC_10g51040</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g51140</i>	20.1.3.12	Cytoskeleton organisation	P-loop containing nucleoside triphosphate hydrolases superfamily protein	Putative minus-end-directed kinesin ATPase
<i>FxaC_10g51240</i>	15.3.5.7.2	RNA biosynthesis	DNA binding;ATP binding;nucleic acid binding;binding;helicases;ATP binding;DNA binding;helicases	Putative DNA helicase chromatin remodeling SNF2 family
<i>FxaC_10g48230</i>	35.2	Not assigned	Sucrase/ferredoxin-like family protein	Putative thioredoxin-like ferredoxin
<i>FxaC_10g49210</i>	35.2	Not assigned	Ribosomal protein L4/L1 family	Putative ribosomal protein L4/L1c

Gene	MapMan BINcode	Description	BLASTP_TAIR	BLASTP_TREMBL
<i>FxaC_10g49490</i>	35.1	Not assigned	Emsy N Terminus (ENT)/ plant Tudor-like domains-containing protein	Putative ENT domain, protein EMSY-LIKE, plant
<i>FxaC_10g49480</i>	18.4.2.2	Protein modification	MAPK/ERK kinase kinase 1	Putative mitogen-activated protein kinase kinase kinase STE-STE11 family
<i>FxaC_10g50410</i>	50.2.4	Enzyme classification	UDP-glycosyltransferase 73B4	Glycosyltransferase
<i>FxaC_10g50610</i>	35.2	Not assigned	NA	NA
<i>FxaC_10g51100</i>	16.11.3.2	RNA processing	Nuclear pore complex protein	Uncharacterized protein

Annex 8. Additional information about phylogenetic analysis of lipoxygenases. The species used in this study are *Arabidopsis thaliana* (At), *Cucumis melo* (Cm), *Cucumis sativum* (CR and CS), *Fragaria × ananassa* (Fa), *Fragaria vesca* (Fv), *Malus × domestica* (Md), *Olea europaea* (Oe), *Prunus dulcis* (Pd), *Prunus persica* (Pp), *Pisum sativa* (Ps), *Solanum lycopersicum* (Sl or TomLOX), *Solanum tuberosum* (St) and *Vitis vinifera* (Vv). ID refers to the gene identification from GDR, TAIR, Sol Genomics or NCBI databases. Aa is the number of amino acids of the complete sequence. The presence or absence of signal peptide was predicted with TargetP2.0 software.

Protein name	Type	Clade	ID	Aa	Signal peptide
AtLOX1	9-LOX	9-A	AT1G55020.1	859	No signal peptide
AtLOX2	13-LOX	13-F	AT3G45140.1	896	Chloroplast transit peptide
AtLOX3	13-LOX	13-E	AT1G17420.1	919	Chloroplast transit peptide
AtLOX4	13-LOX	13-E	AT1G72520.1	926	Chloroplast transit peptide
AtLOX5	9-LOX	9-A	AT3G22400.1	886	No signal peptide
AtLOX6	13-LOX	13-C	AT1G67560.1	917	Chloroplast transit peptide
CmLOX10	13-LOX	13-F	KT613843	902	Chloroplast transit peptide
CmLOX13	13-LOX	13-F	KT613844	906	Chloroplast transit peptide
CRLOX-1	13-LOX	13-A	U36339	877	No signal peptide
CSLBLOX	13-LOX	13-A	CAA63483.1	878	No signal peptide
FaLOX1	9-LOX	9-A	FxaC_13g32900.1	884	No signal peptide
FaLOX10	13-LOX	13-F	FxaC_12g02830.1	306	Nuclear export signal
FaLOX11	13-LOX	13-D	FxaC_12g40801.1	157	No signal peptide
FaLOX12	13-LOX	13-C	FxaC_13g21340.1	919	Chloroplast transit peptide
FaLOX13	13-LOX	13-C	FxaC_14g13650.1	920	Chloroplast transit peptide
FaLOX14	9-LOX	9-A	FxaC_14g24250.1	884	No signal peptide
FaLOX15	9-LOX	9-D	FxaC_14g32480.1	207	No signal peptide
FaLOX16	9-LOX	9-D	FxaC_14g32510.1	167	Nuclear localization signal
FaLOX17	13-LOX	13-C	FxaC_15g14640.1	920	Chloroplast transit peptide
FaLOX18	9-LOX	9-A	FxaC_15g22470.1	884	No signal peptide
FaLOX19	9-LOX	9-D	FxaC_15g31550.1	860	No signal peptide
FaLOX2	13-LOX	13-F	FxaC_10g48980.1	619	No signal peptide
FaLOX20	9-LOX	9-D	FxaC_16g09940.1	860	No signal peptide
FaLOX21	9-LOX	9-A	FxaC_16g15090.1	516	Nuclear export signal
FaLOX22	13-LOX	13-C	FxaC_16g18570.1	920	Chloroplast transit peptide
FaLOX23	13-LOX	13-H	FxaC_17g35570.1	892	Chloroplast transit peptide
FaLOX24	13-LOX	13-E	FxaC_21g24910.1	914	Chloroplast transit peptide
FaLOX25	13-LOX	13-E	FxaC_22g21600.1	914	Chloroplast transit peptide
FaLOX26	9-LOX	9-C	FxaC_22g62500.1	918	Chloroplast transit peptide

Protein name	Type	Clade	ID	Aa	Signal peptide
FaLOX27	9-LOX	9-D	FxaC_22g62501.1	867	No signal peptide
FaLOX28	13-LOX	13-E	FxaC_23g03550.1	914	Chloroplast transit peptide
FaLOX29	9-LOX	9-C	FxaC_23g37150.1	875	Chloroplast transit peptide
FaLOX3	13-LOX	13-F	FxaC_9g03100.1	950	Chloroplast transit peptide
FaLOX30	9-LOX	9-D	FxaC_23g37160.1	918	No signal peptide
FaLOX31	9-LOX	9-D	FxaC_24g10010.1	479	Nuclear localization signal
FaLOX32	9-LOX	9-C	FxaC_24g10011.1	423	Nuclear localization signal
FaLOX33	13-LOX	13-G	FxaC_25g00610.1	860	No signal peptide
FaLOX34	13-LOX	13-G	FxaC_25g00640.1	508	Nuclear localization signal
FaLOX35	13-LOX	13-H	FxaC_25g00650.1	921	Chloroplast transit peptide
FaLOX36	13-LOX	13-H	FxaC_26g42540.1	923	Chloroplast transit peptide
FaLOX37	13-LOX	13-G	FxaC_26g42630.1	315	Nuclear localization signal
FaLOX38	13-LOX	13-G	FxaC_28g38200.1	86	Nuclear export signal
FaLOX39	13-LOX	13-G	FxaC_28g38220.1	127	Signal peptide
FaLOX4	13-LOX	13-F	FxaC_9g03120.1	959	Chloroplast transit peptide
FaLOX40	13-LOX	13-G	FxaC_28g38230.1	880	No signal peptide
FaLOX41	13-LOX	13-H	FxaC_28g38250.1	126	Nuclear export signal
FaLOX42	13-LOX	13-H	FxaC_28g38260.1	137	Chloroplast transit peptide
FaLOX43	13-LOX	13-H	FxaC_28g38880.1	195	No signal peptide
FaLOX44	13-LOX	13-H	FxaC_28g38881.1	250	No signal peptide
FaLOX45	13-LOX	13-G	FxaC_28g38900.1	880	No signal peptide
FaLOX46	13-LOX	13-G	FxaC_28g38920.1	87	Signal peptide
FaLOX47	13-LOX	13-G	FxaC_28g38930.1	138	No signal peptide
FaLOX48	13-LOX	13-B	FxaC_5g19730.1	826	Nuclear localization signal
FaLOX49	13-LOX	13-B	FxaC_7g16340.1	844	No signal peptide
FaLOX5	9-LOX	9-A	FxaC_14g24250.1	884	No signal peptide
FaLOX6	13-LOX	13-F	FxaC_10g48990.1	959	Chloroplast transit peptide
FaLOX7	13-LOX	13-D	FxaC_11g07030.1	257	Nuclear localization signal
FaLOX8	13-LOX	13-F	FxaC_11g45700.1	948	Chloroplast transit peptide
FaLOX9	13-LOX	13-F	FxaC_11g45710.1	959	Chloroplast transit peptide
FvLOX1	13-LOX	13-B	FvH4_2g17660.1	802	No signal peptide
FvLOX10	13-LOX	13-E	FvH4_6g34970.1	914	Chloroplast transit peptide
FvLOX11	13-LOX	13-G	FvH4_7g00290.1	514	No signal peptide
FvLOX12	13-LOX	13-G	FvH4_7g00380.1	880	No signal peptide
FvLOX13	13-LOX	13-G	FvH4_7g00420.1	639	Nuclear localization signal
FvLOX14	13-LOX	13-H	FvH4_7g00440.1	818	Nuclear localization signal

FvLOX2	13-LOX	13-D	FvH4_3g09140.1	769	Nuclear localization signal
FvLOX3	13-LOX	13-F	FvH4_3g43530.1	950	Chloroplast transit peptide
FvLOX4	9-LOX	9-D	FvH4_4g09530.1	857	No signal peptide
FvLOX5	13-LOX	13-C	FvH4_4g24180.1	919	Chloroplast transit peptide
FvLOX6	13-LOX	13-H	FvH4_5g05680.1	767	No signal peptide
FvLOX7	13-LOX	13-H	FvH4_5g26250.1	892	Chloroplast transit peptide
FvLOX8	9-LOX	9-D	FvH4_6g06090.1	863	No signal peptide
FvLOX9	9-LOX	9-C	FvH4_6g06100.10	919	Chloroplast transit peptide
MdLOX12	13-LOX	13-D	MD10G1294000	446	Nuclear localization signal
MdLOX1a	9-LOX	9-D	MD09G1069500	809	No signal peptide
MdLOX1b	9-LOX	9-D	MD04G1204600	963	Chloroplast transit peptide
MdLOX1c	9-LOX	9-D	MD04G1204500	862	No signal peptide
MdLOX1d	9-LOX	9-D	MD12G1218300	793	No signal peptide
MdLOX2a	13-LOX	13-E	MD17G1167100	906	Chloroplast transit peptide
MdLOX2b	13-LOX	13-E	MD09G1180900	905	Chloroplast transit peptide
MdLOX3a	9-LOX	9-A	MD13G1191200	879	No signal peptide
MdLOX3b	9-LOX	9-A	MD16G1191300	884	No signal peptide
MdLOX4	9-LOX	9-C	MD04G1204100	930	Chloroplast transit peptide
MdLOX5a	13-LOX	13-H	MD07G1003900	920	Chloroplast transit peptide
MdLOX5b	13-LOX	13-H	MD02G1317800	914	Chloroplast transit peptide
MdLOX5c	13-LOX	13-H	MD07G1004500	678	No signal peptide
MdLOX5d	13-LOX	13-H	MD07G1003600	446	Nuclear localization signal
MdLOX6a	13-LOX	13-F	MD03G1021200	940	Chloroplast transit peptide
MdLOX6b	13-LOX	13-F	MD11G1023100	984	Chloroplast transit peptide
MdLOX7a	9-LOX	9-C	MD04G1204000	868	No signal peptide
MdLOX7b	9-LOX	9-C	MD12G1218000	922	Chloroplast transit peptide
MdLOX7c	9-LOX	9-C	MD12G1218200	885	Chloroplast transit peptide
MdLOX7d	9-LOX	9-C	MD04G1204200	933	Chloroplast transit peptide
MdLOX8a	13-LOX	13-C	MD16G1113200	924	Chloroplast transit peptide
MdLOX8b	13-LOX	13-C	MD13G1114000	920	Chloroplast transit peptide
MdLOX9	13-LOX	13-D	MD05G1339300	909	Chloroplast transit peptide
MdLOX9a	13-LOX	13-B	MD05G1156900	600	No signal peptide
OeLOX	9-LOX	9-B	ACG56281	864	No signal peptide
OeLOX2	13-LOX	13-E	ACD43484	913	Chloroplast transit peptide
PdLOX1	9-LOX	9-D	CAB94852	862	No signal peptide
PdLOX2	9-LOX	9-D	CAD10779	862	No signal peptide
PpLOX04	13-LOX	13-H	Prupe.2G005500.1	920	Chloroplast transit peptide

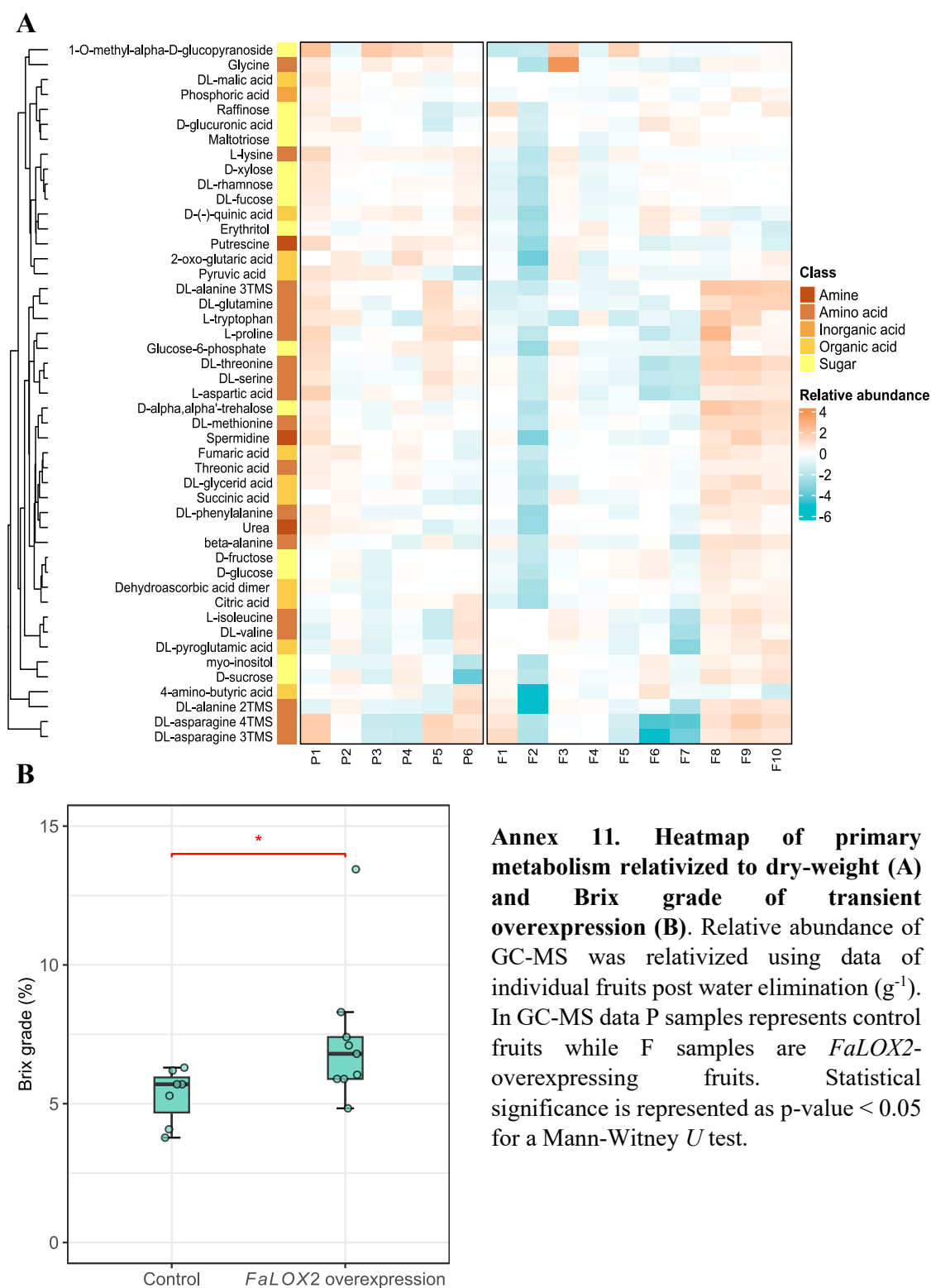
Protein name	Type	Clade	ID	Aa	Signal peptide
PpLOX1.1	9-LOX	9-D	Prupe.6G324200.1	946	Chloroplast transit peptide
PpLOX1.2	9-LOX	9-D	Prupe.6G324600.1	916	Chloroplast transit peptide
PpLOX12	9-LOX	9-D	Prupe.6G324400.1	862	No signal peptide
PpLOX2.1	13-LOX	13-E	Prupe.3G039200.1	912	Chloroplast transit peptide
PpLOX2.2	13-LOX	13-D	Prupe.4G047800.1	906	Chloroplast transit peptide
PpLOX3.1	9-LOX	9-A	Prupe.1G011400.1	881	No signal peptide
PpLOX5.1	13-LOX	13-H	Prupe.2G005300.1	917	Chloroplast transit peptide
PpLOX5.2	13-LOX	13-H	Prupe.2G005800.1	905	Chloroplast transit peptide
PpLOX6.1	13-LOX	13-F	Prupe.6G018700.1	947	Chloroplast transit peptide
PpLOX7.1	9-LOX	9-C	Prupe.6G324100.1	923	Chloroplast transit peptide
PpLOX7.2	9-LOX	9-C	Prupe.6G324300.1	933	Chloroplast transit peptide
PpLOX8.1	13-LOX	13-C	Prupe.1G232400.1	920	Chloroplast transit peptide
PpLOX9.1	13-LOX	13-B	Prupe.8G189000.1	839	No signal peptide
PsLOXN2	13-LOX	13-A	CAG44502	826	No signal peptide
SILOX10	13-LOX	13-F	Solyc12g011040	892	Chloroplast transit peptide
SILOX11	13-LOX	13-C	Solyc05g014790	911	Chloroplast transit peptide
SILOX12	9-LOX	9-B	Solyc01g099210	863	No signal peptide
SILOX13	13-LOX	13-F	Solyc01g006560	902	Chloroplast transit peptide
SILOX14	9-LOX	9-A	Solyc09g075870	854	No signal peptide
SILOX6	9-LOX	9-A	Solyc09g075860	877	No signal peptide
SILOX7	9-LOX	9-B	Solyc01g099200	841	No signal peptide
SILOX8	9-LOX	9-B	Solyc08g029000	861	No signal peptide
SILOX9	9-LOX	9-B	Solyc01g099180	854	No signal peptide
StLOXH1	13-LOX	13-F	CAA65268	899	Chloroplast transit peptide
StLOXH3	13-LOX	13-E	CAA65269	914	Chloroplast transit peptide
TomLOXA	9-LOX	9-B	Solyc08g014000	860	No signal peptide
TomLOXB	9-LOX	9-B	Solyc01g099190	858	No signal peptide
TomLOXC	13-LOX	13-F	Solyc01g006540	896	Chloroplast transit peptide
TomLOXD	13-LOX	13-E	Solyc03g122340	908	Chloroplast transit peptide
TomLOXE	9-LOX	9-B	Solyc01g099160	862	No signal peptide
VvLOXA	13-LOX	13-F	ACZ17391	901	Chloroplast transit peptide
VvLOXC	9-LOX	9-B	ACZ17392	859	No signal peptide
VvLOXO	13-LOX	13-E	ACZ17393	916	Chloroplast transit peptide

Annex 9. Interface predicted template modeling (ipTM) and predicted template modeling (pTM) of several proteins. These parameters were obtained from AlphaFold3 server were pTM indicated the quality of the prediction for the model and ipTM indicated the quality of the interaction between the protein and the ligand selected. All lipoxygenase proteins, except for FaLOX10, were modelled with a Fe^{3+} as ligand. GFP: green fluorescent protein; TID: TurboID.

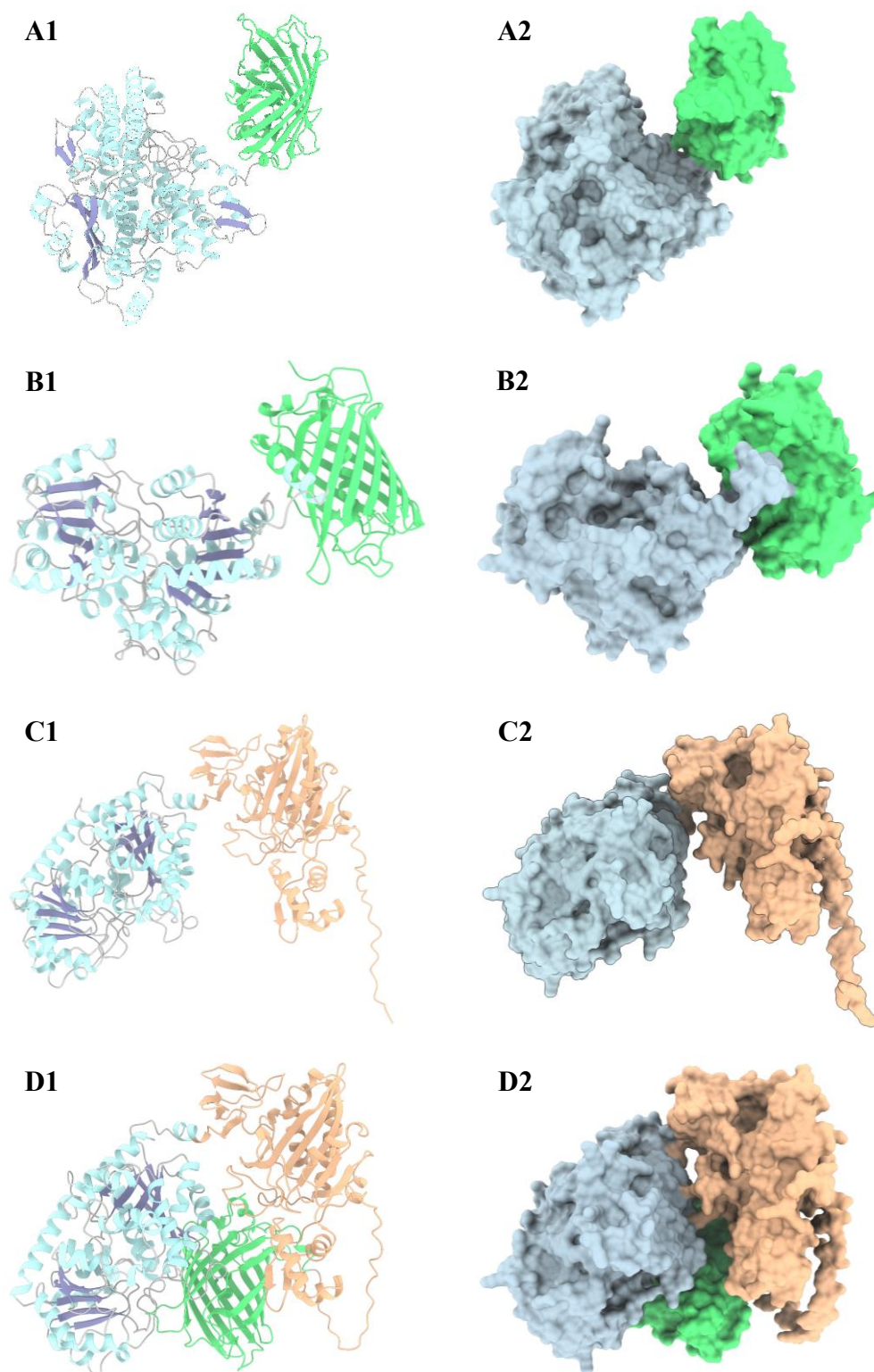
Protein name	N° amino acids	ipTM	pTM
FaLOX2	619	0.99	0.91
FvLOX3	950	0.99	0.84
FaLOX10	306	-	0.94
FaLOX3	950	0.99	0.86
TomLOXC	896	0.99	0.88
FaUDPSS	476	-	0.94
FaUDP	477	-	0.94
FaUGT1	483	-	0.94
FvUGT2	483	-	0.94
AtUGT72B1	480	-	0.94
FaLOX-GFP	873	-	0.74
GFP-FaUDPSS	724	-	0.68
GFP-TID-FaUDPSS	1,075	-	0.56
TID-FaUDPSS	837	-	0.60

Species/Abbrv																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																

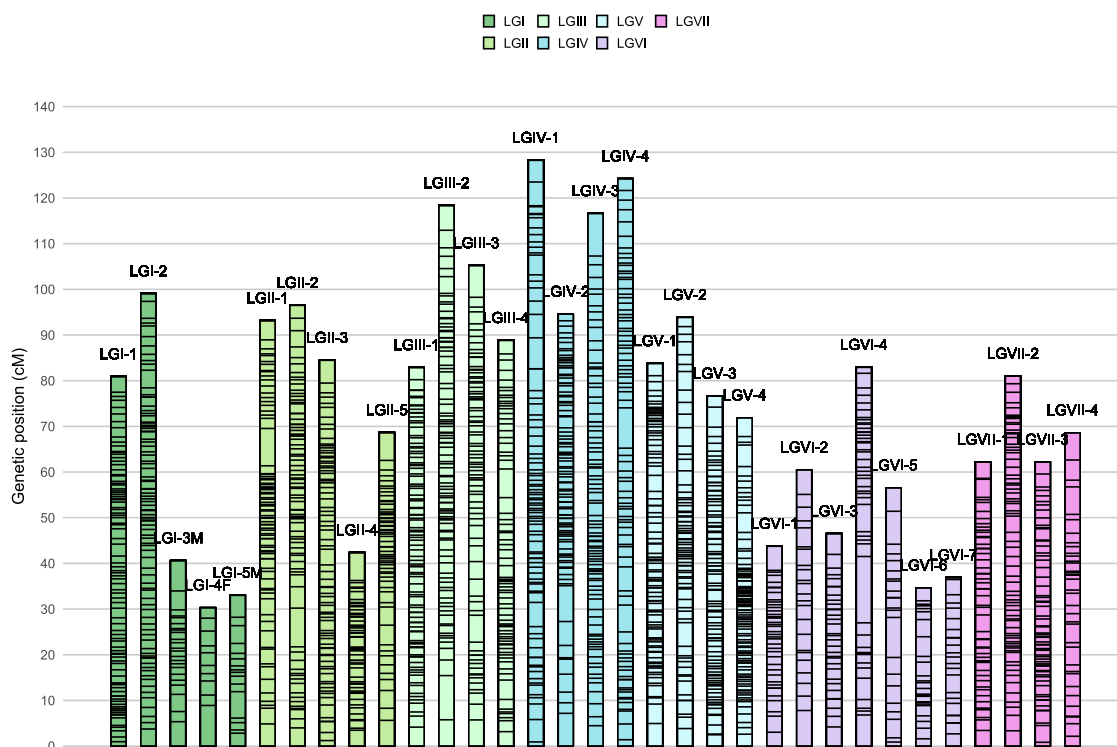
285



Annex 11. Heatmap of primary metabolism relativized to dry-weight (A) and Brix grade of transient overexpression (B). Relative abundance of GC-MS was relativized using data of individual fruits post water elimination (g^{-1}). In GC-MS data P samples represents control fruits while F samples are *FaLOX2*-overexpressing fruits. Statistical significance is represented as p-value < 0.05 for a Mann-Witney *U* test.



Annex 12. Tridimensional structure of fusion proteins used for confocal microscopy and TurboID proximity labelling techniques. Both schematic (X1) and surface (X2) visualization were represented in these images. (A) FaLOX-GFP; (B) GFP-FaUDPSS; (C) TID-FaUDPSS; (D) GFP-TID-FaUDPSS. GFP: green fluorescent protein (green-colored); TID: TurboID (orange-colored). Quality modelling parameters for these proteins can be consulted on **Annex 9**.

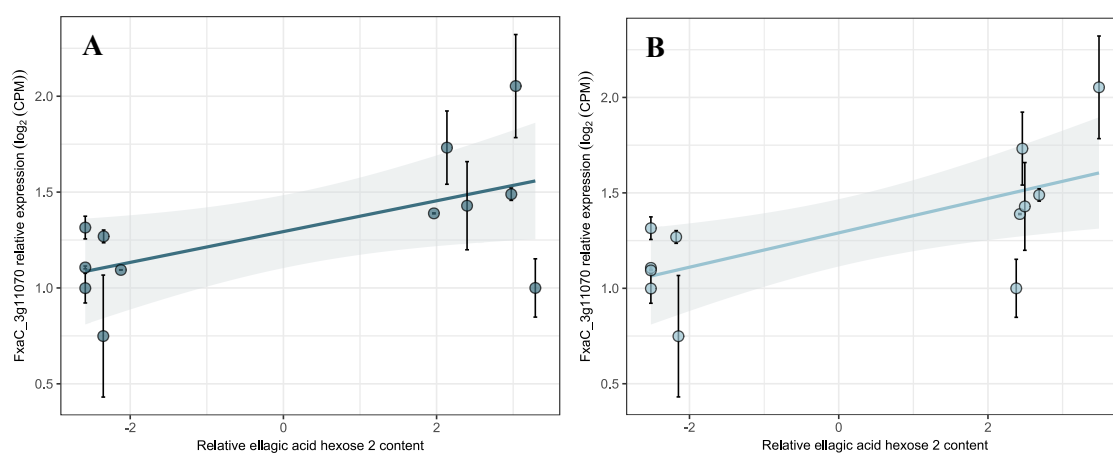


Annex 13. Representation of ‘232’ × ‘1392’ linkage maps. Seven different linkage groups (LG) with subdivisions are represented. Horizontal lines within each chromosome represent genetic markers used in the mapping analysis. Distances between chromosome positions are represented as cM.

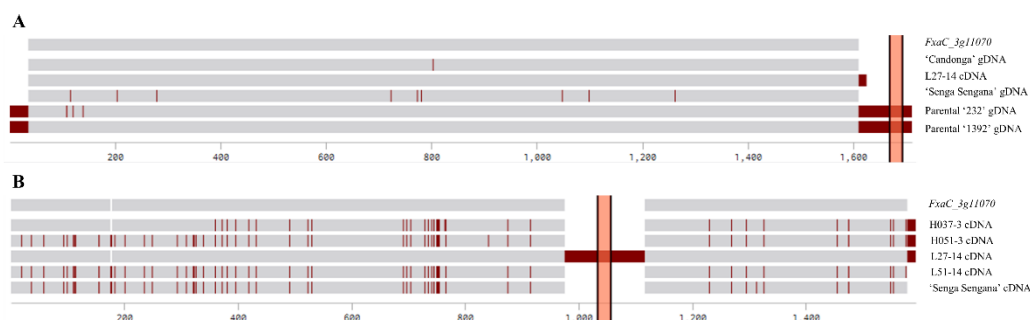
Annex 14. summary of the linkage maps of the segregating population ‘232’ × ‘1392’. Each subdivision of the seven linkage groups (LG) was represented with correspondent number of genetic markers and size. Marker density was calculated as marker/cm. linkage groups (LG) were assigned according to the nomenclature of *Fragaria vesca* v4 genome.

LG	Markers	Size (cM)	Density
LGI-1	93.00	80.88	1.15
LGI-2	90.00	99.04	0.91
LGI-3M	24.00	40.60	0.59
LGI-4F	11.00	30.36	0.36
LGI-5M	17.00	33.02	0.51
LGII-1	83.00	93.13	0.89
LGII-2	80.00	96.50	0.83
LGII-3	77.00	84.47	0.91
LGII-4	54.00	42.34	1.28
LGII-5	58.00	68.62	0.85
LGIII-1	96.00	82.88	1.16
LGIII-2	55.00	118.33	0.46
LGIII-3	81.00	105.20	0.77
LGIII-4	73.00	88.83	0.82

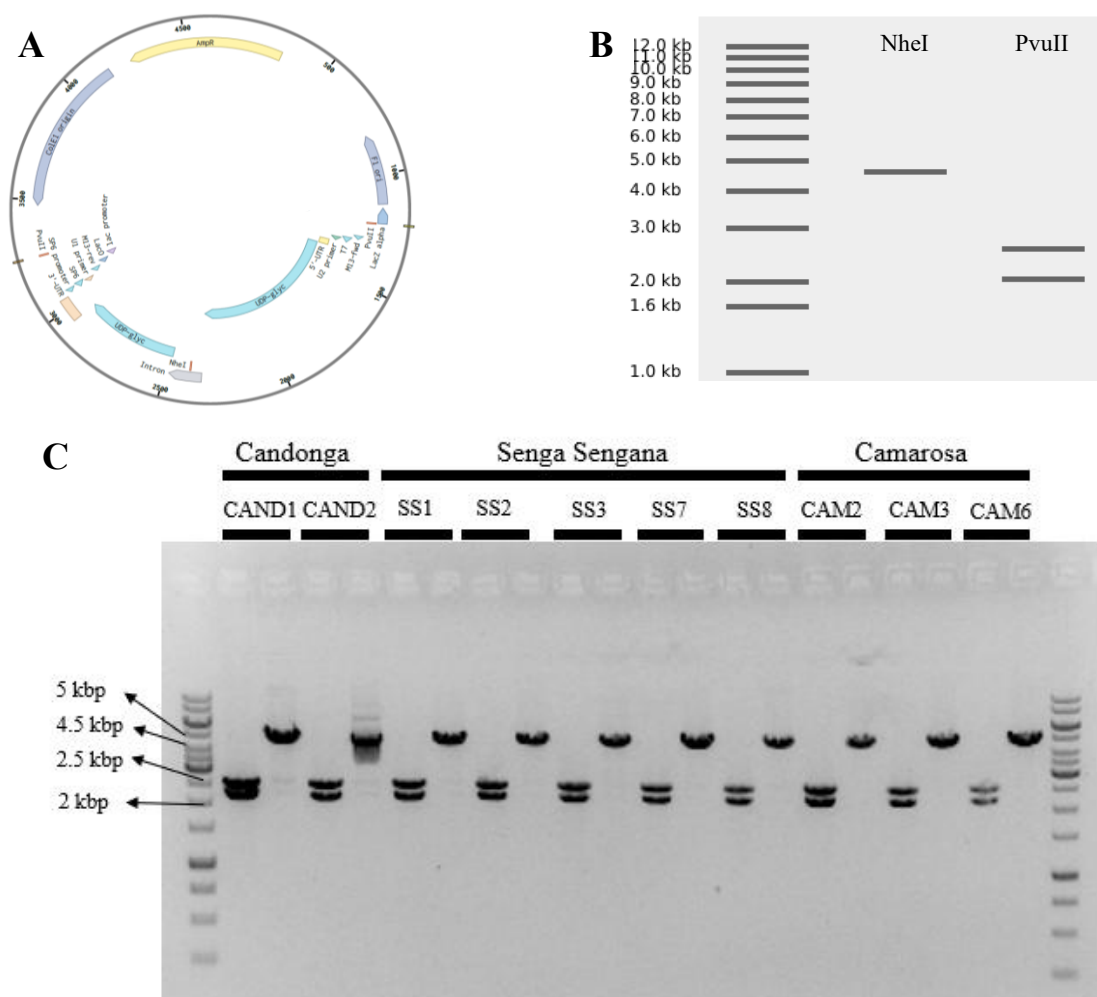
LGIV-1	87.00	128.25	0.68
LGIV-2	96.00	94.56	1.02
LGIV-3	88.00	116.61	0.75
LGIV-4	85.00	124.20	0.68
LGV-1	92.00	83.77	1.10
LGV-2	73.00	93.86	0.78
LGV-3	92.00	76.63	1.20
LGV-4	99.00	71.87	1.38
LGVI-1	41.00	43.76	0.94
LGVI-2	24.00	60.48	0.40
LGVI-3	34.00	46.50	0.73
LGVI-4	48.00	82.93	0.58
LGVI-5	23.00	56.55	0.41
LGVI-6	26.00	34.64	0.75
LGVI-7	24.00	36.88	0.65
LGVII-1	53.00	62.18	0.85
LGVII-2	70.00	80.99	0.86
LGVII-3	64.00	62.17	1.03
LGVII-4	38.00	68.59	0.55
Total	2,049	2,489.62	



Annex 15. Linear correlation of *FxaC_3g11070* gene expression and ellagic acid hexose 2 content. Data obtained from secondary metabolite extraction of Goodberry F₁ individuals. In two different locations. (A) France expression data, $R^2 = 0.4566$; (B) Spain expression data, $R^2 = 0.4566$.



Annex 16. Nucleotide sequence alignments of *FxaC_3g11070*. (A) Alignment of genomic sequences from different individuals; (B) Alignment of fruit cDNA sequences from different individuals. Mismatches are highlighted as vertical lines in red.



Annex 17. Restriction assay of *FxaC_3g11070* with genomic DNA from different strawberry cultivars. (A) Genomic DNA of *FxaC_3g11070* (UDP-glyc) was cloned into pGEM-T Easy vector (PvuII and NheI restriction sites was highlighted in orange). (B) Virtual digestion of the construct with PvuII (2.0 and 2.6 kbp) and NheI (4.6 kbp). (C) Real digestion of different red fruits extractions (CAND, SS and CAM). NheI only possessed one cut site on the intron sequence of the candidate gene.

Annex 18. Additional information about phylogenetic analysis of glucosyltransferases. The species used in this study are *Arabidopsis thaliana* (At), *Camilla sinensis* (Cs), *Disopyros kaki* (Dk), *Fragaria × ananassa* (Fa), *Fragaria vesca* (Fv), *Solanum lycopersicum* (Sl) and *Vitis vinifera* (Vv). ID refers to the gene identification from GDR, TAIR, Sol Genomics or NCBI databases. Aa is the number of amino acids of the complete sequence. The presence or absence of signal peptide was predicted with TargetP2.0 software

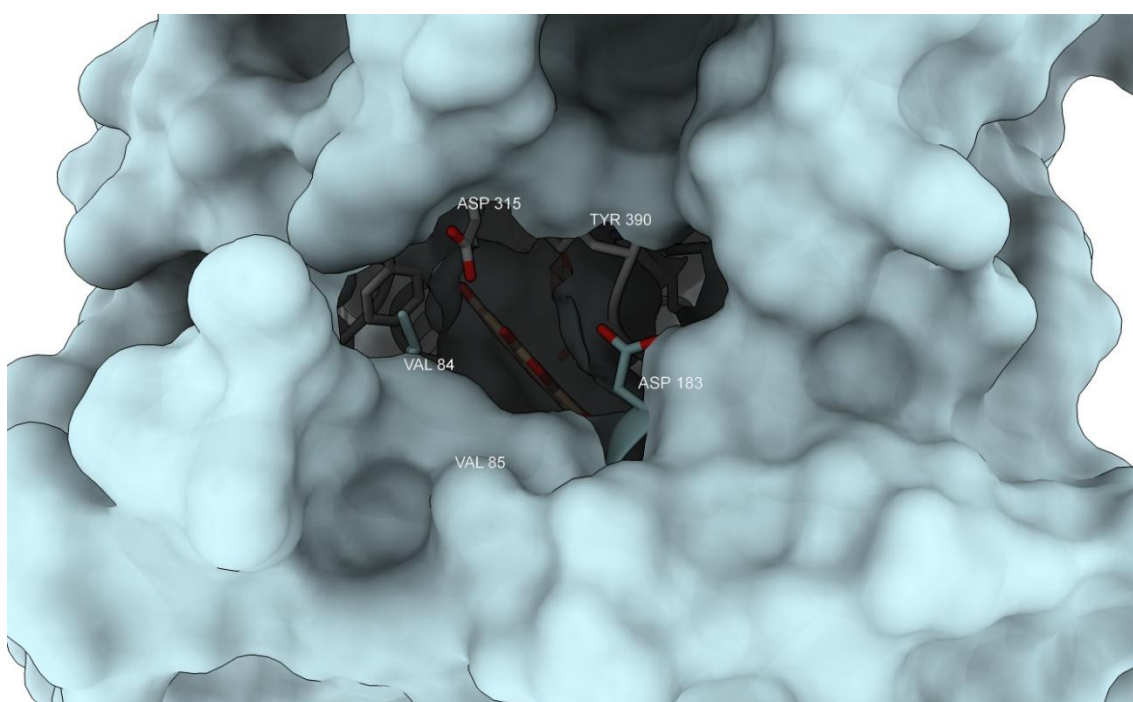
Gene	Clade	ID	Aa	Peptide
AtUGT71B1	E	AT3G21750.1	473	No signal peptide
AtUGT71C5	E	AT1G07240.1	480	No signal peptide
AtUGT72B1	E	AT4G01070.1	480	No signal peptide
AtUGT73B1	D	AT4G34138.1	488	No signal peptide
AtUGT73B2	D	AT4G34135.1	483	Nuclear localization signal
AtUGT73C1	D	AT2G36750.1	491	No signal peptide
AtUGT73C6	D	AT2G36790.1	495	No signal peptide
AtUGT74C1	L	AT2G31790.1	457	Nuclear localization signal
AtUGT74F2	L	AT2G43820.1	449	Nuclear localization signal
AtUGT75B2	L	AT1G05530.1	455	No signal peptide
AtUGT75C1	L	AT4G14090.1	456	Nuclear localization signal
AtUGT76B1	H	AT3G11340.1	447	No signal peptide
AtUGT76C2	H	AT5G05860.1	450	No signal peptide
AtUGT78D1	F	AT1G30530.1	453	Nuclear localization signal
AtUGT78D2	F	AT5G17050.1	460	Nuclear localization signal
AtUGT78D3	F	AT5G17030.1	459	Nuclear localization signal
AtUGT79B1	A	AT5G54060.1	468	No signal peptide
AtUGT79B2	A	AT4G27560.1	455	No signal peptide
AtUGT79B3	A	AT4G27570.1	453	No signal peptide
AtUGT79B4	A	AT3G29630.1	448	Nuclear localization signal
AtUGT79B6	A	AT5G54010.1	453	No signal peptide
AtUGT82A1	F	AT3G22250.1	461	Nuclear localization signal
AtUGT84A1	L	AT4G15480.1	490	No signal peptide
AtUGT85A1	G	AT1G22400.1	489	No signal peptide
AtUGT86A1	J	AT2G36970.1	490	No signal peptide
AtUGT87A2	J	AT2G30140.1	455	Nuclear localization signal
AtUGT88A1	E	AT3G16520.3	462	No signal peptide
AtUGT89B1	B	AT1G73880.1	473	No signal peptide
AtUGT89C1	B	AT1G06000.1	435	Nuclear localization signal
AtUGT90A1	C	AT2G16890.2	478	No signal peptide
AtUGT91A1	A	AT2G22590.1	470	No signal peptide
AtUGT92A1	C	AT5G12890.1	488	No signal peptide
CsUGT78A14	F	KP682360.1	459	No signal peptide
DkFGT	F	AB435084.1	459	No signal peptide
FaF3GT	F	FxaC_28g00600.t1	465	Nuclear localization signal
FaF3OGT	F	FxaC_28g00600.t1	463	Nuclear localization signal
FaGT1	F	FxaC_26g03070.t1	466	Nuclear localization signal
FaGT2	L	FxaC_6g04890.t1	555	Nuclear localization signal
FaGT3	E	FxaC_22g17370.t1	478	No signal peptide

Gene	Clade	ID	Aa	Peptide
FaGT4	A	FxaC_1g14300.t1	478	No signal peptide
FaGT6	E	FxaC_21g18240.t1	479	No signal peptide
FaGT7	D	FxaC_8g33670.t1	487	No signal peptide
FaUDP	E	FxaC_3g11070.t1	477	No signal peptide
FaUDPSS	E	FxaC_3g11070.t1	476	No signal peptide
FaUGT1	E	FxaC_1g15671.1	483	No signal peptide
FaUGT2	D	FxaC_2g07670.t1	571	No signal peptide
FaUGT3	L	FxaC_2g11400.t1	177	No signal peptide
FaUGT4	E	FxaC_2g19510.t1	485	No signal peptide
FaUGT5	L	FxaC_2g22670.t1	477	No signal peptide
FaUGT6	G	FxaC_2g27710.t1	479	Nuclear localization signal
FaUGT7	G	FxaC_2g27740.t1	271	Nuclear localization signal
FaUGT8	L	FxaC_3g00842.t1	496	Nuclear localization signal
FaUGT9	L	FxaC_3g00843.t1	461	Nuclear localization signal
FaUGT10	E	FxaC_3g15620.t1	480	No signal peptide
FaUGT11	E	FxaC_3g15740.t1	205	Peroxisomal targeting signal
FaUGT12	G	FxaC_3g23520.t1	183	Peroxisomal targeting signal
FaUGT13	G	FxaC_3g23570.t1	464	No signal peptide
FaUGT14	E	FxaC_4g02230.t1	475	No signal peptide
FaUGT15	G	FxaC_4g16910.t1	287	Nuclear localization signal
FaUGT16	L	FxaC_4g39600.t1	401	Nuclear localization signal
FaUGT17	A	FxaC_5g17470.t1	691	No signal peptide
FaUGT18	L	FxaC_5g20100.t1	468	Nuclear localization signal
FaUGT19	A	FxaC_5g26110.t1	676	No signal peptide
FaUGT20	L	FxaC_6g02491.t1	461	Nuclear localization signal
FaUGT21	L	FxaC_6g04890.t1	555	Nuclear localization signal
FaUGT22	D	FxaC_6g05640.t1	969	No signal peptide
FaUGT23	D	FxaC_6g05641.t1	504	No signal peptide
FaUGT24	A	FxaC_10g09490.t1	435	Nuclear localization signal
FaUGT25	H	FxaC_10g12210.t1	442	No signal peptide
FaUGT26	E	FxaC_10g14050.t1	504	No signal peptide
FaUGT27	J	FxaC_10g44980.t1	491	No signal peptide
FaUGT28	J	FxaC_10g45070.t1	491	No signal peptide
FaUGT29	D	FxaC_10g50410.t1	214	No signal peptide
FaUGT30	G	FxaC_11g16320.t1	356	No signal peptide
FaUGT31	H	FxaC_11g38930.t1	452	Nuclear localization signal
FaUGT32	H	FxaC_11g38940.t1	384	Nuclear localization signal
FaUGT33	H	FxaC_11g38970.t1	234	Nuclear localization signal
FaUGT34	J	FxaC_11g41520.t1	456	Nuclear localization signal
FaUGT35	J	FxaC_11g41521.t1	431	No signal peptide
FaUGT36	J	FxaC_11g41530.t1	449	No signal peptide
FaUGT37	J	FxaC_12g06810.t1	456	No signal peptide
FaUGT38	E	FxaC_12g08170.t1	478	No signal peptide
FaUGT39	H	FxaC_12g08870.t1	444	Nuclear localization signal
FaUGT40	A	FxaC_12g14740.t1	286	Signal peptide

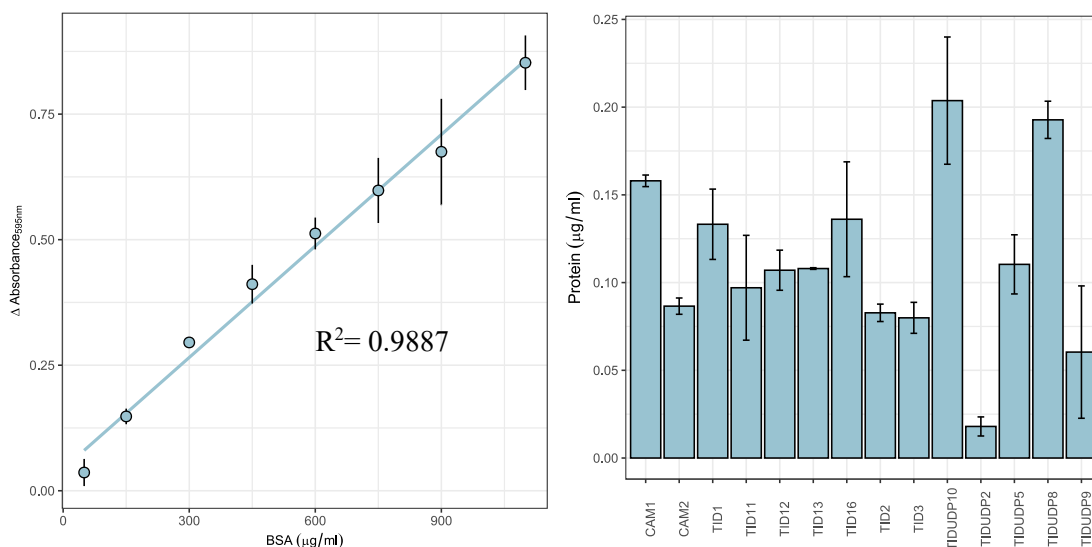
FaUGT41	G	FxaC_12g32560.t1	482	Nuclear localization signal
FaUGT42	A	FxaC_12g36620.t1	452	Nuclear localization signal
FaUGT43	A	FxaC_12g48140.t1	302	Signal peptide
FaUGT44	A	FxaC_13g02640.t1	393	Signal peptide
FaUGT45	L	FxaC_13g04351.t1	140	No signal peptide
FaUGT46	C	FxaC_13g13900.t1	485	No signal peptide
FaUGT47	C	FxaC_13g15890.t1	489	No signal peptide
FaUGT48	A	FxaC_13g23660.t1	324	Signal peptide
FaUGT49	L	FxaC_13g30440.t1	516	No signal peptide
FaUGT50	L	FxaC_13g30460.t1	473	Nuclear localization signal
FaUGT51	L	FxaC_14g26750.t1	473	Nuclear localization signal
FaUGT52	L	FxaC_14g26760.t1	475	Nuclear localization signal
FaUGT53	L	FxaC_14g26770.t1	516	No signal peptide
FaUGT54	A	FxaC_14g26790.t1	281	Signal peptide
FaUGT55	L	FxaC_14g28180.t1	463	Nuclear localization signal
FaUGT56	D	FxaC_28g12750.t1	207	No signal peptide
FaUGT57	E	FxaC_28g20520.t1	473	Nuclear localization signal
FaUGT58	B	FxaC_28g30760.t1	691	Nuclear localization signal
FaUGT71K3	E	FxaC_10g43100.t1	478	No signal peptide
FvUGT1	E	FvH4_1g13990.t1	491	No signal peptide
FvUGT2	E	FvH4_1g14013.t1	483	No signal peptide
FvUGT71A33	E	FvH4_6g39430.t1	478	No signal peptide
FvUGT71A34	E	FvH4_6g39410.t1	515	No signal peptide
FvUGT71A35	E	FvH4_3g10810.t1	475	No signal peptide
FvUGT71W2	E	FvH4_6g39311.t1	475	No signal peptide
FvUGT75T1	L	FvH4_7g22821.t1	454	Nuclear localization signal
SIUGT91A6	A	Solyc11g010740.3.1	483	No signal peptide
SIUGT91A7	A	Solyc11g010760.1.1	474	No signal peptide
SIUGT91R1	A	Solyc11g010790.1.1	470	No signal peptide
SIUGT91R2	A	Solyc11g010790.1.1	470	No signal peptide
SIUGT91R3	A	Solyc11g010810.1.1	470	No signal peptide
VvGT1	F	AAB81682.1	456	No signal peptide
VvGT5	F	BAI22846.1	459	No signal peptide
VvGT6	F	BAI22847.1	458	No signal peptide
VvGT14	G	XM_002285734.2	473	No signal peptide
VvGT16	G	XP_002285779	475	No signal peptide

Annex 19A. Distances between amino acids from S1, S2 and P pockets in FaUDPSS protein model. Amino acids selected were closely related with pocket surfaces. Distances were measured as the diameter assuming circular shape of each pocket using ChimeraX.

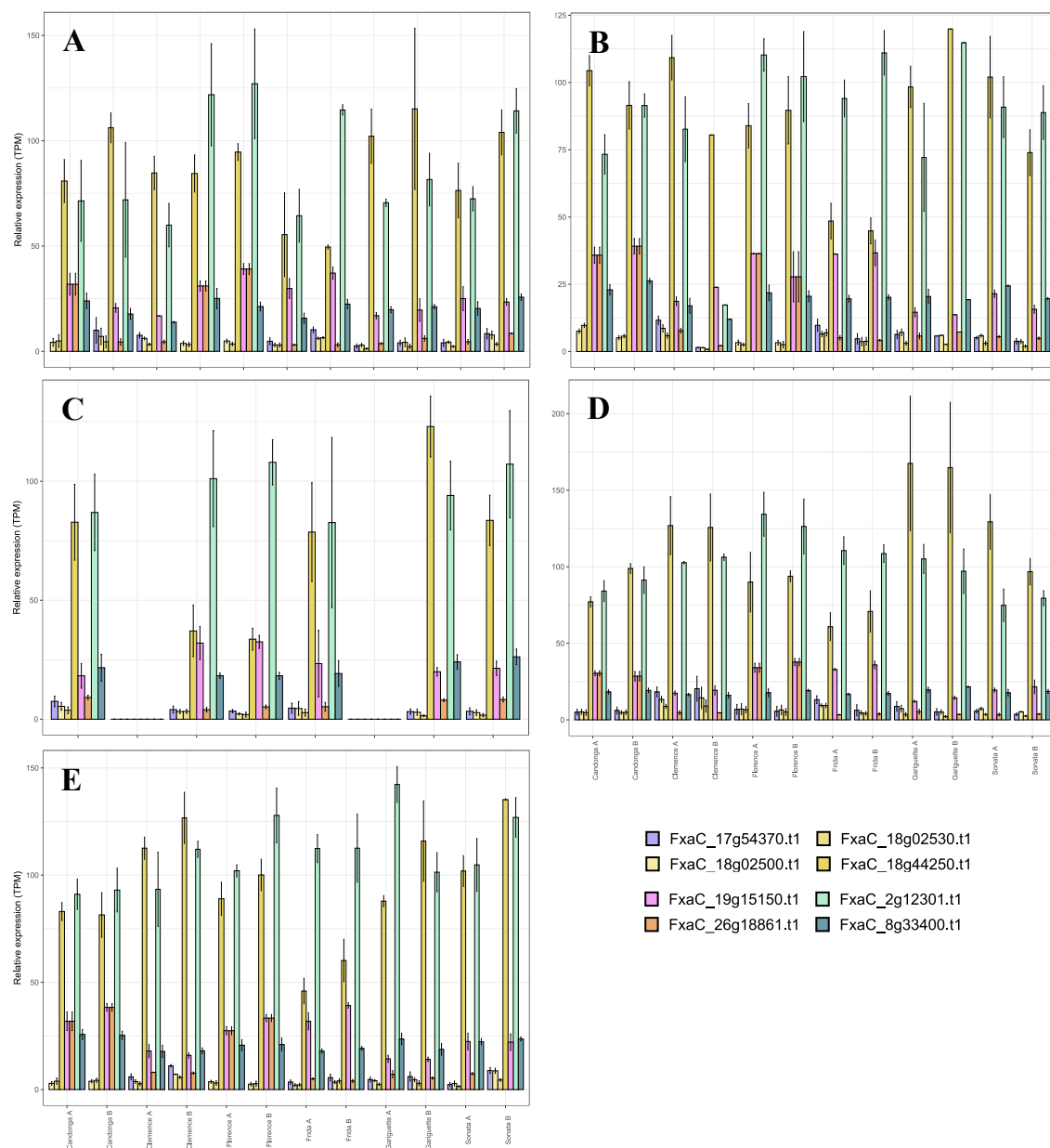
Amino acid 1	Amino acid 2	Diameter (Å)	Radius (Å)	Radius (nm)	Area (nm ²)
S1 pocket					
Arg251	Trp350	5.88	2.94	0.29	0.27
Arg248	Trp350	7.35	3.68	0.37	0.42
Arg248	Pro352	7.80	3.98	0.40	0.50
S2 pocket					
Asp315	Val85	7.15	3.58	0.36	0.40
Val84	Asp183	7.90	3.95	0.40	0.49
Val84	Tyr390	7.03	3.52	0.35	0.39
P pocket					
Phe319	Glu51	6.29	3.15	0.32	0.31
Ser48	Phe319	8.20	4.10	0.41	0.53
Phe318	Glu51	6.54	3.27	0.33	0.34



Annex 19B. Example of selected amino acids for calculating S1 pocket area of FaUDPSS protein. Important amino acids that form the cavity of S1 pocket are highlighted and represented as ball and stick models.



Annex 21. Quantification of protein concentrations in strawberry fruits used for TurboID proximity labelling assay. CAM: ‘Camarosa’ fruits untreated; TID1: ‘Camarosa’ fruits agroinfiltrated with free TurboID; TIDUDP: ‘Camarosa’ fruits agroinfiltrated with TurboID-FaUDPSS fusion protein.



Annex 22. Gene expression of some TurboID candidates in several commercial strawberry varieties and in several locations. (A) France, (B) Germany, (C) Italy, (D) Norway and (E) Poland. Relative gene expression in transcript per million (TPM).