

TECHNISCHE UNIVERSITÄT MÜNCHEN
Professur für Biotechnologie der Naturstoffe

FORMATION OF ELLAGIC ACID PRECURSORS IN
FRAGARIA

KATJA HÄRTL

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan
für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur
Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
genehmigten Dissertation.

VORSITZENDER: Prof. Dr. Harald Luksch

PRÜFER DER DISSERTATION: 1. Prof. Dr. Wilfried Schwab

2. Prof. Dr. Brigitte Poppenberger-Sieberer

Die Dissertation wurde am 23.01.2017 bei der Technischen Universität München ein-
gereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernäh-
rung, Landnutzung und Umwelt am 25.04.2017 angenommen.

ACKNOWLEDGMENTS

I'd like to take the opportunity to say "*Thank You*" to everyone who supported me during my PhD. Time flies - and so did four years go by so quickly.

Heartfelt gratitude is dedicated to my supervisor Professor Dr. Wilfried Schwab. Firstly, you accepted me to join the BiNa-Lab team and guided my research. Secondly and more importantly, you gave me the freedom to contribute my own ideas while motivating me through challenging situations, not to mention your constant support and eternal optimism. I underwent quite a development, both professionally and personally.

Maybe not a single LC-MS measurement would have been possible without Dr. Thomas Hoffmann. You showed me how to analyze the data, how to repair the machines, and how to setup my labeling experiment. But not only because of all this you deserve my sincere appreciation, you were always ready to listen to problems of any sort, and gave me valuable advice. The same can be said about Dr. Fong-Chin Huang and Dr. Thilo Fischer. Thank you all very much.

Furthermore, I'd like to acknowledge the whole Bina-Team. Thank you all for the nice working atmosphere. Elisabeth, Kate, and Guangxin, thank you for all your support. I had and will continue to have a great time in our girls-only office. Dear Kate thank you very much for editing the language of my thesis. Chuankui thank you for kindly providing the *UGT71K3* construct. Katrin, Doreen, and Friede thank you for your friendship and all your help. When I started back in 2012 you introduced me to the general workflow, integrated me into the team and explained, supported, encouraged... I miss you guys. Heike, Ruth, Mechthild, Kilian and Hannelore thank you for your support and help in administrative and organizational issues, which are very important for a PhD student's well being.

Special thanks go to my collaboration partners Antje Feller and Alisandra Denton. I had a great time in Italy and Aachen. Your kind help with experiments and manuscripts I will never forget.

I'm also grateful to the department of Animal Physiology and Immunology at TUM in particular, Dr. Melanie Spornraft for providing the Bioanalyzer and for a comprehensive introduction to RNA-Sequencing and R.

Moreover, I thank Prof. Dr. Poppenberger and Prof. Dr. Luksch for their participation as members of my examination board, the DFG and TUM for funding, and last but not least, a big thanks to all friends for constant support.

I dedicate this thesis to my dear, supporting, patient, and motivating family: my parents, my parents-in-law, and above all, my husband Fabian. I love you.

ABSTRACT

Phenols, a vast class of plant secondary metabolites, are derived from the shikimate and phenylpropanoid pathway. Among them are tannins, which can be classified in condensed tannins and hydrolyzable tannins. Amongst the latter are ellagitannins/ellagic acid (ETs/EA), considered as a group of particularly bio-active polyphenols. *In planta* they act in defense against bacteria, fungi, and viruses, but are also of utmost importance for the public health. They have been found to scavenge free radicals, exhibit antioxidant properties, and guard off diseases such as hormonal cancers or cardiovascular disorders. Despite their universally acknowledged significance for human health, surprisingly their biosynthesis in plants remains elusive. Therefore, this doctorate thesis concentrates on the identification and characterization of genes and encoding enzymes participating in the formation of pentagalloylglucose (PGG), the first known precursor of ETs/EA. Strawberry, utilized here as the plant model organism, is known to contain high amounts of ETs/EA.

The first intermediate compound of the ETs/EA biosynthetic pathway, is β -glucogallin which arises from an ester bond formation between gallic acid and glucose. This reaction is enabled by UDP-glucose dependent glucosyltransferases (UGTs). Here, five UGTs from strawberry and raspberry were identified that catalyze the formation of β -glucogallin. Subsequent elucidation of the substrate specificity revealed promiscuous functions, and determination of the kinetic constants confirmed their role as gallic acid UGTs. By site-directed mutagenesis, three amino acids with notable effect on the substrate preference, but positioned outside of the activity center of the protein, were identified. The secondary metabolites were analyzed in strawberry fruits of different developmental stages to determine the distribution of the ET precursors. Green fruits were

found to be the main source. Finally, injection of deuterium-labeled gallic acid into green fruits of stable transgenic plants, confirmed the UGT function *in vivo*.

In subsequent steps, additional gallic acid residues are attached to the free hydroxy-groups of the glucose core of β -glucogallin to yield pentagalloylglucose (PGG). These acyl-transfers are probably mediated by serine carboxypeptidase-like (SCPL) acyltransferases (ATs). Following phylogenetic comparison of the strawberry genome to known ATs from other plants, several candidate genes were found and considered. The expression pattern of putative ATs is supposedly concomitant to the metabolic profile of β -glucogallin, as the reaction can only proceed upon the availability of the substrate in the fruit tissue. Therefore, it was attempted to narrow down the list of genes by performing RNA-sequencing (RNA-seq) and metabolic profiling of strawberry fruits in progressive ripening stages. One red-fruited and two natural white-fruited strawberry genotypes were selected. Global analysis of the transcriptomes showed that variance in gene expression is highest within the intermediate ripening stages. Meanwhile, metabolite profiling revealed that already green tissues of the red- and white-fruited varieties feature a particular polyphenol profile. Key polyphenol genes, such as anthocyanidin glucosyltransferase were found to be down-regulated in the white genotypes compared to the red genotype. Likewise, a transcription factor of the bHLH class, and a putative glutathione S-transferase were among the most highly differentially expressed genes. Furthermore, genes associated with flavor formation and fruit softening seemed to interact with the biosynthetic pathway of the polyphenols, as they exhibited coordinate expression.

Consequently, three putative AT genes were selected for subsequent *in vitro* and *in vivo* characterization. Two candidates were successfully cloned in an appropriate expression vector and recombinant protein expression was attempted in *Saccharomyces cerevisiae*. However, no recombinant protein and no *in vitro* activity could be detected. Alternatively, the two genes were over-expressed in ripening strawberry fruits, and

elevated EA concentrations in the transgenics substantiated their *in planta* function as putative galloyltransferases.

The results provide the foundation for the breeding of strawberry with improved health benefits and the biotechnological production of bio-active plant metabolites.

ZUSAMMENFASSUNG

Die große und heterogene Klasse der phenolischen Naturstoffe wird im Shikimisäure- und Phenylpropanoid-Stoffwechselweg gebildet und zählt zur Gruppe der pflanzlichen Sekundärmetabolite. Dazu gehören auch die Tannine, die wiederum in kondensierte und hydrolysierbare Tannine unterteilt werden. Ellagitannine/Ellagsäure (ETs/EA) zählen zu den hydrolysierbaren Tanninen. Ihnen werden diverse gesundheitsfördernde Wirkungen zugeschrieben. So können sie zum Beispiel freie Radikale abfangen, antioxidative Wirkungen entfalten und vor Krankheiten wie Krebs und kardiovaskulären Störungen schützen. Die biologische Bedeutung dieser Verbindungen ist seit längerem bekannt und wurde bereits mehrfach bestätigt. Der genetische Hintergrund ihrer Biosynthese in der Pflanze ist jedoch wenig erforscht. Ziel dieser Arbeit war es daher die Gene und kodierenden Enzyme zu identifizieren, die an der Bildung des ersten bekannten ET-Vorläufermoleküls Pentagalloylglucose (PGG) beteiligt sind. Da Erdbeerfrüchte reich an bioaktiven Polyphenolen sind, wurde die Pflanze als Modellorganismus gewählt.

Durch Veresterung von Gallussäure und Glucose wird β -Glucogallin gebildet, das erste Intermediat der ET-Biosynthese. Diese Reaktion wird durch UDP-Glucose abhängige Glucosyltransferasen (UGTs) vermittelt. Im Rahmen dieser Doktorarbeit wurden fünf UGTs aus Erdbeere und Himbeere identifiziert, welche die Bildung von β -Glucogallin katalysieren. Die rekombinanten Enzyme zeigten Promiskuität gegenüber einer Reihe von Substraten und die Bestimmung der kinetischen Konstanten bestätigte ihre Funktionen als Gallussäure-UGTs. Durch positionsgerichtete Mutagenese konnten drei Aminosäuren identifiziert werden, die zwar außerhalb des aktiven Zentrums der Proteine lagen, sich aber dennoch deutlich auf die Substratpräferenz auswirkten. Zusätzlich wurde die Konzentration der

ET-Vorläufermoleküle in Erdbeeren verschiedener Reifestadien quantifiziert. Grüne, unreife Früchte konnten als Hauptquelle ermittelt werden. Abschließend bestätigte die Injektion stabil-isotopenmarkierter Gallussäure in grüne Früchte transgener Pflanzen die *in vivo* Funktion der UGTs.

Im Biosyntheseweg von PGG werden nachfolgend weitere Gallussäuremoleküle mit freien Hydroxylgruppen des Glucoserestes von β -Glucogallin verestert. Die Übertragung dieser Acylgruppen wird vermutlich von Serine Carboxypeptidase-ähnlichen (SCPL) Acyltransferasen (ATs) katalysiert. Um Kandidatengene zu identifizieren wurde ein phylogenetischer Vergleich zwischen bekannten ATs anderer Pflanzen und dem *F. vesca* Genom durchgeführt, wodurch einige Gene für eine weiterführende funktionelle Charakterisierung in Betracht gezogen werden konnten. Das Expressionsmuster putativer ATs korreliert wahrscheinlich mit der zeitlichen und räumlichen Verfügbarkeit ihres vermuteten Substrates, dem β -Glucogallin. Durch einen kombinierten Ansatz aus einer RNA-Sequenzierung (RNA-seq) und einer Quantifizierung der Sekundärmetabolite wurde die Expression der putativen ATs mit der Konzentration an Substrat in Erdbeeren verschiedener Reifestadien in Verbindung gebracht. So konnte die Liste der möglichen Kandidatengene verkleinert werden. Drei Walderdbeervarietäten wurden hierfür ausgewählt: Eine, die rot-reife Früchte bildet und zwei, die weiß-reife Früchte tragen. Nach Analyse der Transkriptome zeigte sich, dass die differentielle Genexpression in intermediären Früchten am höchsten ist. Die Quantifizierung der Metabolite dagegen offenbarte signifikante Unterschiede zwischen den weißen und roten Fruchtsorten bereits im grünen Entwicklungsstadium. Das Expressionsprofil bekannter Gene des Anthocyanin/Flavonoid-Biosynthesewegs, wie zum Beispiel der Anthocyanidin Glucosyltransferase, war in den weißen Genotypen im Vergleich zum roten Genotyp signifikant herunterreguliert. Ebenso zeigte ein Transkriptionsfaktor und eine putative Glutathion S-Transferase differentielle Expression. Des Weiteren konnte festgestellt werden, dass assoziierte Gene der Aromaproduktion und des Weichwerdens der überreifen

Frucht mit dem Biosyntheseweg der Polyphenole interagieren, da ihre Expressionsprofile aufeinander abgestimmt waren.

Im Folgenden wurden drei putative ATs für die *in vitro* und *in vivo* Charakterisierung ausgewählt. Zwei der Kandidaten konnten erfolgreich in einen geeigneten Expressionsvektor kloniert und in *Saccharomyces cerevisiae* transformiert werden. Jedoch konnte keine heterologe Expression der Proteine nachgewiesen werden. Alternativ dazu wurden beide Gene in reifenden Erdbeerfrüchten überexprimiert. In den transgenen Früchten konnte eine erhöhte Konzentration an EA ermittelt werden. Dieses Ergebnis bekräftigt die Annahme, dass die Kandidatengene *in vivo* als Galloyltransferasen agieren könnten.

Die Forschungsergebnisse schaffen die Grundlage für die Züchtung von Erdbeersorten mit verbessertem gesundheitlichen Nutzen und die biotechnologische Produktion bioaktiver Pflanzenmetabolite.

CONTENTS

1 INTRODUCTION	1
1.1 The strawberry plant	1
1.1.1 Origin	1
1.1.2 Fruit ripening	2
1.1.3 Plant secondary metabolism	3
1.2 Metabolite biosynthesis downstream of the shikimate pathway	4
1.2.1 The phenylpropanoid pathway	4
1.2.2 The lignin pathway	4
1.2.3 The flavonoid/anthocyanin pathway	6
1.2.4 Biosynthesis of tannins	7
1.3 From Ellagitannins to Ellagic Acid	13
1.3.1 Biosynthesis of Ellagic Acid	13
1.3.2 Physicochemical Properties	13
1.3.3 Antimicrobial activities	15
1.3.4 Health benefits	16
1.4 Aims of the thesis	17
2 MATERIAL	19
2.1 Plant material	19
2.2 Chemicals	19
2.3 Prokaryotic and eukaryotic microorganisms	20
2.4 Primers	20
2.5 Equipment	23
2.5.1 Liquid chromatography ultraviolet electrospray ionization mass spectrometry (LC-UV-ESI-MSn, LC-MS)	23
2.5.2 General Equipment	23
2.6 Software and internet resources	28
3 METHODS	29
3.1 Basic techniques	29

3.1.1	Extraction of secondary metabolites from strawberry and blueberry tissues and quantification by LC-MS	29
3.1.2	RNA-Isolation from strawberry fruits	30
3.1.3	cDNA synthesis	30
3.1.4	Polymerase chain reaction (PCR)	30
3.1.5	Isolation of plasmid DNA from <i>E. coli</i>	32
3.1.6	Digestion of DNA by restriction endonucleases	32
3.1.7	Agarose gel electrophoresis	33
3.1.8	Ligation of DNA fragments	34
3.1.9	Transformation of plasmid DNA in microorganisms	34
3.1.10	Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)	35
3.1.11	Western Blot	37
3.2	Characterization of gallate-UDP-glucose dependent glucosyltransferases (UGTs)	38
3.2.1	Construction of UGT plasmids	38
3.2.2	Heterologous expression and purification of recombinant gallate-UGTs in <i>E. coli</i>	39
3.2.3	Radio-labeled enzyme assays and kinetics	39
3.2.4	Injection of deuterium-labeled gallic acid into strawberry fruits	40
3.2.5	Identification and quantification of metabolites by LC-MS analysis	40
3.3	Transcriptome and metabolome analysis of <i>F. vesca</i>	40
3.3.1	Fruit sample preparation	41
3.3.2	Analysis of <i>F. vesca</i> metabolites	41
3.3.3	Bioinformatic analysis and evaluation of the sequencing data	41
3.3.4	Validation of candidate gene expression by real-time (RT)-PCR	43
3.4	Characterization of putative acyltransferases (ATs)	44

3.4.1	Construction of AT plasmids	44
3.4.2	Heterologous expression and purification of recombinant ATs in <i>S. cerevisiae</i>	45
3.4.3	Enzyme assays and product detection via LC-MS	45
3.4.4	Over-expression of candidate ATs in straw- berry fruits	46
3.4.5	Validation of AT over-expression by RT- PCR	47
4	RESULTS	49
4.1	Formation of galloyl glucose	49
4.1.1	Identification of putative gallic acid UGTs from <i>F. vesca</i> , <i>F. x ananassa</i> and <i>R. idaeus</i>	49
4.1.2	Enzymatic UGT activity	50
4.1.3	Determination of the kinetic values	53
4.1.4	Site-directed mutagenesis of FaGT2*	53
4.1.5	Metabolite Analysis in <i>F. x ananassa</i> and <i>F.</i> <i>vesca</i>	53
4.1.6	<i>In vivo</i> confirmation of FaGT2* and FaGT2 activity	56
4.2	Transcriptome and metabolome analysis of <i>F.</i> <i>vesca</i>	60
4.2.1	Metabolite profiling	61
4.2.2	Transcriptome profiling - global data anal- ysis	64
4.2.3	Transcriptome profiling - differentially ex- pressed genes	64
4.2.4	Transcriptome profiling - flavonoid/an- thocyanin pathway	71
4.2.5	Transcriptome profiling - flavor formation and fruit softening	71
4.2.6	Expression of <i>FvGT2</i> and <i>FvGT5</i> in <i>F.</i> <i>vesca</i>	73
4.3	Formation of 1,2,3,4,6-pentagalloylglucose	75
4.3.1	Identification of AT gene candidates	75

4.3.2	<i>In vitro</i> characterization of acyltransferase candidates	76
4.3.3	Up-regulation of FvAT1 and 3 in strawberry fruits	77
5	DISCUSSION	83
5.1	Catalysis of galloyl glucose by UGTs	83
5.1.1	Gallic acid UGTs in <i>F. vesca</i> , <i>F. x ananassa</i> , and <i>R. idaeus</i>	83
5.1.2	Enzymatic activity and biochemical characterization	84
5.1.3	Correlation of UGT expression pattern and metabolite levels	85
5.1.4	FaGT2/FaGT2* <i>in vivo</i> activity	86
5.2	Woodland strawberry metabolome and transcriptome	87
5.2.1	Metabolite Analysis	87
5.2.2	Global transcriptome data analysis and evaluation of differential gene expression	88
5.2.3	Transcript levels of functionally confirmed anthocyanin/flavonoid biosynthesis genes	91
5.2.4	Interaction of polyphenol metabolism, fruit flavor formation and softening	92
5.3	Identification and characterization of putative ATs	93
5.3.1	Trans-esterification mediated by serine carboxypeptidase-like (SCPL) acyltransferases	93
5.3.2	Production of recombinant FvAT1 and 3	94
5.3.3	Transient over-expression of candidate ATs in strawberry fruits	95
6	CONCLUSIONS AND OUTLOOK	97
	BIBLIOGRAPHY	99
	PUBLICATIONS OF THE AUTHOR	117

LIST OF FIGURES

Figure 1	Schematic illustration of the shikimate, ellagic acid, phenylpropanoid, flavonoid and anthocyanin pathway.	5
Figure 2	Representative examples and subunits of "true" plant polyphenols	7
Figure 3	Possible biosynthesis pathways of gallic acid	10
Figure 4	UGT reaction mechanism	11
Figure 5	AT reaction mechanism	12
Figure 6	Possible pathway of ellagic acid biosynthesis in strawberry	14
Figure 7	Protein sequence alignment of gallate-UGTs	51
Figure 8	FaGT2* mutant substrate screen	55
Figure 9	Heatmap of relative concentration of GA, β-glucogallin, and EA in strawberry tissues	57
Figure 10	<i>In vivo</i> confirmation of FaGT2/FaGT2* activity	58
Figure 11	Identification of gallic acid 4-O-glucoside	59
Figure 12	Woodland strawberry varieties analyzed by metabolite and transcriptome profiling	60
Figure 13	PCA of untargeted metabolite profiling data	62
Figure 14	Targeted metabolite quantification	63
Figure 15	Global gene expression analysis	65
Figure 16	Transcript profiles of flavonoid/anthocyanin pathway genes	72
Figure 17	Transcript levels of genes associated to flavor formation and fruit softening	74
Figure 18	Transcriptome data of <i>FvGT2</i> and 5	75

- Figure 19 Transcript levels of *FvAT1, 2 and 3* 76
Figure 20 Protein sequence alignment of SCPL acyl-transferases AT2.1-2.4 78
Figure 21 Over-expression of *FvAT1* and *3* in fruits of *F. × ananassa* 80
Figure 22 Metabolite levels in *FvAT1*, *FvAT3* and *GUS* transfected fruits 81
Figure 23 Chemical structures of sinapic acid and gallic acid 95

LIST OF TABLES

Table 1	Gallate-UGT primers	21
Table 2	AT primers	22
Table 3	RT-PCR primers	23
Table 4	HPLC system details	24
Table 5	MS system details	25
Table 6	General laboratory equipment	26
Table 7	PCR reaction mixture	31
Table 8	Standard PCR protocol	31
Table 9	Touchdown PCR protocol	32
Table 10	Restriction enzyme reaction mixture	33
Table 11	50× TAE Buffer recipe	33
Table 12	10× Orange dye recipe	34
Table 13	Cultivation media	36
Table 14	SDS-PAGE buffers	36
Table 15	Western Blot buffers	38
Table 16	RT-PCR reaction mix	43
Table 17	RT-PCR cycling conditions	44
Table 18	Breaking buffer	45
Table 19	MMA medium	47
Table 20	UGT substrate screen	52
Table 21	Kinetic UGT values	54
Table 22	Gene candidates significantly down-regulated in the white genotypes	67
Table 23	Gene candidates significantly up-regulated in the white genotypes	69

ACRONYMS

AAT	alcohol acyltransferase
ABC	ATP-binding cassette
ANR	anthocyanidin reductase
ANS	anthocyanidin synthase
AP	alkaline phosphatase
Arg	arginine
Asp	aspartate
AT	1-O-acylglucose dependent acyltransferase
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bHLH	basic helix-loop-helix
4CL	4-coumaroyl-CoA ligase
CA4H	cinnamic acid 4-hydroxylase
CAD	cinnamyl alcohol dehydrogenase
CAZymes	carbohydrate-active enzymes
CCR	cinnamoyl CoA reductase
cDNA	complementary DNA
CHS	chalcone synthase
CHI	chalcone isomerase
CoA	coenzyme A
CTAB	cetyltrimethylammonium bromide
Cys	cysteine
DAD	diode array detector
dATP	deoxyadenosine triphosphate
DFR	dihydroflavonol reductase
DMF	dimethylformamide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dpa	days post anthesis
dsDNA	double stranded DNA
EA	ellagic acid
EDTA	ethylenediaminetetraacetic acid
e.g.	exempli gratia, for example

ET	ellagitannin
EtOH	ethanol
F ₃ H	flavonoid 3-hydroxylase
FEM	Fondazione Edmund Mach
FGT	flavonoid glucosyltransferase
FLS	flavonol synthase
fw	forward
GA	gallic acid
GDR	Genome Database for Rosaceae
Glu	glutamate
GST	glutathione S-transferase
GT ₁	anthocyanidin glucosyltransferase
GT ₂	(hydroxy)cinnamic acid and (hydroxy)benzoic acid glucosyltransferase
GUS	β-glucuronidase
HCl	hydrochloride
HHDP	3,4,5,3',4',5'-hexahydroxydiphenoyl
His	histidine
HPLC	high performance liquid chromatography
HW ₄	Hawaii 4
ICC	ion charge control
IgG	immunoglobulin
Ile	isoleucine
KCl	potassium chloride
LAR	leucoanthocyanidin reductase
LB	Luria-Bertani
LC-MS	liquid chromatography electro-spray ionization mass spectrometry
Leu	leucine
Lys	lysine
MeOH	methanol
Met	methionine
MMA	modified MacConkey agar
mRNA	messenger RNA
MRSA	methicillin-resistant <i>Staphylococcus aureus</i>
MS	mass spectrometry
MT	malonyl transferase

mQ	milli-Q, ultrapure, and deionized water
MYB	myeloblastosis
NaCl	sodium chloride
NBT	nitroblue tetrazolium
NCBI	National Center for Biotechnology Information
Ni	nickel
OD	optical density
OPR	12-oxophytodienoic acid reductase
Orange G	disodium 7-hydroxy-8-[(E)-phenyldiazenyl]-1,3-naphthalenedisulfonate
ORF	open reading frame
PAL	phenylalanine ammonia-lyase
PCA	principle component analysis
PCR	polymerase chain reaction
PGG	1,2,3,4,6-pentagalloylglucose
pH	potentia hydrogenii
Phe	phenylalanine
PMSF	phenylmethylsulfonyl fluoride
Pro	proline
PSPG	plant secondary product glycosyltransferase
PVDF	polyvinylidene fluoride
QR	quinone oxidoreductase
RdV	Reine des Vallees
rev	reverse
RNA	ribonucleic acid
RNA-seq	RNA-sequencing
RPM	reads per million
RT	room temperature
RT-PCR	real-time PCR
SCPL	serine carboxypeptidase-like
SC-U	synthetic complete minimal medium for yeast cultivation without uracil
SDH	shikimate dehydrogenase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	serine
SGR	Strawberry Genomic Resources

SMS	Sequence Manipulation Suite
SNP	single nucleotide polymorphism
SOC	super optimal broth with catabolite repression
TAE	TRIS-acetate-EDTA
TF	transcription factor
Thr	threonine
TRIS	2-amino-2-(hydroxymethyl)-1,3-propanediol
Trp	tryptophane
TUM	Technical University of Munich
Tyr	tyrosine
UDP	uridine-5'-diphosphate
UGT	UDP-glucose dependent glucosyltransferase
UV	ultra violet
Val	valine
YW	Yellow Wonder

INTRODUCTION

1.1 THE STRAWBERRY PLANT

1.1.1 *Origin*

There is evidence for cultivation of native strawberries, even in the ancient roman empire (Darrow, 1966). The Romans may have valued these fruits for their sweet flavor and their appealing, red-colored fruits. Today, a large number of species exist, with diploid ($2n=2x=14$) woodland strawberry *Fragaria vesca* being considered as the most widespread one (Hancock, 1999).

Even early Romans collected strawberries from the mountain side.

However, its fruits are considerably small when compared to the progeny of the garden strawberry *F. × ananassa*. Approximately 300 years ago, the domestication of *F. × ananassa* started in Brittany with an accidental hybridization event between *F. chiloensis* and *F. virginiana*, two species that originated from the American continent. Its origin was first described in 1766 by the French botanist Antoine Nicholas Duchesne, who named the plant after the pineapple (ananas) to acknowledge the distinctive shape and flavor of the fruits. The hybridization resulted in an octoploid ($2n=8x=56$) plant that grew considerably larger fruits, although yields and quality were still poor. The first propagation initiative was started in 1817 by Thomas A. Knight. He used clones of both ancestral varieties in his crossings and produced large, hard fruits (Darrow, 1966; Hancock, 1999; Pearl, 1928; Wilhelm and Sagen, 1974). Since then breeding has been promoted further, and today numerous varieties of garden strawberry exist (Horvath et al., 2011). In fact, with an average production of 7,739,622 t in 2013 it has become one of the most cultivated berries worldwide (FAOSTAT, 2013).

Along with other famous edible fruit bearing genera like *Malus* (apple), *Prunus* (e.g. plum, cherry), *Pyrus* (pear) and

Rubus (e.g. raspberry, blackberry), the genus *Fragaria* belongs to the family of *Rosaceae* in the order of *Rosales*. Fruits and seeds of this family can occur in different forms, for example drupes (*Prunus*) and in the case of strawberries accessory fruits, which consist of achenes (seeds) and receptacle (pulp) (Folta and Gardiner, 2009).

1.1.2 *Fruit ripening*

Plant hormones regulate the fruit development.

Strawberry is categorized as a non-climacteric fruit, as the ripening process is independent of an increased ethylene respiration and biosynthesis as such (Giovannoni, 2004). In contrast, a recent study reported a potential relationship between ethylene production and strawberry fruit ripening (Iannetta et al., 2006). Further clarification is necessary.

The ripening of garden strawberry fruits can be subdivided into five to seven stages: small green, big green, green-white, white, turning, ripe and over-ripe (Zhang et al., 2011). At the beginning, cells begin to divide and expand which leads to a progressive enlargement of the receptacle, accompanied by seed and embryo formation (Aharoni, 2002). This is mainly regulated by the plant hormone auxin whose level increases concomitant with cell division (Manning, 1998; Perkins-Veazie, 1995). Meanwhile, water and sugars accumulate. During the process of enlargement the level of chlorophylls decreases, and the fruits turn from green to white (Zhang et al., 2011). Subsequently, the level of auxin decreases and the level of abscisic acid increases, while accumulating anthocyanins change the color from white to red (Jia et al., 2011; Symons et al., 2012). Additionally, cellulases and pectin-methylesterases reduce the firmness of the fruit through de-esterification and decomposition of cell wall components (Civello, 1999; Rosli et al., 2004). Thus, fruit development accompanies the changing compositions of secondary metabolites.

1.1.3 Plant secondary metabolism

The study of secondary plant products was perhaps started by Friedrich Wilhelm Sertürner in 1806 (Hartmann, 2007). He was able to isolate an organic acid from the poppy plant. After precipitation, the newly formed white crystals had an anesthetizing effect much like opium. He called the substance "*principium somniferum*". His results were published in several journals, including Trommsdorff's *Journal der Pharmazie* (Sertürner, 1805, 1806) and *Annalen der Physik* (Sertürner, 1817). A few years later however, he called the substance morphine. In doing so, he started a new field of research called *Natural Product Chemistry*. Nowadays, as corresponding enzymes and genes are known (Kutchan et al., 1988; Zenk, 1991), and scientists begin to understand that biosynthetic pathways can interact (Burbulis and Winkel-Shirley, 1999), the term *Plant Secondary Metabolism* is also commonly used.

Unlike primary metabolism, secondary metabolism is less essential for plant growth and development. Nevertheless, it is important for the survival of a single or a group of individual plants. Because the products of secondary metabolism can be modified and adapted, plants can actively react to changing environmental conditions or even compensate for calamities and stress-factors (Hartmann, 2007). Strawberries (*F. × ananassa*), for example, accumulate anthocyanins and phenolics under long-term salt stress conditions. It was shown that plants with higher contents of these metabolites were less susceptible to salt stress (Keutgen and Pawelzik, 2008). However, the boundary between primary and secondary metabolism is not so easily determined as recent studies suggest. Junctions via specific metabolites exist (Fits and Memelink, 2000; Glawischnig et al., 2004).

Secondary products enable plants to adapt to their environment.

1.2 METABOLITE BIOSYNTHESIS DOWNSTREAM OF THE SHIKIMATE PATHWAY

Several secondary metabolites from strawberry (*F. vesca* and *F. × ananassa*) were analyzed within the scope of this thesis, all of which derive from the shikimate pathway (Figure 1).

Starting from 3-dehydroshikimic acid, the major downstream metabolic pathways are the phenylpropanoid, the lignin, the flavonoid/anthocyanin and the biosynthetic pathway of condensed and hydrolyzable tannins.

1.2.1 *The phenylpropanoid pathway*

Phenylalanine is the junction between primary and secondary metabolism.

The junction between primary and secondary metabolism in this case is phenylalanine, where the carbon flow is directed towards the production of the activated 4-coumaroyl-CoA thioester (Ferrer et al., 2008). This process consists of three steps: First, cinnamic acid is produced by deamination of phenylalanine via phenylalanine ammonia-lyase (PAL) (Liao et al., 1996). Then, a hydroxyl group is introduced in the *para*-position of the phenyl ring by cinnamic acid 4-hydroxylase (CA4H), whereupon 4-coumaric acid is formed (Guerra et al., 2013). Finally, 4-coumaroyl-CoA ligase (4CL) catalyzes the formation of a thioester bond between the carboxyl group of 4-coumaric acid and the thiol group of coenzyme A (CoA), thus producing the above-mentioned 4-coumaroyl-CoA (Ehlting et al., 1999). From these reaction steps, simple (e.g. hydroxycinnamic acids) and more complex phenylpropanoids such as lignins are derived (Dixon et al., 2002).

1.2.2 *The lignin pathway*

In planta lignin stabilizes the cell wall during secondary growth. It does so, by cross-linking cellulose, hemicellulose and pectin (Weng and Chapple, 2010). Lignin is formed by polymerization of 4-coumaryl alcohol, coniferyl alcohol and sinapyl al-

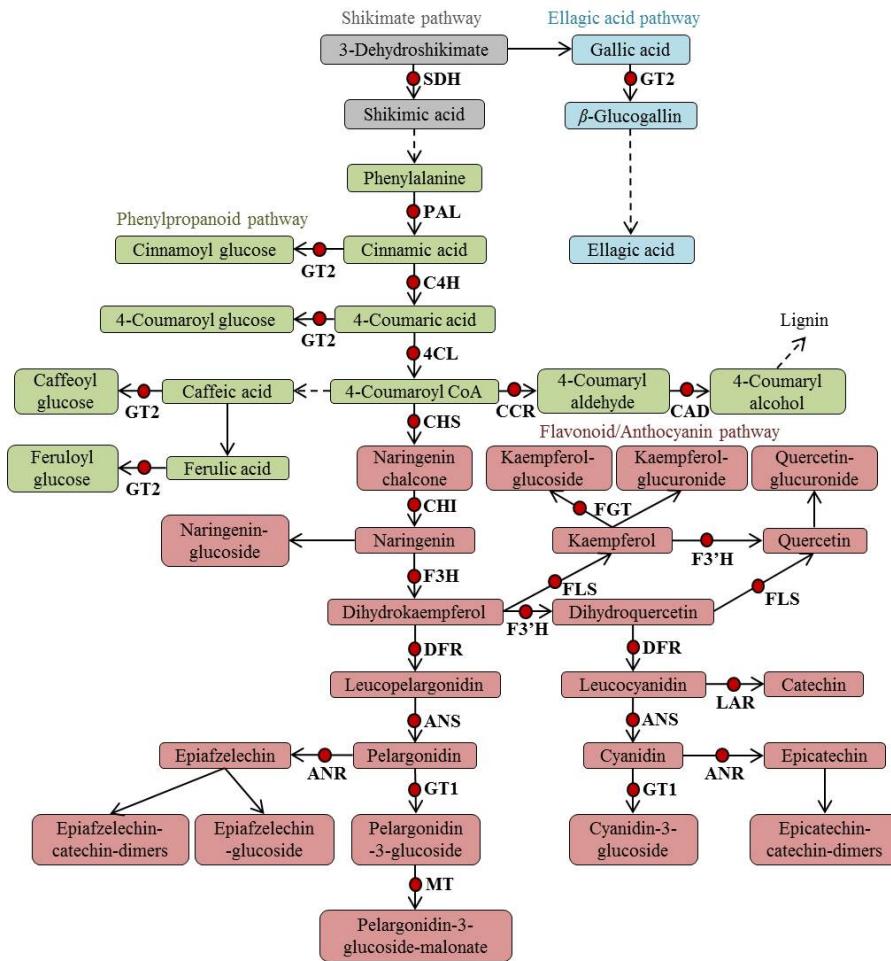


Figure 1: Schematic illustration of the shikimate, ellagic acid, phenylpropanoid, flavonoid and anthocyanin pathway (modified from Medina-Puche et al., 2014). Red dots indicate confirmed enzymatic mechanisms: ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; CA4H, cinnamic acid 4-hydroxylase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; CHI, chalcone isomerase; CHS, chalcone synthase; 4CL, 4-coumaroyl-CoA ligase; DFR, dihydroflavonol reductase; FGT, flavonoid glucosyltransferase; F3'H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; FLS, flavonol synthase; GT1, anthocyanidin glucosyltransferase; GT2, (hydroxy)cinnamic acid and (hydroxy)benzoic acid glucosyltransferase; LAR, leucoanthocyanidin reductase; SDH, shikimate dehydrogenase; MT, malonyl transferase; PAL, phenylalanine ammonia lyase.

cohol, which branch out from the phenylpropanoid pathway via reduction of 4-coumaroyl-CoA to coumaryl aldehyde, catalyzed by cinnamoyl CoA reductase (CCR). A subsequent second reduction step, catalyzed by cinnamyl alcohol dehydrogenase (CAD), leads to the formation of coumaryl alcohol (Brill et al., 1999; Larsen, 2004).

1.2.3 *The flavonoid/anthocyanin pathway*

4-Coumaroyl-CoA assumes a key position in the shikimate derived pathways.

4-Coumaroyl-CoA is the branching point of the shikimate derived pathways, which outlines its key position in the metabolic pathways presented here. From there, two enzymes are responsible for the formation of the basic C-15 flavonoid structure: chalcone synthase (CHS) and chalcone isomerase (CHI). In the strawberry, silencing of *CHS* leads to reduced levels of anthocyanins and a subsequent increase of simple phenylpropanoids such as hydroxycinnamic acids (Hoffmann et al., 2006).

The basic C-15 structure produced by CHS and CHI consists of three aromatic rings (A, B and C, Koes et al., 2005). The central B ring is oxidized by flavonoid 3-hydroxylase (F3H) and the resulting dihydrokaempferol, belonging to the class of flavanonols, is further processed. It is hydroxylated by flavonone 3'-hydroxylase (F3'H) to form dihydroquercetin, which is oxidized by flavonol synthase (FLS) to produce flavonols, such as kaempferol and quercetin (Forkmann, 1991). Simultaneously, dihydroquercetin is reduced by dihydroflavonol reductase (DFR) to leucoanthocyanidins, whereupon colored anthocyanins are produced by the action of anthocyanidin synthase (ANS) (e.g. pelargonidin, cyanidin and delphinidin, Abrahams et al., 2003; Wang et al., 2013). Pelargonidin-3-O-glucoside is the major anthocyanin in the strawberry, and is thereby contributing to the appealing red color of ripe fruits (Almeida et al., 2007). In being part of the human diet, anthocyanins are believed to possess a large number of health-promoting, physiological effects (Kong et al., 2003). They can, for example, scavenge free radicals and suppress the proliferation of cancer cells

(Hui et al., 2010); properties that they have in common with tannins.

1.2.4 Biosynthesis of tannins

The name "tannin" refers to the profession of "tanning". In former times extracts of oak barks and galls, a rich natural source of tannins, were used to turn raw animal hides into leather or to produce ink (Niemetz and Gross, 2005). Tannins belong to the class of polyphenols. According to Quideau et al., 2011, there are three classes of "true" plant polyphenols: phlorotannins (I), proanthocyanidins (II) and gallo-/ellagitannins (III) (Figure 2).

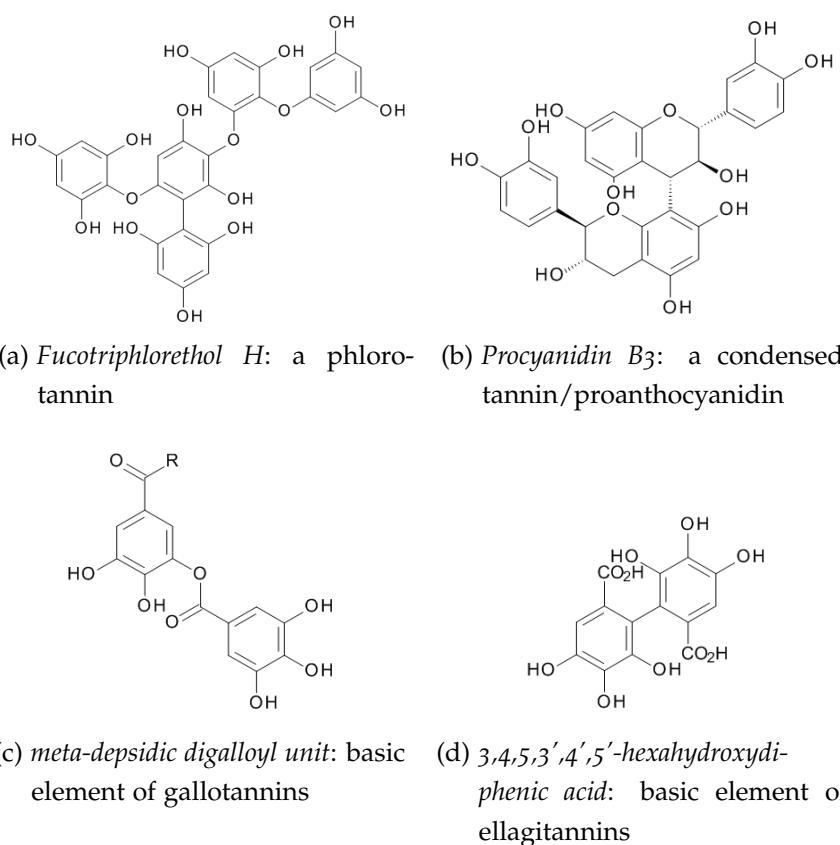


Figure 2: Representative examples and subunits of "true" plant polyphenols. (modified from Quideau et al., 2011 and Niemetz and Gross, 2005)

(I) PHLOROTANNINS derive from polymers of phloroglucinol and are mainly found in brown algae, where they are involved in defense mechanisms against bacteria and herbivores (Arnold and Targett, 1998). In contrast to proanthocyanidins, gallotannins and ellagitannins, they don't have tanning properties, but still can be counted as "true" polyphenols because of their structure and origin (fucotriphloethol H, Figure 2a). A recent study suggests that acyl derivatives of phloroglucinols are also natural metabolites of strawberry fruit (Song et al., 2015b). The authors hypothesize that CHS (Section 1.2.3), mainly responsible for the conversion of naringenin chalcone to naringenin, exhibits dual functionality in acting also as valerenone synthase.

(II) PROANTHOCYANIDINS are also called condensed tannins. They are mostly polymers of flavan-3-ol units such as (+)-catechin (*2,3-trans*) and (-)-epicatechin (*2,3-cis*) (Xie and Dixon, 2005). Catechin and epicatechin are formed by the action of leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR), respectively (Xie et al., 2003; Zhao et al., 2010). In strawberry, procyanidin dimer B₃ appears to be the most abundant condensed tannin (Figure 2b, Pascual-Teresa et al., 2000).

Gallic acid is the basic phenolic unit of hydrolyzable tannins.

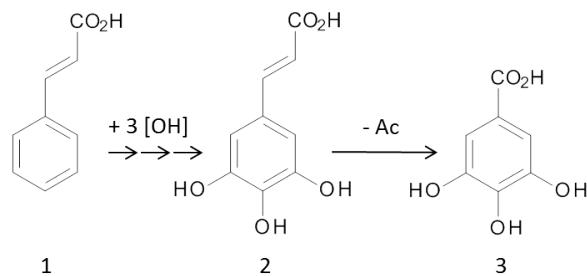
(III) GALLO-/ ELLAGITANNINS , the so called hydrolyzable tannins, are composed of esters of simple sugar molecules (β -D-glucose, mostly) and gallic acid (GA) (Haslam and Cai, 1994). However, the biosynthesis of GA has been a subject of debate since decades. In total, there are three proposed routes (Figure 3). The first hypothesis presupposes β -oxidation of 3,4,5-trihydroxycinnamic acid (Figure 3a, Zenk, 1964). Albeit, 3,4,5-trihydroxycinnamic acid has never been identified in Nature in contrast to its O-methyl ether sinapic acid, which is an intermediate of lignin biosynthesis (Section 1.2.2, Haslam, 1989). The second route assumes that instead of "the missing cinnamic acid" (Quideau, 2009; p. 99), caffeic acid is degraded to protocatechuic acid, and subsequently hydroxylated to GA (Figure 3b, El-Basyouni et al., 1964). But evidence for involvement

of caffeic acid was never found. Therefore, the third approach is the most convincing one. It was also put forward in the 1960s (Conn and Swain, 1961; Cornthwaite and Haslam, 1965), and supports oxidation of the *enol*-form of 3-dehydroshikimic acid (Figure 3c). More recent publications, including enzyme studies of *Betula pubescens* (downy birch), contributed evidence to support this concept (Ossipov et al., 2003; Werner et al., 1997; Werner et al., 2004). Final *in vitro* and *in vivo* evidence was brought in grapevine (*Vitis vinifera*) that gallic acid can be produced from 3-dehydroshikimic acid by the action of shikimate dehydrogenase (SDH) (Bontpart et al., 2016).

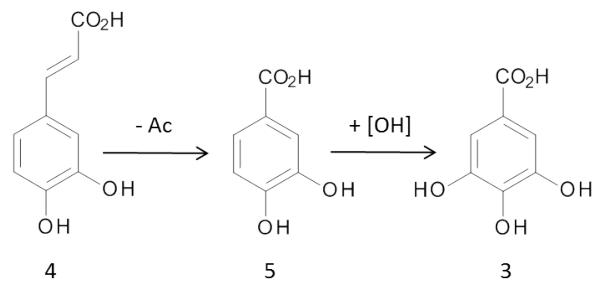
The biosynthesis of gallo-/ellagitannins continues by formation of β -glucogallin (galloyl glucose, 1-O-galloyl- β -D-glucopyranose) through esterification of GA and uridine-5'-diphosphate glucose (UDP-glucose) (Niemetz and Gross, 2001). This reaction is catalyzed by inverting UDP-glucose dependent glucosyltransferases (UGTs) as enzyme studies from *Quercus robur* (pedunculate oak) demonstrated (Gross, 1982; Weisemann et al., 1988). Inverting UGTs facilitate a direct displacement S_N2 -like reaction, in which the stereochemistry of the anomeric reaction center of the donor substrate is inverted, as exemplarily shown for the formation of β -glucogallin (Figure 4). Today, genes coding for 1-O-acylglucose glucosyltransferases have been identified in *V. vinifera* (Khater et al., 2012), *Q. robur* (Mittasch et al., 2014), *F. × ananassa*, *F. vesca* and *Rubus idaeus* - raspberry (Schulenburg et al., 2016a). Clearly, this emphasizes the general importance of this mechanism in the plant kingdom. β -Glucogallin gives rise to 1,2,3,4,6-pentagalloylglucose (PGG) by position-specific consecutive trans-esterification reactions presumably catalyzed by 1-O-acylglucose dependent acyltransferases (ATs) (Figure 5, Cammann et al., 1989; Hagenah and Gross, 1993). It thereby assumes a pivotal role in acting as an acyl-donor and an acyl-acceptor, simultaneously.

PGG is the precursor of gallotannins and ellagitannins, respectively. It is here, where the biosynthetic pathways split. Gallotannins are formed by continuing additions of galloyl units forming *meta*-depside groups (Figure 2c). For

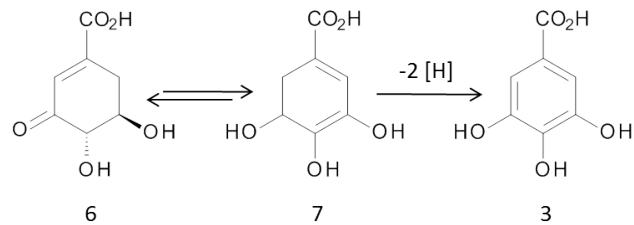
*Downstream of
penta-
galloylglucose, the
biosynthetic
pathways split.*



(a) *Route one:* conversion of 3,4,5-trihydroxy-cinnamic acid (2) to gallic acid (3) starting from cinnamic acid (1)



(b) *Route two:* degradation of caffeic acid (4) to protocatechuic acid (5) and adjacent hydroxylation to gallic acid (3)



(c) *Plausible route three:* oxidation of the *enolform* (7) of 3-dehydroshikimic acid (6) to yield gallic acid

Figure 3: Possible biosynthesis pathways of gallic acid. (modified from Quideau, 2009)

the biosynthesis of ellagitannins (ETs), oxidation reactions are suggested that presumably yield 3,4,5,3',4',5'-hexahydroxydiphenoyl (HHDP) moieties (Niemetz and Gross, 2005). HHDP units consist of linked galloyl residues (Figure 2d).

Ellagic acid (EA), the dilactone of HHDP, is released from ETs by hydrolysis (Niehaus and Gross, 1997).

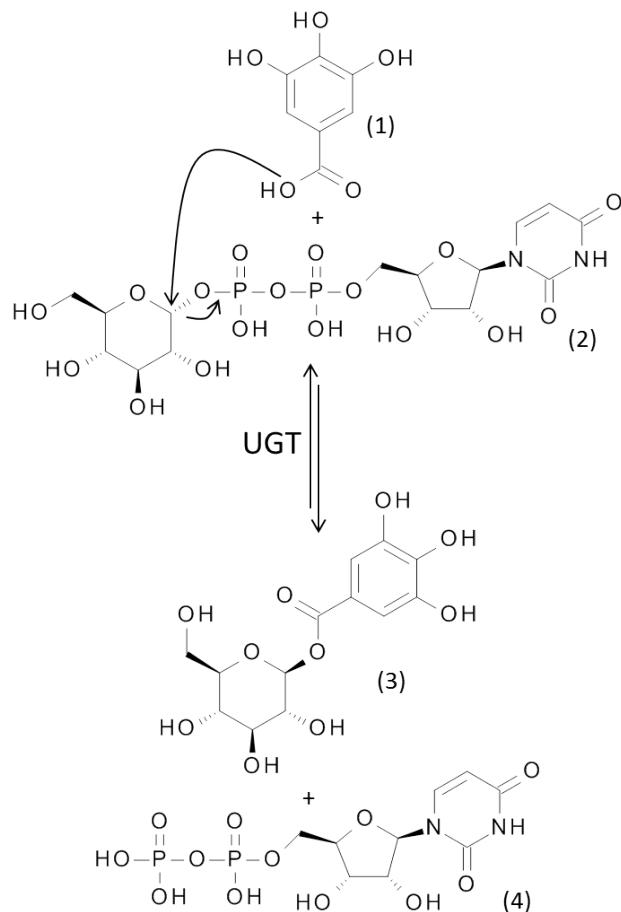


Figure 4: UGT conveyed direct displacement S_N2 -like reaction of the acceptor substrate GA (1) and the donor substrate UDP-glucose (2) to yield the product β -glucogallin (3) and the leaving group UDP (4).

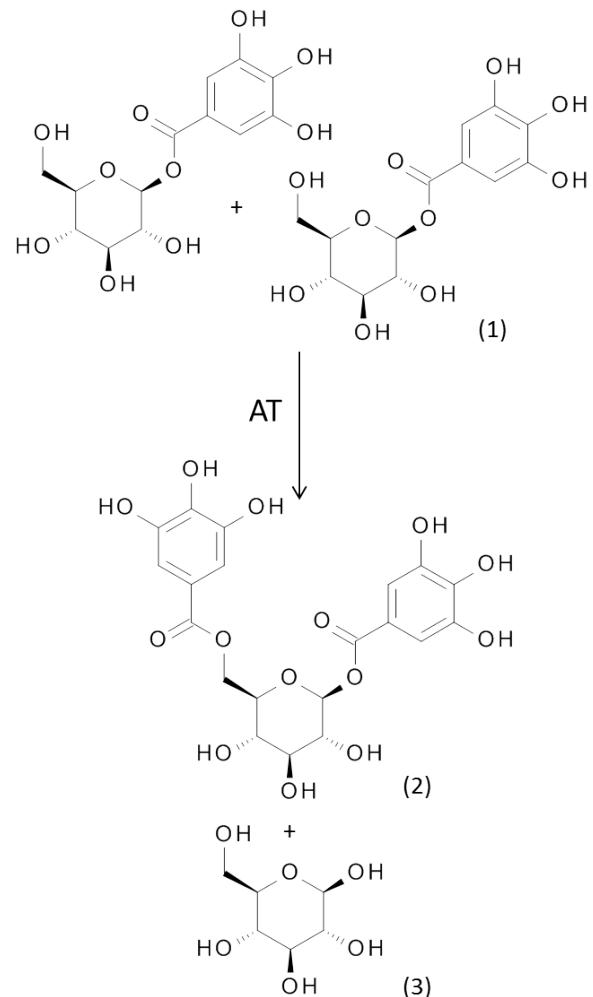


Figure 5: AT conveyed formation of 1,6-di-galloyl glucose (2) from two molecules of β -glucogallin (1), whereupon glucose (3) is released.

1.3 FROM ELLAGITANNINS TO ELLAGIC ACID

Often, a plant can be characterized according to its ET composition (Quideau, 2009). Usually, one ET stands out as the main component, a feature that seems to have evolved through evolution (Okuda et al., 2000). Geraniin, for example, is the principal ET in *Geranium thunbergii* (10 % of the dry weight; Okuda et al., 1980). It was among the first to be studied by X-ray crystallography, thus allowing a detailed structural analysis (Luger et al., 1998). The discovery of Geraniin paved the way for elucidation of numerous other plant ETs. In the woodland as well as in the garden strawberry agrimoniin was identified as the major ET, although it was at first confused with its structural relative sanguinin H-6, caused by matching LC-MS fragmentation patterns (Vrhovsek et al., 2012). It may therefore be assumed that in the genus *Fragaria* EA could well be derived from agrimoniin.

Plants show a specific ellagitannin composition.

1.3.1 Biosynthesis of Ellagic Acid

In *Tellima grandiflora* (fringe cups), PGG is oxidized to the monomeric ET tellimagrandin II by a laccase-type phenol oxidase (Niemetz and Gross, 2001) that introduces a C-C bond between two aromatic rings of the galloyl residues (biaryl bond). In a second step, tellimagrandin II is oxidized further to cornusiin E, a dimeric ET (Niemetz and Gross, 2003). Agrimoniin is also a dimeric ET. Its biosynthesis could therefore follow the same mechanism (Figure 6). It is believed that HHDP is then formed by complete hydrolysis of the ester bonds, whereupon EA would arise by subsequent lactonization (Quideau, 2009).

1.3.2 Physicochemical Properties

The more galloyl groups are added to the glucose core of hydrolyzable tannins, the higher the hydrophobicity. Of all the precursor molecules, PGG showed the lowest ability to dissolve in water (Tanaka et al., 1997b). Interestingly, the solubility rises

Formation of C-C bonds between galloyl groups increases water solubility.

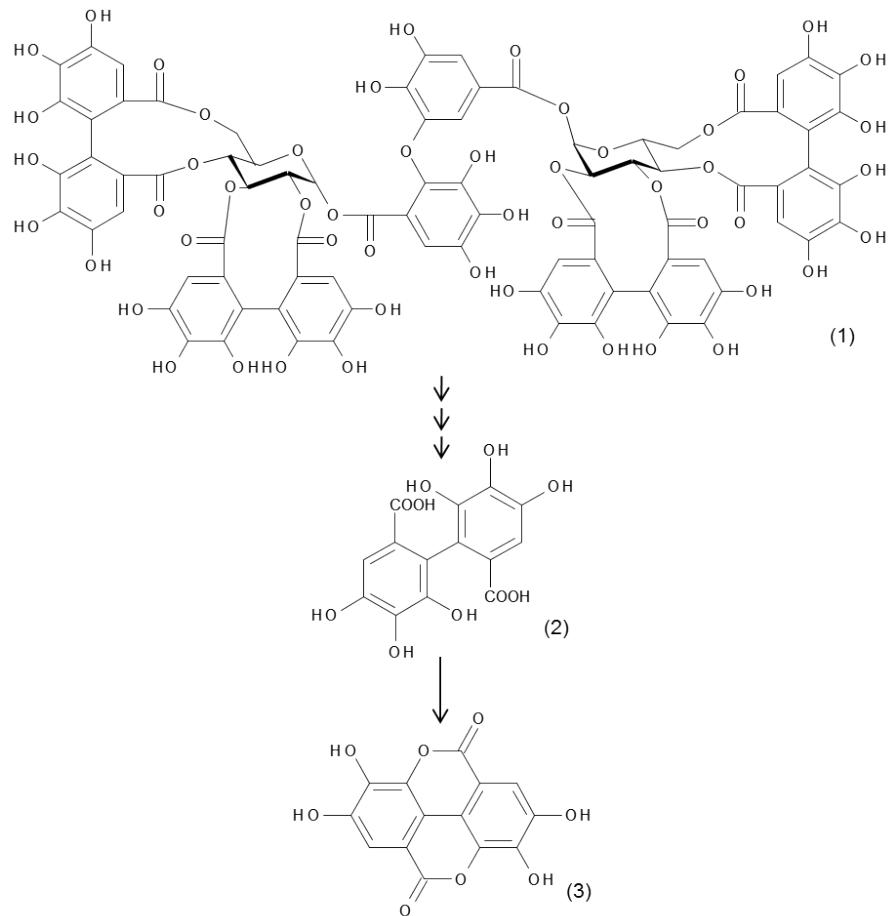


Figure 6: Possible pathway of EA biosynthesis in strawberry. By hydrolysis of ester bonds, agrimoniin (1) is converted to HHDP (2). Upon lactonization, EA (3) is formed. (modified from Quideau, 2009 and Niemetz and Gross, 2005)

again after coupling of galloyl units via biaryl bonds (Nonaka et al., 1980). Thus, sanguin H-6, the structural relative of agrimoniin (Section 1.3), was found to be much more hydrophilic than PGG (Tanaka, 1985). In aqueous media, monomeric and dimeric ETs, as well as PGG can associate with themselves or with co-existing compounds (Tanaka et al., 1997a,b). The molecules are thought to connect via hydrophobic and hydrogen bonding (Cai et al., 1990).

Due to both, hydrophobic association and formation of biaryl bonds, the water solubility increases. It might be for this reason that plants like strawberry can accumulate ETs in higher concentrations. They are enriched for defense against herbivores, because their associative ability allows to precipitate proteins from saliva (Charlton et al., 2002a,b). This denaturation effect creates an unpleasant astringent taste, and can also cause digestive problems (Edelmann and Lendl, 2002; Mehansho et al., 1987). Not only plant predators are put off by the physicochemical properties of ETs, but also microorganisms.

1.3.3 Antimicrobial activities

Various combinations of monomeric and dimeric hydrolyzable tannins have been tested on bacteria, fungi and parasites. It seems that whether or not the ETs show an effect depends on the kind of microorganism they are applied to. For example, against *Pseudomonas aeruginosa*, *Escherichia coli* and *Bacillus subtilis* only few antibacterial properties could be observed (Kolodziej et al., 2000).

On the contrary, when exposed to tellimagrandin I the propagation of *Helicobacter pylori* was suppressed (Funatogawa et al., 2004). *H. pylori* can cause gastritis and ulcers in the human intestinal tract (Ernst and Gold, 2000). If its spreading were to be inhibited without affecting ordinary enterobacteria like *E. coli*, ETs could be an effective and mild remedy. Some ETs showed a distinct effect against methicillin-resistant *Staphylococcus aureus* (MRSA) strains (Shiota et al., 2004). They not only lowered

the production rate, but also inhibited the bacterial penicillin-binding protein 2a, which is suggested to be responsible for the resistance of MRSA to antibiotics. It was also shown that hydrolyzable tannins deplete the parasite *Leishmania donovani* (Kolodziej et al., 2001).

1.3.4 Health benefits

Ellagitannins
possess
anticarcinogenic,
antimicrobial and
antioxidative
properties.

When we talk about the possible health benefits of ETs/EA, we have to take into account how relevant for the human diet they really are. In fact, there are a number of foods available that show a remarkably high concentration. Pomegranate juice, for instance, can contain up to 1900 mg L⁻¹ (300 mg per serving) of ETs, which makes up for an antioxidant activity that is twice as high as in red wine (Gil et al., 2000). Walnuts are also a rich source, providing 802 mg of ETs per 50 g of tissue (Anderson et al., 2001). Of the ET-rich berries, raspberry, cloudberry and strawberry show the highest concentrations decreasing in aforementioned order (Koponen et al., 2007). But regarding food popularity, strawberries outdistance them all (Section 1.1.1), and are therefore a most important source of ETs/EA.

Numerous studies report on the health-promoting effects of ETs/EA upon oral application of ET-rich food to test subjects. In animal models (rats and mice, mostly), they showed inhibiting effects on breast, prostate and colon cancer (Giampieri et al., 2012; Harris et al., 2001; Malik et al., 2005). In human patients, administration of pomegranate juice led to a reduced effect of risk factors linked to cardiovascular and neurodegenerative disorders, and ETs are thought to be of significance for long-term health protection (Fuhrmann and Aviram, 2006; Landete, 2011). Furthermore, like most polyphenols, they possess a strong antioxidant activity. This enables them to scavenge free radicals and to prevent oxidation processes (Quideau et al., 2011). All these authors attribute the health-promoting effects to the ETs/EA present in the administered food. Albeit, specific

compounds that could be responsible for the positive impacts are never mentioned.

This raises, of course, the question about the bioavailability of ETs/EA *in vivo* or rather if, to what extent, and where they are metabolized. Generally, they are hydrophobic compounds that dissolve poorly in aqueous media (Section 1.3.2). In the gut, ETs release free EA, of which only a small rate is absorbed in the stomach and the small intestine (Espín et al., 2007). There, it is further disintegrated to urolithins (dibenzopyranones), by bacteria native to the gut (Selma et al., 2014). From the small intestine they are probably passed on to the liver, and subsequently, excreted via bile and urine (Ito et al., 2008). An accumulation in organ tissues is not reported. However, the amount of EA absorbed from food depends strongly on the constitution of single individuals and differs greatly from one patient to another (Quideau, 2009). But the connection between an ET-rich diet and promoted health has been confirmed repeatedly. Thus, more research is necessary to improve the knowledge about digestion and metabolism of ETs/EA.

1.4 AIMS OF THE THESIS

While the importance of ETs and EA is widely recognized, little is known about the biosynthetic pathway. Therefore, the primary goal was the identification of genes and enzymes participating in the biosynthesis of PGG, the proposed precursor of ETs. Strawberry, which contains high amounts of bio-active polyphenols, in particular ETs/EA, was employed as a plant system. In particular, this thesis concentrates on cloning, expression and characterization of putative UGTs catalyzing the formation of β -glucogallin. Furthermore, transcriptome and metabolite profiling was performed, in order to select AT genes presumably conveying the formation of PGG. The candidate ATs were then cloned, and their activity was analyzed *in vitro*, and *in vivo*.

2

MATERIAL

2.1 PLANT MATERIAL

For this study, *F. vesca* cv. Reine des Vallees (RdV), Yellow Wonder (YW) and Hawaii 4 (HW4), *F. × ananassa* cv. Calypso and Elsanta, and stable transgenic *FaGT2*-silenced (Calypso background) strawberry plants were grown at the Call Unit for plant research of TUM School of Life Sciences in Freising, Germany. Since their generation by Lunkenbein et al., 2006a in our laboratory, the transgenic lines were rejuvenated regularly by collecting clones from stolons. Fruits were collected from April until August 2013, 2014 and 2015, freeze-dried directly after harvest and stored at -20°C until further processed. *R. idaeus* cv. Tu-lameen and *F. vesca* (unknown accession) were grown on the Campus of Fondazione Edmund Mach (FEM) in Vigalzano di Pergine, Italy. Fruits were collected by Antje Feller¹ in summer 2013 (*F. vesca*) and 2014 (*R. idaeus*), freeze-dried directly after harvest and stored at -80°C until further processed. Moreover, blueberries (presumably *Vaccinium myrtillus*) were bought at a local supermarket and employed under equal conditions.

2.2 CHEMICALS

All chemicals were obtained either from Roth (Karlsruhe, Germany), Sigma-Aldrich (Steinheim, Germany) or Fluka (Steinheim, Germany) except when otherwise stated.

¹ Department of Food Quality and Nutrition, IASMA Research and Innovation Center, FEM, Via E. Mach 1, 38010 San Michele all'Adige, TN, Italy

2.3 PROKARYOTIC AND EUKARYOTIC MICROORGANISMS

The following organisms were used for cloning and expression of candidate genes:

- ❖ *Escherichia coli* XL1-Blue (Agilent Technologies, Waldbronn, Germany)
Genotype *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^qZΔM15 Tn10 (Tet^r)*]
- ❖ *E. coli* NEB 10-beta (New England Biolabs, Frankfurt, Germany)
Genotype *Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 gale15 e14- φ8odlacZΔM15 recA1 relA1 endA1 nupG rpsL (Str^R) rph spoT1 Δ(mrr-hsdRMS-mcrBC)*
- ❖ *E. coli* JM109 (Promega, Mannheim, Germany)
Genotype *endA1 recA1 gyrA96 thi hsdR17 (rk- mk⁺) relA1 supE44 Δ(lac-proAB)* [F' *traD36 proAB laqI^qZΔM15*]
- ❖ *E. coli* BL21 (DE3) pLysS (Promega, Mannheim, Germany)
Genotype *fhuA2 [lon] ompT gal [dcm] ΔhsdS*
- ❖ *Saccharomyces cerevisiae* INVSc1 (Invitrogen, Karlsruhe, Germany)
Genotype *MATa his3D1 leu2 trp1-289 ura3-52 MAT his3D1 leu2 trp1-289 ura3-52*
- ❖ *Agrobacterium tumefaciens* AGLo (Lazo et al., 1991)
Genotype *EHA101 pTiBo542ΔT-region Mop⁺*

2.4 PRIMERS

All primers were ordered from Eurofins MWG Synthesis GmbH (Ebersberg, Germany). Table 1 contains the primers for cloning and mutagenesis of the gallate-UGTs and Table 2 the primers for cloning of the ATs. Table 3 holds all real-time PCR (RT-PCR) primers.

Table 1: Primers used for cloning and mutagenesis of gallate-UGTs in the pGEX4T-1 vector system

GENE	DESCRIPTION	SEQUENCE
<i>FaGT2</i> *	BamHI fw	CGCGGATCCATGGGTTCCGAATCATTGGTT
	NotI rev	ATAGTTAGCGGCCGCTTACGACTCGACTAGTTCAAC
	R230S fw	GAACTCGAGAGCGAGATCATCGAGTACATGGCTCGTTG
	R230S rev	GATGATCTCGCTCTCGAGTTCTGGAAAGTGTCCATCAA
	E420D_I422V fw	ATCCCTAGGGATGAGGTAGAGAAGTGCTGCTGGAGGCG
	E420D_I422V rev	GCACTTCTCTACCTCATCCCTAGGGATCACCTGTCTC
<i>FvGT2</i>	BamHI fw	TAAGCAGGATCCATGGGTTCCGAATCATTGGTT
	NotI rev	CTTAGCGGCCGCTTACGATTGACTAGTTCAACC
<i>RiGT2</i>	BamHI fw	TAAGCAGGATCCATGGGTTCCGAATCATTGGTC
	EcoRI rev	TGCTTAGAATTCTCAAATAACCAGTCAACCTTCC
<i>FaGT5</i>	EcoRI fw	AAAGAATTCATGGGTTCTGTGGGATCTG
	XhoI rev	AAACTCGAGAGCAATAGAACGTCTAGCAATCTC

Table 2: Primers used for cloning of ATs

GENE	VECTOR	DESCRIPTION	SEQUENCE
<i>FvAT1</i>	pGEM-T	fw	ATGGGTTCGGAAGCAAGCTCAGTATGC
	pGEM-T	rev	TCACCCCTGCGACTTGACCATCAACG
	pYES2	BamHI fw	CGGGATCCAACACAATGTCTTCGGAAGCAAGCTCA
	pYES2	XhoI rev	CCGCTCGAGTCACCCTGCGACTTGACCATC
	pYES2	XhoI His-Tag rev	CCGCTCGAGTCATGATGATGATGATGCCCTGCGACTTG
	pBI12135S2x	BamHI fw	CGCGGATCCATGGGTTCGGAAGCAAGCTCA
	pBI12135S2x	Ecl136II rev	TTCGAGCTCTCACCCCTGCGACTTGACCATC
<i>FvAT3</i>	pGEM-T	UTR fw	GCCGTACCAAGATTGATTTCATTAG
	pGEM-T	UTR rev	AGGTGCCAACCTTTTCCCTCATAA
	pYES2	BamHI fw	AAAGAATTATGGTTCTGTGGGATCTG
	pYES2	XhoI rev	CCGCTCGAGTCACAGAGGATAATAAGCCAA
	pYES2	XhoI His-Tag rev	CCGCTCGAGTCATGATGATGATGATGCAGAGGATAATA
	pBI12135S2x	BamHI fw	CGCGGATCCATGGCGGGTCTGTGTTA
	pBI12135S2x	Ecl136II rev	TTCGAGCTCTCACAGAGGATAATAAGCCAACCA

Table 3: Primers used for RT-PCR of candidate genes

GENE	ORIENTATION	SEQUENCE
<i>FvAT1</i>	fw	CCGGTGATTGCCCTCAAC
	rev	CTTCAACGGGCCTATTTGGT
<i>FvAT3</i>	fw	CCCATGAAGCTCCAGGGATA
	rev	GCAACAAGTCCAATGTCGTCC
18-26S spacer	fw	ACCGTTGATTGCACAATTGGTCATCG
	rev	TACTGCGGGTCGGCAATCGGACG

2.5 EQUIPMENT

2.5.1 *Liquid chromatography ultraviolet electro-spray ionization mass spectrometry (LC-UV-ESI-MS_n, LC-MS)*

Relevant secondary metabolites were analyzed and quantified by LC-MS. System details are provided in Table 4 and Table 5.

2.5.2 *General Equipment*

General laboratory equipment is alphabetically listed in Table 6.

Table 4: HPLC system used for analysis and quantification of secondary metabolites

COMPONENT	DETAILS
HPLC	Agilent 1100 Series (Agilent Technologies Inc., Santa Clara, US-CA)
Pump	Quaternary Pump G1311A (Agilent)
Injector	Sample Injector G1313A (Agilent)
Column 1	Security Guard Cartridges C18 4 × 2 mm (Phenomenex, Aschaffenburg, Germany)
Column 2	Luna 3 µm C18 (2) 100 Å 150 × 2.0 mm (Phenomenex)
Column temperature	25 °C
Mobile phase	A: water + 0.1 % formic acid B: methanol + 0.1 % formic acid
Flow rate	0.2 mL min ⁻¹
Gradient	0-50 % B in 30 min, 50-100 % B in 5 min, 100 % B for 15 min, 100-0 % B in 5 min, 0 % B for 10 min
DAD	Wavelength Detector G1314A (Agilent) set to 280 nm

Table 5: MS system used for analysis and quantification of secondary metabolites

COMPONENT	DETAILS
MS	Bruker Daltonics Esquire 3000 ^{plus} Ion Trap (Bruker, Bremen, Germany)
Spraygas	30 p.s.i. nitrogen
Drygas	330 °C, 9 L min ⁻¹ nitrogen
Resolution	13,000 <i>m/z</i> per sec
Scan range	100-800 <i>m/z</i>
Polarity	alternating positive/negative
Ion accumulation	until ICC target achieved 20,000 (positive), or 10,000 (negative), or the max. time of 200 msec was reached
Capillary voltage	± 4,000 V
End plate voltage	± 500 V
Skimmer	40 V
Capillary exit	121 V
MS/MS	auto-tandem MS ²
Collision gas	4 × 10 ⁻¹ mbar helium
Collision voltage	1 V

Table 6: General laboratory equipment

UNIT	DEVICE
Agarose gel electrophoresis	MIDI (Carl Roth GmbH & Co. KG, Karsruhe, Germany) MAXI (neoLab, Heidelberg, Germany)
Autoclave	Systec V 95 (Systec, Wettenberg, Germany)
Balances	SCALTEC SPB61 (SCALTEC Instruments GmbH, Heiligenstadt, Germany) TP 214 (Denver Instrument, Bohemia, US-NY) Scout Pro SPU4001 (Ophaus, Pine Brook, US-NJ)
Blotting chamber	Semy Dry Blotter (Biostep GmbH, Jahnsdorf, Germany)
Centrifuges	Sigma K415, Sigma 1-14, Sigma 2K15 (Sigma, Osterode am Harz, Germany) Eppendorf 5415R, MiniSpin (Eppendorf AG, Hamburg, Germany)
Clean bench	Hera Safe (Heraeus Holding GmbH, Hanau, Germany)
Freeze dryer	Savant ModulyoD (Thermo Fisher Scientific Inc., Waltham, US-MA)
Incubator	Forma Steri-Cycle CO ₂ (Thermo Fisher Scientific Inc., Waltham, US-MA)
PCR	Primus 96 advanced (Peqlab Biotechnologie, Erlangen, Germany) SensoQuest labcycler (SensoQuest GmbH, Göttingen, Germany) Step One Plus incl. Step One 2.1 Software (Applied Biosystems, Foster City, US-CA)
pH-Meter	CG 820 (Schott Geräte GmbH, Mainz, Germany)
Photometer	Nicolet evolution 100 (Thermo Fisher Scientific Inc., Waltham, US-MA)
RT-PCR System	StepOnePlus™ (Applied Biosystems®, Foster City, US-CA)
Scintillation counter	Tri-Carb2800TR (PerkinElmer Life and Analytical Sciences, Shelton, US-CT)

Table 6: Equipment (continued).

UNIT	DEVICE
SDS-PAGE	MINI-Vertical Electrophoresis Unit (Carl Roth GmbH & Co. KG, Karlsruhe, Germany)
Shakers	Orbital Shaker Polymax 1040 (Heidolph Instrument GmbH, Schwabach, Germany) Shaker DOS-10L (neoLab, Heidelberg, Germany) Labinco Rotary Mixer (Labinco BV, Breda, Netherlands)
Spectrophotometer	Nanodrop 1000 (Peqlab Biotechnologie, Erlangen, Germany)
Thermoblock	Thermomixer comfort (Eppendorf AG, Hamburg, Germany)
Ultrasonic probe	Bandelin UW2070 / HD2070 (Bandelin Electronic GmbH & Co. KG, Berlin, Germany)
UV-Transilluminator	G:Box incl. GeneSnap 6.07.04 Software (Syngene, Cambridge, England)
Vacuum concentrator	Christ RVC 2-18 (Christ, Osterode am Harz, Germany)
Voltage generators	Consort E835, CV245 (Consort nv, Turnhout, Belgium) BioRad PowerPac 200 (BioRad; Hercules, US-CA)
Vortex mixer	Vortex-Genie® 2 (Scientific Industries Inc., Bohemia, US-NY)
Water bath	Julabo HC5 / 7 (JULABO GmbH, Seelbach, Germany) Sonorex RK103H (Bandelin Electronic GmbH Co. KG, Berlin, Germany)
Water polisher	Purelab® Classic UVF MK2 (Elga Labwater, Celle, Germany)

2.6 SOFTWARE AND INTERNET RESOURCES

The following software was used:

- ❖ Accelrys Draw 4.1 free version (<http://accelrys.com/>)
- ❖ Citavi 4 (<http://www.citavi.de/de/index.html/>)
- ❖ DataAnalysis 6.2, QuantAnalysis 6.2 (Bruker, Bremen, Germany)
- ❖ GENEIOUS Pro 5.5.6 (<http://www.geneious.com/>)
- ❖ MiKTEX 2.9 (<http://www.miktex.org/>)
- ❖ R 3.2.2 (R Core Team, 2015)
- ❖ R Studio (<http://www.rstudio.com/>)
- ❖ StepOneTM Software 2.1 (Applied Biosystems, Foster City, US-CA)
- ❖ TeXnicCenter (<http://www.texniccenter.org/>)

The following internet resources were used:

- ❖ Genome Database for Rosaceae (GDR)
(<http://www.rosaceae.org/>)
- ❖ OligoAnalyzer 3.1
(<http://www.eu.idtdna.com/calc/analyzer/>)
- ❖ Phenol-Explorer (<http://phenol-explorer.eu/>)
- ❖ Primer-BLAST
(<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>)
- ❖ Strawberry Genomic Resources (SGR)
(<http://www.bioinformatics.townson.edu/strawberry/>)
- ❖ Sequence Manipulation Suite (SMS)
(http://www.bioinformatics.org/sms2/rev_comp.html/)
- ❖ National Center for Biotechnology Information (NCBI)
(<http://www.ncbi.nlm.nih.gov/>)

3

METHODS

Some methods used in this thesis are part of articles that have been already published in peer-reviewed journals. In such cases as these, descriptions are held short, and due references are made in the text.

3.1 BASIC TECHNIQUES

3.1.1 *Extraction of secondary metabolites from strawberry and blueberry tissues and quantification by LC-MS*

Sample extraction and LC-MS analysis were carried out according to Ring et al., 2013. In short, fruits were frozen separately in liquid nitrogen and lyophilized to complete dryness before, in the case of strawberry, the achenes were separated from the pulp, and each sample was ground individually to fine powder by mortar and pestle. Blueberries were processed as whole fruits. Secondary metabolites were extracted from 3-5 replicates of 50 mg fruit powder each by addition of methanol supplemented with 0.2 mg mL⁻¹ of internal standard (Biochanin A and 4-Methylumbelliferyl-β-D-glucuronide). After extraction, the methanol was removed in a vacuum concentrator and the residue re-dissolved in water. Samples were run on a methanol-water gradient (Table 4), and the metabolites were identified according to their mass spectra and retention time, quantified and expressed as per mil equivalents of the dry weight (% equ. dw).

3.1.2 RNA-Isolation from strawberry fruits

The fruits were frozen in liquid nitrogen and grinded individually by mortar and pestle. Total RNA was isolated from the fine powder according to the CTAB protocol (Liao et al., 2004). DNA was removed by digestion with DNase I (Thermo Fisher Scientific Inc., Waltham, US-MA) according to the manufacturers instructions. Yields and integrity was determined by spectrophotometry, agarose gel electrophoresis or chip-based capillary electrophoresis (Bioanalyzer 2100, Agilent Technologies Inc., Santa Clara, US-CA). Chip-based capillary electrophoresis was performed under the direction of Dr. Melanie Spornraft at the Department of Animal Physiology and Immunology¹.

3.1.3 cDNA synthesis

A total of one µg of RNA was used as input per cDNA synthesis reaction. Either Oligo(dT)₂₀ primers (50 µM) or random hexamers (50 ng/µL) or an equimolar mixture of both were used for annealing. Reverse transcription was carried out with the SuperScript® III First-Strand Synthesis System for RT-PCR (Invitrogen™, Thermo Fisher Scientific Inc., Waltham, US-MA), while following the producers specifications.

3.1.4 Polymerase chain reaction (PCR)

The generally applied reaction mixture is depicted in Table 7. Two polymerase chain reaction (PCR) programs were used and employed for different purposes. Table 8 illustrates the general method used for gene amplification, which was also utilized to determine the presence of insert DNA in plasmid clones. For colony PCR, individual transformants were directly added to the reaction mixtures and used as DNA templates. If genes of interest were difficult to amplify, a touchdown cycling program would be applied, to increase specificity and sensitivity

¹ Technical University of Munich, Weihenstephaner Berg 3, 85354 Freising

according to Korbie and Mattick, 2008. This was achieved by a gradual reduction of the annealing temperature from 65.0 to 60.5 °C in 0.5 °C steps and additional 25 cycles at 58 °C annealing temperature (Table 9). Gene-specific primers (Table 1 and Table 2) and Phusion® DNA polymerase were used in all reactions except when testing plasmid clones, where *Taq* polymerase was employed (both New England Biolabs Inc., Ipswich US-MA).

Table 7: Generally employed PCR reaction mixture

COMPONENT	FINAL CONC.	FINAL VOL.
DNA template	50 ng	variable
Primer fw (10 pmol μL^{-1})	0.5 μM	1 μL
Primer rev (10 pmol μL^{-1})	0.5 μM	1 μL
dNTPs (10 mM)	500 μM	1 μL
5× Phusion® GC buffer	1x	4 μL
DNA polymerase	1 unit	0.5 μL
mQ		to 20 μL

Table 8: Standard PCR protocol for gene amplification

STEP	TEMP.	TIME	CYCLES
Initial denaturation	95 °C	5 min	1
Denaturation	95 °C	30 sec	35
Annealing	55-65 °C	30 sec	
Elongation	72 °C	30 sec per kb	
Final elongation	72 °C	10 min	1
Hold	4 °C	∞	1

Table 9: Touchdown PCR protocol for improvement of gene amplification

STEP	TEMP.	TIME	CYCLES
Initial denaturation	96 °C	2 min	1
Denaturation	96 °C	30 sec	1
Annealing	65.0-60.5 °C	30 sec	
Elongation	72 °C	30 sec per kb	
Denaturation	96 °C	30 sec	25
Annealing	58 °C	30 sec	
Elongation	72 °C	30 sec per kb	
Final elongation	72 °C	5 min	1
Hold	4 °C	∞	1

3.1.5 Isolation of plasmid DNA from E. coli

Plasmid DNA was isolated with the Wizard® Plus SV Minipreps DNA Purification System (Promega, Madison, US-WI). No alterations were made to the manufacturers protocol. Yields were quantified by spectrophotometry.

3.1.6 Digestion of DNA by restriction endonucleases

Restriction endonucleases are enzymes that cut DNA at specific recognition sites. Their digestive ability was used to prepare DNA inserts and vectors for subsequent ligation, or to test plasmid DNA regarding cloning success. Respective enzymes were used in FastDigest™ quality (Thermo Fischer Scientific Inc., Waltham, US-MA) and applied in standard reaction mixtures (Table 10) that were incubated at 37 °C for one hour or overnight. Reaction products were evaluated by separation of the DNA fragments on agarose gels.

Table 10: Composition of restriction enzyme reaction mixture

COMPONENT	VOLUME
DNA	variable (up to 1 µg)
10× FastDigest™ buffer	2 µL
Restriction enzyme	1 µL
mQ	to 20 µL

3.1.7 Agarose gel electrophoresis

Separation and visualization of DNA or RNA was carried out by agarose gel electrophoresis. By boiling, 1 % of agarose was dissolved in 1x TAE Buffer (Table 11) and slightly cooled before Roti®-GelStain (5 µL / 100 mL agarose) was added to the solution. Gels were run at 120 V for an appropriate duration (30 min for 1.5 kb DNA fragment). GeneRuler™ 1 kb DNA Ladder (Thermo Fischer Scientific Inc., Waltham, US-MA) was used as a marker. When necessary, loading dye (Table 12) was added to the samples prior to pocket filling. Extraction of DNA from gels was performed with the QIAquick® Gel Extraction Kit (Qiagen N.V., Venlo, Netherlands) according to the distributed protocol.

Table 11: Composition of 50× TAE buffer (pH 8)

COMPONENT	VOLUME
TRIS	242 g
Acetic acid	51.1 mL
EDTA (0.5 M)	100 mL
mQ	to 1 L

Table 12: Composition of 10× Orange dye (pH 8)

COMPONENT	VOLUME
Orange G	0.21 % (w/v)
EDTA	100 mM
Glycerol	50 % (v/v)

3.1.8 Ligation of DNA fragments

All genes were sub-cloned in the pGEM®-T Easy vector system (Promega, Madison, US-WI) first to increase the efficiency of subsequent cloning experiments. The pGEM®-T Easy vector is linearized and features a 3'-terminal thymidine overhang at both ends. Therefore, DNA inserts were prepared by incubation with dATP (0.5 mM) and *Taq* DNA polymerase (5 units) in 1x ThermoPol® reaction buffer (all Thermo Fischer Scientific Inc., Waltham, US-MA) for 30 min at 72 °C before ligation was performed according to the manufacturers instructions. After thus prepared, inserts were cut from the sub-cloning vector by appropriate restriction enzymes and ligated in expression vectors by T4-DNA ligase (Promega, Madison, US-WI). Insert-vector ratios were calculated, and ligation reactions were performed as specified by the manufacturer.

3.1.9 Transformation of plasmid DNA in microorganisms

E. coli cells were mixed with ligation reactions and set on ice for 30 min before they were heat-shocked for 45 sec at 42 °C in a water bath. Afterwards, they were cooled on ice for 5 min, supplemented with SOC medium (Hanahan, 1983) and incubated at 37 °C and 150 rpm for 1.5 h. Transformants were plated on LB-agar (Bertani, 1951) containing appropriate antibiotics and incubated at 37 °C overnight.

S. CEREVISIAE INVsc1 cells were transformed with the *S. c.* EasyComp™ Transformation Kit (Invitrogen™, Thermo Fischer Scientific Inc., Waltham, US-MA), plated on SC-U minimal agar (prepared according to the pYES2 vector manual; Invitrogen™, Thermo Fischer Scientific Inc., Waltham, US-MA) and incubated for 48 h at 30 °C.

A. TUMEFACIENS cells were gently mixed with 1 µg of plasmid DNA and set on ice for 5 min before they were frozen in liquid nitrogen for 5 min, followed by a 5 min heat-shock step at 37 °C. Then 1 mL of Luria-Bertani (LB) medium was added and the cells were shaken at 150 rpm for 2-4 h at 28 °C and finally plated on selective LB agar (48 h, 28 °C).

Media compositions are shown in Table 13.

3.1.10 *Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)*

As described by Laemmli, 1970, proteins were separated by discontinuous SDS-PAGE according to their molecular weight. A maximum of 10 µg protein per sample was mixed with 1× Roti®-Load 1 and denatured at 95 °C for 5 min, before the gels were loaded. Protein concentrations were determined by Bradford assay (Bradford, 1976), and samples were run in 1× running buffer on 5 % stacking gels and 10 % running gels at 120 V and 400 mA for approximately 1.5 h. 5 µL of PageRuler™ Plus Prestained Protein Ladder (Thermo Fischer Scientific Inc., Waltham, US-MA) was used as a marker.

After electrophoresis, the gel was removed from the tray and the proteins were stained overnight at room temperature with a colloidal coomassie solution (Kang et al., 2002). If necessary, excess color was removed with a destaining solution (20 % EtOH and 10 % acetic acid in ddH₂O). Buffers and solutions were prepared in mQ water and are listed in Table 14.

Table 13: Media for the cultivation of microorganisms

MEDIUM	COMPONENT	COMPOSITION
LB medium/agar	NaCl Yeast extract Peptone/Tryptone Agar (optional) pH	10 g L ⁻¹ 5 g L ⁻¹ 10 g L ⁻¹ 15 g L ⁻¹ 7
SOC medium	NaCl Yeast extract Peptone/Tryptone KCl MgCl ₂ MgSO ₄ × 4 H ₂ O Glucose pH	0.5 g L ⁻¹ 5 g L ⁻¹ 20 g L ⁻¹ 2.5 mM 10 mM 10 mM 20 mM 7
SC-U medium/agar	Yeast nitrogen base Amino acid mix A [*] Amino acid mix B ⁺ Glucose/Galactose Agar (optional)	0.67 % (w/v) 0.01 % (w/v) 0.005 % (w/v) 2 % (w/v) 2 % (w/v)

^{*} Adenine, Arg, Cys, Leu, Lys, Thr, Trp⁺ Asp, His, Ile, Met, Phe, Pro, Ser, Tyr, Val

Table 14: Buffers and solutions used for SDS-PAGE

BUFFER	COMPONENT	COMPOSITION
10× Running buffer	TRIS Glycin SDS	30 g L ⁻¹ 144 g L ⁻¹ 10 g L ⁻¹
Colloidal coomassie	Al ₂ (SO ₄) ₃ × 16 H ₂ O EtOH Phosphoric acid Coomassie G250	50 g L ⁻¹ 100 mL L ⁻¹ 25 mL L ⁻¹ 0.2 g L ⁻¹

3.1.11 Western Blot

Via Western Blot, proteins can be detected by antibodies that bind specifically to a target protein. After SDS-PAGE, the proteins are pulled from the gel by an electric current and transferred onto a polyvinylidene fluoride (PVDF) membrane (15 V, 0.8 mA/cm² membrane, 1 h). Before the blotting was started, the membrane was activated in 80 % EtOH. To ensure electrical conductivity the whole transfer was carried out in semi-dry blotting buffer (Table 15). Subsequently, the membrane was incubated in blocking buffer (1 h, RT), before a specific antibody was added (1 h, RT). The gallate-UGTs were expressed with an N-terminal GST-Tag. Therefore, they were detected by a combination of a primary anti-GST-antibody and an alkaline phosphatase (AP) coupled secondary anti-mouse-IgG. The putative acyltransferases were expressed with a C-terminal His-Tag and detected by an anti-6x His-Tag[®]-antibody (Abcam, Cambridge, UK), which was already coupled to an AP. After incubation with antibodies, the membrane was washed three times in washing buffer and the blotted proteins were detected by the colorimetric conversion of 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT) catalyzed by the AP.

Table 15: Buffers used for Western Blot

BUFFER	COMPONENT	COMPOSITION
Semi-dry blotting buffer	TRIS Glycin MeOH	3 g L ⁻¹ 15 g L ⁻¹ 200 mL L ⁻¹
Washing buffer	TRIS NaCl Tween 20	2.42 g L ⁻¹ 8.18 g L ⁻¹ 1 mL L ⁻¹
Detection buffer	TRIS NaCl MgCl ₂ BCIP NBT	100 mM 100 mM 3 mM 5 % (w/v) in mQ 5 % (w/v) in 70 % DMF

3.2 CHARACTERIZATION OF GALLATE-UDP-GLUCOSE DEPENDENT GLUCOSYLTRANSFERASES (UGTs)

3.2.1 Construction of UGT plasmids

As described in Schulenburg et al., 2016a, *FaGT2**, *FaGT2*, *FaGT5*, and the mutants *FaGT2*_R230S* and *FaGT2*_E420D_I422V* were cloned into the pGEX4T-1 vector for expression in *E. coli*. *FvGT2* and *RiGT2* were cloned into the pGEX4T-1 vector by Antje Feller, who provided both constructs within the scope of our cooperation and helped with the processing of the respective enzyme assays. Site-directed mutagenesis of the *FaGT2** sequence was performed according to the QuikChange™ protocol (Agilent Technologies Inc., Santa Clara, US-CA). *VvgGT1* was obtained according to Khater et al., 2012, also cloned into the pGEX4T-1 vector and used as a positive control. All transformants were checked by sequencing of the complete insert (Eurofins Genomics). A construct of pGEX4T-1 containing

UGT_{71K3} was kindly provided by Dr. Chuankui Song² and employed under identical conditions.

3.2.2 *Heterologous expression and purification of recombinant gallate-UGTs in E. coli*

All UGT candidate genes were expressed in *E. coli* BL₂₁ (DE₃) pLysS, tailed with an N-terminal GST-Tag, and purified as specified in Schulenburg et al., 2016a. In short, the cells were disrupted by sonication, and the crude protein solution was supplemented with 10 µM of protease inhibitor phenylmethylsulfonyl fluoride (PMSF). Subsequently, the recombinant fusion proteins were purified by affinity chromatography. Yields were quantified (Bradford, 1976), and the presence of the heterologously expressed enzymes was proved by SDS-PAGE and Western Blot.

3.2.3 *Radio-labeled enzyme assays and kinetics*

The UGTs analyzed here produce acyl esters of (hydroxy)benzoic and (hydroxy)cinnamic acids and UDP-glucose. Before substrate screens and kinetic measurements were conducted, optimal reaction conditions were determined (100 mM TRIS-HCl, 30 °C, 2 µg of recombinant enzyme, pH 7). Then, the formation of 1-O-acyl esters was carried out using a mixture of labeled UDP-[U-¹⁴C]-D-glucose (specific activity: 300 mCi mmol⁻¹; Biotrend Chemikalien GmbH, Köln, Germany) and unlabeled UDP-D-glucose as co-substrates. The reaction products were separated by extraction with 1-butanol and their radioactivity was determined by liquid scintillation counting. The kinetic data were determined with increasing substrate and fixed co-substrate concentrations. An empty vector protein extract was used as negative control. To minimize inconsistencies in the data, kinetic constants were cal-

² Associate Professorship of Biotechnology of Natural Products, Technical University of Munich, Liesel-Beckmann-Straße 1, 85354 Freising

culated from Hanes-Woolf plots and non-linear regression of the hyperbolic Michaelis-Menten equation. Only values that matched both calculations were included in the Results section. Hanes-Woolf plots alone might exaggerate incorrect measures of substrate concentration (e.g. pipetting errors) because the substrate concentration is used on both axes. A more detailed description can be found in the publication of Schulenburg et al., 2016a.

3.2.4 *Injection of deuterium-labeled gallic acid into strawberry fruits*

In vivo activity of the GT2 enzymes was determined by injection of 10 mM deuterium(d₂)-labeled gallic acid (3,4,5-trihydroxybenzoic-2,6-d₂ acid) into small green fruits of stable transgenic *FaGT2*-silenced strawberry plants and *F. × ananassa* cv. Calypso control plants. Fruits were collected after 24 h of incubation time, freeze-dried and stored at -20 °C.

3.2.5 *Identification and quantification of metabolites by LC-MS analysis*

The levels of gallic acid, d₂-gallic acid, β-glucogallin, d₂-β-glucogallin, gallic acid glucoside, d₂-gallic acid glucoside, digalloyl-glucose, d₂-digalloyl-glucose, ellagic acid and d₂-ellagic acid were quantified as described in Section 3.1.1 and Schulenburg et al., 2016a.

3.3 TRANSCRIPTOME AND METABOLOME ANALYSIS OF *f. vesca*

The following section describes the methods applied according to Härtl et al., 2017a.

3.3.1 *Fruit sample preparation*

Fruits of two white-fruited *F. vesca* varieties YW and HW4, and one red-fruited variety RdV were collected in three ripening stages (green: \approx 10 dpa, intermediate: \approx 25 dpa, and ripe: \approx 35 dpa). Before the achenes were removed from the fruity pulp, the fruits were frozen in liquid nitrogen. Three strawberry varieties (RdV, YW and HW4), three ripening stages (green, intermediate and ripe) and two tissues (achenes and receptacle) equal 18 samples subjected to metabolite quantification and transcriptome analysis. For RNA isolation material of 10 fruits per sample was pooled to ensure high reliability regarding the ripening stage. Total RNA was prepared as described in Section 3.1.2. Metabolite extraction and LC-MS measurements were performed as specified in Section 3.1.1, and three to five biological replicates per sample were processed.

3.3.2 *Analysis of F. vesca metabolites*

Untargeted metabolite profiling was done in R according to literature to reveal metabolites accumulating differently (Benton et al., 2010; Smith et al., 2006; Tautenhahn et al., 2008). A p-value ≤ 0.01 was considered significant. Hierarchical clustering was achieved employing the complete linkage method and Spearman distance metric. Visualization of the data variance in principle component analysis (PCA) plots was performed according to Ritchie et al., 2015. Targeted metabolite profiling was conducted via the internal standard method as described in Section 3.1.1, and the compounds were identified using authentic reference chemicals and the in-house database.

3.3.3 *Bioinformatic analysis and evaluation of the sequencing data*

Total, DNA-free RNA was sent to Eurofins Genomics (Ebersberg, Germany) where RNA-seq was carried out on an Illumina HiSeq2000 platform (Illumina Inc., San Diego, US-CA),

and the cDNA library was generated from isolated poly(A)-tailed 3'-RNA-fragments. RNA-seq data were handled on a locally installed Galaxy server (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010) employing the Application Programming Interface (API; Sloggett et al., 2013), the Galaxy Data Manager (Sloggett et al., 2013), the Galaxy ToolShed (Blankenberg et al., 2014b), and built-in reference data (Blankenberg et al., 2014a). Quality trimming of the single end reads was performed using the Trimmomatic tool applying default settings (Bolger et al., 2014). Via TopHat (Kim et al., 2013) in default settings the trimmed reads were mapped to the *F. vesca* reference genome (Tennessen et al., 2014). Mapped reads were counted with HTSeq-count (Anders et al., 2015) supplemented with gene prediction information (Tennessen et al., 2014). Thus created raw counts were normalized for library size (Anders and Huber, 2010), and a cutoff of 20 reads summed across all samples was set. Genes showing a lower transcript number were considered as not expressed and excluded from further analyses. Finally, the remaining normalized read counts were scaled to per million scale to yield reads per million (RPM).

Hierarchical clustering and subsequent PCA plot generation was performed as described in Section 3.3.2, except in advance variance stabilization was used to transform the data set (Huber et al., 2002). Gene annotations were assigned by MapMan BINs (Lohse et al., 2014) and *F. vesca* gene ontologies (Darwish et al., 2015; Shulaev et al., 2011). Finally, differential gene expression was determined in edgeR employing the general linear models (Robinson et al., 2010), and after calculation of the false discovery rate (FDR; Benjamini and Hochberg, 1995) a cutoff of FDR<0.05 was set to determine significant genes.

The normalization of the raw read counts and the differential expression analysis was performed by Alisandra Denton³ within the scope of our cooperation. Furthermore, Alisandra computed Figure 16 based on Figure 1 and provided an R-script for the creation of Figure 17, Figure 18 and Figure 19.

³ Institute for Biology I, RWTH Aachen University, Worringer Weg 3, 52074 Aachen, Germany

3.3.4 Validation of candidate gene expression by real-time (RT)-PCR

The transcript profiles of GT1 and MYB10 were validated by RT-PCR. cDNA was synthesized as described in Section 3.1.3. To observe dsDNA synthesis, the reaction mixtures (Table 16) were prepared with the SensiMix™ SYBR HI-ROX Kit (Bioline Ltd, London, UK) containing SYBR Green. UBC9 was employed as reference gene. Primer sequences were adopted from Griesser et al., 2008 and Lin-Wang et al., 2014. In advance, cDNA concentration, primer concentration and annealing temperature were optimized (GT1 gene12591: 1× cDNA, 61 °C; MYB10 gene31413: 0.1× cDNA, 57 °C; UBC9 gene12591: 0.1× cDNA, 55 °C; 400 nM primers). Double-strand DNA amplification was generated in 40 cycles, followed by a melting curve analysis to determine the PCR specificity (Table 17). All analyses were performed in triplicates and relative expression was determined according to known protocols (Pfaffl, 2001).

Table 16: RT-PCR reaction mixture

COMPONENT	FINAL CONC.	FINAL VOL.
cDNA template	variable	2 µL
Primer fw (10 pmol µL ⁻¹)	0.4 µM	0.8 µL
Primer rev (10 pmol µL ⁻¹)	0.4 µM	0.8 µL
Master mix	2×	10 µL
mQ		to 20 µL

Table 17: Thermal cycling conditions for determination of relative gene expression

STEP	TEMP.	TIME	CYCLES
Initial denaturation	95 °C	2 min	1
Denaturation	95 °C	5 sec	40
Annealing	55/57/61 °C	10 sec	
Elongation	72 °C	20 sec	
Melting curve step 1	95 °C	15 sec	1
Step 2	60 °C	60 sec	
Step 3	95 °C	15 sec	

3.4 CHARACTERIZATION OF PUTATIVE ACYLTRANSFERASES (ATs)

3.4.1 Construction of AT plasmids

Putative ATs were amplified from cDNA of green achenes of *F. vesca* cv. YW with primers introducing BamHI and XhoI restriction sites (Table 2). For *in vitro* enzyme characterization, the genes were cloned into the pYES2 yeast expression vector (Invitrogen™, Thermo Fischer Scientific Inc., Waltham, US-MA) and transformed into the *S. cerevisiae* strain INVSc1 as described in Section 3.1.9. Additionally, the sequences were amplified with primers introducing a C-terminal His-Tag.

For *in vivo* characterization, the putative ATs were amplified with primers introducing BamHI and Ecl136II restriction sites, cloned into the pBI12135S2x vector (Hoffmann et al., 2006), and transformed in the *A. tumefaciens* strain AGLo (Lazo et al., 1991).

3.4.2 Heterologous expression and purification of recombinant ATs in *S. cerevisiae*

The putative ATs were expressed in SC-U medium containing 2 % galactose as described by Huang et al., 2010. After harvesting, the cells were resuspended in breaking buffer containing 1 mM PMSF (Table 18, prepared according to Huang and Schwab, 2011), and the crude protein was extracted by vortexing with glass beads as described by Huang and Schwab, 2011. Yields were quantified (Bradford, 1976) and the protein was kept on ice or frozen at -20 °C supplemented with 50 % glycerol (v/v) until further processed. The presence of recombinant proteins was confirmed by SDS-PAGE and Western Blot (Section 3.1.10, Section 3.1.11).

If needed, the His-Tag flagged heterologously expressed proteins were enriched via affinity chromatography using the Profinity™ IMAC Ni-Charged Resin (Bio-Rad Laboratories Inc., Hercules, US-CA) while following the producers manual.

Table 18: Breaking buffer recipe for purification of recombinant ATs

COMPONENT	COMPOSITION
Sodium phosphate buffer	50 mM
EDTA	1 mM
Glycerol	5 % (v/v)
PMSF	1 mM
pH	7.5

3.4.3 Enzyme assays and product detection via LC-MS

To determine the *in vitro* AT activity, enzyme assays were performed. Reaction mixtures of 200 µL in total contained 100 mM buffer (TRIS-HCl, sodium phosphate or citric acid), crude protein extract and 1 mM of substrate. Incubation took place at RT or 30 °C overnight and the products were extracted by ad-

dition of one volume of ethyl acetate, 1 min of vortexing and 2 min of centrifugation (full speed). The supernatant was transferred to a new tube, the residue was extracted a second time and the supernatants combined. Subsequently, the ethyl acetate was removed in a vacuum concentrator and the reaction products were resolved in 50 µL of mQ water (1 min vortexing, 5 min sonication, 2 min centrifugation at full speed). Product detection in clear supernatants was carried out via LC-MS as described in Section 3.1.1 and in the publication of Ring et al., 2013.

3.4.4 Over-expression of candidate ATs in strawberry fruits

As published by Hoffmann et al., 2006, AGLo cells containing the pBI12135S2x-AT vectors were grown in LB at 28 °C with appropriate antibiotics (100 µg mL⁻¹ Kanamycin, 100 µg mL⁻¹ Rifampicin) until OD₆₀₀ reached 0.8. Then, the cells were harvested and resuspended in a modified MacConkey agar (MMA) medium (Table 19, prepared according to Spolaore et al., 2001). The resulting *Agrobacterium* suspension was injected into green-white fruits (approximately 14 days post pollination) of *F. × ananassa* cv. Elsanta. A pBI12135S2x construct containing a β-glucuronidase (GUS)-intron gene (Hoffmann et al., 2006) was processed under identical conditions and used as a control. In *A. tumefaciens* the intron from the *F. × ananassa* quinone oxidoreductase (QR) suppresses gene expression. But in plant cells the intron is spliced out, and a functional mRNA is produced.

Fruits injected with transformants were kept for 14 days attached to the plant, harvested subsequently and cut in half. One half was frozen in liquid nitrogen and stored at -80 °C until RNA was extracted for reverse transcription and RT-PCR measurements. The other half was lyophilized for 2 days, ground to fine powder by mortar and pestle and subjected to metabolite extraction and LC-MS analysis (Section 3.1.1).

Table 19: MMA medium recipe for *A. tumefaciens* infiltration

COMPONENT	COMPOSITION
Murashige and Skoog salts	4.3 g L ⁻¹
Morpholine ethanesulphonic acid (pH 5.6)	10 mM
Sucrose	20 g L ⁻¹

3.4.5 Validation of AT over-expression by RT-PCR

After the candidate ATs and the GUS-intron control gene were over-expressed in ripening strawberry fruits by *A. tumefaciens* mediated gene transfer, the relative gene expression was determined by RT-PCR. Total RNA was isolated and cDNA was synthesized according to Section 3.1.2 and Section 3.1.3. RT-PCR was carried out as described in Section 3.3.4 employing specific primers listed in Table 3. All reactions were run in three technical replicates on 15 to 20 biological replicates, and relative expression was normalized against a 18-26S interspacers as described in Griesser et al., 2008.

4

RESULTS

4.1 FORMATION OF GALLOYL GLUCOSE

Here, the identification and characterization of five UGTs from woodland strawberry, garden strawberry, and raspberry are reported that catalyze the formation of β -glucogallin. Metabolite analyses of fruits in different ripening stages were performed, which enabled the identification of the main source tissue of gallic acid, β -glucogallin and ellagic acid. Finally, an *in vivo* experiment confirmed the *in planta* function.

The results of this section have already been published in a peer-reviewed journal (Schulenburg et al., 2016a).

4.1.1 Identification of putative gallic acid UGTs from *F. vesca*, *F. x ananassa* and *R. idaeus*

Recently, the first gallic acid UGT genes have been identified in *V. vinifera* (Khater et al., 2012). The closest homologous sequences in the garden strawberry are *FaGT2* and *FaGT5*. Although, *FaGT2* encodes a cinnamoyl and 4-coumaroyl glucose ester forming UGT (Landmann et al., 2007; Lunkenstein et al., 2006a) it was considered a candidate gene. Through amplification of the *FaGT2* open reading frame (ORF) from *F. × ananassa* cv. Elsanta cDNA an allelic sequence *FaGT2** to the original *FaGT2* version was obtained. It comprised 18 single nucleotide polymorphisms (SNPs) translating to eight amino acid substitutions (Figure 7). In addition to the *FaGT5* gene both *FaGT2* versions were characterized, as well as two mutants of *FaGT2** generated by site-directed mutagenesis. Furthermore, the GT2-orthologs *F. vesca FvGT2* and *R. idaeus RiGT2* were drawn on for functional verification across the Rosaceae plant family. *VvgGT1*

The analyzed UGTs are homologous proteins.

(Khater et al., 2012) was employed as a positive control. All GT2 proteins share a sequence homology of >80 % with VvGT1 (Figure 7).

4.1.2 Enzymatic UGT activity

All UGTs converted gallic acid.

The substrate specificity of the UGT enzymes was determined *in vitro* with a set of 18 (hydroxy)benzoic and (hydroxy)cinnamic acid derivatives, whereas a mixture of labeled UDP-[U-¹⁴C]-D-glucose and unlabeled UDP-D-glucose served as donor substrates (Table 20). The acceptor substrates were chosen because they are known precursors of ETs/EA (e.g. gallic acid, 4-hydroxybenzoic acid, and protocatechuic acid), or because of their structural similarity to these precursors (e.g. salicylic acid and gentisic acid), while others have been reported in fruits of the *Rosaceae* plant family (e.g. cinnamic acid and *p*-coumaric acid, Lunkenbein et al., 2006a). All GT2-homologous proteins converted GA to β-glucogallin. FaGT2 and FaGT2* preferred benzoic acid derivatives, while FvGT2 and RiGT2 showed a higher affinity towards cinnamic acid derivatives. Sinapic acid was the best substrate for FaGT5 (100 % relative activity), although, it also converted veratric acid efficiently. However, FaGT5 showed only little activity towards GA. VvgGT1 preferred *m*-coumaric acid, but also exhibited high activity towards veratric and sinapic acid.

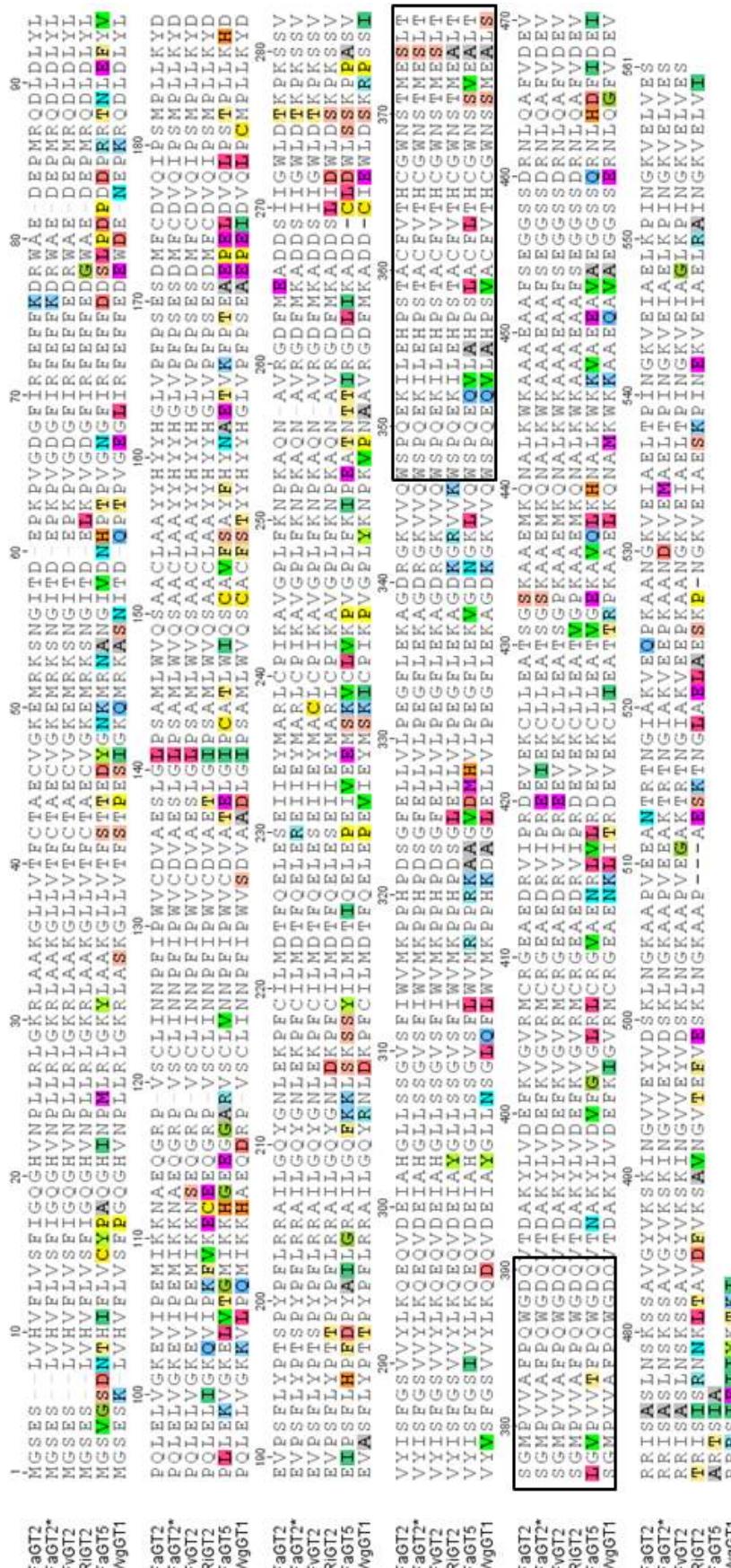


Figure 7: Protein sequence alignment of gallate-UGTs FaGT2*, FaGT2, FvGT2, RiGT2, FaGT5 and VvgGT1. The ClustalX function integrated in the GENEIOUS Pro 5.5.6 software (Biomatters; <http://www.genieous.com/>, accessed August 2015) was employed in default parameters, to generate this alignment. The PSPG box is boxed (Figure adopted from Schulenburg et al., 2016a).

Table 20: Relative activity of the UGT enzymes towards 18 acceptor substrates (adopted from Schulenburg et al., 2016a).

SUBSTRATES	FaGT2*	FaGT2	FvGT2	RiGT2	FaGT5	VvgGT1
sorbic acid	32	66	45	81	13	93
salicylic acid	0	88	1	4	1	6
3-Hydroxybenzoic acid	99	95	80	77	3	39
4-Hydroxybenzoic acid	100	89	77	89	15	69
Protocatechuic acid	92	95	71	68	6	61
Gentisic acid	0	81	5	15	1	13
Gallic acid	34	89	51	43	5	42
Vanillic acid	99	98	83	78	21	69
Veratric acid	76	100	82	64	92	94
Syringic acid	87	94	75	59	13	86
Cinnamic acid	38	80	57	93	14	31
<i>o</i> -Coumaric acid	76	87	82	99	11	82
<i>m</i> -Coumaric acid	95	63	100	90	21	100
<i>p</i> -Coumaric acid	53	75	95	100	48	90
Caffeic acid	65	81	92	91	17	80
Ferulic acid	50	5	61	65	36	61
3,4-Dimethoxycinnamic acid	85	55	88	75	34	75
Sinapic acid	89	17	93	78	100	99

The relative activities (%) were determined by radio-chemical analysis with UDP-[U-¹⁴C]-D-glucose, and refer to the highest level of radioactivity extracted for every enzyme. FaGT2* 100 % ≡ 1.0 nkat mg⁻¹; FaGT2 100 % ≡ 1.4 nkat mg⁻¹; FvGT2 100 % ≡ 1.3 nkat mg⁻¹; RiGT2 100 % ≡ 1.9 nkat mg⁻¹; FaGT5 100 % ≡ 2.5 nkat mg⁻¹; VvgGT1 100 % ≡ 1.4 nkat mg⁻¹. Empty vector control (<1 %) was subtracted from the relative concentrations.

4.1.3 Determination of the kinetic values

Enzyme kinetics were determined for the six UGTs with selected acceptor substrates (Table 21). The obtained kinetic values confirmed a high affinity of all UGTs towards GA except FaGT5, with RiGT2 showing the highest enzyme efficiency in producing β -glucogallin ($k_{\text{cat}}/K_M = 20,000 \text{ s}^{-1} \text{ M}^{-1}$). FaGT5 clearly preferred sinapic acid ($k_{\text{cat}}/K_M = 15,555 \text{ s}^{-1} \text{ M}^{-1}$) above the other substrates tested.

4.1.4 Site-directed mutagenesis of FaGT2*

To determine the effect of the amino acid substitutions in FaGT2 and FaGT2*, three positions were reversed in the FaGT2* sequence by site-directed mutagenesis. Two positions were selected because of their vicinity to the PSPG box (Glu420 and Ile422). The third position was chosen for its closeness to the proposed active site of the protein (Arg230, Bönisch et al., 2014a). In mutant FaGT2*-Glu420Asp-Ile422Val glutamate and isoleucine were exchanged by aspartate and valine, respectively. In mutant FaGT2*-Arg230Ser arginine was mutated to serine at position 230. The substrate specificity of both mutants was determined for the set of 18 substrates (Section 4.1.2). The relative activity of the two mutants was calculated and the activity of FaGT2* subtracted (Figure 8). Both mutants showed an increased activity towards GA (35 % and 38 % for FaGT2*-Arg230Ser and FaGT2*-Glu420Asp-Ile422Val, respectively), *o*-coumaric (24 % and 39 %) and caffeic acid (25 % and 31 %). On the other hand, the activity towards sorbic acid decreased (35 % and 24 %).

Activity towards gallic acid increased.

4.1.5 Metabolite Analysis in *F. x ananassa* and *F. vesca*

Pulp and achenes of *F. x ananassa* cv. Calypso, and *F. vesca* cv. YW were analyzed by LC-MS in the small green, big green, white, turning, and ripe developmental stage to determine the

The metabolite concentration was highest in green tissues.

Table 21: Kinetic values of the GT2-homologous enzymes determined for the substrates gallic acid, 4-hydroxybenzoic acid, vanillic acid, syringic acid, cinnamic acid, sinapic acid and UDP-glucose (adopted from Schulenburg et al., 2016a).

ENZYME	SUBSTRATE	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\text{s}^{-1} \text{ M}^{-1}$)
FaGT2*	Gallic acid	96 ± 3.6	0.8 ± 0.04	8,333
	4-Hydroxybenzoic acid	315 ± 8.4	2.0 ± 0.06	6,349
	Vanillic acid	364 ± 12.8	2.1 ± 0.04	5,769
	Syringic acid	262 ± 7.0	1.8 ± 0.07	6,870
	Sinapic acid	35 ± 3.0	1.2 ± 0.04	34,285
	UDP-glucose	378 ± 2.3	3.8 ± 0.07	10,052
FaGT2	Gallic acid	264 ± 9.9	1.8 ± 0.01	6,818
	4-Hydroxybenzoic acid	899 ± 12.3	2.2 ± 0.17	2,447
	Vanillic acid	621 ± 12.5	2.1 ± 0.04	3,381
	Syringic acid	874 ± 6.2	2.6 ± 0.1	2,974
	Sinapic acid	83 ± 4.8	2.5 ± 0.17	30,120
	UDP-glucose	492 ± 5.0	4.9 ± 0.15	9,959
FvGT2	Gallic acid	232 ± 8.3	1.1 ± 0.05	4,741
	UDP-glucose	318 ± 11.9	2.2 ± 0.18	6,918
RiGT2	Gallic acid	80 ± 8.1	1.6 ± 0.05	20,000
	UDP-glucose	417 ± 16.7	6.6 ± 0.07	15,827
FaGT5	4-Hydroxybenzoic acid	515 ± 6.5	0.5 ± 0.04	9,708
	Vanillic acid	94 ± 1.5	0.9 ± 0.02	9,574
	Syringic acid	518 ± 10.0	0.5 ± 0.07	9,652
	Sinapic acid	45 ± 1.1	0.7 ± 0.02	15,555
	UDP-glucose	529 ± 3.0	4.5 ± 0.23	8,506
VvgGT1	Gallic acid	72 ± 3.5	1.1 ± 0.03	15,277
	4-Hydroxybenzoic acid	73 ± 2.0	2.0 ± 0.07	27,397
	Vanillic acid	60 ± 4.7	1.8 ± 0.10	30,000
	Syringic acid	155 ± 2.0	2.0 ± 0.01	12,903
	Sinapic acid	52 ± 2.5	1.9 ± 0.12	36,538
	UDP-glucose	439 ± 16.9	10.1 ± 0.23	22,779

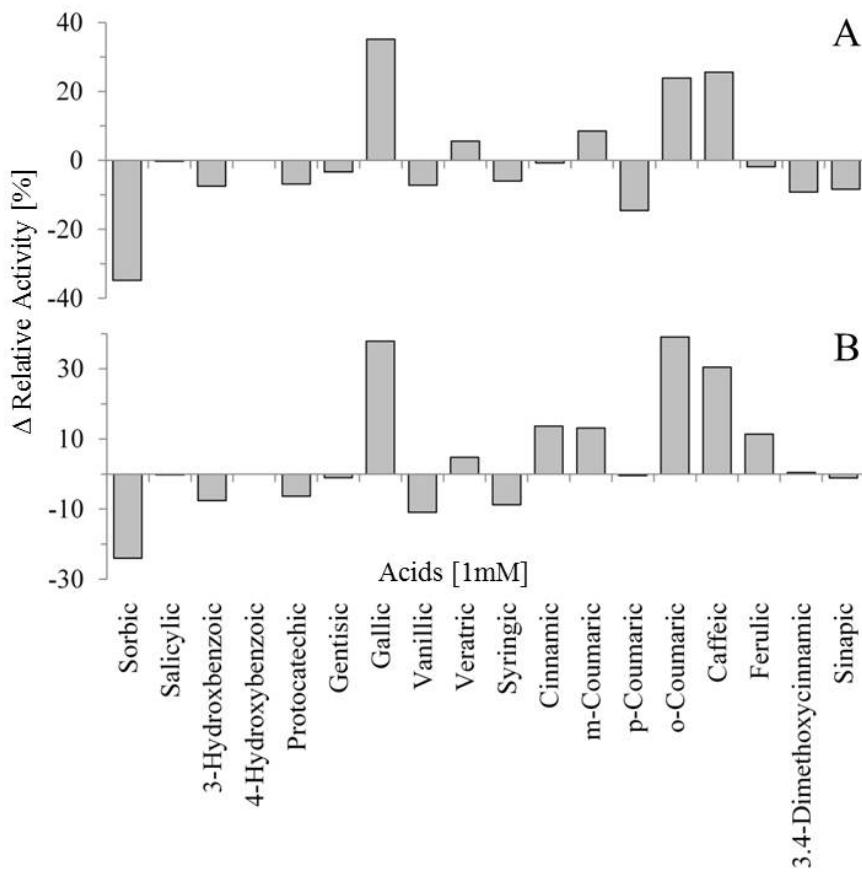


Figure 8: The difference (Δ) between the relative activities of the mutants FaGT2*-Arg230Ser (A) and FaGT2*-Glu420Asp-Ile422Val (B) was calculated against the activities of FaGT2* (Figure adopted from Schulenburg et al., 2016a).

main source tissue of GA, β -glucogallin and EA across fruit ripening (Figure 9). All three metabolites were strongly enriched in green achenes of garden and woodland strawberry, and concomitant with the fruit ripening the levels decreased. The same trend was observed in the receptacle tissues. However, the overall level of analyzed metabolites was higher in the achenes than in the receptacle.

4.1.6 In vivo confirmation of *FaGT2** and *FaGT2* activity

EA production was reduced in the transgenics.

The *in vivo* activity of *FaGT2*/*FaGT2** as gallic acid UGTs was confirmed by injection of aqueous solutions of 2,6-d₂-gallic acid into *F. × ananassa* cv. Calypso control fruits, and fruits of stable transgenic *FaGT2*-silenced Calypso plants. Levels of *FaGT2* transcripts in the fruits of the transgenic line corresponded to only 34 % of the levels in control plants (Lunkenbein et al., 2006b). After one day, the levels (% equ. dw) of d₂-gallic acid, d₂- β -glucogallin, d₂-gallic acid 4-O-glucoside, d₂-di-galloyl glucose and d₂-ellagic acid were quantified by LC-MS (Figure 10). A higher level of d₂-gallic acid, and d₂-gallic acid 4-O-glucoside was found in the *FaGT2*-silenced fruits in comparison to control fruits. In contrast, the level of d₂-ellagic acid was reduced in the transgenic fruits. The enhanced level of labeled d₂-gallic acid 4-O-glucoside in *FaGT2*-silenced fruit suggested the existence of an additional UGT forming a glucoside of gallic acid rather than a glucose ester. Thus, we screened our strawberry fruit UGT library and found UGT71K3, which has been shown to glucosylate pyrogallol a structural homologue of gallic acid (Song et al., 2016b). When UGT71K3 was incubated *in vitro* with UDP-glucose and gallic acid, it readily formed gallic acid 4-O-glucoside. Subsequently, the product was unambiguously identified by LC-MS in comparison with blueberry extract utilized as authentic reference material according to literature (Schuster and Herrmann, 1985, Figure 11). Interestingly, the concentration of d₂- β -glucogallin and d₂-di-galloyl glucose remained constant.

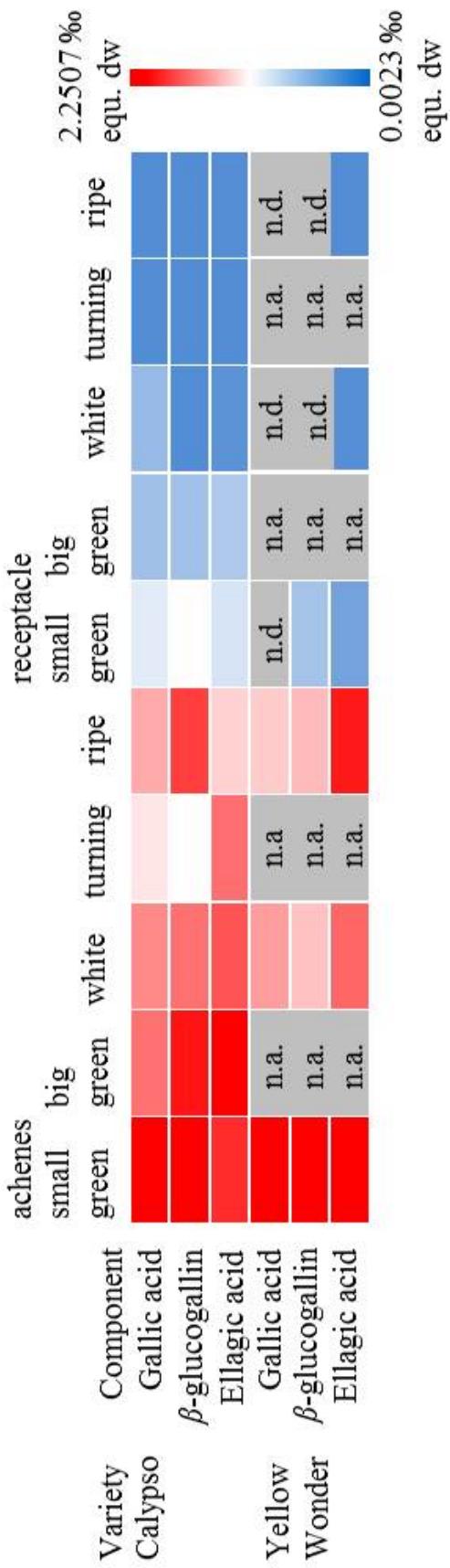


Figure 9: Color code visualization of relative concentrations (per mil equivalents of the dry weight, % equ. dw) of GA, β -glucogallin and EA in five developmental stages of *F. × ananassa* cv. Calypso and *F. vesca* cv. YW fruit tissue. Achenes were removed from the receptacle post-harvest, and both tissues were analyzed separately. n.a., not analyzed; n.d. not detectable; n=3–5 biological replicates (Figure adopted from Schulenburg et al., 2016a).

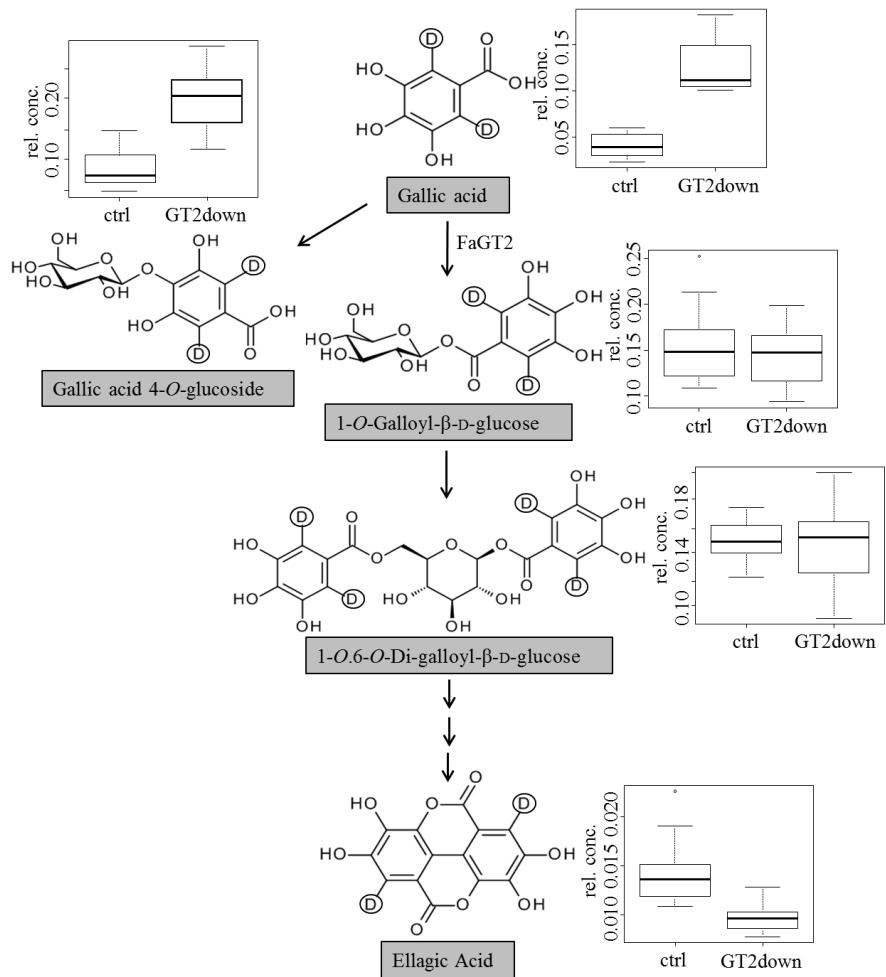


Figure 10: Schematic view of the suggested ellagic acid biosynthesis pathway including molecular structures of quantified metabolites. The box plots show the relative concentration (% equ. dw) of deuterium-labeled compounds in control fruits (ctrl), and in *FaGT2*-silenced fruits (GT2down). Circles highlight the d_2 residues. (Figure adopted from Schuilenburg et al., 2016a).

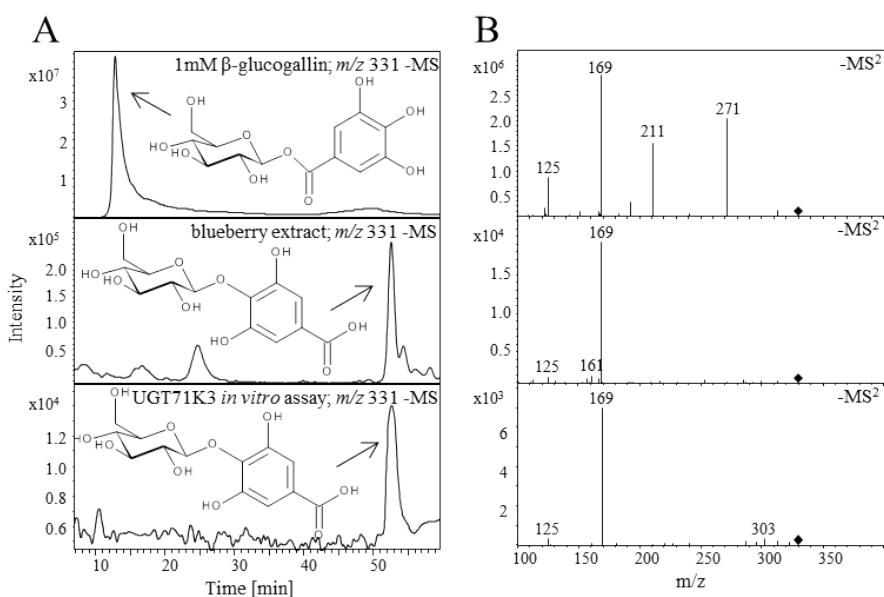


Figure 11: Identification of gallic acid 4-O-glucoside in comparison with authentic reference material. LC-MS extracted ion chromatograms of m/z 331 -MS (A) and corresponding - MS^2 spectra (B) of 1 mM β -glucogallin authentic chemical, an extract of blueberry containing authentic gallic acid 4-O-glucoside (Schuster and Herrmann, 1985), and product formed by recombinant UGT71K3 from UDP-glucose and gallic acid.

4.2 TRANSCRIPTOME AND METABOLOME ANALYSIS OF *F. vesca*

To gain a deeper understanding of the biosynthesis of secondary metabolites, we performed a combined comparative approach of metabolite profiling by LC-MS and transcriptome profiling by RNA-seq in two naturally white-fruited *F. vesca* genotypes (YW and HW4) and one red-fruited variety (RdV). The natural color mutants provide the opportunity to determine how altering transcript levels affect the production of anthocyanins/flavonoids. Of every strawberry variety three ripening stages (green, intermediate, and ripe) were sampled, and the achenes (seeds) were separated from the receptacle (fruity pulp) and analyzed separately (Figure 12).

The results of this section have been published in a peer-reviewed journal (Härtl et al., 2017a).

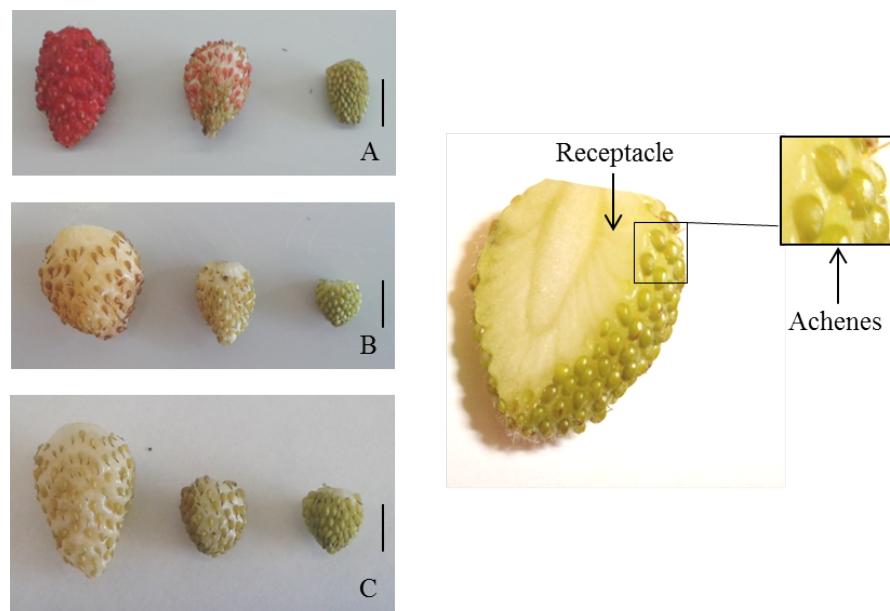


Figure 12: Fruits of *F. vesca* varieties RdV (A), YW (B) and HW4 (C) in the three ripening stages ripe (left), intermediate (middle) and green (right). Cross-section of a green *F. vesca* cv. HW4 fruit (D) with arrows indicating the tissues (receptacle and achenes) separated before RNA and metabolite extraction. Scale bars = 5 mm. (adopted from Härtl et al., 2017a)

4.2.1 Metabolite profiling

Metabolite analyses of phenols, phenylpropanoids, flavonoids, proanthocyanidins, and anthocyanins were performed by LC-MS in receptacles and achenes of green, intermediate, and ripe fruits of *F. vesca* cv. RdV, YW and HW4. Untargeted metabolite analysis revealed that 271 metabolites accumulated significantly different in the fruits of RdV, YW and HW4. When hierarchical clustering was performed, the receptacle samples of the three genotypes clustered according to their ripening stage (Figure 13). All achene samples formed one cluster. Consequently, the secondary metabolites conveying the color to the receptacle of the red genotype is not the primary variance in the data.

On the other hand, the targeted quantification of phenols, phenylpropanoids, flavonoids, proanthocyanidins, and anthocyanins revealed that every ripening stage and tissue is characterized by a particular group of secondary metabolites (Figure 14). Ripe achenes and receptacle of RdV showed high concentrations of the anthocyanins pelargonidin glucoside, pelargonidin glucoside malonate, and cyanidin glucoside conveying the red color (Figure 12), whereas in tissues of YW and HW4 those compounds were missing completely. Green achenes of RdV were dominated by gallic acid, gallic acid glucose ester, and ellagic acid, and intermediate achenes contained high amounts of flavonoids. The levels of secondary metabolites in tissues of YW and HW4 mostly deviated strongly from the respective tissues of the red genotype suggesting spatiotemporal specificities. For example, the concentration of flavonoids was significantly reduced in green achenes of YW and HW4 compared to green of RdV, while phenols accumulated to comparable levels. Furthermore, differences between the two white genotypes were detected as flavonoids and proanthocyanidins were strongly enriched almost exclusively in green receptacle of YW except kaempferol glucoside and catechin.

Specific groups of metabolites characterized the ripening stages.

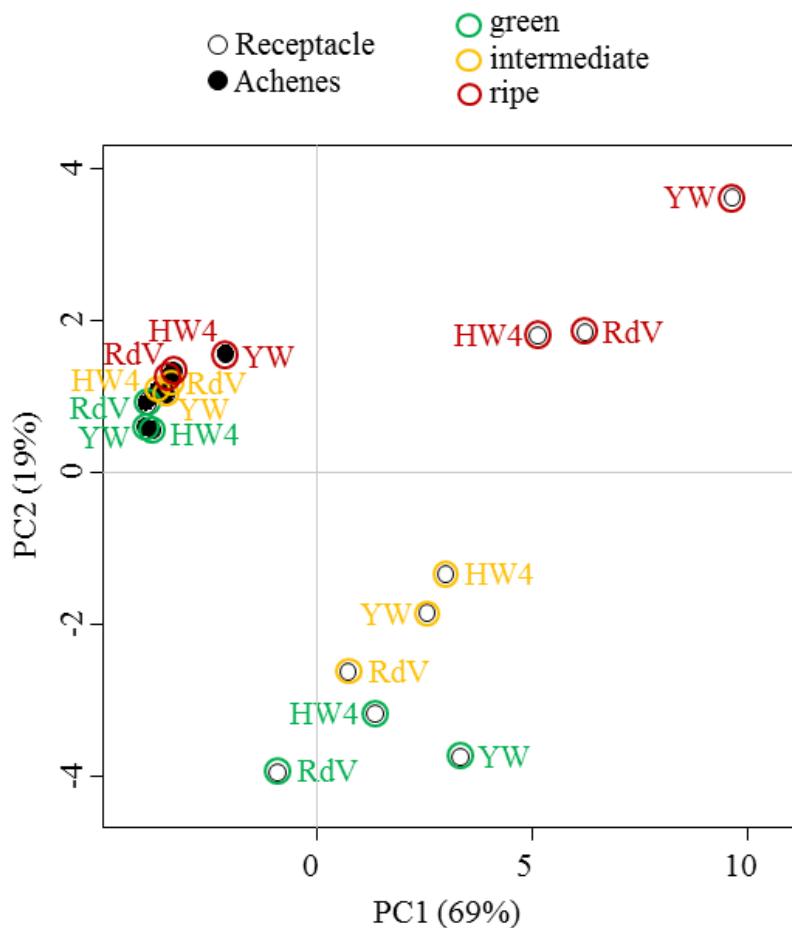


Figure 13: Untargeted analysis of secondary metabolites in receptacle and achenes of strawberry varieties RdV, YW and HW4 of three developmental stages (green, intermediate and ripe). Three to five biological replicates per data point. The variance in the data is depicted by PCA. (adopted from Härtl et al., 2017a)

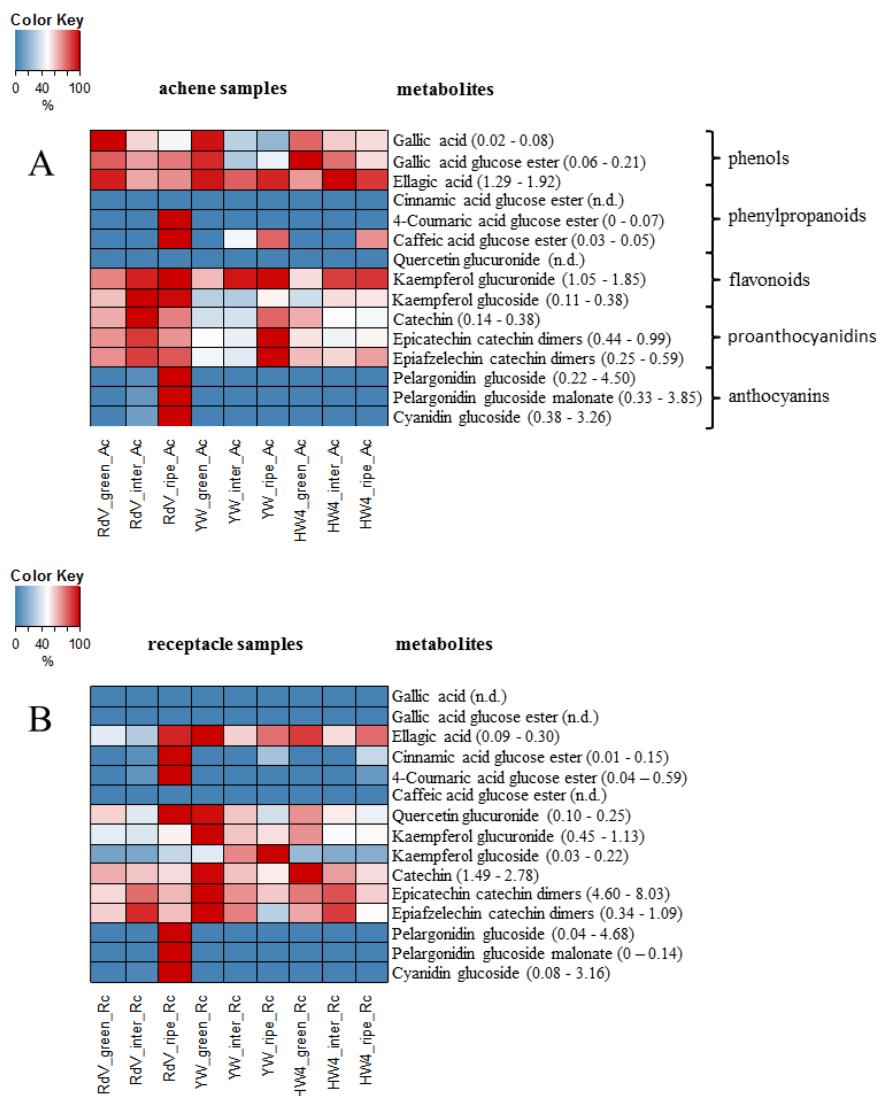


Figure 14: Quantification of secondary metabolites by LC-MS in achene and receptacle of the strawberry varieties RdV, YW and HW4 of three developmental stages (green, intermediate and ripe). Heat maps show levels expressed as per mil equivalents of the dry weight (relative concentration), with lowest levels shown in blue, and highest levels in red. Individual min. and max. values are given in parentheses. (A) Color code presentation of metabolite levels in achene (Ac) tissues of different ripening stages (green, intermediate and ripe), and *F. vesca* varieties (RdV, YW and HW4). (B) Metabolite levels in receptacle (Rc) tissues. n.d. not detected. Three to five replicates were analyzed. (adopted from Härtl et al., 2017a)

4.2.2 Transcriptome profiling - global data analysis

*Intermediate
achenes and
receptacle of the
white varieties
grouped with green
tissues.*

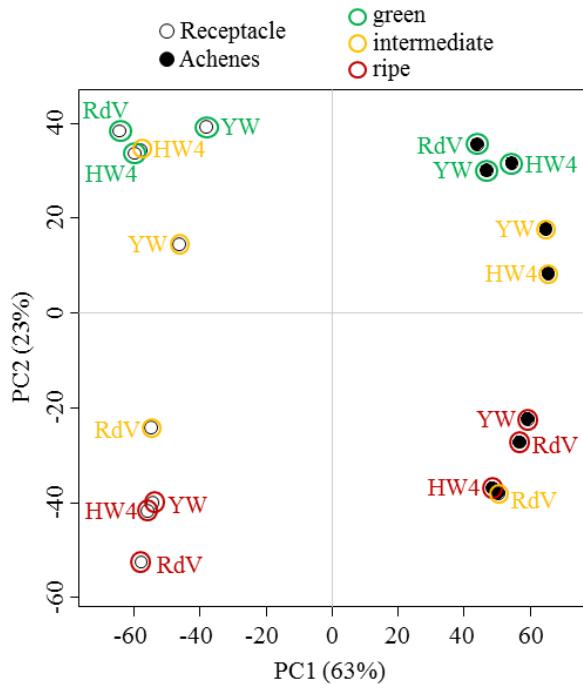
Global mRNA sequencing of the receptacle and achenes of red-fruited *F. vesca* cv. RdV and both white-fruited *F. vesca* varieties YW and HW₄ was performed to investigate the differential accumulation of transcripts. RNA-seq yielded 19208 expressed genes out of 33673 annotated *F. vesca* sequences. The 500 genes showing highest expression were employed for hierarchical clustering and subsequent PCA to visualize the primary variances in the data. According to the PCA, the identity of the tissue (achenes or receptacle) was the strongest factor (Figure 15a). Additionally, the green and ripe developmental stages grouped together, respectively, forming clusters clearly set apart from each other. Interestingly, the intermediate tissues of RdV were more closely related to the ripe tissues of all three genotypes, as against intermediate tissues of YW and HW₄ that clustered with the green tissues.

In order to assess the conservation of expressed genes between genotypes a Venn graph was generated. Therefore, three groups were created by summing up the RPM values of each gene across the samples originating from the same genotype, which equaled six samples per group (RdV, YW and HW₄, respectively). The majority of genes was commonly expressed among the genotypes. In contrast, only few genes were unique (Figure 15b). Consequently, differential expression is more likely to cause the lack-of-color phenotype in the white varieties than presence/absence of expression.

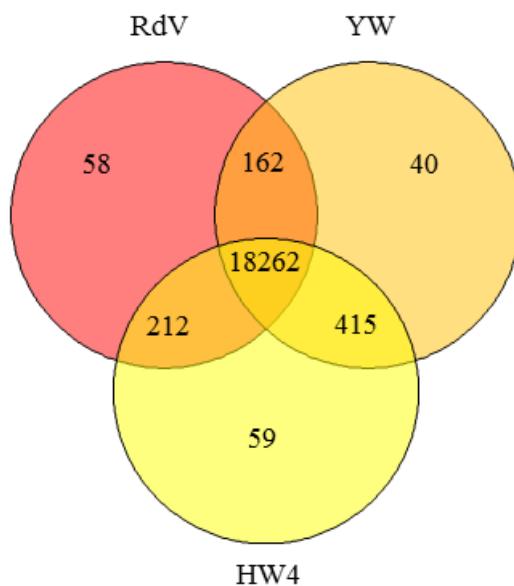
4.2.3 Transcriptome profiling - differentially expressed genes

*Five confirmed
anthocyanin
biosynthesis genes
were differentially
expressed.*

Differential gene expression between the red and the white genotype was determined to find candidates possibly responsible for the loss-of-function phenotype. Overall, 33 genes were significantly down-regulated in the white genotypes (YW and HW₄) compared to the red genotype, among them early (gene26826: chalcone synthase *CHS*, gene21346: chalcone iso-



(a) PCA of top 500 transcripts in *F. vesca* cv. RdV, YW and RdV in two tissues (achenes and receptacle) and three developmental stages (green, intermediate and ripe).



(b) Venn graph showing the number of commonly and uniquely expressed genes among genotypes.

Figure 15: Global gene expression analysis among samples. (adopted from Härtl et al., 2017a)

merase *CHI*, and gene14611: flavanone 3-hydroxylase *FHT*) and late (gene15176: dihydroflavonol reductase *DFR*, gene32347: anthocyanin synthase *ANS*, and gene12591: anthocyanin 3-O-glucosyltransferase *GT1*) functionally confirmed anthocyanin biosynthesis genes (Table 22). Furthermore, gene31672 a predicted GST was among the down-regulated transcripts, as well as two oxophytodienoate reductase-like transcripts (gene12477 and gene25083).

On the other hand, 31 genes were significantly up-regulated in the white genotypes (Table 23). Currently, most candidates lack a functional annotation but are nonetheless featuring a high fold change and overall transcript number, such as gene30960 ($\log_{2}FC=4.6$; 109.65 RPM in RdV, 743.31 RPM in YW and 4,114.10 RPM in HW4). Interestingly, candidate gene27422, functionally annotated as basic helix-loop-helix (bHLH) class ORG2-like transcription factor (TF), is exclusively expressed in the white genotypes and showed the second most highest fold change ($\log_{2}FC=10.3$). Moreover, a predicted ATP-binding cassette (ABC) transporter (gene32014) was among the significantly up-regulated genes.

Table 22: Fold change (logFC) and overall sum of normalized RPM of genes significantly down-regulated in the white genotypes *F. vesca* cv. YW and HW4 compared to the red genotype RdV. (adopted from Härtl et al., 2017a)

GENE ID	LOGFC	SUM RDV	SUM YW	SUM HW4	PREDICTION	CONFIRMED FUNCTION
gene00395	-1.9	731.88	315.17	150.83	protein ZINC INDUCED FACILITATOR-LIKE 1-like (LOC101299619)	
gene01839	-1.9	10,330.44	3,810.76	2,474.66	probable cinnamyl alcohol dehydrogenase 1 (LOC101292655)	
gene04905	-3.7	97.83	50.39	1.57	receptor-like protein 12 (LOC101309371)	
gene05464	-7.8	44.47	1.54	0.00	uncharacterized sequence	
gene06602	-1.7	2,566.42	922.01	709.59	crocin glucosyltransferase, chloroplastic-like (LOC101309923)	
gene07846	-4.2	34.46	46.60	0.00	pentatricopeptide repeat-containing protein At1g12300, mitochondrial-like (LOC101315323)	
gene08163	-5.0	58.15	30.00	0.00	uncharacterized sequence	
gene10142	-5.0	1,112.15	113.26	156.68	2-alkenal reductase NADP(+)-dependent-like (LOC101302097)	
gene12477	-6.1	19.72	0.00	0.24	12-oxophytidienoate reductase 3-like (LOC101293338)	
gene12565	-4.2	1,995.58	591.07	186.81	S-noroclaurine synthase-like (LOC101292845)	
gene12591	-9.2	2,341.19	9.70	5.62	anthocyanin 3-O-glucosyltransferase 2 (LOC101300000)	GT1 (Griesser et al., 2008)
gene12759	-3.4	309.99	369.76	3.36	putative F-box/LRR-repeat protein 23 (LOC101304436)	
gene12884	-4.0	393.87	4.14	94.47	dirigent protein 1-like (LOC101292468)	
gene13009	-4.3	75.94	12.35	2.78	F-box protein CPR30-like (LOC101302499)	
gene14611	-2.4	2,516.77	664.53	763.09	naringenin, 2-oxoglutarate 3-dioxygenase (LOC101300182)	FHT (Almeida et al., 2007)
gene15176	-3.6	715.23	307.29	133.18	bifunctional dihydroflavonol 4-reductase / flavanone 4-reductase (DFR) (LOC101293459)	DFR (Miosic et al., 2014)

Table 22: Down-regulated candidate genes (continued).

GENE ID	LOGFC	SUM RDV	SUM YW	SUM HW4	PREDICTION	CONFIRMED FUNCTION
gene16103	-1.4	771.39	317.72	267.59	pyridoxal kinase (LOC101304704)	
gene16795	-4.7	57.72	53.17	0.00	uncharacterized sequence	
gene17181	-5.7	7.74	0.09	0.00	lysine histidine transporter-like 8 (LOC101292649)	
gene20261	-8.0	45.42	2.53	0.00	TMV resistance protein N-like (LOC101312392)	
gene21346	-2.2	1,208.57	728.57	318.92	probable chalcone-flavonone isomerase 3 (LOC101305307)	<i>CHI</i> (Forkmann, 1991)
gene23269	-2.7	220.80	276.56	3.82	uncharacterized sequence	
gene24010	-4.6	54.15	7.10	9.93	uncharacterized sequence	
gene24179	-1.7	647.85	201.67	160.11	aspartic proteinase Asp1 (LOC101314219)	
gene25083	-4.9	32.68	0.48	7.99	12-oxophytodienoate reductase 2-like (LOC101297812)	
gene26826	-3.2	1,308.73	407.09	245.52	polyketide synthase 1 (LOC101298456)	<i>CHS</i> (Song et al., 2015b)
gene27955	-4.9	69.19	5.79	0.49	receptor-like serine/threonine-protein kinase SD1-8 (LOC101306554)	
gene30678	-4.1	187.53	178.71	0.63	transmembrane protein 184 homolog DDB-G0279555-like (LOC105350765)	
gene31672	-7.2	677.95	17.03	0.49	glutathione S-transferase F11-like (LOC101294111)	
gene32347	-3.3	1,160.82	253.07	205.30	leucoanthocyanidin dioxygenase (LOC101308284)	<i>ANS</i> (Almeida et al., 2007)
gene32421	-3.0	642.27	91.66	87.91	protein P21-like (LOC101300343)	
gene32435	-2.5	551.19	197.64	91.74	short-chain dehydrogenase/reductase zb-like (LOC101296098)	
gene33838	-1.4	290.82	187.98	56.97	AMP deaminase-like (LOC101301583)	

Table 23: Fold change (logFC) and overall sum of normalized RPM of genes significantly up-regulated in the white genotypes *F. vesca* cv. YW and HW4 compared to the red genotype RdV. (adopted from Härtl et al., 2017a)

GENE ID	LOGFC	SUM RDV	SUM YW	SUM HW4	PREDICTION
geneo1275	6.7	0.00	6.97	8.01	uncharacterized sequence
geneo3760	5.3	3.68	66.58	87.32	ceramide-1-phosphate transfer protein (LOC101298698)
geneo4372	4.4	0.88	1.06	150.22	mitochondrial saccharopine dehydrogenase-like oxidoreductase At5g39410 (LOC101312472)
geneo7901	2.4	126.31	632.81	441.40	18.1 kDa class I heat shock protein-like (LOC101300322)
geneo8062	7.1	0.00	9.46	9.99	CDT1-like protein b (LOC101298288)
geneo8537	2.5	81.78	338.72	403.12	uncharacterized LOC101304935 (LOC101304935)
geneo9254	4.5	2.60	7.46	441.69	uncharacterized LOC101298543 (LOC101298543)
gene12602	6.4	0.00	0.47	98.59	uncharacterized LOC101291726 (LOC101291726)
gene12786	3.8	3.87	7.82	138.37	B3 domain-containing transcription factor VRN1-like (LOC101297387)
gene13191	1.5	1,126.99	2,885.38	3,653.64	heat shock protein 83 (LOC101307345)
gene13320	2.1	62.67	207.20	336.93	BCL2-associated athanogene 3 (BAG3)
gene16235	2.5	69.40	231.74	471.49	homeobox-leucine zipper protein ATHB-6-like (LOC101309384)
gene16479	5.0	0.42	15.96	18.85	cysteine synthase-like (LOC101302477)
gene16510	4.4	2.11	26.95	40.82	uncharacterized LOC101294957 (LOC101294957)
gene19533	7.2	0.00	3.52	34.73	putative receptor-like protein kinase At4go0960 (LOC105350176)
gene20844	9.1	0.00	14.51	99.44	uncharacterized LOC105353058 (LOC105353058)
gene20847	10.7	5.77	1,038.42	5,629.92	calmodulin-interacting protein 111-like (LOC101311429)
gene24034	5.2	4.63	67.52	141.62	uncharacterized LOC101294593 (LOC101294593)

Table 23: Up-regulated candidate genes (continued).

GENE ID	LOGFC	SUM RDV	SUM YW	SUM HW4	PREDICTION
gene24512	11.3	0.00	158.67	224.58	uncharacterized LOC101301427 (LOC101301427)
gene24545	2.8	56.32	334.93	225.68	uncharacterized LOC101302298 (LOC101302298)
gene24775	5.8	0.00	3.06	10.38	uncharacterized LOC101301427 (LOC101301427)
gene24779	7.1	0.00	7.66	18.03	uncharacterized LOC101302918 (LOC101302918)
gene26609	2.9	15.71	74.02	140.03	dolichyl-phosphate beta-glucosyltransferase (LOC101312675)
gene27422	10.3	0.00	24.45	425.64	transcription factor ORG2-like (LOC101309207)
gene27944	6.0	0.00	0.31	67.35	uncharacterized LOC101309177 (LOC101309177)
gene27945	4.9	0.25	0.28	208.49	uncharacterized LOC101305101 (LOC101305101)
gene28620	8.7	0.00	20.26	42.71	trifunctional UDP-glucose 4,6-dehydratase/UDP-4-keto-6-deoxy-D-glucose 3,5-epimerase / UDP-4-keto-L-rhamnose-reductase RHM1-like (LOC101302674)
gene29781	7.2	0.00	15.39	11.14	anthocyanidin 3-O-glucoside 2"-O-glucosyltransferase-like (LOC101310006)
gene30676	7.2	0.00	0.09	521.85	uncharacterized sequence
gene30960	4.6	109.65	743.31	4,114.10	uncharacterized sequence
gene32014	5.9	0.00	5.63	3.06	ABC transporter C family member 10-like (LOC101302270)

4.2.4 Transcriptome profiling - flavonoid/anthocyanin pathway

Building upon the fact that five biochemically confirmed anthocyanin biosynthesis genes were among the differentially expressed candidates, the whole flavonoid/anthocyanin pathway was examined starting from phenylalanine (Figure 16). According to their expression profiles the genes could be assigned to four coordinately regulated groups. (I) The first group consisting of *PAL*, *CA4H*, *4CL*, *FLS*, and *FGT* showed high transcript number in green achenes of all genotypes and intermediate achenes of the two white varieties YW and HW4. (II) Transcript abundance of the second group (*CHS*, *CHI*, *F3H*, *DFR*, and *ANS*) was elevated in all samples of RdV but reduced in tissues of YW and HW4 except green achenes, and green and intermediate receptacle. (III) The third group showed the most significant differences in expression between the red- and white-fruited varieties. *GT1* and a *GST* were almost exclusively expressed in intermediate and ripe tissues of RdV. (IV) *F3'H*, *ANR*, and *LAR* formed the fourth group, which is characterized by high transcript abundance in green tissues of all three genotypes and intermediate receptacle of HW4. The expression profile of *GT2* did not fit into one of the four groups, as its transcripts were primarily abundant in green achenes of all genotypes and ripe receptacle of RdV.

The flavonoid/anthocyanidin pathway genes could be assigned to four coordinate groups.

4.2.5 Transcriptome profiling - flavor formation and fruit softening

The hindered anthocyanin pathway may affect other fruit ripening related biosynthetic pathways. Therefore, transcript abundance of genes acting in fruit flavor formation and softening was analyzed (Figure 17). Pinene synthase and hydroxylase (Aharoni et al., 2004) are associated with volatile terpene formation, whereas acyltransferases FcAAT1 and SAAT (Aharoni et al., 2000; Cumplido-Laso et al., 2012) take part in ester production. QR (Raab et al., 2006) and eugenol synthase (Aragüez et al., 2013) facilitate furaneol and eugenol formation, respec-

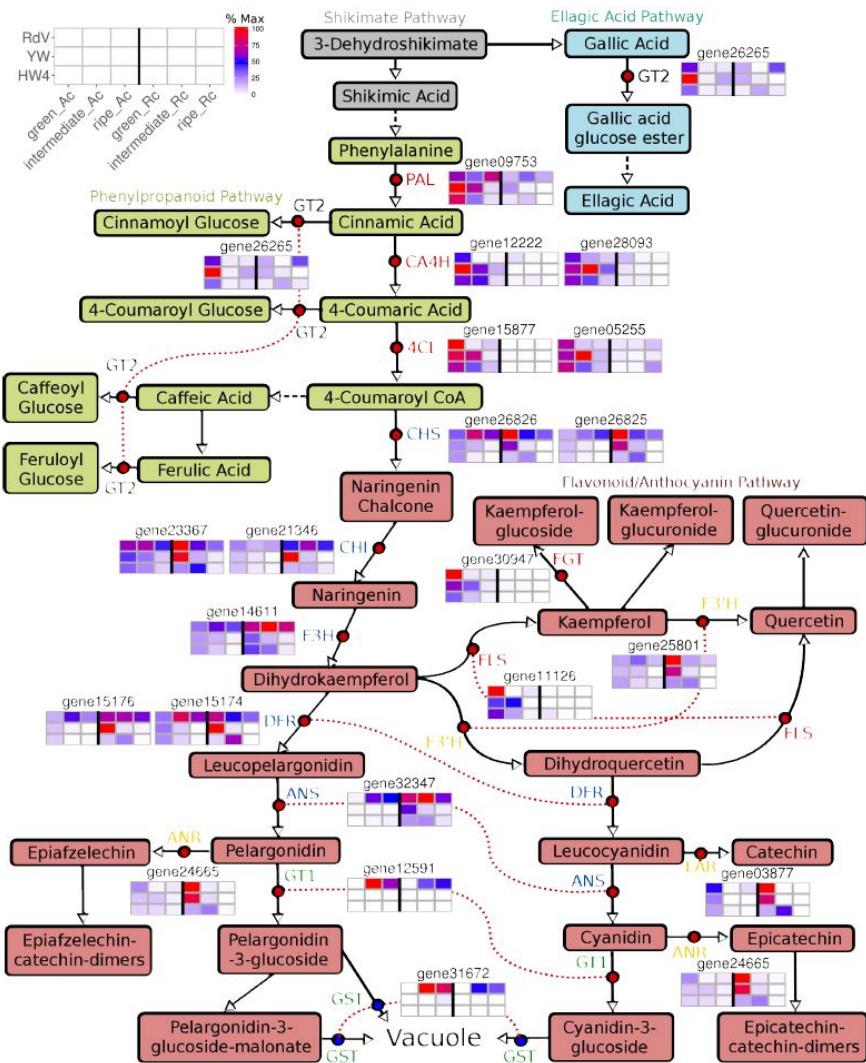


Figure 16: Scheme of the flavonoid/anthocyanin pathway in strawberry fruit. Red dots indicate biochemically characterized enzymes in strawberry fruit: ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; CA4H, cinnamic acid 4-hydroxylase; CHI, chalcone isomerase; CHS, chalcone synthase; 4CL, 4-coumaroyl-CoA ligase; DFR, dihydroflavonol reductase; FGT, flavonoid glucosyltransferase; F₃H, flavanone 3-hydroxylase; F_{3'}H, flavonoid 3'-hydroxylase; FLS, flavonol synthase; GT1, anthocyanidin glucosyltransferase; GT2, (hydroxy)cinnamic acid and (hydroxy)benzoic acid glucosyltransferase; LAR, leucoanthocyanidin reductase; PAL, phenylalanine ammonia lyase. Blue dots indicate putative GST, glutathione S-transferase. Identical genes are connected by a dotted red line when adjacent. Enzymes shown in the same color are co-regulated. Heatmaps show relative transcript levels (% Max) of genes in receptacle (Rc) and achenes (Ac) of *F. vesca* cv. RdV, YW and HW4 at the green, intermediate and ripe developmental stage. (adopted from Härtl et al., 2017a)

tively, and pectin esterase (Castillejo et al., 2004), pectate lyase (Benítez-Burraco et al., 2003), polygalacturonase (Redondo-Nevado et al., 2001) and beta-galactosidase (Paniagua et al., 2016) control fruit softening.

The corresponding genes showed coordinate expression, which correlated to the fruit ripening in receptacles of all three *F. vesca* varieties. Their expression almost always peaked in ripe receptacles. However, in intermediate receptacle of the red genotype RdV transcripts were already abundant, in contrast to the respective tissue of the white genotypes. Here, transcripts only increased in ripe receptacles and for some genes the values even exceeded those reached in RdV. Apparently, the ripening process in the white varieties is slowed down, as expression of genes related to flavor and cell wall degradation is delayed.

Gene expression was delayed in the white varieties.

4.2.6 Expression of *FvGT2* and *FvGT5* in *F. vesca*

By means of the transcriptome analysis mRNA-expression of the functionally characterized UGTs *FvGT2* (gene26265) and *FvGT5* (gene26249, Section 4.1) was evaluated. Transcript levels corresponding to *FvGT2* were highest in green achenes of all three woodland strawberry genotypes, decreased towards the intermediate stage and increased again slightly in ripe achenes (Figure 18). The same pattern was found in the receptacle samples of the red-fruited variety RdV. On the contrary, *FvGT2*-expression in receptacles of YW and HW4 did not reach a second peak in the ripe developmental stage.

The transcript levels of *FvGT5* exhibited considerable differences in the achenes of the red- compared to the two white-fruited varieties. In achenes of RdV the expression pattern was alike the one observed for *FvGT2* showing a two-phase pattern with increased levels in the green and ripe developmental stages, whereas in achenes of HW4 a ripening correlated trend was observed. In YW increased *FvGT5*-expression was only de-

*In RdV *FvGT2* expression exhibited a biphasic pattern.*

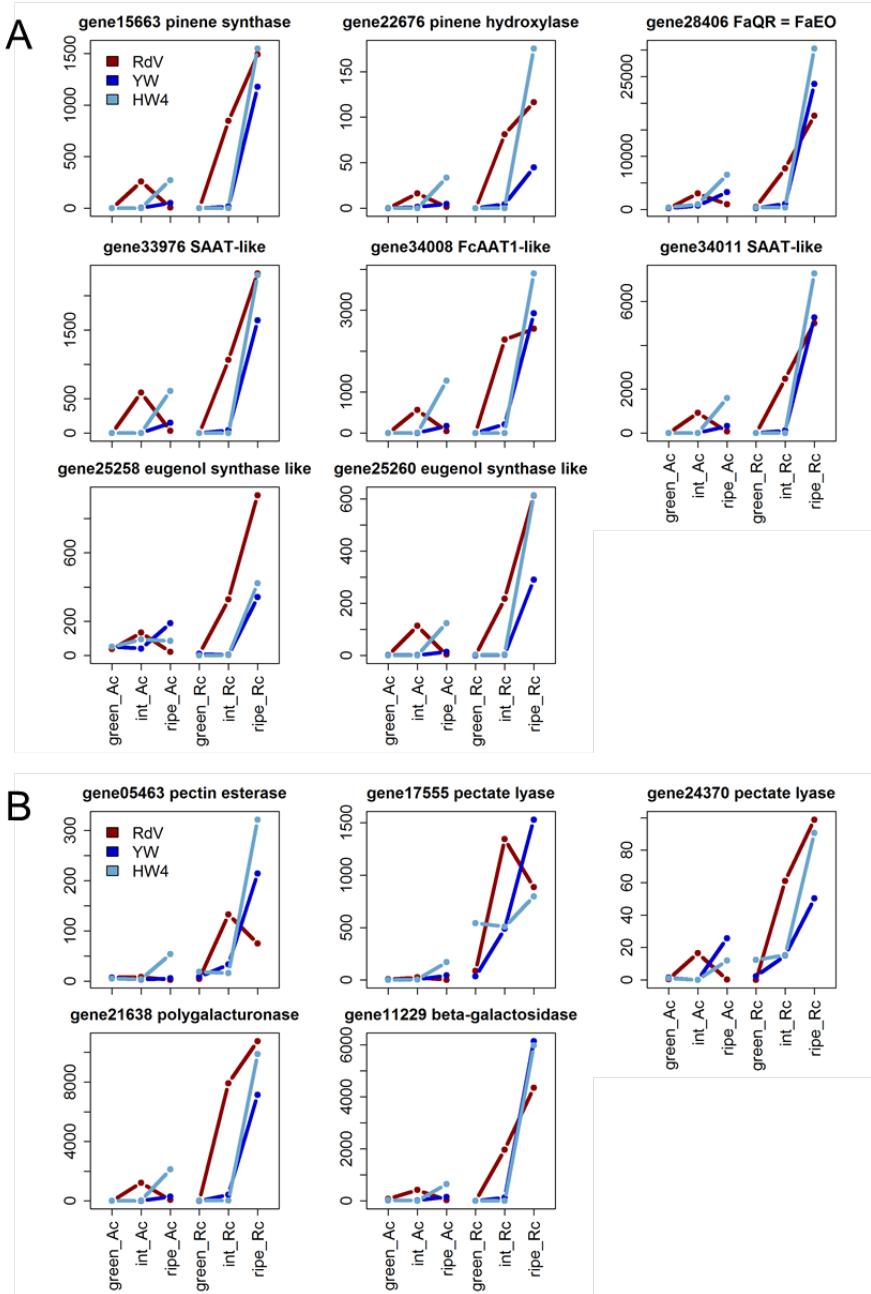


Figure 17: Transcript levels (normalized RPM) of genes encoding enzymes involved in (A) flavor formation and (B) fruit softening in receptacle (Rc) and achenes (Ac) of *F. vesca* cv. RdV, YW and HW4 at the green, intermediate (int) and ripe developmental stage. (adopted from Härtl et al., 2017a)

tected in intermediate achenes. The overall levels in receptacles did not show remarkable changes among the three genotypes.

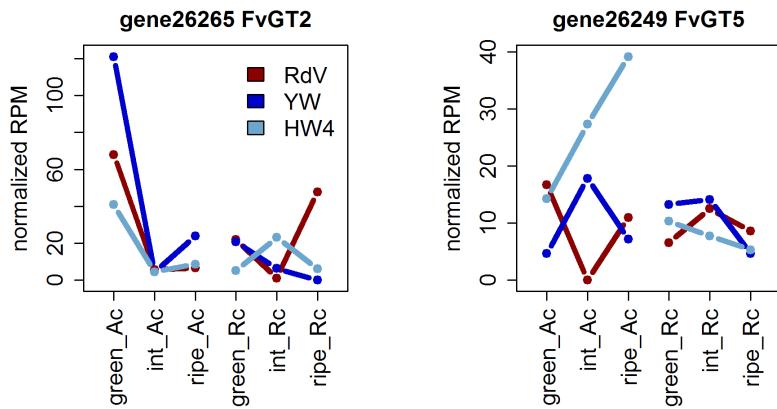


Figure 18: Transcript levels of *FvGT2* and *5* (normalized RPM) in receptacle (Rc) and achenes (Ac) of woodland strawberry varieties RdV, YW, and HW4 at the green, intermediate (int) and ripe fruit stage.

4.3 FORMATION OF 1,2,3,4,6-PENTAGALLOYLGLUCOSE

4.3.1 Identification of AT gene candidates

It has been shown in wild tomato (*Lycopersicon pennellii*), thale cress (*Arabidopsis thaliana*) and rape (*Brassica napus*) that hydrolases of the serine carboxypeptidase type can catalyze the formation of ester bonds (Lehfeldt et al., 2000; Li and Steffens, 2000; Li et al., 1999; Milkowski et al., 2004; Shirley et al., 2001). These serine carboxypeptidase-like (SCPL) acyltransferases are able to facilitate acyl transfer between two molecules of 1-O- β -glucose esters, without involvement of CoA-thioesters (Milkowski and Strack, 2004). PGG, the precursor of EA is supposedly formed from multiple molecules of β -glucogallin by trans-esterification reactions, which could well be mediated by SCPL acyltransferases. Metabolite analyses in garden and woodland strawberry (Section 4.1.5, Section 4.2) revealed green unripe fruit tissue as main source of EA, GA, and β -glucogallin. To search for candidates with high sequence homology, the sequences of known acyltransferases were blasted against the *F.*

Homologues with high expression in green tissues were selected.

vesca genome. Additionally, by analysis of the RNA-seq data set the list of candidate genes could be further narrowed down. Finally, the genes FvAT1 (gene25429), FvAT2 (gene25430), and FvAT3 (gene25433) were selected, because according to the transcriptome data their expression is highest in green, unripe tissue and decreases during fruit ripening in at least one of the three strawberry varieties (Figure 19).

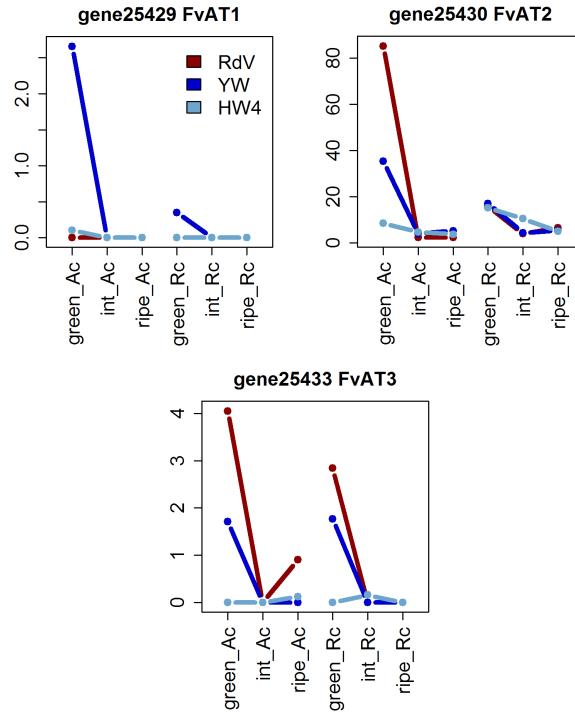


Figure 19: Transcript levels of *FvAT1*, 2 and 3 (normalized RPM) in receptacle (Rc) and achenes (Ac) of woodland strawberry varieties RdV, YW and HW4 at the green, intermediate (int) and ripe developmental stage.

4.3.2 In vitro characterization of acyltransferase candidates

Protein expression could not be detected.

Subsequently, the candidate ATs were cloned from cDNA of green achenes. Different sets of primers were tested, as well as various polymerases, cDNA templates and PCR cycling conditions. The *FvAT1* and *FvAT3* genes were successfully amplified, while *FvAT2* resisted cloning. Meanwhile, the *F. vesca* gene models were updated on the NCBI database ac-

cording to the new strawberry annotation release 101. During the process, the 5,661 bp long *FvAT2* sequence was split into four individual genes *FvAT2.1* (XM 004289722, gi 101305898), *FvAT2.2* (XM 004289723, gi 101306187), *FvAT2.3* (XM 011459530, gi 105349275) and *FvAT2.4* (XM 011459533, gi 101307458) comprising a length of 1,404 to 1,419 bp (Figure 20). The protein sequences of *FvAT2.1-2.4* share a pairwise identity of more than 61 %, while *FvAT2.1-2.3* even share more than 77 % homology. Consequently, new primers were designed to match the newly set coding regions. However, still no sequences could be obtained.

On the other hand, the *FvAT1* and β genes were successfully cloned into the pYES2 expression vector and transformed into the *S. cerevisiae* strain INVSc1 (Section 3.4.1). The protein expression was induced by adding galactose, but the recombinant *FvAT1* and β proteins could neither be detected by SDS-PAGE nor by Western blot. Thereupon, different incubation temperatures, expression times and galactose concentrations were tested to improve the production of the recombinant proteins. When these attempts did not lead to increased expression, the ATs were re-cloned without His-tag to ensure correct protein folding. This approach failed either.

4.3.3 Up-regulation of *FvAT1* and β in strawberry fruits

In parallel to the attempts to characterize *FvAT1* and β *in vitro*, both sequences were cloned into the pBI12135S2x vector and transformed into the *A. tumefaciens* strain AGLo (Section 3.4.1). The transformed *Agrobacteria* were injected into ripening *F. × ananassa* cv. Elsanta fruits to induce gene-transfer mediated over-expression of the candidates. *Agrobacteria* transformed with a pBI121 vector containing a GUS intron gene served as a control (Hoffmann et al., 2006). After two weeks, the fully ripened fruits were harvested, frozen and ground to fine powder. Aliquots of the fruit powder were used to isolate the RNA, to reverse transcribe it, and to validate the *FvAT1*

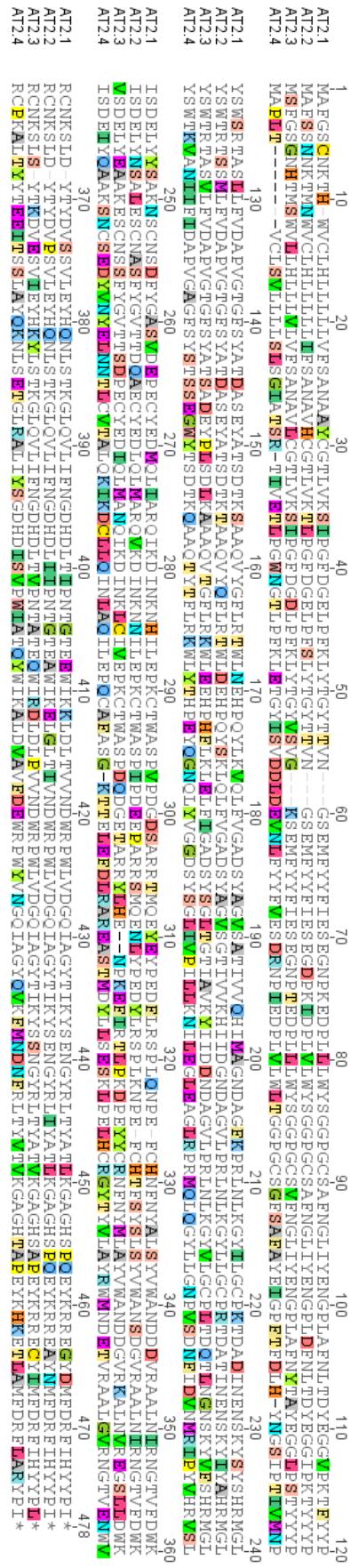


Figure 20: Protein sequence alignment of SCPL ATs FvAT2.1, FvAT2.2, FvAT2.3, FvAT2.4. The ClustalX function integrated in the GENEOUS Pro 5.5.6 software (Biomatters; <http://www.geneious.com/>, accessed August 2016) was employed in default parameters, to generate this alignment.

and β over-expression by RT-PCR. On the other hand, the same fruit powder was subjected to extraction and quantification of β -glucogallin, di-galloyl glucose and EA.

Via RT-PCR the over-expression of both ATs was confirmed, albeit an effect on the phenotype was not observed (Figure 21). Untargeted metabolite analysis revealed 89 compounds that accumulated significantly different in the *FvAT1* and β over-expressing fruits compared to the *GUS* intron control fruits. Out of these compounds, only EA and di-galloyl glucose could be identified according to their retention time, mass spectra and corresponding MS² spectra. Subsequent targeted analysis of β -glucogallin, di-galloyl glucose and EA yielded low and constant levels of β -glucogallin in the *FvAT1* and β over-expressed, as well as in the control fruits (Figure 22). Di-galloyl glucose was found at an equal level in the *FvAT1*- and *GUS*-treated fruits but was clearly reduced in the *FvAT3*-transfected tissues. The relative EA concentration was in both, *FvAT1* and β , significantly increased compared to the control.

EA production was elevated in transfected fruits.

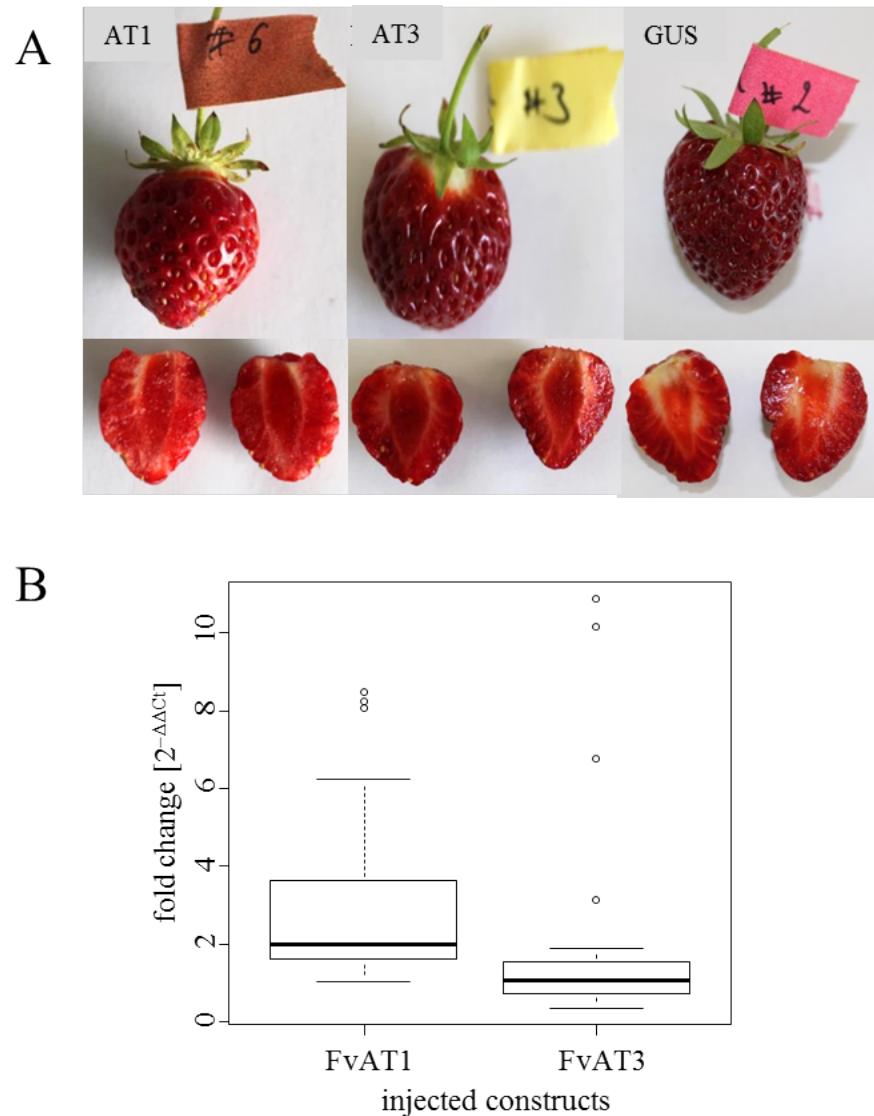


Figure 21: Over-expression of *FvAT1* and β . (A) Phenotype of *F. × ananassa* cv. Elsanta fruits, transfected with over-expression constructs of *FvAT1*, β and GUS control. (B) Fold change ($2^{-\Delta\Delta Ct}$) of *FvAT1* and β expression in the transgenic fruits is corrected by the AT transcript level in the control fruits.

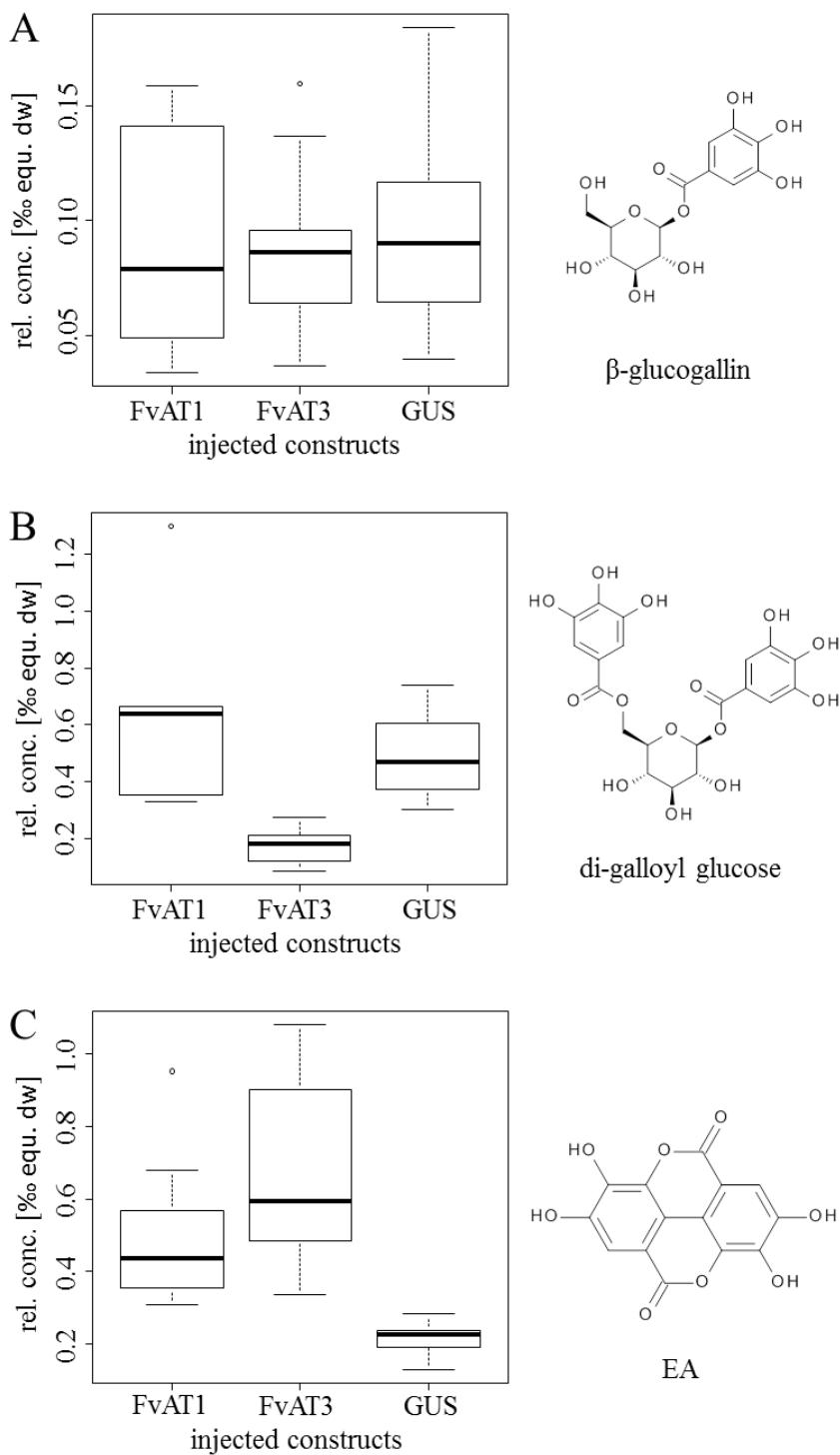


Figure 22: Relative concentration (rel. conc.) of β -glucogallin (A), di-galloyl glucose (B) and EA (C) in *FvAT1*, *FvAT3* and *GUS* transfected *F. × ananassa* cv. Elsanta fruits in %_o equivalents of the dry weight (%_o equ. dw). n=15-20 fruits.

5

DISCUSSION

5.1 CATALYSIS OF GALLOYL GLUCOSE BY UGTS

5.1.1 *Gallic acid UGTs in F. vesca, F. x ananassa, and R. idaeus*

UGTs that transfer sugar units from activated donor substrates to specific acceptor compounds form a vast class of carbohydrate-active enzymes (CAZymes) (André et al., 2014). They can be classified according to their structural fold, and according to their mechanism of stereo-chemical bond formation at the anomeric carbon atom of the product (Lairson et al., 2008). FaGT₂, FaGT₂^{*}, FvGT₂, RiGT₂, FaGT₅ and VvgGT₁ belong to the family 1 UGTs, as they glycosylate small lipophilic molecules (<http://www.cazy.org/>, accessed July 2016; Bowles et al., 2005; Sawada et al., 2005). They are inverting enzymes, which means that they permute the stereo-chemistry at the anomeric center of the donor substrates. Furthermore, they adopt the GT-B protein fold forming the horseshoe-like active cleft with two so-called Rossmann folds facing each other (Moréra et al., 2001; Vrielink et al., 1994). Within the active cleft the β -glucogallin producing UGTs comprise the PSPG box, a conserved C-terminal motif that is associated to plant family 1 UGTs and interacts with the sugar donor (Bowles et al., 2006).

It has been shown that UGTs that cluster into phylogenetic groups often exhibit similar functionality (Gachon et al., 2005; Ross et al., 2001). The candidate enzymes analyzed yield 1-O-acyl-esters of (hydroxy)benzoic and (hydroxy)cinnamic acids, while showing a high degree of structural relatedness. Thus, the phylogenetic classification scheme can be applied. However, the GT₂-homologous enzymes from strawberry and raspberry possess a C-terminal extension of \approx 80 amino acids (Figure 7); a feature they have in common with a hydroxybenzoic

acid UGT from *Q. robur* (UGT84A13, Mittasch et al., 2014) that also catalyzes β -glucogallin. The significance of the additional sequence remains elusive.

5.1.2 Enzymatic activity and biochemical characterization

The GT2-enzymes showed a clear bias for tri-functionalized acceptor substrates, as they developed high affinity (low K_M) for gallic acid, sinapic acid and syringic acid (Table 21). RiGT2 and VvgGT1 produced β -glucogallin with the highest enzyme efficiency (k_{cat}/K_M) of $20,000\text{ s}^{-1}\text{ M}^{-1}$ and $15,277\text{ s}^{-1}\text{ M}^{-1}$, respectively. These values clearly do not meet the specificity constant of UGT84A13 from *Q. robur*, which converts gallic acid with an efficiency of $25,952\text{ s}^{-1}\text{ M}^{-1}$ (Mittasch et al., 2014). However, they exceed the published values of VvgGT1-VvgGT3 from *V. vinifera* of $4,231$, $3,194$ and $3,873\text{ s}^{-1}\text{ M}^{-1}$, respectively (Khater et al., 2012). In contrast to the GT2-enzymes, FaGT5 preferred sinapic acid as acceptor substrate. Sinapic acid can be found in strawberry fruits in high amounts ($450\text{ }\mu\text{g g}^{-1}$ dw; Russell et al., 2009), as well as its glucose derivatives (Fait et al., 2008). The sinapic acid glucose derivatives were found to accumulate in receptacle tissues along fruit ripening, and are thought to contribute to the development of nutrient transport vessels (Fait et al., 2008; Humphreys and Chapple, 2002). Albeit, the authors did not distinguish between glucosides and glucose esters. In contrast, Lunkenstein et al., 2006a quantified sinapoyl-D-glucose ester exclusively in white and turning strawberry fruits. However, the authors did not separately analyze receptacle and achenes. Despite the accumulation of sinapic acid glucose derivatives being a matter of discussion, one of the putative *in vivo* FaGT5 functions might be the formation of precursors for the biosynthesis of complex phenylpropanoids such as lignins.

Site-directed mutagenesis of FaGT2* demonstrated that amino acids can also have a notable effect on the substrate specificity of an enzyme, even if they are not directly part

of the active cleft (Figure 8). Similar observations have been noted by Bönisch et al., 2014b for a monoterpenol UGT from *V. vinifera*. Both FaGT2* mutants exhibited enhanced activity towards a number of substrates including gallic acid, *o*-coumaric and caffeic acid (Figure 8). Interestingly, the double mutant FaGT2*-Glu420Asp-Ile422Val showed improved turnover rates compared to the single mutant FaGT2*-Arg230Ser. It seems that the gradual amino acid exchanges procured a step-wise adjustment of the FaGT2* enzyme specificity to the FaGT2 activity (Table 20).

5.1.3 Correlation of UGT expression pattern and metabolite levels

Green achenes have been identified as the main source tissue of GA, β -glucogallin and EA. Accordingly, in the red-fruited strawberry variety RdV *FvGT2*-mRNA (gene26265) was most abundant in green achene tissue, decreased towards the white stage, and reached a second peak in ripe receptacle tissue (Figure 18). In addition, the garden strawberry *FaGT2* has a ripening correlated expression, as well as a preference for the acceptor substrate cinnamic acid (Lunkenbein et al., 2006a). Therefore, the GT2-enzymes seem to have two functions: in green fruits they glycosylate gallic acid to produce the precursor of ellagic acid and, as the level of ETs/EA decreases in ripe fruits, they glycosylate cinnamic acid. On the contrary, in the receptacle samples of the white-fruited *F. vesca* genotypes a two-phase pattern of *FvGT2*-expression was not detected, which is reflected in the metabolite data as a lower level of cinnamic acid glucose ester was detected in ripe receptacles of the white-fruited varieties compared to the corresponding RdV tissue (Figure 14). Several UGTs are known that show promiscuity regarding their substrate specificity (Hall and Luca, 2007), and it is believed that this ability is a result of constant adaption to the ever changing environment (Ulusu, 2015). Evolution seems to also have broadened the putative *in vivo* function of FvGT5 that could even diverge among genotypes. Especially in ach-

enes of YW transcript levels were high, as opposed to low levels of sinapic acid glucose derivatives in strawberry achenes (Fait et al., 2008). Consequently, in achenes of YW the FvGT5 enzyme could catalyze the esterification of other metabolites such as veratric acid, which was also esterified *in vitro* with a high relative activity (92 %, Table 20). While veratric acid has been detected in blueberry (Zadernowski et al., 2005), its presence in strawberry has not been verified until now. However, the structurally related methoxyeugenol exhibits a similar methylation pattern and has been identified in ripe strawberry fruit (Hoffmann et al., 2011).

5.1.4 *FaGT2/FaGT2** *in vivo* activity

After injection of d₂-labeled gallic acid (Section 4.1.6) into green fruits of *F. × ananassa* cv. Calypso control plants and fruits of stable transgenic *FaGT2*-silenced plants, the levels of EA precursors were quantified. Levels of d₂-GA were elevated in the transgenic fruits, most likely due to a diminished *FaGT2* activity. On the other hand, d₂-EA formation was reduced albeit not completely. Lunkenbein et al., 2006a, who propagated the transgenic strawberry plants, recorded 34 % residual expression for *FaGT2*, which may explain the reduced formation of EA. *FaGT5* showed only base activity towards GA but probably also contributed to the production of β-glucogallin (Table 20). Interestingly, the levels of d₂-β-glucogallin and d₂-di-galloyl glucose remained unchanged in the transgenic fruits compared to control fruits. The remaining *FaGT2* activity seems to perpetuate the flux towards EA but with severely reduced efficiency. This led to reduced levels of EA and the formation of a new compound not reported in strawberry fruit until now. The novel metabolite was identified as gallic acid 4-O-glucoside (Schuster and Herrmann, 1985). *In vitro* assays showed that this compound is synthesized by UGT71K3 (Figure 11), an enzyme that has been shown to also glucosylate the structurally related compound pyrogallol (Song et al., 2016b). Consequently, the excess

d₂-GA was directed towards glucoside production as a result of the ester formation being inhibited.

5.2 WOODLAND STRAWBERRY METABOLOME AND TRANSCRIPTOME

The diploid woodland strawberry *F. vesca* is an ideal system to study the genetic basis of important traits such as color in the *Rosaceae* family, because in addition to a number of red-fruited strawberry varieties also natural white-fruited varieties are commercially available (Kaiser et al., 2016). In order to reveal metabolic and genetic differences in receptacle and achenes of the red-fruited *F. vesca* genotype RdV and two white-varieties HW4 and YW comparative metabolite profiling and transcriptome analyses were performed.

5.2.1 Metabolite Analysis

Untargeted and targeted metabolite profiling (Section 4.2.1) of RdV tissues showed significant spatial and temporal differences in secondary metabolite concentration already at an early ripening stage. Polyphenol accumulation in RdV receptacle samples exhibited two peaks, as proanthocyanidins were most abundant at the intermediate stage and anthocyanins, quercetin glucuronide, phenylpropanoids and EA at the ripe developmental stage. In achene tissues the metabolite accumulation followed a tri-phasic profile. Phenols were increased at the green stage, flavonoids (except quercetin glucuronide) and proanthocyanidins at the intermediate stage, and anthocyanins at the ripe stage. In contrast, the polyphenol levels in tissues of the white varieties showed a mixed pattern, and also differences between YW and HW4 could be observed.

Four anthocyanins convey the color in red strawberries: pelargonidin 3-glucoside, pelargonidin 3-glucoside 6'-malonate, pelargonidin 3-rutinoside and cyanidin 3-glucoside (Kosar et al., 2004; Muñoz et al., 2011; Simirgiotis et al., 2009; Tuli-

pani et al., 2008). Consistent with the results of the targeted metabolite profiling (Figure 14), those compounds were significantly reduced in the naturally white-fruited varieties (Cheel et al., 2005; Salvatierra et al., 2013; Simirgiotis and Schmeda-Hirschmann, 2010). However, this difference does not account for the primary variance in metabolite accumulation between RdV and the two white-fruited genotypes YW and HW4, as PCA grouped the tissues according to the developmental stage rather than the variety (Figure 13).

5.2.2 Global transcriptome data analysis and evaluation of differential gene expression

In accordance with the metabolite profiles, global RNA-seq data analysis also indicated early disturbances in the ripening process of YW and HW4, as variance in gene expression of these white-fruited genotypes compared to red-fruited RdV is highest within the intermediate ripening stages (Figure 15).

Major anthocyanin/flavonoid biosynthesis genes were found differentially expressed (Table 22). Among the down-regulated genes were *CHS* and *CHI*, known to encode enzymes that catalyze the formation of naringenin, the first product featuring the characteristic flavonoid C-15 ring structure (Koes et al., 2005). *CHS* assumes a key position in the pathway, as its transient and stable down-regulation led to color mutants in fruits and flowers of various plant species (Deroles et al., 1998; Que et al., 1998), including the strawberry (Hoffmann et al., 2006; Lunkenbein et al., 2006b; Song et al., 2015b). Furthermore, *F3H* and *DFR* expression was reduced, whose encoded proteins are responsible for the oxidation of naringenin to dihydrokaempferol (*F3H*, Forkmann, 1991) and subsequent reduction to leucoanthocyanidins (*DFR*, Wang et al., 2013). Likewise, transcripts of *ANS* and *GT1* were significantly down-regulated, thereby impairing the production of colored anthocyanidins (Abrahams et al., 2003) and their glucosylation (Griesser et al., 2008). The resulting glucosylated anthocyanins,

which are responsible for the color of flowers and fruits, are accumulated in the plant vacuoles (Grotewold, 2006). Moreover, transcript abundance of a putative *GST* (gene31672) was significantly reduced in the white genotypes (Table 22). It has been shown in several plants, including *Arabidopsis* (Kitamura et al., 2004), maize (Marrs et al., 1995) and petunia (Mueller et al., 2000) that GSTs can presumably act as carriers for the transport of anthocyanins to the vacuole. Also, *GST* Solyc02g081340 from tomato (*Solanum lycopersicum*), orthologous to gene31672, was associated with anthocyanin synthesis (Qiu et al., 2016; Tohge et al., 2015; Zhang et al., 2016). Therefore, the *GST* candidate gene31672 might act in pigment transport. Along with the many characteristics of the candidate gene, its expression profile showed striking co-regulation to *GT1* (Figure 16) and thus validates its putative function. Furthermore, the two putative 12-oxophytodienoic acid reductases (OPRs) gene12477 and gene25083 were among the down-regulated transcripts. OPR3 from *A. thaliana* catalyzes the peroxisome located conversion of (9*S*, 13*S*)-12-oxophytodienoic acid to 3-2(2'(*Z*)-pentyl)cyclopentane-1-octanoic acid, a crucial step in the biosynthetic pathway of the phytohormone jasmonic acid and its derivatives (Sanders et al., 2000; Stintzi and Browse, 2000). Jasmonates derive from cyclopentanone and are synthesized in the octadecanoid pathway (Delker et al., 2006). They are known to function in plant development and growth (Cheong and Choi, 2003) but also in defense mechanisms against insects and fungal pathogens (McConn et al., 1997; Vijayan et al., 1998). Recently, the jasmonic acid derivative methyl jasmonate has been associated to fruit ripening and coloration in grapevine (*V. vinifera*, Jia et al., 2016). Methyl jasmonate-treated developing fruits exhibited increased levels of anthocyanin pigments as well as elevated transcripts of ripening-related genes, such as *PAL*, *DFR*, *CHS*, *CHI*, *F3H*, *FGT*, and *GST*. Interestingly, in addition to softening- and aroma-related genes also the jasmonic acid biosynthesis gene OPR3 was among the up-regulated transcripts. Likewise, our results suggest a relationship between fruit coloration and expression

of the putative OPRs gene12477 and gene25083. However, the expression of a functional jasmonic acid carboxyl methyl-transferase is negatively correlated with the strawberry fruit development (Preuß et al., 2014).

Differential expression analysis also revealed a number of genes significantly up-regulated in the white genotypes (Table 23) such as gene27422, an annotated bHLH TF. It is known that the expression of anthocyanin biosynthesis genes in plants is regulated by a protein complex consisting of myeloblastosis (MYB) family, bHLH and WD40-repeat TFs (Allan et al., 2008; Jaakola, 2013). Limited knowledge is available about plant bHLH class TFs. They function in anthocyanin biosynthesis, as well as in phytochrome signaling, fruit dehiscence and carpel/epidermal development (Schaart et al., 2013; Vallarino et al., 2015). Therefore, the TF presumably encoded by candidate gene27422 could regulate pigment formation in strawberry. MYB proteins have been studied more intensively. In strawberry, MYB1 (gene09407) was characterized as transcriptional repressor of the anthocyanin biosynthesis (Aharoni et al., 2001; Kadomura-Ishikawa et al., 2015; Salvatierra et al., 2013). According to our data the respective gene did not exhibit differential expression. Therefore, MYB1 might not account for the lack-of-color phenotype in the white-fruited strawberry varieties analyzed here. As activator of pigment formation in strawberry, MYB10 (gene31413) was recently introduced (Lin-Wang et al., 2014; Medina-Puche et al., 2014). Moreover, it was suggested that a single amino acid substitution in the MYB10 protein could be responsible for the lack-of-color phenotype in YW and HW4 (Hawkins et al., 2016). However, the transcript and metabolome profiling (Section 4.2.1 and Section 4.2.2) revealed variations in the ripening process of these genotypes already at the green stage, where MYB10 transcripts are very rare. Consequently, it can be concluded that additional MYB TFs might govern the biosynthesis pathway of anthocyanins in strawberry. In addition, a predicted ABC transporter C family member (gene32014) was found significantly up-regulated. Recently, it has been shown that ABC transporters, in particular

from the subfamily C, are able to mediate the transport of anthocyanins into the vacuole (Buer et al., 2007; Francisco et al., 2013; Goodman et al., 2004). The reason for the up-regulation of this candidate ABC transporter in strawberry genotypes lacking anthocyanins remains elusive. However, the minor overall transcript levels of gene32014 detected in RdV, YW and HW4 (0.00 RPM, 5.63 RPM and 3.06 RPM, respectively; Table 23) suggest other ABC transporter genes to be involved in the transport of strawberry pigments. Amongst the up-regulated genes were candidates equally high expressed in both white genotypes, such as gene30960 and gene13191 but also candidates exclusively expressed in either one or the other, such as gene30676 and gene04372. This indicates not only transcription-related differences among the red- and white-fruited varieties but also among the two white genotypes.

5.2.3 Transcript levels of functionally confirmed anthocyanin/flavonoid biosynthesis genes

The anthocyanin/flavonoid biosynthesis genes could be classified in coordinately regulated groups according to the expression profiles. Early pathway genes (*PAL*, *CA4H* and *4CL*), pivotal for overall polyphenol precursor formation, were mainly expressed in immature seeds of the red and white fruit varieties, respectively (Figure 16). Co-regulation of these genes has also been reported in parsley leaves and cell cultures when exposed to stress conditions (Logemann et al., 1995). Downstream genes *CHS*, *CHI*, *F3H*, *DFR* and *ANS*, involved in the provision of precursors for proanthocyanidin, flavonoid and anthocyanin formation, were co-expressed amongst each other. This has been also shown in apple fruits for *CHS*, *F3H*, *DFR* and *ANS* during ripening-related peel coloration (Honda et al., 2002). The transcripts of these flavonoid genes accumulated mostly in green pulp of YW, where they contribute to the high levels of flavonoids and proanthocyanidins (Figure 14). The co-regulated transcript profiles of *GT1* and *GST* were striking.

Anthocyanidin UGTs are most critical for fruit coloration in various plants, including strawberry (Griesser et al., 2008), grape (Kobayashi et al., 2001), apple (Honda et al., 2002) and litchi (Wei et al., 2011). Low expression of these genes result in absence or extremely low pigment concentrations. *GT2* was mainly expressed in green tissues of all genotypes and red receptacle of RdV. It might assume a dual function as GA UGT in early fruit development (Schulenburg et al., 2016a) and (hydroxyl)cinnamic acid UGT in late stages (Lunkenbein et al., 2006a).

5.2.4 *Interaction of polyphenol metabolism, fruit flavor formation and softening*

A connection between fruit flavor and polyphenol metabolism has been observed frequently, as phenolic compounds can serve as precursor molecules in flavor biosynthesis (Koeduka et al., 2006). Consequently, the expression levels of the confirmed flavor biosynthesis genes pinene synthase and hydroxylase (gene15663 and gene22676; Aharoni et al., 2004), alcohol acyltransferases (AATs) (gene33976, gene34011 and gene34008; Aharoni et al., 2000; Cumplido-Laso et al., 2012), QR (gene28406; Raab et al., 2006) and eugenol synthases (gene25258 and gene25260; Aragüez et al., 2013) were analyzed in the red and the two white fruit varieties (Figure 17). It appears that expression of these flavor genes is delayed in the white genotypes, as transcripts were only abundant in ripe receptacles of YW and HW4, in contrast to RdV, where expression already increased in intermediate receptacles. This finding is consistent with the PCA analysis (Figure 15a), as intermediate tissues of the white varieties clustered with green tissues and intermediate tissues of RdV with ripe tissues. Moreover, the expression profiles of the functionally characterized fruit softening genes pectin esterase (gene05463; Castillejo et al., 2004), pectate lyases (gene17555 and gene24370; Benítez-Burraco et al., 2003), polygalacturonase (gene21638; Redondo-

Nevado et al., 2001) and β -galactosidase (gene11229; Paniagua et al., 2016) corresponded with this assumption. Above all, the genes related to fruit flavor formation and softening are co-regulated.

5.3 IDENTIFICATION AND CHARACTERIZATION OF PUTATIVE ATS

5.3.1 *Trans-esterification mediated by serine carboxypeptidase-like (SCPL) acyltransferases*

The formation of β -glucogallin from GA and UDP-glucose was first shown in enzyme assays containing protein extracts from oak leaves (Gross, 1983a). Interestingly, within the same preparations the authors observed the accumulation of di- and tri-galloyl glucose molecules - a reaction that would normally require energy-rich co-factors. Consequently, it was assumed that in this reaction β -glucogallin serves as "activated" compound, thereby acting as acyl-donor and -acceptor likewise (Gross, 1983b). Furthermore, it was revealed that this co-factor-free mechanism applies to every acyl-transfer step from β -glucogallin towards the formation of PGG (Cammann et al., 1989; Gross and Denzel, 1991; Hagenah and Gross, 1993; Schmidt et al., 1987).

Most acyl-transfers in plants are conveyed by BAHD family acyltransferases. They acylate a wide range of natural compounds and use CoA-activated thioesters as donors (D'Auria, 2006). As PGG formation does not require additional co-factors, other enzymes must have acyl-transfer activity, too. Recently, it was reported that proteins of the serine carboxypeptidase type can act as acyltransferases (Milkowski et al., 2004; Mugford et al., 2009; Shirley et al., 2001). Despite their close relatedness to proteases they do not exhibit peptidase activity but mediate acylation instead (Mugford and Osbourn, 2010). Apparently, constant adaption to changing environmental conditions has enabled plants to diversify protein functions. Similarly,

FaGT2 holds a dual function (Section 5.1.3). As an example, *A. thaliana* acyltransferase SMT catalyzes the formation of 1,2-di-O-sinapoyl- β -D-glucose from two molecules of 1-O-sinapoyl- β -D-glucose, while releasing free D-glucose (Stehle et al., 2008a,b). The chemical structures of sinapic acid and gallic acid resemble each other (Figure 23). Therefore, the production of di-, tri-, tetra-, and pentagalloylglucose could follow the same mechanism. Like serine carboxypeptidases, SCPL acyltransferases feature a α/β hydrolase fold and a conserved protein sequence motif (serine, histidine, and aspartate), which is referred to as "catalytic triad" (Stehle et al., 2009). However, SCPL acyltransferases have developed a slight alteration to the catalytic mechanism along evolution that distinguishes them from ordinary proteases. In order to incorporate the complex acyl glucose substrates, the network of hydrogen bonds in the active cleft was modified (Mugford and Milkowski, 2012). In contrast to the proteases, which possess the "GESYA" (glycine, glutamate, serine, tyrosine and alanine) penta-peptide motif, acyltransferases seem to have adapted a conserved "GDSYS" (glycine, aspartate, serine, tyrosine and serine) sequence (Fraser et al., 2005; Mugford and Osbourn, 2010; Mugford et al., 2009; Stehle et al., 2009). Of the acyltransferases analyzed here, FvAT2.4 and FvAT β contain the "GDSYS" motif, whereas FvAT1 shows the sequence "SDSYA". Site-directed mutagenesis of the FvAT1 motif could provide information about functionality and substrate specificity.

5.3.2 Production of recombinant FvAT1 and β

Expression of recombinant FvAT1 and β was not detected. It has been suggested that SCPL acyltransferases are secreted to other cell compartments (e.g. vacuole), and that therefore, a complex mechanism of post-translational modifications is necessary for correct protein folding (Mugford and Milkowski, 2012). For this reason, the eukaryotic host organism yeast was employed albeit without success. One possibility to improve

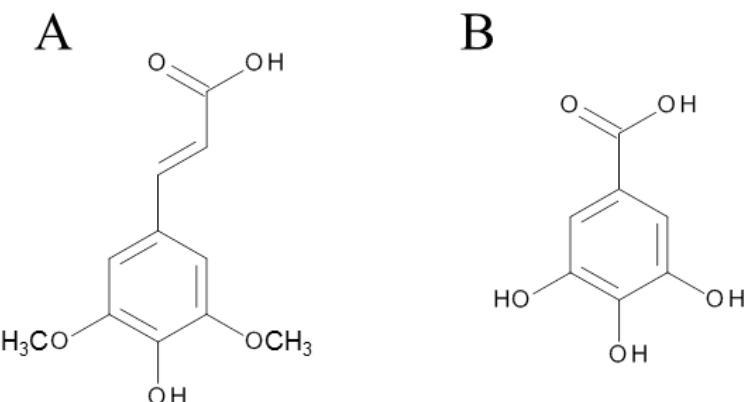


Figure 23: Comparison of the chemical structures of sinapic acid (A) and gallic acid (B).

the recombinant protein expression would be to tailor the sequences to the requirements of the host. The production of recombinant AtSMT in *S. cerevisiae*, for example, a sinapoyl-transferase from *A. thaliana* was improved after optimization of the codon usage and after removal of sequence motifs that disturbed the protein expression in yeast (Stehle et al., 2008b). Furthermore, other eukaryotic hosts such as *Nicotiana benthamiana* should be tested to allow correct folding.

5.3.3 Transient over-expression of candidate ATs in strawberry fruits

Via LC-MS measurements, the levels of β -glucogallin, di-galloyl glucose, and EA were determined in fruits carrying over-expression constructs of *FvAT1*, *FvAT3* and a *GUS* intron control. The overall concentration of β -glucogallin was low in all tested tissues (Figure 22). As precursor molecule, it is subjected to transformation into EA, which causes constant consumption of the intermediate (Section 5.1.4). The amount of di-galloyl glucose was decreased in *FvAT3* over-expressing fruits, but unchanged in *FvAT1* over-expressing fruits compared to the control. On the other hand, EA clearly accumulated in both transient transgenics. Until now, it is not known how many enzymes are involved in the formation of PGG. In

theory, every trans-esterification step could be mediated by an individual acyltransferase. Our results suggest that FvAT3 might be involved in the production of a higher intermediate, e.g. tri- or tetra-galloyl glucose, thereby reducing the di-galloyl glucose concentration. FvAT1 might catalyze other steps of the trans-glucosylation cascade. These results provide a first indication of the *in planta* function of FvAT1 and 3. However, further characterization is necessary, and special emphasis should be laid on the detection of *in vitro* activity.

6

CONCLUSIONS AND OUTLOOK

Over the past decade, natural plant products have become of major interest for the pharmaceutical, food and cosmetic industry. Polyphenols such as EA/ETs along with their precursors are of special importance to the public health. Moreover, they contribute to an effective protection against human diseases, exhibit antioxidative, antimicrobial and antifungal properties and are believed to delay skin aging. Indeed, recent findings strongly support the commonly held theory that these compounds have significant pharmacological and physiological properties. Ultimately, the goal would be to breed species naturally containing a high amount of applicable bio-active plant polyphenols, and make it available within a consumers reach. However, too high concentrations of EA/ETs would lead to unpalatable fruits due to their ability to precipitate proteins. Therefore, a balanced aroma and polyphenol profile would be crucial.

Considering the limited availability of sources containing EA and ETs, strawberry is an optimal choice. Firstly, it is a particularly rich source of these health-promoting polyphenols. Secondly, it is easily propagated as it possesses a comparatively short life cycle in comparison to other source plants, such as pomegranate and walnut. Lastly, its appealing deep-red color and sweet flavor is much valued by consumers all over the world. But before modification of the content of these polyphenols in plants can be enabled, the biosynthetic pathway of EA/ETs needs to be elucidated completely. This thesis contributes to its clarification: Firstly, the enzymes catalyzing the first known precursor of hydrolyzable tannins β -glucogallin were identified and characterized in garden and woodland strawberry, and raspberry. Secondly, extensive metabolome and transcriptome analysis enabled a deeper insight into the

biosynthetic pathway of secondary metabolites and allowed the selection of candidate genes presumably involved in the formation of PGG, the subsequent precursor of EA/ET biosynthesis. Thirdly, *in vivo* characterization of these candidate genes substantiated their putative function. However, further studies are necessary to complete the story. Special emphasis should be laid on the characterization of the enzymes facilitating the formation of di-, tri-, tetra- and pentagalloylglucose, as recombinant protein expression of FvAT₁ and β could not be achieved.

BIBLIOGRAPHY

- Abrahams, S; Lee, E; Walker, AR; Tanner, GJ; Larkin, PJ; Ashton, AR** (2003). "The Arabidopsis TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development." In: *Plant J* 35 (5), pp. 624–636.
- Aharoni, A** (2002). "Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays." In: *J Exp Bot* 53 (377), pp. 2073–2087.
- Aharoni, A; De Vos, C. H. Ric; 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." In: *Plant J* 28 (3), pp. 319–332.
- Aharoni, A et al.** (2000). "Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays." In: *Plant Cell* 12 (5), pp. 647–661.
- Aharoni, A et al.** (2004). "Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species." In: *Plant Cell* 16 (11), pp. 3110–3131.
- Allan, AC; Hellens, RP; Laing, WA** (2008). "MYB transcription factors that colour our fruit." In: *Trends Plant Sci* 13 (3), pp. 99–102.
- Almeida, JRM et al.** (2007). "Characterization of major enzymes and genes involved in flavonoid and proanthocyanidin biosynthesis during fruit development in strawberry (*Fragaria × ananassa*).". In: *Arch Biochem Biophys* 465 (1), pp. 61–71.
- Anders, S; Huber, W** (2010). "Differential expression analysis for sequence count data." In: *Genome Biol* 11 (10), R106.
- Anders, S; Pyl, PT; Huber, W** (2015). "HTSeq—a Python framework to work with high-throughput sequencing data." In: *Bioinformatics* 31 (2), pp. 166–169.
- Anderson, KJ; Teuber, SS; Gobeille, A; Cremin, P; Waterhouse, AL; Steinberg, FM** (2001). "Walnut polyphenolics inhibit in vitro human plasma and LDL oxidation." In: *J Nutr* 131 (11), pp. 2837–2842.
- André, I; Potocki-Véronèse, G; Barbe, S; Moulis, C; Remaud-Siméon, M** (2014). "CAZyme discovery and design for sweet dreams." In: *Curr Opin Chem Biol* 19, pp. 17–24.
- Aragüez, I et al.** (2013). "Eugenol production in achenes and receptacles of strawberry fruits is catalyzed by synthases exhibiting distinct kinetics." In: *Plant physiology* 163 (2), pp. 946–958.

- Arnold, TM; Targett, NM** (1998). "Quantifying in situ rates of phlorotannin synthesis and polymerization in marine brown algae." In: *J Chem Ecol* 24 (3), pp. 577–595.
- El-Basyouni, SZ; Chen, D; Ibrahim, RK; Neish, AC; Towers, G** (1964). "The biosynthesis of hydroxybenzoic acids in higher plants." In: *Phytochemistry* 3 (4), pp. 485–492.
- Benítez-Burraco, A; Blanco-Portales, R; Redondo-Nevado, J; Bellido, M. Luz; Moyano, E; Caballero, JL; Muñoz-Blanco, J** (2003). "Cloning and characterization of two ripening-related strawberry (*Fragaria × ananassa* cv. Chandler) pectate lyase genes." In: *J Exp Bot* 54 (383), pp. 633–645.
- Benjamini, Y; Hochberg, Y** (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." In: *J Roy Stat Soc B Met* 57 (1), pp. 289–300.
- Benton, HP; Want, EJ; Ebbels, TMD** (2010). "Correction of mass calibration gaps in liquid chromatography-mass spectrometry metabolomics data." In: *Bioinformatics* 26 (19), pp. 2488–2489.
- Bertani, G** (1951). "A method for detection of mutations, using streptomycin dependence in *Escherichia coli*." In: *Genetics* 36 (6), p. 598.
- Blankenberg, D; Johnson, JE; Taylor, J; Nekrutenko, A** (2014a). "Wrangling Galaxy's reference data." In: *Bioinformatics* 30 (13), pp. 1917–1919.
- Blankenberg, D; Kuster, Gv; Bouvier, E; Baker, D; Afgan, E; Stoler, N; Taylor, J; Nekrutenko, A** (2014b). "Dissemination of scientific software with Galaxy ToolShed." In: *Genome Biol* 15 (2), p. 403.
- Blankenberg, D; Kuster, Gv; Coraor, N; Ananda, G; Lazarus, R; Mangan, M; Nekrutenko, A; Taylor, J** (2010). "Galaxy: a web-based genome analysis tool for experimentalists." In: *Curr Protoc Mol Biol* 89, pp. 19.10.1–19.10.21.
- Bolger, AM; Lohse, M; Usadel, B** (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." In: *Bioinformatics* 30 (15), pp. 2114–2120.
- Bönisch, F; Frotscher, J; Stanitzek, S; Rühl, E; Wüst, M; Bitz, O; Schwab, W** (2014a). "A UDP-glucose:monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome." In: *Plant Physiol* 165 (2), pp. 561–581.
- (2014b). "Activity-based profiling of a physiologic aglycone library reveals sugar acceptor promiscuity of family 1 UDP-glucosyltransferases from grape." In: *Plant Physiol* 166 (1), pp. 23–39.
- Bontpart, T et al.** (2016). "Two shikimate dehydrogenases, VvSDH3 and VvSDH4, are involved in gallic acid biosynthesis in grapevine." In: *J Exp Bot* 67 (11), pp. 3537–3550.
- Bowles, D; Isayenkova, J; Lim, EK; Poppenberger, B** (2005). "Glycosyltransferases: managers of small molecules." In: *Curr Opin Plant Biol* 8 (3), pp. 254–263.

- Bowles, D; Lim, EK; Poppenberger, B; Vaistij, FE** (2006). "Glycosyltransferases of lipophilic small molecules." eng. In: *Annu Rev Plant Biol* 57, pp. 567–597.
- Bradford, MM** (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." In: *Anal Biochem* 72 (1-2), pp. 248–254.
- Brill, EM; Abrahams, S; Hayes, CM; Jenkins, Colin L. D.; Watson, JM** (1999). "Molecular characterisation and expression of a wound-inducible cDNA encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne (*Medicago sativa* L.)" In: *Plant Mol Biol* 41 (2), pp. 279–291.
- Buer, CS; Muday, GK; Djordjevic, MA** (2007). "Flavonoids are differentially taken up and transported long distances in *Arabidopsis*." In: *Plant Physiol* 145 (2), 10.1104/pp.107.101824, pp. 478–490.
- Burbulis, IE; Winkel-Shirley, B** (1999). "Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway." In: *P Natl Acad Sci USA* 96 (22), pp. 12929–12934.
- Cai, Y; Gaffney, SH; Lilley, TH; Magnolato, D; Martin, R; Spencer, CM; Haslam, E** (1990). "Polyphenol interactions. Part 4. Model studies with caffeine and cyclodextrins." In: *J Chem Soc* (12), p. 2197.
- Cammann, J; Denzel, K; Schilling, G; Gross, GG** (1989). "Biosynthesis of gallotannins. β -Glucogallin-dependent formation of 1,2,3,4,6-pentagalloylgucose by enzymatic galloylation of 1,2,3,6-tetragalloylgucose." In: *Arch Biochem Biophys* 273 (1), pp. 58–63.
- Castillejo, C; de la Fuente, José Ignacio; Iannetta, P; Botella, MÁ; Valpuesta, V** (2004). "Pectin esterase gene family in strawberry fruit: study of FaPE1, a ripening-specific isoform." In: *J Exp Bot* 55 (398), pp. 909–918.
- Charlton, AJ; Baxter, NJ; Khan, ML; Moir, AJG; Haslam, E; Davies, AP; Williamson, MP** (2002a). "Polyphenol/peptide binding and precipitation." In: *J Agr Food Chem* 50 (6), pp. 1593–1601.
- Charlton, AJ; Haslam, E; Williamson, MP** (2002b). "Multiple conformations of the proline-rich protein/epigallocatechin gallate complex determined by time-averaged nuclear Overhauser effects." In: *J Am Chem Soc* 124 (33), pp. 9899–9905.
- Cheel, J; Theoduloz, C; Rodríguez, J; Saud, G; Caligari, Peter D. S.; Schmeda-Hirschmann, G** (2005). "E-cinnamic acid derivatives and phenolics from Chilean strawberry fruits, *Fragaria chiloensis* ssp. *chiloensis*." In: *J Agr Food Chem* 53 (22), pp. 8512–8518.
- Cheong, JJ; Choi, YD** (2003). "Methyl jasmonate as a vital substance in plants." In: *Trends Genet* 19 (7), pp. 409–413.
- Civello, PM** (1999). "An expansin gene expressed in ripening strawberry fruit." In: *Plant Physiol* 121 (4), pp. 1273–1279.
- Conn, EE; Swain, T** (1961). "Biosynthesis of gallic acid in higher plants." In: *Chem Ind* 18, pp. 592–593.

- Cornthwaite, D; Haslam, E** (1965). "548. Gallotannins. Part IX. The biosynthesis of gallic acid in *Rhus typhina*." In: *J Chem Soc*, pp. 3008–3011.
- Cumplido-Laso, G** et al. (2012). "The fruit ripening-related gene FaAAT2 encodes an acyl transferase involved in strawberry aroma biogenesis." In: *J Exp Bot* 63 (11). 10.1093/jxb/ers120, pp. 4275–4290.
- Darrow, GM** (1966). *The strawberry. History, breeding and physiology*. New York: Holt, Rinehart and Winston.
- Darwish, O; Shahan, R; Liu, Z; Slovin, JP; Alkharouf, NW** (2015). "Re-annotation of the woodland strawberry (*Fragaria vesca*) genome." In: *BMC genomics* 16, p. 29.
- D'Auria, JC** (2006). "Acyltransferases in plants: a good time to be BAHD." In: *Curr Opin Chem Biol* 9 (3), pp. 331–340.
- Delker, C; Stenzel, I; Hause, B; Miersch, O; Feussner, I; Wasternack, C** (2006). "Jasmonate biosynthesis in *Arabidopsis thaliana* - enzymes, products, regulation." In: *Plant Biol* 8 (03), pp. 297–306.
- Deroles, SC; Bradley, J. Marie; Schwinn, KE; Markham, KR; Bloor, S; Manson, DG; Davies, KM** (1998). "An antisense chalcone synthase cDNA leads to novel colour patterns in *lisianthus* (*Eustoma grandiflorum*) flowers." In: *Mol Breeding* 4 (1), pp. 59–66.
- Dixon, RA; Achnine, L; Kota, P; Liu, CJ; Reddy, MSS; Wang, L** (2002). "The phenylpropanoid pathway and plant defence-a genomics perspective." In: *Mol Plant Pathol* 3 (5), pp. 371–390.
- Edelmann, A; Lendl, B** (2002). "Toward the optical tongue: Flow-through sensing of tannin–protein interactions based on FTIR spectroscopy." In: *J Am Chem Soc* 124 (49), pp. 14741–14747.
- Ehlting, J; Buttner, D; Wang, Q; Douglas, CJ; Somssich, IE; Kombrink, E** (1999). "Three 4-coumarate Coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms." In: *Plant J* 19 (1), pp. 9–20.
- Ernst, PB; Gold, BD** (2000). "The disease spectrum of *Helicobacter pylori*: The immunopathogenesis of gastroduodenal ulcer and gastric cancer." In: *Annu Rev Microbiol* 54, pp. 615–640.
- Espín, JC; González-Barrio, R; Cerdá, B; López-Bote, C; Rey, AI; Tomás-Barberán, FA** (2007). "Iberian pig as a model to clarify obscure points in the bioavailability and metabolism of ellagitannins in humans." In: *J Agric Food Chem* 55 (25), pp. 10476–10485.
- 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." In: *Plant Physiol* 148 (2), pp. 730–750.
- FAOSTAT** (2013). *Food and agriculture organization of the United Nations Statistics Division*. <http://faostat3.fao.org/browse/Q/QC/E> [accessed in January 2016].

- Ferrer, JL; Austin, MB; Stewart, C; Noel, JP** (2008). "Structure and function of enzymes involved in the biosynthesis of phenylpropanoids." In: *Plant Physiol Biochem* 46 (3), pp. 356–370.
- Fits, L van der; Memelink, J** (2000). "ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism." In: *Science* 289 (5477), pp. 295–297.
- Folta, KM; Gardiner, SE**, eds. (2009). *Genetics and genomics of Rosaceae*. New York: Springer New York.
- Forkmann, G** (1991). "Flavonoids as flower pigments. The formation of the natural spectrum and its extension by genetic engineering." In: *Plant Breeding* 106 (1), pp. 1–26.
- Francisco, RM** et al. (2013). "ABCC1, an ATP binding cassette protein from grape berry, transports anthocyanidin 3-O-glucosides." In: *Plant Cell* 25 (5). 10.1105/tpc.112.102152, pp. 1840–1854.
- Fraser, CM; Rider, LW; Chapple, C** (2005). "An expression and bioinformatics analysis of the *Arabidopsis* serine carboxypeptidase-like gene family." In: *Plant Physiol* 138 (2), pp. 1136–1148.
- Fuhrmann, B; Aviram, M** (2006). "Protection against cardiovascular diseases." In: *Pomegranates. Ancient Roots to Modern Medicine*. Ed. by NP Seeram; RN Schulman; D Heber. 1st ed. Boca Raton: Taylor & Francis Group, pp. 63–80.
- Funatogawa, K; Hayashi, S; Shimomura, H; Yoshida, T; Hatano, T; Ito, H; Hirai, Y** (2004). "Antibacterial activity of hydrolyzable tannins derived from medicinal plants against *Helicobacter pylori*." In: *Microbiol Immunol* 48 (4), pp. 251–261.
- Gachon, CMM; Langlois-Meurinne, M; Saindrenan, P** (2005). "Plant secondary metabolism glycosyltransferases: the emerging functional analysis." In: *Trends Plant Sci* 10 (11), pp. 542–549.
- Giampieri, F; Tulipani, S; Alvarez-Suarez, JM; Quiles, JL; Mezzetti, B; Battino, M** (2012). "The strawberry: Composition, nutritional quality, and impact on human health." In: *Nutrition* 28 (1), pp. 9–19.
- Giardine, B** et al. (2005). "Galaxy: A platform for interactive large-scale genome analysis." In: *Genome Res* 15 (10), pp. 1451–1455.
- Gil, MI; Tomás-Barberán, FA; Hess-Pierce, B; Holcroft, DM; Kader, AA** (2000). "Antioxidant activity of pomegranate juice and its relationship with phenolic composition and processing." In: *J Agr Food Chem* 48 (10), pp. 4581–4589.
- Giovannoni, JJ** (2004). "Genetic regulation of fruit development and ripening." In: *Plant Cell* 16, pp. 170–180.
- Glawischnig, E; Hansen, BG; Olsen, CE; Halkier, BA** (2004). "Camalexin is synthesized from indole-3-acetaldoxime, a key branching point between primary and secondary metabolism in *Arabidopsis*." In: *P Natl Acad Sci USA* 101 (21), pp. 8245–8250.

- Goecks, J; Nekrutenko, A; Taylor, J** (2010). "Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences." In: *Genome Biol* 11 (8), R86.
- Goodman, CD; Casati, P; Walbot, V** (2004). "A multidrug resistance-associated protein involved in anthocyanin transport in Zea mays." In: *Plant Cell* 16 (7). 10.1105/tpc.022574, pp. 1812–1826.
- Griesser, M et al.** (2008). "Redirection of flavonoid biosynthesis through the down-regulation of an anthocyanidin glucosyltransferase in ripening strawberry fruit." In: *Plant Physiol* 146 (4), pp. 1528–1539.
- Gross, GG** (1982). "Synthesis of β -glucogallin from UDP-glucose and gallic acid by an enzyme preparation from oak leaves." In: *FEBS Lett* 148 (1), pp. 67–70.
- Gross, GG** (1983a). "Partial purification and properties of UDP-glucose:vanillic acid 1-O-glucosyl transferase from oak leaves." In: *Phytochemistry* 22 (10), pp. 2179–2182.
- (1983b). "Synthesis of mono-, di- and trigalloyl- β -D-glucose by β -glucogallin-dependent galloyltransferases from oak leaves." In: *Z Naturforsch* 38c, pp. 519–523.
- Gross, GG; Denzel, K** (1991). "Biosynthesis of gallotannins. β -Glucogallin-dependent galloylation of 1, 6-digalloylglucose to 1, 2, 6-trigalloylglucose." In: *Z Naturforsch C* 46 (5-6), pp. 389–394.
- Grotewold, E** (2006). *The science of flavonoids*. New York: Springer.
- Guerra, FP; Wegrzyn, JL; Sykes, R; Davis, MF; Stanton, BJ; Neale, DB** (2013). "Association genetics of chemical wood properties in black poplar (*Populus nigra*).". In: *New Phytol* 197 (1), pp. 162–176.
- Hagenah, S; Gross, GG** (1993). "Biosynthesis of 1,2,3,6-tetra-O-galloyl- β -D-glucose." In: *Phytochemistry* 32 (3), pp. 637–641.
- Hall, D; Luca, Vd** (2007). "Mesocarp localization of a bi-functional resveratrol/hydroxycinnamic acid glucosyltransferase of Concord grape (*Vitis labrusca*).". In: *Plant J* 49 (4), pp. 579–591.
- Hanahan, D** (1983). "Studies on transformation of *Escherichia coli* with plasmids." In: *J Mol Biol* 166 (4), pp. 557–580.
- Hancock, JF** (1999). *Strawberries*. Vol. 11. Crop production science in horticulture. Oxon et al.: CABI Pub.
- Harris, GK et al.** (2001). "Effects of lyophilized black raspberries on azoxymethane-induced colon cancer and 8-hydroxy-2'-deoxyguanosine levels in the Fischer 344 rat." In: *Nutr Cancer* 40 (2), pp. 125–133.
- Härtl, K; Denton, A; Franz-Oberdorf, K; Hoffmann, T; Spornraft, M; Usadel, B; Schwab, W** (2017a). "Early metabolic and transcriptional variations in fruit of natural white-fruited *Fragaria vesca* genotypes." In: *Sci Rep* 7 (Article number:45113).
- Hartmann, T** (2007). "From waste products to ecochemicals: Fifty years research of plant secondary metabolism." In: *Phytochemistry* 68 (22-24), pp. 2831–2846.

- Haslam, E; Cai, Y** (1994). "Plant polyphenols (vegetable tannins): Gallic acid metabolism." In: *Nat Prod Rep* 11, p. 41.
- Haslam, E** (1989). *Plant polyphenols: Vegetable tannins revisited*. Cambridge et al.: Cambridge University Press.
- Hawkins, C; Caruana, J; Schiksnis, E; Liu, Z** (2016). "Genome-scale DNA variant analysis and functional validation of a SNP underlying yellow fruit color in wild strawberry." In: *Sci Rep* 6, p. 29017.
- Hoffmann, T; Kalinowski, G; Schwab, W** (2006). "RNAi-induced silencing of gene expression in strawberry fruit (*Fragaria x ananassa*) by agroinfiltration: a rapid assay for gene function analysis." In: *Plant J* 48 (5), pp. 818–826.
- Hoffmann, T; Kurtzer, R; Skowranek, K; Kießling, P; Fridman, E; Pichersky, E; Schwab, W** (2011). "Metabolic engineering in strawberry fruit uncovers a dormant biosynthetic pathway." In: *Metab Eng* 13 (5), pp. 527–531.
- Honda, C** et al. (2002). "Anthocyanin biosynthetic genes are coordinately expressed during red coloration in apple skin." In: *Plant Physiol Biochem* 40 (11), pp. 955–962.
- Horvath, A** et al. (2011). "Structured diversity in octoploid strawberry cultivars. Importance of the old European germplasm." In: *Ann Appl Biol* 159 (3), pp. 358–371.
- Huang, FC; Schwab, W** (2011). "Cloning and characterization of a 9-lipoxygenase gene induced by pathogen attack from *Nicotiana benthamiana* for biotechnological application." In: *BMC Biotechnol* 11 (1), p. 30.
- Huang, FC; Studart-Witkowski, C; Schwab, W** (2010). "Overexpression of hydroperoxide lyase gene in *Nicotiana benthamiana* using a viral vector system." In: *Plant Biotech J* 8 (7), pp. 783–795.
- Huber, W; Heydebreck, Av; Sultmann, H; Poustka, A; Vingron, M** (2002). "Variance stabilization applied to microarray data calibration and to the quantification of differential expression." In: *Bioinformatics* 18 (Suppl 1), pp. 96–104.
- Hui, C; Bin, Y; Xiaoping, Y; Long, Y; Chunye, C; Mantian, M; Wenhua, L** (2010). "Anticancer activities of an anthocyanin-rich extract from black rice against breast cancer cells in vitro and in vivo." In: *Nutr Cancer* 62 (8), pp. 1128–1136.
- Humphreys, JM; Chapple, C** (2002). "Rewriting the lignin roadmap." In: *Curr Opin Plant Biol* 5 (3), pp. 224–229.
- Iannetta, PPM; Laarhoven, LJ; Medina-Escobar, N; James, EK; McManus, MT; Davies, HV; Harren, FJM** (2006). "Ethylene and carbon dioxide production by developing strawberries show a correlative pattern that is indicative of ripening climacteric fruit." In: *Physiol Plantarum* 127 (2), pp. 247–259.

- Ito, H; Iguchi, A; Hatano, T** (2008). "Identification of urinary and intestinal bacterial metabolites of ellagitannin geraniin in rats." In: *J Agr Food Chem* 56 (2), pp. 393–400.
- Jaakola, L** (2013). "New insights into the regulation of anthocyanin biosynthesis in fruits." In: *Trends Plant Sci* 18 (9), pp. 477–483.
- Jia, HF; Chai, YM; Li, CL; Lu, D; Luo, JJ; Qin, L; Shen, YY** (2011). "Abscisic acid plays an important role in the regulation of strawberry fruit ripening." In: *Plant Physiol* 157 (1), pp. 188–199.
- Jia, H** et al. (2016). "Jasmonic acid involves in grape fruit ripening and resistant against Botrytis cinerea." In: *Funct Integr Genomics* 16 (1), pp. 79–94.
- Kadomura-Ishikawa, Y; Miyawaki, K; Takahashi, A; Noji, S** (2015). "RNAi-mediated silencing and overexpression of the FaMYB1 gene and its effect on anthocyanin accumulation in strawberry fruit." In: *Biol Plant* 59 (4), pp. 677–685.
- Kaiser, R; Mageney, V; Schwefel, K; Vollmers, D; Krüger, A; Horn, R** (2016). "Genotyping of red and white fruited strawberry (*Fragaria L.*) accessions and hybrids based on microsatellite markers and on the genetic diversity in the allergen genes fra a 1 and fra a 3." In: *Genet Resour Crop Ev* 63 (7), pp. 1203–1217.
- Kang, D; Gho, YS; Suh, M; Kang, C** (2002). "Highly sensitive and fast protein detection with coomassie brilliant blue in sodium dodecyl sulfate-polyacrylamide gel electrophoresis." In: *Bull Korean Chem Soc* 23 (11), pp. 1511–1512.
- Keutgen, AJ; Pawelzik, E** (2008). "Quality and nutritional value of strawberry fruit under long term salt stress." In: *Food Chem* 107 (4), pp. 1413–1420.
- Khater, F; Fournand, D; Vialet, S; Meudec, E; Cheynier, V; Terrier, N** (2012). "Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis." In: *J Exp Bot* 63 (3), pp. 1201–1214.
- Kim, D; Pertea, G; Trapnell, C; Pimentel, H; Kelley, R; Salzberg, SL** (2013). "TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions." In: *Genome Biol* 14 (4), R36.
- Kitamura, S; Shikazono, N; Tanaka, A** (2004). "TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*." In: *Plant J* 37 (1), pp. 104–114.
- Kobayashi, S; Ishimaru, M; Ding, CK; Yakushiji, H; Goto, N** (2001). "Comparison of UDP-glucose:flavonoid 3-O-glucosyltransferase (UGFT) gene sequences between white grapes (*Vitis vinifera*) and their sports with red skin." In: *Plant Sci* 160 (3), pp. 543–550.

- Koeduka, T et al.** (2006). "Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester." In: *P Natl Acad Sci USA* 103 (26), pp. 10128–10133.
- Koes, R; Verweij, W; Quattrocchio, F** (2005). "Flavonoids: A colorful model for the regulation and evolution of biochemical pathways." In: *Trends Plant Sci* 10 (5), pp. 236–242.
- Kolodziej, H; Kayser, O; Kiderlen, AF; Ito, H; Hatano, T; Yoshida, T; Foo, LY** (2001). "Antileishmanial activity of hydrolyzable tannins and their modulatory effects on nitric oxide and tumour necrosis factor-alpha-release in macrophages in vitro." In: *Planta Med* 67 (9), pp. 825–832.
- Kolodziej, H; Kayser, O; Latté, KP; Kiderlen, AF** (2000). "Enhancement of antimicrobial activity of tannins and related compounds by immune modulatory effects." In: *Plant Polyphenols 2*. Ed. by **GG Gross; RW Hemingway; T Yoshida; SJ Branham**. Boston and MA: Springer US, pp. 575–594.
- Kong, JM; Chia, LS; Goh, NK; Chia, TF; Brouillard, R** (2003). "Analysis and biological activities of anthocyanins." In: *Phytochemistry* 64 (5), pp. 923–933.
- Koponen, JM; Happonen, AM; Mattila, PH; Törrönen, AR** (2007). "Contents of anthocyanins and ellagitannins in selected foods consumed in Finland." In: *J Agri Food Chem* 55 (4), pp. 1612–1619.
- Korbie, DJ; Mattick, JS** (2008). "Touchdown PCR for increased specificity and sensitivity in PCR amplification." In: *Nat Protoc* 3 (9), pp. 1452–1456.
- Kosar, M; Kafkas, E; Paydas, S; Baser, KHC** (2004). "Phenolic composition of strawberry genotypes at different saturation stages." In: *J Agr Food Chem* 52 (6), pp. 1586–1589.
- Kutchan, TM; Hampp, N; Lottspeich, F; Beyreuther, K; Zenk, MH** (1988). "The cDNA clone for strictosidine synthase from *Rauvolfia serpentina* DNA sequence determination and expression in *Escherichia coli*." In: *FEBS Lett* 237 (1-2), pp. 40–44.
- Laemmli, UK** (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." In: *Nature* 227 (5259), pp. 680–685.
- Lairson, LL; Henrissat, B; Davies, GJ; Withers, SG** (2008). "Glycosyltransferases: Structures, functions, and mechanisms." In: *Annu Rev Biochem* 77, pp. 521–555.
- Landete, JM** (2011). "Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health." In: *Food Res Int* 44 (5), pp. 1150–1160.
- Landmann, C; Fink, B; Schwab, W** (2007). "FaGT2: A multifunctional enzyme from strawberry (*Fragaria × ananassa*) fruits involved in the metabolism of natural and xenobiotic compounds." In: *Planta* 226 (2), pp. 417–428.

- Larsen, K** (2004). "Molecular cloning and characterization of cDNAs encoding cinnamoyl CoA reductase (CCR) from barley (*Hordeum vulgare*) and potato (*Solanum tuberosum*).” In: *J Plant Physiol* 161 (1), pp. 105–112.
- Lazo, GR; Stein, PA; Ludwig, RA** (1991). "A DNA transformation–competent *Arabidopsis* genomic library in *Agrobacterium*." In: *Nat Biotechnol* 9 (10), pp. 963–967.
- Lehfeldt, C; Shirley, AM; Meyer, K; Ruegger, MO; Cusumano, JC; Viitanen, PV; Strack, D; Chapple, C** (2000). "Cloning of the SNG1 gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism." In: *Plant Cell* 12 (8), pp. 1295–1306.
- Li, AX; Steffens, JC** (2000). "An acyltransferase catalyzing the formation of diacylglycerol is a serine carboxypeptidase-like protein." In: *P Natl Acad Sci USA* 97 (12), pp. 6902–6907.
- Li, AX; Eannetta, N; Ghangas, GS; Steffens, JC** (1999). "Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase." In: *Plant Physiol* 121 (2). 10.1104/pp.121.2.453, pp. 453–460.
- Liao, YC; Li, HP; Kreuzaler, F; Fischer, R** (1996). "Nucleotide sequence of one of two tandem genes encoding phenylalanine ammonia-lyase in *Triticum aestivum*." In: *Plant Physiol* 112 (3), p. 1398.
- Liao, Z; Chen, M; Guo, L; Gong, Y; Tang, F; Sun, X; Tang, K** (2004). "Rapid isolation of high-quality total RNA from taxus and ginkgo." In: *Prep Biochem Biotech* 34 (3), pp. 209–214.
- Lin-Wang, K; McGhie, TK; Wang, M; Liu, Y; Warren, B; Storey, R; Espley, RV; Allan, AC** (2014). "Engineering the anthocyanin regulatory complex of strawberry (*Fragaria vesca*).” In: *Front Plant Sci* 5, pp. 1–14.
- Logemann, E; Parniske, M; Hahlbrock, K** (1995). "Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley." In: *P Natl Acad Sci USA* 92 (13), pp. 5905–5909.
- Lohse, M et al.** (2014). "Mercator: A fast and simple web server for genome scale functional annotation of plant sequence data." In: *Plant Cell Environ* 37 (5), pp. 1250–1258.
- Luger, P; Weber, M; Kashino, S; Amakura, Y; Yoshida, T; Okuda, T; Beurskens, G; Dauter, Z** (1998). "Structure of the tannin geraniin based on conventional X-ray data at 295 K and on synchrotron data at 293 and 120 K." In: *Acta Crystallogr Sect B Struct Sci* 54 (5), pp. 687–694.
- Lunkenbein, S; Bellido, M; Aharoni, A; Salentijn E. M., J; Kaldenhoff, R; Coiner, HA; Muñoz-Blanco, J; Schwab, W** (2006a). "Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose:cinnamate glucosyltransferase from strawberry." In: *Plant Physiol* 140 (3), pp. 1047–1058.

- Lunkenbein, S; Coiner, H; De Vos, C. H. Ric; Schaart, JG; Boone, MJ; Krens, FA; Schwab, W; Salentijn, EMJ** (2006b). "Molecular characterization of a stable antisense chalcone synthase phenotype in strawberry (*Fragaria × ananassa*)."
In: *J Agric Food Chem* 54 (6), pp. 2145–2153.
- Malik, A; Afaq, F; Sarfaraz, S; Adhami, VM; Syed, DN; Mukhtar, H** (2005). "Pomegranate fruit juice for chemoprevention and chemotherapy of prostate cancer."
In: *Proc Natl Acad Sci USA* 102 (41), pp. 14813–14818.
- Manning, K** (1998). "Isolation of a set of ripening-related genes from strawberry. Their identification and possible relationship to fruit quality traits."
In: *Planta* 205 (4), pp. 622–631.
- Marrs, KA; Alfenito, MR; Lloyd, AM; Walbot, V** (1995). "A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2."
In: *Nature* 375 (6530), pp. 397–400.
- McConn, M; Creelman, RA; Bell, E; Mullet, JE; Browse, J** (1997). "Jasmonate is essential for insect defense in *Arabidopsis*.
In: *Proc Natl Acad Sci USA* 94 (10), pp. 5473–5477.
- Medina-Puche, L et al.** (2014). "MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria × ananassa* fruits."
In: *J Exp Bot* 65 (2), pp. 401–417.
- Mehansho, H; Ann, DK; Butler, LG; Rogler, J; Carlson, DM** (1987). "Induction of proline-rich proteins in hamster salivary glands by isoproterenol treatment and an unusual growth inhibition by tannins."
In: *J Biol Chem* 262 (25), pp. 12344–12350.
- Milkowski, C; Baumert, A; Schmidt, D; Nehlin, L; Strack, D** (2004). "Molecular regulation of sinapate ester metabolism in *Brassica napus*: Expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation."
In: *Plant J* 38 (1), pp. 80–92.
- Milkowski, C; Strack, D** (2004). "Serine carboxypeptidase-like acyltransferases."
In: *Phytochemistry* 65 (5), pp. 517–524.
- Miosic, S et al.** (2014). "Dihydroflavonol 4-reductase genes encode enzymes with contrasting substrate specificity and show divergent gene expression profiles in *Fragaria* species."
In: *PLOS ONE* 9 (11), e112707.
- Mittasch, J; Böttcher, C; Frolova, N; Bönn, M; Milkowski, C** (2014). "Identification of UGT84A13 as a candidate enzyme for the first committed step of gallotannin biosynthesis in pedunculate oak (*Quercus robur*)."
In: *Phytochemistry* 99, pp. 44–51.
- Moréra, S; Larivière, L; Kurzeck, J; Aschke-Sonnenborn, U; Freemont, PS; Janin, J; Rüger, W** (2001). "High resolution crystal structures of T4 phage β-glucosyltransferase: Induced fit and effect of substrate and metal binding."
In: *J Mol Biol* 311 (3), pp. 569–577.
- Mueller, LA; Goodman, CD; Silady, RA; Walbot, V** (2000). "AN9, a petunia glutathione S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein."
In: *Plant Physiol* 123 (4), pp. 1561–1570.

- Mugford, ST; Milkowski, C** (2012). "Serine carboxypeptidase-like acyltransferases from plants." In: *Method Enzymol* 516, pp. 279–297.
- Mugford, ST; Osbourn, A** (2010). "Evolution of serine carboxypeptidase-like acyltransferases in the monocots." In: *Plant Signal Behav* 5 (2), pp. 193–195.
- Mugford, ST** et al. (2009). "A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats." In: *Plant Cell* 21 (8), pp. 2473–2484.
- Muñoz, C; Sánchez-Sevilla, JF; Botella, MA; Hoffmann, T; Schwab, W; Valpuesta, V** (2011). "Polyphenol composition in the ripe fruits of *Fragaria* species and transcriptional analyses of key genes in the pathway." In: *J Agric Food Chem* 59 (23), pp. 12598–12604.
- Niehaus, JU; Gross, GG** (1997). "A gallotannin degrading esterase from leaves of pedunculate oak." In: *Phytochemistry* 45 (8), pp. 1555–1560.
- Niemetz, R; Gross, GG** (2001). "Gallotannin biosynthesis: β -Glucogallin: hexagalloyl 3-O-galloyltransferase from *Rhus typhina* leaves." In: *Phytochemistry* 58 (5), pp. 657–661.
- (2003). "Ellagitannin biosynthesis. Laccase-catalyzed dimerization of tellimagrandin II to cornusin E in *Tellima grandiflora*." In: *Phytochemistry* 64 (7), pp. 1197–1201.
 - (2005). "Enzymology of gallotannin and ellagitannin biosynthesis." In: *Phytochemistry* 66 (17), pp. 2001–2011.
- Nonaka, GI; Harada, M; Nishioka, I** (1980). "Eugenin, a new ellagittannin from cloves." In: *Chem Pharm Bull* 28 (2), pp. 685–687.
- Okuda, T; Mori, K; Hatano, T** (1980). "The distribution of geraniin and mal-lotusinic acid in the order geranaiales." In: *Phytochemistry* 19 (4), pp. 547–551.
- Okuda, T; Yoshida, T; Hatano, T** (2000). "Correlation of oxidative transformations of hydrolyzable tannins and plant evolution." In: *Phytochemistry* 55 (6), pp. 513–529.
- Ossipov, V; Salminen, JP; Ossipova, S; Haukioja, E; Pihlaja, K** (2003). "Gallic acid and hydrolysable tannins are formed in birch leaves from an intermediate compound of the shikimate pathway." In: *Biochem Syst Ecol* 31 (1), pp. 3–16.
- Paniagua, C; Blanco-Portales, R; Barceló-Muñoz, M; García-Gago, JA; Waldrön, KW; Quesada, MA; Muñoz-Blanco, J; Mercado, JA** (2016). "Anti-sense down-regulation of the strawberry β -galactosidase gene Fa β Gal4 increases cell wall galactose levels and reduces fruit softening." In: *J Exp Bot* 67 (3), pp. 619–631.
- Pascual-Teresa, Sd; Santos-Buelga, C; Rivas-Gonzalo, JC** (2000). "Quantitative analysis of flavan-3-ols in Spanish foodstuffs and beverages." In: *J Agric Food Chem* 48 (11), pp. 5331–5337.
- Pearl, RT** (1928). *The history of the cultivated strawberry*. Wye: South-Eastern Agricultural College.

- Perkins-Veazie, P** (1995). "Growth and ripening of strawberry fruit." In: *Hortic Rev* 17 (8), pp. 267–297.
- Pfaffl, MW** (2001). "A new mathematical model for relative quantification in real-time RT-PCR." In: *Nucleic Acids Res* 29 (9), 45e–45.
- Preuß, A; Augustin, C; Figueroa, CR; Hoffmann, T; Valpuesta, V; Sevilla, JF; Schwab, W** (2014). "Expression of a functional jasmonic acid carboxyl methyltransferase is negatively correlated with strawberry fruit development." In: *J Plant Physiol* 171 (15), pp. 1315–1324.
- Qiu, Z; Wang, X; Gao, J; Guo, Y; Huang, Z; Du, Y** (2016). "The tomato Hoffman's Anthocyaninless gene encodes a bHLH transcription factor involved in anthocyanin biosynthesis that is developmentally regulated and induced by low temperatures." In: *PLOS ONE* 11 (3), e0151067.
- Que, Q; Wang, HY; Jorgensen, RA** (1998). "Distinct patterns of pigment suppression are produced by allelic sense and antisense chalcone synthase transgenes in petunia flowers." In: *Plant J* 13 (3), pp. 401–409.
- Quideau, S** (2009). *Chemistry and biology of ellagitannins. An underestimated class of bioactive plant polyphenols*. Hackensack and NJ: World Scientific.
- Quideau, S; Deffieux, D; Douat-Casassus, C; Pouységu, L** (2011). "Plant polyphenols: Chemical properties, biological activities, and synthesis." In: *Angew Chem Int Edit* 50 (3), pp. 586–621.
- R Core Team** (2015). "R: A language and environment for statistical computing." In: *R Foundation for Statistical Computing, Vienna, Austria*. URL <http://www.R-project.org>.
- Raab, T; López-Ráez, JA; Klein, D; Caballero, JL; Moyano, E; Schwab, W; Muñoz-Blanco, J** (2006). "FaQR, required for the biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase." In: *Plant Cell* 18 (4), pp. 1023–1037.
- Redondo-Nevado, J; Moyano, E; Medina-Escobar, N; Caballero, J. L.; Muñoz-Blanco, J** (2001). "A fruit-specific and developmentally regulated endopolygalacturonase gene from strawberry (*Fragaria × ananassa* cv. Chandler)." In: *J Exp Bot* 52 (362), pp. 1941–1945.
- Ring, L et al.** (2013). "Metabolic interaction between anthocyanin and lignin biosynthesis is associated with peroxidase FaPRX27 in strawberry fruit." In: *Plant Physiol* 163 (1), pp. 43–60.
- Ritchie, ME; Phipson, B; Di Wu, Hu, Y; Law, CW; Shi, W; Smyth, GK** (2015). "Limma powers differential expression analyses for RNA-seq and microarray studies." In: *Nucleic Acids Res* 43 (7), e47.
- Robinson, MD; McCarthy, DJ; Smyth, GK** (2010). "edgeR: A Bioconductor package for differential expression analysis of digital gene expression data." In: *Bioinformatics* 26 (1), pp. 139–140.
- Rosli, HG; Civello, PM; Martínez, GA** (2004). "Changes in cell wall composition of three *Fragaria × ananassa* cultivars with different softening rate during ripening." In: *Plant Physiol Biochem* 42 (10), pp. 823–831.

- Ross, J; Li, Y; Lim, EK; Bowles, DJ** (2001). "Higher plant glycosyltransferases." In: *Genome Biol* 2 (2), pp. 3004–1.
- Russell, WR; Labat, A; Scobbie, L; Duncan, GJ; Duthie, GG** (2009). "Phenolic acid content of fruits commonly consumed and locally produced in Scotland." In: *Food Chem* 115 (1), pp. 100–104.
- Salvatierra, A; Pimentel, P; Moya-León, MA; Herrera, R** (2013). "Increased accumulation of anthocyanins in *Fragaria chiloensis* fruits by transient suppression of FcMYB1 gene." In: *Phytochemistry* 90, pp. 25–36.
- Sanders, PM; Lee, PY; Biesgen, C; Boone, JD; Beals, TP; Weiler, EW; Goldberg, RB** (2000). "The *Arabidopsis* DELAYED DEHISCENCE1 gene encodes an enzyme in the jasmonic acid synthesis pathway." In: *Plant Cell* 12 (7), pp. 1041–1061.
- Sawada, S et al.** (2005). "UDP-glucuronic acid:anthocyanin glucuronosyl transferase from red daisy (*Bellis perennis*) flowers. Enzymology and phylogenetics of a novel glucuronosyltransferase involved in flower pigment biosynthesis." In: *J Biol Chem* 280 (2), pp. 899–906.
- Schaart, JG et al.** (2013). "Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria × ananassa*) fruits." In: *New Phytol* 197 (2), pp. 454–467.
- Schmidt, SW; Denzel, K; Schilling, G; Gross, GG** (1987). "Enzymatic synthesis of 1, 6-digalloylglucose from β-glucogallin by β-glucogallin: β-glucogallin 6-O-galloyltransferase from oak leaves." In: *Z Naturforsch C* 42 (1-2), pp. 87–92.
- Schulenburg, K; Feller, A; Hoffmann, T; Schecker, JH; Martens, S; Schwab, W** (2016a). "Formation of β-glucogallin, the precursor of ellagic acid in strawberry and raspberry." In: *J Exp Bot* 67 (8), pp. 2299–2308.
- Schuster, B; Herrmann, K** (1985). "Hydroxybenzoic and hydroxycinnamic acid derivatives in soft fruits." In: *Phytochemistry* 24 (11), pp. 2761–2764.
- Selma, MV; Beltran, D; Garcia-Villalba, R; Espin, JC; Tomas-Barberan, FA** (2014). "Description of urolithin production capacity from ellagic acid of two human intestinal Gordonibacter species." eng. In: *Food Funct* 5 (8). Journal Article Research Support, Non-U.S. Gov't, pp. 1779–1784.
- Sertürner, FW** (1805). "[no title]." In: *Journal der Pharmacie für Aerzte und Apotheker* 13, pp. 229–243.
- (1806). "Darstellung der reinen Mohnsäure (Opiumsäure) nebst einer Chemischen Untersuchung des Opiums mit vorzüglicher Hinsicht auf einen darin neu entdeckten Stoff und die dahin gehörigen Bemerkungen." In: *Journal der Pharmacie für Aerzte und Apotheker* 14, pp. 47–93.
- (1817). "Über das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandtheile des Opiums." In: *Ann Phys* 55, pp. 56–89.

- Shiota, S; Shimizu, M; Sugiyama, J; Morita, Y; Mizushima, T; Tsuchiya, T** (2004). "Mechanisms of action of corilagin and tellimagrandin I that remarkably potentiate the activity of β -lactams against methicillin-resistant *Staphylococcus aureus*." In: *Microbiol Immunol* 48 (6).
- Shirley, AM; McMichael, CM; Chapple, C** (2001). "The sng2 mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase." In: *Plant J* 28 (1), pp. 83–94.
- Shulaev, V** et al. (2011). "The genome of woodland strawberry (*Fragaria vesca*)."
In: *Nat Genet* 43 (2), pp. 109–116.
- Simirgiotis, MJ; Schmeda-Hirschmann, G** (2010). "Determination of phenolic composition and antioxidant activity in fruits, rhizomes and leaves of the white strawberry (*Fragaria chiloensis* spp. *chiloensis* form *chiloensis*) using HPLC-DAD-ESI-MS and free radical quenching techniques."
In: *J Food Compost Anal* 23 (6), pp. 545–553.
- Simirgiotis, MJ; Theoduloz, C; Caligari, Peter D. S.; Schmeda-Hirschmann, G** (2009). "Comparison of phenolic composition and antioxidant properties of two native Chilean and one domestic strawberry genotypes."
In: *Food Chem* 113 (2), pp. 377–385.
- Sloggett, C; Goonasekera, N; Afgan, E** (2013). "BioBlend: Automating pipeline analyses within Galaxy and CloudMan."
In: *Bioinformatics* 29 (13), pp. 1685–1686.
- Smith, CA; Want, EJ; O'Maille, G; Abagyan, R; Siuzdak, G** (2006). "XCMS: Processing mass spectrometry data for metabolite profiling using non-linear peak alignment, matching, and identification."
In: *Anal Chem* 78 (3), pp. 779–787.
- Song, C; Ring, L; Hoffmann, T; Huang, FC; Slovin, J; Schwab, W** (2015b). "Acylphloroglucinol biosynthesis in strawberry fruit."
In: *Plant Physiol* 169 (3), pp. 1656–1670.
- Song, C; Zhao, S; Hong, X; Liu, J; Schulenburg, K; Schwab, W** (2016b). "A UDP-glucosyltransferase functions in both acylphloroglucinol glucoside and anthocyanin biosynthesis in strawberry (*Fragaria × ananassa*)."
In: *Plant J* 85 (6), pp. 730–742.
- Spolaore, S; Trainotti, L; Casadore, G** (2001). "A simple protocol for transient gene expression in ripe fleshy fruit mediated by *Agrobacterium*."
In: *J Exp Bot* 52 (357), pp. 845–850.
- Stehle, F; Brandt, W; Schmidt, J; Milkowski, C; Strack, D** (2008a). "Activities of *Arabidopsis* sinapoylglucose:malate sinapoyltransferase shed light on functional diversification of serine carboxypeptidase-like acyltransferases."
In: *Phytochemistry* 69 (9), pp. 1826–1831.
- Stehle, F; Brandt, W; Stubbs, MT; Milkowski, C; Strack, D** (2009). "Sinapoyltransferases in the light of molecular evolution."
In: *Phytochemistry* 70 (15–16), pp. 1652–1662.

- Stehle, F; Stubbs, MT; Strack, D; Milkowski, C** (2008b). "Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism." In: *FEBS J* 275 (4), pp. 775–787.
- Stintzi, A; Browse, J** (2000). "The *Arabidopsis* male-sterile mutant, opr3, lacks the 12-oxophytodienoic acid reductase required for jasmonate synthesis." In: *Proc Natl Acad Sci USA* 97 (19), pp. 10625–10630.
- Symons, GM; Chua, YJ; Ross, JJ; Quittenden, LJ; Davies, NW; Reid, JB** (2012). "Hormonal changes during non-climacteric ripening in strawberry." In: *J Exp Bot* 63 (13), pp. 4741–4750.
- Tanaka, T** (1985). "Tannins and related compounds. Part 28. Revision of the structures of sanguinins H-6, H-2, and H-3, and isolation and characterization of sanguinin H-11, a novel tetrameric hydrolysable tannins and seven related tannins from *Sanguisorba officinalis*." In: *J Chem Res (M)*, pp. 2001–2029.
- Tanaka, T; Jiang, ZH; Kouno, I** (1997a). "Structures and biogenesis of rhoipteleanins, ellagitannins formed by stereospecific intermolecular C-C oxidative coupling, isolated from *Rhoiptelea chiliantha*." In: *Chem Pharm Bull* 45 (12), pp. 1915–1921.
- Tanaka, T; Zhang, H; Jiang, ZH; Kouno, I** (1997b). "Relationship between hydrophobicity and structure of hydrolyzable tannins, and association of tannins with crude drug constituents in aqueous solution." In: *Chem Pharm Bull* 45 (12), pp. 1891–1897.
- Tautenhahn, R; Böttcher, C; Neumann, S** (2008). "Highly sensitive feature detection for high resolution LC/MS." In: *BMC Bioinformatics* 9, p. 504.
- Tennessen, JA; Govindarajulu, R; Ashman, TL; Liston, A** (2014). "Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps." In: *Genome Biol Evol* 6 (12), pp. 3295–3313.
- Tohge, T et al.** (2015). "Ectopic expression of snapdragon transcription factors facilitates the identification of genes encoding enzymes of anthocyanin decoration in tomato." In: *Plant J* 83 (4), pp. 686–704.
- Tulipani, S et al.** (2008). "Antioxidants, phenolic compounds, and nutritional quality of different strawberry genotypes." eng. In: *J Agric Food Chem* 56 (3), pp. 696–704.
- Ulusu, NN** (2015). "Evolution of enzyme kinetic mechanisms." In: *J Mol Evol* 80 (5-6), pp. 251–257.
- Vallarino, JG et al.** (2015). "Central role of FaGAMYB in the transition of the strawberry receptacle from development to ripening." In: *New Phytol* 208 (2), pp. 482–496.
- Vijayan, P; Shockley, J; Lévesque, C. André; Cook, R. James; Browse, J** (1998). "A role for jasmonate in pathogen defense of *Arabidopsis*." In: *Proc Natl Acad Sci USA* 95 (12), pp. 7209–7214.
- Vrhovsek, U; Guella, G; Gasperotti, M; Pojer, E; Zancato, M; Mattivi, F** (2012). "Clarifying the identity of the main ellagitannin in the fruit of

- the strawberry, *Fragaria vesca* and *Fragaria ananassa* Duch." In: *J Agric Food Chem* 60 (10), pp. 2507–2516.
- Vrielink, A; Rüger, W; Driessen, HP; Freemont, PS** (1994). "Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose." In: *EMBO J* 13 (15), pp. 3413–3422.
- Wang, H; Fan, W; Li, H; Yang, J; Huang, J; Zhang, P** (2013). "Functional characterization of dihydroflavonol-4-reductase in anthocyanin biosynthesis of purple sweet potato underlies the direct evidence of anthocyanins function against abiotic stresses." In: *PLOS ONE* 8 (11), e78484.
- Wei, YZ; Hu, FC; Hu, GB; Li, XJ; Huang, XM; Wang, HC** (2011). "Differential expression of anthocyanin biosynthetic genes in relation to anthocyanin accumulation in the pericarp of *Litchi chinensis* Sonn." In: *PLOS ONE* 6 (4), e19455.
- Weisemann, S; Denzel, K; Schilling, G; Gross, GG** (1988). "Enzymatic synthesis of 1-O-phenylcarboxyl- β -d-glucose esters." In: *Bioorg Chem* 16 (1), pp. 29–37.
- Weng, JK; Chapple, C** (2010). "The origin and evolution of lignin biosynthesis." In: *New Phytol* 187 (2), pp. 273–285.
- Werner, I; Bacher, A; Eisenreich, W** (1997). "Retrobiosynthetic NMR studies with ^{13}C -labeled glucose. Formation of gallic acid in plants and fungi." In: *J Biol Chem* 272 (41), pp. 25474–25482.
- Werner, RA; Rossmann, A; Schwarz, C; Bacher, A; Schmidt, HL; Eisenreich, W** (2004). "Biosynthesis of gallic acid in *Rhus typhina*: Discrimination between alternative pathways from natural oxygen isotope abundance." In: *Phytochemistry* 65 (20), pp. 2809–2813.
- Wilhelm, S; Sagen, JE** (1974). *A history of the strawberry from ancient gardens to modern markets*. Berkeley and CA: University of California, Division of Agricultural Sciences.
- Xie, DY; Dixon, RA** (2005). "Proanthocyanidin biosynthesis—still more questions than answers?" In: *Phytochemistry* 66 (18), pp. 2127–2144.
- Xie, DY; Sharma, SB; Paiva, NL; Ferreira, D; Dixon, RA** (2003). "Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis." In: *Science* 299 (5605), pp. 396–399.
- Zadernowski, R; Naczk, M; Nesterowicz, J** (2005). "Phenolic acid profiles in some small berries." In: *J Agric Food Chem* 53 (6), pp. 2118–2124.
- Zenk, MH** (1964). "Notizen. Zur Frage der Biosynthese von Gallussäure." In: *Z Naturforsch* 19b (1), pp. 485–492.
- Zenk, MH** (1991). "6. Chasing the enzymes of secondary metabolism. Plant cell cultures as a pot of gold." In: *Phytochemistry* 30 (12), pp. 3861–3863.
- Zhang, J; Wang, X; Yu, O; Tang, J; Gu, X; Wan, X; Fang, C** (2011). "Metabolic profiling of strawberry (*Fragaria x ananassa* Duch.) during fruit development and maturation." In: *J Exp Bot* 62 (3), pp. 1103–1118.

- Zhang, L; Huang, Z; Wang, X; Gao, J; Guo, Y; Du, Y; Hu, H** (2016). "Fine mapping and molecular marker development of anthocyanin absent, a seedling morphological marker for the selection of male sterile 10 in tomato." In: *Mol Breed* 36 (8), p. 107.
- Zhao, J; Pang, Y; Dixon, RA** (2010). "The mysteries of proanthocyanidin transport and polymerization." In: *Plant Physiol* 153 (2), pp. 437–443.

PUBLICATIONS OF THE AUTHOR

ARTICLES

- Härtl, K; Denton, A; Franz-Oberdorf, K; Hoffmann, T; Spornraft, M; Usadel, B; Schwab, W (2017a). "Early metabolic and transcriptional variations in fruit of natural white-fruited *Fragaria vesca* genotypes." In: *Sci Rep* 7 (Article number:45113).
- Härtl, K; Huang, FC; Giri, A; Franz-Oberdorf, K; Frotscher, J; Schwab, W (2017b). "Glucosylation of smoke-derived volatiles in grapevine (*Vitis vinifera*) is catalyzed by a promiscuous resveratrol glucosyltransferase. [accepted manuscript]." In: *J Agric Food Chem*.
- Härtl, K; Kalinowski, G; Hoffmann, T; Preuss, A; Schwab, W (2017c). "RNAi-mediated endogene silencing in strawberry fruit: detection of primary and secondary siRNAs by deep sequencing." In: *Plant Biotechnol J* 15 (5), pp. 658–668.
- Schulenburg, K; Feller, A; Hoffmann, T; Schecker, JH; Martens, S; Schwab, W (2016a). "Formation of β -glucogallin, the precursor of ellagic acid in strawberry and raspberry." In: *J Exp Bot* 67 (8), pp. 2299–2308.
- Song, C; Hong, X; Zhao, S; Liu, J; Schulenburg, K; Huang, FC; Franz-Oberdorf, K; Schwab, W (2016a). "Glucosylation of 4-hydroxy-2,5-dimethyl-3(2H)-furanone, the key strawberry flavor compound in strawberry fruit." In: *Plant Physiol* 171 (1), pp. 139–151.
- Song, C; Zhao, S; Hong, X; Liu, J; Schulenburg, K; Schwab, W (2016b). "A UDP-glucosyltransferase functions in both acylphloroglucinol glucoside and anthocyanin biosynthesis in strawberry (*Fragaria × ananassa*).". In: *Plant J* 85 (6), pp. 730–742.
- Song, C; Le, G; Liu, J; Zhao, S; Hong, X; Schulenburg, K; Schwab, W (2015a). "Functional characterization and substrate promiscuity of UGT71 glycosyltransferases from strawberry (*Fragaria × ananassa*).". In: *Plant Cell Physiol* 56 (12), pp. 2478–2493.

PRESENTATIONS AND POSTERS

- Schulenburg, K; Huang, FC; Schwab, W (2016b). *Plant glucosyltransferases for the biotechnological production of small molecule glucosides*. Poster. Winner of 1st Poster Prize. Garching: Industrial Biotechnology Forum.

- Schulenburg, K; Franz-Oberdorf, K; Schwab, W** (2015). *What's the difference between a white and a red strawberry.* Poster. Freising: qPCR and NGS Symposium.
- Schulenburg, K; Schwab, W** (2015). *Transiente Überexpression von Acyltransferasen in F. × ananassa cv. Elsanta.* Presentation. Freising: GHL Symposium Gewächshauslaborzentrum Dürnast.
- Franz, K; Schulenburg, K; Schwab, W** (2013). *Biosynthese der Naturstoffe in Erdbeeren.* Presentation. Freising: GHL Symposium Gewächshauslaborzentrum Dürnast.
- Schulenburg, K; Schwab, W** (2013). *Formation of ellagic acid precursors in Fragaria.* Poster. Tübingen: Deutsche Botanikertagung.